Supramolecular Architectures and Mimics of Complex Natural Folds Derived from Rationally Designed α-Helical Protein Structures

by

Nathan Albert Tavenor

B.S. Chemistry, Butler University, 2011

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Nathan A. Tavenor

It was defended on

May 18, 2017

and approved by

Danith H. Ly, Professor, Department of Chemistry, Carnegie Mellon University

Jill E. Millstone, Associate Professor, Department of Chemistry

Sunil K. Saxena, Professor, Department of Chemistry

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

Supramolecular Architectures and Mimics of Complex Natural Folds Derived from Rationally Designed α-Helical Protein Structures

Nathan A. Tavenor, PhD

University of Pittsburgh, 2017

Protein-based supramolecular polymers (SMPs) are a class of biomaterials which draw inspiration from and expand upon the many examples of complex protein quaternary structures observed in nature: collagen, microtubules, viral capsids, etc. Designing synthetic supramolecular protein scaffolds both increases our understanding of natural superstructures and allows for the creation of novel materials. Similar to small-molecule SMPs, protein-based SMPs form due to self-assembly driven by intermolecular interactions between monomers, and monomer structure determines the properties of the overall material. Using protein-based monomers takes advantage of the self-assembly and highly specific molecular recognition properties encodable in polypeptide sequences to rationally design SMP architectures.

The central hypothesis underlying our work is that α-helical coiled coils, a well-studied protein quaternary folding motif, are well-suited to SMP design through the addition of synthetic linkers at solvent-exposed sites. Through small changes in the structures of the cross-links and/or peptide sequence, we have been able to control both the nanoscale organization and the macroscopic properties of the SMPs. Changes to the linker and hydrophobic core of the peptide can be used to control polymer rigidity, stability, and dimensionality. The gaps in knowledge that this thesis sought to fill on this project were 1) the relationship between the molecular structure of the cross-linked polypeptides and the macroscopic properties of the SMPs and 2) a means of creating materials exhibiting multi-dimensional net or framework topologies.

Separate from the above efforts on supramolecular architectures was work on improving backbone modification strategies for an α-helix in the context of a complex protein tertiary fold. Earlier work in our lab had successfully incorporated unnatural building blocks into every major secondary structure (β-sheet, α-helix, loops and β-turns) of a small protein with a tertiary fold. Although the tertiary fold of the native sequence was mimicked by the resulting artificial protein, the thermodynamic stability was greatly compromised. Most of this energetic penalty derived from the modifications present in the α-helix. The contribution within this thesis was direct comparison of several α-helical design strategies and establishment of the thermodynamic consequences of each.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

LIST OF EQUATIONS

PREFACE

There are a lot of people I would like to thank who taught and gave me support to make this whole thing possible. Graduate school has been quite the journey, and I don't think I truly knew what to expect when I started it. There have been times of truly exciting breakthroughs and others of great frustration. Thank you to everyone who has supported and taught me through it all. I have learned and grown a lot from all of you.

Seth, you have been a wonderful advisor. I'm grateful that you were always available when I got stuck or didn't know what to do with ambiguous data. I have gained a lot from your creativity, scientific rigor and helpful criticism. I appreciate being given the freedom to explore new characterization techniques and modeling endeavors even if they didn't always pan out. Thank you for your patience with me as I stuck around longer than I think either of us would have preferred.

I thank Drs. Danith Ly, Jill Millstone, and Sunil Saxena for serving on my PhD thesis committee. Dr. Saxena it has also been wonderful collaborating with your lab on two projects resulting in one paper and hopefully a second. I would also like to thank Dr. Millstone again for being my original research proposal mentor and Drs. Seth Childers and Haitou Liu for serving on my proposal committee. Dr. Millstone, I especially appreciate your enthusiasm and helpful critiques in developing my proposal and writing skills.

I would also like to thank the graduate students who came before me—Drs. George Lengyel, Conor Haney, Zach Reinert, and Kaylyn Oshaben—I learned so much about peptide synthesis and purification from all of you. Thank you all for teaching me and helping me work through research problems in my early years here. I always looked up to you all and have missed you. Conor and George—I will never forget both of your favorite phrases of encouragement "it's going to fail" and "unbridled pessimism"—it was fun to see you both again at ACS conferences. Zach, it was great to be able to collaborate with you briefly on the GB1 project and parse through the complexities of protein folding thermodynamics. Kaylyn, it was fun to at sit back-to-back desks with you for four years and share the joys and frustrations of the materials project together.

To the younger students in the lab—Halina, Kelly, Hajira, Haley, Chino, Jackie, Kim, and Shilpa—I appreciate having had you all as colleagues and will miss working with you. If I can leave you any wisdom it is this: don't take research failures as personal failures—research can be frustrating enough as is. Instead, always ask why something turned out unexpectedly—sometimes you may stumble upon a hidden treasure. Also, take time for yourselves and learn to separate yourselves from your work. Thanks also to my two undergraduates Carly Parris and Matt Murnin who were both enthusiastic and quick learners. Matt, I am especially thankful for all the hours you spent setting up crystal trays—after you left they led to a great breakthrough resulting in a *JACS* paper.

To my collaborators in the Saxena lab—Dr. Ishara Silva and Matt Lawless—thank you both for all of the EPR experiments you ran for me. I could not have done these projects without you. I always enjoyed our conversations together trying to parse through the data interpretation—even if it was confusing at times! You were both a pleasure to work with and I wish you success on all your future endeavors.

A special thanks is also needed for all of the support staff within the Chemistry Department and University. Thank you to all the office and maintenance workers who enable our labs to run smoothly. Thank you to the mass spec, NMR, X-Ray, supercomputing and materials characterization facilities for maintaining all of the shared equipment. Also, thank you to the machine and electronics shops for the help maintaining and repair our instrumentation when they breakdown—you guys have been lifesavers! Thanks also to the support staff at the Advanced Photon Source, Argonne National Laboratory for making remote collection of crystallography data possible.

Thank you to all my family, friends and mentors who have supported me throughout the years. I am sorry I cannot name you all, but I would not be the person I am today without you. I especially appreciate my parents for instilling in me a love of learning from an early age by teaching me to read, taking us to museums, and encouraging us to excel; but not faulting us when we sometimes failed. A special thanks to my undergraduate professors, especially Dr. Samide, who gave me my first foray into research and encouraged me to pursue graduate studies. I am also very grateful to the Christian community at the Upper Room Church; I praise God for all of your love, friendship, prayers, and conversations.

Last, but most important to me, I would like to thank my wife, Kristen, for all of her support. I know I can be difficult to deal with when I'm stressed from projects not working out. Thank you for joining me here in Pittsburgh on this often crazy adventure that has been our lives the last few years. You've been there to celebrate the successes with me and encouraged me when I want to give up. You're the love of my life and I truly appreciate having by my side through the ups and downs of grad school and look forward to our future together that results from all of this.

1.0 INTRODUCTION

1.1 SUPRAMOLECULAR POLYMERS

Supramolecular polymers (SMPs) are a class of materials that are generated through a series of defined non-covalent interactions among monomer units, in contrast to classical polymers which are held together by covalent bonds.[1-2] Any non-covalent force can be used to drive supramolecular polymerization, but those that have been most widely used are van der Waals (i.e. hydrophobic interactions in aqueous solvent), H-bonding, host-guest, metal chelation, and charge-charge interactions. The strength of each of these forces can be highly dependent on environment, allowing SMPs to be assembled and disassembled by external stimuli. The great potential application of SMPs as "smart" materials and their relative ease of recyclability arise from this responsiveness to environment.[2] Another appealing feature of supramolecular polymers is that they are self-healing, due to the ability to reform supramolecular interactions after mechanical deformation.[3] A key limitation of SMPs lies in the difficulty of developing materials that are both environmentally responsive and exhibit the mechanical stability of conventional polymers.

While supramolecular complexes between specific host-guest molecule pairs have long been studied[4]—both in biotic and abiotic contexts—the first non-biological SMP was not described until more recently.[2, 5] This SMP was formed by mixing together bifunctional diamidopyridines with uracil derivatives (Figure 1).

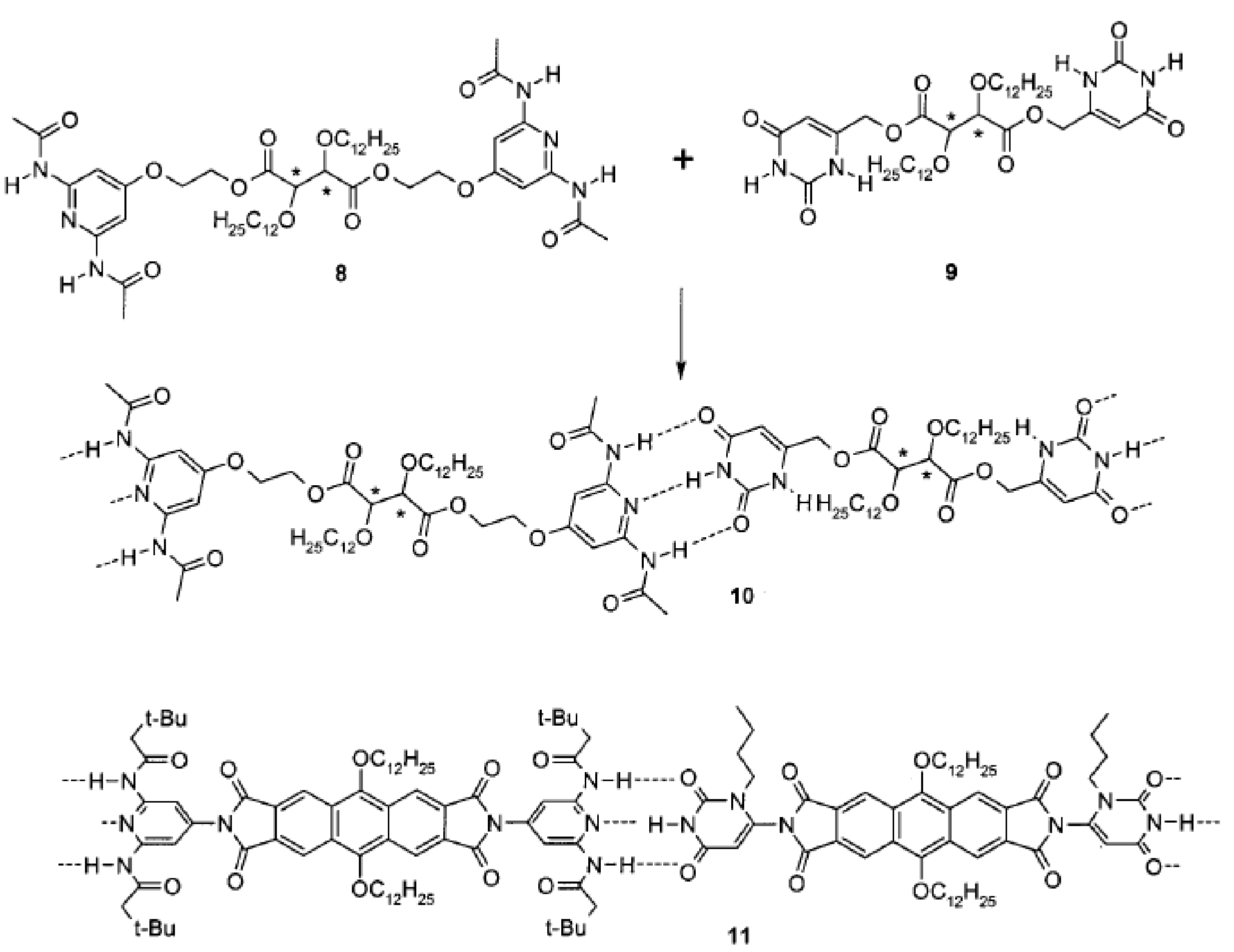


Figure 1. Supramolecular polymers based on uracil.

(Adapted with permission from reference [5]. Copyright © 1990 Verlag GmbH & Co. KGaA, Weinheim.)

Formation of the SMP occurs in a similar manner to nucleic acid complexation, driven by complementary H-bonds between acceptors and donors on each molecule. Polymer-like properties were observed upon mixing that neither monomer exhibits alone. Note the linker domains in each monomer joining the two ends of each supramolecular unit. The use of linkers to join together moieties with strong non-covalent interactions is a recurring motif in supramolecular polymers and play a prominent role in the materials discussed in this thesis (see Chapters 2 and 4).

This early work was followed by the seminal design of a self-assembling monomer based on the UPy motif (Figure 2).[6]

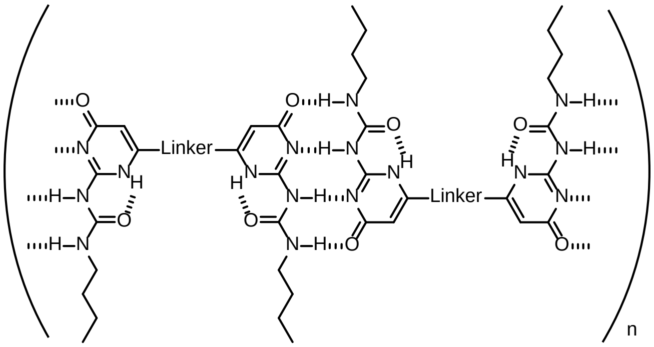


Figure 2. Hydrogen bonding between bifunctional UPy (2-ureido-pyrimidone) monomers with a variable linker.

The UPy motif utilizes multiple hydrogen bonds surrounded by a hydrophobic pocket formed by proximal alkyl chains to create a very strong affinity between monomers. Cross-linked UPy moieties undergo supramolecular polymerization to form nanofibers. Recent work has produced a supramolecular polymer blend of the UPy motif alongside metal coordination resulting in a system with two disassembly phase transitions (one for each supramolecular interaction).[7]

Another prominent H-bonding motif that has found use in supramolecular polymerization is based around a benzenetricarboxamide (BTA) core (Figure 3).[8]

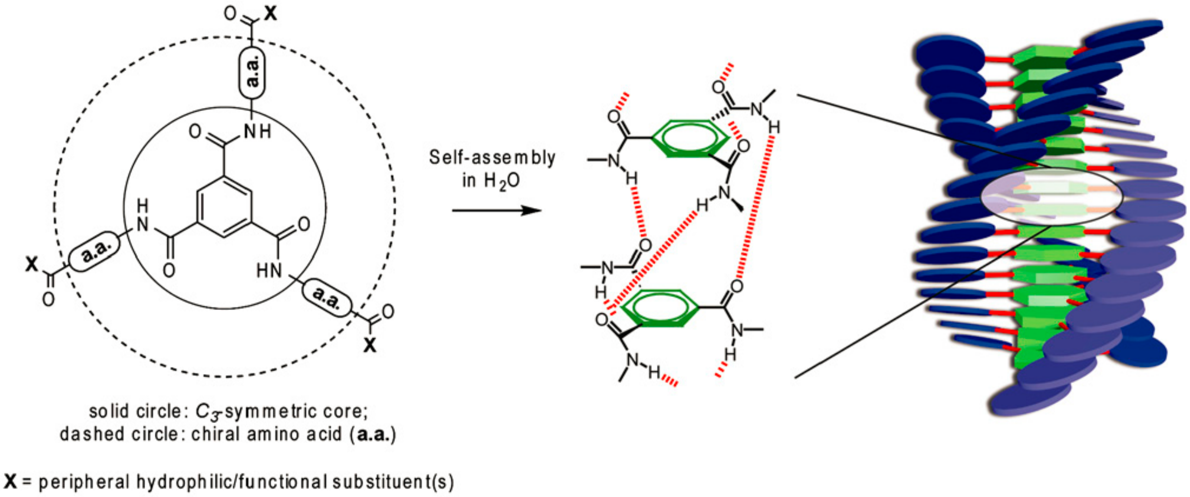


Figure 3. Structure of the BTA core with a tunable spacer and cartoon assembly of the supramolecular structure. Reproduced with permission from Ref. [9]; Copyright 2010 National Academy of Sciences.

BTA derivatives self-assemble through a combination of π-π stacking between the benzene rings and H-bonding between the amides. The amide moieties can be functionalized to fine tune the properties of the resulting supramolecular polymer. As an example, monomers tagged with different fluorophores have been mixed together in order to study the kinetics of supramolecular fiber formation using stochastic optical reconstruction microscopy (STORM).[10]

Stacking among aromatic groups is another commonly used strategy in supramolecular polymerization. Due to the hydrophobic nature of the interaction, suitable monomers can easily self-assemble in aqueous solution, and the resulting polymers are soluble if the periphery is functionalized with polar groups. The alignment of π-orbitals in these assemblies has been used to develop supramolecular polymers capable of charge transfer to create photoconductive nanowires and semiconductors.[11] This motif has also been used develop photocatalytic systems to produce hydrogen with chromophore amphiphiles.[12-13]

Studies on the host-guest chemistry of synthetic molecules began five decades ago with investigations of crown ethers binding metal cations.[14] These efforts advanced to include organic molecule guests nearly a decade later.[15] SMPs built by host-guest interactions allow control over directionality and high specificity within the system.[1-2, 4, 16] They may be composed of either two different homoditopic monomers or a single heteroditopic monomer with host functionality on one end and guest functionality on the other (Figure 4).

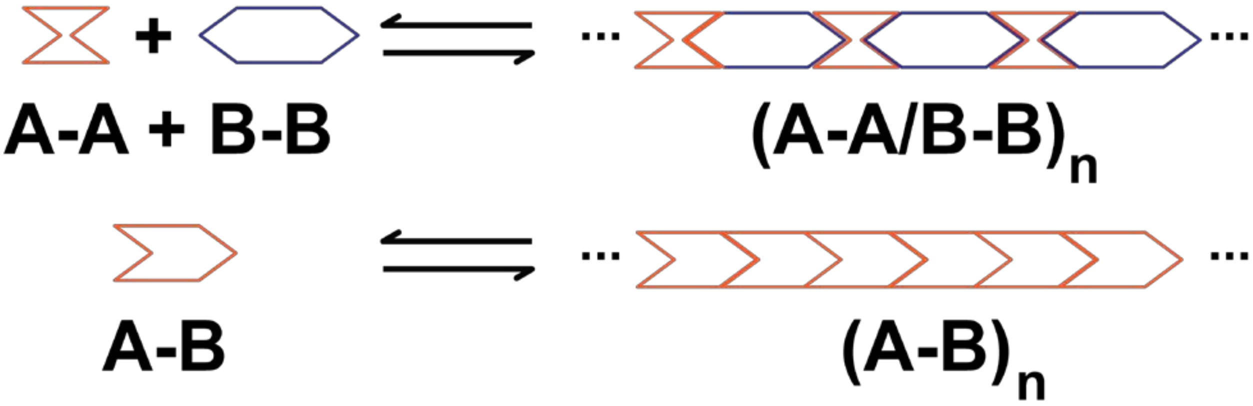
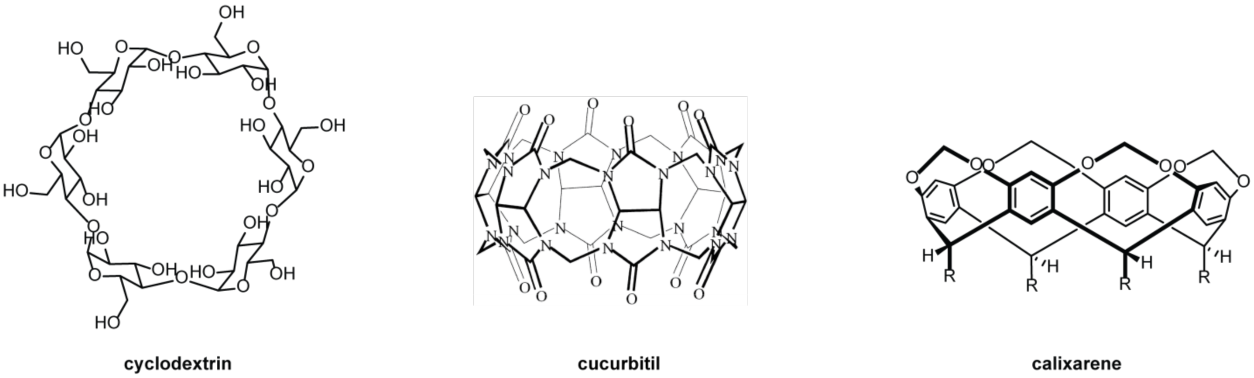


Figure 4. Homoditopic (A-A + B-B) and heteroditopic (A-B) type supramolecular polymerization.

Other stoichiometries and configurations are also possible depending on the structure and degree of functionalization of the linker core (i.e. dendrimer cores).[16] Host functional groups in SMPs are usually cyclic molecules like cyclodextrins, cucurbitils, and calixarene (Figure 5), each of which binds a specific type of guest molecule.[1]



cyclodextrin cucurbitil calixarene

Figure 5. Example structures of commonly used host molecules: cyclodextrin, cucurbitil, and calixarene. Structure of cucurbitil reproduced with permission of the International Union of Crystallography.[17]

Cyclodextrins are macrocycles composed of glucose monomers and can bind a variety of guests including hydrocarbons, polar compounds (carboxylic acids and amines), and hydrophilic polymers.[18] Cucurbitils may bind either one or two guest molecules at a time, the identity of which depends on ring size; smaller cucurbitils bind gas molecules, metal ions, and aliphatic amines, while larger rings enable the binding of bulkier aromatic, adamantyl, and fullerene derivatives.[19] Calixarene binds porphyrin or viologen. Different host-guest systems can be readily combined together to produce more complex SMPs due to the orthogonality of the interactions.[4]

SMPs formed through metal coordination are an important subset of host-guest systems.[4] Here, two or more organic ligands (the hosts) form a coordination bond with a metal ion (the guest). Of particular interest to the work presented in Chapter 4, is the field of supramolecular metallopolymers (or coordination polymers). This encompasses everything from simple 1-D materials to highly ordered 3-D metal organic frameworks (MOFs). Both metallopolymers and metal organic frameworks harness the strong forces of metal coordination to direct the formation of supramolecular assemblies.[20-22]

One of the first coordination polymers base on an organic ligand was formed through binding of Cu(I) by 4,4′,4″,4‴-tetracyanotetraphenylmethane.[23-24] Ligand moieties used to generate supramolecular polymers through metal coordination typically bind metal ions in a 2:1 or 3:1 (ligand:metal) ratio with pyridyl, carboxylate, and alkynyl functionalities being the most commonly used motifs.[23] These ligand groups are usually connected through a linker—the flexibility and length of which can be used to tune molecular properties. Polymers built through metal coordination have a high degree of modularity and tunability since most ligands will bind many different metals. Recently, orthogonal binding of two different metals into a supramolecular polymer has been achieved.[25] Metal coordination has also been combined with host-guest chemistry to produce supramolecular polymers with two orthogonal interactions to produce a great array of 1-D and 2-D assemblies.[4]

The first MOF was described two decades ago exhibiting high thermal stability and the ability to host aromatic guest molecules.[26] Since that time, the structural modularity of simple multi-dentate organic ligands chelating metal ions has been successfully utilized to produce diverse crystalline materials with tunable structure and properties. The variety of MOFs resulting from this structural control has led to applications in areas such as catalysis, gas capture, and sensing.[27-31]

Highlighted below are a few recent examples of functional MOF architectures that have been developed (Figure 6).

