Modular Supramolecular Biomaterials Based on a Coiled-Coil Scaffold

by

Kaylyn Marie Oshaben

B.S. Chemistry, Saint Francis University, 2010

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kaylyn Marie Oshaben

It was defended on

January 22nd, 2016

and approved by

Subha Das, Associate Professor, Departmental of Chemistry, Carnegie Mellon University

Nathaniel Rosi, Professor, Department of Chemistry

Sunil Saxena, Professor, Departmental of Chemistry

Dissertation Advisor: W. Seth Horne, Associate Professor, Department of Chemistry

MODULAR SUPRAMOLECULAR BIOMATERIALS BASED ON A COILED-COIL

SCAFFOLD

Kaylyn Marie Oshaben, PhD

University of Pittsburgh, 2016

Nature uses proteins and nucleic acids to form a wide array of structural motifs. Chemists have applied these motifs to the rational design of supramolecular biomaterials. Intricate assemblies of fibers, nets and spheres have been synthesized and characterized; however, existing approaches often lack fine control over size and morphology. In an effort to address this limitation, we have developed a system based on self-assembly of a modular subunit consisting of two α-helical peptides, which self associate to form a coiled coil, attached at their midpoints by a small organic linking group.

We found that the linker identity not only impacted the flexibility of the assemblies but linker length was important to maintaining the folding of the peptides in the subunit. Our subunit design also allowed us to examine if assembly size could be controlled by changes to coiled-coil stability through sequence mutations. Supramolecular polymer growth models show assembly size can be controlled by changes in the association affinity of the monomer. We designed, synthesized and characterized a series of coiled coils with

varying folded stabilities to use in the subunits and observed that assembly size increased when the stability of the coiled coil is increased.

With the impact of the components of our self-assembling subunits characterized, we began examining if added functionality fluctuated with changes to the subunit. We developed a synthetic scheme for attaching a donor fluorophore and used a capping peptide labeled with an acceptor fluorophore to study Förster resonance energy transfer in the dimeric coiled coil and larger assemblies.

Finally, we observed that GCN4-p1, a well studied dimeric coiled coil, crystallized as either a dimer or trimer depending on the crystallization conditions. We carried out an extensive panel of solution-phase experiments to determine if the trimeric oligomerization state exists as measureable population. We found the solution conditions impact the preferred oligomerization state in the GCN4-p1 sequence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

PREFACE

How and where do you start acknowledging the people who helped you accomplish something as challenging as earning a Ph.D.? Honestly, words are not an adequate expression for the gratitude I feel, but words are what I have to work with and I'm going try my best.

First, I have to thank my advisor Dr. W. Seth Horne—who has not only acted as a fantastic teacher and mentor over the past 5 years, but without his willingness to take me on as a graduate student, I doubt I would have found a lab to call home. After not being selected for a research group, I was tasked with conducting a second round of somewhat awkward interviews. During my interview with Seth, he asked why I hadn't looked at the group in my first round of interviews. I don't remember exactly what my response was, but I can guarantee it, along with my pink hair, did not instill confidence in my seriousness and abilities as a scientist pursuing an advanced degree. Regardless the reasoning behind Seth's decision to open a spot up in the lab for me, joining the Horne lab was probably the most important point in my graduate career. Seth provided me with an engaging and challenging project along with what felt like endless amounts of guidance, patience and understanding over the past several years. Thank you, Seth, for your dedication and for being a model for the type of scientist I want to be.

I could not have survived graduate school without being surrounded by a great set of lab mates. I have to thank Dr. George Lengyel and Dr. Conor Haney for their always honest critiques and answering "if I were (insert a lab supply), where would I live" a countless number of times; Dr. Zach Reinert for always lending a listening ear to my ramblings about science or the Penguins, and Ken, Nate, Halina, Kelly, Hajira, Haley and Chino for making the lab an enjoyable place to

walk into each day. To my two undergrads, Kelsey Wiggins and Laura Pinksavage, thank you for the countless hours spent in front of the semi-prep. I wish you the best.

Next, I owe a huge thank you to my family—Mom, Dad, Sandi, Bob, Bubba and Grandpa. Your support and love mean the world to me even if you have no idea why I care so much about the "curly cue protein thingies"—hopefully this finally explains why. Power Heather, I made it through—who would have guessed? You, my friend, you. Kesslak, Bourdess, Chuck, Jim, my high school chemistry teacher Mr. Obarsky, Hot Metal Bridge Faith Community, all of my house mates over the years and friends, both old and new, thank you for your endless encouragement, kind words and late nights of both work and play.

Finally, I'd like to thank my two main squeezes—Ryan and Schrödinger (my cat). You always put your faith in me even when I questioned every choice I was making or when I did completely irrational things like crying over spilt coffee. I couldn't have done this without your limitless love and support. Thank you.

1.0   INTRODUCTION

1.1   SUPRAMOLECULAR POLYMERS

A polymer, in the simplest terms, is a macromolecule made up of a repeating monomer or subunit. Supramolecular polymers are defined as a directional and reversible polymer that forms by non-covalent driving forces.[1] This broad definition allows a wide range of molecules to be classified as supramolecular polymers. This class of polymers can exhibit properties similar to their covalently bonded counterparts, but differ in the increased capacity to be responsive to stimuli such as temperature[2], pH[3] and light.[4, 5] The responsiveness of supramolecular polymers can arise from the influence of either an external stimulus on the monomer unit, which is seen in conventional polymers, or from the disruption of the non-covalent force driving polymer assembly. The latter mechanism is unique to supramolecular polymers, and it also allows for easy recycling of monomer.

Supramolecular polymers derive their interesting properties from the ordered but dynamic interactions of their monomers. The time scale of the non-covalent interactions is important for maintaining these properties: too short of a time scale and no ordered polymer exists, while too long of a time scale causes the polymer to have properties closer to conventional covalent

polymers. Because of their dynamic interactions, supramolecular polymers have found use in fields such as self-healing materials[2] and biomedical applications[6, 7]. Polymers that are repeatedly stressed, mechanically or otherwise, degrade over time and their degradation can cause catastrophic failures. Self-healing covalent polymers usually require a harsh or toxic chemical to begin molecular rearrangement for repairs. Supramolecular polymers can use both their dynamic bonds and their responsiveness to external stimuli, such as temperature changes, to initiate rearrangement making the self-healing process less hazardous.

Self-healing in polymers is a biomimetic process of what many organisms do naturally when injured. Thus, it should be no surprise that supramolecular polymers have also found use in biomedical applications. In this context, the supramolecular polymer serves as a scaffold for the growth of another material. Peptide-amphiphiles, short peptide sequences with an extended hydrocarbon chain on one end, function as a supramolecular scaffold for biomineralization[6] by self-assembling into nanorods and fibers. The peptide sequence exposed to the surroundings can be tuned for interaction with a specific molecule. This type of scaffold has been used to generate new blood vessels,[8] guide cellular differentiation and promote bone and cartilage regrowth.[6, 9]

Supramolecular polymers represent an interesting and complex class of materials. Their classification has been based on three different principles: 1) the nature of the non-covalent driving force, 2) the type of monomer used and 3) the mechanism of polymer formation.[10] The classification scheme for supramolecular polymers has evolved as the field has expanded over the last two decades.

The initial classification of supramolecular polymers focused solely on the driving force behind polymer assembly and separated polymers into groups based on hydrogen bonding[11], π-π interactions[12], hydrophobic interactions and metal-ligand binding[13]. This type of grouping allowed

for easy classification of supramolecular polymers and enabled a direct link between association affinity of the driving interaction and the size of the polymer formed. However, classification based strictly on driving force ignores other important underlying mechanisms and factors in polymer growth. As the field of supramolecular polymers expanded, the need arose for a more comprehensive system of classification that included the other factors in assembly formation.

A second level of classification in supramolecular polymers takes the monomer identity and interaction type into consideration. As alluded to earlier, the monomers used in supramolecular polymers span a range of structures and complexities. On one end of the complexity spectrum, monomers can consist of small molecules capable of directional non-covalent interactions. More complex monomers may be composed of larger, more intricate
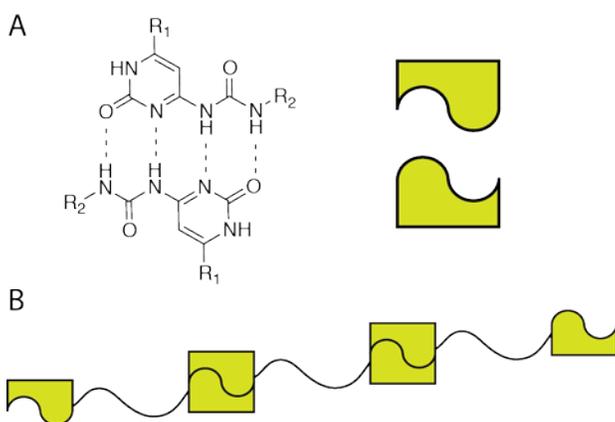
Figure 1. Hydrogen bond patterning in UPy molecules

Hydrogen bonding in Upy molecules (a) and a cartoon representation of the formation of linear supramolecular polymers when two UPy molecules are covalently linked (b). This is an example of an A-A interaction.

macromolecules or a subunit type design such as functionalized ureidopyrimidinone derivatives

(Figure 1).[11] When monomers have the same identity and assemble using the same driving force, the interaction is considered A-A type. Ureidopyrimidinone derivatives are an example of A-A interaction type monomers. Monomers may also be classified as A-B type. A-B interaction monomers have the same chemical identity but complimentary assembly moieties; The Hamilton Wedge is an example.[14] The wedge (N^1,N^3-bis(6-butyramidopyridin-2-yl)-5-hydroxyisophthalamide) contains six hydrogen bonding groups that point inward and can interact with cyanuric acid and other barbiturates (Figure 2). Systems have also been reported that have a single type of interaction but two different bifunctionalized monomers. Crown ethers and the formation of pseudoroxtanes have been used to generate A-B type interactions where a single
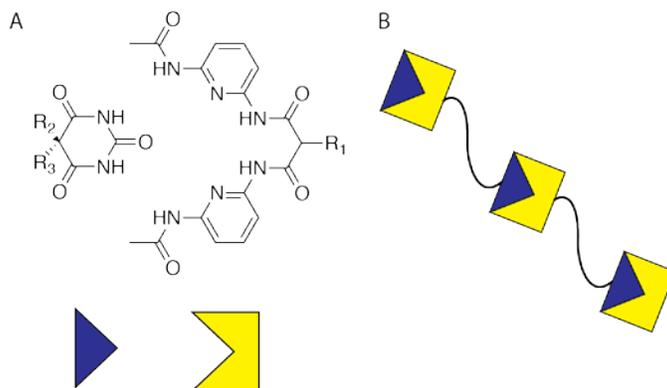
Figure 2. The Hamilton Wedge supramolecular receptor.

The wedge contains 6 patterned hydrogen bond sites pointed inward (donor-acceptor-donor-donor-acceptor-donor) that can interact with complimentary small molecules (a). A short linker can connect the two pieces of the wedge to form a supramolecular polymer (b). This is an example of an A-B interaction.

crown ether-charged amine is used.[15] These types of monomers have also been designed so they

4

assemble to form self-sorting co-block polymers (Figure 3).[16]
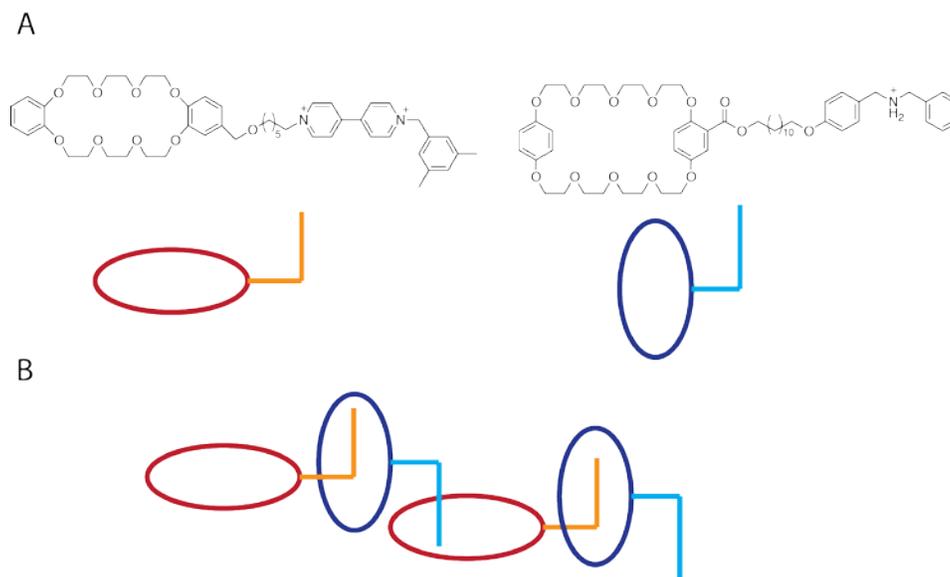
Figure 3. Examples of pseudoroxtane supramolecular polymers.
Structures of two different pseudoroxtanes and their cartoon representations (a). These two psedoroxtanes are self sorting and form co-block supramolecular polymers in solution (b).

While categorization by monomer type provides a slightly better way to sort supramolecular polymers, this scheme falls short of providing a clear description of the polymers' behavior. Based on the diversity of monomer types, supramolecular polymers can form by an assortment of growth mechanisms. Common growth mechanisms fall into three categories: isodesmic, cooperative and ring-and-chain.[17] Each category has thermodynamic properties that help to distinguish it from the other growth mechanisms.

Isodesmic supramolecular polymerization describes a system where the addition of a single monomer is identical over all steps of the linear polymerization process. The addition of monomers

to the growing polymer chain in this model does not change the association affinity of the monomers and equally decreases the $\Delta G°$ for the polymer after each addition (Figure 4). Systems that follow the this model tend to be the simplest systems to characterize and
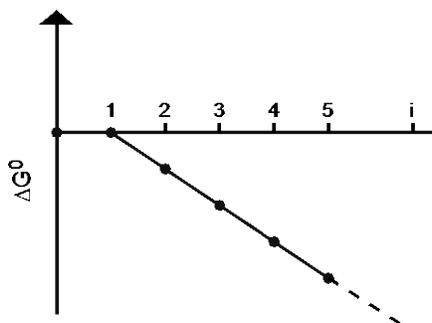
Figure 4. Growth mechanism in isodesmic supramolecular polymers.

The change in free energy in an isodesmic self-associating supramolecular polymer per monomer addition. Each addition of a monomer contributes an equal negative value to the $\Delta G°$.

manipulate because the degree of polymerization (size of the polymer formed) is a function of the association affinity and the concentration of the monomer in solution.[18] Many examples of this type of supramolecular polymer are based off of the ureidopyrimidinone molecule. Ureidopyrimidinone (UPy) forms four high affinity self-complimentary hydrogen bonds and the UPy moiety has been linked together by a variety of molecules.[19, 20] UPy-based supramolecular polymers are responsive to external stimuli such as temperature and pH. However, they suffer from limitations shared by other supramolecular polymers such as the need for high concentration of monomer if the association affinity is low (the association constant of UPy is close to 10^{-7}M in

CDCl$_3$)[21]. While concentration can be increased easily in many cases, the monomer will always be the most abundant species in solution by number, and increasing concentration in isodesmic systems increases the polydispersity of the assemblies. Also, care must be taken when measuring the degree of polymerization in any supramolecular system, as small changes in solution conditions can alter this property.

The second growth mechanism possible in supramolecular polymers is the cooperative association model. Unlike the isodesmic model, the cooperative association model consists of two stages of polymer growth governed by two association constants.[22] The first stage of growth can be separated into two categories, nucleation and downhill, depending on the value of $\Delta G°$ for the growing polymer chain. Cooperative nucleation association systems are classified by an initial increase in $\Delta G°$ of the polymer and a lag in the formation of an actual supramolecular polymer. The lag comes from the time for a nucleation or seeding event to occur before a barrier is reached and the polymerization follows an isodesmic model (Figure 5a). The second cooperative association model is the downhill model where $\Delta G°$ of polymer formation decreases at every step but the magnitude per addition changes at some point (Figure 5b). The nucleation steps in the downhill model have a lower association affinity than the elongation phase (which then follows an isodesmic model of association). The main difference between these two cooperative association models is identity of the least stable species; in the nucleation model, the
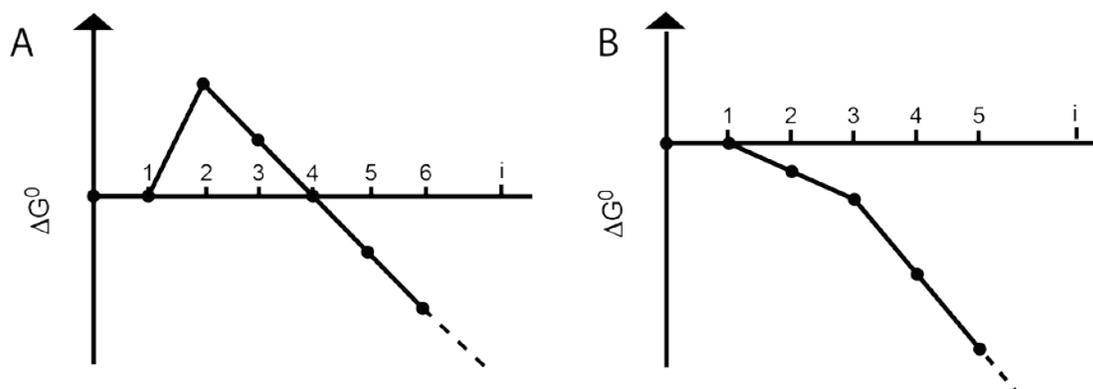
Figure 5. Cooperative supramolecular growth mechanisms.

The change in free energy in a cooperative nucleation (a) and downhill (b) self-associating supramolecular polymer system. In the cooperative nucleation system, polymer formation is unfavorable until a nucleus of a certain size is reached. In a cooperative downhill system, there is a change in the amount of free energy per monomer addition at some point.

nucleated "seed" is the least stable species, while in the downhill model, the monomer is the least stable. The driving force behind the cooperativity in these systems varies depending on the identity of the monomer and can be caused by electronic, hydrophobic and structural effects.[1] Structural effects arise from a change in the association affinity of the monomer due to the growing polymer chain interacting with itself (i.e. the association affinity of a monomer forming a helix changes after a turn of the helix is completed due to interactions arising from the structured helix).

An interesting example of cooperative supramolecular polymerization is found in engineered coiled-coil based fibers. The coiled coil (discussed extensively in section 1.3) is a protein motif were two α-helices fold and associate due to the burial of a hydrophobic interface.[23] Additional charged groups were engineered onto sides of the motif to promote lateral intermolecular electrostatic interactions, which served as the nucleation event for the formation of fibers through the end-to-end association of the coiled coils (Figure 6).[24]

8
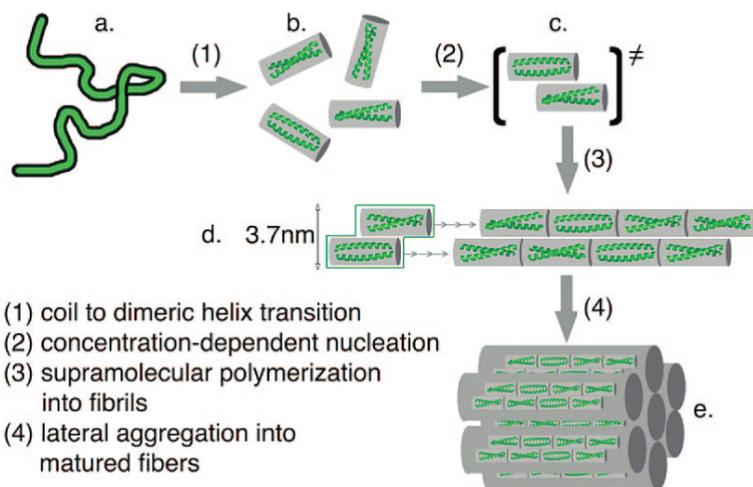
Figure 6. Cooperative downhill supramolecular polymer made of coiled coil peptide fibers. The peptide first folds into a dimeric coiled coil through the burial of a hydrophobic interface. Designed electrostatic interactions cause the coils to associate side-by-side and end-to-end forming fibers. The fibers can then associate with other fibers via electrostatic interactions. Reprinted with permission from Dong, H.; *et al.*, *J. Am. Chem. Soc.*, **2008**, *130*, 13691-13695. Copyright 2008 American Chemical Society.

A final growth mechanism possible in supramolecular polymers is the ring-and-chain association model. This model describes a system where there is an equilibrium between linear and cyclic species in solution. Monomers formed by this type of mechanism tend to have a flexible linker between the associating end groups, and they show complex equilibria. There is a critical concentration below which ring structures are favored and above which linear polymers begin to form.[17] The critical concentration is dependent on the length and flexibility of monomer itself. The previously described pseudoroxtane-forming crown ethers illustrate this type of equilibrium.[25]

Supramolecular polymers are an interesting class of materials with great potential due to their tunability and responsiveness to a variety of stimuli. The majority of the supramolecular polymer field has focused on synthetic small molecules. A relatively small number of supramolecular polymers using biomolecules as scaffolds have been reported in the literature such as cyclic peptides[26] and coiled-coil fibers[23]. Many of the driving forces behind protein and nucleic acid folding are the same that drive the formation of supramolecular polymers, and there are many methods available to determine and control association affinities for these biomolecules. However, the potential complexity introduced due to the sequence-structure relationships has limited their consideration for use in supramolecular polymers. An expanding field focused on protein-based materials may help to fill the gap in knowledge needed to utilize these biomolecules in supramolecular polymers.

1.2   PROTEIN-BASED MATERIALS

Nature uses wonderfully designed architectures to create functional and dynamic self-assembling materials out of proteins,[27] nucleic acids,[28] and carbohydrates.[25] Scientists have worked with these building blocks or their mimetics to create novel materials with designed functionality. Most of the early work in self-assembling biomaterials utilized DNA and RNA because of the well-defined interaction patterns of nucleic acids.[28, 29] Only more recently has the more diverse family of peptides and proteins been used to create novel architectures.[30]

In simple terms, protein sequence defines folded structure and folded structure defines function. Proteins primarily use the 20 naturally occurring amino acids as sequence building

blocks. Synthetic biologists interested in making protein-based materials have these building blocks and many other unnatural variants to choose from in their designs. The diversity of available amino acids, the lack of definitive rules for sequence-structure relationships and synthetic inaccessibility of larger proteins present ongoing challenges in the *de novo* design of protein-based materials. Much of the work in the field has focused on using naturally occurring protein structural motifs or slight modifications thereof to create new materials.

Some naturally occurring proteins fold to form larger supramolecular architectures or discreet oligomers, and these provide useful templates for designing new materials. [27, 31, 32, 33] The advantage of using a naturally occurring protein as a starting point is the knowledge of folding properties and the ability to insert rational mutations. A great example of a protein-based material from nature is the tobacco mosaic virus (TMV) coat protein, which folds to form disks or rods depending on the pH of the solution.[31] TMV has been used to create a light harvesting biomaterial (Figure 7).[34] A site-specific mutation of Ser→Cys in the TMV coat protein provided a functional handle for the attachment of donor and acceptor fluorophores for Förster resonant energy transfer (FRET) studies. While naturally occurring protein structures offer many advantages, they also have some disadvantages to their use in materials. First, many naturally occurring proteins are not accessible to chemical synthesis due to their sizes, limiting the types of amino acids available for insertion. Second, because the sequence-structure relationship is already determined, extensive modification or modification of specific residues may disrupt or completely eliminate folding and assembly.
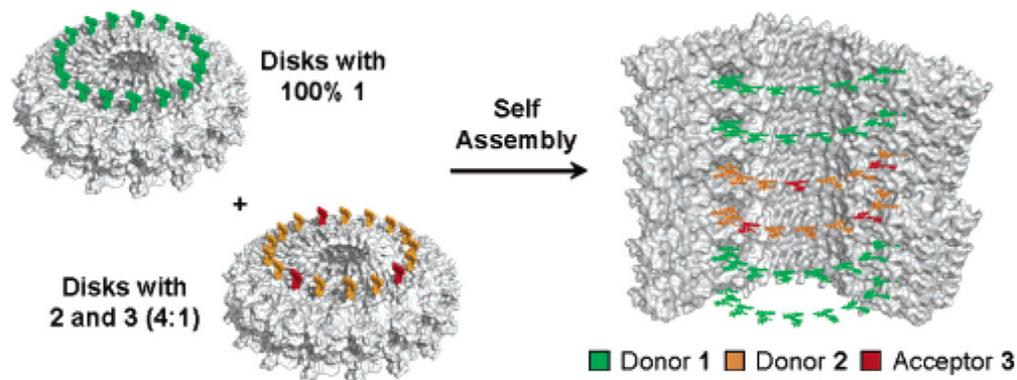
Figure 7. A light harvesting supramolecular polymer designed from TMV.
The protein from the TMV was modified with two donors and one acceptor fluorophore and allowed to assemble
into disks with varying ratios of the donors and acceptors. The disks were then mixed to form a supramolecular
polymer and the energy transfer efficiency measured by FRET. Adapted with permission from Miller, R. A.; *et al.,
J. Am. Chem. Soc.*, **2007**, *192*, 3104-3109. Copyright 2007 American Chemical Society.

An area that bridges chemical biology and materials and has attracted considerable interest in recent years is the construction of fusion proteins to create nm-scale objects using rational design and computational methods. As an example, fusion proteins can be constructed from two different domains connected through a short linker.[32] The protein domains are chosen for specific symmetry elements to guide the formation discrete nanostructures (i.e. a 2-fold and 3-fold fusion protein would form oligomers in multiples of 6). Designing such fusion proteins is challenging, but several methods have given rise to discrete oligomers. Early efforts focused on rational modification of existing proteins. Protein cages and elongated fibers using a dimeric and a trimeric protein domains have been assembled by their linkage through a short α-helix (Figure 8a and 8b).[32] The proteins were chosen by searching the protein data bank (PDB) for dimer and trimer forming sequences that began or ended with an α-helix. The proteins were then computationally linked and analyzed before a final fusion oligomer was chosen. More recent efforts have successfully designed specific

12

nanoshapes such as cubes using the same methodology (Figure 8c).[35] In addition to computation-aided rational design, some examples of *de novo* proteins used in conjunction with natural proteins to construct fusion proteins exist[36]; however, the *de novo* design of proteins is too demanding for widespread use.
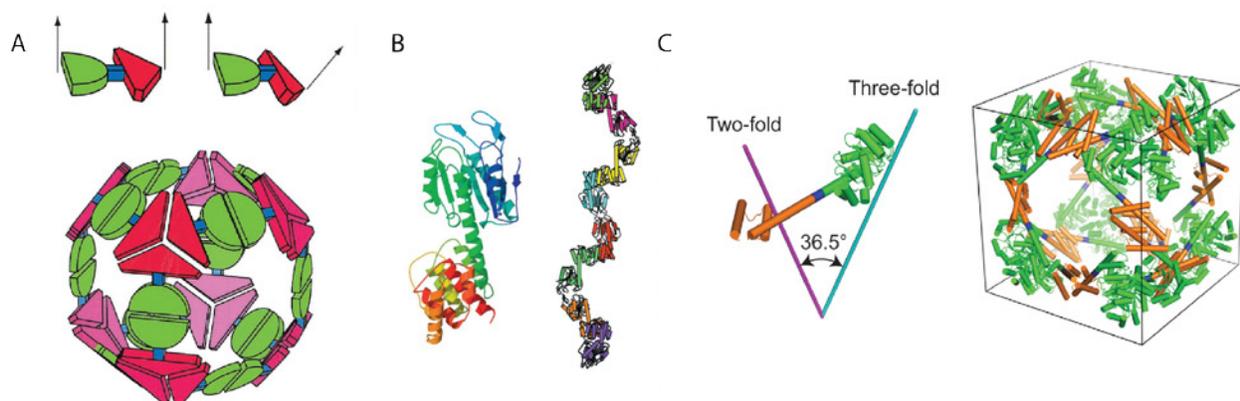
Figure 8. Examples of fusion proteins used in making nano architectures. Early work (a and b) focused exploring design elements such as protein identity and helical linker impact. More recent work (c) focuses on designing predictable architectures, such as cubes, based on the symmetry elements in the fusion protein. Figure adapted with permission from Padilla, J. E.; *et al.*, *Proc. Nat. Acad. Sci. U.S.A.,* **2001**, *98*, 2217-2221. (c) adapted with permission from Macmillan Publishers LTD: NATURE CHEMISTRY Lai, T.; *et al.*, *Nat. Chem.*, **2014**, *6*, 1065-1071. Copyright 2014.

.

A more computationally demanding design of protein nanostructures has been developed that does not require fusion proteins but instead re-engineers two oligomer forming proteins to dock to one another. In this approach, natural oligomeric proteins with the desired intrinsic symmetry are mutated to create the specific building blocks required for the nanoshape by

designing new protein-protein interfaces (PPIs) between them.[37] The oligomer or oligomer sets chosen are guided by the desired architecture. The amino acids that will compose the new PPIs are computationally determined by first aligning the new interfaces based on their topology. The models used for this step contain only up to the C_β carbon of the side chains so the fit is not
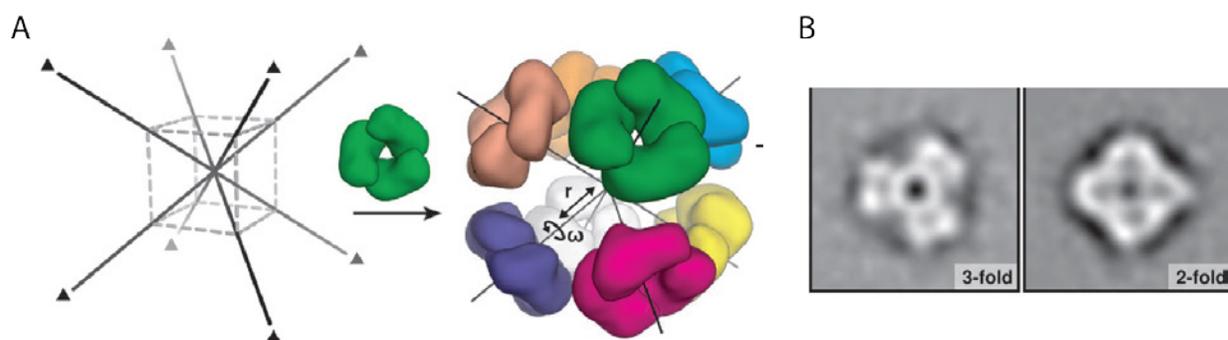
Figure 9. Nano-cubes made of a single, interface-engineered trimeric protein.
The alignment of the 3-fold symmetry needed to form a cube (a). TEM images of the 2- and 3-fold symmetry in the designed nano-cubes (b). Adapted from King, N. P.; *et al.*, *Science*, **2012**, *336*, 1171-1174. Reprinted with permission from AAAS.

biased by the identity of the amino acids. The amino acids side chains at the new PPI are designed back in to create a low energy PPI that also drives folding. This method has successfully created discrete nanostructures with one and two sets of protein oligomers (Figure 9).[38]

Protein-based material formation is not limited to natural oligomer forming sequences and the generation of new PPIs; other non-covalent interactions have been utilized.[33] Protein-ligand interactions, such as biotin and strepavidin, have been used as the driving force behind self-assembling biomaterials. A C_4 symmetric adolase was functionalized with a single biotin

molecule, allowing for an interaction with streptavidin to form cross-like structures.[27] Further functionalization of the adolase and streptavidin with biotin created supramolecular assemblies that formed in two dimensions (Figure 10). Combination of other non-covalent interactions have
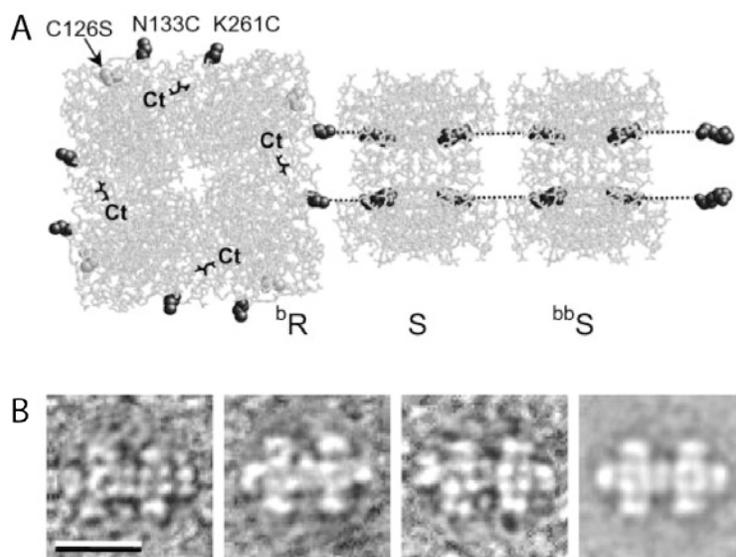
Figure 10. Biomaterials formed by protein-ligand interactions.

A tetrameric adolase with 4-fold symmetry was modified with biotin molecules (a). The adolase binds strepavidin to form cross-like structures. Strepavidin was also modified to contain biotin, allowing for grid like architectures to be achieved (b). Figure adapted from Ringler, P.; Schulz, G. E.; *Science*, **2003**, *302*, 106-109. Reprinted with permission from AAAS.

also allowed for the formation of protein arrays. Protein crystals combined designed hydrophobic and electrostatic interactions, His-metal binding, and small molecule dimerization in a protein to form highly ordered arrays that were controlled by the strength and types interactions present.[39]

This allowed for some control over the shape and mechanism of formation by only changing solution conditions such as pH and metal concentration or identity.

One of the most extensively studied and interesting materials that have bridged the gap between supramolecular polymers and biomaterials are hydrogels. Hydrogels are entangled polymer networks that respond to external stimuli, such as temperature[40], pH[41], ion concentration change[42] and light[43], by changing volume. Hydrogels find use in many common products such as diapers, contact lenses, drug delivery systems, and agriculture.[44] However, very few of these compounds are biocompatible or biodegradeable. Protein based hydrogel polymers have been utilized to address this problem. A synthetic peptide was designed to form a hydrogel with two helix regions separated by an Ala/Gly rich region.[45] The C-terminus of the peptide contained a Cys residue as a means to link two of these peptides together through a disulfide bond. Driven by the hydrophobic effect, the two helices form an intra- or interchain coiled coil. The close proximity and two-point interaction (covalent disulfide bond and non-covalent coiled-coil interaction) in these chains caused the formation of a hydrogel that is responsive to both temperature and pH (Figure 11). The sequence of one of the helices, which originally had a high percentage of acidic residues, was mutated to contain a higher amount of basic residues. The increased electrostatic interaction between the acidic and basic residues in the coiled coil raised the sol-gel transition to a higher temperature. This is an early and simple example of how coiled coil topology can be tuned to control material properties. The topology and use of coiled-coil motifs in materials will be discussed extensively in the next section.
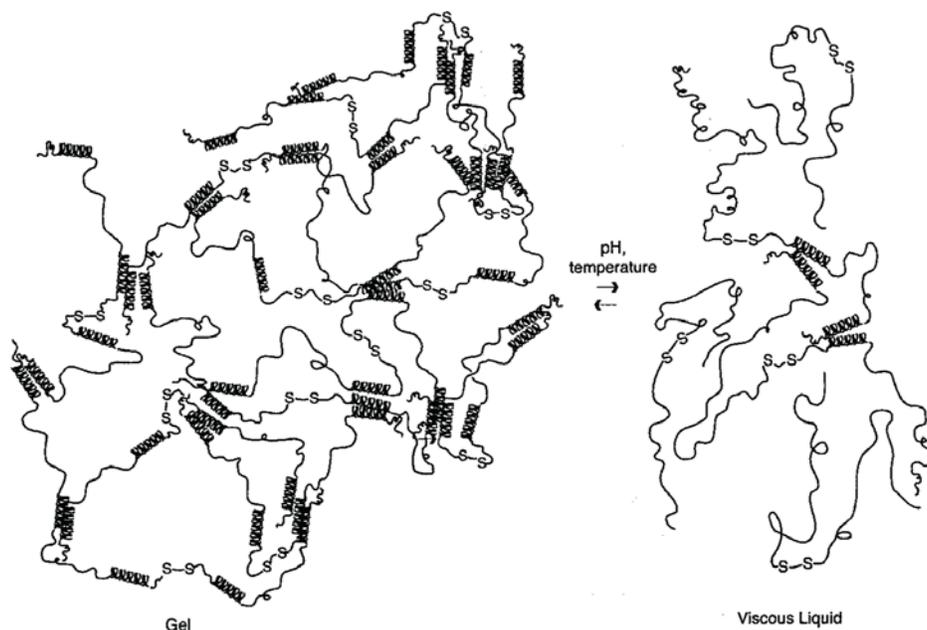
Figure 11. Temperature and pH responsive coiled-coil based hydrogels.