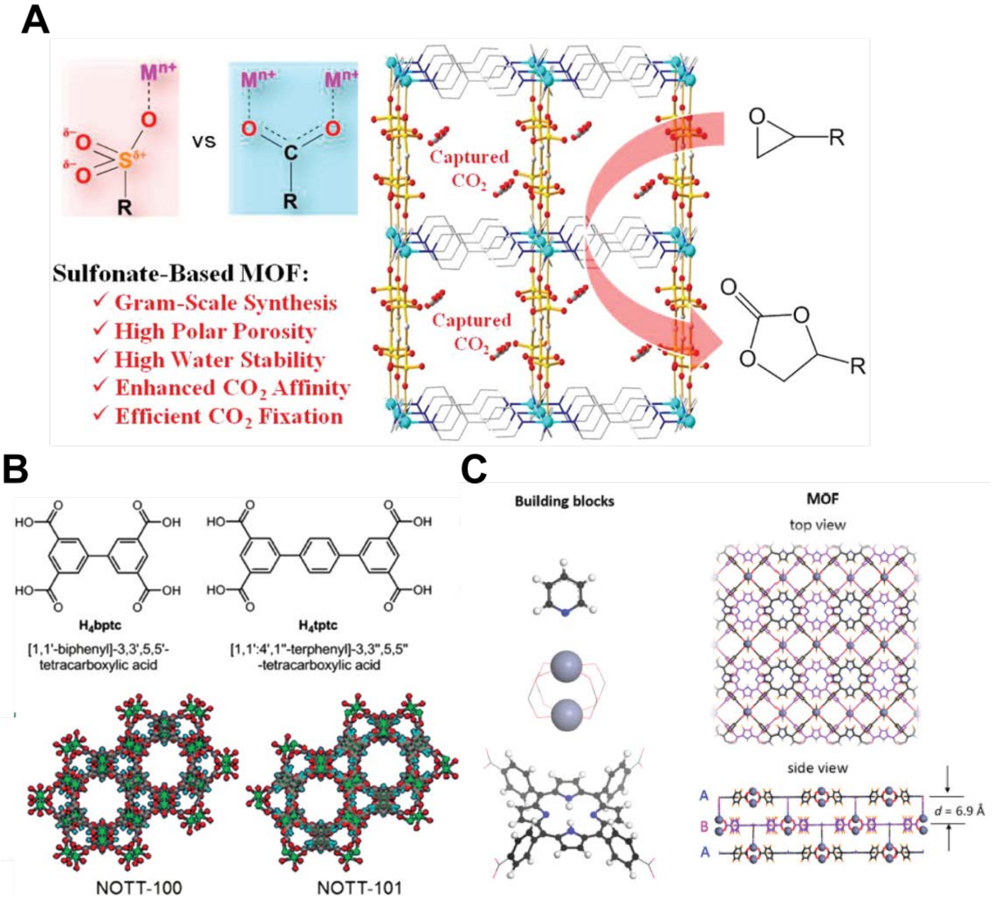


Figure 6. Examples of functional MOF architectures.

A) A MOF with the ability to convert CO_2 into carbonates. Reprinted with permission from *Chem. Mater.*, **2016**, *28* (17), pp 6276–6281. Copyright 2016 American Chemical Society. B) Thin film MOF architectures capable of gas sensing. Adapted with permission from reference [32]. Copyright © 2016 Verlag GmbH & Co. KGaA, Weinheim. C) Schematic of a MOF with potential use in solar cells. Reprinted with permission from *ACS Appl. Mater. Interfaces*, **2016**, *8* (45), pp 30863–30870. Copyright 2016 American Chemical Society.

A MOF based on Cu^{2+} with bipyridyl and 1,2-ethanedisulfonate ligands was reported with the ability to both capture and convert CO_2 into carbonates by reacting them with epoxides.[33] Methods for growing MOFs composed of aromatic carboxylic acids and Cu^{2+} ions in thin films has resulted in devices with the ability to sense water and volatile organic compounds.[32] MOFs designed for use in solar cells have been reported utilizing 3-D architectures composed porphyrin/pyridine linkers coordinated to $Zn(OAc)_2$ with a thin film of palladium porphyrin as exciton acceptor.[34] The great utility of MOFs for a wide range of applications has inspired the development of MOF-like architectures from protein ligands.

1.1.1 Naturally Occurring Supramolecular Polymers

Nature largely utilizes macromolecules to construct supramolecular materials which ultimately give rise to the great array of forms and functions of life. Living organisms can be considered as highly complex supramolecular machines. The emergent properties unique to life arising from the organization and interaction of these supramolecular systems demonstrate that the whole is more than merely the summed properties of the components. Examples of important biological supramolecular materials based on lipids and proteins are detailed below. Peptide-based materials are discussed in more detail in Section 1.2. Although not found in nature, supramolecular materials that self-assemble into programmable shapes have been developed from DNA using the complementarity and specificity of its bases.[35] This field of DNA "origami" has advanced to the point where almost any shape can be drawn that software will then prescribe the requisite DNA strands to make.[36-37]

Lipids are a class of biomacromolecules that form 2-D and 3-D supramolecular polymers in the form of micelles and plasma membrane bilayers.[38] Lipids are composed of hydrophobic tails with polar head groups. Membrane bilayer properties can be altered between solid gels, liquid crystals, and liquid rafts by tuning the type and composition of the lipids present. These materials are used to contain and separate biological components into compartments and for signal tranduction. This facilitates differentiation between cells and organelles along with transportation of cellular cargo.

Compared to lipids, proteins are a much more structurally diverse class of biomacromolecule. They play myriad structural and functional roles both within and outside of cells. SMP forming proteins are usually structural elements.[39] A few examples include (Figure 7): amyloid fibrils (involved in neurodgenerative disease),[40] collagen (a component of extra-cellular matrix support),[41] myosin/actin filaments (responsible for contractile motion and transport within cells),[42] and keratin (vital to epithelial protection and support).[43]

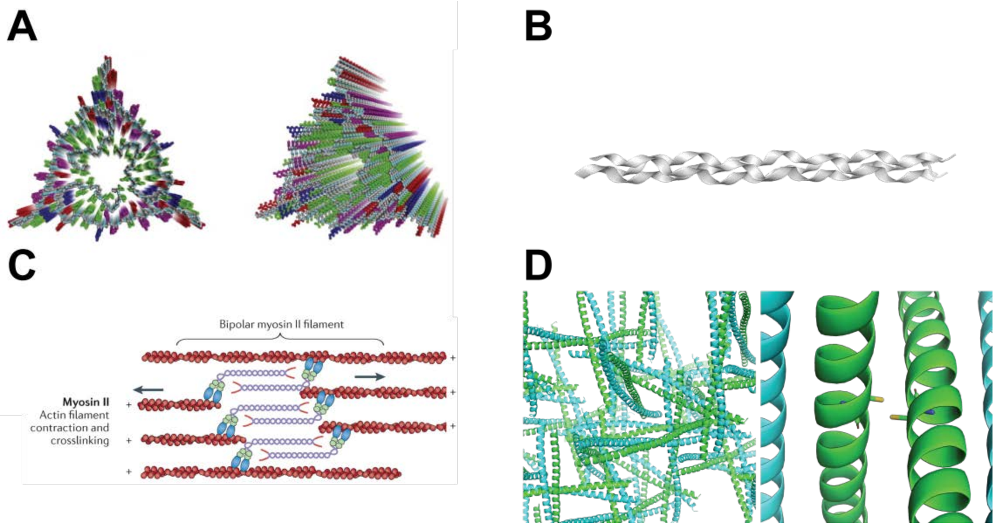


Figure 7. Examples of supramolecular fibers formed by proteins.

A) Structure of $A\beta_{1\text{-}40}$ fibers derived from solid state NMR. Reprinted from *Cell*, 154, Jun-Xia Lu, Wei Qiang, Wai-Ming Yau,Charles D. Schwieters, Stephen C. Meredith, Robert Tycko, Molecular Structure of β-Amyloid Fibrils in Alzheimer's Disease Brain Tissue, Pages 1257-1268, Copyright 2013, with permission from Elsevier. B) Cartoon structure of collagen (PDB ID: 1cag) derived from a 1.9 Å resolution X-ray crystal structure.[44] C) Schematic of the supramolecular interaction between myosin II and actin filaments for contractile function. Reprinted by permission from Macmillan Publishers Ltd: *Nature Reviews Neuroscience* reference [45], copyright 2013. D) Crystal structure (PDB ID: 3TNU) of an α-keratin heterodimer; (close-up) Cys residues in proximity for disulfide formation.[43]

These different classes of supramolecular assemblies arise from different secondary structure elements: amyloid fibrils are composed of extended β-sheets,[40] collagen is composed of a trimer of polyproline type II helices,[41] the actin monomer is a globular protein,[42] and keratin is built from a dimeric α-helix.[43] Proteins may also be embedded in lipid membranes as part of a heterogeneous supramolecular material.[46] Many of these membrane proteins play important functional roles such as: membrane channels, signal receptors, and energy transformers.[46]

Living organisms demonstrate that supramolecular materials can be used to produce multifunctional organic machines. The burgeoning field of synthetic biology will rely heavily on understanding the design principles behind these systems. The ability to create supramolecular materials with tunable and multifunctional properties opens up a world of possibilities limited only by the human imagination and the laws of nature.

1.2 PROTEIN-BASED SUPRAMOLECULAR MATERIALS

Protein-protein interactions are a powerful means of generating supramolecular materials. Proteins and peptides (short synthetically-accessible fragments of proteins) are particularly well-suited as a basis for material self-assembly due to their modularity and structural diversity.[39, 47] At the nanometer-scale level in living organisms we observe a dizzying array of supramolecular protein-based machines and scaffolds. A few examples of their many functions include: energy collection (photosystem II),[48] structural (cytoskeleton),[42] membrane transport (aquaporin),[46] supramolecular enzyme complexes (mitochondrial respiratory complexes).[49] Many designed materials draw inspiration from these structures that have been observed in nature.

Non-covalent interactions between the functional groups of the primary amino acid sequence and with solvent molecules control both the structure and ultimately function of polypeptides.[50] These interactions can give rise to local folding patterns (secondary structure). Interactions between secondary structure domains can generate complex folds (tertiary structure) and supramolecular complexes (quaternary structure). Quaternary structures can range in size from as small as a few nanometers to as large as micrometers (cytoskeletal fibers, collagen, viral capsids).[51]

Even with the vastness of polypeptide sequence space available (20^n for the canonical amino acids, where n is the peptide length), there are only a few recurring secondary structure motifs: α-helix, β-sheet, β-turn, loop, and 3_{10}-helix (Figure 8).

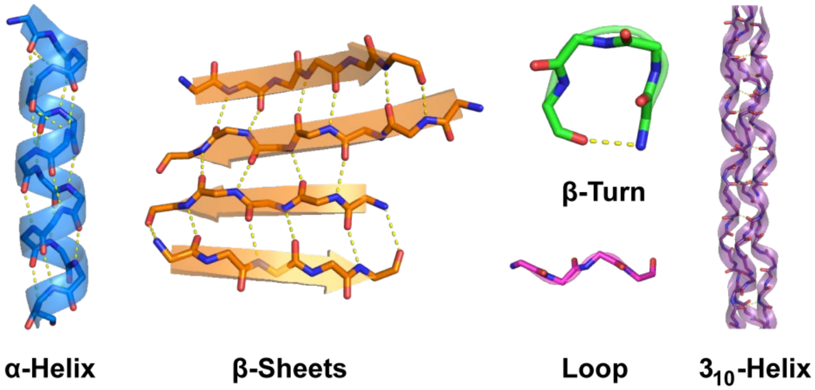


Figure 8. Common protein secondary structural motifs. Hydrogen bonds between main chain amides are depicted as dashed yellow lines.

These secondary structures are defined by different patterns of hydrogen bonding between or within strands. α-Helices arise from a 13-membered H-bonded ring, while β-hairpin turns and 3_{10}-helices arise from a 10-membered ring. β-sheets occur from peptide strands in extended conformations and may H-bond either between sections of the same sequence or intramolecularly (i.e. amyloid fibrils). Amides in loops typically only H-bond with solvent and are not useful in generating supramolecular polymers, except as linkers between different domains.

Secondary structures can combine in numerous ways to form tertiary folds or supramolecular quaternary structures. Control over quaternary structure, since it can be the basis for supramolecular polymerization, is of particular interest to the development of peptide-based materials.[52] Folding is not a prerequisite for peptide self-assembly, and interesting supramolecular materials have been developed from short, non-secondary structure forming peptides. These short peptides can be used to form supramolecular polymers similar to other small molecules through H-bond networks and aromatic stacking,[53] or metal chelation to form peptide MOFs.[54] The focus here will be on the self-assembly of larger oligomers where folding is an important part of supramolecular assembly. Common design motifs used to construct protein-based supramolecular polymers include: coiled coils, amyloid fibrils, protein-ligand interactions, viral capsids, and metal chelation.[52, 55]

Coiled coils are a quaternary structure that has found wide use in the construction of peptide-based materials.[52] Coiled coils are assemblies formed between 2-7 α-helices;[56] their design rules are described in more detail below (Section 1.3). Early work engineering coiled-coil based materials resulted in strategies for forming supramolecular fibers.[57] "Sticky-end" coiled coils that self-assemble into fibers (Figure 9) can be generated by offsetting the charge-charge interactions peripheral to the hydrophobic core[58] or rotation of the hydrophobic core to a different helical face.[59]

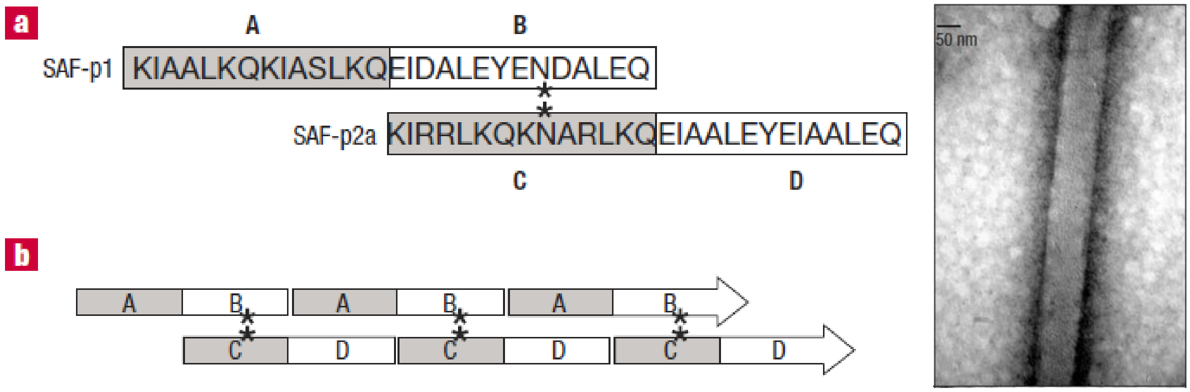


Figure 9. "Sticky-ended" coiled coil heterodimers designed to form supramolecular fibers.

(Reprinted by permission from Macmillan Publishers Ltd: *Nature Materials* reference [59], copyright 2003.)

Fibrils can also be created through judicious placement of salt bridges and residues with low steric bulk.[60-61] The similarity among the above examples is that they utilize exclusively naturally occurring amino acids without any post-synthetic modification of the oligomers. Expanding beyond this design constraint can have significant benefits. One technique to form larger assemblies from coiled coils is to join two or more strands together with a covalent linker, so that upon folding the subunits self-polymerize. This strategy has been used to form both linear supramolecular polymers[62-63] as well as large self-assembling cages.[64]

Peptide amphiphiles are short peptides functionalized with long alkyl tails (typically at the N-terminus). In this manner, they are similar to lipids being composed of a hydrophobic domain and a more polar head group (Figure 10).

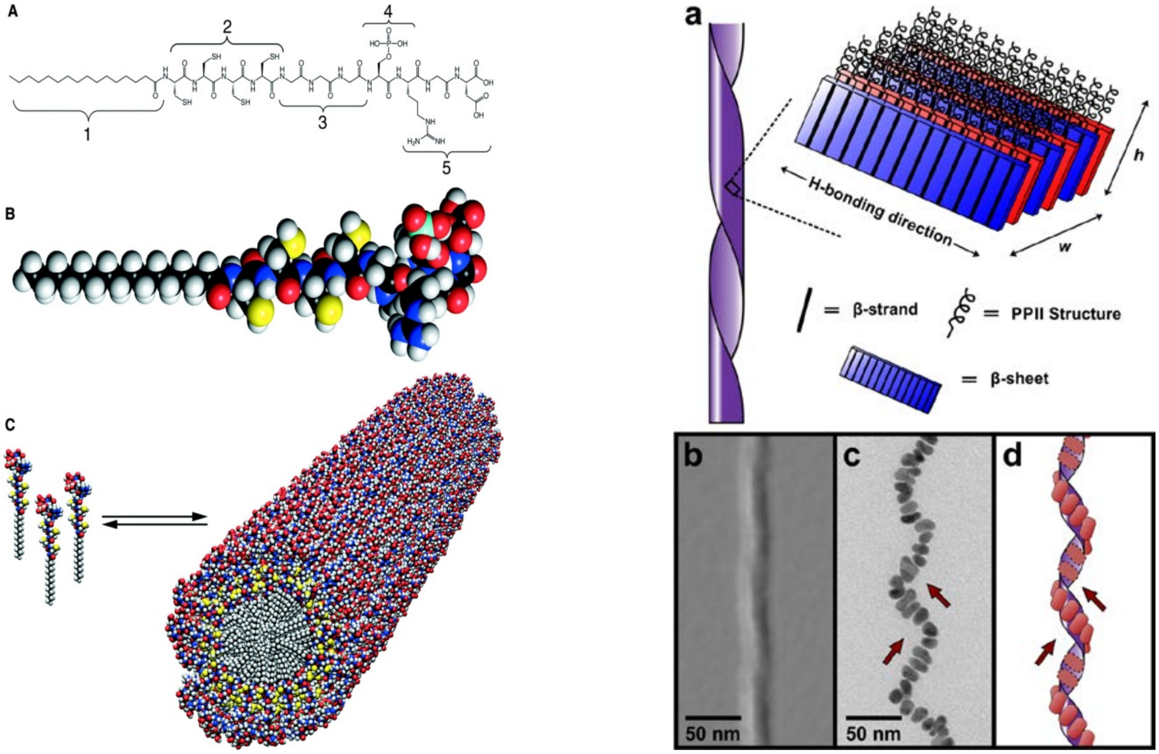


Figure 10. Peptide amphiphile self-assembly.

Self-assembly of a peptide amphiphile nanofiber (left, reproduced from reference [65], Copyright 2002 National Academy of Sciences) and templation of a helical nanoparticle composite (right, reprinted with permission from Merg, A. D., et al. *J. Amer. Chem. Soc.* **2016**, *138* (41), pp 13655–13663).[66]

Unlike lipids, the head group in a peptide amphiphile can dictate properties through folding and self-assembly. Self-assembly of these molecules results in large fibers with the alkyl chains on the interior and peptide head groups exposed to water;[65] the final morphology of which can also be varied through solvent effects.[67] Peptide amphiphiles have also been used with great success to template the formation of composite materials, such as superhelical Au nanoparticle arrays[33] imparting chirality to the optical properties of the nanoparticles.[66]

Another class of peptide-based SMP are cyclic peptides with alternating D,L stereochemistry, which can assemble into columnar nanotubes (Figure 11).[68-69]

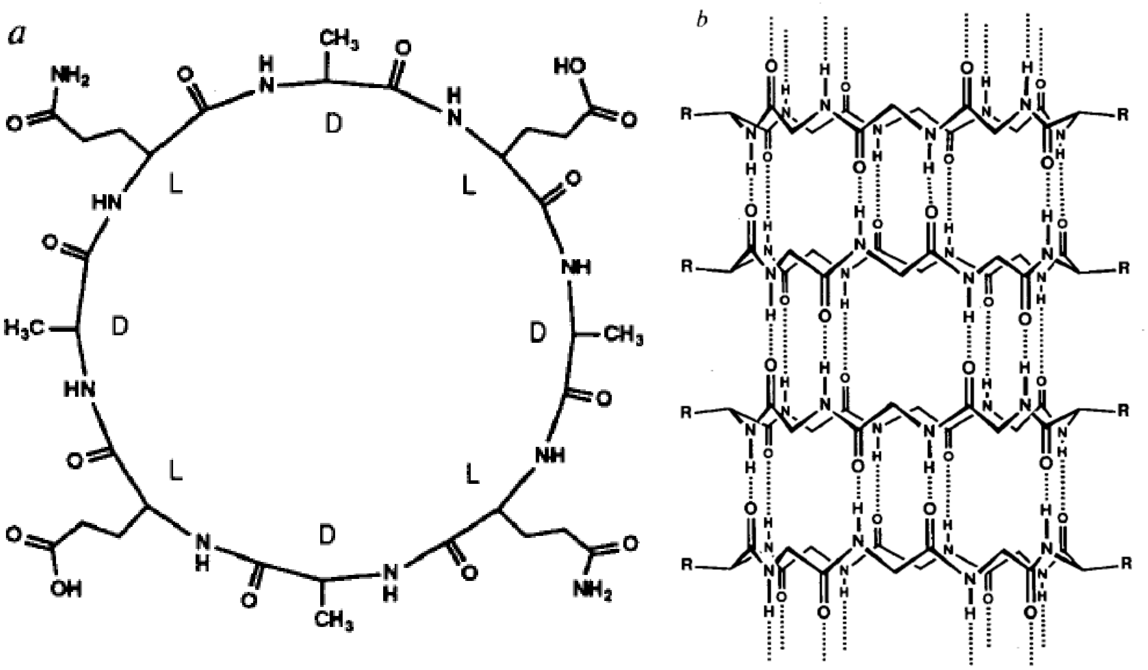


Figure 11. 2D-structure of a eight residue cyclic peptide *cyclo*[-(D-Ala-Glu-D-Ala-Gln)$_2$-] and self assembly into a tubular column through anti-parallel stacking of the peptides mediated by amide H-bonding.

Reprinted by permission from Macmillan Publishers Ltd: *Nature* **1993,** *366*, 324 - 327,[69] copyright 1993.

The cyclic D,L-peptides are fixed into a rigid secondary structure regardless of the amino acid sequence. They self-assemble into columnar fibers through backbone amide hydrogen bonds and laterally through non-covalent interactions among side chains. Nanotube properties can be tuned through the choice of side chains including unnatural amino acids and chromophores to impart electronic properties.[70] Cyclic peptide nanotubes show promising applications for antibiotics, ion channels biosensors, gene delivery, photoactive supramolecular polymers, and electronic materials.[71-72]

Due to the ability of metals to impart important redox and charge transfer properties, protein designs incorporating metals have been long studied.[73] Recent efforts have focused on expanding their application to direct the self-assembly of protein-based supramolecular architectures.[74-75] Pioneering work has shown that expressed proteins bearing metal-coordinating side chains can be used to this end (Figure 12A).[55, 76]

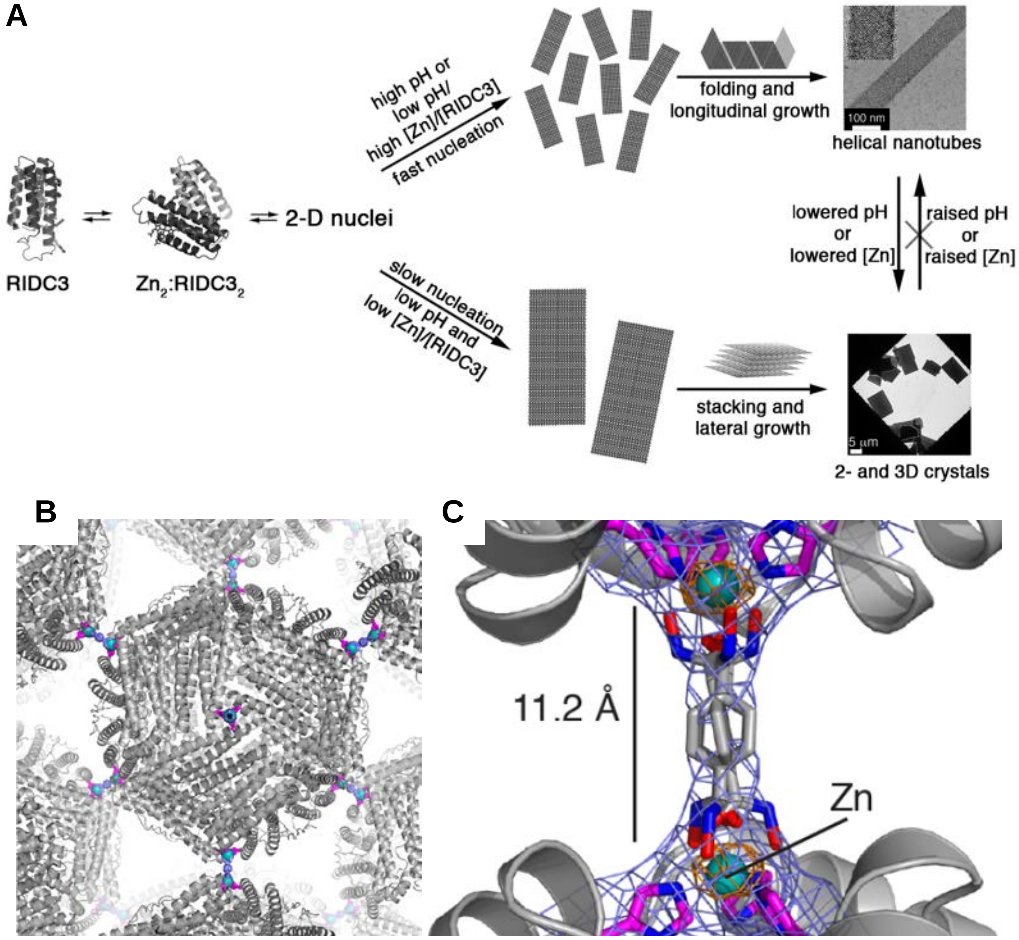


Figure 12. Metal-directed self assembly of protein supramolecular architectures.