The coiled-coil can associate intra- or interchain to create a gel. The viscosity of the solution was responsive to pH and temperature change. Adapted from Petka, W. A., *et al.*, *Science,* **1998**, 281, 389-392. Reprinted with permission from AAAS.

The use of proteins and peptides opens up many new possibilities in tunable and functionalized materials. The diverse array of amino acids and available secondary, tertiary, and quaternary structures provide a palette of starting points for materials and design. They also offer the ability to add a variety of functional groups through selective amino acid reactions. Yet, there are still obstacles facing protein- and peptide-based materials. As mentioned earlier, natural proteins provide an excellent starting point for material synthesis; however, modifications to natural sequences may not be tolerated. Extensive modification, including the design of new PPIs, is computationally demanding and there is no guarantee the designed protein will fold as projected. An additional limitation with many of the discussed examples is synthetic accessibility—most are

17

large proteins that must be expressed. The variety and efficiency of unnatural amino acid incorporation via recombinant methods has improved over the past decade but insertion of multiple unnatural amino acids is still not a trivial task. Smaller and *de novo* peptides have found some use in biomaterials; however, their shorter sequence allows for less sequence variation without disruption of the desired fold. There remains an unmet need for a synthetically accessible, yet tunable, protein scaffold for material design.

1.3   COILED COILS AND THEIR USE IN MATERIALS

Coiled-coil proteins are coded for by approximately 3% of protein-coding genes[46] and are described as two or more α-helices whose folding is driven by the burial of a hydrophobic core.[47] Coiled-coil peptides vary greatly in length with the shortest stable motifs consisting of chains with ~21 residues. This motif also exhibits a wide range of tunable properties, such as stability and oligomerization state[48, 49, 50], that are based on modifications to a simple heptad repeat. The majority of the work determining the properties of coiled-coil folding has been performed by modifying naturally occurring sequences[51, 52, 53, 54]; however, a number of *de novo* coiled coils have been designed[55, 56] and used in materials.[57, 58, 59, 60] The coiled-coil motif provides an interesting scaffold to use in the creation of supramolecular materials: its folding is dependent on a non-covalent interaction, providing an assembly driving force, and modifications to sequence can both control the degree of polymerization and introduce functionality.

The coiled-coil motif is generally described as some number of greater than two α-helices whose folding is driven by the formation of a hydrophobic core. To discuss the sequence-structure relationship in coiled coils, the simplest, dimeric case will be used. As stated before, coiled-coil

peptides derive many of their interesting properties from the placement of certain types of residues within the 7-residue heptad repeat, designated in a helical wheel by *abcdefg*, of the α-helix (Figure 12). The heptad repeat signifies the number of residues required to return to directly below the original residue position on the helix. The heptad repeat also creates a residue pattern on the α-helix: the residues at the *a* and *d* positions line the hydrophobic core; *e* and *g* residues flank the hydrophobic core; and the *b*, *c* and *f* residues are solvent exposed.

Figure 12. Helical wheel diagrams of dimeric, trimeric and tetrameric coiled coils.

This diagram shows the relative positions of the amino acids in the heptad repeat. In the dimeric coiled-coil, the *a* and *d* positions are buried in the hydrophobic core, and the *e* and *g* positions are still in close enough proximity to interact and influence the orientation of the assembly. In the higher the oligomerization states, the *e* and *g* residues

are closer in proximity. The positions are labeled as *a*-orange, *b*-yellow, *c*-green, *d*-blue, *e*-purple, *f*-grey, *g*-red; the hydrophobic-core residues are bolded.

Arguably, the most influential residue positions to coiled-coil folding are *a* and *d* due to their impact on stability, oligomerization state and orientation.[48] As the hydrophobicity of the residues in these two positions increases, the stability of the coiled coil also tends to increase. GCN4-p1, which is well studied dimeric coiled coil found in a yeast DNA transcription factor[61], has Val in all of the *a* positions (except for an Asn at position 16) and Leu in all of the *d* positions. When all of the *a* residues were mutated to Ile (one additional methylene group), the stability of the peptide increased from 53°C to >100°C while remaining dimeric (Figure 12).

The *a/d* positions also influence the oligomerization state of the coiled coil. Crick proposed a "knobs in holes" packing geometry for parallel dimeric coiled coils.[62] The packing geometry for the parallel dimer places the *a* residues oriented slightly out of the hydrophobic core (parallel packing) and the *d* residues are pointed into the core (perpendicular packing). Parallel geometry discriminates little for the type of side chains placed at *a*, but bulky and β-branched residues are disfavored at the perpendicular packing *d* positions.[49] In tetramers, the packing geometries are reversed, favoring residues such as Ile at *d* and Leu at *a*. Trimers exhibit a third type of packing for both the a and d layers in which the C_α-C_β's of the residues in form an ~60° angle (acute packing) to one another (Figure 13a).
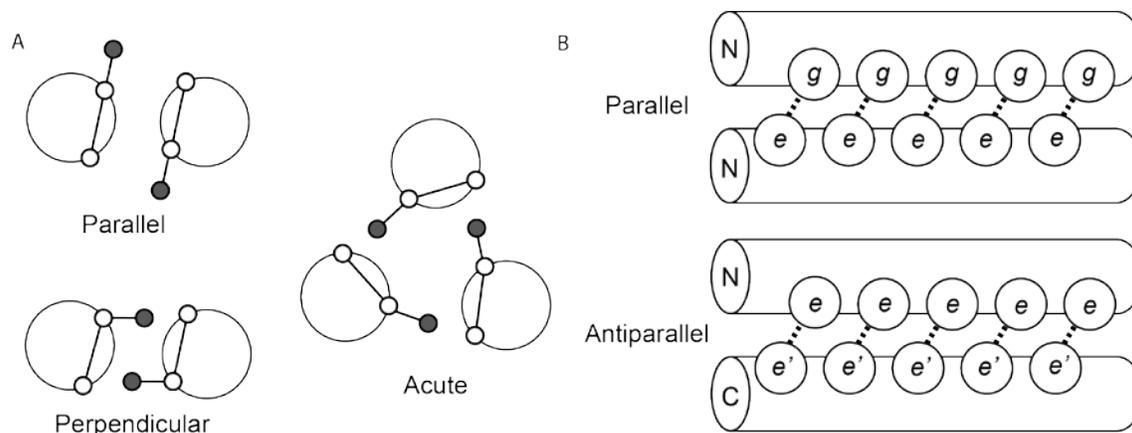
Figure 13. Common core packing arrangements of coiled-coil sequences

Parallel packing (a), which orients the vector along C_α-C_β slightly out of the hydrophobic core, occurs at the *a* position in dimers and *d* position in tetramers. Perpendicular packing orients the vector along C_α-C_β slightly into the hydrophobic core. This packing arrangement occurs at the *d* position in dimers and the *a* position in tetramers. The packing geometries help to explain the preference for certain residue types in the *a* and *d* positions. Acute packing occurs in trimers with the the vector along C_α-C_β at a 60˚ angle with the peptide backbone of the other helices. The *e* and *g* positions align differently depending on the orientation of the helices. (b) The charge pairs of these residues can help to stabilize the overall fold of the coil.

While rules for determining oligomerization state provide a good starting template for designing coiled coils, they are not absolute. Hydrophilic residues placed at core positions can influence oligomerization state preference but at the expense of stability. GCN4-pVL, a GCN4-p1 point mutant with Asn_{16}→Val, has a melting temperature over 100˚C but does not exhibit a favored oligomerization state.[48] In general, the placement of Asn at *a* favors dimers, but decreases the stability of the coiled-coil fold. The residues at the *e* and *g* positions can help to stabilize the preferred oligomerization state by forming polar contacts across helices but do not typically have enough influence to significantly alter the oligomerization state.

The *e* and *g* positions play minor roles in oligomerization state determination, but demonstrate a larger role in the orientation of the helices. In a parallel dimer, the *e* residue of one helix can interact with the *g* residue of the other helix; whereas the *e* (or *g*) residues of both helices interact in antiparallel dimers and tetramers[63] (Figure 13b). Thus, the orientation of the helices can be influenced by strategically placing charged residues at the *e/g* positions of the coiled coil. Examples of designed antiparallel coiled coils are the EZ and KZ sequences.[53] The heterotetrameric coiled coil places Glu at the *e* and *g* positions of the EZ sequence and Lys at the *e* and *g* positions in the KZ sequence causing the sequences to preferentially adopt an antiparallel tetrameric arrangement with respect to one another to form fibers (Figure 14). The other three.
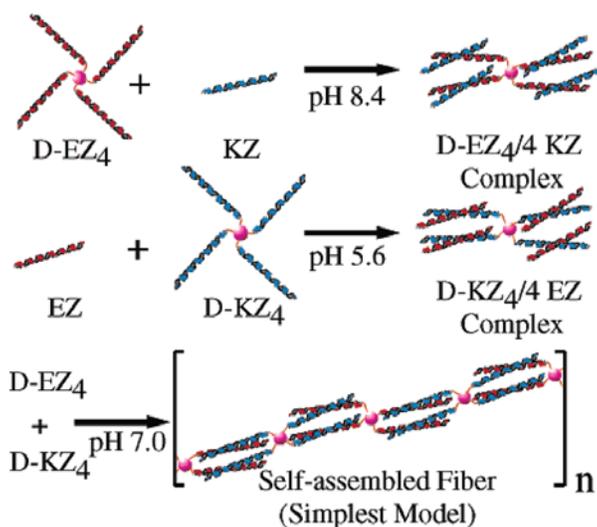
Figure 14. High-aspect ratio fibers formed by coiled coil hubs.

An antiparallel heterotetramer is formed by the acidic EZ sequence and the basic KZ sequence. The ends of the coiled coils can be linked together with a tetrameric linker to form high aspect ratio fibers. Figure adapted from with permission from Zhou, M., *et al*, *J. Am. Chem. Soc*. **2003**, *126*, 734-735. Copyright 2003 American Chemical Society.

22

positions in the helical wheel (*b, c* and *f*) are solvent exposed in dimers but can have some interactions in larger oligomers. The solvent exposed positions can be used to promote interhelical association by non-covalent interactions or add functionality through covalent attachment. Using these design strategies, a wide range of coiled-coil topologies with varying lengths, stabilities, oligomerization states and orientations have been prepared and characterized.

The many options to tune the properties of coiled coils with sequence modification makes this motif an excellent starting point for constructing dynamic, tunable materials. Initial efforts in using coiled coils in materials focused on the creation of hydrogels[45] (discussed in section 1.2) or elongated fibers with the coiled coils interacting in an end-to-end fashion.[53] The first example of creating elongated, coiled-coil fibers used a 21 residue *de novo* tetrameric coiled coil that associated at the termini through electrostatic interactions to form fibrils and then the fibrils interacted to form larger fibers (Figure 15a).[64] Recent work using similar end-to-end interactions and a series of coiled coils with oligomerization states up to seven has created peptide nanofibers and porous nanotubes.[65] Careful engineering of the solvent-exposed sites on the coiled coils allowed for the single nanotubes to associate and form larger bundles of fibers (Figure 15b).[66] In addition to end-to-end coiled-coil assemblies, work has been done using hanging-end coils in the formation of fibers.[23, 24] The use of coiled-coil ropes and fibers has gone beyond exploration of design principles to applications such as scaffolds for tissue engineering and assembly of hybrid materials.[67]
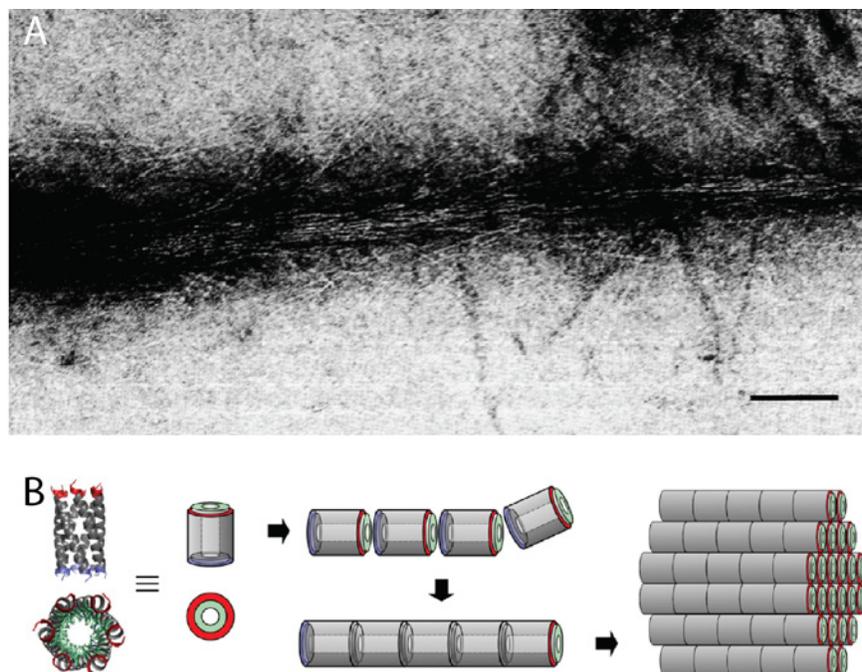
Figure 15. TEM image of the fibers formed from one of the first published *de novo* tetrameric coiled coils.

The coiled coils were proposed to associate through the charges at the N- and C-termini (a). Recent work used

similar end-to-end association to design nanotubes and fibers made from coiled coils with oligomerization states up

to seven. Figure (a) adapted with permission from the Proceedings of the Japan Academy, Ser B.: Kojima, S.; *et al.*,

Proc. Japan Acad., **1997**, *73*, 7-11. Figure (b) reprinted with permission from Burgess, N. C.; *et al.*, *J. Am. Chem.*

Soc., **2015**, *137*, 10554-10562. Copyright 2015 American Chemical Society.

The use of coiled coil peptides in materials have moved into the design of more complex architectures. Net, cage[60] and polyhedra[68] coiled-coil designs have great potential to be used in applications such as drug delivery, catalysis and directed nanoparticle synthesis.[7] Many different architectures have been reported in the very recent literature. A combination of homo- and heteromeric coiled coil dimers and trimers linked by a disulfide bond formed a self-assembling coiled-coil cage (Figure 16).[60] A pyramidal polyhedron consisting of six different coiled-coil dimers in a single chain has also been characterized.[36] These architectures, while excellent

examples of the versatile roles coiled coils can assume in materials, have limited accessibility to modification without either a major input of computational design or structural characterization. The design of a coiled-coil material that is modular and tunable would address this problem and expand the scope of coiled-coil based materials.
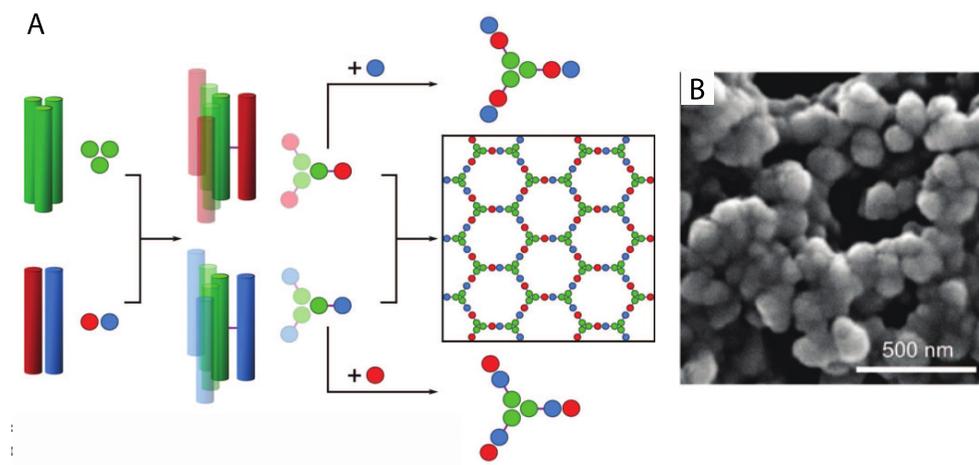
Figure 16. Protein cage assemblies driven by coiled-coil self association.

Two sets of coiled coils, a homomeric trimer and heteromeric dimer, were linked at a solvent exposed site by a disulfide bond (a). This mixture could be capped with the complimentary dimeric coil or form a lattice when the two different cross-linked species were mixed in aqueous solutions. The placement of the disulfide linkage gave rise to cage like assemblies in solution as seen by SEM (b). From Fletcher, J. M.; *et al.*, *Science*, **2012**, *340*, 595-599. Reprinted with permission from AAAS.

Coiled coils offer a unique scaffold for materials with a variety of properties that are tunable by sequence adjustments. The ability to easily tune their association affinity and their cooperative folding, driven by the burial of a hydrophobic core, make them ideal candidates for

bridging the fields of protein-based materials and supramolecular polymers. Very few examples exist of coiled coils utilized for this purpose. Linking the coiled coils, either at their ends or a mid-point, creates a monomer subunit, and their self-association through burial of the hydrophobic core provides the driving force for polymerization. The association affinity of the coiled coils should remain constant with the polymer formation following an isodesmic association model. Because of this, the degree of polymerization could be modulated by sequence modifications. In addition to the potential to control size, the morphology could also be adjusted by mutations to the coiled coil sequence. While certainly feasible, *de novo* design of coiled coils would not be needed with the diverse palate of coiled coils already characterized in the literature.

1.4   PROJECT GOALS

The major challenge in working with proteins as a basis for materials is the balance between controlling the final folded structure and synthetic accessibility. Large protein domains often have well-defined and complex folds; because of the complex folding, however, modification the without disrupting the established folding pattern becomes challenging. Shorter sequences offer ease of synthesis and are amenable to diverse chemical modifications. These advantages come at the cost of predicting and controlling the overall fold of the system. When proteins have been successfully used to create materials, the ability to easily vary the components of the system to fine tune properties such as size and morphology are not generally designed into the system.

Here, we aim to develop a platform for preparing protein-based supramolecular biomaterials in which size and morphology can be controlled by changing by the properties of the monomeric building block. To exert fine control over these materials, we propose to build a subunit

26

consisting of two α-helical coiled coils linked at a solved exposed residue by an organic bridge. By using a modular synthetic approach, the interplay among linker length and rigidity, coiled-coil stability, assembly size, and morphology can be elucidated. In our design, the subunit can easily be altered to change the size and morphology of the material. Once these relationships are established, the impact of assembly size and flexibility on function can be tested by the addition of a functional moiety on to the coiled-coil scaffold.

1.4.1 Creation of a Modular Coiled-Coil Subunit

Early work in the use of coiled coils in materials focused on forming high aspect ratio fibers through either blunt or over hanging end-to-end interactions (see section 1.3). These materials formed as intended; however, the method of assembly either completely eliminated or significantly limited the possibility of modifying the size and morphology of the resulting material without redesigning the base unit. Our aim is to develop a modular synthetic subunit consisting of two α-helical coiled coils linked at a solvent exposed site by a small organic bridge that folds to form supramolecular polymers in aqueous solutions. The design of the subunit allows us to investigate how the linker and coiled-coil components of the subunit are translated to changes in the larger polymer.

We first developed a synthesis of the subunit and demonstrated the effects of linker identity on polymer properties. Three linkers were chosen to examine the effect of linker length and rigidity—a short disulfide bond and two linkers that were identical in length but differing in rigidity. Too short a linker created unfavorable interactions between the connected coiled coils, while longer linkers allowed the peptide dimers to fold independently. The polymers formed by the two longer linkers exhibited differences in apparent length corresponding to differences in the

rigidity of the linker. The impact of coiled-coil association affinity was examined with two previously published and four newly characterized sequences where mutations were restricted to residues in the hydrophobic core. The four variants were subjected to biophysical characterization assuring their solution behavior matched the two previously characterized peptides before being used in the subunit. The series of subunits with the hydrophobic core mutated coiled coils generally followed the assembly size trend predicted by a supramolecular association model showing assembly size can be modulated by small changes to the subunit. Our data suggest the design of a modular subunit consisting of coiled-coil forming peptides covalently linked by solvent exposed sites would allow for a variety of self-assembling architectures to be prepared and explored for a wide range of applications.

1.4.2   Oligomerization State Switching in GCN4-p1

Although the study of sequence-structure relationships in coiled coils has been active for close to twenty years, it remains a dynamic field of research that even now affords new insights. The dimeric yeast transcription factor coiled coil, GCN4-p1, was the first extensively studied coiled-coil sequence characterized at high resolution and has long been described as "well behaved" as it is exhibits a consistent oligomerization state, stability and orientation in a variety of solution conditions. Following up on the observation of a trimeric crystal form of a designed hydrophobic core mutant, we explored the possibility that the native GCN4-p1 sequence could also adopt a trimeric oligomerization state. Crystallization efforts yielded structures of GCN4-p1 in both dimeric and trimeric assemblies. Solution phase studies showed mixed results for oligomerization state depending on the method and solution conditions. Molecular dynamic simulations implied

the dimeric and trimeric folded states are similar in energy and the oligomerization state of GCN4-p1, and other coiled coils, is context dependent.

1.4.3   Towards Functional Materials

Our initial work focused on the design and synthesis of a modular subunit consisting of two α helical peptides connected through a small organic linker that fold to form larger assemblies in dilute aqueous buffer. We showed how the properties of subunit components—the linker and peptide sequence—impacted the properties of the larger assemblies. Our goal in this work was to examine if there is a synthetic route to add functionality to the assemblies and if the added functionality properties fluctuate with changes to the subunit. We chose to add functionality in the form of light harvesting groups. Light-harvesting energy transfer requires specific spacing of chromophores and fluorophores for high efficiency and we felt our self-assembled materials would provide an excellent scaffold for fluorophore placement. Energy transfer would be studied by placing a donor fluorophore on each of the peptides in the subunit and an acceptor fluorophore on an uncrosslinked peptide, which would be used as a capping group. This allows us to not only probe the energy transfer efficiency between the donor and acceptor, but to also look at how energy is delocalized across a series of donors in the assembly, called the antenna effect. We expect the differences in flexibility imparted by changing linkers could impact the energy transfer dynamics.

2.0   CREATION OF A MODULAR COILED-COIL SUBUNIT

Work detailed in this chapter has been published as:

Staples, J. K.; Oshaben, K. M.; Horne, W. S.; "A Modular Synthetic Platform for the Construction of Protein-Based Supramolecular Polymers via Coiled-Coil Self-Assembly." *Chemical Science*, **2012**, 3, 3387-3392

Oshaben, K. M.; Horne, W. S.; "Tuning Assembly Size in Peptide Based Supramolecular Polymers by Modulation of Subunit Assembly Size." *Biomacromolecules*, **2014**, 15, 1436-1442

In Nature, proteins and nucleic acids are used to create an array of complex supramolecular structures from simple building blocks[27, 30] and scientists look to these architectures for inspiration in the rational design of synthetic biomaterials.[69] While protein-based assemblies are more common in biological systems, nucleic acids have found greater use in designer biomaterials because of the well-characterized relationship between DNA sequence and folding.[28, 29, 70, 71] The sequence-folding relationship in proteins is complex and has few distinct rules, making designing well-defined architectures a challenging process. Naturally occurring proteins, while well folded, are not typically synthetically accessible and may not tolerate sequence mutations. Shorter peptides, with greater sequence control and ease of synthesis, may not have a defined fold. Despite these limitations, *de novo* designed proteins have given rise to hydrogels[45], disks[34], spheres[60] and high aspect-ratio fiber[52, 53] supramolecular assemblies.

The α-helical coiled coil, because of it's well-studied sequence-structure relationships and synthetic accessibility, is an excellent basis for the creation of novel biomaterials (discussed

extensively in section 1.3). Coiled-coil supramolecular assemblies have used primarily either end-to-end or overlapping interactions to drive folding.[64, 66] We envisioned an alternative approach based on a modular subunit, consisting of two α-helical coiled coils forming peptides attached at their midpoint by a small organic linker, that will fold in dilute aqueous conditions (Figure 17). The design for our subunit was inspired by the supramolecular polymer field. An issue many synthetic biomaterials suffer from is the inability to adjust the size or morphology with only minor changes to the monomeric building block. The size of the assemblies is generally controlled by the addition of capping agents and this method of size control may not be precise enough for applications where assembly size is critical such as light harvesting devices. In an isodesmic supramolecular polymer, the size of the assembly is a function of the concentration of the monomer in solution and the association affinity of the monomer.[18] A larger value for either the concentration or association affinity leads to larger assemblies. By tethering together two of the same coiled-coil forming peptides, we create an monomer with an A-A type interaction that should follow an isodesmic self-association model driven by the cooperative association and folding of the coiled-coil domains in the subunit. It is well documented that the folded stability of coiled coils, which is coupled to the association affinity, can be changed by mutations in the hydrophobic core.[48, 72] The oligomerization state can also be controlled with similar mutations. Changes in the oligomerization state and to the linker structure could be imparted as a means modify the morphology of the assemblies. Thus, we hypothesized that a library of subunits based on only a few key peptide sequences and linker structures would enable us to tune the material size and shape to the desired application.
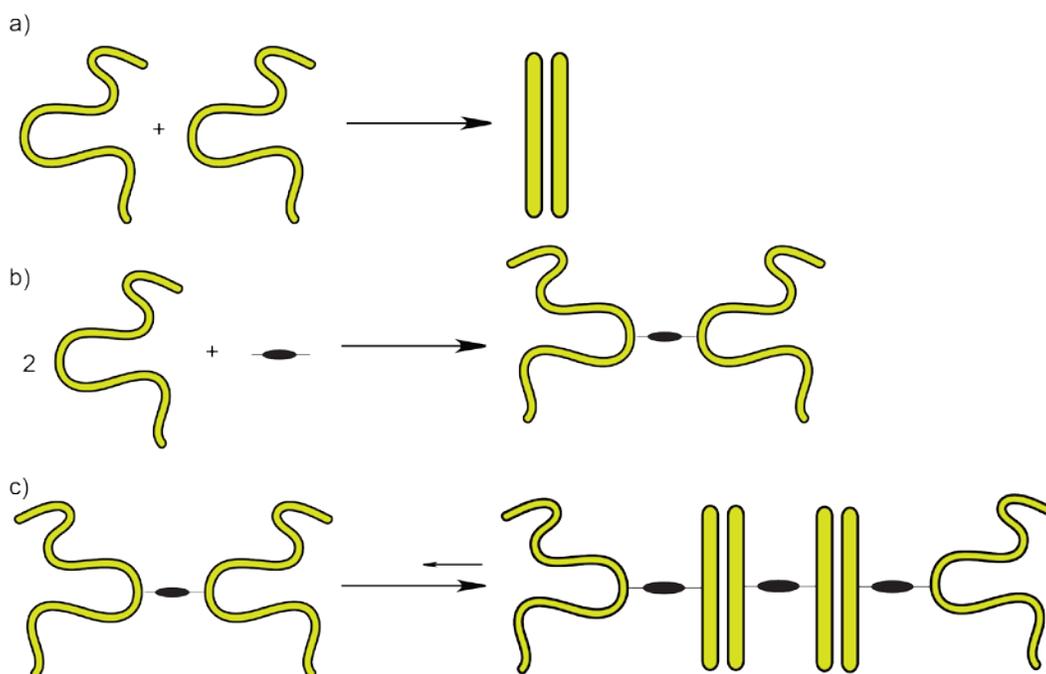
Figure 17. The creation of a modular, self-assembling subunit.

First, the self-associating coiled-coil peptide motif was chosen (a). Second, we envisioned a subunit where a small

organic molecule would be used to cross link the two chains (b). Finally, assembly formation would be driven by the

folding of the coiled coil to form supramolecular polymers (c).

2.1   DESIGN AND SYNTHESIS OF COILED-COIL CROSSLINKED SUBUNIT

Two known 33-residue coiled-coil sequences and three linking strategies were initially explored

to test the synthetic accessibility and solution folding properties of our designed subunit. The two

peptides, commonly known as GCN4-p1 (**1**) and GCN4-pIL (**2**), are derived from the yeast DNA

transcription factor GCN4 (Figure 18).[48] These two peptides differ by four of their *a*-position

residues in the hydrophobic core: GCN4-p1 has three Val and one Asn, while GCN4-pIL has four

Ile. While the folds of these two peptides are nearly identical, the subtle difference in the sequence

gives rise to a significant (>1.7 kcal/mol) decrease in the folded stability between **1** and **2**. Because

of the cooperative folding of coiled-coil peptides, folded stability is directly related to association affinity.
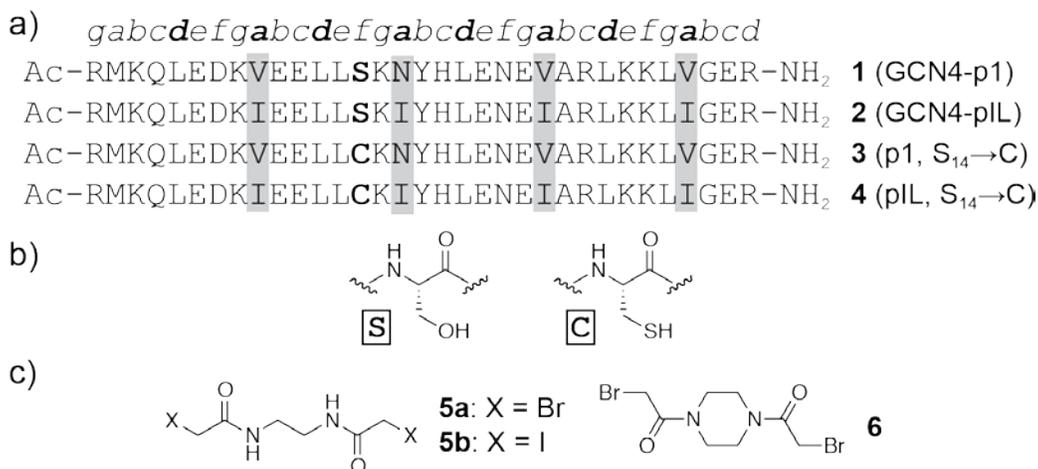
Figure 18. Peptide sequences and linker structures used in subunits. Sequences of peptides **1**–**4**. (a) Mutation of Ser to Cys inserts a functional handle for the formation of the cross-link (b). Structures of organic linkers **5a**, **5b**, and **6** (c).

Peptides **3** and **4** are analogues of peptides **1** and **2** respectively, with a Ser → Cys mutation at a solvent exposed site midway along the chain (Figure 18). The Cys, which is the only Cys in the sequence, provides a functional handle to connect the two helices through chemoselective ligation to create a subunit. As mentioned previously, coiled-coil peptides have been linked primarily through their termini to create self-assembling materials.[24, 53] Placing the cross-link at the mid-point of the coiled coil should afford us maximum control over assembly properties through modification of not only peptide sequence but also identity of the linking group. Peptides **1-4** were prepared by Fmoc solid phase peptide synthesis (SPPS), purified by preparative high

performance liquid chromatography (HPLC) and their identity confirmed by matrix assisted laser desorption ionization mass spectrometry (MALDI-MS). A table of masses can be found in the experimental section (Table 4).

Two linking strategies for subunit formation were examined involving the Cys thiol: oxidation of Cys to form a disulfide bond and nucleophilic addition to an α-haloacetamide to form a thioether. Three linkers with α-haloacetamides were synthesized (**5a**, **5b**, **6**) to probe the effect of linker length and rigidity on solution phase assembly properties and the effect of the halide leaving group on ligation reaction efficiency. We synthesized bis-bromoacetamide linkers, **5a** and **6**, from ethylenediamine and piperazine, and we also synthesized bis-iodoacetamide linker, **5b**, from ethylenediamine. Utilizing thiol reactive chemistry allowed us to develop a modular synthetic approach based on the chemoselective ligation of unprotected peptides in solution to prepare cross-linked subunits. This approach allows for the rapid combination of a handful of linkers and peptides to form a variety of subunits.

The ligation reactions were performed in dilute aqueous conditions with small amount of organic co-solvent (Figure 19). Oxidation of **3** in aqueous DMSO at room temperature yielded disulfide cross-linked subunit **7**. Peptides **3** and **4** were used in combination with linkers **5a**, **5b** and **6** to form thioether linked subunits **8**-**10**. Our initial reaction conditions for the formation of the thioether were focused on carefully controlling the ratio of peptide to linker by using a 2:1 peptide to linker ratio with a slight (5%) excess of peptide. If the ratio of peptide to linker varied too much from this ratio, undesired products would form. A large excess of peptide would lead
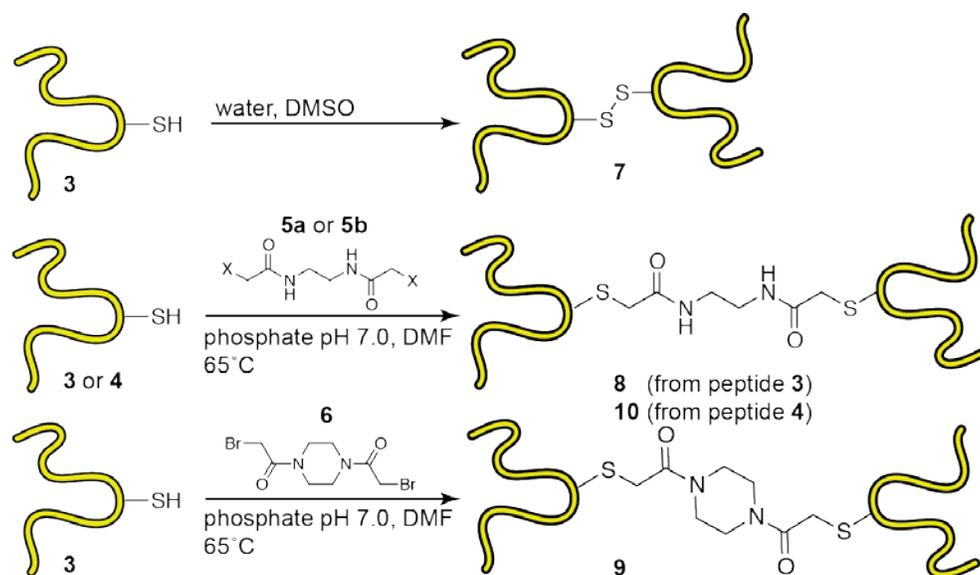
Figure 19. Synthesis of cross-linked subunits **7**–**10**.

Subunits **8** and **10** are based on the ethylenediamine linker and subunit **9** uses the piperazine based linker.

to the formation of disulfide-linked subunits and unreacted starting material. More problematic is having an excess of linker, which leads to a side product where a linker molecule forms only a single thioether bond (half reacted product—Figure 20, compound c). Based on the observed byproducts and the reasoning behind their formation, we hypothesized improved crude purities could be achieved if we added the linker in two aliquots. This approach should help to maintain a slight excess of peptide throughout the reaction. The original reaction conditions had 250 µM peptide in 25 mM phosphate buffer pH 7.0. One aliquot (0.5 equivalent) of linker in acetonitrile was added, and a second aliquot (0.5 equivalent) was added after 30 minutes. The reaction proceeded for 4 hours before being quenched and then analyzed by HPLC (Figure 20a). The
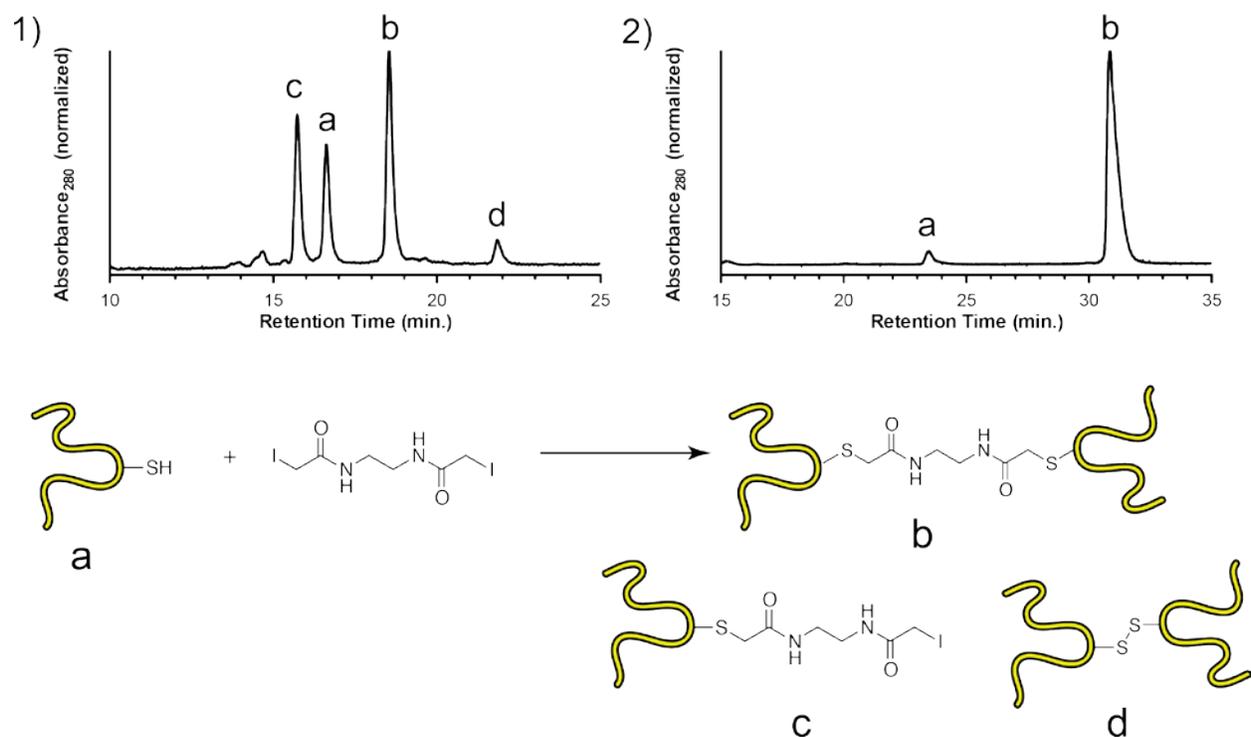
Figure 20. Optimization of the crosslinking reaction mixture.