A) Zn^{2+}-directed protein assembly of an engineered version of cytochrome$_3$ (RIDC3) into helical nanotubes or 2D and 3D crystals. Reprinted by permission from Macmillan Publishers Ltd: *Nature Chemistry* reference [76], copyright 2012. B) Crystal structure of T122H ferritin in the presence of Zn^{2+} and linker benzene-1,4-dihydroxamic acid. C) Close-up of linker coordination to Zn^{2+} bridging two ferritin proteins. Adapted from ref. [77] through an open access *ACS Editors' Choice* License. Copyright 2015 American Chemical Society.

Incorporating metal-bining amino acids (like His) into the C3 symmetric vertices of the octahedral ferritin quaternary structure can be used to bind a cognate small-molecule as a linker and form highly ordered lattices (Figure 12B,C).[77]

Metal coordination has also been used for the supramolecular organization of shorter synthetic peptides that form defined quaternary structures. These architectures are notable for coupling the orthogonal forces of highly specific peptide-directed association with metal binding[78] in a fully synthetic polypeptide ligand. Collagen has been functionalized with bipyridine to direct the formation of fibers, discs, and hollow spheres using divalent Fe, Cu, or Co ions (Figure 13).[79-81]

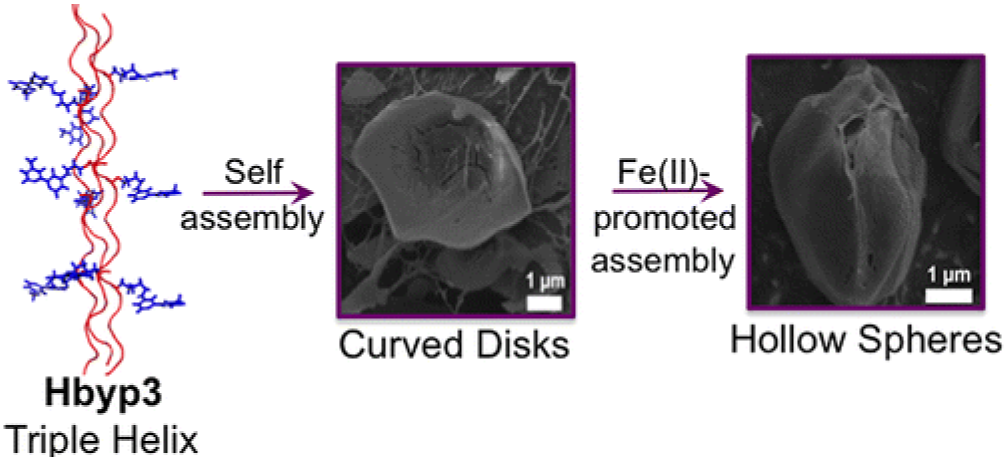


Figure 13. Schematic of bypiridine-modified collagen triple helix (Hbyp3) which self assemble into either curved disks (no metal ion) or hollow spheres (with a divalent ion, i.e. Fe^{2+}).

Reprinted with permission from Przybyla, D. E. et al. *J. Am. Chem. Soc.*, **2013**, *135* (9), 3418–3422.[81] Copyright 2013 American Chemical Society.

Synthetic coiled-coils have also been modified with a variety of metal-binding moieties to produce supramolecular polymer and crystalline assemblies.[82-84] Expansion upon this theme to form highly-ordered one, two, and three dimensional crystalline assemblies and supramolecular polymers is the focus of Chapter 4.

1.3 ALPHA-HELICAL COILED COILS

Coiled coils are protein structural motifs consisting of two or more α-helices bound together in a bundled quaternary structure. Biologically, they function in a variety of roles such as DNA transcription factors, structural components, molecular motors, and receptors.[85] Structurally, coiled coils may be parallel or antiparallel, homo- or heterooligomeric, and contain between two to seven helices in a single bundle. Each of these characteristics can be programmed by the amino acid sequence.[86] As with all α-helices, coiled-coil forming sequences are described using a helical wheel to designate the alignment of each residue in the amino acid sequence (Figure 14).

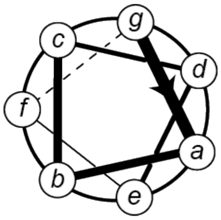


Figure 14. Helical wheel diagram for an α-helix.

In α-helices, two full turns of the helix are completed by a sequence of seven residues known as the heptad repeat. Each position around the α-helix is designated with a lower case letter (*a-b-c-d-e-f-g*). This feature of α-helices means that an alternating pattern of *i, i+3, i+7, etc.* residues share the same helical face and causes a coiled coil to have a left-handed superhelical twist.[87] By convention, the *a* and *d* heptad positions are used to indicate the hydrophobic face of the helix, which forms the buried core of the multi-helix coiled-coil assembly.[87] More complex hydrophobic patterns like the hendecad and pentadecad also exist and give rise to coiled coils with a nearly parallel and right-handed supercoil, respectively.[86] Due to the presence of a hydrophobic face within the folded α-helix, coiled-coil quaternary structure is directly coupled to the folding of the individual oligomers and monomeric helix is rarely observed. Observation of helical structure implies coiled coil formation.[88]

Although the coiled-coil motif was first hypothesized by Crick in 1953 from the X-ray diffraction pattern of α-keratin,[89] the first coiled-coil crystal structure was not fully solved until 1991 when the structure of the leucine zipper region of the yeast transcription factor GCN4 was published (Figure 15).[87]

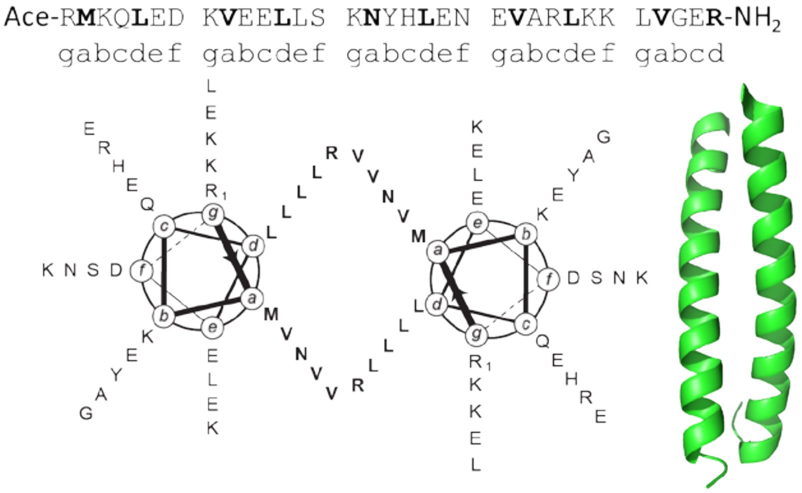


Figure 15: Sequence (top), helical wheel diagram (left) and crystal structure (right, PDB ID= 4DMD) of GCN4p1. Hydrophobic core packing residues are highlighted in bold.

This crystal structure demonstrated that the GCN4 leucine zipper, termed GCN4p1, was a parallel coiled-coil homodimer exhibiting a "knobs-into-holes"[89] packing motif of the hydrophobic *a* and *d* residues.[87] Although GCN4p1 was thought to only exist as a dimer, twenty years later it was discovered that under certain conditions it will adopt a parallel trimer fold.[90]

Subsequent studies since the first crystal structure of GCN4p1 have greatly increased our understanding of the coiled-coil folding motif. Varying chain length established that at least three heptads are necessary to form stable folded dimers.[91] Mutations to the hydrophobic core demonstrated that oligomerization state and stability can be controlled simply by changing the residues at the *a* and *d* heptad positions.[92-94] A seminal study in the mid- 1990's determined the effects of systematically mutating these positions with Ile, Val, and Leu and created a nomenclature for describing GCN4p1 mutants where the letters after the 'p' indicate residues present at the *a* and *d* positions, respectively.[92] For example, mutating all *a* positions to Ile and all *d* positions to Leu leads to the peptide "GCN4pIL". Studies on pairing preferences of coiled-coil interfaces have shown that β-branched side chains (Ile and Val) at *d* positions largely favor trimers and tetramers while leucine at this position was more promiscuous. Mutants pIL (parallel dimer), pII (parallel trimer), and pLI (parallel tetramer) exhibited single oligomerization states while the other mutants gave rise to mixtures.[92] More recently, sequences have been designed that assemble to form a pentamer,[95] hexamer,[61] heptamer[96] and dodecamer.[97] Typically, these higher order helix bundles (hexamer and above) are the result of mutating the *e* and *g* positions to alanines allowing room for the required close packing arrangement or mutations to the hydrophobic core. Pentamers are unusual in that they are created from larger hydrophobic residues at the *a* and *d* heptad positions like phenylalanine[98] or tryptophan.[95]

Placing a hydrophilic residue in the core at an *a* heptad position was shown to de-stabilize the folded stability but improve oligomerization state and topology specificity.[92, 99] In general, a single Asn residue at an *a* position favors dimers while a single Gln residue favors trimers;[92, 100-

[101] however, Asn at the second *a* position can create an atypical three-state unfolding pathway in which the coiled coil changes from dimer to trimer to monomer with an increase in temperature.[102]

The *e* and *g* heptad positions also play a significant role in determining coiled-coil oligomerization state and the formation of homo- vs. hetero- assemblies. Observing from the GCN4p1 crystal structure that the *e* and *g* positions engage in salt bridges spanning the coiled-coil interface[87] and knowing that the naturally occuring Fos and Jun peptides formed a heterodimer,[103] researchers designed a *de novo* set of heterodimeric pLL peptides designated BASEp1 and ACIDp1 (Figure 16).[104]

BASEp1: **Ac-AQLKKKL QALKKKN AQLKWKL QALKKKL AQ-NH_2**
ACIDp1: **Ac-AQLEKEL QALEKEN AQLEWEL QALEKEL AQ-NH_2**

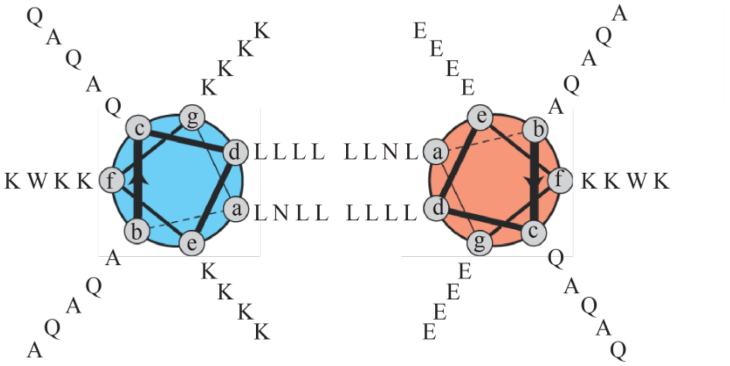


Figure 16: First de novo designed heterodimer coiled-coil system comprising of two leucine zippers one with all glutamates (carboxylic acids) at the e and g positions (ACIDp1, red) and one with all lysines (primary amines) at the e and g positions (BASEp1, blue). Helical wheels are depicted facing one another as they would in the coiled coil.

These sequences differ only at *e* and *g* positions: all Lys residues at these positions in BASEp1 and all Glu residues at corresponding positions in ACIDp1. Under physiological conditions, the homodimers are destabilized through charge-charge repulsion while the heterodimer is favored through complementary charge interactions at the *e/g* interface.[104] These two peptides are the basis for the heterodimer system used later to characterize the subunits of our coiled-coil based supramolecular polymer system (see Section 2.2). A later report described the crystal structure of a *de novo* designed heterotrimer that preferentially formed over the competing homo- and heterodimers.[101] More recently, a parallel heterotetramer has been designed.[105] Unnatural residues have also been used to control coil-coil pairing by using complementary, variable length *e* and *g* positions[106] or changing the size of the hydrophobic core residues.[107-108]

The knowledge of the coiled-coil structural motif that has developed over the past few decades has greatly increased our understanding of protein folding and allowed for a wide range of rationally designed structures. A number of excellent reviews[86, 88, 98, 109-110] delve more into the details of how various mutations affect coiled-coil folding and function. Recently, researchers have created a *de novo* set of homooligomeric coiled-coil peptides four heptads in length with full biophysical and structural characterization.[111] These sequences were intended for use by others as a starting point for further engineering and we made use of these in the work described in Chapter 4. With this plethora of information available it is now largely possible to choose a desired coiled-coil structure and then develop a sequence for that particular application.

1.4 PROTEIN HELICES CONTAINING BACKBONE MODIFICATION

In addition to their importance in materials, α-helices also have an important role in biochemistry. α-Helices make up approximately 40% of the protein secondary structure found in nature and are often found involved in protein-protein interfaces (PPIs).[112-113] PPIs are involved in many cell signaling pathways and disease states, but are often considered "un-druggable" by small molecules since they typically encompass large, flat surface areas.[114] Thus, mimicry of α-helical structure is important to many biomedical applications. Several strategies for mimicking α-helices exist including peptide stapling[115] as well as the use of non-peptide scaffolds,[112] and foldamers.[116] A number of excellent reviews have been written on helix mimicry,[112, 116-118] but the focus here will be precedent involving the types of residues examined in Chapter 3.

The field of folded peptide mimetics or "foldamers"[119] seeks to develop oligomers with mixed or wholly unnatural backbones that fold to form defined structures in aqueous solution. One motivation for incorporating unnatural backbone units into peptides is to improve stability to proteolytic degradation, a key limitation of peptide therapeutics.[120]

Helical secondary structure has been particularly well-studied as a foldamer design motif. An array of strategies has been used to create backbones that adopt helical structures. Some of these strategies are only modest modifications of the α-amino acid residues found in natural peptides (Figure 17): addition of an extra methylene unit between amides (β-residues), methylation of the alpha carbon (C_α-methyl residues), amide→thioamide substitution, amide→ester substitution, amide replacement with a 1,2,3-triazole[121] and transposition of the side chain from C_α to N (peptoids).[117, 122]

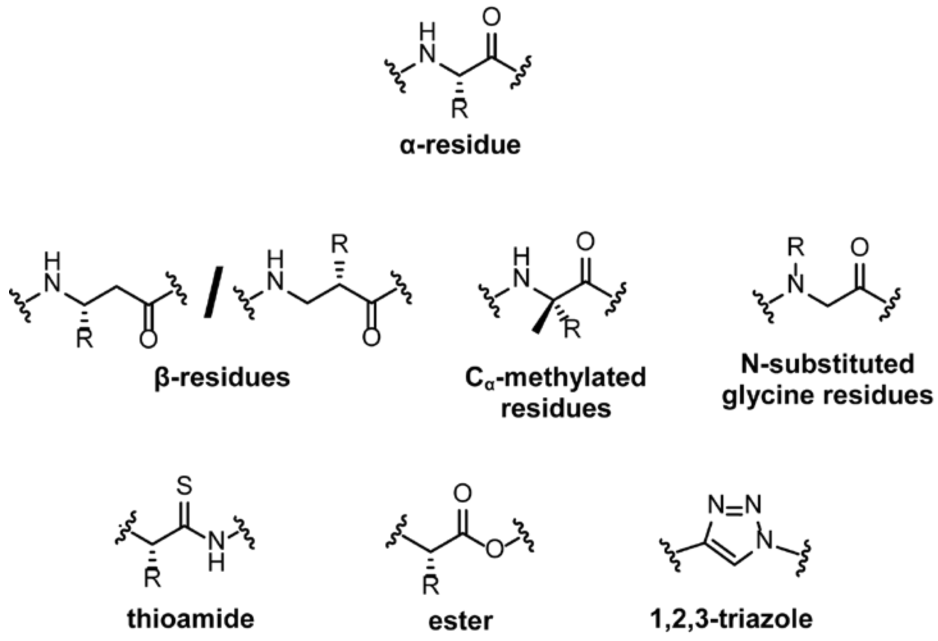


Figure 17. Commonly used helix backbone modifiers in peptide mimicry; the canonical α-residue backbone is depicted at the top.

More exotic helical designs have also been developed including urea-based backbones[123] and conjugated aromatic polymers.[124]

Examining prior work involving β-residues, helices have been developed using both mixed α/β backbones as well as pure β backbones. Peptides made up exclusively of β-residues (β-peptides) can form helices similar to peptides with pure α-backbones, but with wider diameters and larger intramolecular H-bonding rings.[125] Early work with pure β backbones showed oligomers composed of the cyclic β-residues aminocyclohexanoic acid (ACHC)[126] and aminocyclopentanoic acid (ACPC)[127] formed stable helices in organic solvents (Figure 18).

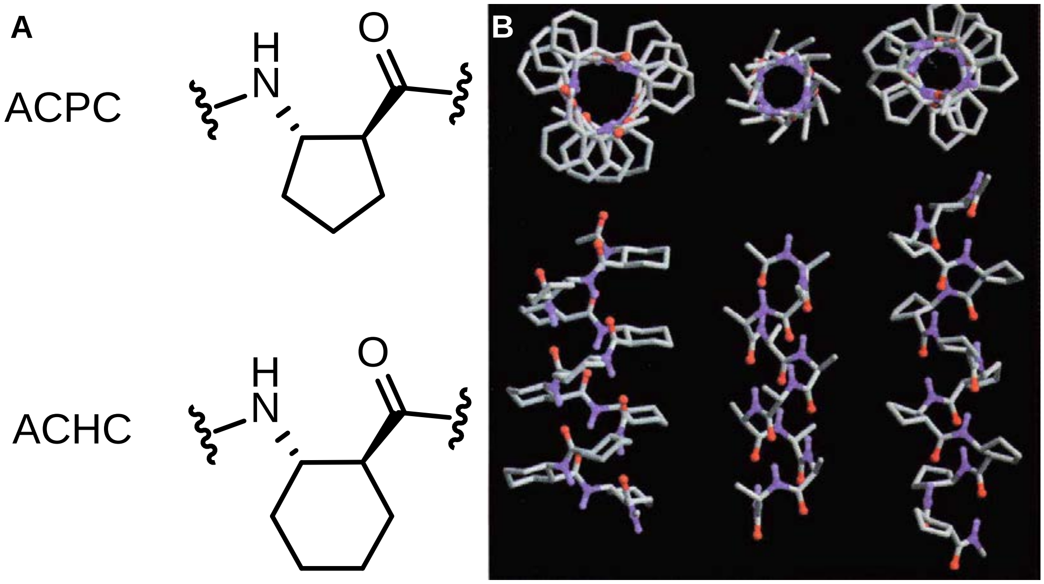


Figure 18. Oligomers composed of acyclic β-residues.

A) Structures of ACPC and ACHC residues. B) Idealized ACHC decamer (left), alanine decamer (middle), and ACPC decamer (right). Adapted by permission from Macmillan Publishers Ltd: *Nature* Reference [127], copyright 1997.

Reports of oligomers composed of acyclic β-residues that fold in water quickly followed.[128-129] Seminal work in mixed α/β backbones resulted in one of the first foldamer quaternary structures characterized at high resolution, a coiled coil (Figure 19).[130]

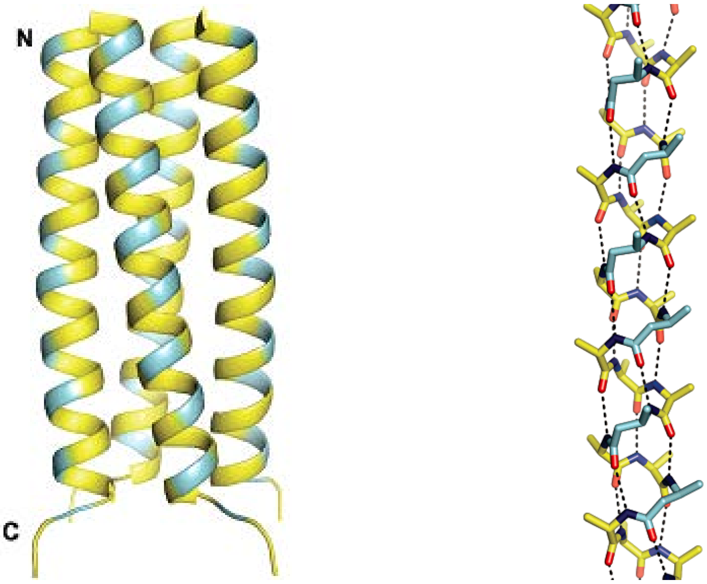


Figure 19. Crystal structure of a tetrameric helical bundle formed by a mixed α/β-peptide Cartoon (left) and backbone H-bonding structure (right) with α-residues in yellow and β-residues in cyan; side chain atoms are omitted (PDB: 2OXK).. Adapted with permission from Horne, W. S. et al. *J. Am. Chem. Soc.*, **2007**, *129* (14), pp 4178–4180.[130] Copyright 2007 American Chemical Society.

The disadvantage of mixed backbone helices is that their folds are less thermodynamically stable than the native forms.[131] Patterning of α- and β-residues within the helix can be used to control helical handedness.[132] Control over left/right handed helix propensity was achieved by altering the sequence length of peptides composed of *cis*-aminocyclicpentanoic acid (ACPC) residues with alternating L/D α-residues in various patterns (ααβ, αααβ, αβααβ, and ααβαααβ).[132]

Methylation of the alpha carbon in amino acids is another well-studied modification for promoting helical folds. C_α-methylated residues, though rare, are naturally occurring amino acids found in some fungi which use them in the synthesis of non-ribosomal peptide antibiotics.[133] Their propensity to form helical structures was first observed with aminoisobutyric acid (Aib), the C_α-methylated version of Ala.[134] Aib was found to have a more restricted Ramachandran plot than L-α-residues with dihedral angles preferring only helical or fully extended regions.[134] Helices containing Aib are more stable than those formed by longer Ala peptides,[135] and Aib was used successfully to increase folded stability of the full-length protein thermolysin.[136] It has long been assumed that the increased thermodynamic stability that Aib imparts arises from a decrease in the entropy of folding based on its restricted conformational freedom.[136-137] This assumption had never been tested with a full thermodynamic analysis until work described in Chapter 4.[138]

1.5 PROJECT GOALS

1.5.1 Origin of Linker Flexibility in Coiled-Coil Based Supramolecular Polymers

In prior work from the Horne lab, homodimeric coiled-coil based SMPs (CC-SMPs) were constructed by cross-linking two peptide strands of GCN4p1 with organic linkers at the *f*-position of the second heptad.[62] Unlike previous coiled-coil cross-linked fibers, where the linker was parallel to the fiber axis,[63] linkers in this system are perpendicular due to cross-linking at an internal position as opposed to the termini. Two symmetric linkers were examined in prior work: one based on an ethylene diamide (EDA) core and another based on a piperizine diamide (PIP) core; synthesis of the cross-linked species was achieved through an S_N2 reaction between a Cys thiol and primary alkyl halides at either end of the linkers. Importantly, the linkers separate the sulfurs by 8 atoms allowing for independent folding of the two cross-linked coiled-coil domains, which was not the case for disulfide cross-linked variants.[62] These linkers differ by only two carbons, yet SMPs containing them exhibited different physical properties. Note the rotational constraint of the Pip linker as opposed to the EDA linker between the amides (Figure 20).