Semi-preparative HPLC traces of the original cross-linking reaction (1) on a 30%-40% acetonitrile/0.01% TFA gradient and the optimized procedure (2) on a 30%-35% acetonitrile/0.01% TFA gradient. Below is the cross-linking reaction and products we observed. In the original reaction conditions, while there is a large product peak (b), there are also large amounts of linker half reacted with peptide (c) and unreacted peptide (a), and some evidence disulfide formation (d). The optimized conditions (2) show only a small amount of starting material (a) and a large product peak (b).

products of this method contained a mixture of unreacted starting material, subunit, disulfide linked subunit and half reacted product (Figure 20).

To optimize the cross-linking reaction further, we examined how the frequency and number of linker aliquot additions was well as the organic co-solvent affected product distribution. Dioxane, DMF, and THF were examined as alternatives to acetonitrile as co-solvents, but little to no difference was observed in the distribution of products. DMF was chosen moving forward as it led to slightly improved solubility of the linker over the other co-solvents and a distribution of products that favored the subunit. The number and frequency of linker addition was varied from a single aliquot to five aliquots added over an hour and a half. Larger volumes and less frequent linker additions led to the formation of undesired side products. We found that five aliquots of linker added at 15 minute intervals increased the amount of subunit formed while decreasing the amount of half reacted product. The cross-linking reaction still required several hours to complete. We heated the reaction mixture to 65°C to decrease the reaction time. The increase in reaction temperature also causes almost full denaturation of the coiled coil, which likely increases accessibility of the thiol group. These two factors were significant contributors to the increase in product formation and the decrease in reaction time to two hours.

The final optimized cross-linking reaction proceeded as follows: one aliquot (a fifth of the total volume) of linker was added to 150 µM peptide in 25 mM phosphate buffer pH 7.0 at 65°C. Four more aliquots of linker were added at 15 minute intervals over the next hour. After 2 hours, the reaction was quenched with a mixture of water and acetonitrile with 0.1% TFA. Following this protocol peptide **3** was reacted with linkers **5a** and **5b** to form subunit **8**. In these reactions, halide leaving group had some effect on reaction efficiency with the iodide linker (**5b**) showing slightly superior results to the bromide linker (**5a**). Peptide **3** was also reacted with linker **6** to form subunit **9**, and peptide **4** was reacted with linker **5b** to generate subunit **10**. Subunits **7-10** were purified using preparative HPLC, and the identity of each confirmed by MALDI-MS.

2.2   CHARACTERIZATION OF COILED-COIL SUBUNITS WITH VARYING LINKER LENGTHS AND COILED-COIL STABILITIES

The design of the subunit described in the previous section was motivated by the desire to easily control assembly properties. The modular synthetic route we developed enabled the combination of a small number of peptide and linker building blocks to prepare a variety of subunits. Working within that design, the series of subunits **7**-**10** were characterized to examine the effect of linker length and rigidity as well as coiled-coil stability on solution-phase assembly properties. To assess the impact of individual subunit components on the assembly behavior, we needed to examine the effect on folding and thermodynamic properties of the coiled-coil domain when placed in the context of a subunit and the formation, if any, and size of supramolecular assemblies in aqueous solution.

Circular dichroism (CD) spectroscopy was used to investigate the folded conformation and thermodynamic stability of the coiled-coil domains in each subunit. All experiments were carried out with 100 µM subunit in 10 mM HEPES pH 7.0. The subunits (**7**, **8** and **9**) were first compared to the folding behavior of peptide **3**, the native-like GCN4-p1 Ser→Cys mutant (Figure 21). CD scans of the subunits derived from that sequence (**7**, **8** and **9**) retained similar α-helical content to peptide **3** (indicated by the magnitude of the negative peaks at 208 nm and 222 nm). Thermal melts monitored by CD show that subunits **8** and **9** have a cooperative unfolding transition with a melting temperature (T_m) of ~60˚C, which is similar to peptide **3**. Disulfide linked subunit **7** exhibited a very shallow unfolding transition and a T_m ~20˚C lower than subunits **8** and **9**. The lower T_m
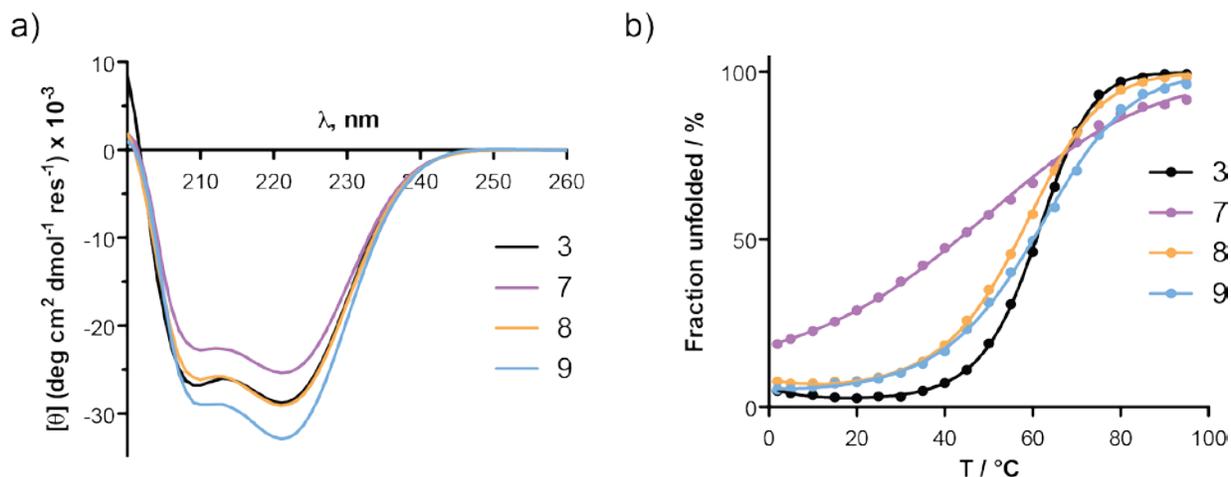
Figure 21. Biophysical characterization of the subunits using circular dichroism. CD scans at 25 °C (a) and thermal melts (b) for peptide **3** and cross-linked subunits **7**, **8**, and **9**. All samples are 100 µM concentration in 10 mM HEPES, pH 7.0. Each thermal melt was fit to a two-state unfolding transition to obtain melting temperature (T_m) values.

indicates a decrease in the stability of the coiled-coil folded state. We hypothesize the short disulfide linkage, which has only 4 atoms between the peptide backbone C_α's, causes crowding of the adjacent chains in the folded state.[73] This crowding allows for one of the α-helices to fold but disfavors the folding of the second helix in the subunit. The linkers used in subunits **8** and **9** each have 12 atoms between C_α separating the peptide backbones. The extra distance between the backbones restores the independent folding of the two α-helices in the subunit. Due to the unfavorable interactions and destabilization introduced by the short disulfide linkage, subunit **7** was not characterized further.
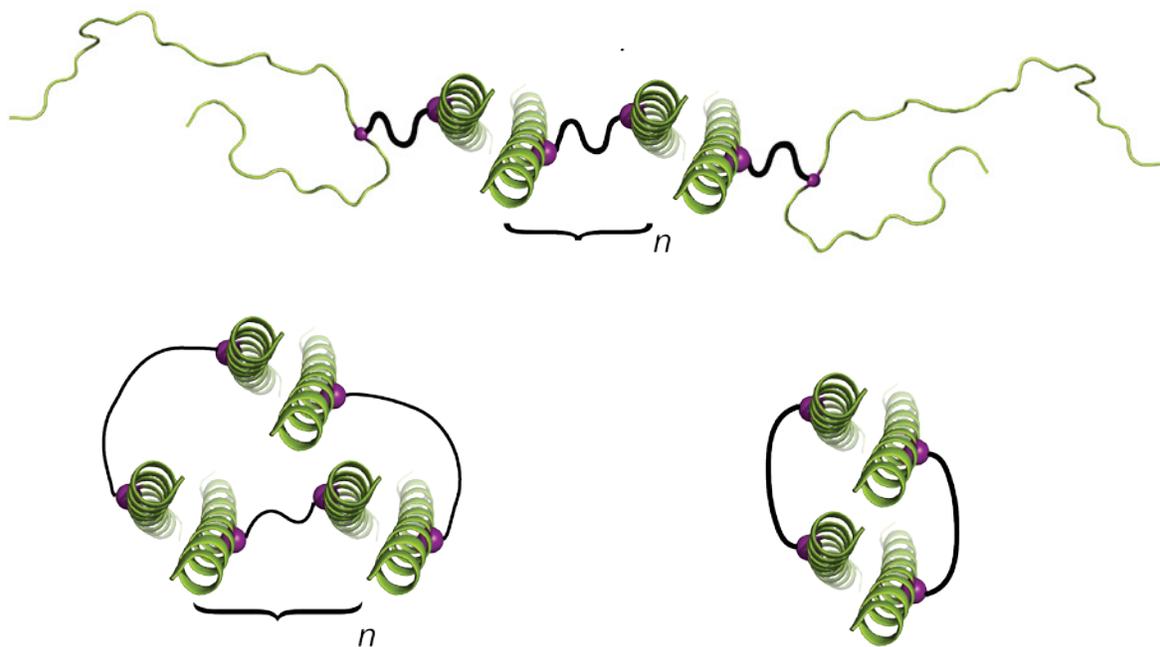
Figure 22. Possible assemblies formed by the cross-linked subunit.

The top image is the desired linear assembly. The bottom left is a cyclic assembly that could be formed by the

growing chain associating with itself. The bottom right image is the smallest cyclic assembly that could form.

However, this assembly should be disfavored due to the short linker length.

The coupling of coiled-coil folding and association was one of the key qualities in choosing

this motif for use in a supramolecular polymer.[61] The observation of a stable, α-helical structure

by CD indicates the formation of the coiled coil in solution. It should be noted, however, that due

to the homomeric nature of the coiled-coil interaction, several assembly possibilities arise that

include not only the predicted linear assembly, but also cyclic variants formed by one end of the

growing polymer associating with the other end of the same chain (Figure 22). The formation of

small (2 subunit) cyclic assemblies should be disfavored due to the linkers being too short to form

an intramolecular bridge to accommodate well-folded coiled coils.

To measure the size of the supramolecular assemblies, we employed a combination of gel permeation chromatography (GPC) and dynamic light scattering (DLS). Subunits **8**, **9** and **10** were subjected to GPC at a concentration of 100 μM subunit and eluted from a Superdex 200 column at room temperature with 10 mM HEPES pH 7.0 with 150 mM NaCl. The column was calibrated using protein molecular weight standards. Peptide **1** was analyzed as a control and eluted at a volume that corresponds to the expected dimeric fold. Subunits **8** and **9** eluted as supramolecular assemblies with a distribution of sizes (Figure 23). The number of subunits in the assemblies for each sample varied from 1-50 with the maximum UV absorbance occurring between 15-20 units. The exact value of these numbers may be skewed as the assemblies are in an equilibrium under elution conditions and could be subject to dilution effects. Moreover, the protein standards used for the calibration curve are globular as opposed to our expected oblong an assemblies. In terms of supramolecular polymer classification, these subunits function as A-A type monomers. Assuming an isodesmic association model, the number of subunits in the assembly would be analogous to the degree of polymerization (n).[10] Additional polymer properties were calculated from the GPC chromatographs (Table 1).
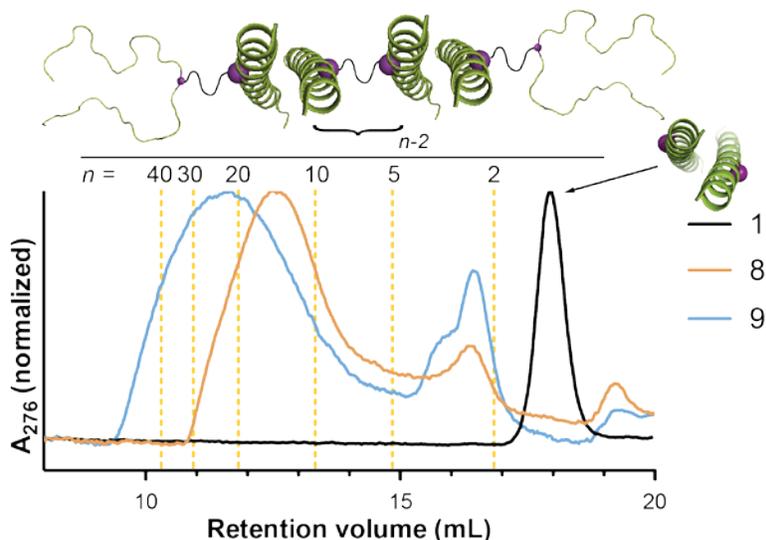
Figure 23. Gel permeation chromatography data for peptide **1** and cross-linked subunits **8** and **9**. Each injection consisted of a 100 µL sample of 100 µM concentration peptide eluted at room temperature with 50 mM HEPES, 150 mM NaCl, pH 7.0. Dotted lines indicate the predicted elution volumes for molecular weights corresponding to the indicated n values based on calibration of the column to protein molecular weight standards.

Table 1. Properties of supramolecular assemblies formed by cross-linked subunits **8** and **9**

Subunit	n [a]	M_n [a] (Da)	M_w [a] (Da)	PDI[a]
8	15	94000	112000	1.19
9	21	132000	182000	1.38

a Average degree of polymerization (n), number average molecular weight (M_n), weight average molecular weight (M_w), and polydispersity index (PDI) were determined for the major peaks in the GPC chromatograms shown in Figure 22.

Subunits **8** and **9** differ by only two atoms in the linker, and each has the same number of atoms separating the peptide backbones; however, there are significant differences in the observed assemblies in solution. Subunit **9** appears to give rise to assemblies that have more subunits than **8**. The only difference in these 8 kDa macromolecules are the three additional degrees of torsional freedom in the ethylenediamine-based linker versus the piperazine-based linker. An isodesmic self-association model (which would apply to this system) predicts only the association affinity of

42

the monomer unit (our designed subunit) would change the assembly size. The expected value of n is 14 given the association affinity of GCN4-p1 reported in the literature ($K_D = 0.48$ µM).[74] The GPC data for subunit **8** suggest $n = 15$, while for subunit **9**, $n = 21$, which is larger than predicted by this model. We initially hypothesized that the rigidity imparted by using a cyclic structure in the linker of the subunit affects the assembly properties. The rationale for this hypothesis is based on extrapolating that rigidity across the entire assembly, giving rise to an apparently larger polymer chain. In later work by a colleague in the group, in collaboration with the Saxena lab at the University of Pittsburgh, this hypothesis was rigorously tested with double electron electron resonance (DEER) spectroscopy and molecular dynamics simulations to explore the impact of linker rigidity on supramolecular assembly size. Data from this study supports the idea that the small amount of additional rigidity in subunit **9** could account for the apparently larger assembly size observed in the GPC profile (Figure 24).[75]

Recall, we also synthesized subunit **10**, which is a combination of the ethylenediamine based linker and the high-stability GCN4-p1 variant GCN4-pIL. When this sample was injected on the GPC, nothing eluted off the column. This sequence may be more aggregation prone in the context of our subunit due to its high folded stability causing longer assemblies that become entangled in one another. As discussed previously, the association affinity, which is related to the
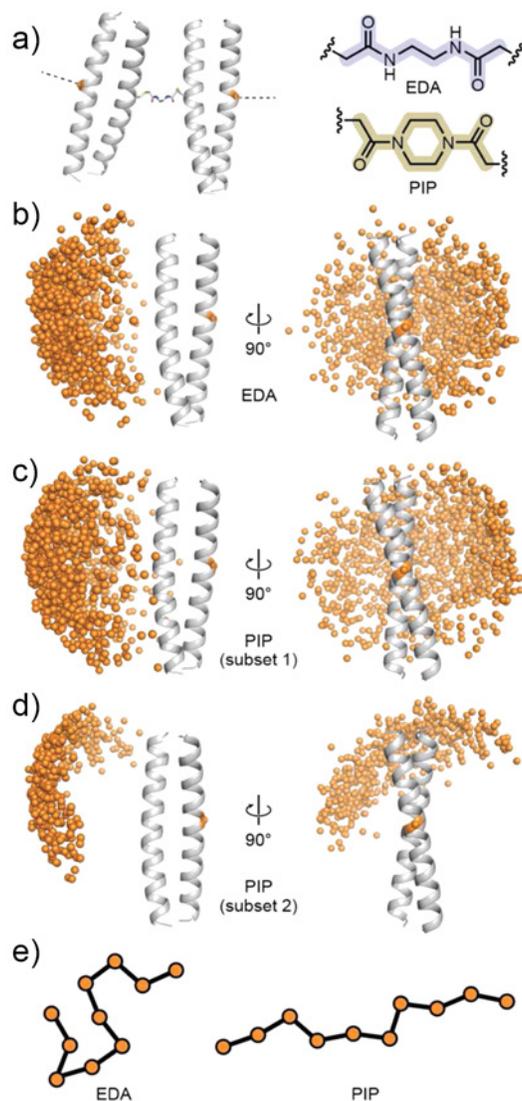
Figure 24. Model of a capped subunit with Ser_{14} C_α on the outer helices shown as orange spheres (a).

Based on MD simulations, predicted location of the next Ser_{14} C_α for a subunit with an EDA linker (b) or a PIP

linker (c, d). For each panel, models are superimposed based on one of the two coiled coils, and the position of the

second coiled coil is indicated by a single orange sphere, which is would be the next subunit in the growing chain.

(e) Simplified schematic showing different chain stiffness in the supramolecular polymers based on the EDA versus

PIP linker. Figure reprinted with permission from Tavenor, N. A., *et al*, *J. Phys. Chem. B.,* **2014**, *118*, 9881-9889.

Copyright 2014 American Chemical Society.

folded stability in coiled coils, in an isodesmic supramolecular polymer controls the degree of polymerization. We explored the roll of coiled coil association affinity more fully in subsequent work (see section 2.4).

The use of GPC to determine assembly size in dynamic systems has limitations. The subunits in this system are in a dynamic equilibrium under the elution conditions used and can be prone to dilution effects while on the column. The use of globular protein standards for the molecular weight calibration curve may also skew the data. We employed DLS as a complimentary method to GPC to measure the size of the assemblies in solution. While DLS has its own set of simplifying assumptions about particle shape and behavior, the data from this experiment would allow us to qualitatively compare the sizes subunits **8** and **9** without inference from dilution effects. We obtained DLS data for subunits **8** and **9** at 200 µM in 50 mM HEPES at pH 7.0 with 150 mM NaCl. The results showed qualitative agreement with our GPC data. Subunit **9** with the piperazine based linker gave rise to particles with a larger apparent hydrodynamic radius (R_H = 41 nm) than subunit **8** with the more flexible ethylenediamine based linker (R_H = 25 nm). The polydispersity of both samples was high (PDI = 0.45-0.49), which also agrees with the broad distributions seen in the GPC chromatographs. DLS also allowed us to gather some solution phase data for subunit **10**. The assembly size in **10** was larger than the other two subunits (R_H = 140 nm); however, interpretation of this data should be approached with caution as the potential presence of aggregation prone species complicates DLS data analysis.

The GPC chromatographs of the subunits, in addition to the broad peaks of the large assemblies, have smaller peaks corresponding to assemblies of 2-4 subunits. We offer two possible explanations for this observation. First, the possibility exists that the subunits are forming cyclic oligomers by one end of the growing chain associating with the other end (see Figure 22). The

45

formation of cyclic species is reasonable based on the mechanism of self-assembly. Subunit **8** would be expected to more readily form small, cyclic assemblies based on the flexibility of the linker; however, subunit **9** forms larger amounts of the smaller assemblies. Subunit **9** has two propagation lengths depending on the linker conformation.[75] It is possible that once smaller assemblies form with subunit **9**, the confirmation of the more rigid linker traps the subunits in a cyclic assembly. The second possible reason for the appearance of the smaller assemblies is the presence of impurities in the form of peptide **3** or peptide **3** with half reacted linker (see Figure 20). While the samples are >95% pure by analytical HPLC, even trace amounts of impurities could act as a capping agent to the growing supramolecular assembly chain. We performed a series of experiments to test this hypothesis.

The CD data from subunit **8** suggests that the coiled-coil domains fold independently of one another. If some amount of peptide **1** was added to a sample of subunit **8**, peptide **1** could associate with the assembly chain and act as a capping group by stopping assembly growth. As the amount of **1** in the mixture is increased, the size of the assemblies would continue to decrease. We carried out a series of mixing experiments with **8** and **1** (keeping the total concentration of peptide constant). The GPC data confirm our expectations, showing a systematic change in size distribution towards smaller assemblies with increased amount of peptide **1**. This data also raise the possibility of mixing as a potentially useful tool in controlling assembly size and properties.
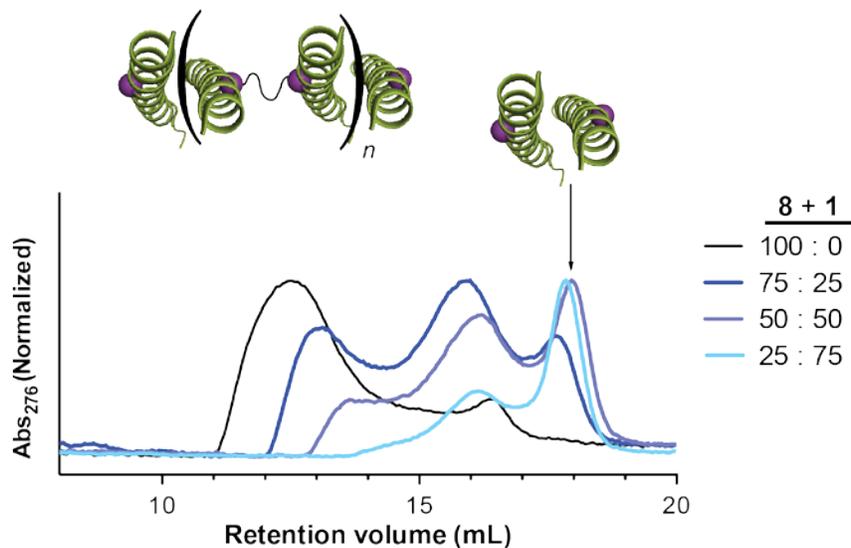
Figure 25. Gel permeation chromatography data for peptide **1** mixed with **8**.

This data shows the potential ability to control assembly size by mixing cross-linked subunits and capping agents

different ratios

2.3 DESIGN OF COILED-COIL SEQUENCES WITH A RANGE OF FOLDED STABILITIES

One motivation for using the coiled-coil motif in our subunits is its cooperative folding mechanism. Coiled coils fold through the burial of a hydrophobic core, and the burial of this core drives the formation of the quaternary structure. The coupling of folding and self-association in this motif allows us to use easily obtainable thermal stability data as a comparison between coiled coil association affinities. In our pursuit of a modular subunit that affords us control over the size and morphology of the resulting supramolecular polymer, we discovered a potential limitation with subunits containing the very thermally stable GCN4-pIL sequences is their propensity to aggregate in an uncontrolled manner. The two coiled-coil sequences we used

to examine the impact of coiled-coil stability on degree of polymerization had a difference in folded stability of greater than 2 kcal/mol. Part of the motivation for choosing the GCN4-p1 sequence as a starting point in the design of modular synthetic biomaterials was the extensive literature on the relationship between sequence and peptide properties. Several studies have looked specifically at how mutations to the *a* and *d* positions in the hydrophobic core affect the preferred oligomerization state and thermal stability of this particular coiled coil.[76, 77, 78, 79, 80] Most combinations of Val, Leu and Ile have been used in the GCN4-p1 core (most of these studies also mutated Asn$_{16}$) with the exception of pVV and pIV. This is presumably because the bulkiness of the Val residues would not be well tolerated at the *d* position and Val insertion at either core position generally causes a loss of specificity in oligomerization state preference.[48] With these peptides, the residue replacement at the *a* and/or *d* position was universal. To our knowledge, no prior report has systematically changed single residues in the hydrophobic core of GCN4-p1 to examine the impact on folded stability. We aimed to synthesize a series of mutants of the GCN4-p1 sequence that systematically varies the core composition by single amino acid substitutions for more precise size control of our supramolecular polymers.

Our goal is to design and synthesize a series of hydrophobic core mutants of GCN4-p1 with systematically altered association affinity. As stated in the previous section, the sequence differences in GCN4-p1 (**1**) and GCN4-pIL (**2**) are four *a* position residues in the hydrophobic core. The removal of the polar contact created by Asn and the increase in hydrophobicity with three Val→Ile mutations accounts for the increase in the folded stability of GCN4-pIL. To destabilize the hydrophobic core, we to replaced Val$_9$ with aminobutyric acid (Abu), which has one fewer methylene groups than Val, to generate peptide

11. To increase hydrophobicity and increase folded stability, Ile was inserted systematically at Val 9, 21 and 30 giving rise to peptides **12**, **13** and **14** with one, two or three Ile in the core (Figure 26).
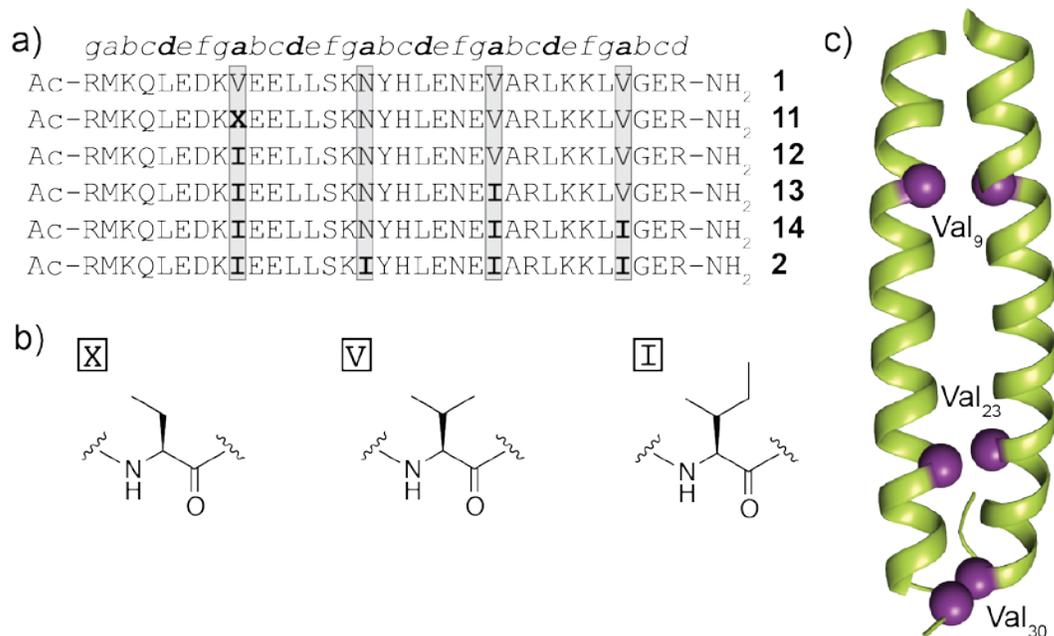
Figure 26. (a) Sequences of GCN4-p1 (**1**) and hydrophobic core mutants **1-2, 11−14** (a); Structures of valine (Val, V), 2-aminobutyric acid (Abu, X), and isoleucine (Ile, I) residues are shown (b). Crystal structure of GCN4-p1 (1, PDB 4DMD) with the positions of the mutated residues in the dimeric coiled coil indicated (c).

We had several choices when we designed the core mutants of GCN4-p1. Both Leu and Ile could increase the hydrophobicity of the core and either the *a* or *d* position could be mutated. Ile is strongly discriminated against at the *d* position due to the preferred perpendicular packing in a dimeric oligomerization state, eliminating the possibility of mutating residues at the *d* positions.[48,]

[79] Ile was chosen to replace Val instead of Leu because of packing preferences at the *a* position. Previous studies that replaced the *a* position of the GCN4-p1 hydrophobic core (including Asn$_{16}$) with Leu reported a change in oligomerization state from a dimer to a tetramer along with a change from parallel to perpendicular packing geometry with these mutations.[48] Ile's influence at the *a* position in coiled-coil dimers has been debated with opinions varying from the *a* position having no preference for Ile to having a slight preference.[76] We also took into consideration the previously reported GCN4-pIL sequence's dimeric oligomerization state and increased thermal stability. In peptides **11-14**, the Asn at position 16 was not mutated to help the sequences specify a dimeric fold.[81] Peptides **11-14** were synthesized by solid phase methods, purified by HPLC and their identity confirmed by MALDI-MS (Table 4 in experimental).
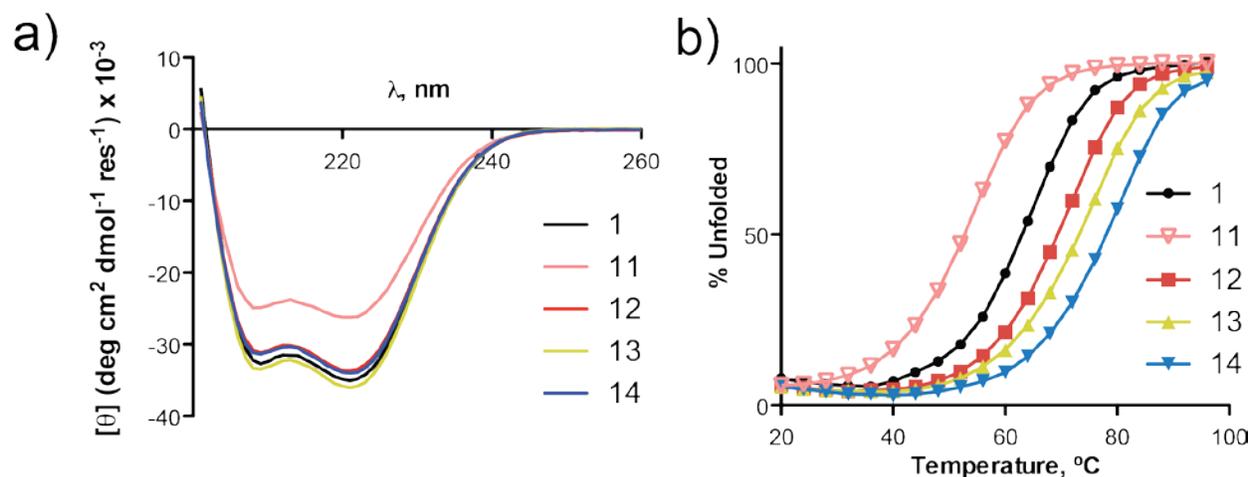
Figure 27. CD scans at 20 °C (a) and thermal melts (b) of peptides **1** and **11-14** at 100 μM concentration in 10 mM pH 7 phosphate buffer.

The mutations to the hydrophobic core are minimally disruptive to overall helicity but systematically alter thermodynamic stability of the quaternary fold.

CD scans of peptides **11**-**14** in pH 7.0 phosphate buffer were used to assess how the introduced mutations affected the α-helical fold of the coiled coil. Peptide **12**-**14** retained similar helicity as the native GCN4-p1 sequence (Figure 27). Peptide **11** showed a slightly decreased helicity compared to GCN4-p1, which is attributed to partial fraying at the N-terminus due to increased disorder around the Abu in the hydrophobic core. CD thermal melts confirm our hypothesis that modulating the hydrophobicity of the coiled-coil core can be used to adjust the thermal stability of coiled coil in the designed sequences (Table 2). Each peptide had a similar two-state cooperative thermal denaturation curve with thermal unfolding midpoints (T_m) ranging from 53˚C-78˚C at 100 μM. The changes in T_m correspond to a 0.4-0.7 kcal mol^{-1} stabilization

Table 2. Circular dichroism data for peptides **1** and **11**-**14**

Peptide	$[\theta]_{222}$ (deg cm^2 dmol^{-1} res^{-1}) x 10^3	T_m (˚C)[a]	$\Delta\Delta G_{fold}$ (kcal mol^{-1})[b]
1	-35.5	62.3 ± 0.3	--
11	-23.7	52.7 ± 0.1	+ 1.1 ± 0.4
12	-33.1	69.4 ± 0.2	- 0.7 ± 0.04
13	-35.9	73.0 ± 0.3	- 1.1 ± 0.05
14	-33.7	77.9 ± 0.2	- 1.7 ± 0.07

[a] Midpoint of the thermal melt of 100 μM peptide in 10 mM phosphate buffer, pH 7. [b] Change in folding free energy with respect to **1**.

per Val→Ile substitution in **12**-**14** and a 1.1 kcal mol^{-1} destabilization in the Val→Abu mutation in **11**. GPC results suggested that the inserted mutations did not change the oligomerization state of the mutants with each peptide adopting a dimeric fold in solution. Peptides **11**-**14** eluted at the same volume when injected at a concentration of 100 μM in pH 7.0 phosphate buffer with 150 mM NaCl (Figure 28). A known trimeric coiled coil, GCN4-pII[51], was also injected and the peak was clearly resolvable.
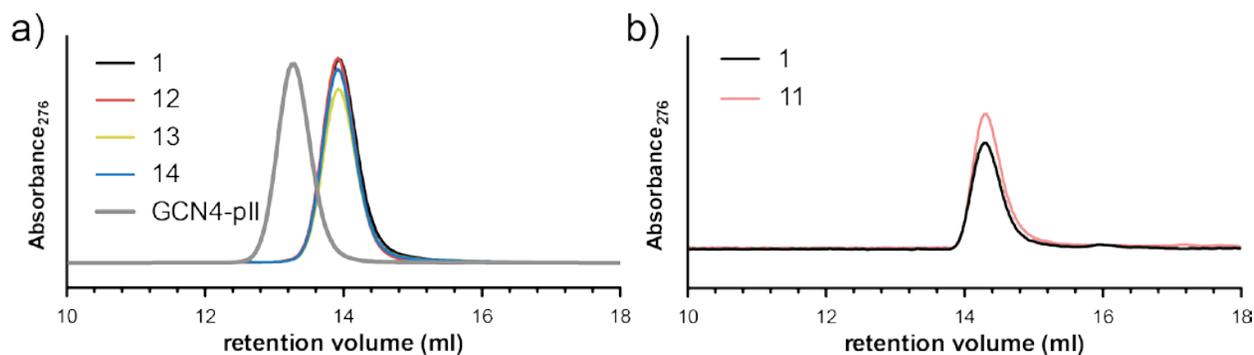
Figure 28. Gel permeation chromatography of peptide 1, GCN4-pII, a trimeric coiled coil, and the mutant peptides confirming the dimeric oligomerization state. Peptide 11 was examined on a different day than peptides 12-14 and therefore plotted independently**.**

2.4   HIGH-RESOLUTION STRUCTURAL ANALYSIS OF CORE MUTANTS

In order to probe the fine structure of the hydrophobic core mutants, we grew diffraction quality crystals of peptides **11**-**14** by hanging drop vapor diffusion and solved their structures by X-ray crystallography to 2.2 Å resolution or better (Table 3). All of the mutants crystallized in the C_2 space group with a single dimeric coiled coil making up the asymmetric unit. Comparison of the overall fold of the four new mutants to the previously published wild-type GCN4-p1 showed the quatenary fold was not altered.