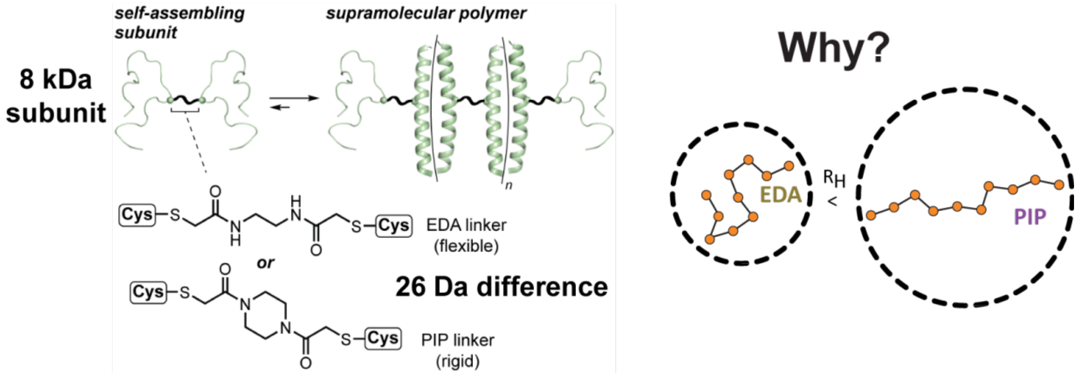


Figure 20. Schematic of supramolecular polymer self-assembly from a subunit cross-linked with either the Pip or EDA linker.

Non-denaturing gel permeation chromatography (GPC) indicated that the ring-constrained Pip linker produced assemblies with a larger hydrodynamic radius while the more flexible EDA linker produced assemblies with a smaller hydrodynamic radius.[62] Whether or not the polymerization number (*n*) is actually significantly different could not be accurately determined by this method. The more likely reason for the observed difference in hydrodynamic radius is the flexibility of the EDA linker causing CC-SMPs composed from it to be more compact and "spaghetti-like" while the those composed of the Pip linker are straighter "hair-like" fibers. In this thesis, we report the molecular origin of these differences through alternate biophysical characterization methods of structure and dynamics in the CC-SMP backbone (Chapter 2).

1.5.2 Design of Coiled-Coil Based Supramolecular Metallopolymers

An alternative to covalently cross-linking coiled coils with organic linkers to promote supramolecular assembly is to do so through metal coordination. This strategy has been used extensively in the supramolecular polymer and metal organic framework fields and provided inspiration for this work.[23, 54, 139-141] The driving goal of this work was to design a series of coiled-coil peptides that assemble into supramolecular architectures driven by metal coordination and tunable by peptide sequence (Figure 21).

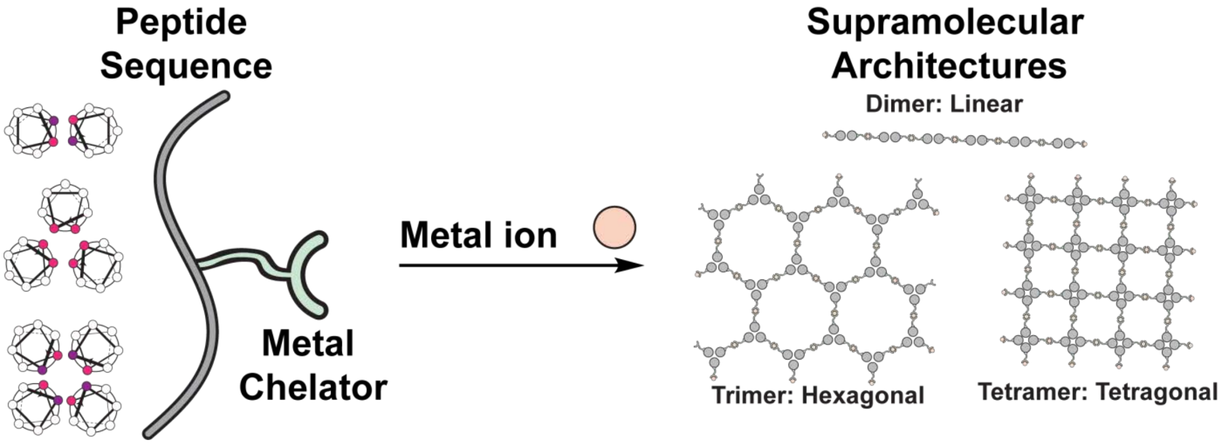


Figure 21. Schematic for generating supramolecular architectures from peptide metal coordination. Question marks indicate tunable aspects of the system.

Incorporation of a metal chelating group into the peptide allows for the formation of supramolecular materials by the simple addition of a metal ion at an appropriate stoichiometric ratio. This assembly process can be easily tuned by the number of equivalents and identity of the metal.

We sought to control the morphology of the final supramolecular architecture through judicious placement of the metal-chelating site and oligomerization state of the coiled coil. The design of peptide sequences that assemble in the presence of metal to form crystalline and solution phase assemblies was pursued in parallel. Negative results were fed back into the design process to inform the development of subsequent systems. The result was a novel class of highly ordered metallopeptide architectures with analogies to MOF-based materials (Chapter 3).

1.5.3 Improving Helix Backbone Modification Strategies

The B1 domain of the *Streptococcal* G protein (GB1) had been previously used by our lab to elucidate the design rules of heterogeneous backbone modification within the context of a tertiary fold.[142] GB1 is a synthetically accessible protein containing all of the major secondary structural motifs (α-helices, β-sheets, β-turns, and loops). Backbone substitutions were found for each structural element that allowed the protein to retain its native fold.[142] Modification of the helix, while tolerated, proved to be one of the most de-stabilizing substitutions. Subsequent full thermodynamic analysis of proteins incorporating acyclic and cyclic β-residues in the helix sought to understand the origin of this effect.[131]

Although multiple strategies have been developed for helix backbone modification, thermodynamic comparison between them within the same system is lacking (Figure 22).

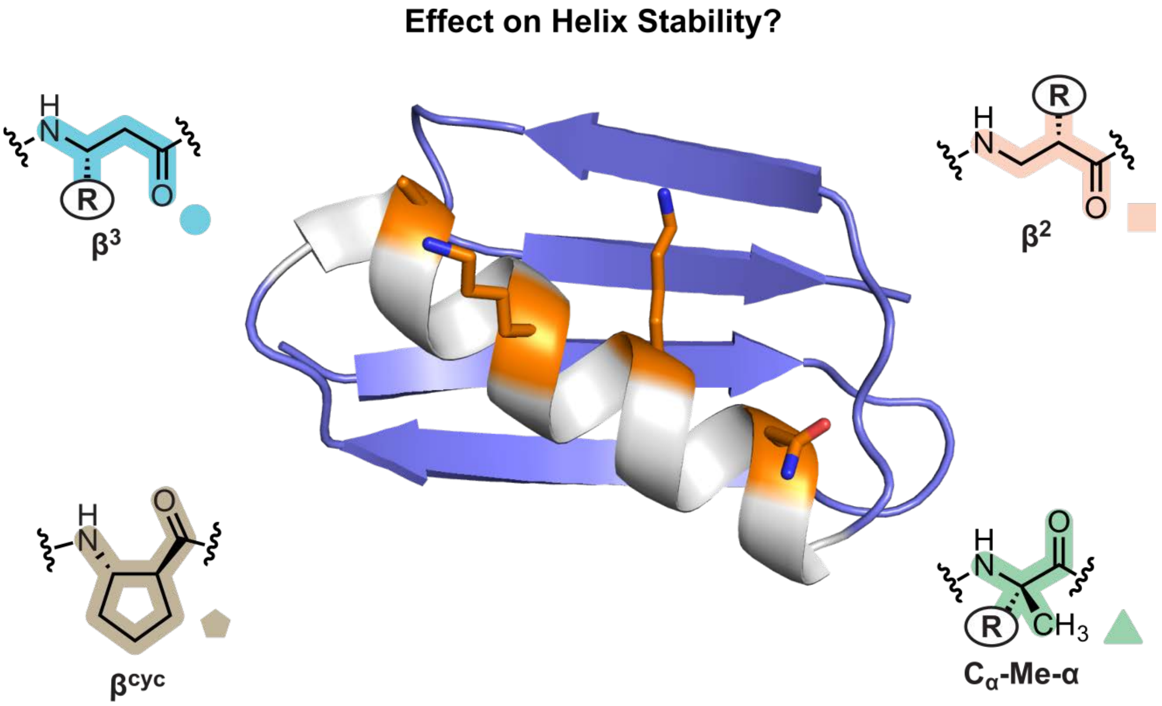


Figure 22. How does helix backbone modification effect the folded stability of a tertiary fold? Crystal structure of the GB1 tertiary fold (PDB ID = 2QMT) with mutation sites highlighted in orange; unnatural amino acid types compared surround it.

We turned once again to the helix of GB1 to explore this question. Calculations suggested that certain regioisomers of β-residues may form thermodynamically more stable helices than others.[143] It has also long been known that methylation of the alpha carbon strongly promotes the helical conformation.[144] By careful thermodynamic analysis of each substitution type we were able to directly compare these different strategies for helix backbone modification alongside those examined by the lab previously (Chapter 4).

2.0 ORIGIN OF LINKER FLEXIBILITY IN COILED-COIL BASED SUPRAMOLECULAR POLYMERS

The work in this chapter has been previously published as:

> Tavenor, N. A.; Silva, K. I.; Saxena, S.; Horne, W. S. "Origins of Structural Flexibility in Protein-Based Supramolecular Polymers Revealed by DEER Spectroscopy." *J. Phys. Chem. B.* **2014**, 118, 9881-9889.

The crystallography experiments in this chapter were performed by Dr. W. Seth Horne. Electron spin resonance data were collected and analyzed by Dr. K. Ishara Silva. The computational modeling was supported in part by the University of Pittsburgh Center for Simulation and Modeling through the supercomputing resources provided. A particular acknowledgement of thanks goes to Dr. Albert DeFusco for his help in the use of those resources.

Prior work on protein-based supramolecular polymers (SMPs) generated by coiled coils cross-linked by a short organic linker revealed that a subtle change in the structure of the linker altered the apparent size of the SMP.[62] Upon changing the linker from a linear ethylenediamine (EDA) core to a cyclic piperizine (PIP), an increase in assembly size was observed by gel permeation chromatography (Figure 23) even though the strength of the non-covalent forces holding the assembly together (peptide domain dimerization) remained unchanged.[62]

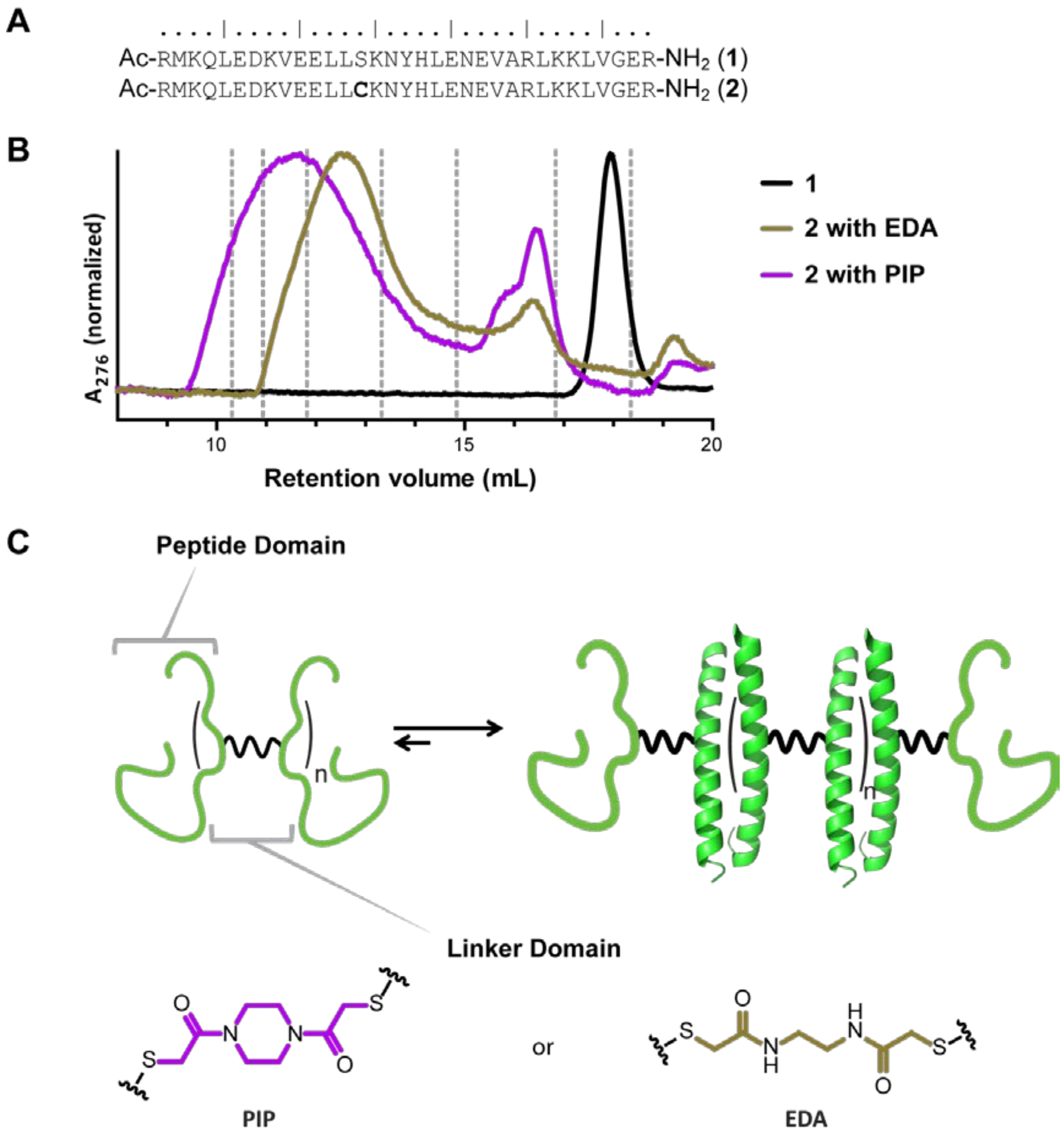


Figure 23. Coiled-coil based supramolecular polymers with modular linkers.

A) Sequences of GCN4p1 **(1**) and S_{14}→C mutant (**2**). B) Gel permeation chromatograms of peptide **1** and **2** cross-linked with either the EDA (gold) or PIP (magenta) linker (GPC data adapted from Ref. [62] with permission from The Royal Society of Chemistry). C) Supramolecular assembly of a cross-linked peptide and structures of the linkers.

An open question from this work was the physical basis for the difference in apparent size of the assemblies generated from EDA- and PIP-based linkers. We endeavored to answer this question through a detailed measurement of the conformational range of each linker.

With the aim of better understanding the role of the linker in dictating the supramolecular polymer properties, we utilized a pulsed electron paramagnetic resonance (EPR) technique known as double electron-electron resonance (DEER) spectroscopy to measure the probability distance distributions between paramagnetic spin-labels across a cross-linked coiled coil subunit. DEER measurements interrogate the dipolar relaxation signal between two paramagnetic centers from which a distance probability distribution can be extracted.[145] This signal appears as an decaying sine wave. Period length is related to the distance between spin-labels (longer periods implies longer distances) and signal decay rate is related to the distribution of distances present (a slower decay implies a narrower distance distribution).[146] DEER was particularly well-suited to this task as it is sensitive to the dipolar resonance between paramagnetic electrons within 1.5-16 nm of one another.[146-149] This is approximately an order of magnitude greater than the distance detected by the analogous NMR technique of measuring NOEs.[145]

Applying DEER to the study of linker flexibility in the coiled coil supramolecular assemblies discussed above was acccomplished in three stages: (1) determining the ideal electron spin label for the system, (2) designing a system where a single subunit could be isolated and analyzed, and (3) constructing a molecular model based on the experimental results.

2.1 COMPARISON OF SPIN-LABEL STRATEGIES

Spin labels applied in EPR can be any moiety that contains an unpaired electron. This unpaired electron may either be in the form of a paramagnetic metal ion or an organic radical. An ideal spin label must be able to be incorporated into a molecule of interest without perturbing its structure and also contribute minimal conformational ambiguity in the position of the spin relative to the system of interest. The most widely used organic spin-label functional group is a nitroxide radical composed of an oxygen bonded to a nitrogen flanked by two pairs of geminal dimethyl groups. This moiety can be incorporated into peptides and proteins in different ways; the two most common are as an MTSSL-modified Cys (denoted as ρ) and as a TOAC residue (denoted as Σ), depicted below (Figure 24).

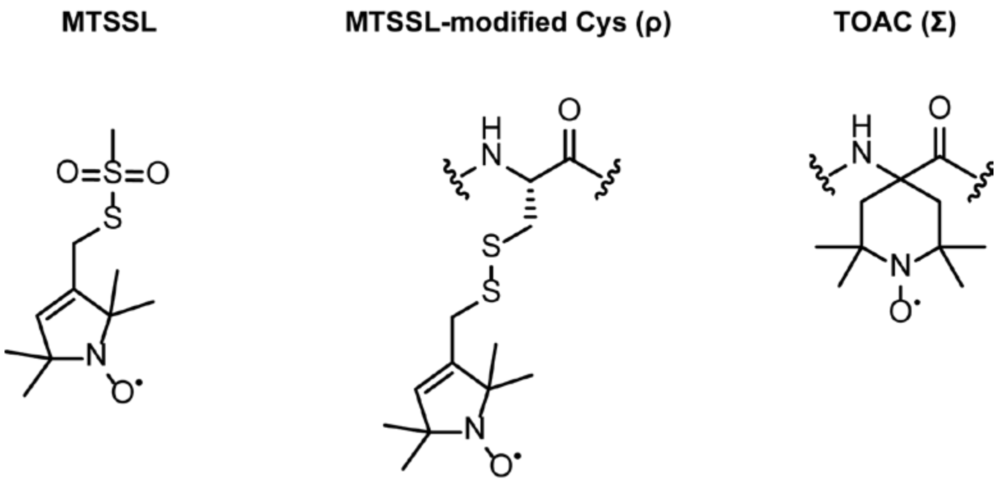


Figure 24. Structures of two commonly used nitroxide spin-labeling residues in peptides: MTSSL-modified Cys (ρ, left) and TOAC (Σ, right).

Incorporation of a nitroxide into a synthetic peptide may be achieved by either coupling the modified residue on resin or through post-synthetic modification of a functional side chain (e.g. Cys). The latter approach is also readily applied to biologically expressed proteins.[146] Incorporation on resin during solid-phase peptide synthesis (SPPS) widens the repertoire of modifications available to include highly rigid labels like the amino acid TOAC. TOAC's high steric hindrance decreases the nucleophilicity of the amino group and significantly increases the difficulty of peptide synthesis.[150] Post-synthetic modification typically utilizes MTSSL attachment through quantitative disulfide formation with a Cys residue.[151]

Structurally, the MTSSL moiety is less likely to perturb the folding of a protein compared to TOAC and may be placed at any solvent-exposed site. A disadvantage to MTSSL labeling is the number of freely rotatable side-chain bonds; these increase the ambiguity in the final structural analysis. The restricted conformational freedom of TOAC means that it can only be incorporated at sites that fall within its allowed dihedral range, which correspond to β-bend, α-helix and 3_{10}-helix.[152] This restricted freedom is advantageous for a spectroscopic probe since the functional group measured in the analysis (the nitroxide) is tightly coupled to the feature of interest (the peptide backbone).

Since these two spin-labeling strategies (TOAC and MTSSL) have been used extensively,[153-154] we compared the two methods to determine which would be best-suited to the study of the supramolecular polymer system at hand (*vide supra*). The parent peptide sequence GCN4p1 (**1**) was modified by either mutation of $S_{14} \rightarrow$ Cys (**2**) for MTSSL attachment generating peptide **3** or mutation of $E_{10} \rightarrow \Sigma$ for TOAC incorporation generating peptide **5** (Figure 25). Peptide **4** is a control sequence for TOAC incorporation in which a simplified analogue is used in place of the spin-label.

. . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . .

Ac-RMKQLEDKVEELLSKNYHLENEVARLKKLVGER-NH_2 **(1)**
Ac-RMKQLEDKVEELL**C**KNYHLENEVARLKKLVGER-NH_2 **(2)**
Ac-RMKQLEDKVEELL**ρ**KNYHLENEVARLKKLVGER-NH_2 **(3)**
Ac-RMKQLEDKV**X**ELLSKNYHLENEVARLKKLVGER-NH_2 **(4)**
ABA-RMKQLEDKV**Σ**ELLSKNYHLENEVARLKKLVGER-NH_2 **(5)**

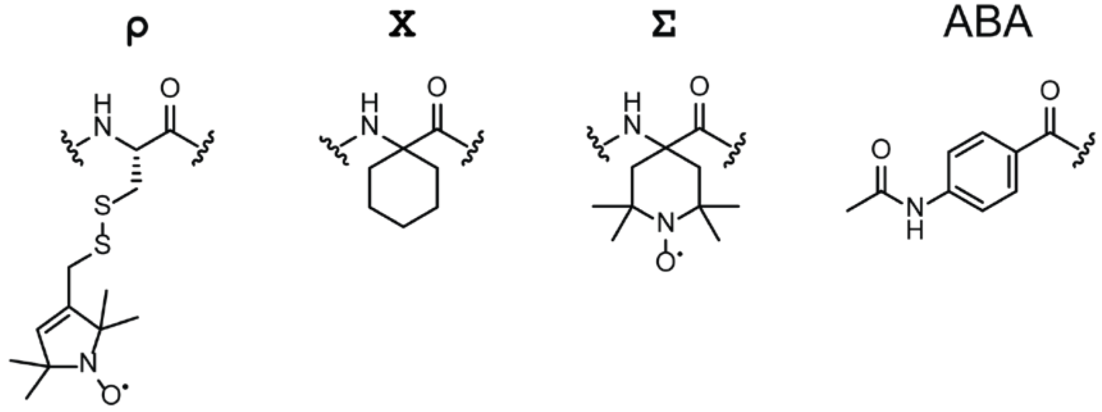


Figure 25. Sequences of peptides **1-5** with structures of unnatural residues MTSSL-modified Cys (ρ), Ahx (X), TOAC (Σ), and acetamidobenzoyl N-terminal cap (ABA).

Circular dichroism (CD) thermal melts and X-ray crystallography were used to study the perturbation to the folded structure while DEER measurements were made to compare the conformational flexibility in the folded coiled-coil quaternary structure.