An area of particular interest in these structures was the packing arrangement of the mutated residues in the hydrophobic core. Prior work on coiled coils have assessed how packing geometry affects the residue preference at the *a* and *d* positions. As explained in section 1.3, parallel packing in the hydrophobic core—where the C_α-C_β of the residue is oriented slightly out of the hydrophobic core—accommodates bulky, β-branched residues (such as Val and Ile) better

	11	12	13	14
PDB ID	4NIZ	4NJ0	4NJ1	4NJ2
Space Group	C2	C2	C2	C2
unit cell				
a, b, c (Å)	82.3, 30.1, 27.8	83.2, 30.4, 27.8	83.4, 30.5, 27.9	39.4, 36.7, 47.6
α, β, γ (°)	90, 101.3, 90	90, 102.0, 90	90, 102.1, 90	90, 104.8, 90
resolution (Å)	28.32-2.00 (2.07-2.00)	40.69–1.90 (1.97–1.90)	28.55–2.00 (2.07–2.00)	24.89–2.20 (2.28–2.20)
Total Reflections	15493	17995	11581	10822
Unique	4580	5462	4582	3367
Reflections R_{merge}	0.03 (0.09)	0.05 (0.26)	0.03 (0.09)	0.09 (0.29)
(%)	30.3 (4.7)	20.4 (2.2)	28.3 (4.8)	8.2 (1.8)
I / σ completeness	98.7 (94.8)	99.3 (97.6)	96.2 (86.5)	98.5 (96.5)
(%) redundancy	3.4 (2.1)	3.3 (2.2)	2.5 (1.7)	3.2 (3.3)
Refinement				
resolution (Å)	27.31-2.00	27.18–1.90	27.31–2.00	24.98–2.20
R_{work} / R_{free} (%)	21.2 / 24.3	25.9 / 28.3	24.2 / 25.5	24.8 / 30.0
avg. B factor (Å^2)	20.1	28.8	25.7	28.6
RMSD				
bond lengths (Å)	0.005	0.003	0.002	0.01
bond angles (°)	0.62	0.65	0.33	1.3

than perpendicular packing sites where residue C_α-C_β's are pointed directly into the core. Coiled coils based on the GCN4-p1 sequence tend to favor parallel packing at *a* position residues (primarily Val) and perpendicular packing at *d* position residues (primarily Leu).[48] Design of peptide **11**-**14** took these packing preferences into consideration when Abu and Ile were chosen to replace Val. Analysis of the crystal structures of peptides **11**-**14** shows the mutated residues effectively pack into the hydrophobic core while maintaining the same side chain orientation relative to the wild-type structure (Figure 29). The additional CH_2 groups of Ile in peptides **12**-**14** are sufficiently buried in the hydrophobic core to have minimal effects on the surface physical properties of the folded protein. Analysis of Asn_{16} residue packing showed the polar contact this residue makes is also maintained in all of the mutated sequences.
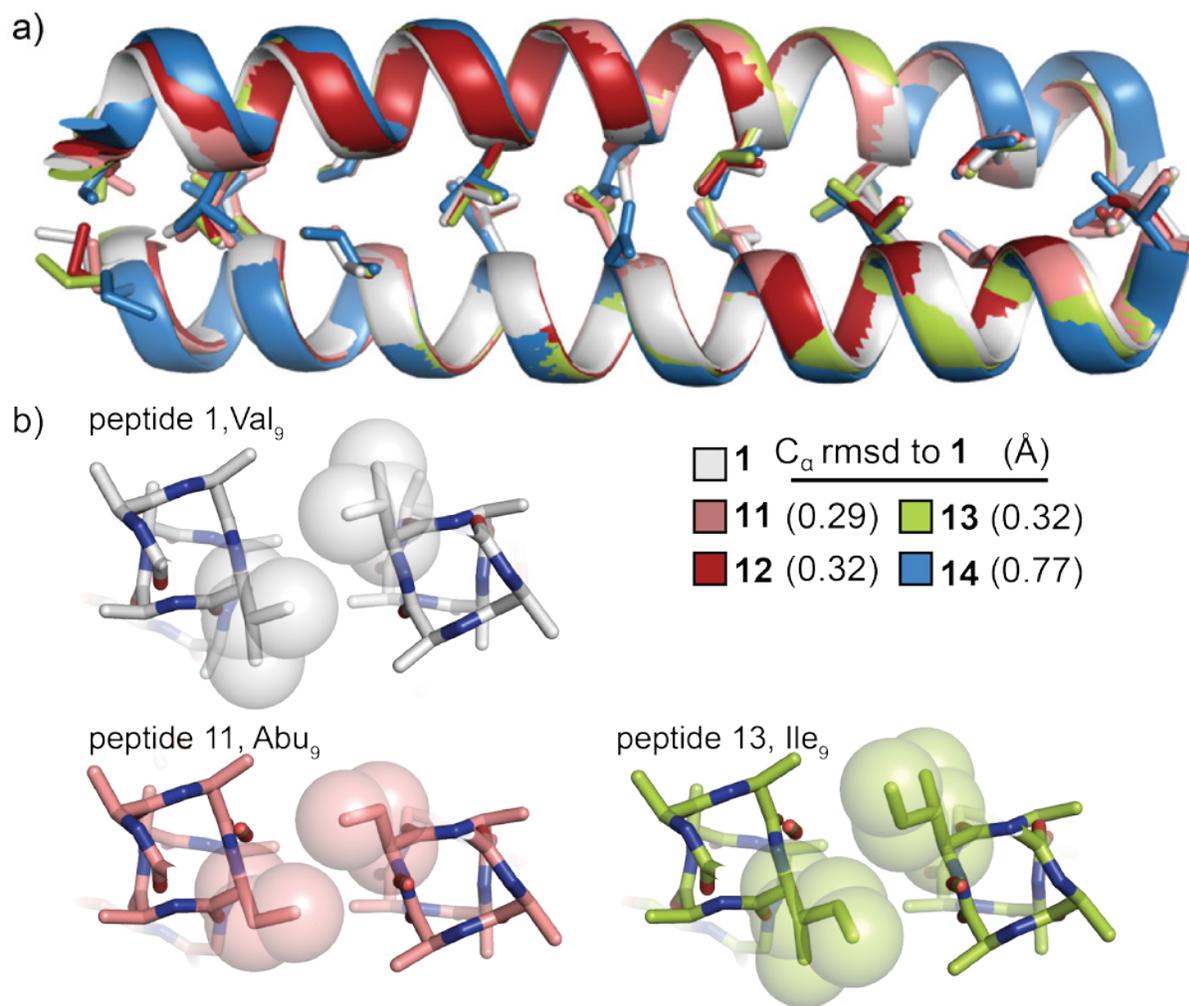
Figure 29. Comparison of X-ray crystal structures of peptides 1 and 11-14. (a) Overlay of residues 2−30 with calculated backbone RMSDs to peptide **1**. (b) Comparison of parallel packing at a heptad position 9 in the hydrophobic core when occupied by a Val, Abu, or Ile residue.

2.5   SOLUTION PHASE DETERMINATION OF SUBUNIT SIZE

With a set of well-characterized GCN4-p1 mutants, we set out to examine the effect of modified coiled-coil stability on supramolecular polymer size. We incorporated these peptides in to the branched subunit to test our hypothesis that increasing or decreasing the association affinity of the coiled coil in the subunit will tune the size of the assemblies. Sequences **15-18**, Ser→Cys point mutations of sequences **11-14**, were synthesized and subjected to the previously described ligation reaction using the bromoacetamide piperazine-based linker (Figure 30). The subunits (**19-22**) from these reactions were purified using semi-preparative HPLC. Products were identified by MALDI-MS, and the purity confirmed by analytical HPLC (Table 4 in experimental).

a)
```
Ac-RMKQLEDKXEELLCKNYHLENEVARLKKLVGER-NH2  15 (19)
Ac-RMKQLEDKIEELLCKNYHLENEVARLKKLVGER-NH2  16 (20)
Ac-RMKQLEDKIEELLCKNYHLENEIARLKKLVGER-NH2  17 (21)
Ac-RMKQLEDKIEELLCKNYHLENEIARLKKLIGER-NH2  18 (22)
```

b)

19-22
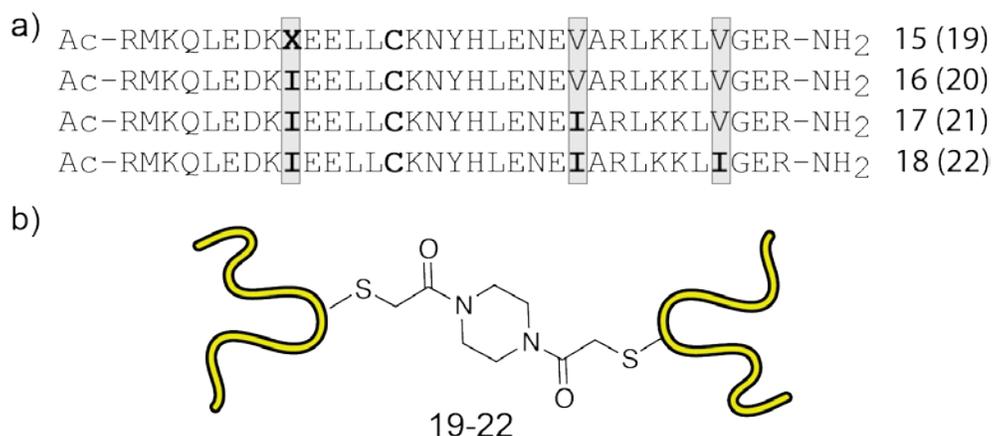
Figure 30. The sequences of the hydrophobic core mutants with a Ser→Cys mutation (a).

The first numbers correspond to the free Cys peptide. The numbers in the parentheses are the piperazine subunits

based on those peptides (b). X = Abu

55

DLS was used to measure the apparent hydrodynamic radius of the supramolecular assemblies formed by cross-linked subunits **19**-**22** in solution. Each was analyzed in 50 mM HEPES buffer pH 7.0 with 150 mM NaCl at a concentration of 200 µM subunit. Each sample was prepared from a stock solution and, after the addition of buffer, filtered to remove any dust or particulate matter. The formation of the assemblies in each sample was monitored over time by DLS. Subunit **9**, based on the GCN4-p1 sequence, and subunit **19**, based on the destabilizing Val→Abu peptide, both appeared to stabilize after 3 hours (Figure 31a). Subunits **20** and **21**, with one or two Val→Ile insertions, reached equilibrium on a longer time scale, but reached a stable assembly size by 8 hours. Subunit **22**, based on the triple Val→Ile mutant, showed signs of turbidity after 6 hours and did not reach a stable assembly size on the time scale of the experiment. After 24 hours at room temperature, significant aggregation was apparent in the sample. It is worth noting that the model used to analyze the raw DLS data (diameter of a hard sphere with equivalent scattering behavior) differs from the actual nature of the assemblies (long, flexible chains); however, it provides qualitative data about assembly size without dilution effect that would be present if GPC were used to measure size.

Comparison of the apparent hydrodynamic radius data from DLS measurements of subunits **19**-**22** confirms our hypothesis that modulation the thermodynamic stability of the coiled coil by rational mutations to the hydrophobic core is an effective way of controlling size in our supramolecular polymers. We compared of the apparent hydrodynamic radius of subunit **8** and subunits **19**-**22** with relative association free energy change in the isolated coiled-coil peptide (**1** and **11**-**14**) in each subunit (Figure 31b). For subunits **8**, **19** and **20**, each stabilizing core mutation increased the apparent hydrodynamic radius by ~10 nm per substitution. Subunits
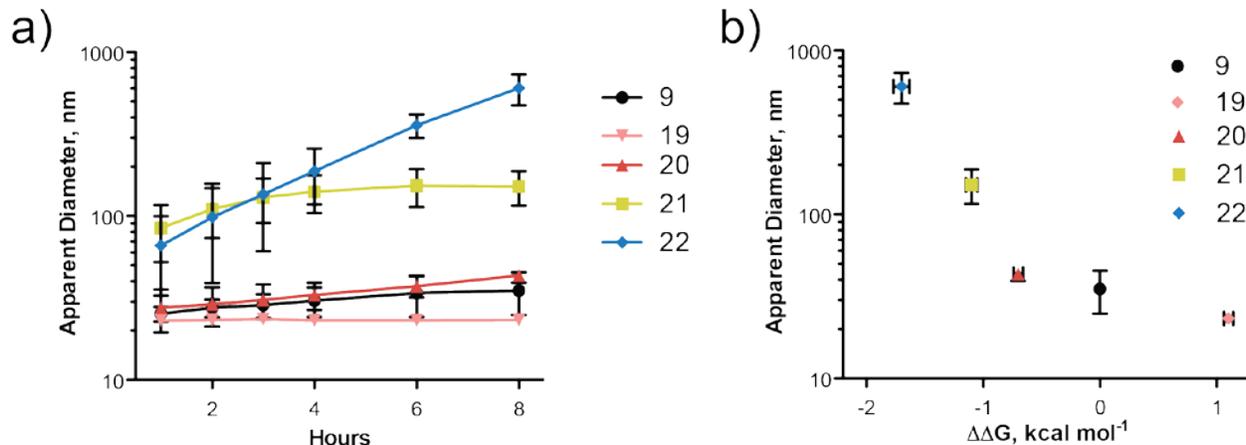
Figure 31. Time-dependent DLS data monitoring the self-assembly of subunits **9** and **19**-**22** (a). Plot of apparent assembly size by DLS at 8 h against free energy of association of the coiled coil domain ($\Delta\Delta G_{assoc}$ vs 1) estimated by biophysical analysis of **11**-**14** (b).

20-**22** showed change in assembly size of >100 nm per stabilizing core substitution. Initial interpretation of this data would suggest a cooperative assembly mechanism may more accurately describes this set of subunits; however, we believe the nonlinear relationship between assembly size in subunits **20**-**22** and the degree of stabilization most likely arises due to the larger supramolecular polymers having a higher tendency to associate laterally to form larger aggregates.

2.6 ASSEMBLY MORPHOLOGY

In order to directly visualize the assemblies formed with these subunits and to test for the possibility of inter-chain association, we employed negative-stain transmission electron microscopy (TEM) of subunit **8** (Figure 32). The TEM results showed high aspect-ratio fibers,

consistent with the expected mechanism of assembly. The diameter of the fibers (6 nm) is in good agreement with the predicted diameter based on measurements taken from the native GCN4-p1 crystal structure. Examples of inter-chain association can clearly be seen with two or more chains bundling together. The lateral interaction of multiple chains, a form of nonideal behavior in supramolecular polymers, would become more apparent when the more stable subunits form larger chains. The additional lateral association in the most stable subunit **22** is likely a cause of the observed uncontrolled aggregation
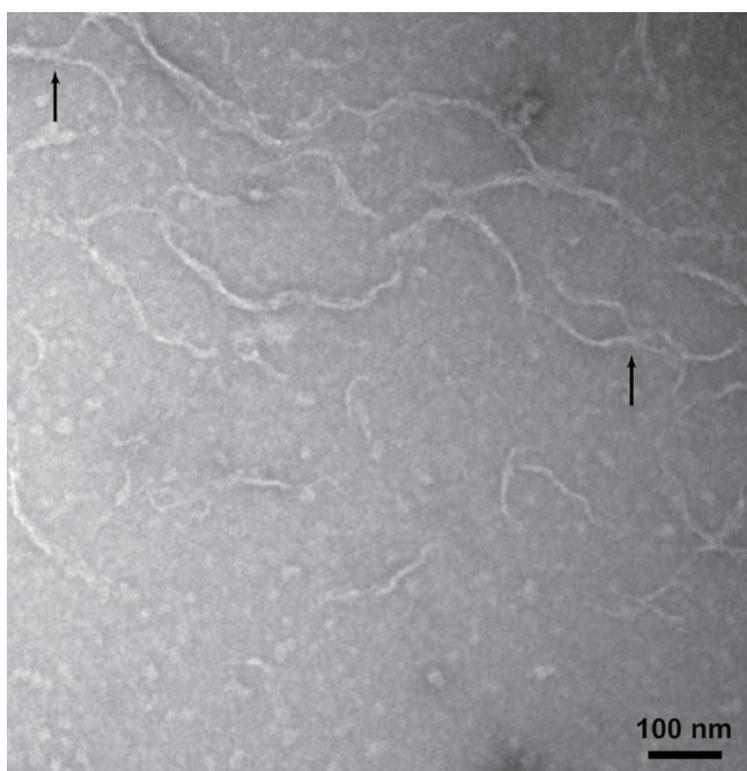
Figure 32. Negative-stain transmission electron microscopy image of the fibrous assemblies formed by subunit 9.

Arrows indicate examples of apparent interchain association.

2.7 SUMMARY

In summary, we have demonstrated a modular, convergent method for the synthesis of branched subunits consisting of an α-helical coiled-coil peptide linked at its midpoint by an organic bridge. These subunits are capable of forming supramolecular assemblies in dilute aqueous solutions. The method allows for easy substitution of either linker or peptide, and the chemoselective thioether formation used as the ligation reaction permits the synthesized peptides to be replaced with material produced by heterogeneous expression. By systematically varying the components of the subunit, we were able to determine the dependence of assembly properties on linker structure and coiled-coil association affinity.

Two characteristics of the linking group—length and rigidity—affect assembly size. Too short of a linker causes the solvent exposed side chains of adjacent coils to come in close proximity to one another, disrupting the tertiary structure as seen in subunit **7**. The rigidity of the linker also plays an important role in assembly size. The ethylenediamine and piperazine based linkers differ only in rigidity. The difference in rigidity of the linker is extrapolated to the overall assembly properties indicated by a larger apparent hydrodynamic radius. Because the self-association process is dynamic, combinational control over the process can be exerted. The addition of monomeric peptide to cross-linked subunit can be used as a tool to modify the distribution of sizes depending on the ratio of subunit to free coiled-coil peptide.

We also studied the effect of increased folded stability on the assembly size using a known variant of GCN4-p1. While the more stable GCN4-pIL peptide gave rise to larger surpamolecular chains, problems with assembly solubility and aggregation prevented further biophysical characterization. This supports our hypothesis that increasing the association affinity of the coiled coil will have an effect on the length of the chains; however, it also indicated the need for more

moderate increases in the folded stability to balance it with finely controlling the size of assemblies while remaining soluble.

To achieve greater control over the size of the supramolecular assemblies through association affinity, we created a series of hydrophobic core mutants that systematically altered the stability of the coiled coil. The mutation of Val at an *a* position to either Ile or Abu translates to a change in association free energies ranging from 0.4 kcal mol^{-1} to ~ 1.7 kcal mol^{-1} while still retaining the native folded state. Solution-phase biophysical measurements confirm the equilibrium is not shifted away from the dimeric oligomerization state and high-resolution crystal structures show parallel dimeric coiled coils for all of the mutants. The close packing of the hydrophobic core remains intact when a methyl group is either removed, in the case of Abu, or added with of the bulkier Ile residue. Importantly, the polar contact between Asn$_{16}$ is not disrupted.

The above set of coiled-coil forming peptides was used to prepare a family of self-assembling subunits through cross-linking at a solvent-exposed Cys introduced into each sequence. Solution analysis of the resulting subunits by DLS indicated a direct correlation between changes in the peptide association free energy and supramolecular polymer size. Transmission electron microscopy confirms the fibrous morphology of the materials formed. This family of soluble supramolecular polymers of tunable size will find use in our ongoing work to prepare functional assemblies based on this system. Moreover, our results suggest that maintaining ideal behavior in larger linear assemblies may be possible through introduction of capping groups in the coiled-coil domains that disfavor interchain association in the supramolecular polymers through electrostatic or steric effects. The series of well-characterized peptides with tunable association affinity adds to the library of building blocks that can be used in the rational design of peptide-based supramolecular biomaterials.

2.8   EXPERIMENTAL

2.8.1   Synthesis of Peptides

Peptides were prepared by Fmoc solid-phase synthesis using manual microwave-assisted protocols[82] or in automated fashion on a Protein Technologies Tribute Automated Synthesizer. NovaPEG Rink Amide resin was used to prepare the C-terminal amide. Peptides were cleaved from resin by treatment with 94% trifluoroacetic acid, 2.5% water, 2.5% ethanedithiol and 1% triisopropylsilane solution for 2 to 6 hours. After the peptide was cleaved from resin, it was precipitated from the filtered cleavage solution by addition of ~40 mL cold diethyl ether. The precipitate was pelleted by centrifugation and the ether decanted. The peptide pellet was suspended in a mixture of 0.1% TFA in water and 0.1% TFA in acetonitrile for purification. Peptides were purified by HPLC on a C_{18} preparative column using 0.1% TFA in water and 0.1% TFA in acetonitrile gradients. HPLC fractions containing the product were combined, frozen, and lyophilized. Peptide identity was confirmed by mass spectrometry using a Voyager DE Pro MALDI-TOF instrument. All peptides were >95% pure by analytical HPLC on a C_{18} column.

Table 4. MALDI masses of peptides and subunits synthesized in Chapter 2

Peptide	[M + H]+ m/z (avg.)	
	Calculated	Observed
1	4038.7	4036.9
3	4054.7	4054.7
4	4095.9	4096.7
7	8106.4	8105.4
8	8248.6	8248.3
9	8274.6	8274.7
10	8330.8	8333.0
11	4024.6	4022.4
12	4052.7	4051.2
13	4066.7	4065.8
14	4080.7	4079.7
15	4040.7	4037.2
16	4068.7	4067.1
17	4082.7	4081.1
18	4096.8	4095.4
19	8246.6	8243.7
20	8302.7	8297.6
21	8330.7	8325.2
22	8358.7	8355.8

2.8.2   Synthesis of Bromoacetamide Ethylendiamine Linker (5a)

Protocol adapted from the literature[83] and compound prepared by J.K. Staples. Ethylene diamine (0.668 mL, 10 mmol) was dissolved in 10 mL chloroform, and the solution cooled on an ice bath. The flask was equipped with two addition funnels, one containing bromoacetyl bromide (2.613 mL, 30 mmol) in 12 mL chloroform and the other containing potassium carbonate (3.3 g, 23.9 mmol) in 12 mL water. The contents of the addition funnels were added simultaneously to the stirred reaction over 15 minutes. After addition was complete, the ice bath was removed, and the reaction allowed to continue at room temperature for 2 h. The aqueous and organic layers were separated and the aqueous phase extracted with ethyl acetate (4 x 100 mL). The combined organic layers were then dried over magnesium sulfate and the solvent removed under vacuum to give 1.66

g (55% yield) of the product as a white solid. ^{1}HNMR (400 MHz, D$_2$O) δ = 3.84 (4H, s), 3.34 (4H, s). ^{13}CNMR (100 MHz, DMSO-d$_6$) δ = 166.2, 38.5, 29.5. HRMS *m/z* calculated for C$_6$H$_{11}$Br$_2$N$_2$O$_2$ [M+H]$^+$: 300.9178; found: 300.9182.

2.8.3 Synthesis of Iodoacetamide Ethylenediamine Linker (5b)

Protocol adapted from literature.[84] Ethylene diamine (0.67 mL, 10 mmol) was added to 200 mL of 0.05 M NaOH. Iodoacetyl chloride (1.79 mL, 25 mmol) in 50 mL 1,2-dichloroethane was then added to the mixture and allowed to vigorously stir at room temperature for 10 minutes. A pale yellow precipitate was recovered after vacuum filtration. The precipitate was then washed with water to remove impurities, vacuum filtered and dried under vacuum to give 1.30 g (33.1% yield) of white powder. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ = 8.28 (2H, s), 3.61 (4H, s), 3.09 (4H, s); ^{13}C NMR (125 MHz, DMSO-d$_6$) δ = 168.3, 39.0, 1.2; HRMS m/z calculated for C$_6$H$_9$I$_2$N$_2$O$_2$ [M−H]−: 394.8753; found: 394.8748.

2.8.4 Synthesis of Iodoacetamide Piperazine Linker (6)

Protocol adapted from the literature[84], [85] and compound prepared by J.K. Staples. Piperazine (0.8614 g, 10 mmol) was dissolved in 10 mL chloroform and the solution cooled on an ice bath. The flask was equipped with two addition funnels, one containing bromoacetyl bromide (2.613 mL, 30 mmol) in 12 mL chloroform and the other containing potassium carbonate (3.3 g, 23.9 mmol) in 12 mL water. The contents of the addition funnels were added simultaneously to the stirred reaction over 15 minutes. The ice bath was removed, and the reaction allowed to continue at room temperature for an additional 2 h. The aqueous and organic layers were separated

and the aqueous phase extracted with ethyl acetate (4 x 100 mL). The combined organic layers were then dried over magnesium sulfate and the solvent removed under vacuum to give 2.54 g (77% yield) of the product as a white solid. HNMR (400 MHz, D_2O) δ = 4.05 (4H, s), 3.65 (8H, m); ^{13}CNMR (100 MHz, DMSO-d_6) δ = 165.0, 164.9, 45.7, 45.3, 41.5, 41.2, 28.0, 27.9; HRMS m/z calculated for $C_8H_{13}Br_2N_2O_2$ [M+H]$^+$ 326.9344; found: 326.9378.

2.8.5 Synthesis of Disulfide Subunit (7)

A 3 mL solution (1.8 mg/mL, ~222 µM) of peptide **3** was prepared in 0.1 M phosphate buffer, pH 8 with 15% v/v DMSO. The reaction was allowed to proceed with stirring for 24 h, after which the formation of a precipitate was observed and analysis by HPLC showed significant depletion of the starting material. The solution was diluted with an additional 3 mL of 15% v/v aqueous DMSO to dissolve the precipitate, lyophilized, and redissolved in a 20% aqueous acetonitrile with 0.1% TFA. The crude material was purified by preparative HPLC; the identity and purity of the final product were verified by analytical HPLC and MALDI-MS, respectively.

2.8.6 Synthesis of Ethylenediamine Subunits (8 and 10)

A solution of **3** or **4** was prepared in deionized water and the concentration determined by UV absorption at 276 nm in 6.0 M guanadinum chloride, pH 7.0. A fresh solution of linker was prepared in approximately 5 mL of DMF (1.5 mM final concentration of linker). A 4 mL solution was prepared consisting of 150 µM peptide in 25 mM phosphate buffer, pH 7.0. One aliquot of the linker solution (40 µL) was added immediately. The reaction was placed in water bath at approximately 65°C and allowed to stir for 2 hours with four additional aliquots of linker being

added in 15 minute intervals for the first hour (200 µL linker total). The reaction was quenched

after 2 hours with a 25% acetonitrile/0.01% TFA solution and then purified using molecular weight

centricon spin filters (3000 MWCO) and semi-preparative HPLC. Centricon filters were used to

remove the DMF linker co-solvent and concentrate the subunit. Equal amounts of reaction solution

and water were added to the spin filter, and the mixture was spun at 6000 RPM for 30 minutes.

Additional filtration spins were performed until the reaction mixture had been washed with ~3X

volumes of water. Identity and purity of the final product was confirmed by analytical HPLC and

MALDI-MS.

2.8.7 Synthesis of Piperazine Subunits (9 and 19-22)

A solution of **3** was prepared in deionized water and the concentration determined by UV

absorption at 276 nm in 6.0 M guanadinum chloride, pH 7.0. A fresh solution of linker was

prepared in approximately 5 mL of DMF (1.5 mM final concentration of linker). A 2.75 mL

solution was prepared consisting of 150 µM peptide in 25 mM phosphate buffer, pH 7.0 with 1%

acetonitrile. One aliquot of the linker solution was added immediately. The reaction was placed in

water bath at approximately 65°C and allowed to stir for 2 hours with four additional aliquots of

linker being added in 15 minute intervals for the first hour (137.5 µL linker total). To aid with

solubility, two aliquots of acetonitrile were added in 30 minute intervals (total of 3% acetonitrile).

The reaction was quenched after 2 hours with a 25% acetonitrile/0.01% TFA solution and then

purified using molecular weight centricon spin filters (3000 MWCO) and semi-preparative HPLC.

Centricon filters were used to remove the DMF linker co-solvent and concentrate the subunit.

Equal amounts of reaction solution and water were added to the spin filter, and the mixture was

spun at 6000 RPM for 30 minutes. Additional spin filtering was performed until the reaction

mixture had been washed with ~3X volumes of water. Identity and purity of the final product was confirmed by analytical HPLC and MALDI-MS.

2.8.8 Circular Dichroism Scans and Melts

Measurements were taken on an Olis DSM17 Circular Dichroism Spectrometer using 0.1 cm quartz cuvettes. Peptide concentration was determined by UV absorbance at 276 nm (ε = 1450 M^{-1} cm^{-1})[86] from the single Tyr residue in the GCN4-p1 sequence. Samples of 100 µM peptide in buffer were prepared and scanned from 200 nm to 260 nm in 1 nm increments, an integration time of 5 seconds and a bandwidth of 2 nm at 20°C. A buffer blank was used to correct each spectrum and baseline molar ellipticity at 260 nm. Variable temperature CD was taken of the subunits by monitoring molar ellipticity at 222 nm from 0-96°C in 5°C increments with a 5 minute equilibration time between data points and an integration time of 5 seconds. Variable temperature CD of the mutants was taken by monitoring molar ellipticity at 222 nm from 20-96°C in 4°C increments with a 2 minute equilibration time between data points and an integration time of 5 seconds. Thermal melt data was fit to a two- state unfolding model to obtain the melting temperature (T_m).[87] From the melting temperatures, percent unfolded was calculated:

Equation 1

$$\%unfolded = \left[1 - \frac{[\theta]_T - [\theta]_{unfolded}}{([\theta]_{fold} + T \times M_{fold}) - [\theta]_{unfold}} \right] \times 100$$

where $[\theta]_T$ is the molar elipticity at 222 nm at the temperature of the measurement, $[\theta]_{unfold}$ is the molar elipticity at 222 nm of the fully unfolded peptide obtained from the two state denaturation fit, $[\theta]_{fold}$ is the molar elipticity at 222 of the fully folded peptide, T is the temperature in degrees Celsius, and M_{fold} is the slope of the fully folded measurements at low temperatures

from the two-state denaturation fit. The use of % unfolded versus molar ellipticity to analyze the thermal denaturation makes the change in signmodial shape easier to visualize.

2.8.9   Calculation of ΔΔG

A value for ΔΔG for the mutant peptides was calculated using[88]:

Equation 2

$$\Delta\Delta G = \Delta H_{T_m} \frac{\Delta T_m}{T_{m,w.t.}}$$

In the above equation, ΔH_{Tm} is the enthalpy obtained for the wild type peptide from the thermal melt analysis, $T_{M,\,w.t.}$ is the melting temperature of the wild type obtained from the thermal melt, ΔT_m is the difference in melting temperatures of the mutant and wild type calculated by:

Equation 3

$$\Delta T_m = T_{m,mut.} - T_{m,w.t.}$$

Error propagation in the calculation of ΔΔG is obtained from equation below where the $\delta\{T_{m,w.t.}\}$ value is obtained as standard deviations of the two-state thermal denaturation curve fit.

Equation 4

$$\delta\{\Delta\Delta G\} = \Delta\Delta G \sqrt{\left(\frac{\delta\{\Delta H_{T_m}\}^2}{\Delta H_{T_m}}\right) + \left(\frac{\delta\{T_{m,w.t.}\}^2}{T_{m,w.t.}}\right) + \left(\frac{\delta\{\Delta T_m\}^2}{\Delta T_m}\right)}$$

The value for $\delta\{\Delta T_m.\}$ was obtained from equation, where error values are obtained from the two-state thermal denaturation fit. The value for $\delta\{\Delta H_{Tm}.\}$ was calculated from the same equation by using the correlating $\delta\{H\}$ values.

Equation 5

$$\delta\{\Delta T_m\} = \sqrt{(\delta\{\Delta T_{m,w.t.}\})^2 + \delta(\{\Delta T_{m,mut}\})^2}$$

2.8.10 Gel Permeation Chromatography

GPC of the subunits was preformed on a Superdex 200 10/300 column (10 x 300 mm, 24 mL bed volume, 13 μm average particle size, GE Healthcare). The column was equilibrated with 0.15 M NaCl in 0.05 M HEPES, pH 7.0. Subunits were loaded onto the column (100 μL of sample with 100 μM subunit) and eluted at a flow rate of 0.65 mL/min. A molecular weight calibration curve was obtained by fitting elution volumes of 1 mg/mL solution of ferritin, BSA, ovalbumin and aprotinin.

GPC of the GCN4-p1 mutants was carried out on a Superdex 75 10/300 column (10 x 300 mm, 24 mL bed volume, 13 mm average particle size, GE Healthcare). The column was equilibrated with 0.15 M NaCl in 0.05 M sodium phosphate, pH 7.0. Peptides were loaded onto the column (100 mL sample at 100 μM concentration in equilibration buffer) and eluted at a flow rate of 0.8 mL/min. A molecular weight calibration curve was obtained by fitting the elution volumes of 1 mg/mL solutions of BSA, ovalbumin, aprotinin, a 17-residue synthetic peptide (Ac-YEAAAKEAAAKEAAAKA-NH$_2$), and vitamin B12.

2.8.11 Dynamic Light Scattering

Data were collected on a Malvern Zetasizer Nano ZS90 (Malvern Instruments, Westborough, MA) dynamic light scattering instrument with a 632.8 nm laser at a fixed angle of 90° and a constant temperature of 25°C. Subunit samples (200 μM peptide in 150 mM NaCl, 50 mM HEPES pH 7) were prepared by adding 10X buffer to concentrated stock solutions of subunit in water. The samples were filtered using 0.22 μm filters and measurements were taken every hour for eight

hours. Measurements were taken in a low volume quartz cuvette with a path length of 1 cm. Three measurements of at least 25 runs were taken for each sample.

2.8.12 Crystallization, Diffraction Data Collection and Structure Determination of Core Mutants

Crystallization was carried out using the hanging drop vapor diffusion method. Drops were prepared by mixing 0.7 or 0.5 µL of peptide stock (10 mg/mL in water) with 0.7 µL of buffer and allowed to equilibrate over a well containing 700 µL of buffer solution. Crystals of peptide **11** were obtained from a well buffer containing 0.05 M sodium acetate, 0.1 M sodium citrate tribasic pH 5.6, 20% w/v PEG 4000. A single crystal was flash frozen in liquid N_2 after being soaked in the mother liquor supplemented with 15% glycerol. Crystals of peptide **12** were obtained from a well buffer composed of 0.3 M sodium acetate pH 4.6, 0.1 M sodium citrate tribasic pH 5.6, and 25% w/v PEG 4000. A single crystal was flash frozen in liquid N_2 after being soaked in the above buffer supplemented with 25% v/v glycerol. Crystals of peptide **13** were obtained from a well buffer composed of 0.1 M sodium acetate pH 4.6, 0.1 M sodium citrate tribasic pH 5.6, and 25% w/v PEG 4000. A single crystal was flash frozen in liquid N_2 after being soaked in the above buffer supplemented with 10% v/v glycerol. Crystals of peptide **14** were obtained from a well buffer composed of 0.15 M sodium citrate tribasic pH 5.6, 20% v/v 2-propanol, and 15% PEG w/v 4000. A single crystal was flash frozen in N_2 after being soaked in the above buffer supplemented with 10% v/v glycerol. Diffraction data were collected on Rigaku Saturn 944 CCD or Rigaku Raxis HTC detector using CuKα radiation. d*TREK was utilized to index, integrate, and scale the collected data.

Structure refinement was carried out using CCP4[89] and Phenix[90]. The structures were solved by molecular replacement using previously published structures of GCN4-p1 (PDB 2ZTA[91], 4DMD[81]) as search models. A combination of refinement programs were used to complete the structure: Refmac[92] and Phenix for automated refinement, Coot[93] for manual model building, ARP/wARP[94] for solvent building, and Phenix for construction of composite omit maps.

2.8.13 Transmission Electron Microscopy

TEM was carried out on an FEI Morgagni 268 electron microscope at an accelerating voltage of 80 kV. Samples of subunit **9** (200 μM) were prepared in 50 mM HEPES buffer pH 7.2. The solution was syringe filtered through a 0.22 μm pore size filter and allowed to equilibrate for 8 h. The equilibrated sample was diluted to 50 μM, dropped onto a carbon Formvar coated 300-mesh grid (Electron Microscopy Science), and allowed to stand. After 10 min, the drop on the grid was diluted with 6 μL of water, excess liquid was wicked off, and the remaining material stained with 2% uranyl acetate for 1 min. The grids were allowed to dry overnight in open air and then stored in a desiccator prior to imaging.

3.0   OLIGOMARIZATION STATE SWITCHING IN GCN4-P1 SEQUENCES

Work detailed in this chapter has been published as:

Oshaben, K. M.; Salari, R. M.; McCaslin, D. R.; Chong, L. T.; Horne, W. S., "The Native
GCN4-p1 Sequence Does Not Uniquely Specify a Dimeric Oligomerization State."
Biochemistry, **2012**, 51, 9581-9591

GCN4-p1 is a well-studied dimeric coiled coil found at the C-terminus of the yeast transcription factor GCN4.[48, 61, 91] In literature studying GCN4-p1, no evidence has indicated an oligomerization state other than dimer for the native sequence. The strong dimeric preference and ability to change coiled-coil stability by mutations to the hydrophobic core—without changing oligomerization state—was a key component in choosing it as a base sequence for our materials work (Chapter 2). The mutants gave similar solution biophysical data as the native GCN4-p1 sequence, but in the process of obtaining high-resolution crystal structures of the core mutants, we observed the oligomerization state of the mutant peptides in the crystal lattice switched between dimeric and trimeric forms depending on the crystallization conditions. The oligomerization state switch observed in the mutants prompted us to re-evaluate the native sequence's ability to adopt different folds.

The hydrophobic core composition of GCN4-p1 based coiled coils have been previously analyzed to assess which *a* and *d* position residues favor dimeric, trimeric and tetrameric oligomerization states.[48, 49, 95] These analyzes suggest that having predominately Val at the *a* positions and Leu at the *d* positions, GCN4-p1 should have poor discrimination between oligomerization states. The addition of Asn at position 16 in the hydrophobic core pushes the

equilibrium to favor the dimer.[96, 97, 98] The results discussed in this section show that Asn is not sufficient to prevent GCN4-p1 from adopting a trimeric fold. These findings were quite surprising as the GCN4-p1 sequence is thought to code exclusively for the dimeric oligomerization state and no other variation had been reported in the literature.

3.1   STRUCTURE DETERMINATION OF NATIVE GCN4-P1 SEQUENCES

Two high-resolution crystal structures were obtained from the crystallization of GCN4-p1 in different buffer conditions. Based on indexing of the diffraction pattern, neither crystal form corresponded to the same lattice as the previously published structure (PDB ID: 2ZTA).[91] We obtained X-ray diffraction data for each crystal and solved the structures to 2.0 Å and 2.2 Å resolution (Table 5). The higher resolution structure was solved using molecular replacement with the native dimer GCN4-p1 structure as the search model. The second crystal form failed to give reasonable results when using the dimeric GCN4-p1 structure as the search model. Molecular replacement was repeated with a single α-helix and a parallel trimeric coiled coil was found in the asymmetric unit. The molecular replacement was repeated with GCN4-pII, a trimeric mutant of GCN4-p1[97], confirming the trimeric oligomerization state.

Although the crystal lattice of the dimer we obtained is different from the published GCN4-p1 structure, the overall coiled-coil quaternary fold remained largely the same (C_α RMSD = 0.53 Å). Differences in the two structures are found primarily in the conformations of some solvent exposed side chains and the C-terminal carboxamide. The C-terminal tail tends to be disordered past Gly_{31} and this region is not consistently resolved in the electron density

Table 5. Crystallographic data collection and refinement data for the GCN4-p1 dimer and trimer.

	GCN4-p1 dimer	GCN4-p1 trimer
Data Collection		
unit cell		
a, b, c (Å)	83.4, 30.5, 27.8	61.2, 34.4, 78.1
α, β, γ (°)	90, 104.5, 90	90, 139.7, 90
resolution (Å)	40.4–2.0 (2.07–2.00)	50.5–2.2 (2.28–2.20)
R_{merge} (%)	5.8 (25.3)	9.4 (22.1)
$I / \sigma(I)$	15.5 (2.2)	9.7 (3.2)
completeness (%)	99.8 (100.0)	99.5 (100.0)
redundancy	3.8 (2.4)	2.7 (2.8)
Refinement		
resolution (Å)	40.4–2.00	50.54–2.20
no. reflections	4432	5198
R_{work} / R_{free} (%)	23.5 / 26.5	22.6 / 28.7
no. atoms	567	903
avg. B factor (Å^2)	23.9	24.7
RMSD		
bond lengths (Å)	0.016	0.010
bond angles (°)	1.8	1.4

A residue with significant impact on the oligomerization state in the GCN4 series of coiled coils is Asn_{16}.[91, 96, 98, 99] This statement is supported in the literature by (1) the formation of a polar contact in an otherwise hydrophobic core of the dimer[99] and (2) the substitution of only Asn_{16} can lead to changes in the favored oligomerization state. [96, 97, 98, 99] To gain insight into the role of Asn_{16} in oligomerization state preference, we examined the hydrophobic core surrounding this residue in both the dimeric and trimeric structures of GCN4-p1. The dimeric coiled-coil core is tightly packed, burying a surface area of 1810 Å^2, and has a single polar contact between the Asn_{16} residues (Figure 33a). The trimeric form of GCN4-p1 is also tightly packed, burying a surface area of 4000 Å^2, but a cavity of 112 Å^3 is observed in the hydrophobic core (Figure 33b). Although excluded from the solvent, two waters are contained within this cavity and are
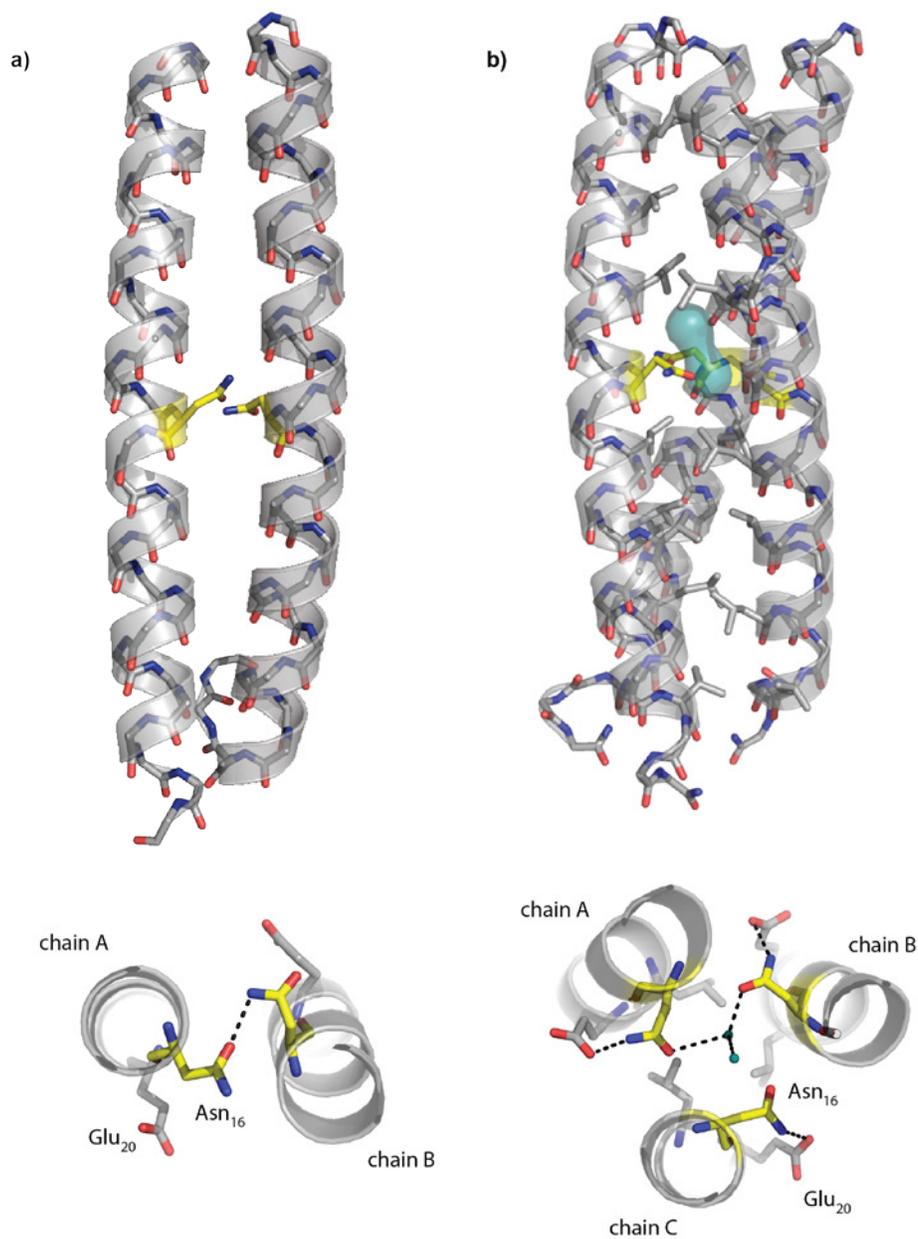
Figure 33. The hydrophobic core of native GCN4-p1 in a different crystal lattice retains the tight packing and polar contact found in the 2ZTA structure (a).

The trimeric GCN4-p1 hydrophobic core sequesters two water molecules away from the solvent at Asn_{16}. The Asn residues form a polar contact with the water and also form an addition intra-chain contact with Glu_{20} that may help to stabilize the trimeric oligomerization state (b). A side view of the trimer structure with the sequestered water molecules highlighted in blue.

stabilized in the pocket by a series of hydrogen bonds. One water molecule forms hydrogen bonds with the carboxamines of two of the Asn_{16} residues, while the second forms a hydrogen bond above the plane of the Asn_{16} residues with the first water (Figure 33a). All of the Asn_{16} residues are involved in intrachain hydrogen bonds with Glu_{20}, providing additional stabilization of the trimeric fold. This series of hydrogen bonds is thought to help stabilize the polar water in an otherwise nonpolar core.

While the trimeric structure of the native GCN4-p1 sequence is the first reported in the literature, five other point mutations to this sequence at Asn_{16} have caused dimer to trimer oligomerization state switches. Four of these mutations—Ser, Thr, Gln and (*S*)-aminoburtyric acid (Abu)—have high-resolution crystal structures. [96, 97, 98, 99] The fifth mutation, Asn→Val, poorly discriminated between dimeric and trimeric species in solution and did not crystalize.[56] We performed a backbone overlay of our trimeric GCN4-p1 structure and those of the four published crystal structures of the trimeric point mutations to assess structural similarity (Figure 34). The backbone structure homology was high across the series (C_α RMSD = 0.55 Å-0.85 Å) with the most similar structure being the Asn→Thr mutation. Idealized superhelix parameters, calculated from the C_α positions, indicate that the GCN4-p1 trimer is more tightly wound than the other coiled coils due to a small pitch but is similar otherwise.

The hydrophobic core composition of GCN4-p1, predominately Val at *a* and Leu at *d*, is known to be poorly discriminating between oligomerization states.[49, 51, 98] The core variant GCN4-pVL, where Asn is replaced by Val, has been shown to populate a mixture of both dimeric and trimeric assembly states in solution.[48] Placement of Asn in the hydrophobic core has been proposed to shift the equilibrium and help to specify the dimeric state in solution. The data presented here challenge that point, showing that the trimeric state is still accessible even with a
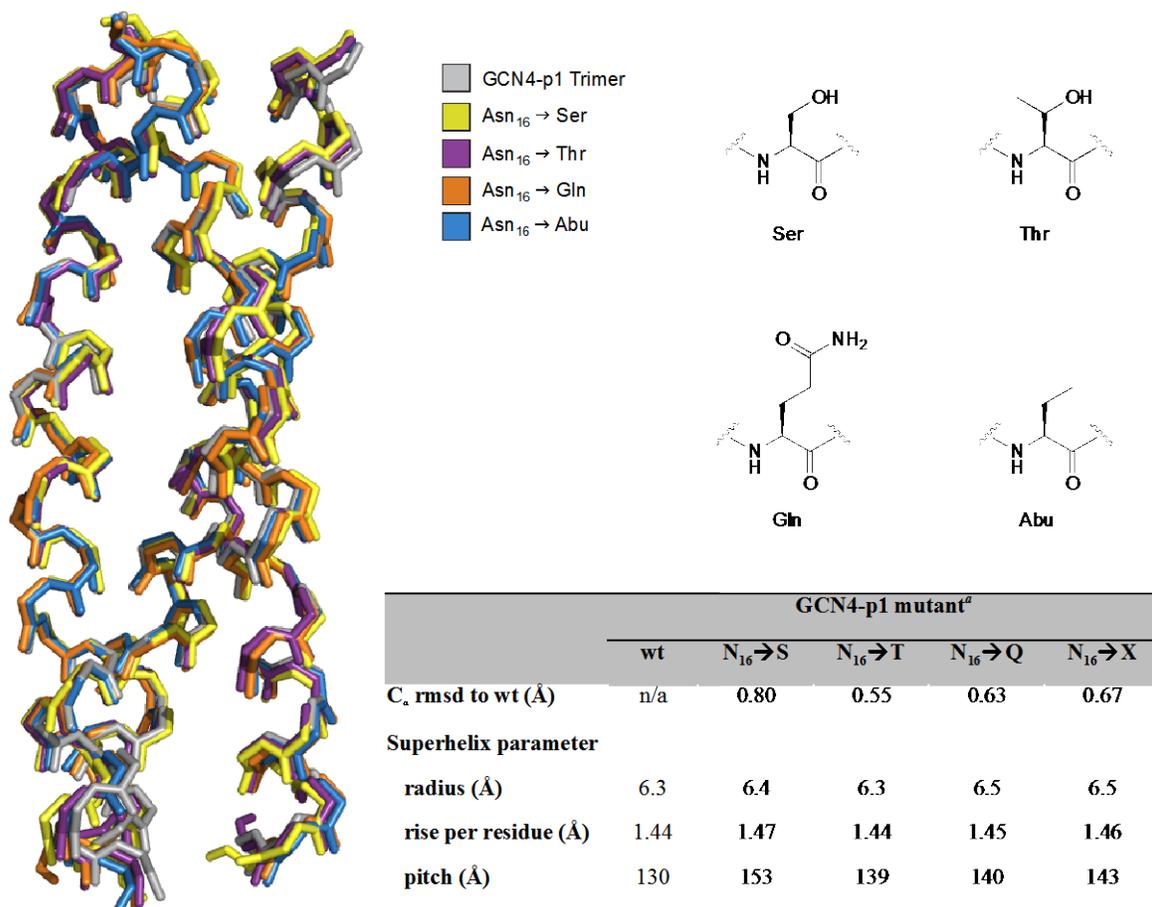
Figure 34. Overlays of the four published point mutants of GCN4-p1 and our trimer reveal little variation in the location of the backbone atoms.

Comparisons of superhelix parameters show no major discrepancies in the overall fold of these coiled coils.

	GCN4-p1 mutant[a]				
	wt	$N_{16} \rightarrow S$	$N_{16} \rightarrow T$	$N_{16} \rightarrow Q$	$N_{16} \rightarrow X$
C_a rmsd to wt (Å)	n/a	0.80	0.55	0.63	0.67
Superhelix parameter					
radius (Å)	6.3	6.4	6.3	6.5	6.5
rise per residue (Å)	1.44	1.47	1.44	1.45	1.46
pitch (Å)	130	153	139	140	143

polar Asn residue in the hydrophobic core. Bioinformatics studies have looked at the influence of residue identity at the *a* position in discriminating between dimeric and trimeric coiled-coil oligomerization states. The conclusions of these studies confirm Asn in the hydrophobic core favors a dimeric oligomerization state. However, a search of an online database of coiled coils[100]

in the PDB indicates this control is not absolute, as 9 homotrimeric parallel folds were found with

an Asn at an *a* position. In most of these proteins, a pocket is formed in the proximity of the Asn

residue and usually contains an ordered water or ion—similar to the GCN4-p1 trimer structure

(Table 6). The formation of this pocket could contribute to the destabilization of trimeric species

relative to the dimeric species where the hydrophobic core is well packed.

Table 6. PDB entries with an *a* position Asn that are homotrimeric parallel coiled coils.[a]

PDB ID[b]	Resolution (Å)[c]	Core Asn Residue Position #	Volume (Å^3)	Molecule in Core
1EBO	3.00	85	49	Chloride
1WP7	2.20	155	50	Water
1WT6	1.60	37	100	Water
		51	34	Water
2FYZ	2.20	133	50	Water
		148	120	Nothing
2IEQ	1.75	14	206	Water
2POH	2.10	63	67	Water
2W6B	2.80	608	35	Nothing

[a] Identified by a search of the CC+ database [b] PDB accession code [c] Refinement resolution of the crystal structure.

As mentioned above, the sequences used in our structure determination contained a C-

terminal amide, whereas the previously published GCN4-p1 structure contains a C-terminal

carboxylic acid. GCN4 based coiled-coils tend to be disordered past Gly$_{31,}$ and the C-terminus is

not usually resolved in electron density. Therefore, the change to the C-terminus seemed unlikely

to cause the oligomerization state switch. To explicitly test the role of the C-terminus in

oligomerization state switching, we synthesized the C-terminal carboxylic acid and crystallized it

in the buffers that gave rise to the dimeric and trimeric structures. A crystal that diffracted in the

same unit cell and symmetry was obtained from the dimeric buffer conditions using the C-terminal

carboxylic acid peptide. Using the trimeric buffer, a crystal form was obtained that was

crystallographically related to the C-terminal amide. The lattice that the C-terminal carboxylic acid GCN4-p1 crystallized in had lower symmetry, with three crystallographically independent coiled-coil trimers present in the asymmetric unit. Neither of these structures were refined. Overall, these data suggest the identity of the C-terminal group has no impact on the oligomerization state preference in GCN4-p1.

3.2   SOLUTION PHASE OLIGOMERIZATION STATE STUDIES

Obtaining crystal structures of proteins, while a challenging endeavor, provide a wealth of information on folding and functional properties that cannot be observed or fully explained from solution phase measurements. However, the process of crystallization requires a unique combination of conditions that can include high concentrations of the desired protein, extremes in pH or temperature, unusual buffers, as well as the addition of crowding agents, salts, metal ions and/or cryo-protectants. These conditions help to establish contacts and interactions needed for the formation of a regular crystal lattice. The crystallization buffer is usually a far cry from the conditions used for many dilute, solution phase characterization techniques and, as such, crystal structures should not exclusively be used to predict solution phase behavior.

While GCN4-p1 and related mutants are some of the most extensively studied coiled-coil sequences, no published literature, to our knowledge, has suggested the GCN4-p1 sequence populates a trimeric oligomerization state in solution; though some have questioned whether it folds by a two-state mechanism.[101] We performed a series of experiments to examine the assumption that the dimer is the favored folded state in benign pH 7.0 buffer and to investigate if there is a measurable amount of trimer under other experimental conditions.

78

Analytical gel permeation chromatography (GPC) was used to compare the oligomerization state of native GCN4-p1[48] and GCN4-pII[48, 51] a known trimeric variant of GCN4-p1, in pH 7.0 phosphate buffer. Both peptides eluted as single peaks (Figure 35). Calibration of the column using globular protein molecular weight standards showed that the molecular weight corresponding to GCN4-p1's elution volume is slightly larger than expected for a dimeric coiled coil ($MW_{app}/MW_{calc} = 2.3$) and molecular weight for the trimeric GCN4-pII matched the expected weight of a trimeric coiled coil ($MW_{app}/MW_{calc} = 3.0$). The variation in the molecular weight of the dimeric species is not taken as an indication of higher order oligomer states as the variation is within the error range (~15%) of analytical GPC and the coiled coil's elongated shape may give a falsely larger molecular weight.
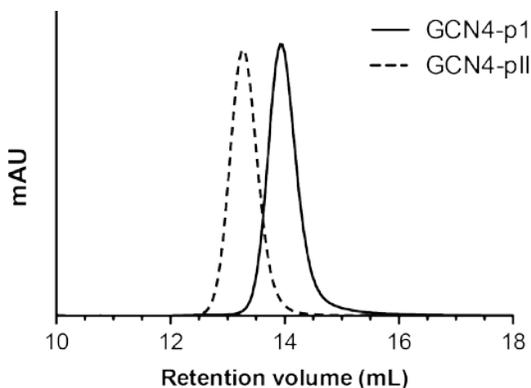
Figure 35. Gel permeation chromatography of GCN4-p1 and GCN4-pII, a trimeric coiled-coil, in 150 mM NaCl, 50mM HEPES pH 7.0 show two distinct retention volumes.

A molecular weight calibration curve indicated the trimer eluted at a volume that corresponds to 3.0 x MW_{calc} and the dimer eluted at a volume 2.3 x MW_{calc}.

We carried out circular dichroism (CD) scans and thermal melts of 100 µM GCN4-p1 in pH 7.0 phosphate buffer, which were consistent with previously reported data. These measurements served as a baseline for the comparison to behavior in solution conditions mirroring the dimer and trimer crystallization buffers. We carried out thermal melts at 100 µM peptide in both crystallization buffers with and without PEG to isolate the effects of precipitant from those of the high salt and/or low pH (Figure 36). The data suggest the coiled coil is slightly destabilized in the both crystallization buffers without PEG relative to pH 7.0 phosphate most likely because of the pH and/or salt content. The addition of PEG to each crystallization buffer increased the stability of the coiled coil. The effect was more pronounced in the trimer buffer (ΔT_m = +8˚C) than the dimer buffer (ΔT_m = +4˚C).
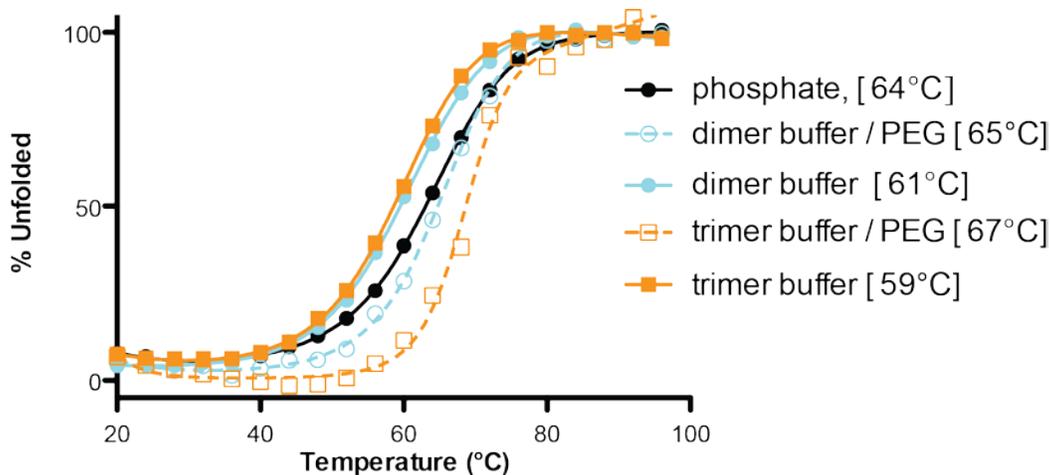
Figure 36. Circular dichroism thermal melts for GCN4-p1 in different buffers.

Solution conditions from top to bottom of the legend are: 10 mM phosphate, pH 7.0; 0.1 M sodium acetate, 0.1 M sodium citrate, pH 5.3; 0.1 M sodium acetate, 0.1 M sodium citrate, 25% w/v PEG 4000, pH 5.3; 0.2 M ammonium sulfate, 0.1 M MES, pH 6.6; 0.2 M ammonium sulfate, 0.1 M MES, 30% w/v PEG 5000 monomethylether pH 6.6. Curves are shown for fits to a two-state unfolding transition, with melting temperatures (T_m) indicated in brackets.

The outlier point at 72 °C in sample 5 was observed in two independent experiments; the origin is not clear, but it's presence does not impact the T_m determination.

Another method that can be used to determine the oligomerization state in proteins where folding requires self-assembly is concentration-dependent CD. Folding and self-assembly are coupled in GCN4-p1.[61] Owing to its high folded stability, GCN4-p1 cannot be fully denatured within the sensitivity limits of CD; therefore, this analysis must be performed under partially denaturing conditions, a method which has been previously applied to study the oligomerization state in coiled coils.[95] The molar ellipticity was monitored at 222 nm in 10 mM phosphate buffer, pH 7.0 with 6 M urea. The concentration of urea chosen was determined to give the best range of folded population over the concentration range of peptide used in the experiment (400 µM to ~ 1 µM). The data were fit using a model that allowed the number of chains in the associated state (n) to float giving a value of $n = 2.26 \pm 0.08$ (Figure 37). An oligomerization state of 2.26 lacks physical meaning, but the result suggests the concentration depending folding does not follow a simple monomer-dimer equilibrium under the conditions of the experiment.

In addition to CD studies, we performed sedimentation equilibrium analytical ultracentrifugation (SE-AU), in collaboration with Darrel McCaslin at the University of Wisconsin Madison, under several buffer conditions to examine their effect on the oligomerization state of GCN4-p1. The buffer conditions examined were: (A) pH 7.0 phosphate buffer with added NaCl (GPC buffer), (B) pH 5.3 acetate with added sodium citrate (dimer crystallization buffer without PEG), (C) pH 6.6 MES buffer with added ammonium sulfate (trimer crystallization buffer without PEG) and (D) pH 7.0 phosphate buffer with added 6 M urea (CD titration buffer). Over the concentration range of this experiment, the analysis suggests there
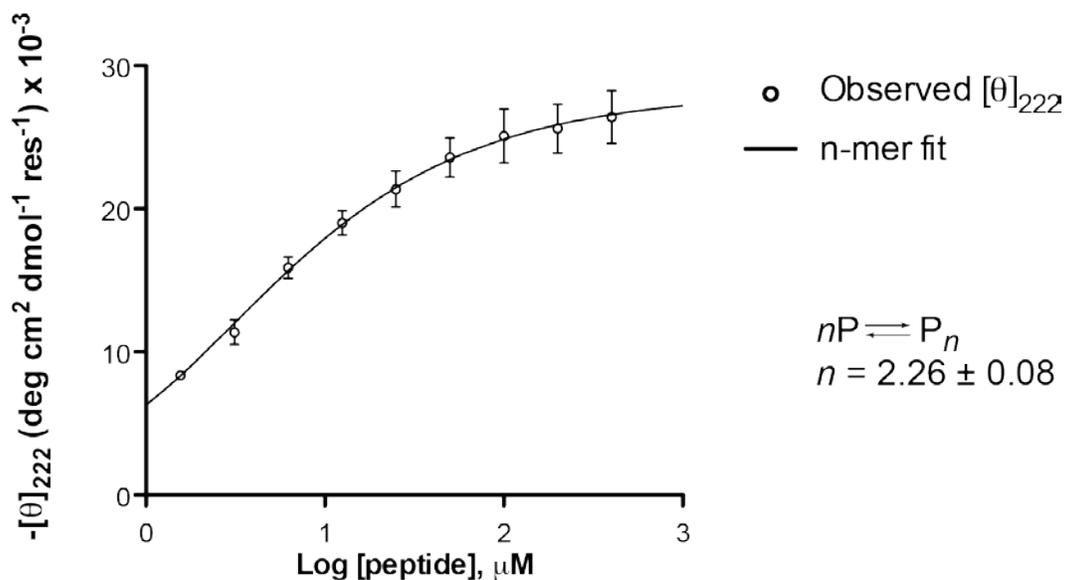
Figure 37. Concentration dependent molar elipticity of GCN4-p1 at 20 °C in 10 mM phosphate, 6 M urea, pH 7.0.

The curve is the best fir of the data to a self-association model of monomer to *n*-mer, where *n* is allowed to float.

there is a single molecular weight species in the examined buffer conditions. A computed partial specific volume was used to identify the dimer as the species in solution (Table 7). However, there is some uncertainty in the treatment of the sample with 6 M urea. In the presence of high concentrations of urea, corrections are need to the partial specific volume to account for the preferential interaction of water with urea.[102] Application of a known method using this calculation leads to an MW_{app}/MW_{calc} = 2.22, which agrees with the concentration dependent CD data. However, the method that corrects the partial specific volume was developed to provide accurate analysis under fully denaturing conditions, which is not the case for our system. Therefore, the dimeric form of GCN4-p1 best describes the preferred oligomerization state under the buffer conditions tested here.

Table 7. Sedimentation equilibrium results for GCN4-p1 under different buffer conditions.

Buffer	MW_{obs} (Da)[a]	MW_{obs}/MW_{calc}[b]
0.05 M phosphate, 0.15 M NaCl, pH 7.0	8120 ± 60	2.01 ± 0.01
0.1 M sodium acetate, 0.1 M sodium citrate, pH 5.3	7550 ± 40	1.87 ± 0.01
0.1 M MES, 0.2 M $(NH_4)_2SO_4$, pH 6.6	8280 ± 60	2.05 ± 0.01
0.05 M phosphate, 6 M urea, pH 7.0	8970 ± 70	2.01 ± 0.02

[a] apparent molecular weight (MW_{obs}) determined from fit to a single-species model; [b] ratio of apparent molecular weight by AU to that of the monomeric peptide (MW_{calc}).

The SE-AU studies presented here imply the salt/buffer components are not enough to lead to a measureable amount of trimer in solution. These experiments, however, excluded PEG from the solution conditions to prevent complications from a sedimenting buffer component. PEG is used as a precipitating agent in crystallization buffers and is often added to solutions as a non-ionic crowding agent to simulate a crowded cellular environment. The role of crowding on protein thermodynamics has been studied extensively in recent years.[103, 104] Theoretical calculations suggest the crowding effect of PEG increases protein association affinity, and the increase in association affinity is greatest for systems with more chains in the assembly (i.e., a crowding agent would have a more pronounced effect on a trimer than a dimer).[104] Thermal melts of GCN4-p1 in the crystallization buffers with added PEG show an increase in stability relative to both crystallization buffers without PEG and pH 7.0 phosphate buffer. The increase in stability is greater in the trimer promoting buffer. This observation suggests the addition of PEG shifts the equilibrium towards trimer under certain buffer conditions. The PEG itself cannot be solely responsible for the formation of the trimeric species as a similar concentration is used in the dimer buffer conditions. We hypothesize the combination of buffer conditions and the crowding effect of PEG contributes to the shift in equilibrium from dimeric to trimeric oligomerization state in the crystallization conditions.

3.3   SUMMARY

Collectively, the data presented suggest the native GCN4-p1 sequence can adopt either a dimeric or trimeric oligomerization state depending on the environment. Consistent with a large body of published work, we found that this coiled coil favors the dimeric state in pH 7.0 phosphate buffer. This state is also favored in pH 5.3 acetate/citrate buffer with or without added PEG. In pH 6.6 MES/ammonium sulfate with added PEG, a high-resolution crystal structure of GCN4-p1 in a trimeric oligomerization state was obtained. We also found evidence of a trimeric coiled coil population in solution under partial denaturing conditions by concentration-dependent CD, however, sedimentation equilibrium experiments suggest a single dimeric species in solution.

The strongest direct evidence we have for the trimeric state of GCN4-p1 is the high-resolution crystal structure. While crystal lattice contacts could help to select for either the dimeric or trimeric oligomerization state over the course of crystallization, an existing amount of either oligomerization state would need to be present to assume this. Examination of different components and conditions of the crystallization buffer helped to determine those that may influence the dimer-trimer equilibrium. The high concentration of protein needed for crystallization could shift the equilibrium towards the trimeric fold via Le Chatelier's principle. This hypothesis is contradicted by the SE-AU data. The fits from the SE-AU, from multiple solution conditions, all indicate a single dimeric species. Although the concentration of peptide in the SE-AU experiments do not reach the concentrations found in the crystallization conditions, the same solution of peptide when mixed with different buffers gave rise to dimeric species, arguing against a concentration dependent equilibrium shift. The lack of trimer in the SE-AU conditions argues a component of the crystallization buffer is responsible for the equilibrium shift. As

discussed in the previous section, PEG acts as a crowding agent and can impact the folded stability of the coiled coils with a greater impact on assemblies with higher oligomerization states.

It is our hypothesis that the GCN4-p1 is on the verge of two folded coiled-coil oligomerization states—a dimeric coiled coil that has been observed previously and a trimeric coiled coil reported here. The preferred oligomerization state is context dependent. Under partially denaturing conditions, the trimer may form to some extent, though the dimer is still favored. The addition of a crowding agent, such as PEG, which mimics a cellular environment, may shift the equilibrium to favor trimers under certain conditions. The biological relevance of the formation of the trimer is strengthened by a paper in which the authors replace the coiled-coil domain of the heat shock transcription factor (HSF), a trimeric DNA binding domain, with the GCN4-p1 sequence.[105] The GCN4-HSF chimeras with the leucine zippers of GCN4-p1 and binding domains of HSF were able to bind three box DNA with affinity similar to wild type. The authors proposed the proximity of the DNA boxes enforced the stochiometry and changed the preferred oligomerization state of GCN4-p1. Our data suggest the crowded cell environment, along with the templating effect of the DNA, may have aided the GCN-HSF chimeras in retaining wild type like functionality.

Through a collaboration with Prof. Lillian Chong and Dr. Reza Salari at the University of Pittsburgh, we were able to gain computational data about the relative stabilities of the two oligomeric states using parallel tempering molecular dynamics. The extent of folding was monitored over a 1 µs simulation using the C_α RMSD from the crystal structure, helical content and chain dissociation as parameters. The two association states show similar melting temperatures within error regardless of the parameter used. The absolute values of the melting temperatures from these simulations are artificially high relative to experimental data due to the implicit solvent

model used. However, we are interested in the relative melting temperatures, so the systematic overestimation of the T_m in the MD calculation does not interfere with our analysis. Data from the MD simulations also allowed us to examine the relative free energies of folding as a function of temperature between the two oligomerization states. Below 45°C, the dimer is the more stable species. Between 45°C and 155°C, which is near the calculated melting temperatures, the dimer and trimer have similar stabilities suggesting both oligomerization states could be populated in this temperature range (Figure 38).
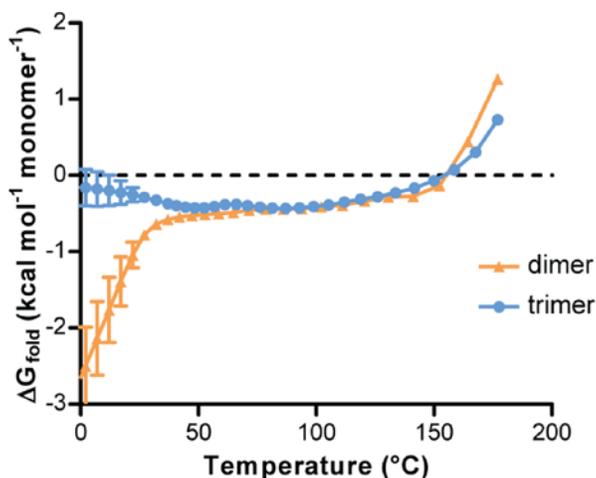
Figure 38. Folding free energies of the dimer vs trimer computed from parallel tempering simulations at various temperatures.

The folded state was defined based on C_α rmsd from the crystal structure. Uncertainty calculations are described in the Methods section.

In summary, we report the folding behavior of GCN4-p1, a well-studied, canonical coiled coil, is more complex than previous appreciated. We obtained a high-resolution crystal structure

of the native sequence in both dimeric and trimeric states. An extensive panel of solution-phase biophysical characterizations suggests the presence of both crystallographically observed oligomerization states under conditions with additional crowding agents, such as PEG, and under partial denaturing conditions. Parallel-tempering molecular dynamics studies found the relative energies of the two folded states to be quite similar. These finding have implications in ongoing efforts to create predictive algorithms for coiled-coil folding. They also impact design and mutational studies where control over coiled-coil oligomerization state is imperative. This work also highlights the importance of environmental conditions on interpreting solution phase biophysical studies. Frequently, biophysical measurements are taken in dilute aqueous solutions at or below room temperature; therefore one must consider the potential consequences of using these biophysical evaluations to infer behavior in crowded and complex cellular environments.