2.1.1 Effect of Spin-Labeling on Peptide Folding

Peptide **2** was synthesized as a variant of GCN4p1 (**1**) with the point mutation $S_{14} \rightarrow C$ in order to facilitate attachment of MTSSL. MTSSL was attached through disulfide formation in aqueous buffer using known methods yielding peptide **3**.[151]

Prior to synthesis of peptide **5**, a structural analogue lacking the nitroxide was made using the related amino acid Ahx (X) at the E_{10} site. This site was chosen for modification based on analysis of a published structure of GCN4p1 (PDB 4DMD)[90] suggesting that TOAC at this position would not perturb the fold. Factors that were considered in this analysis were 1) the possibility steric clashes with other residues, 2) loss of important intermolecular interactions (e.g. salt bridges, hydrogen bonds, hydrophobic core), and 3) proximity to the linker attachment site. Mutation of E_{10} was found to least likely have an impact considering these factors. Due to its increased steric bulk, Ahx required alternative coupling strategies for itself and the two subsequent residues (*vide infra*).

Peptide **4** was crystallized by hanging drop vapor diffusion and yielded a structure (PDB 4TL1) that was solved to 1.80 Å resolution. Comparison with a published structure of the native GCN4p1 coiled coil (PDB 4DMD) resulted in a backbone alignment with a 0.29 Å RMSD between C_α atoms (Figure 26).

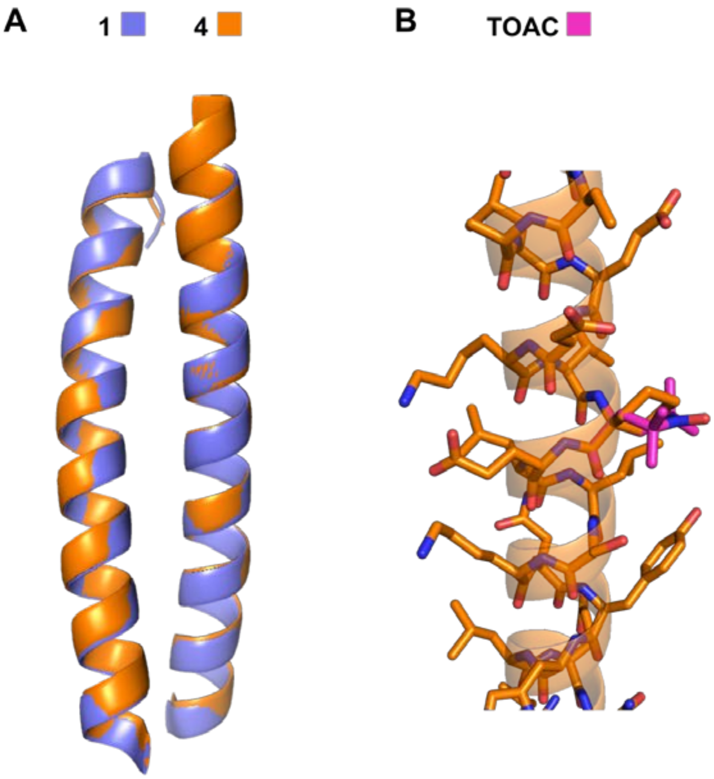


Figure 26. Crystal structure of peptide **4**.

A) Alignment of X-ray crystal structures of peptides **1** (PDB ID = 4DMD) and **4** (PDB ID = 4TL1), C_{α} RMSD = 0.29 Å; B) modeled overlay of the TOAC residue with the Ahx residue.

A CD scan of peptide **4** indicated solution-phase helical structure and thus formation of the coiled-coil dimer (Figure 27).

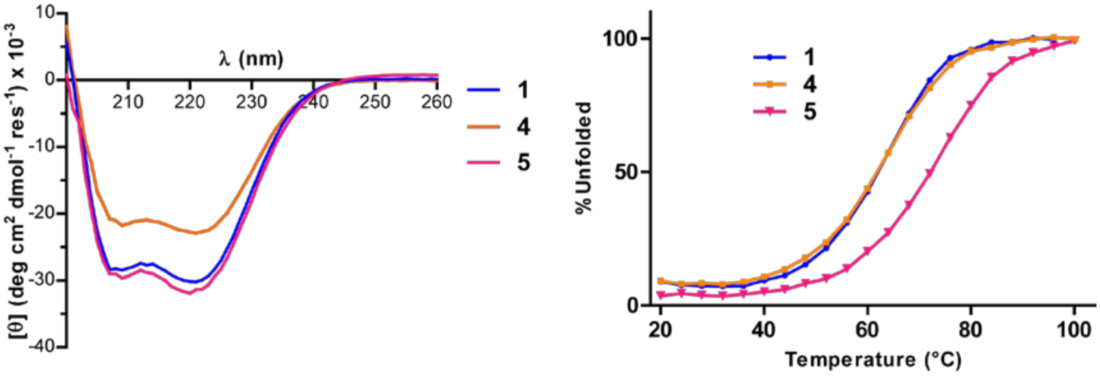


Figure 27. CD scans (left) and melts (right) of peptides **1**, **4**, and **5**

Peptides were measured at 100 μM measured in aqueous HEPES buffer (10 mM pH 7).

The scan of peptide **4** has a lower ellipticity signal than peptide **1** which may indicate slightly less helical content; however the thermal melt indicated that peptide **4** has an identical folded stability as wild type. Confident that incorporation of TOAC at this site would not perturb the folded structure, we synthesized peptide **5**.

Amino acid coupling involving TOAC required alternative strategies due to its extreme steric bulk. Nucleophilicity of the amino group is known to be very weak (pK_a ~ 6.0 for the conjugate acid),[155] and therefore subsequent amino acids do not couple to it easily. Serendipitously, we discovered that Fmoc protection of the amino group is unnecessary since TOAC will not self-polymerize during SPPS like typical amino acids. This discovery greatly reduces the cost for TOAC incorporation since it may be used in synthesis as the free amino acid. The amino acid immediately following TOAC (Val_9) was activated as the acid fluoride in order to achieve effective coupling.[156] The difficulty of this coupling is likely a combined effect of TOAC's low nucleophilicity and Val's sterically hindered carboxylate resulting from the beta-branched side chain. The N-terminus of peptide **5** was amidated with acetamidobenzoic acid to provide a chromophore with a strong extinction coefficient to aid in concentration determination of the peptide by UV-Vis spectroscopy. Due to the acidic nature of the peptide cleavage/deprotection reaction, the nitroxide is converted to a hydroxylamine by protonation of the oxygen. The radical was regenerated by treatment with aqueous ammonia prior to HPLC purification.[150, 157]

Crystallization trials of peptide **5** were unsuccessful; however, a CD scan confirmed the secondary structure was similar to **1**, and thermal melt analysis revealed an improved thermal stability over wild type (Figure 27, right). Convinced that both spin-labeled peptides (**3** and **5**) were structurally identical to the wild type we measured the DEER spectrum of the dimeric coiled-coils formed by each.

2.1.2 Comparison of DEER Spectra

Samples of **3** and **5** for DEER measurements were prepared at 100 µM in 10 mM HEPES buffer, pH 7 with 20% glycerol (v/v). Cryoprotection with glycerol is necessary for flash freezing the samples resulting in a snapshot of the system at close to room temperature.[146] The results of the DEER measurements were as were consistent with expectations based on relative label flexibilitiy. The coiled coil formed by the peptide labeled with MTSSL exhibited a longer and broader probability distribution, while the peptide labeled with TOAC gave rise to a very narrow and shorter probability distribution. Most probable distances for both were consistent with models generated from the crystal structures of related peptides **1** and **4** (Figure 28).

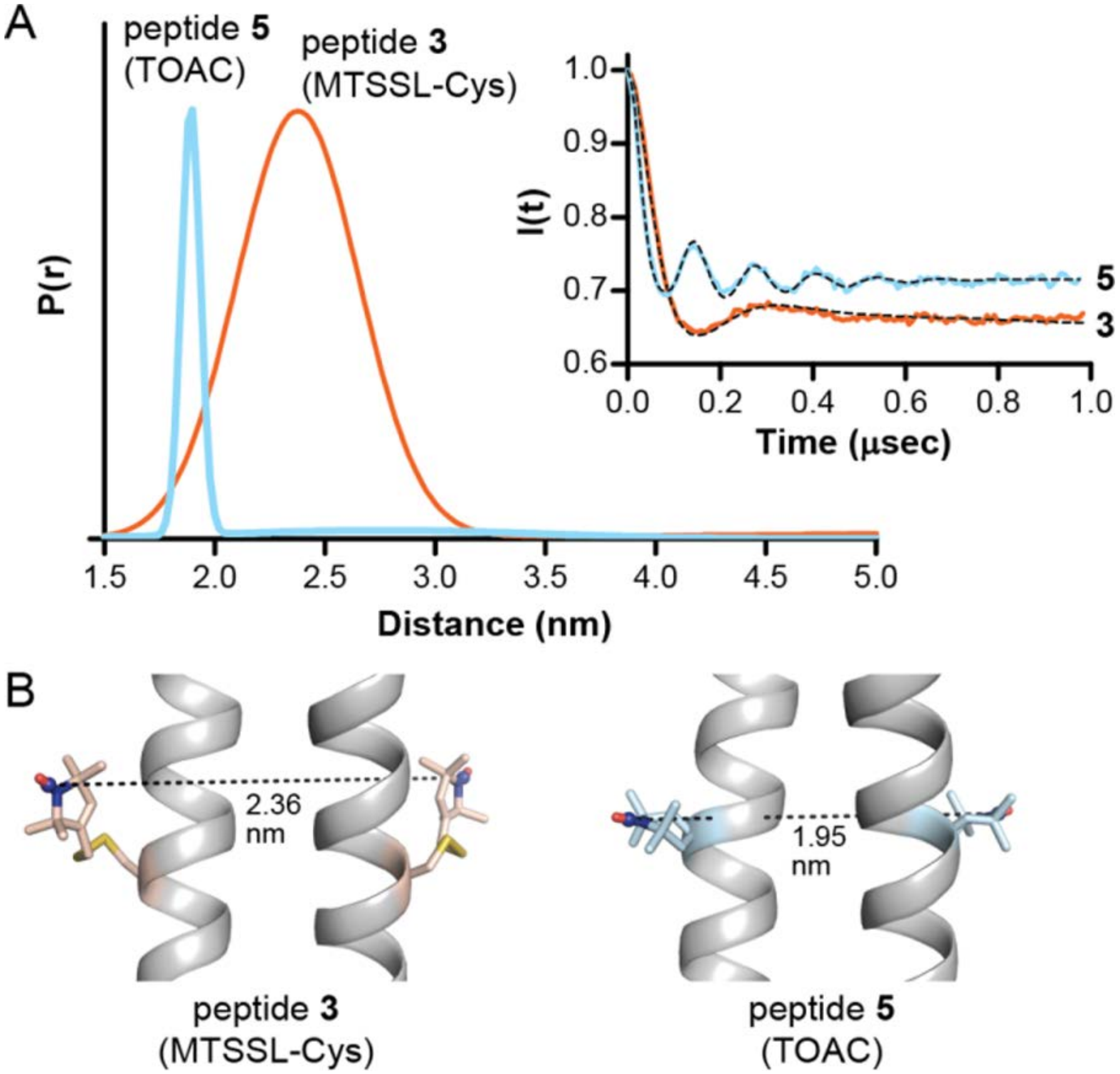


Figure 28. DEER data and structural models for peptides **3** and **5**.

A) Time-domain signals (background subtracted) of peptides **3** and **5** fit with Gaussian distributions for the distance distributions. B) Models of the peptides were constructed by overlaying models of the spin-labels onto X-ray structures of **1** (MTSSL for **3**) and **4** (TOAC for **5**). Adapted from Ref. [158] through an open access *ACS AuthorChoice* License. Copyright 2014 American Chemical Society.

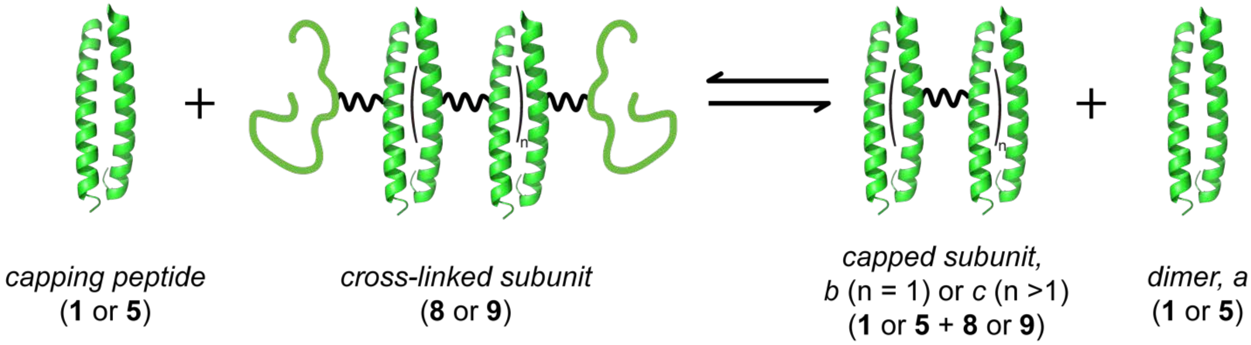
The model for the MTSSL-modified Cys containing coiled coil was built by overlaying the published coordinates of a MTSSL-modified Cys residue from a small helical protein[151] onto Ser_{14} from the X-ray structure of GCN4p1. This model gave a predicted nitroxide-nitroxide distance of

2.36 nm, well within uncertainty of the measured distance (most probable distance of 2.40 nm with a standard deviation of 0.26 nm). The large width of the probability distrbution from the MTSSL-labeled coil coil is caused by the five rotatable bonds connecting it to the backbone. Likewise, a model for peptide **5** was constructed by overlaying published coordinates of TOAC from a short α-helix[159] onto the backbone atoms of Ahx_{10} in the crystal structure of peptide **4**. The predicted nitroxide-nitroxide distance in this model was 1.95 nm, once again in agreement with experiment (most probable distance of 1.87 nm with a standard deviation of 0.07 nm). This narrow distribution is on the order of bond fluctuations. This is important because when the label is placed in the context of the supramolecular polymer any broader distance distributions can be solely attributed to the linker.

Both of the above results are consistent with prior work[160] that demonstrated TOAC gives rise to significantly narrower distance distributions than MTSSL. Even without spectral fitting and transformation to the distance domain, this is apparent in the raw time domain data. Four full periods of the spin relaxation decay are visible for TOAC, whereas only one period is visible with MTSSL. Dampening of the signal results when a larger number of distances are summed to produce the overall signal.[146] Encouraged by the observations that TOAC does not disturb the coiled coil fold and exhibits a very narrow distance distribution, it was chosen as the spin-label for measurements on the supramolecular polymer.

2.2 DESIGN OF A MODULAR DISCRETE SUBUNIT

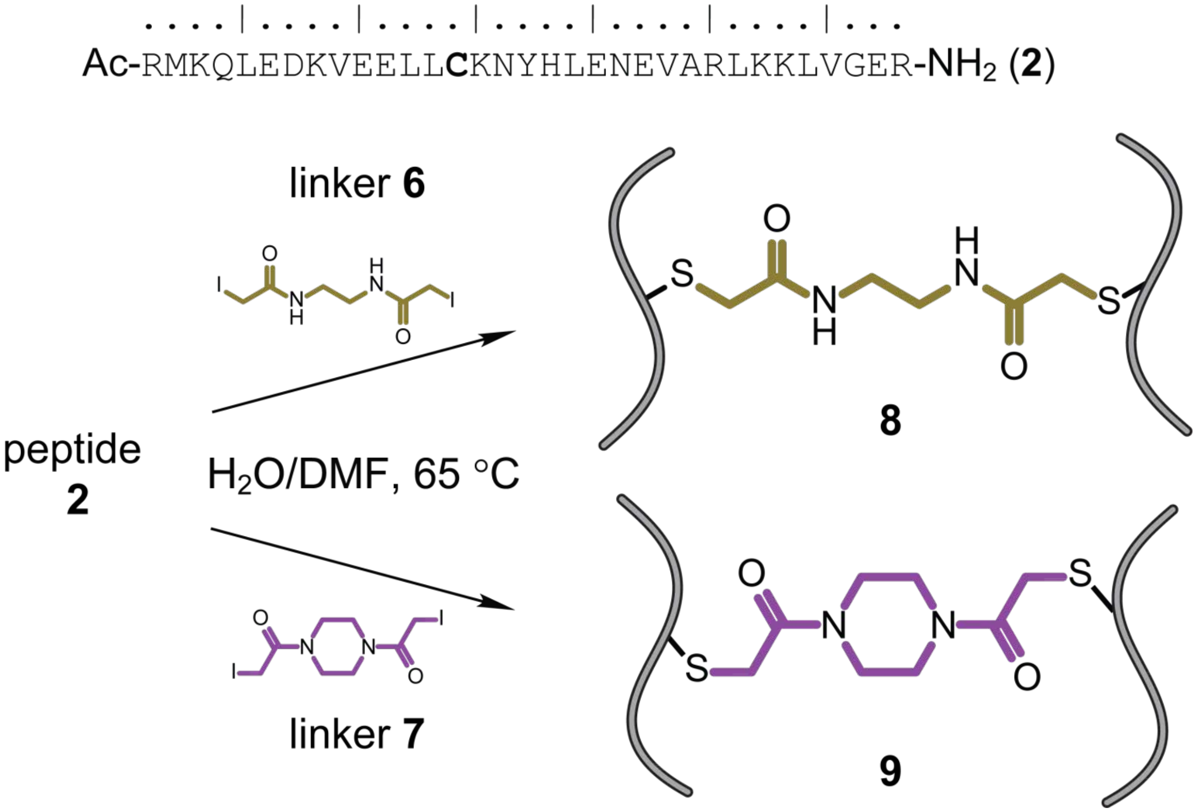
In order to measure the flexibility across a linker in our peptide-based SMPs, it is necessary to devise a system that consists of a single subunit flanked by two spin-labeled peptides. It had been previously observed that the size of the supramolecular assembly could be controlled by addition of a "capping" peptide.[62] This capping peptide consisted of the same primary sequence as the cross-linked subunit minus the linker. Since the intermolecular interactions remain unaltered, this sets up an equilibrium of a statistical distribution of supramolecular species composed of: a) coiled-coil dimers of the uncross-linked capping peptide, b) a single (n = 1) subunit capped by two uncross-linked peptides, and c) longer supramolecular oligomers composed of a variable number (n > 1) of cross-linked subunits terminated with the capping peptide (Scheme 1).



Scheme 1. Statistical supramolecular assembly created by mixing a capping peptide with a cross-linked subunit to generate a statistical mixture of a capped polymer with variable number of subunits and dimer of the capping peptide.

As the concentration of capping peptide increases, the length of species *c* decreases and the population of *a* and *b* increases. A 3:1 ratio of subunit : capping peptide was expected to favor the desired assembly *b* as the major component in solution based on previously reported GPC data.[62] This led us to hypothesize that spin-labeled peptide **5**, though its sequence is altered slightly by TOAC incorporation, could be used to produce the spin-labeled assembly needed for DEER anlaysis.

Cross-linked subunits of peptide **2** were synthesized by reaction of linker **6** (EDA) or **7** (PIP) in aqueous DMF to generate subunits **8** and **9**, respectively (Scheme 2).



Scheme 2. Synthesis of subunits **8** and **9** from peptide **2** and alkyl halide linkers **6** or **7**.

Mixtures of subunits **8** or **9** (100 µM) with TOAC-labeled peptide **5** (33.3 µM) were prepared in deuterated aqueous buffer at pH 7. Solvent deuteration was necessary to lengthen the dipolar relaxation time to enable measurement of longer distances.[146] DEER measurements were expected to give a bimodal distance distribution consisting of species *a* and *b* (the nitroxide-nitroxide distances in species *c* should be too long to be detected). Based on the linker length, it was expected that the distance between nitroxides in species *b* would be approximately 4-5 nm. A peak at 1.9 nm was also expected, as observed for peptide **5** alone.

Results of the DEER measurements for both systems were nearly identical to the measurement on peptide **5** alone (Figure 29).

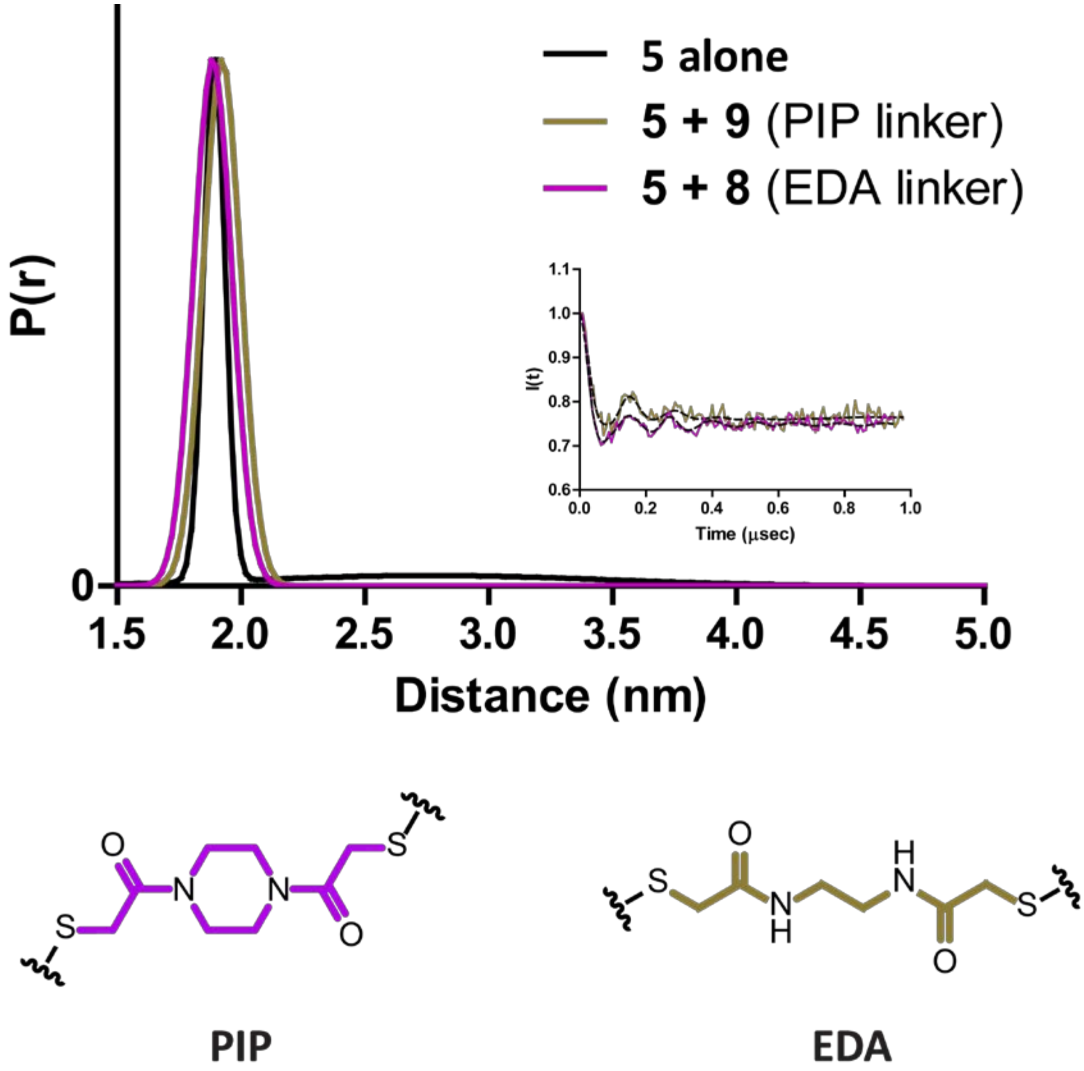


Figure 29. DEER distance distribution results of mixing peptide **5** with cross-linked subunits **8** (EDA, gold) and **9** (PIP, magenta) compared to **5** alone. Time domain data is given in the inset. Adapted from Ref. [158] through an open access *ACS AuthorChoice* License. Copyright 2014 American Chemical Society.