3.4   EXPERIMENTAL

3.4.1   Peptide Synthesis and Purification

Peptides were synthesized by Fmoc solid-phase synthesis using manual microwave-assisted protocols[82] or in automated fashion on a Protein Technologies Tribute Automated Synthesizer. NovaPEG Rink Amide resin was used to prepare the C-terminal amide, and Fmoc-Arg(Pbf)-Wang 100-200 mesh polystyrene Resin was used to prepare C-terminal acid. Peptides were cleaved from resin by treatment with 94% trifluoroacetic acid, 2.5% water, 2.5% ethanedithiol and 1% triisopropylsilane solution for 2 to 4 hours. After the peptide was cleaved from resin, it was precipitated from the filtered cleavage solution by addition of ~40 mL cold diethyl ether. The

precipitate was pelleted by centrifugation and the ether decanted. The peptide pellet was suspended in a mixture of 0.1% TFA in water and 0.1% TFA in acetonitrile for purification. Peptides were purified by HPLC on a C_{18} preparative column using 0.1% TFA in water and 0.1% TFA in acetonitrile gradients. HPLC fractions containing the product were combined, frozen, and lyophilized. Peptide identity was confirmed by mass spectrometry using a Voyager DE Pro MALDI-TOF instrument (monoisotopic [M+H]+ m/z for GCN4-p1: obsd. = 4034.7, calc. 4036.2). All samples were >95% pure by analytical HPLC on a C_{18} column.

3.4.2   Circular Dichroism Scans and Melts

Measurements were taken on an Olis DSM17 Circular Dichroism Spectrometer using 0.1 cm quartz cuvettes. Peptide concentration was determined by UV absorbance at 276 nm (ε = 1450 M^{-1} cm^{-1}) from the single Tyr residue in the GCN4-p1 sequence.[86] Samples of 100 µM peptide in buffer were prepared and scanned from 200 nm to 260 nm in 1 nm increments, an integration time of 5 seconds and a bandwidth of 2 nm at 20°C. A buffer blank was used to correct each spectrum and baseline molar ellipticity at 260nm. Variable temperature CD was taken by monitoring molar ellipticity at 222 nm from 20-96°C in 4°C increments with a 2 minute equilibration time between data points and an integration time of 5 seconds. Thermal melt data was fit to a two-state unfolding model to obtain the melting temperature (T_m).

3.4.3   Concentration Dependent Circular Dichroism

A 400 µM solution of peptide was prepared in 10 mM phosphate buffer, 6 M urea, pH 7.0. Serial two-fold dilutions were made into 10 mM phosphate buffer, 6 M urea, pH 7.0 to generate nine

samples with peptide concentrations ranging from 400 µM to 1.56 µM. Samples from 400 µM to 100 µM were measured in 1 mm quartz cuvettes, 50 µM to 6.25 µM in 2 mm quartz cuvettes, and 3.125 µM and 1.56 µM in 5 mm quartz cuvettes. All of the buffer solution used in the titration was taken from a common stock. CD measurements of each sample were performed on an Olis DSM17 Circular Dichroism Spectrometer. Molar ellipticity of the samples was monitored at 222 nm at 20.0°C using a 2 nm bandwidth and a 5 second integration time. Three independent samples of each concentration were measured to obtain the reported molar elipitcities.

The observed molar ellipticities were fit to the following model[95]:

Equation 6

$$nP_{mon} \rightarrow P_n$$
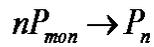
Equation 7

$$K = \frac{P_{mon}{}^n}{P_n}$$

Equation 8

$$P_{tot} = P_{mon} + nP_n$$

Equation 9

$$P_{tot} = P_{mon} + n\left(\frac{P_{mon}{}^n}{K}\right)$$

where are P_{tot}, P_{mon} and P_n the total concentration of peptide, the monomer, and the n-mer respectively, K is the dissociation constant and n is the number of the molecules in the associated state. The concentration dependent molar ellipticity is described by the following equations:

Equation 10

$$[\theta_{obs}] = [\theta_{coil}]f_{coil} + [\theta_{fold}]f_{fold}$$

Equation 11

$$[\theta_{obs}] = [\theta_{coil}](1 - f_{fold}) + [\theta_{fold}]f_{fold}$$

Equation 12

$$f_{fold} = n\frac{P_n}{P_{tot}} = n\frac{P_{mon}{}^n}{KP_{tot}}$$

Equation 13

$$[\theta_{obs}] = [\theta_{coil}]\left(1 - n\frac{P_{mon}{}^n}{KP_{tot}}\right) + [\theta_{fold}]\left(n\frac{P_{mon}{}^n}{KP_{tot}}\right)$$

in which $[\theta_{fold}]$ is the mean molar ellipiticity of the folded peptide, $[\theta_{coil}]$ is the mean molar ellipticity of the random coil, and f is the fraction of peptide in the state specified by the subscript. The value for $[\theta_{coil}]$ was determined experimentally from a thermal melt of GCN4-p1 in 10 mM phosphate buffer, pH 7.0 and 6M urea. The observed fully unfolded baseline was used as the value for $[\theta_{coil}]$. Using Mathematica (Wolfram Research), the best fit parameters values for the K, n, and $[\theta_{fold}]$ were determined using a non-linear fit model.

3.4.4 Gel Permeation Chromatography

GPC was carried out on a Superdex 75 10/300 column (10 x 300 mm, 24 mL bed volume, 13 mm average particle size, GE Healthcare). The column was equilibrated with 0.15 M NaCl in 0.05 M sodium phosphate, pH 7.0. Peptides were loaded onto the column (100 mL sample at 100 µM concentration in equilibration buffer) and eluted at a flow rate of 0.80 mL/min. A molecular weight calibration curve was obtained by fitting the elution volumes of 1 mg/mL solutions of BSA, ovalbumin, aprotinin, a 17-residue synthetic peptide (Ac-YEAAAKEAAAKEAAAKA-NH$_2$), and vitamin B12.

3.4.5   Sedimentation Equilibration Measurements

Sedimentation equilibrium measurements were carried out by Dr. Darrell McCaslin at the University of Wisconsin-Madison using a Beckman Coulter Model XL-A Analytical Ultracentrifuge. GCN4-p1 was prepared by simple dissolution in one of four buffers and used without further manipulation; the initial concentrations in each buffer were determined from spectra as recorded in the centrifuge using an extinction coefficient of 1490 M^{-1} cm^{-1} at 280 nm. The buffers were (A) 0.05 M phosphate, 0.15 M NaCl, pH 7.0 with peptide concentrations of 89, 220, and 415 µM; (B) 0.1 M sodium acetate, 0.1 M sodium citrate tribasic, pH 5.3, 245 µM GCN4-p1; (C) 0.2 M ammonium sulfate, 0.1 M MES, pH 6.6, 269 µM GCN4-p1; (D) 10 mM phosphate, 6 M urea, pH 7.0, 267 µM GCN4-p1. Buffer densities at 20°C were computed using density increment functions[106] as 1.010, 1.004, 1.012, and 1.102 g/mL respectively (a contribution for MES buffer was not available but it is likely that the ammonium sulfate is the dominant contributor). The partial specific volume of GCN4-p1 (0.748 cm^3 g^{-1}) was calculated using consensus values reported for the amino acid residues with a correction applied for the acetyl and carboxamide end groups.[107, 108] In the presence of denaturing concentrations of urea, the partial specific volume can be corrected for specific interactions of urea and water with the protein[102]; this corrections leads to an effective partial specific volume of 0.764 cm^3 g^{-1} for GCN4-p1. However, the data presented here suggest that the protein is in a mostly folded state in 6 M urea, so that the true value may lie closer to 0.748 cm^3 g^{-1}. The molecular weight of GCN4-p1 is 4038 Da including the terminal blocking groups.

Approximately 100 µL of a peptide solution was placed in one sector of a 1.2 cm pathlength, charcoal filled epon centerpiece with ~110 µL of the corresponding buffer added to the reference sector. Gradients were monitored at a nominal wavelength of 276 nm. Samples were

spun at various speeds at 20°C until gradients collected 3 or more hours apart were superimposable. The equilibrium data were analyzed following an approach similar to that described by Laue using software written by D.R.M. for Igor Pro (Wavemetrics, Inc, Lake Oswego, OR).

Under all buffer conditions, a single macromolecular species with a small contribution from non-sedimenting absorbance was able to describe the data. The weight average molecular weights derived from global fits of data in each buffer are summarized in Table 7. The variations are likely a reflection of the computed nature of the partial specific volume in various salts. Plots of the logarithm of the measured absorbance (after subtracting the fitted non-sedimenting absorbance) as a function of squared radial distance from the center of rotation; in such plots, a single species manifests as a series of straight lines with slopes proportional to the weight average molecular weight. The solid lines are based on the fitted weight average molecular weights shown in Table 7 and well account for the available data.

3.4.6   Parallel Tempering Molecular Dynamics

These experiments were performed by Prof. Lillian Chong and Dr. Reza Salari at the University of Pittsburgh. To determine the melting temperatures (T_m) of the dimer and trimer folded states of the GCN4-p1 leucine zipper in isolation, we used parallel tempering molecular dynamics (MD) simulations.[109, 110, 111] Parallel tempering is a widely used replica exchange enhanced sampling technique that involves simultaneously performing "replica" simulations of the system at different temperatures with the aim of making configurations at higher temperatures available to simulations at lower temperatures and vice versa.

To estimate T_m values of the dimer and trimer, three approaches were applied, each using a different order parameter to monitor unfolding in the parallel tempering simulations. In the first approach, an unfolded conformation was defined as having a C_a rmsd from the crystal structure that is more than one standard deviation above the average value at 37 °C in the simulations. The three c-terminal residues that were missing in the crystal structure of the dimer were not included in the rmsd calculations. In the second approach, a conformation was considered unfolded if the number of helical residues was more than one standard deviation below the average value at 37 °C (helical residues were defined as having $\varphi = -60 \pm 30$ and $\psi = -47 \pm 30$[112]). Finally, in the third approach, an unfolded conformation was defined as having at least one of the chains dissociated (i.e., beyond van der Waals distance of 4.5 Å). For each approach, the average fraction unfolded was plotted vs. temperature and the temperature of 50% unfolding was considered as the melting temperature.

To compute the folding free energies of the dimer and trimer at the temperatures of interest, we applied the multistate Bennett acceptance ratio (MBAR) method to our parallel tempering simulations as implemented in the PyMBAR package (https://simtk.org/home/pymbar).[113] Folding free energies were estimated using snapshots collected every 5 ns. The extent of folding was monitored using order parameters with non-discrete values (i.e. C_α rmsd and the helical content). Folding free energies at temperatures that were not included among the temperate replicas of the parallel tempering simulations (i.e. below 37 °C) were estimated by extrapolation using the MBAR method.[113] Uncertainties at each temperature were estimated by the asymptotic covariance matrix of the MBAR estimating equations.[113]

Starting structures for the simulations of the dimer and trimer folded states were prepared using the LEAP module in the AmberTools 1.5 package.[114] Heavy atom coordinates were taken

from the crystal structures of the corresponding oligomeric form. The three C-terminal residues of the dimer had not been resolved in the crystal structure and were added using the PyMol visualization software.[115] Crystallographic water molecules were removed and hydrogen atoms were added using the protonation states present in solution at pH=7. The removal of the two buried water molecules in the trimer crystal structure appears to not substantially affect the stability of the trimer within our implicit solvent model as the structure remained folded in the 500 ns-long standard MD simulation at 20˚C (C_α rmsd values of 2.0 ± 0.4 Å relative to the crystal structure. No cutoff for non-bonding interactions was used. To relieve unfavorable interactions, each starting structure was subjected to energy minimization in two stages, with position restraints applied to the heavy atoms in the first stage and no position restraints in the second stage.

3.4.7 Crystallization, Diffraction Collection and Structure Determination

Crystallization was carried out using the hanging drop vapor diffusion method. Drops were prepared by mixing 0.7 µL of peptide stock (10 mg/mL in water) with 0.7 µL of buffer and allowed to equilibrate at room temperature over a well containing 0.7 mL of buffer solution. Crystals of the GCN4-p1 dimer were obtained from a well buffer composed of 0.1 M sodium acetate pH 4.6, 0.1 M sodium citrate tribasic dihydrate pH 5.6, and 25% w/v PEG 4000. A single crystal was flash frozen in liquid N_2 after being soaked in the above buffer supplemented with 25% v/v glycerol. Crystals of the GCN4-p1 trimer were obtained from a well buffer composed of 0.2 M ammonium sulfate, 0.1 M MES monohydrate pH 6.5 and 30% w/v PEG monomethylether 5000. A single crystal was flash frozen in liquid N_2 after being soaked in the above buffer supplemented with 10% v/v glycerol. The GCN4-p1 C-terminal carboxylic acid dimer was crystallized in 0.3 M sodium acetate pH 4.6, 0.1M sodium citrate tribasic dihydrate pH 5.6, and 20% w/v PEG 4000. A single

crystal was flash frozen in liquid N_2 after being soaked in the parent buffer supplemented with 25% v/v glycerol. The GCN4-p1 C-terminal carboxylic acid trimer was crystallized by mixing 0.7 μL of a 20 mg/ml stock solution and 0.2 μL of the buffer described above for the C-terminal carboxamide trimer. A single crystal was flash frozen in liquid N_2 after being soaked in the parent buffer supplemented with 10% v/v glycerol. Diffraction data were collected on Rigaku Saturn 944 CCD using CuKα radiation. d*TREK was utilized to index, integrate and scale the collected data.

Structure refinement was carried out using CCP4.[89] Phaser was used for molecular replacement and previously published GCN4 coiled-coil derivatives were used as models; the dimer and trimer structures were solved using PDB entries 2ZTA[91] and 1IJ2[96], respectively. A combination of refinement programs were used to complete the structure: Refmac[92] for automated refinement, Coot[93] for manual model building, ARP/wARP[94] for solvent building, and Phenix[90] for construction of composite omit maps. Phenix was also used to compare the symmetry between the trimer crystal forms of the C-terminal carboxamide and C-terminal carboxylic acid ($P2_1$, $a = 34.6$ Å, $b = 58.5$ Å, $c = 101.3$ Å, $g = 90.5°$); this analysis indicated a shared primitive cell between the two lattices. Superhelix parameters and cavity volume were calculated using the TWISTER[116] and the CASTp[117] server, respectively. Buried surface area values were calculated using the PISA server.[117]

4.0 FLUOROPHORE FUNCTIONALIZATION AND ENERGY TRANSFER IN COILED-COIL SUPRAMOLECULAR POLYMERS

We previously demonstrated a modular synthetic route to branched subunits consisting of two α-helical coiled-coil domains connected by a small organic linker. These subunits were shown to fold and self assemble into supramolecular polymers in aqueous solutions. We demonstrated the relationship between linker identity and assembly flexibility[85] as well as the correlation between the association affinity of the coiled coil and assembly size.[118] The control of assembly flexibility and size would be useful in applications where the spatial relationship between functional groups is critical to material function. Here, we designed a series of fluorophore-functionalized peptides and subunits to test our hypothesis that programmed changes to the assembly properties can affect energy delocalization.

4.1 DESIGN OF FLUOROPHORE LABELED PEPTIDES AND SUBUNITS

The addition of fluorophores to our subunits requires consideration of several design elements. Our goal is to examine if energy delocalization is affected by assembly properties. To achieve this, we must have a system with a Förster Resonance Energy Transfer (FRET) donor-acceptor fluorophore pair. The FRET pair must reside within the Förster radius of each other in the assembly, and fluorophore attachment chemistry must be orthogonal to the cross-linking chemistry. Based on the above design considerations, we envisioned placing the donor fluorophore

on the subunit and using a non-cross-linked acceptor labeled peptide as a capping group to measure

FRET and antenna efficiency in the assemblies (Figure 39).
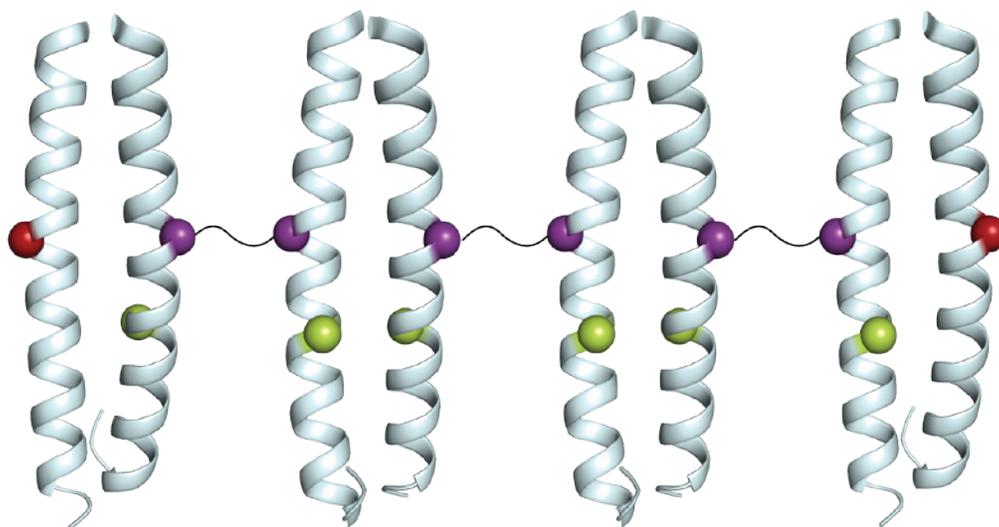
Figure 39. The design of a system based on our self-assembling subunits for examining the changes in FRET with

changing subunit properties.

The purple spheres represent the point of cross-linking attachment; green spheres represent the location of donor

fluorophore on the subunit and the red spheres represent the location of an acceptor fluorophore on a capping

peptide.

The identity of the donor and acceptor fluorophores is a major factor in designing a system

with the maximum energy transfer efficiency possible and should be chosen carefully. Regardless

of the linking chemistry utilized to attach the fluorophores to the peptides, the energy transfer rate,

k_t, is governed by[119]:

Equation 14

$$k_t(r) = \frac{1}{t_D}\left(\frac{R_0}{r}\right)^6$$

where t_D is the decay lifetime of the donor fluorophore in the absence of acceptor, R_0 is the Förster radius of the pair and r is the distance between the fluorophores. The energy transfer rate is heavily dependent on the Förster radius, which is the distance at which half of the acceptor fluorophore are excited by the donor and can be determined by:

Equation 15

$$R_0 = 0.211\left(\kappa^2 n^{-4} Q_D J(\lambda)\right)^{0.5}$$

κ^2 is describes the relative spatial orientation of the two transition dipoles of the fluorophores and is generally assumed to be 2/3, n is the refractive index of the medium, Q_D is the quantum yield of the donor, and $J(\lambda)$ is the spectral overlap of the donor emission band and acceptor absorbance band. A FRET pair with high efficiency will have a donor with a high quantum yield and good emission spectral overlap with the acceptor excitation band. The donor and acceptor must also be positioned on the scaffold so they remain within the Förster radius of the pair, which generally ranges from 20-70 Å. The Förster radius can be experimentally determined but several sources exist for calculated Förster radii of certain fluorophore pairs.[120]

We chose a green-emitting donor fluorescein (ex. 488/em. 530) fluorophore and the red-emitting acceptors tetramethylrhodamine (ex. 557/em. 576) and Texas Red (ex. 596/em. 615). Fluorescein and rhodamine-based dyes have good spectral overlap and Förster radii in the range of 50 Å (Figure 40).[120] This is a reasonable distance as our coiled coil is 46 Å in length. We examined several variants of fluorescein and tetramethylrhodamine dyes in our subunits and capping groups prior making this choice. Other brighter, green emitting dyes, such as Oregon Green, had linking chemistries with difficulties in either ligation to the peptide or purification of

the product, or their manufacture was discontinued. We found direct excitation of the acceptor at the donor excitation wavelength to be problematic with tetramethylrhodamine-based dyes.
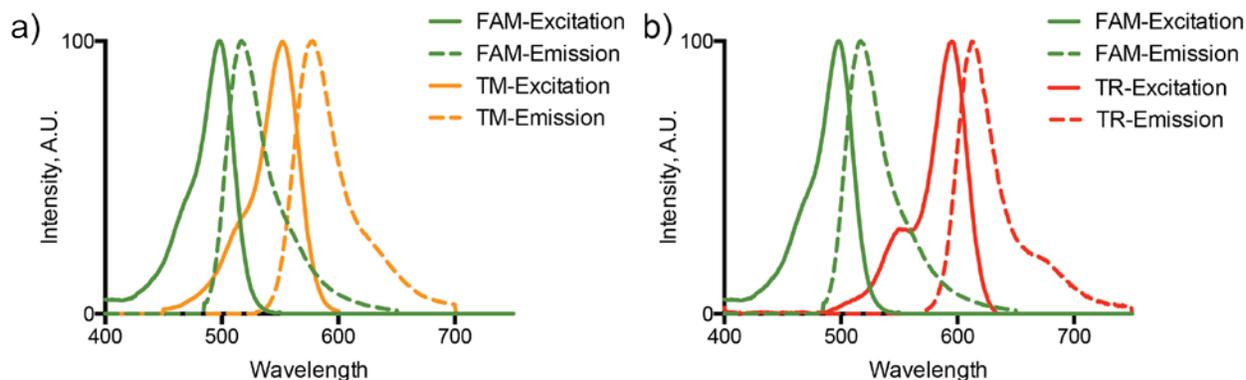
Figure 40. Spectral overlap of fluorescein emission with tetramethylrhodamine (a) and Texas Red (b) excitation band.

The dashed green line in each spectrum is the emission spectrum of fluorescein. There is a large overlap with the tetramethylrhodamine excitation band (solid orange line) (a) and a moderate amount of overlap with the red-shifted Texas Red excitation band (solid red line) (b). Spectral data obtained from the ThermoFisher Spectra Viewer. *www.thermofisher.com*

4.2 PRELIMINARY EXPLORATION OF SYNTHETIC ROUTES

From a design perspective, there are two avenues to the synthesis of a fluorophore-functionalized subunit: subunit synthesis followed by fluorophore functionalization or vice versa. Performing the subunit cross-linking first followed by the attachment of the donor fluorophore (Figure 41) has the advantage of requiring less fluorophore for labeling compared to preparing a large batch of fluorophore-labeled peptide prior to cross-linking. It also minimizes the potential for the

99

fluorophores to degrade by exposure to light. There are many linking chemistries available for fluorophore attachment; however, the reactivities of amino acid side chains and the desire to keep the scaffold amenable to heterologous expression led us to use reactions between thiols and maleimides or halo-acetamides. An additional Cys was added to peptide **3** to replace *e* position Glu$_{20}$—data from the crystal structure of GCN4-p1 suggested replacement of this residue would not disrupt any salt bridges and would be within the Förster radius of the dyes. This Cys required an orthogonal protecting group that was stable to the conditions of peptide cleavage and subunit formation, but could be selectively removed under conditions that did not disrupt the formed cross-link.
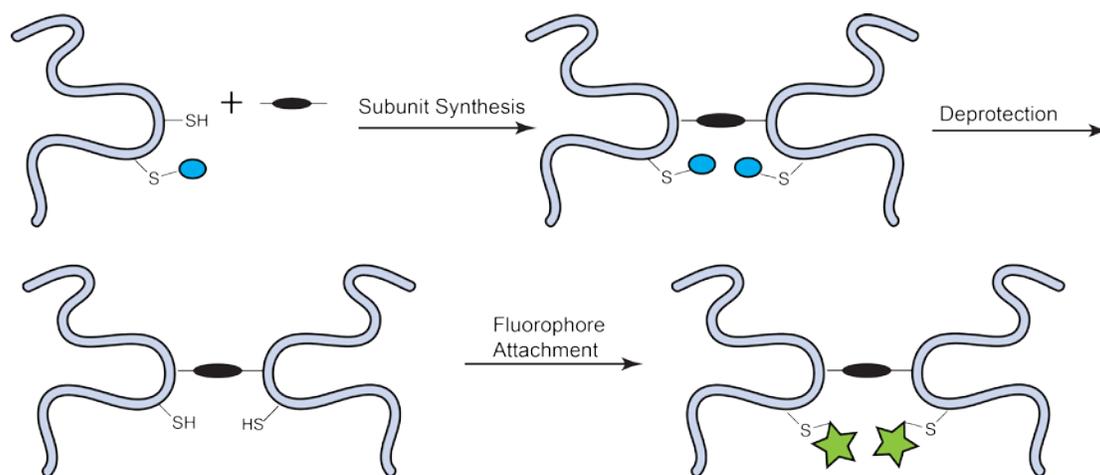
Figure 41. A proposed synthetic route for fluorophore labeled subunits where subunit formation occurs before labeling.

In this approach, an additional Cys with an orthogonal protecting group is introduced by mutating Glu$_{20}$. The subunit is first formed and then the orthogonal protecting group is removed from the subunit. A fluorophore is then attached using a thiol reactive moiety.

We initially tried an acetamidomethyl (Acm) protecting group on Cys_{20}; however, the removal of this group required either harsh reagents[121] or the protecting group removal was low yielding under milder conditions. Next, a 4-methoxybenzyl (Mob) was used as the orthogonal Cys protecting group. The Cys(Mob) containing peptide withstood the cross-linking reaction and the Mob group was easily removed (Figure 42).[122] When the subunit with the additional unprotected Cys was subjected to fluorophore attachment, a mixture of starting material, singly labeled, and doubly labeled subunit was obtained. The complexity of the reaction mixture was further compounded by the fluorophore consisting of a mix of two regioisomers and the maleimide-thiol reaction leading to two stereoisomers for each regioisomer.
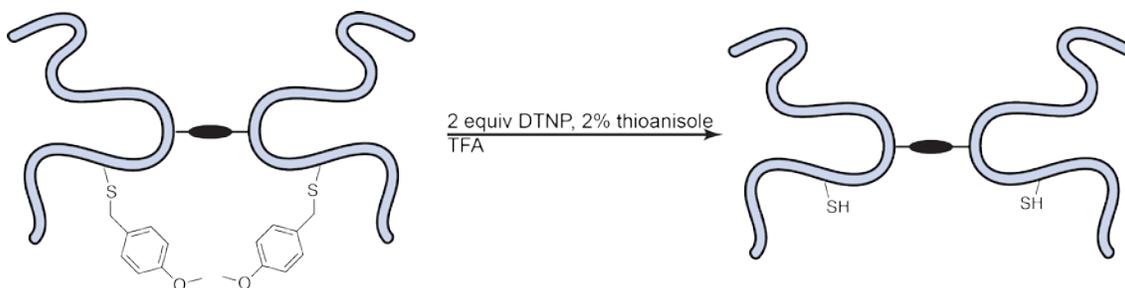
Figure 42. Removal conditions for the MOB protected Cys.

The MOB group was effectively removed under these conditions; however, the attachment of the donor fluorophore and purification of the reaction mixture using the subunit produced in this reaction was problematic.

With the difficulties encountered above, we examined fluorophore attachment to the peptide first, followed by cross-linking (Figure 43). This route has the advantage producing uniformly labeled subunits (assuming no degradation of the fluorophore). The main disadvantage

101

is the greater quantity of fluorophore required for labeling and the need to minimize light exposure during multiple purification and synthesis steps. Instead performing a solution phase coupling of the donor fluorophore to the peptide, we chose to attach the fluorophore on resin to reduce the purification steps and potential exposure to light. Because of the increased amount of fluorophore needed, we chose to attach the fluorophore through a coupling between Lys and relatively inexpensive 5-carboxyfluorescein (5-FAM). The chosen Lys reside would be protected with an Alloc group, which can be selectively cleaved on resin. The Lys deprotection and subsequent fluorophore coupling yielded quantitatively, singly labeled peptide. The labeled peptide was subjected to similar cross-linking conditions as previously described and produced fully labeled subunit. While this method requires more fluorophore and greater care to protect the product from light, the ability to easily obtain completely labeled subunit made this the preferred synthetic route.
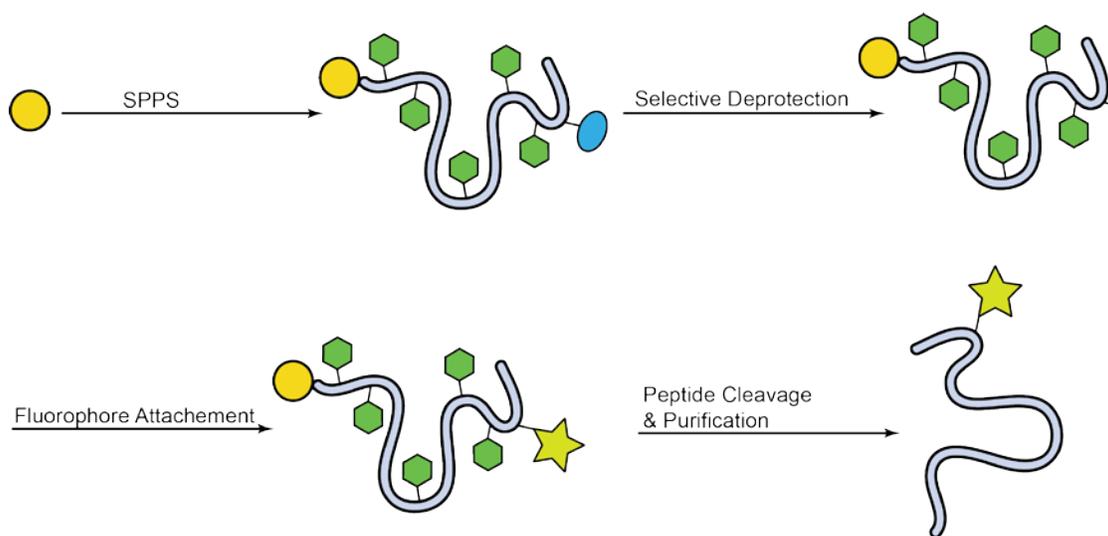
Figure 43. A proposed synthetic route for the synthesis of fluorophore labeled subunit where fluorophore labeling occurs before subunit formation.

Fluorophore labeling is carried out on resin after the peptide is synthesized. An orthogonal protecting group is removed from Lys and the fluorophore then attached. The peptide is cleaved and purified from resin and carried forward into the cross-linking reaction.

The capping peptide with the acceptor fluorophore underwent optimization as well. We first tried attaching tetramethylrhodamine (TM) fluorophore using a bromoacetamide group as the reactive moiety with Cys_{14}. Literature[120] suggested a large excess of dye was needed for efficient labeling; however, we found a slight excess of peptide was necessary for efficient coupling (Figure 44). We used the TM-labeled capping peptide (**23TM**) for initial studies of the dimer and subunit systems. In the course of these experiments, two complicating factor arose that led us to eventually change acceptor fluorophores. Tetramethylrhodamine is known to form dimers at concentrations ranging from 650 μM[123] to ~100 μM[124] in aqueous solutions or when two dyes are in close proximity.[125] The distance between Cys in associated coiled coils is potentially close enough for dimerization to occur. We also observed that significant tetramethylrhodamine emission was present upon excitation at the donor excitation wavelength. Because of these two factors, we chose to switch to a Texas Red (sulforhodamine 101) fluorophore attached through a maleimide moiety (**23TR**). We were able to attach this fluorophore using only a slight excess of fluorophore (Figure 44). As discussed earlier, the use of maleimide as a reactive group towards Cys gives rise to multiple isomers. We were able to collect the different isomers and combined them for use in the fluorescence experiments.

103
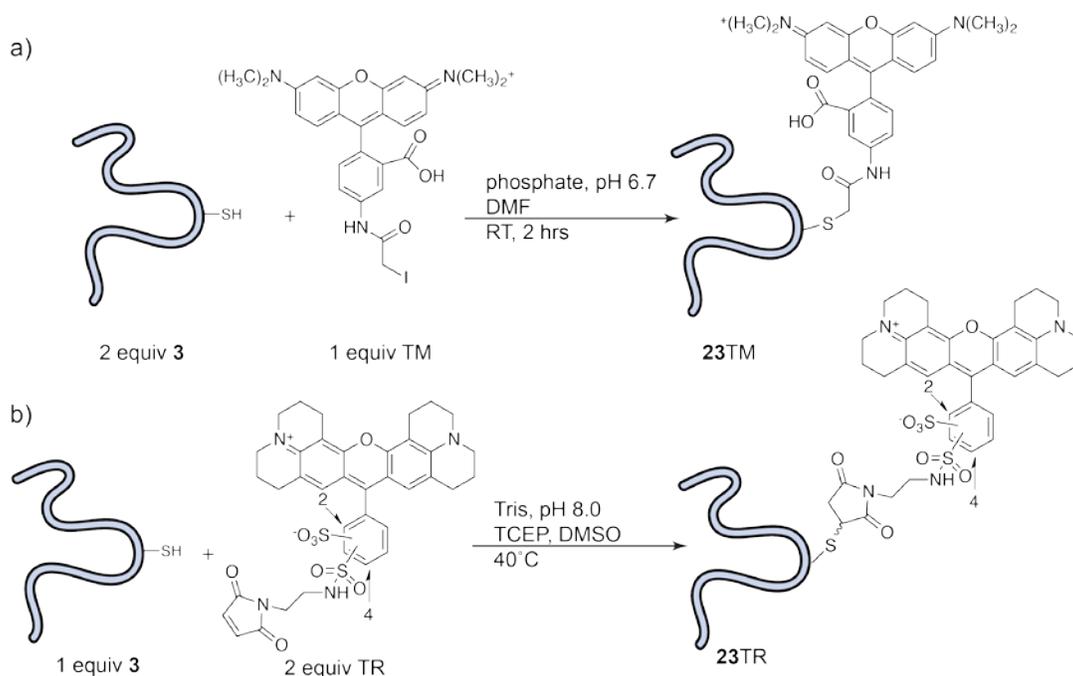
Figure 44. Labeling of peptide **3** with acceptor fluorophores tetramethylrhodamine and Texas Red. 1 equivalent of TMR in DMF was added to 2 equivalents of peptide **3** in phosphate buffer pH 6.7 at room temperature and allowed to react for 2 hours (a). 1 equivalent of **3** in 100 mM Tris pH 8.0 is mixed with 2 equivalents of Texas Red maleimide in DMSO and allowed to react for 1 hour at 40˚C (b).

4.2.1 Fluorescent Subunit Synthesis

With the synthetic route for fluorophore labeled subunits chosen, we had several options for the placement of the donor fluorophore using naturally occurring Lys residues. Lys_3 was chosen for fluorescein attachment because of the direction of SPPS (C-terminus to N-terminus), which would make Lys_3 the most accessible to on resin reactions. Placement of fluorescein at the N-terminus should also be minimally disruptive to peptide folding and subunit formation, and
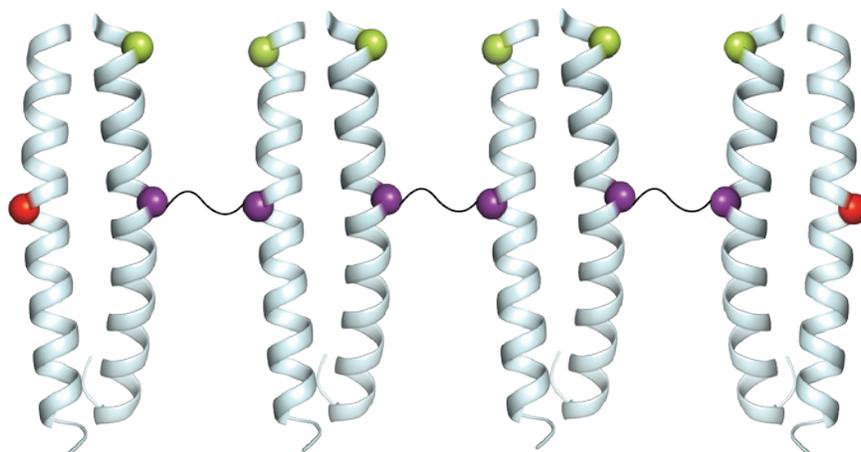
Figure 45. The design of a system based on our self-assembling subunits for examining the changes in FRET with changing subunit properties.