From the absence of any longer distances, it was inferred that the population of desired species *b* is negligible. The lower signal to noise of **5** + **8** and **5** + **9** vs. **5** in isolation implies that some of the nitroxides exist as species *c*, since these spins cannot contribute to the DEER signal. In retrospect, this can be explained by the greater T_m (71 °C) of peptide **5** versus peptide **1** (62 °C) (Figure 27). From this difference in T_m, it can be reasoned that the homodimer of **5** is more stable than the homodimer of **1** and that the heterodimer of the two lies somewhere between. The difference in folded stability of the homodimers versus the heterodimers is likely sufficient to skew the statistical oligomer distribution observed previously upon mixing subunits **8** and **9** with peptide **1**.[62] It was concluded that this strategy was not worth pursuing further and that a new system would have to be devised.

The failure of the first system to produce the desired species in sufficient quantities for observation highlighted the need to create a system where heterodimers dominate over homodimers. Fortunately, the design of heterodimeric coiled coils is well precedented in the literature (see Chapter 1). The sequences of peptides **10**, **11**, and **12** were based on a set of previously described *de novo* designed complementary peptides[104] designated ACIDp1 and BASEp1 (Figure 30).

. . . . | . . . . | . . . . | . . . . | . . . . | . . . . |

Ac–AQLKKKLQALKKKNAQLKWKLQALKKKLAQ–NH_2 **(BASEp1)**

ABA–AQLKKKLQALK**C**KNAQLK**A**KLQALKKKLAQ–NH_2 **(10)**

Ac–AQLEKELQALEKENAQLEWELQALEKELAQ–NH_2 **(ACIDp1)**

ABA–AQLEKELQ**Σ**LEKENAQLE**A**ELQALEKELAQ–NH_2 **(11)**

ABA–A**Σ**LEKELQALEKENAQLEAELQALEKELAQ–NH_2 **(12)**

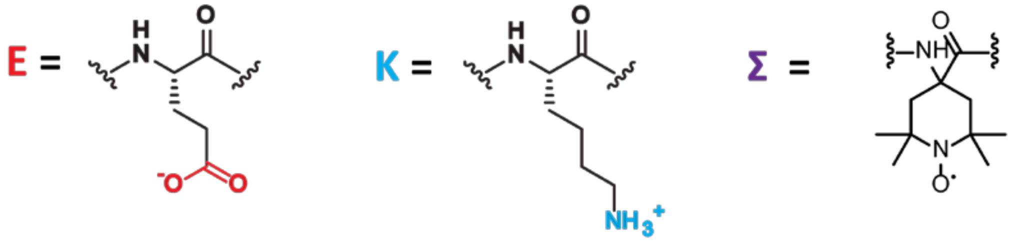


Figure 30. Sequences of Base-p1, Acid-p1 and peptides **10**, **11**, and **12**.

Basic residues are colored in blue, acidic residues in red and nitroxides in purple.

ACIDp1 and BASEp1 are designed such that they only fold as a heterodimer; alone they remain unstructured.

The BASEp1 sequence was modified with a point mutation of K_{12} → C generating peptide **10** with a point of attachment for the linkers. This site is equivalent to the point of attachment in the original system since they are both the *f* position in the second heptad of the helix. The ACIDp1 sequence was altered at positions 2 or 9 with TOAC to generate peptides **12** and **11**, respectively. These positions were chosen because they are at *c* positions in either the first (**12**) or second heptad (**11**). This allows the linker dynamics to be probed at two different positions: one near the point of linker attachment and the other at the N-terminus. Using two different spin-label sites will allow a better model of linker flexibility to be built. All three peptides were also capped at the N-terminus with a 4-acetamidobenzoic acid chromophore to aid in concentration determination.

CD measurements of the three peptides revealed a surprising observation. While the basic peptide (**10**) was poorly structured in isolation, both TOAC-labeled acidic peptides (**11** and **12**) exhibited a significant helical signature (Figure 31).

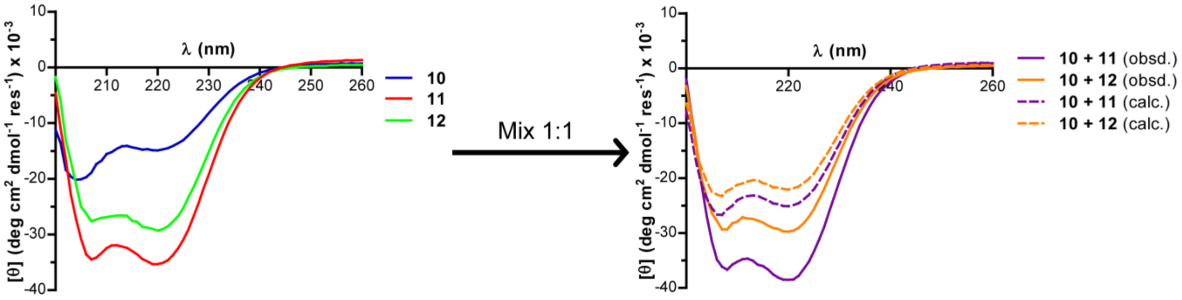
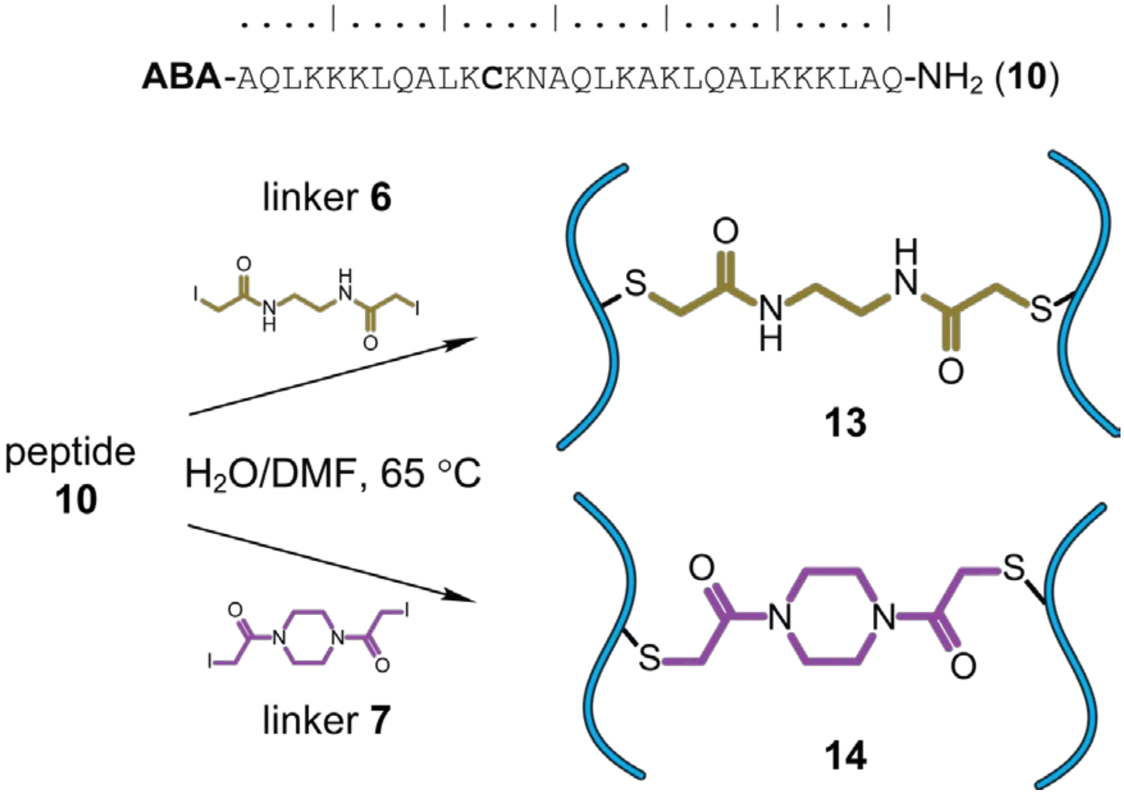


Figure 31. CD scans of peptides **10-12** (100 µM in 10 mM HEPES buffer, pH 7). Mixtures 1:1 of **10** with **11** or **12** (experimental = solid; calculated non-interacting spectra = dashed).

The presence of a significant helical structure for peptides **11** and **12** is surprising since the Glu residues at the *e* and *g* positions destabilizes homodimer formation through charge-charge repulsion,[104] and formation of monomeric helices are disfavored by the large hydrophobic face formed by Leu residues at the *a* and *d* positions[88] (*vide supra* Section 1.3). The observed helical signature of peptides **11** and **12** can be attributed to TOAC's strong preference for helical dihedral angles[152] overcoming the energetic penalty of charge-charge repulsion on the Glu residues allowing the coiled coil to form. Alternatively, these sequences may populate a monomeric helix fold.

Mixtures of the complementary components (**10** + **11** or **10** + **12**) in a 1:1 ratio led to a significant increase in helical signature compared to theoretical non-interacting mixtures (Figure 31). This implies that the heterodimer of **10** + **11** or **10** + **12** is a significantly more stable species in solution than the corresponding homodimers. From these results it can be inferred that mixtures of cross-linked peptides **13** and **14** and the TOAC-labeled peptides will form the desired capped subunits and disfavor the undesired homooligomers.

Subunits **13** and **14** were synthesized from the basic peptide **10** in the same manner as subunits **8** and **9**, respectively (Scheme 3).



Scheme 3. Synthesis of subunits **13** and **14** from peptide **10** and alkyl halide linkers **6** or **7**.

These two subunits were then combined with TOAC-labeled peptides **11** and **12** to form four binary samples. The CD results from the mixtures of peptide **11** and **12** with peptide **10** imply the predominant species in these samples will be the spin-label capped cross-linked subunit. These four mixtures were then subjected to DEER measurements for measurement of the distance distributions across the assembly.

2.3 MEASUREMENT OF LINKER FLEXIBILITY BY DEER SPECTROSCOPY

DEER distance distributions for the four samples differed based on linker identity and spin-label position. All data was fit using a unimodal or bimodal Gaussian distribution. Placement of the spin-label near the cross-linking site (subunits capped with peptide **11**) resulted in a unimodal distance distribution for both the PIP and EDA linker with a most probable distance of ~4.2 nm (Figure 32).

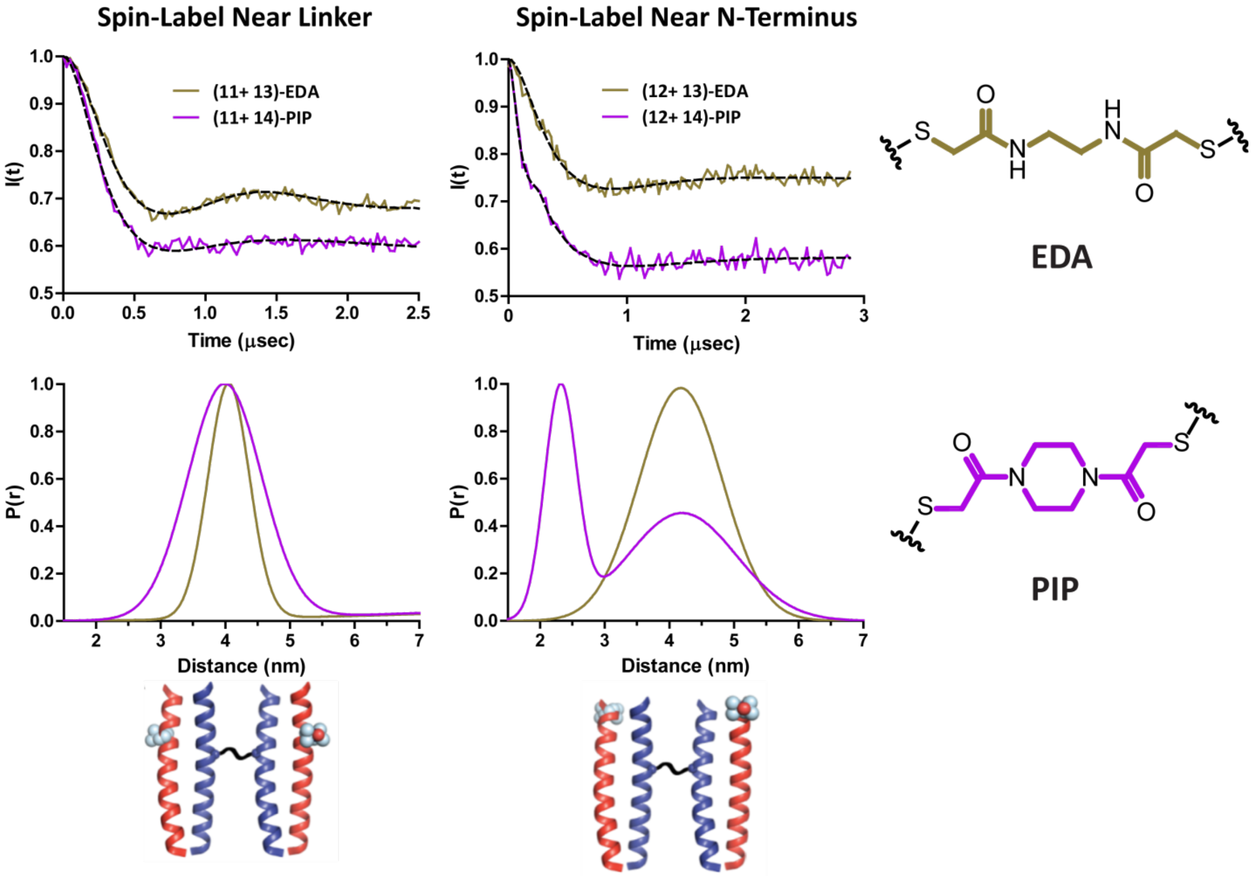


Figure 32. DEER results of the heterodimeric assemblies.

Cartoons of the structures are depicted below denoting the placement of the TOAC spin labels (spheres).

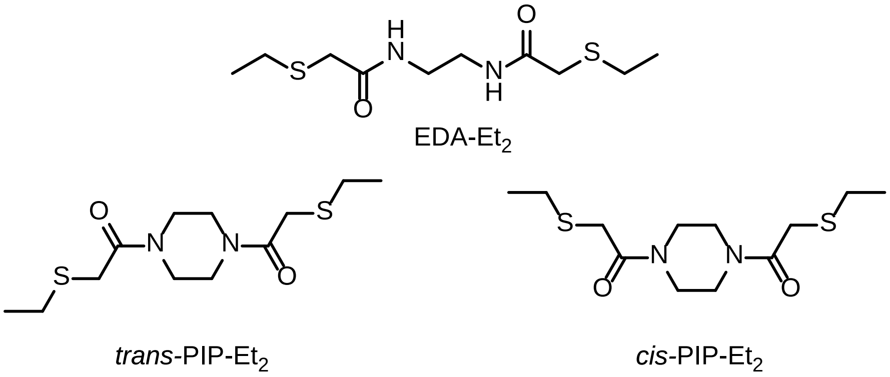
Although the most probable distance is similar for both linkers, the cyclic PIP linker exhibited a wider distance distribution (standard deviation of 0.8 nm for PIP versus 0.6 nm for EDA). Altering the placement of the spin-label to the N-terminus resulted in a unimodal and bimodal distance distribution for the EDA and PIP linkers, respectively (Figure 32). The most probable distance for the EDA linker was again ~4.2 nm; the bimodal distribution of the PIP linker contained most probable distances of ~2.3 nm and ~4.2 nm. The presence of the bimodal distribution can be seen in the time domain data by the presence of a second period. The presence of the second peak was also confirmed by fitting a model with peak suppression. An alternative model fitting procedure, Tikhanov regularization, suggests that a small population with a nitroxide-nitroxide distance of ~2.3 nm may also be possible for the EDA linker.

These results were puzzling since it was expected that the more flexible linker would give rise to a wider range of conformations and therefore a wider distance distribution. Another question posed by the DEER data was the origin of the second sub-population observed for the PIP linker. To gain a greater understanding of the possible conformations giving rise to these distance distributions, models of the subunits informed by the experimental data were built with the aid of molecular dynamics.

2.4 MOLECULAR DYNAMICS-AIDED MODELING OF THE SUPRAMOLECULAR POLYMER

Models of the subunit conformations were built through a combination of molecular dynamics (MD) simulations and static modeling informed by the experimental distance distributions. The modeling strategy was divided into three steps: 1) sample all possible conformations of the linkers using atomic-level molecular dynamics simulations; 2) append crystal structure-derived models of the coiled coils onto the linker conformations, excluding those with steric clashes; and 3) refine the conformational library using the experimental distance distributions. This simplified modeling protocol was chosen over more rigorous MD simulations of the full system due to the high computational demands of the latter. Due to the nature of this modeling algorithm, the resulting ensembles are not equilibrium-weighted, but rather represent the range of possible conformation space that each subunit is likely to adopt.

Conformational sampling of the linkers was performed using the EDA and PIP linker cores capped by a thioethyl moiety on each end as models (Figure 33).



EDA-Et_2

trans-PIP-Et_2

cis-PIP-Et_2

Figure 33. Structures of the molecules used to simulate the linker conformations by MD.

Both the *cis* and *trans* conformers of the PIP linker were simulated separately since tertiary amides do not isomerize quickly enough on the MD timescale. Each linker model was separately simulated by MD for 500 ns in explicit water to obtain the full range of conformations possible.

Coiled-coil models of peptide **1** were altered to include rotamers of Cys at position 14 of one chain of the dimer and TOAC residues at position 4 and 11 of the other chain. One dimer model was then appended to each side of the linkers on all possible conformations generating an ensemble of subunit conformations. A set of custom Python scripts (Appendix B) were used to discard any subunit models that contained steric clashes. The remaining library of conformations was then passed on to the last step.

Final refinement of the conformational library from step 2 removed any models inconsistent with the experimental DEER data. Another custom Python script removed any conformations where the nitroxide-nitroxide distance lay outside of one standard deviation of the most probable distance observed for both spin label positions by DEER. Two sets of models (one for each peak) were generated for the PIP linker due to the bimodal distribution for the N-terminal labeling site. This resulted in ~1000 models for the capped subunits of both the EDA and PIP linkers for the N-terminal peaks at ~4 nm and ~500 models for capped subunit of the PIP linker with a peak at ~ 2.3 nm.

These results provided structural insights into the nature of the differences between the two linkers. Visualization of the ensembles focused on the point of attachment to the next subunit that would be present in the fully assembled supramolecular polymer (Figure 34).

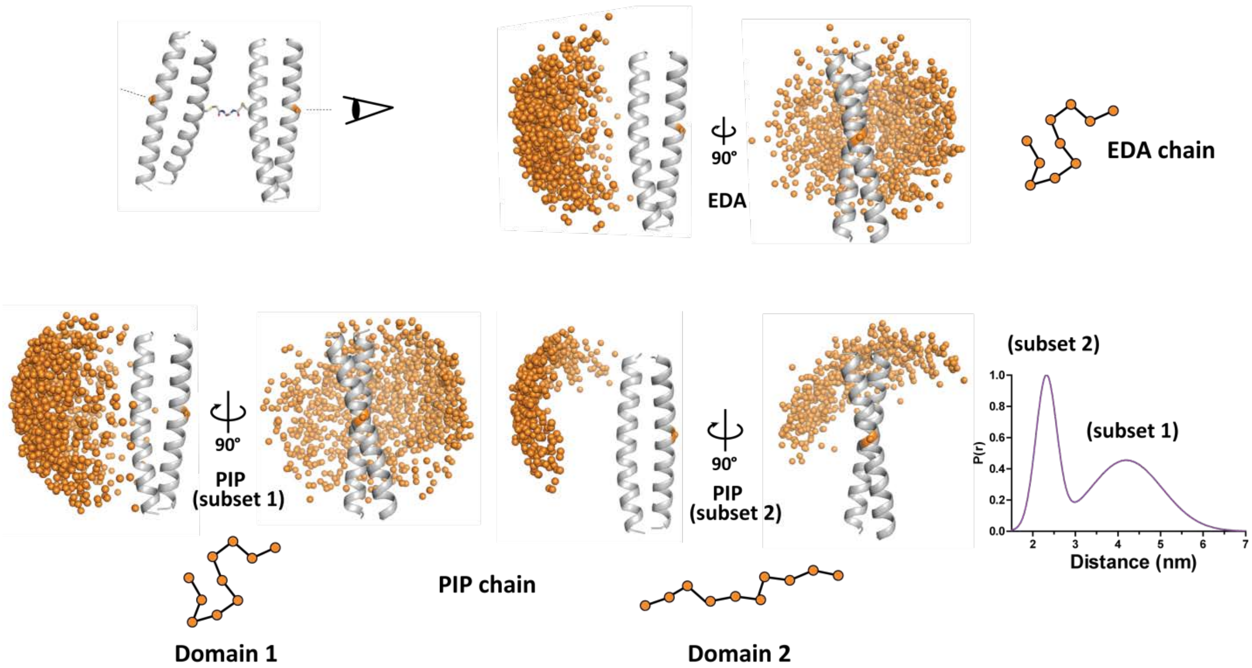


Figure 34. Visualization of the molecular dynamics-aided modeling.

Orange spheres represent the point of attachment to the next subunit propagating along the chain. Adapted from Ref. [158] through an open access *ACS AuthorChoice* License. Copyright 2014 American Chemical Society.

This is represented by displaying the C_α atoms of Ser_{14} as spheres in each subunit. The conformational ensemble generated by the EDA linker is similar to sub-population 1 (N-terminal distance centered around ~4 nm) of the PIP linker. Both ensembles display linker attachment sites over an approximately half-spherical shell. This implies a random orientation of the two coiled coils to one another in these ensembles. Sub-population 2 (N-terminal distance centered around ~2.3 nm) of the PIP linker is markedly different with an ensemble covering a much narrower arc of conformational space. This represents structures where the dimers are roughly parallel to one another across the linker.

It is hypothesized that the orientation of structures in subset 2 of the PIP ensemble gives rise to the macroscopic differences observed between SMPs of the two linkers. Propagation of the supramolecular assembly in roughly the same direction will result in a longer persistence length for the PIP linker thereby leading to a less compact polymer chain compared to the EDA linker. This longer persistence length ultimately gives rise to the larger apparent hydrodynamic diameter observed by GPC. This demonstrates how a small structural difference in linker rigidity between the cross-linked subunits of **8** and **9** can propagate along a supramolecular polymer chains.

2.5 CONCLUSIONS AND FUTURE DIRECTIONS

The work described in this chapter highlights the ability of DEER spectroscopy to probe supramolecular polymer structure. The structural insights gained by restraining MD models with experimental distance distributions were able to uncover the molecular cause of the apparent size difference between supramolecular polymer chains differing by only two CH_2 groups in an 8 kDa subunit. Restricting the conformational freedom of linkers between large macromolecular dimerization domains can be used to increase the apparent size of the final assembly without altering the association affinity of the non-covalent forces.[161] This provides a means of controlling the properties of supramolecular polymer based materials. Replacement of the covalently bound linkers with metal chelating groups is an alternative method to organic linkers to promoting coiled-coil SMP formation. Alteration of the peptide oligomerization state can also be used to control the structure of the supramolecular polymers. Both of these design strategies are explored in Chapter 4.

Follow-up work to this project could refine the structural models presented here. Recently developed techniques involving ensemble-biased metadynamics[162-163] may provide a way to achieve more robust structural models in a reasonable computational time. Metadynamics biases an MD simulation by applying a biasing potential forcing it to explore a wider range of conformations.[164] If this biasing potential is based on experimental restraints, a more accurate structural model can be obtained in a shorter computational time. Simulation techniques such as this combined with appropriate spectroscopic techniques will greatly improve our understanding of complex supramolecular assemblies.

2.6 EXPERIMENTAL

2.6.1 General Information

Solvents and all other reagents were purchased from Acros Organics, Aldrich, Advanced Chem Tech, Baker, EMD, Fluka, or Fisher and used without further purification. NovaPEG Rink Amide Resin, Fmoc-protected amino acids, HATU, PyBOP, and PyBrOP were purchased from Novabiochem. HCTU was purchased from Aapptec or Novabiochem. PyAOP was purchased from AK Scientific. 2,2,6,6-Tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid (TOAC) was purchased from Santa Cruz Biotechnology or Toronto Research Chemicals and (1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)methanethiosulfonate (MTSSL) was purchased from either Toronto Research Chemicals or Enzo Life Sciences. 1-Amino cyclohexanecarboxylic acid (Hex) was purchased from TCI America. Synthesis of Fmoc-Hex-OH was adapted from the literature.[165] The Fmoc-amino acid fluorides, Fmoc-Val-F and Fmoc-Gln-F, were prepared by a known method.[166] Crystallography reagents and tools were purchased from Hampton Research. Deuterated chloroform ($CDCl_3$), dimethyl sulfoxide (d6-DMSO) and glycerol (Glycerol-d3) were purchased from Cambridge Isotopes Laboratory. NMR spectra were recorded on a Bruker

Advance 300 or 400 spectrometer. MALDI of peptides and cross-linked subunits was collected on an AB Sciex Voyager DE Pro MALDI-TOF and small molecule high resolution mass spectrometry on a Thermo Scientific Q-Exactive Orbitrap.

2.6.2 Peptide Synthesis

Peptides were prepared by Fmoc solid-phase peptide synthesis (SPPS) manually with microwave-assisted reactions in a Microwave Assisted Reaction System (MARS, CEM), on a Tribute automated synthesizer (Protein Technologies), or by a combination of the two. All peptides were synthesized on a NovaPEG Rink Amide resin support. Standard coupling solutions consisted of 5 equivalents of Fmoc- protected amino acid with 4.9 equivalents of HCTU and 4.4% (v/v) *N*-methylmorpholine in DMF (automated) or 7.5 equivalents of DIEA in NMP (microwave). Coupling solutions were pre-activated for 2 minutes prior to addition to the resin and vortexed for 45 minutes (automated) or subjected to a 2 minute ramp to 70° C with a 4 minute hold (microwave). Fmoc deprotections were performed with 20% 4-methylpiperidine in DMF by two treatments of 4 minutes vortexing (automated) or one treatment with a 2 minute ramp to 80° C with a 2 minute hold (microwave). Resin was washed 3 times with DMF between steps.

The unnatural amino acids Ahx and TOAC were incorporated using alternative coupling strategies due to the steric hindrance of their amino groups. Fmoc-Ahx-OH was coupled with standard protocols except HCTU was replaced with PyBOP and the two subsequent residues were double-coupled. TOAC was coupled as the free amino acid with PyBOP in place of HCTU in the microwave with double the reaction time. No polymerization of the unprotected TOAC residue was observed, presumably due the high degree of steric hindrance of the N-terminal amine. The residue immediately following TOAC was activated as the acid fluoride[156] and coupled two to three times in the microwave for extended coupling times to ensure complete acylation. The N-terminus of each peptide was either acetylated with a solution of DMF/DIEA/Ac_2O (8:2:1 by volume) for 20 minutes at room temperature (peptides **1**, **2**, **4**, and **5**) or capped with 4-acetamidobenzoic acid using standard microwave coupling and PyBrOP as the activating agent instead of HCTU (peptides **10-12**).

Peptides were cleaved from the resin using various cocktails of TFA with scavengers and reaction times from 3-6 hours. Thiol scavengers were absent from cleavage cocktails of peptides containing nitroxides due to avoid an irreversible side reaction.[150] Cocktail mixtures were as follows (all reagents by volume): TFA/H_2O/EDT/TIS (92/3/3/1) for peptides **1**, **2**, and **4**; TFA/H_2O/anisole/TIS (91/3/3/3) for peptide **5**; TFA/H_2O/EDT/anisole/TIS (90/3/3/3/1) for peptide **10**; and TFA/H_2O/anisole/TIS (85/5/5/5) for peptides **11** and **12**.

Deprotected peptides were precipitated from the cleavage mixtures by addition of cold diethyl ether, centrifuged and decanted to produce a crude solid. Purification was performed by reverse-phase HPLC with a preparative grade C_{18} column (Phenomenex, 150 Å pore size, 10 µm particle size) with gradients between water and acetonitrile with 0.1% TFA (peptides **11-12**) or 1% TFA (peptide **10**). Nitroxide containing peptides (**5**, **11**, and **12**) were treated with 10% ammonium hydroxide prior to purification to regenerate the radical.[150, 157] Peptides **11** and **12** were further purified by ion exchange chromatography on a Mono Q 5/50 GL column (GE Healthcare) with a gradient between 20 mM and 0.5 M ammonium formate, pH 7.7. Peptide **3** (MTSSL labeled) was prepared by disulfide formation of the Cys residue in peptide **2** with MTSSL. To an ice cold solution of 50 µM peptide **2** in 50 mM phosphate, 300 mM NaCl, pH 6.55, MTSSL was added at a final concentration of 500 µM (10 equivalents). The reaction proceeded for 20 minutes and was then purified by preparative HPLC. Purity and identity were confirmed with analytical HPLC (Figure 35) and MALDI-MS (Table 1), respectively.

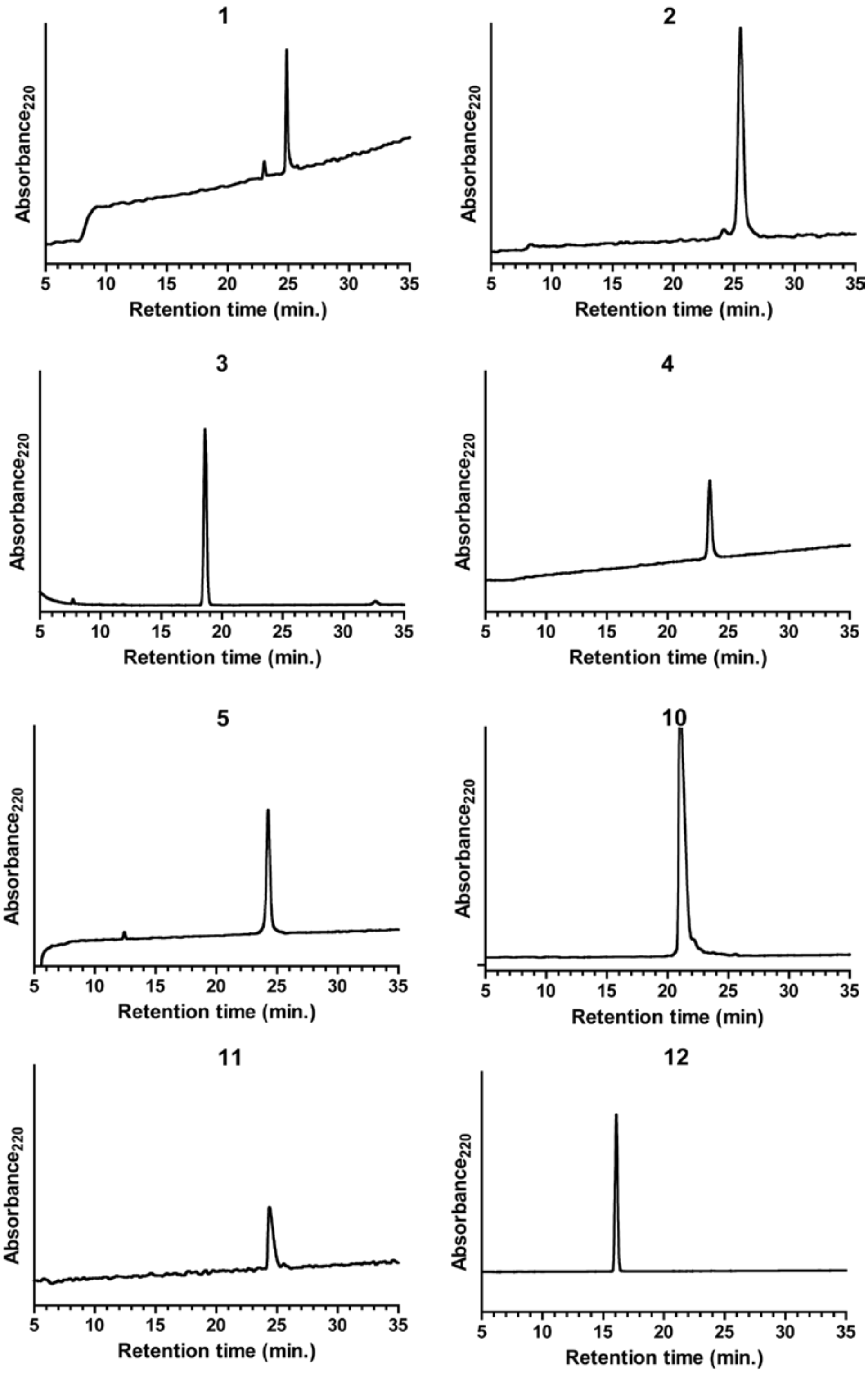


Figure 35. Analytical HPLC chromatograms of purified peptides **1-5** and **10-12**.

Table 1. Calculated and observed average masses of proteins **1-5** and **10-12**.

#	$[M+H]^+$ *m/z* (average)	
	Calculated	**Observed**
1	4038.7	4037.3
2	4054.8	4054.6
3	4239.1	4236.5
4	4034.8	4033.1
5	4224.9	4223.5
10	3534.2	3535.2
11	3696.1	3695.7
12	3641.1	3638.5

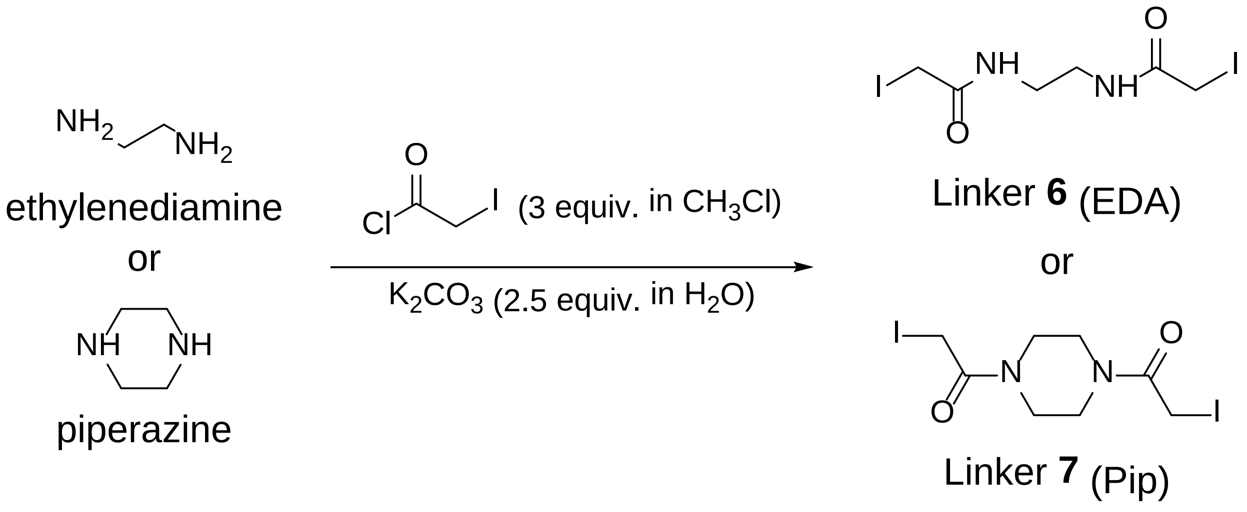
Peptide stock concentrations were determined using UV-Vis spectroscopy using the extinction coefficients listed in Table 2.

Table 2. Chromophores present in peptides and subunits **1-5** and **8-14** with their overall extinction coefficient.

Peptide	*Chromophore(s)*	ε *(M^{-1} cm^{-1}) at the indicated wavelength*
1	Tyr	1,450 at 276 nm
2	Tyr	1,450 at 276 nm
3	Tyr, MTSSL	1,920 at 276 nm
4	Tyr	1,450 at 276 nm
5	Tyr, TOAC, ABA	20,451 at 270 nm
8/9	2x Tyr	2,900 at 276 nm
10	ABA	18,069 at 270 nm
11/12	ABA, TOAC	18,961 at 270 nm
13/14	2x ABA	36,138 at 270 nm

[a] Extinction coefficient contributions for Tyr and ABA were obtained from the literature,[167-168] while those for TOAC and MTSSL-modified Cys were determined by UV measurements on TEMPO and MTSSL, respectively.

2.6.3 Linker Synthesis



Scheme 4. Synthesis of linkers **6** and **7**.

Linker **6** (EDA) was prepared by Dr. Kaylyn Oshaben and synthesized as described;[62] linker **7** (PIP) in a similar manner (Scheme 4). To a stirred solution of piperazine (421 mg, 4.9 mmol, 1 equiv) in 4 mL of chloroform at 0 °C were simultaneously added via addition funnels 1.05 mL of iodoacetyl chloride (14.7 mmol, 3 equiv) in 5 mL of chloroform and 1.68 g of K_2CO_3 (12.25 mmol, 2.5 equiv) in 5 mL of water. The solution was allowed to warm to room temperature and stirred 2 h. The organic layer was isolated by extraction, concentrated, and purified using column chromatography (50% ethyl acetate in acetone). Fractions containing the product were concentrated, re-dissolved in chloroform, and filtered. The filtrate was concentrated and dried under vacuum to afford the product as a pale yellow solid (451 mg, 1.07 mmol, 22% yield). **^{1}H NMR** (400 MHz, DMSO-d_6): δ 3.92 (d, J = 11.6 Hz, 4H), 3.47 (m, 8H). **^{13}C NMR** (100 MHz, DMSO-d_6): δ 166.7, 166.5, 46.0, 45.8, 41.2, 41.1, −1.2, and −1.4. **HRMS** (ESI) m/z calculated for $C_8H_{12}I_2N_2O_2$ $[M + H]^+$: 422.9067; found 422.9078.

2.6.4 Peptide Linker Conjugation

Fresh stock solutions of 2 mM linker **6** or **7** in DMF were prepared fresh for each batch. A 100 μM solution of peptide **2** or **10** in 25 mM phosphate buffer, pH 7 was heated at 60-70 °C. A 0.1 equivalent aliquot of linker **6** or **7** was added every 15 minutes until 0.5 equivalents of linker was reached. The reaction continued for another hour after addition of the final aliquot and was quenched with a solution of 0.1% TFA in 1:1 water/acetonitrile. The product was concentrated by centrifugation through a 3 kDa cutoff filter, washed with ~10 mL of water and re-concentrated to a final volume of ~2 mL. Subunits were purified in the same manner as their starting material peptides. Identity and purity were confirmed by analytical HPLC (Figure 36) and MALDI-MS (Table 3).

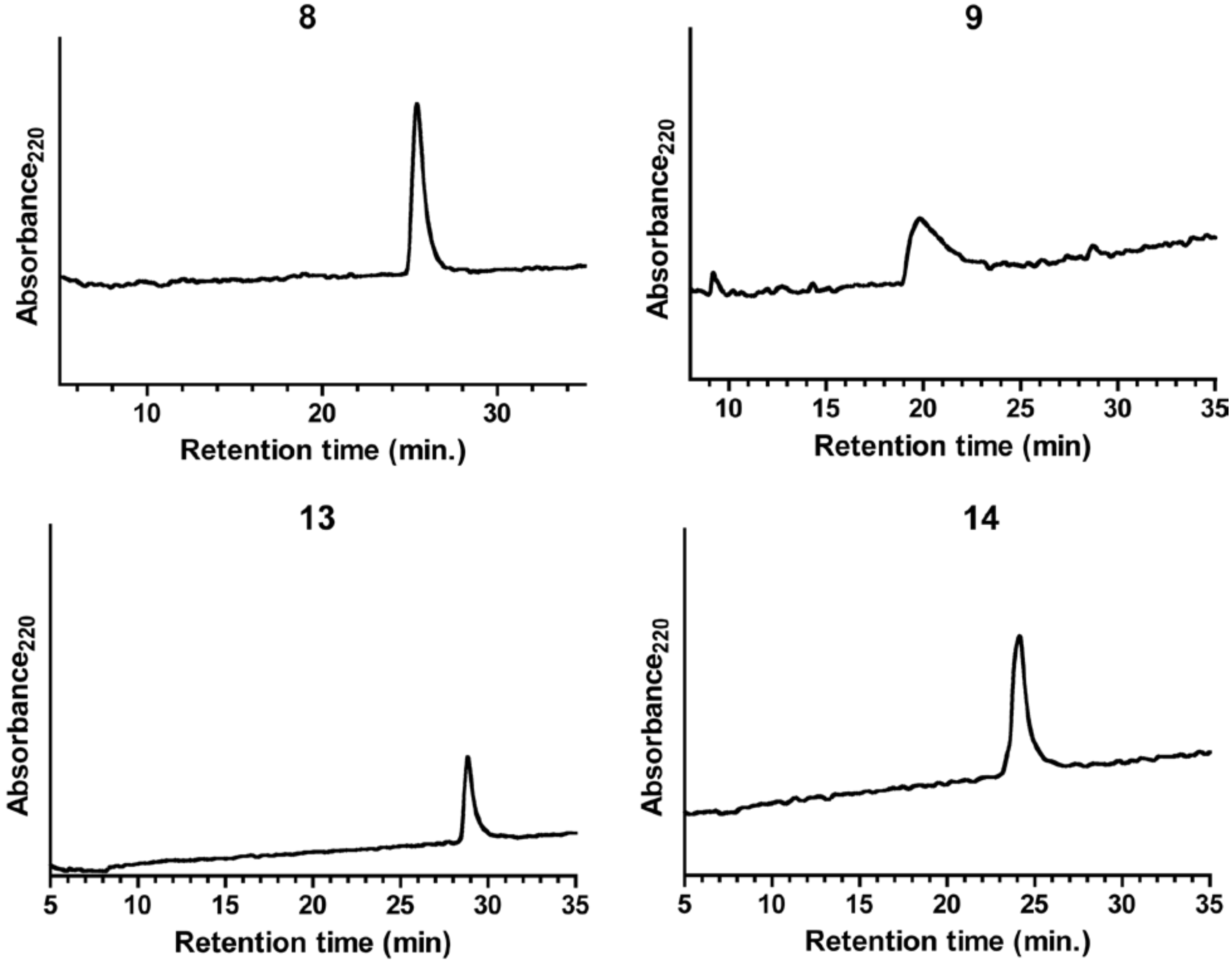


Figure 36. Analytical HPLC chromatograms of purified subunits **8, 9, 13,** and **14**.

Table 3. Calculated and observed average masses of subunits **8, 9, 13,** and **14**.

#	**$[M+H]^+$ *m/z* (average)**	
	Calculated	**Observed**
8	8248.7	8250.3
9	8274.7	8275.4
13	7207.6	7208.8
14	7233.6	7232.6

Stock concentrations of the subunits were determined by UV-Vis spectroscopy using extinction coefficients listed in Table 2.

2.6.5 Circular Dichroism

Circular dichroism (CD) spectroscopy measurements were carried out in 1 mm quartz cuvettes on an Olis DSM 17 CD spectrometer. Peptide stock concentrations were determined by UV-Vis spectroscopy. Buffer components were syringe filtered through 0.22 µm pore filter prior to sample preparation. All measurements were made with a 2 nm bandwidth and 5 second integration time. CD scans were measured from 200-260 nm at 20 °C. Thermal melts were measured at 222 nm and a baseline of 260 nm. Samples were incubated at each temperature for 2 minutes. Thermal melt data was fit to a two-state unfolding model[169] using GraphPad Prism.

2.6.6 Crystallography

Crystallization trials of peptide **4** and **5** were set-up using hanging drop vapor diffusion. Crystals were only obtained with **4**. Peptide stock solution was prepared from lyophilized powder at a concentration of 10 mg/mL, mixed (0.7 µL/ 0.7 µL) with crystallization buffer (0.2 M citrate tribasic, 0.1 M sodium cacodylate pH 6.5, and 30% (v/v) isopropanol) over a well of 0.7 mL of crystallization buffer, and equilibrated at room temperature overnight. A single crystal of peptide **4** was harvested from the crystallization drop, cryoprotected with 30% (v/v) glycerol in crystallization buffer and flash frozen in liquid nitrogen. X-ray diffraction data was collected on a Rigaku/MSCdiffractometer (FR-E generator, VariMax optics, Raxis HTC image plate detector) with Cu K_α radiation. The sample was kept frozen by an X-Stream 2000 low temperature system at 100 K.

Data collection and refinement was performed by Dr. W. Seth Horne. Raw diffraction data was processed with d*TREK. The peptide structure was solved by molecular replacement using

PDB 4DMD as a search model with Phaser.[170] Model refinement was performed with Phenix[171] in conjuction with manual model building in Coot.[172] Data collection and refinement statistics are given below in Table 4.

Table 4. X-ray diffraction data collection and refinement statistics for peptide **4**.

Data Collection	
Unit cell dimensions (Å, °)	a = 83.3, b = 30.4, c = 28.0 $\alpha = \gamma = 90$, $\beta = 102$
Space group	C2
Resolution (Å)	20.48–1.80 (1.86–1.80)
Total observations	19,070
Unique observations	6,386
Redundancy	3.0 (3.0)
Completeness (%)	98.1 (97.1)
I/σ	12.9 (3.0)
R_{merge} (%)	5.5 (17.6)
Refinement	
Resolution (Å)	20.48–1.80
R (%)	18.7
R_{free} (%)	22.8
Avg. B factor (Å^2)	33.4
RMSD	
Bonds (Å)	0.011
Angles (°)	1.1

2.6.7 Double Electron-Electron Resonance (DEER) Spectroscopy Measurements

Samples for DEER measurements of peptides **5** (GCN4p1-TOAC) and **3** (GCN4p1-MTSSL) were prepared in 10 mM HEPES buffer with 20% (v/v) glycerol as a cryoprotectant. Samples of all other peptides were prepared with deuterated solvents (D_2O in place of water and glycerol-*d8* in place of glycerol).

All DEER measurements were performed and processed by Dr. K. Ishara Silva (Saxena Lab, University of Pittsburgh). DEER experiments were performed on a Bruker ElexSysE580 X-band FT/CW spectrometer equipped with a Bruker ER4118-MD5 resonator. Sample temperature was regulated using an Oxford ITC503 temperature controller and an Oxford CLT650 low-loss transfer tube. For each measurement, ~150 μL of the appropriate sample was transferred into a 3 mm inner diameter quartz tube, flash frozen in liquefied MAPP gas, and inserted into a sample cavity pre-cooled to 80 K. The four-pulse DEER experiments were carried out using a pulse sequence of $(\pi/2)\nu_1$-τ_1-$(\pi)\nu_1$-T-$(\pi)\nu_2$-τ_2-$(\pi)\nu_1$-τ_2-echo.[173] The pump frequency ν_2 was set at the maximum of the nitroxide spectrum. The observer frequency ν_1 was offset by ~70 MHz. The length of the $(\pi/2)\nu_1$ and $(\pi)\nu_1$ pulses were 16 ns and 32 ns respectively. The $(\pi)\nu_2$ pulse was set to 16 ns. Step sizes of 8 ns and 16 ns were used to measure shorter and longer distances, respectively for 128 data points. Deuterated solvent and glycerol were used to increase the phase memory time of samples that were expected to give longer distances. Raw DEER data were analyzed using the DEERAnalysis2013[174] software through MATLAB. Data were fit to a Gaussian or double-Gaussian model and the corresponding distance distribution was generated.