The purple spheres represent the point of cross-linking attachment; green spheres represent the location of donor fluorophore on the subunit and the red spheres represent the location of an acceptor fluorophore on a capping peptide.

allow the donor and acceptor fluorophores to remain within their Förster radius (Figure 45). We synthesized a mutant of GCN4-p1 containing Lys_3 protected with an allyloxycarbonyl (Alloc) group using SPPS. The Alloc group was selectively removed on resin by treatment with tetrakis(triphenylphosphine) palladium(0) and phenylsilane in DCM under inert atmosphere.[126] The removal of the Alloc group was confirmed by microcleavage and analysis by MALDI-MS. 5-FAM was coupled to the free amine using an established protocol (Figure 46).[127] From this point on, the FAM labeled peptide (**24**) was protected from light as rigorously as possible. Peptide **24** was cleaved from resin and purified using HPLC. Identity was confirmed by MALDI-MS and purity assessed by analytical HPLC.
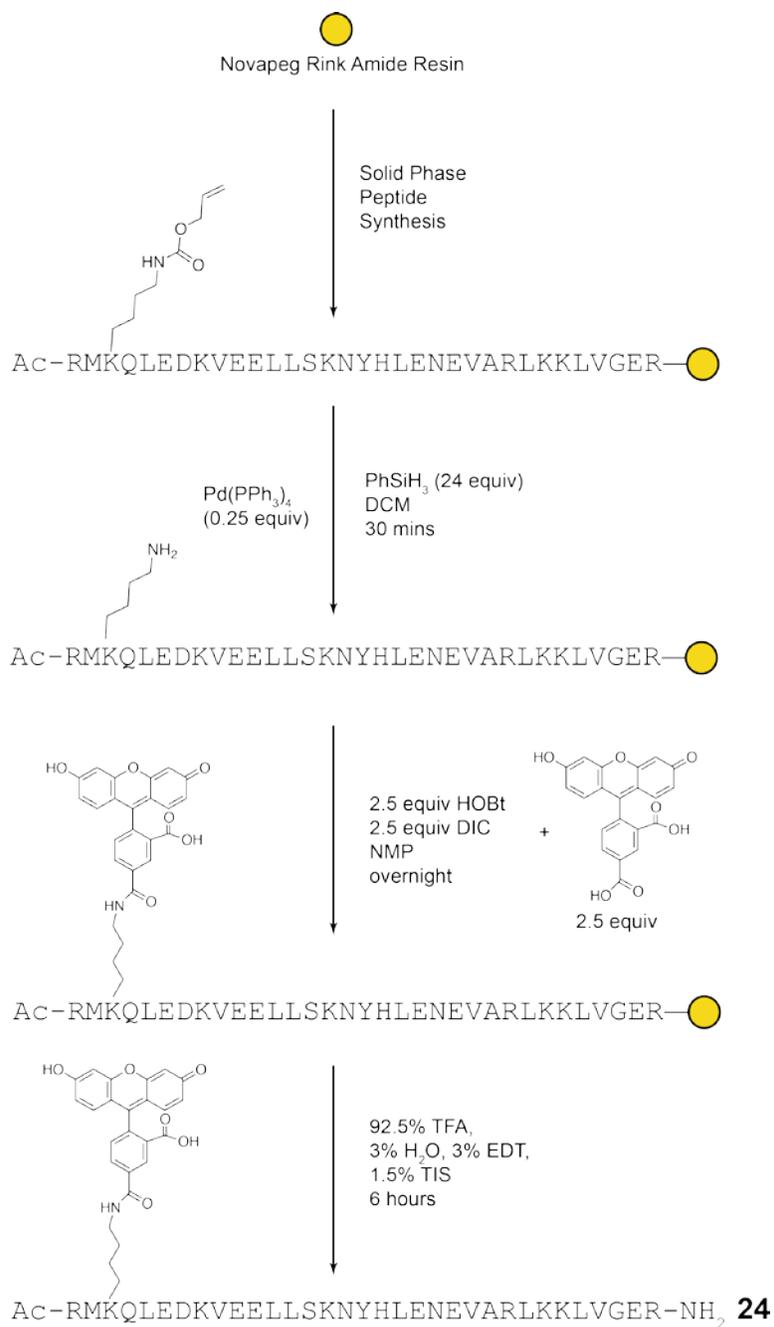
Figure 46. The deprotection of Lys(Alloc) was carried out under inert conditions after the full length peptide was synthesized.

This was followed by 5-carboxyfluorescein attachment using literature precedented conditions. The labeled peptide was cleaved from resin and purified using HPLC.

106

We also synthetized the $Ser_{14}\rightarrow Cys$ mutant of GCN4-p1 with Lys_3 coupled to 5-carboxyfluorescein (peptide **25**) to be crosslinked to form fluorescein labeled subunits. Peptide **25** was subjected to the subunit cross-linking reaction using linkers **5b** and **6**. The cross-linking reactions performed with peptide **25** had a larger amount of disulfide by-product than seen for the peptides from Chapter 2 and showed lower solubility under the established protocol. To aid in solubility, the percentage of acetonitrile was increased to 10%. We found bubbling argon through the reaction mixture for ~10 minutes before adding the first aliquot of linker and lowering the temperature of the water bath to ~50˚C decreased the amount of disulfide formed. The reaction was quenched after 2 hours with a mixture of water and acetonitrile with 0.1% TFA. Spin filters were used to concentrate the products, and subunits **26** and **27** (Figure 47) were purified by semi-preparative HPLC.
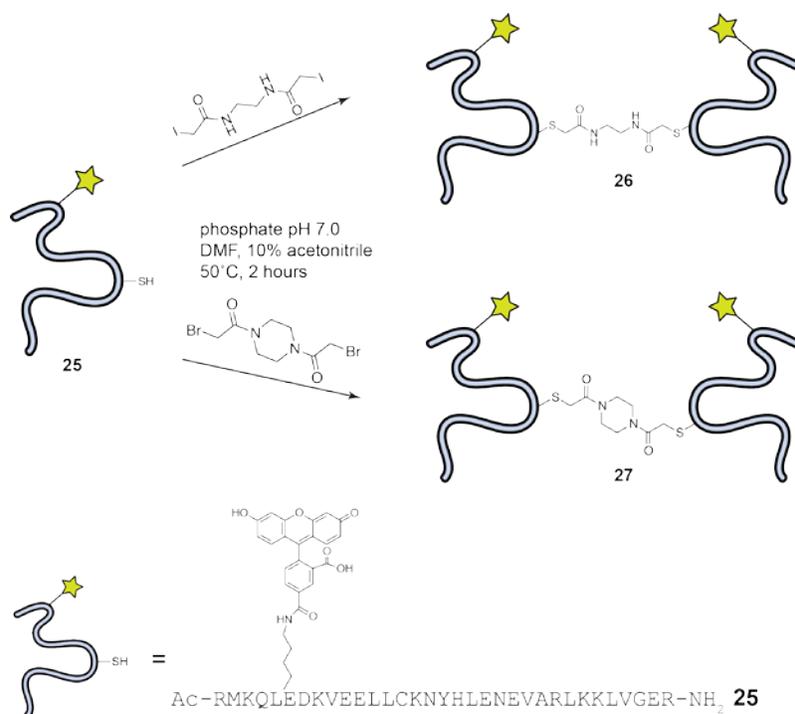
Figure 47. The fluorophore labeled **25** was subjected to the cross-linking reaction with slight modifications to aid in the solubility of the product.

4.3 EXPERIMENTAL DESIGN

Conducting fluorescence experiments requires the balance of several experimental factors including fluorophore pair selection, sample optical density and instrument parameers. The experimental design for our system has an additional level of complexity, as we have to balance the sample optical density while maintaining a concentration where the peptides are fully folded. The native GCN4-p1 coiled coil has a reported K_d of 0.57 µM[74], which should allow us the flexibility to lower the concentration of peptide into a range that is suitable for fluorescence measurements.

Fluorescence spectroscopy is a very sensitive technique and because of this, only small amounts of fluorophore are needed to obtain reliable data. Optical densities (ODs) at the excitation wavelength usually fall in the range of 0.1-0.3, and ODs higher than this range can lead to a non-linear response or distortion of the emission spectra.[119] Fluorescein at an OD of 0.3 in a cell of 1 cm path length is at a concentration of ~ 4 µM for peptide with a single fluorophore. In the case of our subunits, which have two fluoresceins per subunit, the concentration of subunit drops to 2 µM. While the native GCN4-p1 sequence should be fully folded at these concentrations, the addition of a fluorophore to the peptide may have a slight destabilizing effect and the subunit may begin to unfold. To address this problem, we chose a cell with a path length of 2 mm, allowing the concentration of subunit to be around 10 µM. The increase in concentration is also desirable as it better mimics the conditions under which we characterized the assemblies by DLS (200 µM).

The OD of the samples was held around 0.3 for the donor fluorophore ensuring a similar amount of fluorescence from both the dimer and subunit systems would be observed. Fluorescence experiments were carried out on with slit width of 3 x 3 nm at room temperature unless otherwise noted. The donor fluorophore was excited at 488 nm and the acceptor was excited at 557 nm (**23TM**) or 588 nm (**23TR**). Samples containing subunits **26** and **27** were allowed to equilibrate for 4 hours, based on previous DLS data, before any fluorescence measurements were taken to allow the assemblies to reach a stable size.

While studying the fluorescence properties of peptide **24** and subunits **26** and **27**, the observation was made that the subunits showed a significantly decreased donor fluorescence emission compared to peptide **24**. We calculated a value for the Förster radius for self-quenching between two fluorescein molecules of 37 Å (published literature values are around 40 Å[120]). The distance between fluorescein molecules across a coiled coil is approximately 15 Å, and the distance

within a subunit is similar. To test if self-quenching was the cause of the decease in fluorescence emission from fluorescein, we carried out a series of fluorescence experiments at room temperature and at an elevated temperature (60˚C). At the elevated temperature, the coiled coil should be completely dissociated and no assemblies present in solution. This removes the possibility of cross-coil quenching. The fluorescence of peptide **24**, subunit **26** and 5-FAM at an OD of around 0.3 was measured at room temperature and at 60˚C. 5-FAM was included as a control to make sure any changes in fluorescence could be attributed to protein unfolding and not degradation of the dye at high temperatures. The 5-FAM sample, even at room temperature, showed a distorted, red-shifted emission spectrum. Changing the slit width to 1x1 nm restored the expected shape and maxima of the emission spectrum. No decrease in fluorescence emission was observed in the 5-FAM at elevated temperature with the decreased slit width (Figure 48).
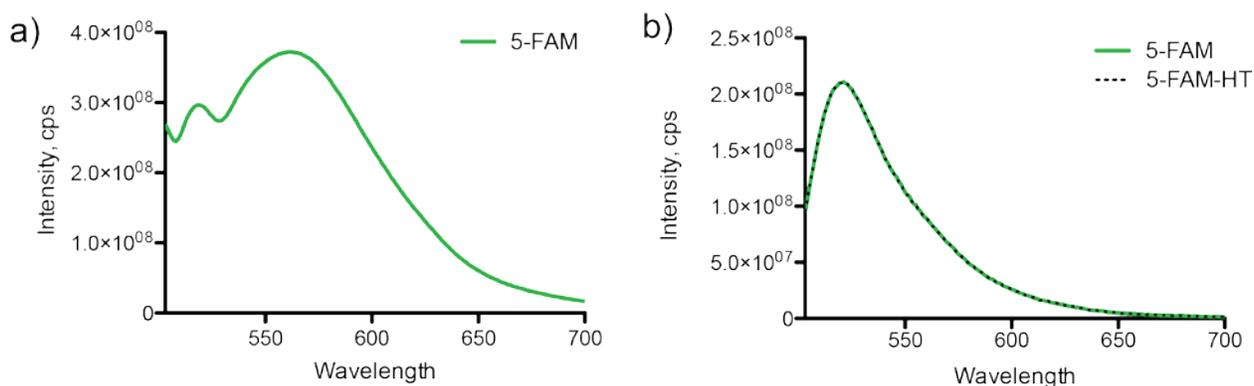
Figure 48. Steady state fluorescence spectra of 5-FAM in 100 mM phosphate pH 7.2 at an OD of ~ 0.3. Using a slit width of 3 x 3 nm, there was significant distortion and red shifting of the maximum emission at room temperature (a). When the slit width was closed to 1 x 1 nm, the distortion and red shift of the emission maxima disappeared (b). Heating the sample to 60˚C did not significantly change the fluorescence emission output (5-FAM-HT). Excitation $\lambda = 488$ nm.

Peptide **24** had the expected emission spectrum at room temperature with a fluorescence emission intensity of 2.8 x 10^8. At the increased temperature however, the fluorescence intensity increased significantly (1.3 x 10^9) and the spectrum also displayed a similar distortion and red shift as the 5-FAM sample. Decreasing the slit width eliminated the distortion while still providing adequate intensity at room temperature.
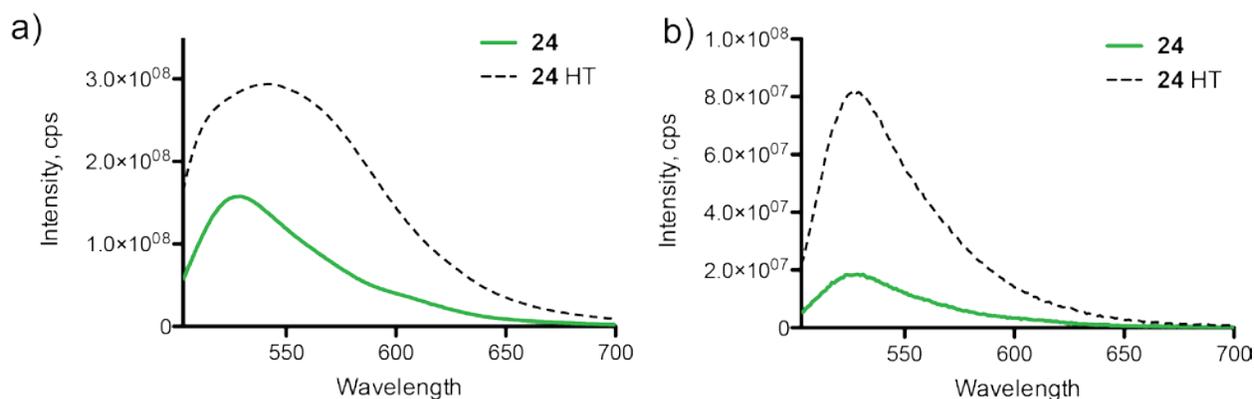
Figure 49. Steady state fluorescence of peptide **24** in 100 mM phosphate pH 7.2 at an OD of ~ 0.3. Measurements taken with a slit width of 3 x 3 nm shows a distortion and red shifting of the emission output at high temperature (a). When the slit width is reduced to 1 x 1 nm, the distortion and red shift in emission maximum is eliminated (b). Excitation λ = 488 nm.

Subunit **26** also showed an increase in fluorescence emission intensity at higher temperatures. No distortion or red shift of the fluorescence maximum was observed when the slit width was 3x3 nm. These findings suggest two sources of self-quenching in our system. First, there is some donor-donor quenching across the coiled coil shown by an increase in donor fluorescence

intensity when peptide **24** is unfolded by heating (Figure 49). We also see an increase in fluorescence intensity when subunit **26** is unfolded (Figure 50). The unfolding of the subunit would eliminated the cross-coil self-quenching; however, the length of the linker holds the denatured peptide chains in close enough proximity that even when the peptides are unfolded quenching could occur. This is confirmed by subunit **26** having some increase in fluorescence emission intensity at an elevated temperature, but not to the degree that the dimer only system increases intensity.
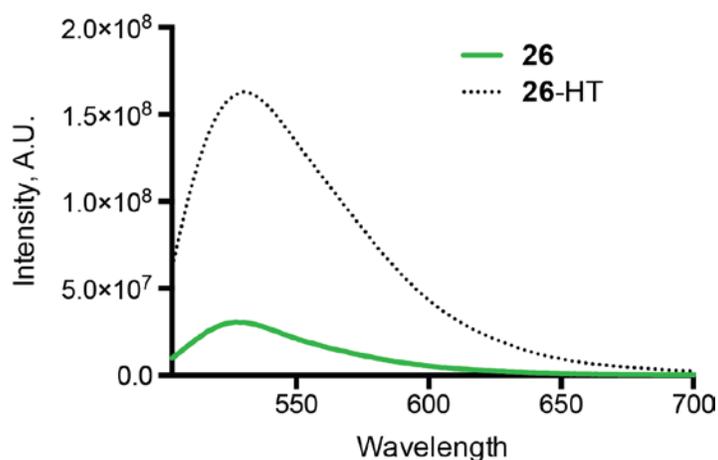
Figure 50. Steady state fluorescence of subunit **26** in 100 mM phosphate pH 7.2 at an OD of ~ 0.3 with 3 x 3 nm slit widths.

The subunit exhibited an increase in fluorescence when heated above the melting temperature of the coiled coil.

Excitation λ = 488 nm.

We carried out similar experiments on peptide **23TR**. The acceptor fluorophore was excited at both 488 and 588 nm with a 3x3 nm slit width at an OD of ~ 0.3. A small decrease in

the fluorescence emission was observed when peptide **23TR** was heated and excited at 588 nm. Texas Red dyes have been shown to have minimal temperature dependence (compared to similar dyes) with higher temperatures leading to small decreases in fluorescence emission (Figure 51). We also examined the potential for this fluorophore to be excited at the donor excitation wavelength (488 nm). Looking at the excitation spectra, 488 nm is at the beginning of Texas Red's excitation band, but the magnitude at 488 nm is less than 10% of the maximum at 594 nm (Figure 51b). Both peptide **23TR** and free Texas Red dye showed significant fluorescence emission when excited at this wavelength. Observing some fluorescence emission was expected; however the intensity of the fluorescence emission was surprising given the excitation spectra. At this time, we do not have an explanation for this behavior.

The final experimental parameter that needed to be established was the ratio of donor labeled subunit to acceptor labeled capping group. Other studies using similar systems vary the ratio of donor to acceptor fluorophore to examine variations in the efficiencies of exciton transfer. We initially tested three ratios of labeled subunit to acceptor peptide: 10:10, 10:5 and 10:1. The concentration of labeled subunit was held constant at ~10 µM and the amount of acceptor peptide changed so the decrease in donor emission could be used as a measure of FRET
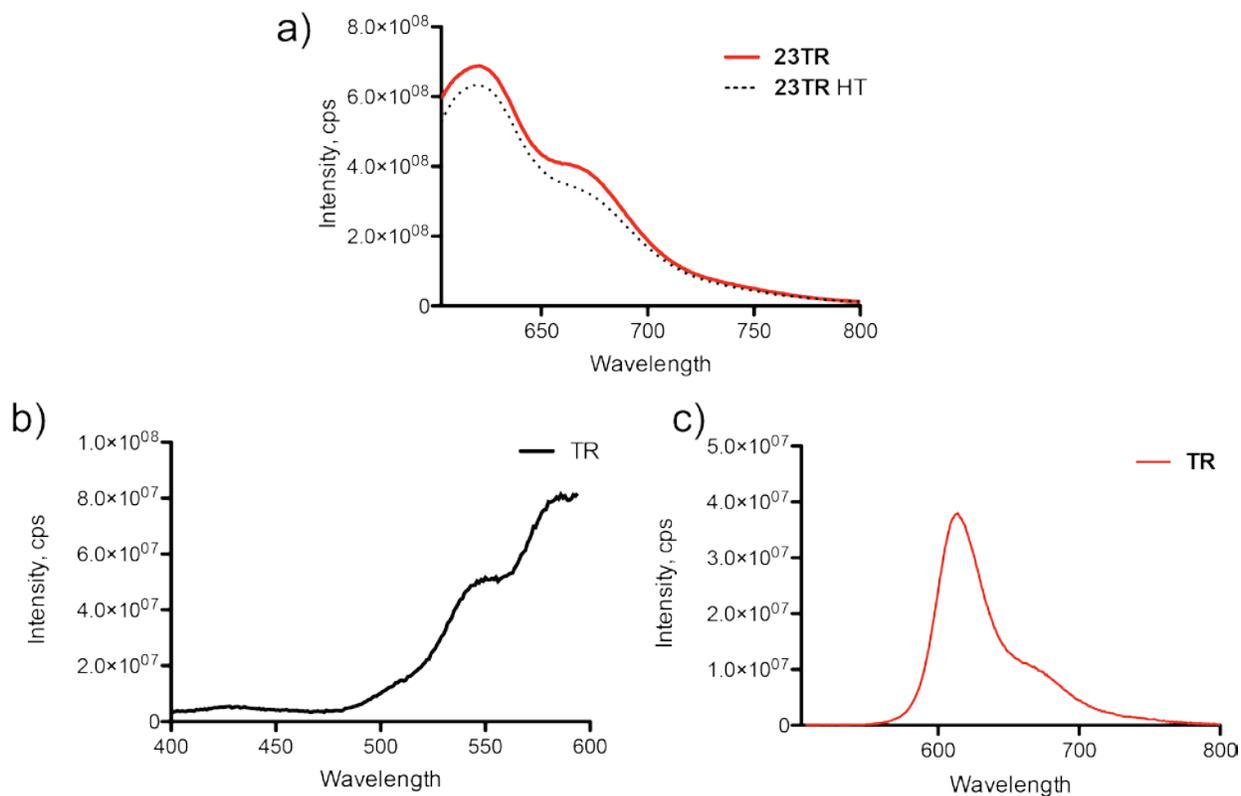
Figure 51. Emission spectra of peptide **23TR** 100 mM phosphate pH 7.2 at an OD of ~ 0.3 with 3 x 3 nm slit width at room temperature and 60˚C.

Peptide **23TR** exhibited a decrease in the fluorescence emission intensity at the higher temperature. Excitation λ = 588 nm. (a). The excitation spectra for the free TR dye in buffered solution shows only a small absorbance at 488 nm. Spectrum recorded at 609 nm (b). Fluorescence emission of TR dye when excited at 488 nm (c).

efficiency. After considering the overall goal of the study to examine how the programmed changes in the assemblies affect energy delocalization, the 10:1 ratio was chosen. The acceptor labeled peptide works a capping agent that stops assembly growth in these samples. Having a smaller amount of capping peptide would allow the larger assemblies to form and better represents the behavior of the assemblies without capping groups and in the more concentrated solutions.

4.4   ENERGY TRANSFER EFFICIENCIES

With the experimental design developed, we next examined if changes in the properties of the subunits impacted the energy delocalization across the assemblies. In this study, we are primarily interested in how the linker flexibility affects the overall FRET efficiency and antenna efficiency, which is a measure of how energy is delocalized along a series of donor fluorophores. We have already shown small changes in the subunit linker flexibility can have an effect on assembly properties. We sought to determine if the behavior of a light-harvesting functional group placed on the subunits is sensitive to these properties.

There are several methods to interpret fluorescence data to analyze FRET efficiency and they fall into two categories based on experimental design: donor quenching and acceptor enhancement. Due to peptides **23TM** and **23TR** exhibiting fluorescence emission when excited at fluorescein's excitation wavelength and some fluorescein emission bleeding into the emission band of both dyes, the donor quenching method was chosen to analyze FRET. The energy transfer efficiency was calculated by[119]:

Equation 16

$$E = 1 - \left(\frac{OD_D I_{DA}}{OD_{DA} I_D} \right)$$

where OD_D is the absorbance at the donor at the donor excitation wavelength, OD_{DA} is the absorbance of the donor and acceptor mixture at the donor excitation wavelength, and I_D and I_{DA} are the fluorescence intensities of the donor with and without the acceptor present. Antenna efficiency was measured by[34]

Equation 17

$$AE = \frac{I_{AEx.488}}{I_{AEx.588}}$$

where $I_{A(ex.\ 488)}$ is the excitation of the acceptor from donor excitation and $I_{A(ex.\ 588)}$ is excitation resulting from direct acceptor excitation.

The dimer system was studied first to assess the experimental design. Peptide **24** was mixed with either peptide **23TM** or **23TR** in a 1:1 ratio. Several concentrations and path lengths were used to examine FRET in the **24** and **23TM** dimer system with concentrations ranging from 25 µM of each peptide to 2.5 µM of each peptide in 100 mM phosphate buffer pH 7.2. The FRET efficiency was around 50% regardless of the concentration with **24** and **23TM**, though caution should be exercised when interpreting fluorescence data with OD higher than 0.3—which is the case in concentrations greater than 15 µM. The observed FRET efficiency of 50% was expected based on the statistical mixture of assemblies in the solution (25% **23TM** dimer, 25% **24** dimer and 50% mixed dimer). Again, it should be noted that peptide **23TM** showed some amount of fluorescence emission when excited at the donor wavelength (488 nm). The 1:1 mixture of **24** and **23TR** had observed FRET values of closer to 30%.

Subunit **26** and **27** were also examined for their FRET efficiency and the antenna effect. Measurements on these samples were carried out with 10 µM of subunit with 1 µM of either peptide **23TM** or **23TR** in 100 mM phosphate buffer pH 7.4 with 3 x 3 nm slit widths (Figure 52). Both samples showed evidence of FRET and we were able to measure antenna efficiencies; however, the values for each subunit fluctuated between measurements. The general trend indicated that the FRET efficiency tended to be higher in subunit **27** but the antenna efficiency
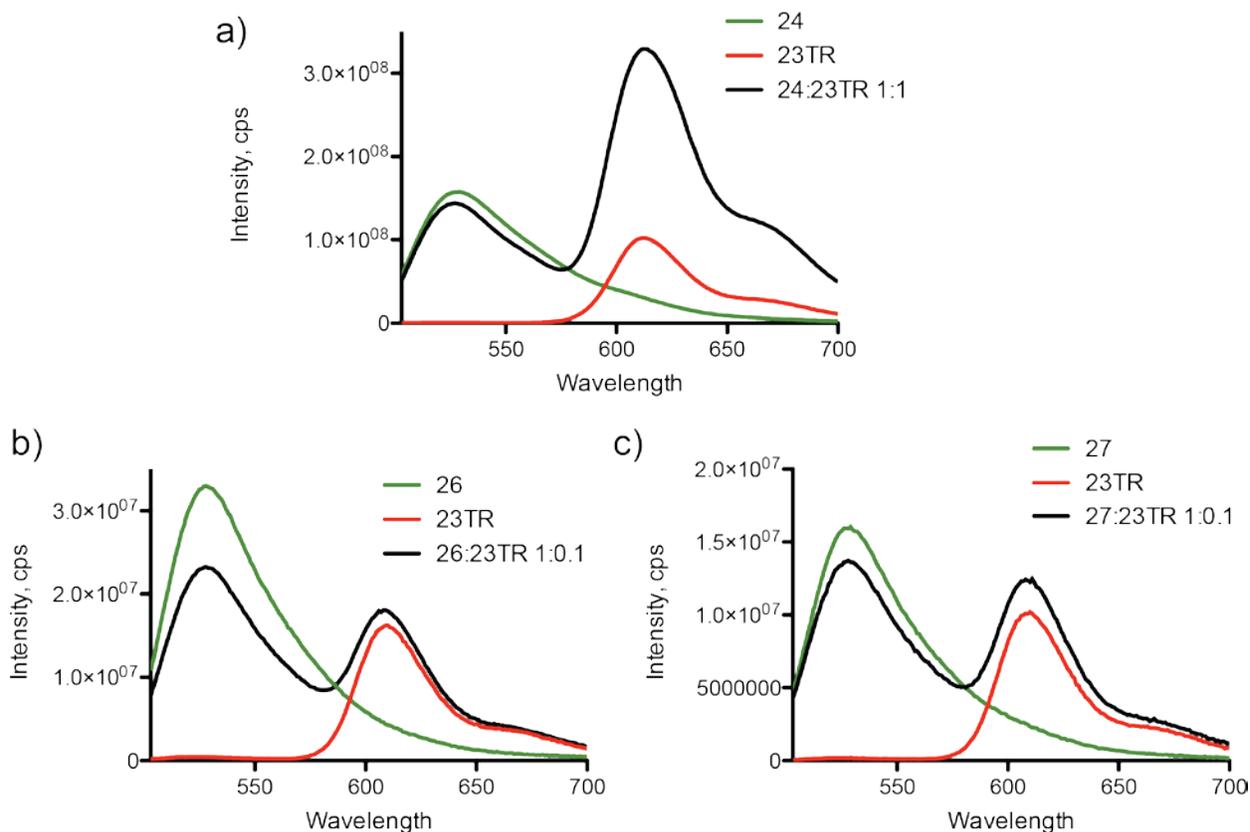
Figure 52. Fluorescence data for the donor and acceptor labeled systems.
Peptides were in 100 mM phosphate buffer pH 7.4 and measurements were taken with 3 nm x 3 nm slit widths. The
dimeric system was measured with a ratio of 1:1 donor:acceptor fluorophore with the optical densities around 0.3
(a). Subunits 26 (b) and 27 (c) were also measured in a 1:0.1 donor subunit to acceptor ratio. The excitation
wavelength for the donor is 488 nm.

was higher in subunit **26**. The flexibility of subunit **26** may allow for more donor fluorophores to
be within the Förster radius leading to an increase in the amount of acceptor emission. At this
point, we do not have a clear explanation of the fluorescence emission data.

4.5   SUMMARY AND FUTURE DIRECTIONS

To conclude, we developed a synthetic route to fluorophore labeled variants of the dimer peptides and our subunits. We also began to explore how the properties of the subunit impacted the energy delocalization in larger assemblies. The synthesis of the fluorophore-functionalized subunits was not trivial and required several design iterations. The optimal synthetic route involved first labeling the peptide and then forming the subunit. Developing a robust characterization protocol for this system also presented a unique challenge that necessitated a balance of the OD of the fluorophores and the concentration of the subunits and peptides.

Fluorescence measurements on the fluorophore labeled dimeric coiled coil indicated the system was able to undergo FRET and had FRET efficiencies ranging from 30% to 50%. The experiments with the larger assemblies were more difficult to interpret. There was some evidence of FRET in both subunits based on donor emission quenching. Subunit **27** generally had a greater FRET efficiency than subunit **26**; however, the differences were small and interpretation of this result should be approached with caution as the large amount of donor-donor quenching may complicate results. Some acceptor emission enhancement was also observed and used to measure the antenna efficiency. In this case, subunit **26** had a higher antenna efficiency than **27**. Again, this data should be interpreted carefully as acceptor emission was present when the samples were excited at the donor excitation wavelength and acceptor enhancement was relatively small.

We chose fluorescein as our donor fluorophore based on our synthetic scheme and the cost of the dye even though its potential to self-quench was known. Research with similar systems used an Oregon Green dye in place of fluorescein; however, Oregon Green is also known to self-quench[128] and in the reported system, the donor fluorophores were placed at similar intervals as ours. The acceptor dyes, tetramethylrhodamine and Texas Red, exhibited fluorescence emission

118

when excited at the donor's excitation wavelength. The degree of acceptor emission when excited at the donor excitation wavelength was low compared to direct excitation of the acceptor. The Texas Red dye had a lower emission intensity at the donor excitation wavelength than the tetramethylrhodamine due to the red shift of its excitation band.

The combination of significant donor self-quenching and acceptor emission when excited at the donor excitation wavelength makes interpretation of the collected fluorescence data, especially of the assemblies, difficult. We chose to analyze the FRET efficiency of the system based on the degree of donor quenching to determine if the changes energy transfer is coupled to subunit identity and assembly properties. We do not know the degree of self-quenching in each of the subunits and therefore, the self-quenching may be masking differences in the donor-acceptor transfer in these assemblies. The self-quenching in our system could be a major factor in the small amount of acceptor enhancement. Exciton transfer between donors along the assembly chain may be occurring, but if the movement of the exciton is stopped through quenching before it can reach the acceptor, we would see an artificially low amount of acceptor emission enhancement. The small degree of acceptor enhancement makes quantitative comparisons of antenna efficiencies our subunit unreliable.

One way to address the problem of donor self-quenching is to use a different method of determining FRET. An alternative method compares the changes in the excitation spectrum to the absorption spectrum and has been used to analyze similar protein systems[34] and dendrimers.[129] Another way to potentially examine the differences caused by the flexibility of the linker is to make longer assemblies through using a coiled coil with a more thermally stable fold. The longer assemblies should increase any differences in energy transfer between the two systems. Future experiments could also incorporate a donor fluorophore that has a lower propensity for self-

quenching or reduce the number of donor fluorophores per subunit from two to one. We were also able to show that there are two sources of donor self-quenching—cross coil and within in the subunit. Self-quenching of the donor fluorophore could be minimized by designing a slightly longer linking group.

4.6   EXPERIMENTAL

4.6.1   Synthesis of 5-carboxyfluorescein Labeled Peptide (24 and 25)

Peptides were synthesized by Fmoc solid phase peptide synthesis (SPPS) using microwave assisted protocols[82] and NovaPEG rink amide resin. Cys and His were coupled at room temperature for 45 minutes. Lys_3 was protected with an Alloc group that was selectively cleaved on resin using tetrakis(triphenylphosphine) palladium (0) (0.24 equiv) and phenylsilane (24 equiv) in DCM under inert conditions.[126] Approximately 1.5 mL of dry DCM was added to the resin under argon and allow to stir. The palladium catalyst was added to 3 mL of dry DCM and stirred until dissolved. The phenylsilane was added to the resin followed by the DCM/palladium mixture. The reaction stirred for 30 minutes. The resin was washed three times with ~2-3 mL of THF, DMF, 5% diisopropylethylamine in DMC and 0.02 M sodium dithiocarbamate in DMF.[130] Attachment of 5-carboxyfluorescein (2.5 equiv) was carried out overnight on resin using HOBt (2.5 equiv) and DIC (2.5 equiv) in NMP.[127] The 5-FAM coupling was protected from light and repeated once. After the second coupling, the resin was washed with 20% piperidine in DMF to remove excess 5-FAM. Washes were repeated until MALDI-MS of microcleavages showed only singly labeled peptide.

Peptides were cleaved from resin by treatment with 92.5% trifluoroacetic acidic, 3% water, 3% ethanedithiol and 1.5% triisopropylsilane solution for 4 to 6 hours. After the peptide was cleaved from resin, it was precipitated using ~40 mL cold diethyl ether. The precipitate was centrifuged and the ether decanted. The peptide pellet was suspended in a 0.1% TFA in water and 0.1% TFA in acetonitrile mixture for purification. Peptides were purified by HPLC on a C_{18} preparative column using 0.1% TFA in water and 0.1% TFA on acetonitrile gradients. HPLC fractions containing the peptide product were combined, frozen and lyophilized. Peptide identity was confirmed by mass spectrometry using a Voyager DE Pro MALDI-TOF instrument. All peptides were >95% pure by analytical HPLC on a C_{18} column.

Table 8. MALDI masses of peptides and subunits synthesized in 4

Peptide	$[M + H]+$ m/z (avg.)	
	Calculated	Observed
23TR	4781.5	4781.7
23TM	4496.9	4495.1
24	4396.9	4394.8
25	4413.0	4411.0
26	8964.2	8960.8
27	8994.2	8990.7

4.6.2 Synthesis of 5-FAM Labeled Subunit (26 and 27)

A solution of **25** was prepared in deionized water and the concentration determined by UV absorption at 490 nm ($\varepsilon = 68,000$ $M^{-1}cm^{-1}$) in 6.0 M guanadinum chloride, pH 7.0. A fresh solution of 1.5 mM linker was prepared DMF. A solution was prepared consisting of 150 μM peptide in 25 mM phosphate buffer, pH 7.0 with 10% acetonitrile. Argon was bubbled through the solution for 10 minutes followed by the addition of one aliquot of the linker solution. The reaction was placed in water bath at approximately 50°C and allowed to stir for 2 hours with four additional aliquots

of linker being added in 15 minute intervals for the first hour. The reaction was quenched after 2 hours with a 25% acetonitrile/0.01% TFA solution and then purified using molecular weight centricon spin filters (3000 MWCO) and semi-preparative HPLC. Centricon filters were used to remove the DMF linker co-solvent and concentrate the subunit. Equal amounts of reaction solution and water were added to the spin filter, and the mixture was spun at 6000 RPM for 30 minutes. Additional spins were performed until the reaction mixture had been washed with ~3X volumes of water. Identity and purity of the final product was confirmed by analytical HPLC and MALDI-MS.

4.6.3 Synthesis of Tetramethylrhodamine Labeled Peptide (23TM)

Peptide **3** was synthesized and purified previously described (section 2.9.1). A 100 µM (2 equiv) of peptide **3** was prepared in 100 mM HEPES buffer pH 6.7. Argon was bubble through the solution for ~ 5 minutes. A fresh solution of 10 mM iodoacetamide tetramethylrhodamine was prepared in DMF. 1 equivalent of dye was added to the peptide in buffer and allowed to stir at room temperature for 2 hours. After 2 hours, the reaction was quenched with a mixture of 0.1% TFA in acetonitrile and water. The reaction mixture was then purified by HPLC on a C_{18} preparative column using 0.1% TFA in water and 0.1% TFA on acetonitrile gradients. HPLC fractions containing the peptide product were combined, frozen and lyophilized. Peptide identity was confirmed by mass spectrometry using a Voyager DE Pro MALDI-TOF instrument. All peptides were >95% pure by analytical HPLC on a C_{18} column.

4.6.4 Synthesis of Texas Red Labeled Peptide (23TR)

Peptide **3** was synthesized and purified previously described (2.9.1). To prepare the Texas Red labeled peptide, a 150 µM solution of peptide **3** was made in 100 mM Tris buffer pH 8.0. TCEP was added (final concentration 1 mM) and this mixture was allowed to sit at room temperature for 10 minutes. A fresh solution of Texas Red maleimide was prepared in DMSO and protected from light. From this point on, reaction vessels were protected from light as best possible. Approximately 2.5 equivalents of the Texas Red were added to the peptide in buffer. The reaction was heated to 35˚C. After 1 hour the reaction was quenched by dilution with 5x-10x mixture of water and acetronitrile with 0.1% TFA. Amicon Ultra centrifugation spin filters (MWCO 3000) were used to remove excess dye and to reduce the final volume to approximately half of the original reaction volume. Peptides were then purified by HPLC on a C_{18} preparative column using 0.1% TFA in water and 0.1% TFA in acetonitrile gradients. HPLC fractions containing the peptide product were combined, frozen and lyophilized. Peptide identity was confirmed by mass spectrometry using a Voyager DE Pro MALDI-TOF instrument. All peptides were >95% pure by analytical HPLC on a C_{18} column.

4.6.5 Steady State Fluorescence

Steady state fluorescence measurements were carried out on a Horiba FluoroMax 3 Fluorescence Spectrometer with slit widths of either 3 x 3 nm or 1 x 1 nm. Stock solutions were made of peptides **23** and **24** and subunits **26** and **27** in deionized water. Their concentrations were measured by diluting 10 µL of the stock solution in 490 µL of 6.0 M guanadinum chloride, pH 7.0. The fluorescein containing peptides were measured at 490 nm with $\varepsilon = 68,000$ $M^{-1}cm^{-1}$ for **24** and

136,000 $M^{-1}cm^{-1}$ for **26** and **27**. Peptide **23TR** was measured at 590 nm with $\varepsilon = 85,000$ $M^{-1}cm^{-1}$ and peptide **23TM** was measured at 556 nm with $\varepsilon = 88,000$ $M^{-1}cm^{-1}$. Samples were prepared in 100 mM phosphate buffer pH 7.2. Optical densities of the samples were taken immediately after preparation in a 2 mm path length low volume quartz cell. Subunit containing samples were allowed to equilibrate for 4 hours before fluorescence measurements were taken. All fluorescence measurements were taken in a 10 cm x 2 mm low volume quartz cell with the 2 mm path length placed in the excitation direction. Higher temperatures were achieved with a cell holder heated by a water jacket. The water bath was heated to 60°C and the temperature of the cell holder allowed to stabilize for 10 minutes before any readings were taken. Once the cell holder was heated, samples were incubated for 3 minutes in the cell and holder before a measurement was taken.

APPENDIX A

NMR SPECTRA

5b

3.084
3.091
3.096

3.615

8.290

4.00
3.88
1.90

ppm

1.22

38.95

168.28

5b

190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 ppm
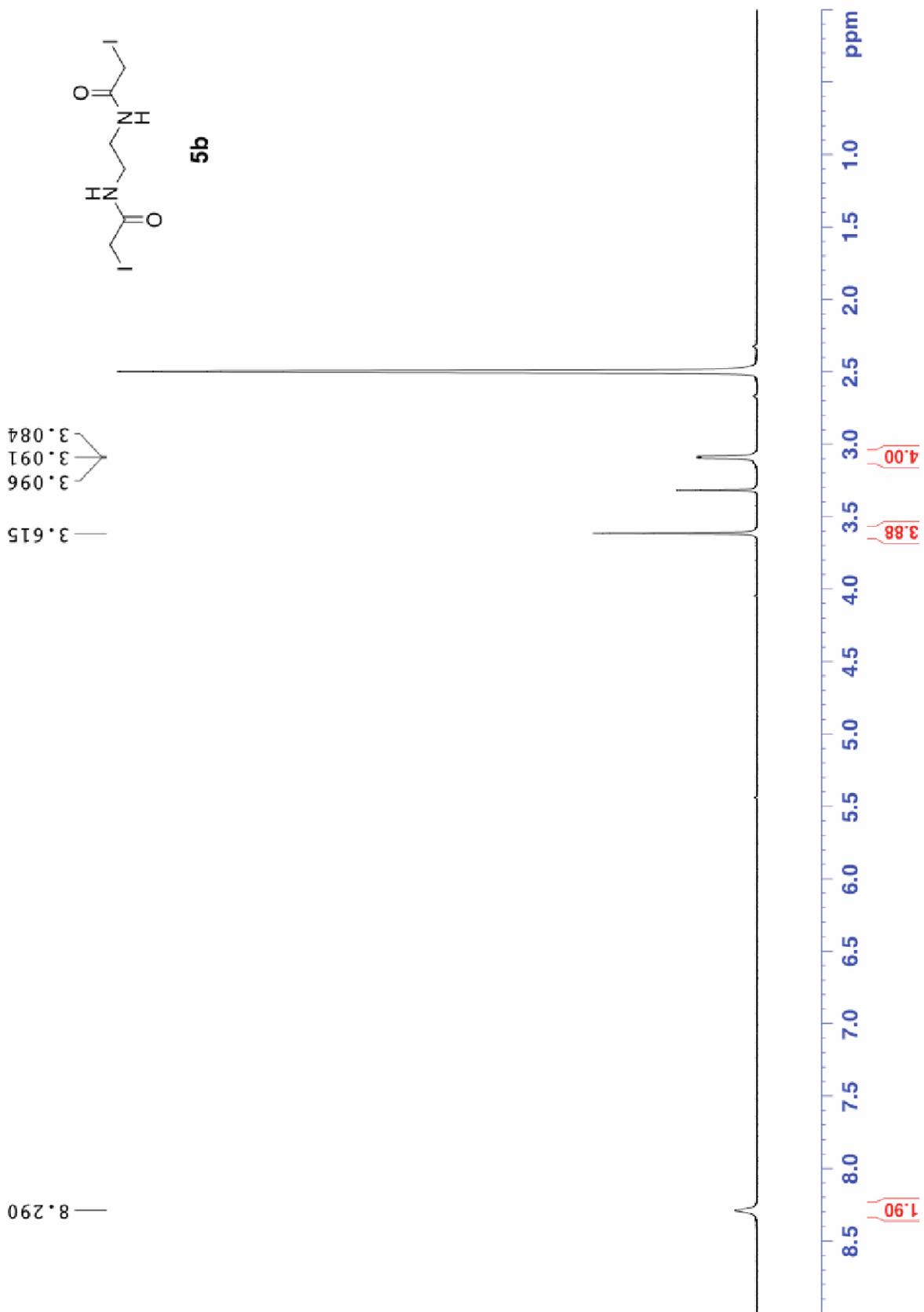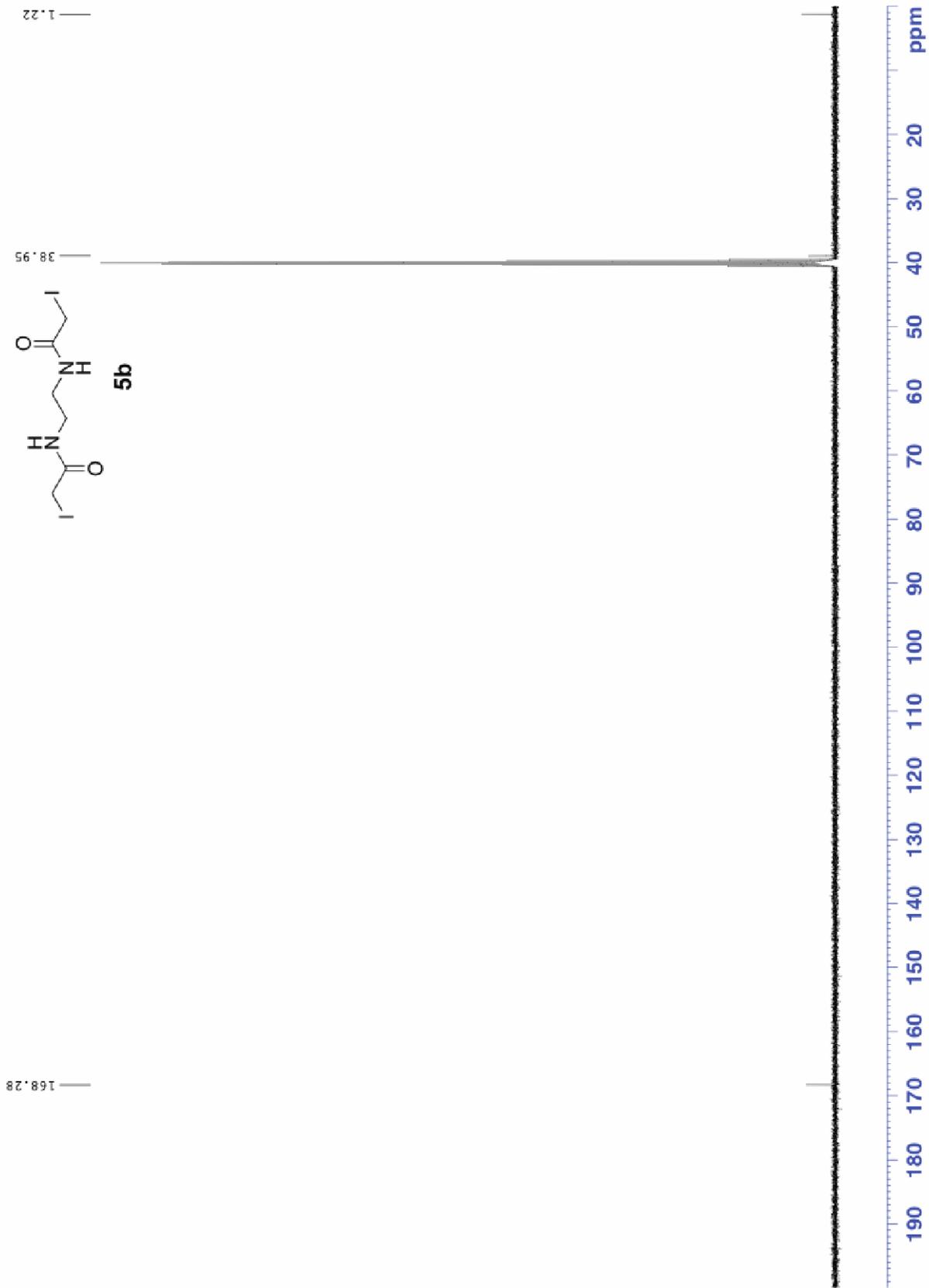
BIBLIOGRAPHY

1. Brunsveld, L.; Folmer, B.; Meijer, E.; Sijbesma, R., Supramolecular polymers. *Chemical Reviews* **2001,** *101* (12), 4071-4098.
2. Cordier, P.; Tournilhac, F.; Soulié-Ziakovic, C.; Leibler, L., Self-healing and thermoreversible rubber from supramolecular assembly. *Nature* **2008,** *451* (7181), 977-980.
3. Yan, X.; Xu, D.; Chi, X.; Chen, J.; Dong, S.; Ding, X.; Yu, Y.; Huang, F., A Multiresponsive, Shape- Persistent, and Elastic Supramolecular Polymer Network Gel Constructed by Orthogonal Self- Assembly. *Advanced Materials* **2012,** *24* (3), 362-369.
4. 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* (7343), 334-337.
5. Aida, T.; Meijer, E.; Stupp, S., Functional supramolecular polymers. *Science* **2012,** *335* (6070), 813-817.
6. Hartgerink, J. D.; Beniash, E.; Stupp, S. I., Self-assembly and mineralization of peptide-amphiphile nanofibers. *Science* **2001,** *294* (5547), 1684-1688.
7. Rajagopal, K.; Schneider, J. P., Self-assembling peptides and proteins for nanotechnological applications. *Current Opinion in Structural Biology* **2004,** *14* (4), 480-486.
8. Rajangam, K.; Behanna, H. A.; Hui, M. J.; Han, X.; Hulvat, J. F.; Lomasney, J. W.; Stupp, S. I., Heparin binding nanostructures to promote growth of blood vessels. *Nano Letters* **2006,** *6* (9), 2086-2090.
9. Hartgerink, J. D.; Beniash, E.; Stupp, S. I., Peptide-amphiphile nanofibers: a versatile scaffold for the preparation of self-assembling materials. *Proceedings of the National Academy of Sciences* **2002,** *99* (8), 5133-5138.
10. De Greef, T. F.; Smulders, M. M.; Wolffs, M.; Schenning, A. P.; Sijbesma, R. P.; Meijer, E., Supramolecular polymerization. *Chemical Reviews* **2009,** *109* (11), 5687-5754.
11. Hirschberg, J. K.; Brunsveld, L.; Ramzi, A.; Vekemans, J. A.; Sijbesma, R. P.; Meijer, E., Helical self-assembled polymers from cooperative stacking of hydrogen-bonded pairs. *Nature* **2000,** *407* (6801), 167-170.
12. Kastler, M.; Pisula, W.; Wasserfallen, D.; Pakula, T.; Müllen, K., Influence of Alkyl Substituents on the Solution-and Surface-Organization of Hexa-p eri-hexabenzocoronenes. *Journal of the American Chemical Society* **2005,** *127* (12), 4286-4296.
13. Velten, U.; Rehahn, M., First synthesis of soluble, well defined coordination polymers from kinetically unstable copper (I) complexes. *Chemical Communications* **1996,** (23), 2639-2640.
14. Berl, V.; Schmutz, M.; Krische, M. J.; Khoury, R. G.; Lehn, J. M., Supramolecular Polymers Generated from Heterocomplementary Monomers Linked through Multiple Hydrogen- Bonding Arrays—Formation, Characterization, and Properties. *Chemistry-A European Journal* **2002,** *8* (5), 1227-1244.

15.   Ashton, P. R.; Baxter, I.; Cantrill, S. J.; Fyfe, M. C. T.; Glink, P. T.; Stoddart, J. F.; White, A. J. P.; Williams, D. J., Supramolecular Daisy Chains. *Angewandte Chemie International Edition* **1998,** *37* (9), 1294-1297.

16.   Wang, F.; Han, C.; He, C.; Zhou, Q.; Zhang, J.; Wang, C.; Li, N.; Huang, F., Self-Sorting Organization of Two Heteroditopic Monomers to Supramolecular Alternating Copolymers. *Journal of the American Chemical Society* **2008,** *130* (34), 11254-11255.

17.   Huang, F.; Scherman, O. A., Supramolecular polymers. *Chemical Society Reviews* **2012,** *41* (18), 5879-5880.

18.   Martin, R. B., Comparisons of indefinite self-association models. *Chemical Reviews* **1996,** *96* (8), 3043-3064.

19.   Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J.; Hirschberg, J. K.; Lange, R. F.; Lowe, J. K.; Meijer, E., Reversible polymers formed from self-complementary monomers using quadruple hydrogen bonding. *Science* **1997,** *278* (5343), 1601-1604.

20.   Wietor, J.-L.; van Beek, D. J. M.; Peters, G. W.; Mendes, E.; Sijbesma, R. P., Effects of Branching and Crystallization on Rheology of Polycaprolactone Supramolecular Polymers with Ureidopyrimidinone End Groups. *Macromolecules* **2011,** *44* (5), 1211-1219.

21.   Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E., Strong dimerization of ureidopyrimidones via quadruple hydrogen bonding. *Journal of the American Chemical Society* **1998,** *120* (27), 6761-6769.

22.   Goldstein, R.; Stryer, L., Cooperative polymerization reactions. Analytical approximations, numerical examples, and experimental strategy. *Biophysical Journal* **1986,** *50* (4), 583.

23.   Smith, A. M.; Banwell, E. F.; Edwards, W. R.; Pandya, M. J.; Woolfson, D. N., Engineering Increased Stability into Self- Assembled Protein Fibers. *Advanced Functional Materials* **2006,** *16* (8), 1022-1030.

24.   Dong, H.; Paramonov, S. E.; Hartgerink, J. D., Self-assembly of α-helical coiled coil nanofibers. *Journal of the American Chemical Society* **2008,** *130* (41), 13691-13695.

25.   Hoshino, T.; Miyauchi, M.; Kawaguchi, Y.; Yamaguchi, H.; Harada, A., Daisy chain necklace: tri [2] rotaxane containing cyclodextrins. *Journal of the American Chemical Society* **2000,** *122* (40), 9876-9877.

26.   Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N., Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **1993,** *366* (6453), 324-327.

27.   Ringler, P.; Schulz, G. E., Self-assembly of proteins into designed networks. *Science* **2003,** *302* (5642), 106-109.

28.   Bandy, T. J.; Brewer, A.; Burns, J. R.; Marth, G.; Nguyen, T.; Stulz, E., DNA as supramolecular scaffold for functional molecules: progress in DNA nanotechnology. *Chemical Society Reviews* **2011,** *40* (1), 138-148.

29.   McLaughlin, C. K.; Hamblin, G. D.; Sleiman, H. F., Supramolecular DNA assembly. *Chemical Society Reviews* **2011,** *40* (12), 5647-5656.

30.   Bromley, E. H.; Channon, K.; Moutevelis, E.; Woolfson, D. N., Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. *ACS chemical biology* **2008,** *3* (1), 38-50.

31.   Lee, L. A.; Niu, Z.; Wang, Q., Viruses and virus-like protein assemblies—Chemically programmable nanoscale building blocks. *Nano Research* **2009,** *2* (5), 349-364.

32. Padilla, J. E.; Colovos, C.; Yeates, T. O., Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proceedings of the National Academy of Sciences* **2001,** *98* (5), 2217-2221.

33. Salgado, E. N.; Radford, R. J.; Tezcan, F. A., Metal-directed protein self-assembly. *Accounts of Chemical Research* **2010,** *43* (5), 661-672.

34. Miller, R. A.; Presley, A. D.; Francis, M. B., Self-assembling light-harvesting systems from synthetically modified tobacco mosaic virus coat proteins. *Journal of the American Chemical Society* **2007,** *129* (11), 3104-3109.

35. Lai, Y.-T.; Reading, E.; Hura, G. L.; Tsai, K.-L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O., Structure of a designed protein cage that self-assembles into a highly porous cube. *Nature Chemistry* **2014,** *6* (12), 1065-1071.

36. Kobayashi, N.; Yanase, K.; Sato, T.; Unzai, S.; Hecht, M. H.; Arai, R., Self-Assembling Nano-Architectures Created from a Protein Nano-Building Block Using an Intermolecularly Folded Dimeric de Novo Protein. *Journal of the American Chemical Society* **2015,** *137* (35), 11285-11293.

37. King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D., Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science* **2012,** *336* (6085), 1171-1174.

38. King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D., Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **2014,** *510* (7503), 103-108.

39. Brodin, J. D.; Ambroggio, X.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two-and three-dimensional crystalline protein arrays. *Nature Chemistry* **2012,** *4* (5), 375-382.

40. O'Leary, L. E.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D., Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nature Chemistry* **2011,** *3* (10), 821-828.

41. Schneider, J. P.; Pochan, D. J.; Ozbas, B.; Rajagopal, K.; Pakstis, L.; Kretsinger, J., Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *Journal of the American Chemical Society* **2002,** *124* (50), 15030-15037.

42. Zhang, R.; Tang, M.; Bowyer, A.; Eisenthal, R.; Hubble, J., A novel pH-and ionic-strength-sensitive carboxy methyl dextran hydrogel. *Biomaterials* **2005,** *26* (22), 4677-4683.

43. Haines, L. A.; Rajagopal, K.; Ozbas, B.; Salick, D. A.; Pochan, D. J.; Schneider, J. P., Light-activated hydrogel formation via the triggered folding and self-assembly of a designed peptide. *Journal of the American Chemical Society* **2005,** *127* (48), 17025-17029.

44. Richter, A.; Paschew, G.; Klatt, S.; Lienig, J.; Arndt, K.-F.; Adler, H.-J. P., Review on hydrogel-based pH sensors and microsensors. *Sensors* **2008,** *8* (1), 561-581.

45. Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A., Reversible hydrogels from self-assembling artificial proteins. *Science* **1998,** *281* (5375), 389-392.

46. Rackham, O. J.; Madera, M.; Armstrong, C. T.; Vincent, T. L.; Woolfson, D. N.; Gough, J., The evolution and structure prediction of coiled coils across all genomes. *Journal of Molecular Biology* **2010,** *403* (3), 480-493.

47. Lupas, A. N.; Gruber, M., The structure of α-helical coiled coils. *Advances in Protein Chemistry* **2005,** *70*, 37-38.

48. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T., A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993,** *262* (5138), 1401-1407.

49. Woolfson, D. N.; Alber, T., Predicting oligomerization states of coiled coils. *Protein Science* **1995,** *4* (8), 1596-1607.

50. Liu, J.; Zheng, Q.; Deng, Y.; Cheng, C.-S.; Kallenbach, N. R.; Lu, M., A seven-helix coiled coil. *Proceedings of the National Academy of Sciences* **2006,** *103* (42), 15457-15462.

51. Harbury, P. B.; Kim, P. S.; Alber, T., Crystal structure of an isoleucine-zipper trimer. *Nature* **1994,** *371*, 80-83.

52. Tsang, B. P.; Bretscher, H. S.; Kokona, B.; Manning, R. S.; Fairman, R., Thermodynamic Analysis of Self-Assembly in Coiled-Coil Biomaterials. *Biochemistry* **2011,** *50* (40), 8548-8558.

53. Zhou, M.; Bentley, D.; Ghosh, I., Helical supramolecules and fibers utilizing leucine zipper-displaying dendrimers. *Journal of the American Chemical Society* **2004,** *126* (3), 734-735.

54. O'Shea, E. K.; Lumb, K. J.; Kim, P. S., Peptide 'Velcro': design of a heterodimeric coiled coil. *Current Biology* **1993,** *3* (10), 658-667.

55. Boyle, A. L.; Bromley, E. H.; Bartlett, G. J.; Sessions, R. B.; Sharp, T. H.; Williams, C. L.; Curmi, P. M.; Forde, N. R.; Linke, H.; Woolfson, D. N., Squaring the circle in peptide assembly: from fibers to discrete nanostructures by de novo design. *Journal of the American Chemical Society* **2012,** *134* (37), 15457-15467.

56. Fletcher, J. M.; Boyle, A. L.; Bruning, M.; Bartlett, G. J.; Vincent, T. L.; Zaccai, N. R.; Armstrong, C. T.; Bromley, E. H.; Booth, P. J.; Brady, R. L., A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synthetic Biology* **2012,** *1* (6), 240-250.

57. Boyle, A. L.; Woolfson, D. N., Rational Design of Peptide‐Based Biosupramolecular Systems. *Supramolecular Chemistry: From Molecules to Nanomaterials* **2012**.

58. Bromley, E. H.; Kuwada, N. J.; Zuckermann, M. J.; Donadini, R.; Samii, L.; Blab, G. A.; Gemmen, G. J.; Lopez, B. J.; Curmi, P. M.; Forde, N. R., The tumbleweed: towards a synthetic protein motor. *HFSP Journal* **2009,** *3* (3), 204-212.

59. Stevens, M. M.; Flynn, N. T.; Wang, C.; Tirrell, D. A.; Langer, R., Coiled‐Coil Peptide‐Based Assembly of Gold Nanoparticles. *Advanced Materials* **2004,** *16* (11), 915-918.

60. Fletcher, J. M.; Harniman, R. L.; Barnes, F. R.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N., Self-assembling cages from coiled-coil peptide modules. *Science* **2013,** *340* (6132), 595-599.

61. O'Shea, E. K.; Rutkowski, R.; Kim, P. S., Evidence that the leucine zipper is a coiled coil. *Science* **1989,** *243* (4890), 538-542.

62. Crick, F. H., The packing of-helices: simple coiled-coils. *Acta Crystallographica* **1953,** *6* (8-9), 689-697.

63. Mason, J. M.; Arndt, K. M., Coiled coil domains: stability, specificity, and biological implications. *ChemBioChem* **2004,** *5* (2), 170-176.

64. Kojima, S.; Kuriki, Y.; YOsHIDA, T.; YAZAKI, K.; MIURA, K.-i., Fibril Formation by an Amphipathic α-Helix-Forming Polypeptide Produced by Gene Engineering. *Proceedings of the Japan Academy. Ser. B: Physical and Biological Sciences* **1997,** *73* (1), 7-11.

65.   Xu, C.; Liu, R.; Mehta, A. K.; Guerrero-Ferreira, R. C.; Wright, E. R.; Dunin-Horkawicz, S.; Morris, K.; Serpell, L. C.; Zuo, X.; Wall, J. S., Rational design of helical nanotubes from self-assembly of coiled-coil lock washers. *Journal of the American Chemical Society* **2013,** *135* (41), 15565-15578.

66.   Burgess, N. C.; Sharp, T. H.; Thomas, F.; Wood, C. W.; Thomson, A. R.; Zaccai, N. R.; Brady, R. L.; Serpell, L. C.; Woolfson, D. N., Modular Design of Self-Assembling Peptide-Based Nanotubes. *Journal of the American Chemical Society* **2015,** *137* (33), 10554-10562.

67.   Bella, A.; Ray, S.; Shaw, M.; Ryadnov, M. G., Arbitrary Self- Assembly of Peptide Extracellular Microscopic Matrices. *Angewandte Chemie International Edition* **2012,** *51* (2), 428-431.

68.   Raman, S.; Machaidze, G.; Lustig, A.; Aebi, U.; Burkhard, P., Structure-based design of peptides that self-assemble into regular polyhedral nanoparticles. *Nanomedicine: Nanotechnology, Biology and Medicine* **2006,** *2* (2), 95-102.

69.   Langford, S. J.; Latter, M. J.; Wilman, B. E.; Bhosale, S. V., Biologically derived supramolecular materials. *Supramolecular Chemistry: From Molecules to Nanomaterials* **2012**.

70.   Bath, J.; Turberfield, A. J., DNA nanomachines. *Nature Nanotechnology* **2007,** *2* (5), 275-284.

71.   Rothemund, P. W., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440* (7082), 297-302.

72.   Thomas, F.; Boyle, A. L.; Burton, A. J.; Woolfson, D. N., A Set of de Novo Designed Parallel Heterodimeric Coiled Coils with Quantified Dissociation Constants in the Micromolar to Sub-nanomolar Regime. *Journal of the American Chemical Society* **2013,** *135* (13), 5161-5166.

73.   Kim, Y. H.; Stites, W. E., Effects of Excluded Volume upon Protein Stability in Covalently Cross-Linked Proteins with Variable Linker Lengths†. *Biochemistry* **2008,** *47* (33), 8804-8814.

74.   Wendt, H.; Baici, A.; Bosshard, H. R., Mechanism of assembly of a leucine zipper domain. *Journal of the American Chemical Society* **1994,** *116* (15), 6973-6974.

75.   Tavenor, N. A.; Silva, K. I.; Saxena, S.; Horne, W. S., Origins of Structural Flexibility in Protein-Based Supramolecular Polymers Revealed by DEER Spectroscopy. *The Journal of Physical Chemistry B* **2014,** *118* (33), 9881-9889.

76.   Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S., The role of position a in determining the stability and oligomerization state of [alpha]-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Science* **1999,** *8* (11), 2312-2329.

77.   Acharya, A.; Ruvinov, S. B.; Gal, J.; Moll, J. R.; Vinson, C., A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: amino acids I, V, L, N, A, and K. *Biochemistry* **2002,** *41* (48), 14122-14131.

78.   Acharya, A.; Rishi, V.; Vinson, C., Stability of 100 homo and heterotypic coiled-coil a-a'pairs for ten amino acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry* **2006,** *45* (38), 11324-11332.

79.   Moitra, J.; Szilák, L.; Krylov, D.; Vinson, C., Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **1997,** *36* (41), 12567-12573.

80. Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C. T.; Hodges, R. S., Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d". *Journal of Molecular Biology* **2000,** *300* (2), 377-402.

81. Oshaben, K. M.; Salari, R.; McCaslin, D. R.; Chong, L. T.; Horne, W. S., The native GCN4 leucine-zipper domain does not uniquely specify a dimeric oligomerization state. *Biochemistry* **2012,** *51* (47), 9581-9591.

82. Price, J. L.; Horne, W. S.; Gellman, S. H., Structural consequences of β-amino acid preorganization in a self-assembling α/β-peptide: fundamental studies of foldameric helix bundles. *Journal of the American Chemical Society* **2010,** *132* (35), 12378-12387.

83. Mishra, A.; Gestin, J.; Benoist, E.; Faivre-Chauvet, A.; Chatal, J., Simplified synthesis of the bifunctional chelating agent 2-(4-aminobenzyl)-1, 4, 7, 10-tetraazacyclododecane-N, N", N'''-tetraacetic acid. *New Journal of Chemistry* **1996,** *20* (5), 585-588.

84. Ozawa, H., Bridging Reagent for Protein II. The Reaction of N, N'-Polymethylene-bis (iodoacetamide) with Cysteine and Rabbit Muscle Aldolase. *Journal of Biochemistry* **1967,** *62* (5), 531-536.

85. Staples, J. K.; Oshaben, K. M.; Horne, W. S., A modular synthetic platform for the construction of protein-based supramolecular polymers via coiled-coil self-assembly. *Chemical Science* **2012,** *3* (12), 3387-3392.

86. Gill, S. C.; Von Hippel, P. H., Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* **1989,** *182* (2), 319-326.

87. Shortle, D.; Meeker, A. K.; Freire, E., Stability mutants of staphylococcal nuclease: large compensating enthalpy-entropy changes for the reversible denaturation reaction. *Biochemistry* **1988,** *27* (13), 4761-4768.

88. Fersht, A., *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. Macmillan: 1999.

89. Collaborative, C. P., The CCP4 suite: programs for protein crystallography. *Acta Crystallographica. Section D, Biological Crystallography* **1994,** *50* (Pt 5), 760.

90. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **2010,** *66* (2), 213-221.

91. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T., X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991,** *254* (5031), 539-544.

92. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography* **1997,** *53* (3), 240-255.

93. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **2004,** *60* (12), 2126-2132.

94. Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A., Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **2008,** *3* (7), 1171-1179.

95. Betz, S.; Fairman, R.; O'Neil, K.; Lear, J.; Degrado, W., Design of two-stranded and three-stranded coiled-coil peptides. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **1995,** *348* (1323), 81-88.

96. Akey, D. L.; Malashkevich, V. N.; Kim, P. S., Buried polar residues in coiled-coil interfaces. *Biochemistry* **2001,** *40* (21), 6352-6360.

97. Knappenberger, J. A.; Smith, J. E.; Thorpe, S. H.; Zitzewitz, J. A.; Matthews, C. R., A buried polar residue in the hydrophobic interface of the coiled-coil peptide, GCN4-p1, plays a thermodynamic, not a kinetic role in folding. *Journal of Molecular Biology* **2002,** *321* (1), 1-6.

98. Gonzalez, L.; Woolfson, D. N.; Alber, T., Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nature Structural & Molecular Biology* **1996,** *3* (12), 1011-1018.

99. Gonzalez, L.; Brown, R. A.; Richardson, D.; Alber, T., Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. *Nature Structural & Molecular Biology* **1996,** *3* (12), 1002-1010.

100. Testa, O. D.; Moutevelis, E.; Woolfson, D. N., CC+: a relational database of coiled-coil structures. *Nucleic Acids Research* **2009,** *37* (suppl 1), D315-D322.

101. Dragan, A. I.; Privalov, P. L., Unfolding of a leucine zipper is not a simple two-state transition. *Journal of Molecular Biology* **2002,** *321* (5), 891-908.

102. Prakash, V.; Timasheff, S. N., The calculation of partial specific volumes of proteins in 8 M urea solution. *Analytical Biochemistry* **1981,** *117* (2), 330-335.

103. Zhou, H.-X.; Dill, K. A., Stabilization of proteins in confined spaces. *Biochemistry* **2001,** *40* (38), 11289-11293.

104. Zhou, H.-X.; Rivas, G.; Minton, A. P., Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics* **2008,** *37*, 375.

105. Drees, B. L.; Grotkopp, E. K.; Nelson, H. C., The GCN4 leucine zipper can functionally substitute for the heat shock transcription factor's trimerization domain. *Journal of Molecular Biology* **1997,** *273* (1), 61-74.

106. Ridgeway, T. P., SL, Computer-aided interpretation of analytical sedimentation data for proteins. *Analytical Ultracentrifugation in Biochemistry and Polymer Science* **1992**, 90-125.

107. Perkins, S. J., Protein volumes and hydration effects. *European Journal of Biochemistry* **1986,** *157* (1), 169-180.

108. Durchschlag, H.; Zipper, P., Calculation of the partial volume of organic compounds and polymers. In *Ultracentrifugation*, Springer: 1994; pp 20-39.

109. Hukushima, K.; Nemoto, K., Exchange Monte Carlo method and application to spin glass simulations. *Journal of the Physical Society of Japan* **1996,** *65* (6), 1604-1608.

110. Hansmann, U. H., Parallel tempering algorithm for conformational studies of biological molecules. *Chemical Physics Letters* **1997,** *281* (1), 140-150.

111. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters* **1999,** *314* (1), 141-151.

112. García, A. E.; Sanbonmatsu, K. Y., α-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proceedings of the National Academy of Sciences* **2002,** *99* (5), 2782-2787.

113. Shirts, M. R.; Chodera, J. D., Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics* **2008,** *129* (12), 124105.

114. Case, D. A.; Darden, T.; Cheatham, T.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R.; Zhang, W.; Merz, K. *Amber 11*; University of California: 2010.

115. Schrödinger, L., The PyMOL Molecular Graphics System, V. 1.3 r1. 2009.
116. Strelkov, S. V.; Burkhard, P., Analysis of α-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. *Journal of Structural Biology* **2002,** *137* (1), 54-64.
117. Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J., CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research* **2006,** *34* (suppl 2), W116-W118.
118. Oshaben, K. M.; Horne, W. S., Tuning Assembly Size in Peptide-Based Supramolecular Polymers by Modulation of Subunit Association Affinity. *Biomacromolecules* **2014,** *15* (4), 1436-1442.
119. Lakowicz, J. R., *Principles of fluorescence spectroscopy*. Springer Science & Business Media: 2013.
120. Haugland, R. P., *The handbook: a guide to fluorescent probes and labeling technologies*. Molecular probes: 2005.
121. Fujii, N.; Otaka, A.; Watanabe, T.; Okamachi, A.; Tamamura, H.; Yajima, H.; Inagaki, Y.; Nomizu, M.; Asano, K., Silver trifluoromethanesulphonate as an S-deprotecting reagent for the synthesis of cystine peptides. *Journal of the Chemical Society, Chemical Communications* **1989,** (5), 283-284.
122. Schroll, A. L.; Hondal, R. J.; Flemer, S., 2, 2′‐ Dithiobis (5‐ nitropyridine)(DTNP) as an effective and gentle deprotectant for common cysteine protecting groups. *Journal of Peptide Science* **2012,** *18* (1), 1-9.
123. Selwyn, J. E.; Steinfeld, J. I., Aggregation of equilibriums of xanthene dyes. *The Journal of Physical Chemistry* **1972,** *76* (5), 762-774.
124. Ajtai, K.; Ilich, P. J.; Ringler, A.; Sedarous, S. S.; Toft, D. J.; Burghardt, T. P., Stereospecific reaction of muscle fiber proteins with the 5'or 6'isomer of (iodoacetamido) tetramethylrhodamine. *Biochemistry* **1992,** *31* (49), 12431-12440.
125. Hamman, B. D.; Oleinikov, A. V.; Jokhadze, G. G.; Bochkariov, D. E.; Traut, R. R.; Jameson, D. M., Tetramethylrhodamine dimer formation as a spectroscopic probe of the conformation of Escherichia coli ribosomal protein L7/L12 dimers. *Journal of Biological Chemistry* **1996,** *271* (13), 7568-7573.
126. Grieco, P.; Gitu, P.; Hruby, V., Preparation of 'side‐ chain‐ to‐ side‐ chain'cyclic peptides by Allyl and Alloc strategy: potential for library synthesis. *The Journal of Peptide Research* **2001,** *57* (3), 250-256.
127. Fischer, R.; Mader, O.; Jung, G.; Brock, R., Extending the applicability of carboxyfluorescein in solid-phase synthesis. *Bioconjugate Chemistry* **2003,** *14* (3), 653-660.
128. Zhuang, X.; Ha, T.; Kim, H. D.; Centner, T.; Labeit, S.; Chu, S., Fluorescence quenching: A tool for single-molecule protein-folding study. *Proceedings of the National Academy of Sciences* **2000,** *97* (26), 14241-14244.
129. Bai, F.; Chang, C.; Webber, S., Photon-harvesting polymers: singlet energy transfer in anthracene-loaded alternating and random copolymers of 2-vinylnaphthalene and methacrylic acid. *Macromolecules* **1986,** *19* (10), 2484-2494.
130. Leelasvatanakij, L.; Aldrich, J., A solid‐ phase synthetic strategy for the preparation of peptide‐ based affinity labels: synthesis of dynorphin A analogs. *The Journal of Peptide Research* **2000,** *56* (2), 80-87.