2.6.8 Molecular Dynamics Aided Modelling

MD simulations were performed using the GROMACS 4.6.3 software package[175] and the AMBER99sb-ildn forcefield.[176] Amber-compatible force field parameters for the EDA and PIP linkers capped with thioethyl groups (EDA-Et_2 and PIP-Et_2) were prepared using Antechamber[177] and acpype.[178] Three separate simulations of the linkers were performed: one for EDA-Et_2, one for PIP-Et_2 with the two amide carbonyls pointing in the same direction (*cis*-PIP-Et_2), and one with the two amide carbonyls pointing in opposite directions (*trans*-PIP-Et_2). The two different starting conformers of PIP were run separately because tertiary amide isomerization is not effectively sampled on the timescale of the simulations. Each system was first energy minimized for 500 ps, equilibrated at 298K in the NVT ensemble for 100 ps, then equilibrated at 298K and 1 atm in the NPT ensemble for 100 ps. Simulations were run for 500 ns in explicit water with the TIP3P model[179] at 298K and 1 atm. A Langevin thermostat was used for temperature control and a Berendsen barostat for pressure control.

We prepared coiled-coil models to append to the two ends of the above set of linker conformers from the published structure of residues 1-30 in the dimer formed by peptide **1** (PDB 4DMD). Coordinates for TOAC in an α-helix[159] were modified to include a virtual atom at the midpoint of the nitroxide N–O bond and incorporated in place of residues 4 and 11 in one chain of the dimer. Residue 14 in the other chain was mutated to Cys, with the side chain rotamer set as one of the two most probable. Residues near the newly introduced Cys were changed to Ala. All side chains other than those listed above were removed. The above procedure generated two models (A and B), each a dimeric coiled coil doubly labeled with TOAC on one helix. The models differ only in the rotamer of the Cys that will be the point of attachment to the linker.

Capped subunit models were generated using a set of custom-written PyMOL scripts. Coiled coil models A and B above were first combined with the three different linker conformational ensembles (10,000 frames each) to generate nine sets of capped subunit structures: EDA-A_2, EDA-AB, EDA-B_2, *cis*-PIP-A_2, *cis*-PIP-AB, *cis*-PIP-B_2, *trans*-PIP-A_2, *trans*-PIP-AB, and *trans*-PIP-B_2. Thioethyl groups at each end of the linker were replaced with a coiled coil by overlay of Cys C_α, C_β, and S_γ atoms in the peptide with the corresponding atoms in the linker. The models were combined based on core linker structure to generate three capped subunit ensembles (EDA, *cis*-PIP, *trans*-PIP). The three capped subunit ensembles were curated to remove structures with steric clashes between helices or a helix and the linker. Steric clashes were defined as peptide backbone C_α atoms coming within 5 Å of one another or van der Waal overlap[180] involving any pair of atoms.

The resulting ensembles were then filtered further to identify structures consistent with the observed DEER data. A model was accepted only if TOAC-TOAC spin-spin distances for both labeling sites fell within one standard deviation of the most probable distance observed in the DEER experiment. Because the N-terminal labeling site for the PIP linker gave a bimodal distribution in the experiment, capped subunit models based on PIP were divided into two sets (subset 1 and subset 2). Conformational ensembles for *cis*-PIP and *trans*-PIP linkers were combined to generate the three conformational ensembles shown in Figure 9: EDA, PIP (subset 1), and PIP (subset 2).

3.0 IMPROVING HELIX BACKBONE MODIFICATION STRATEGIES

The work in this chapter has been previously published as:

> Tavenor, N. A.; Reinert, Z. E.; Lengyel, G. A.; Griffith, B. D.; Horne, W. S. "Comparison of Design Strategies for α-Helix Backbone Modification in a Protein Tertiary Fold." *Chem. Commun.* **2016**, 52, 3789-3792.

Fmoc-β^2-Asn(Dcmp)-OH was synthesized by Dr. Zach E. Reinert and Brian D. Griffith.[138] Fmoc-β^2-Lys(Boc)-OH and Fmoc-β^2-Ala-OH were synthesized by Dr. George A. Lengyel. Proteins **15-20** were synthesized by Dr. Zach E. Reinert. Plasmids and protocols for the expression of GB1 proteins **WT** and **K31A** were generously provided by Dr. Timothy Cunningham and Dr. Sunil K. Saxena (University of Pittsburgh). Expression, purification and characterization of **WT** and **K31A** were performed by Dr. Zach E. Reinert.

Proteins can be thought of as information-rich polymers. These polymers can be described by two sequences: 1) the canonical sequence of amino acid side-chains known to biochemists and 2) the sequence of backbone units displaying those side chains (in natural proteins an oligo-L-α-peptide).[116] Modification of the chemical connectivity of the peptide backbone can improve stability of polypeptides to enzymatic degradation for both structured[181] and unstructured[120] oligomers with the end application being longer lasting peptide therapeutics.[182] Prior efforts to engineer sequences with isolated backbone modifications have largely focused on mimics of α-helix[183-185] and β-sheet[186-189] secondary structures.[116] Prior work in our lab replaced ~20% of the natural backbone in a protein tertiary fold with unnatural amino acids. This resulted in a heterogeneous-backbone that mimicked the native tertiary fold but with significant thermodynamic destabilization of the folded state.[142]

Replacement of individual secondary structural elements within the tertiary fold only incurred minor thermodynamic penalties. Even though the penalty from each modification may be small (~0.5-1.0 kcal/mol), these penalties are additive resulting in a significant overall destabilization when combined.[142] There is therefore a need for backbone-modification strategies that retain or improve upon the stability of the native sequence. One of the secondary structures where the most significant improvement was needed was in the helix.

Inspired by the observation that incorporation of TOAC into coiled coils significantly improved thermal stability over the native sequences (see Chapter 2.2) and the known helix-forming propensity of C_α-methylated residues,[134, 144] we endeavored to determine whether C_α-methylation would have a similar effect in the context of a tertiary fold. In order to place this work into the context of the broader field on α-helix mimicry, the results were compared to design strategies using various regioisomers (both acyclic and cyclic) of β amino acids, some of which were published previously[131] (Figure 37).

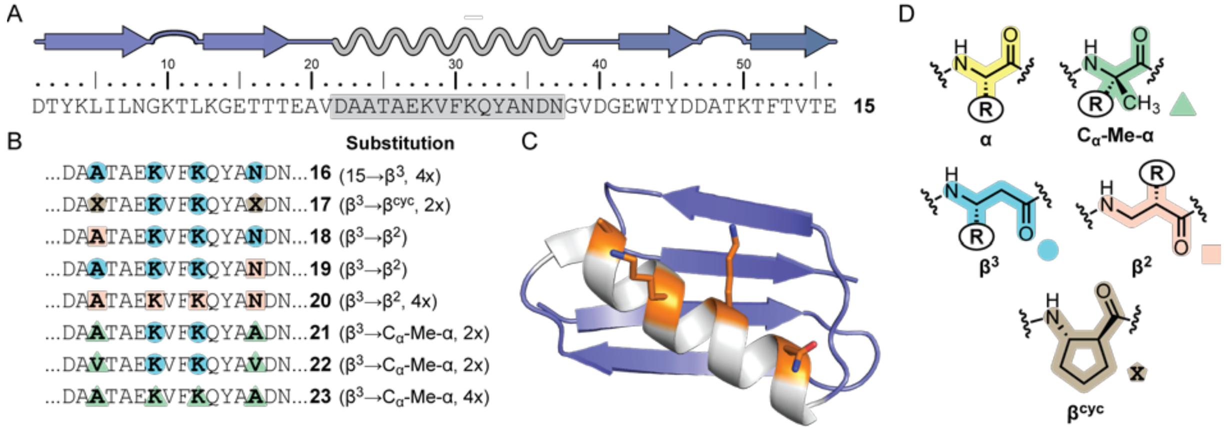


Figure 37. Backbone modification strategies of a helix in the context of a tertiary fold.

A) Primary sequence and secondary structure of Streptococcal protein GB1 (**16**); host protein for helix (gray) backbone modification. B) Helix variants of protein **16**; sequence modifications indicated with the type and quantity of backbone modifications. C) Crystal structure of the GB1 tertiary fold (PDB 2QMT) highlighting the helix (gray) and substituted positions (orange); crystal structure only differs from host sequence at the N-terminus (MQ in crystal vs. DT in **16**). D) Structures of the unnatural amino acid building blocks used for backbone modification with the R group in the building block corresponding to its natural α-amino acid counterpart; residue colors correspond to those highlighted in B. Adapted from Ref. [138] with permission from the Royal Society of Chemistry.

Full thermodynamic analyses and high resolution structural characterization were used to understand the effect that each amino acid class has on the final folded structure. Thermodynamic parameters for cooperative protein folding can be measured by tandem circular dichroism (CD) thermal and chemical denaturant melts.[131, 190] The magnitude of the molar ellipticity is proportional to the percent of protein folded, thus fitting the data in three dimensions (molar ellipticity, temperature, and denaturant concentration) to a multivariate equation allows the extraction of all thermodynamic parameters (see Section 3.4.4).

3.1 BETA RESIDUES

Acyclic β-residues have been used extensively to mimic helices both on their own and in mixed α/β designs.[116, 126] The main advantage of using β-amino acids as a backbone replacement is that the native side-chain can be retained allowing for direct comparison on the impact of the backbone alteration with the native sequence. There are two possible regioisomers for mono-substituted β-residues: β^2 (side-chain adjacent to carbonyl) and β^3 (side-chain adjacent to amide nitrogen). β^3-Residues are more widely utilized than β^2-residues due to their commercial availabilty.

Previous work[131] examining the thermodynamic consequences of incorporating β^3-residues into the helix of protein GB1 had determined that these substitutions are entropically stabilizing, but enthalpically destabilizing with an overall destabilizing effect on the folding free energy. These trends held true when the patterning of β-residues was altered only varying in magnitude. The GB1 mutant (protein **16**) with the best overall stability from that prior reported series[131] was used as a baseline for thermodynamic comparisons of all other proteins in this chapter.

The observed impacts of β^3-residues on the folding of GB1 were puzzling. These building blocks contain an extra rotatable bond compared to α-residues. Therefore, it was hypothesized that β-residues would be entropically destabilizing to the folded structure due to the increased degrees of freedom when moving from a folded to an unfolded state. However, this simplistic analysis ignores the role of medium in folding; proteins do not exist in a vacuum and water plays a large role in determining how they fold. It was suggested that the unexpected thermodynamic effects of β^3-residues were due to differences between solvent interactions in the unfolded versus folded states of the protein.[131] An open question from the past work was whether these trends were unique to β^3-residues or if the same held true for β^2-regioisomers (*vide infra*).

In the same study[131] the impact of cyclic β^3 residues based on aminocyclopentanecarboxylic acid (ACPC) were also examined. β^{cyc}-Residues limit energetically accessible backbone conformational space by incorporating an otherwise freely rotatable bond into a ring.[126] In the GB1 tertiary fold, $\beta^3 \rightarrow \beta^{cyc}$ substitution at positions 24 and 35 (protein **17**) was structurally well accommodated but led to only a modest increase in folded stability compared to protein **16**. It should be noted that cyclization of the β-residues resulted in the lowest degree of entropic destabilization observed for any mutant. However, it was also observed in another mutant that removal of the lysine side-chains significantly decreased the folded stability of the proteins.[131]

The importance of retaining the Lys side-chain at position 35 will be examined in greater detail in Section 3.1.1 (*vide infra*). These results imply that the role of side-chains in determining folded stability is context dependent. This motivated a desire to compare alternative design strategies that retain important side-chain contacts while approaching native protein stability.

3.1.1 Comparison of β^2 with β^3 Residues

Motivation for studying β^2-residues was two-fold: 1) computational analyses[143] suggested that β^2-residues were more thermodynamically favorable to helical folds than β^3-residues in heterogeneous-backbone α/β-peptide helices and 2) restoration of a potentially stabilizing orbital interaction at Asn_{35}[191] by shifting the side-chain from $\beta^3 \rightarrow \beta^2$. Proteins **18-20** (*vide infra*) were synthesized by SPPS (see Section 3.4.2) to test these hypotheses (Figure 37). Proteins **18** and **19** are single $\beta^3 \rightarrow \beta^2$ mutations at the outer β^3 residues: Ala_{24} and Asn_{35}, respectively. Protein **20** substitutes all four β^3 residues with β^2 analogues. Folding thermodynamics in each modified protein was probed by tandem thermal and chemical denaturation CD measurements.[131, 190] High resolution structures were determined by X-ray crystallography (structures of **18** and **19** were solved to 1.95 Å and 1.80 Å resolution, respectively).

Aside from the expected side-chain displacements, analogues **18** and **19** exhibit essentially identical tertiary folds as both natural backbone **15** and analogue **16** bearing an α/β^3 helix (Figure 38).

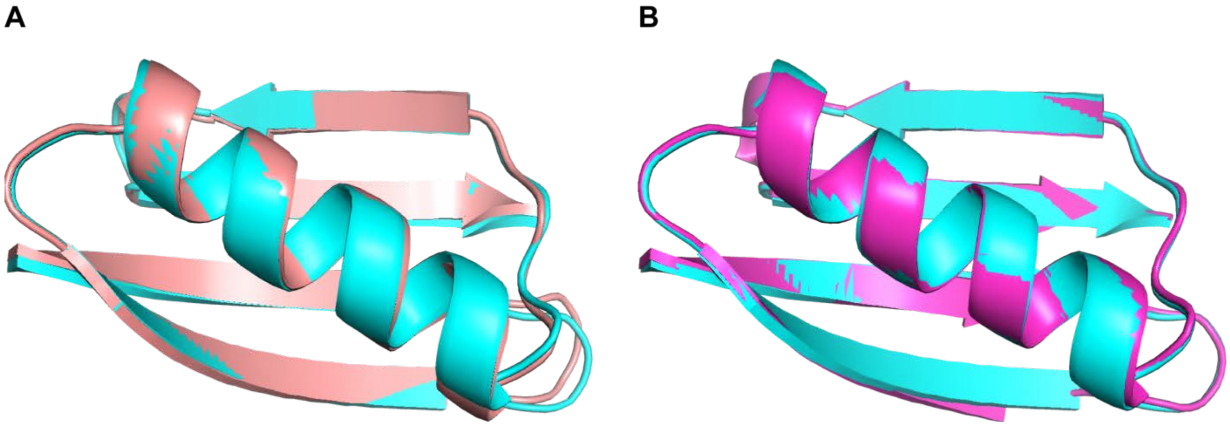


Figure 38. Global alignment of the folded structure of GB1 analogues bearing an α/β^3 helix **18** (A, pink) and **19** (B, magenta) with **16** (light blue).

Analysis of folding thermodynamics (Figure 39) revealed individual $\beta^3 \rightarrow \beta^2$ replacement was neutral (**18**) to slightly destabilizing (**19**). Differences between **18** and **19** were small and similar within experimental uncertainty. Both proteins have an enthalpically stabilized fold compared to **16** that is offset by an entropic penalty (Figure 39). This implies that the computational comparison[143] of the two residue types is likely more accurate when calculating enthalpic than entropic effects. Though the exact origin of the entropy/enthalpy compensation is not clear, changes in the sensitivity of the folded state to chemical denaturant (*m*) and the heat capacity (ΔC_p) difference between the folded and unfolded states (see Section 3.4.4, Table 6) suggest a more compact denatured ensemble in β^2-residue containing variants vs. β^3 counterparts.[131]

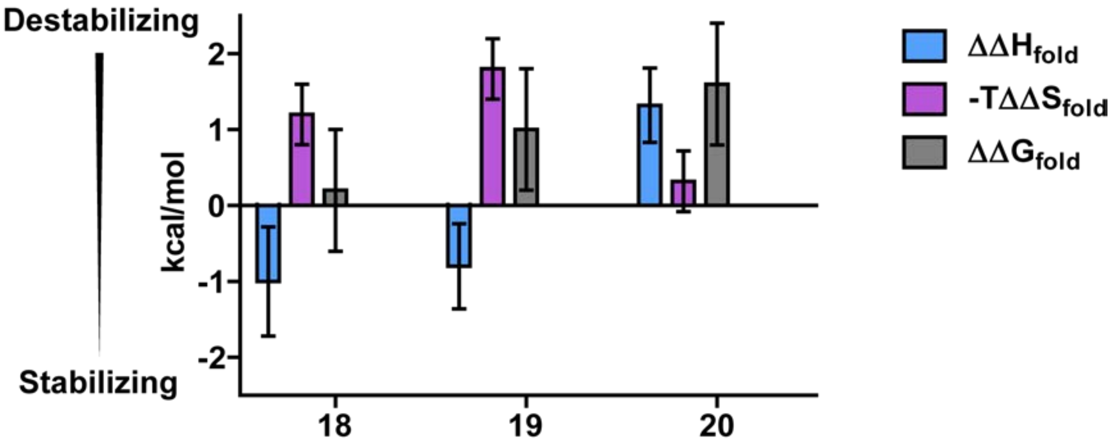


Figure 39. Effect of β^3 to β^2 substitution on the thermodynamics of folding: **18** ($\beta^3 \rightarrow \beta^2$-Ala$_{24}$), **19** ($\beta^3 \rightarrow \beta^2$-Asn$_{35}$), and **20** (all $\beta^3 \rightarrow$ all β^2). All proteins are referenced to the thermodynamic values for protein **16** (all four sites β^3).

Recent literature highlights the importance of local orbital interactions in protein folding thermodynamics.[191] One example is an intraresidue $n \rightarrow \pi^*$ overlap in Asn involving partial donation of a carboxamide C=**O** lone pair into an antibonding orbital from the backbone carbonyl **C**=O.[192] This interaction has been suggested to be worth up to 1.2 kcal mol^{-1} in folding enthalpy.[192] In the crystal structure of wild-type **15** (PDB 2QMT[193]), the α-Asn$_{35}$ is positioned for this orbital overlap to occur, but not in β^3-Asn$_{35}$ from **16**, presumably because the β^3-residue regiochemistry does not allow the side-chain to adopt the necessary orientation (Figure 40).

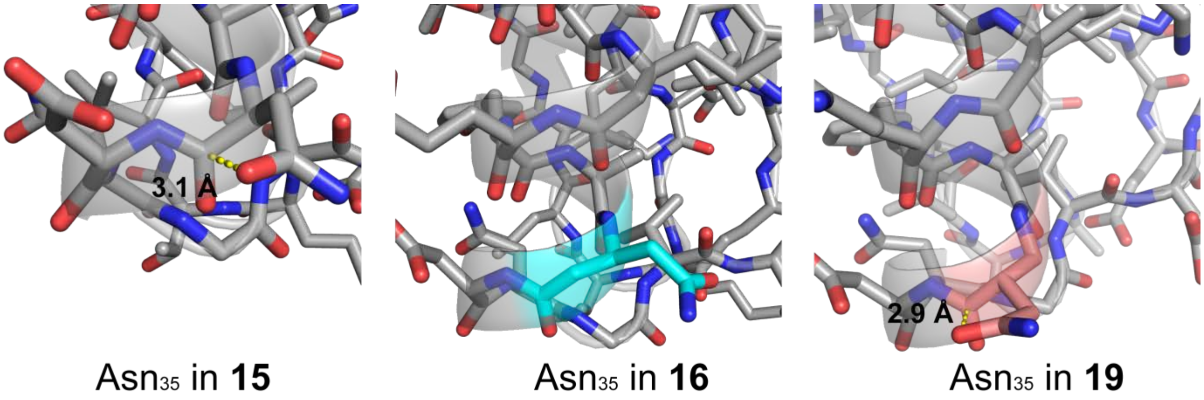


Figure 40. Comparison of crystal structures depicting the side-chain orientation of Asn$_{35}$ with α- (left, **15)**, β^3- (middle, **16**), and β^2- (right, **19)** backbones. Putative n→π* interactions are observed for the α- and β^2- residues, but not the β^3 residue.

In the crystal structure of **19** (PDB 5HG2), two of four chains in the asymmetric unit showed clear evidence for a side-chain to backbone contact at β^2-Asn$_{35}$ analogous to that observed for the α-residue. Although this interaction is observed in the crystal structure, β^3→ β^{2}Asn$_{35}$ substituion was found to be enthalpically equivalent to β^3→ β^2-Ala$_{24}$ substitution where such an interaction is not possible. This implies that while β^2-residues may be better at maintaining native *n*→π* interactions compared to β^3-residues, these contacts do not play a significant role in folded stability in this particular system.

Replacement of all four β^3-residues in **16** with β^2-residues (**20**) resulted in a significant destabilization in folded stability of ~1.6 kcal mol^{-1}, almost entirely enthalpic in origin. Approximately a third (0.6 kcal mol^{-1}) of this destabilization can be explained by β^3→ β^2-Ala$_{24}$ (~0.1 kcal mol^{-1}) and β^3→ β^2-Asn$_{35}$ (~0.5 kcal mol^{-1}) substitutions. The remainder of the destabilization (~1.0 kcal mol^{-1}) likely comes from disruption of important tertiary contacts between the helix and sheet of the protein. Within the wild-type **15** protein fold, Lys$_{31}$ has both a van der Waals interaction with Trp$_{48}$ and forms a salt bridge with Glu$_{27}$ (Figure 41A). Both of these interactions are maintained with β^3-Lys$_{31}$ (Figure 41B), but lost with β^2-Lys$_{31}$. Although crystallization attempts of **20** were unsuccessful, modeling the orientation of the β^2 side-chain supports this conclusion (Figure 41C).

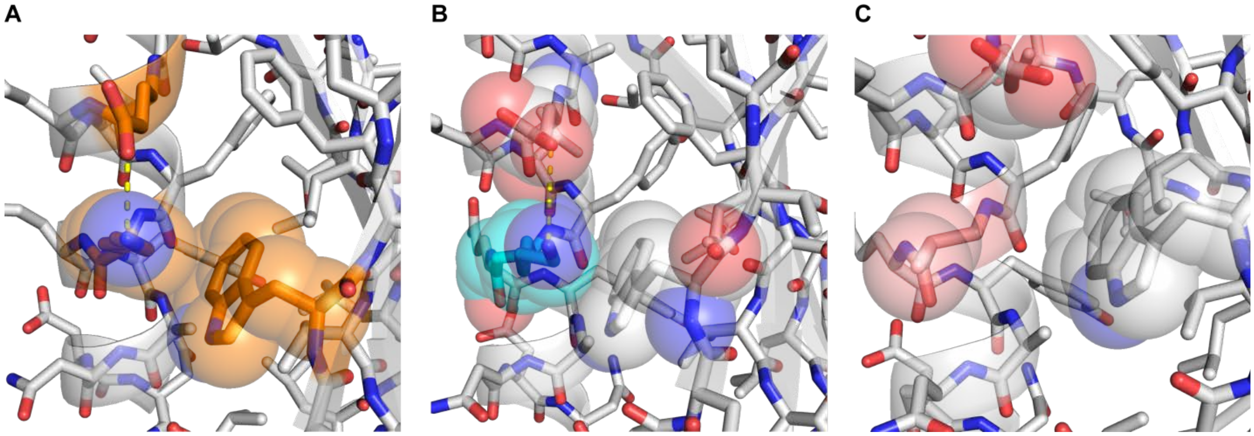


Figure 41. Crystal structures depicting important non-covalent interactions in GB1: a Van der Waals interaction between Lys_{31} and Trp_{48}, and a salt bridge between $Lys_{31}$and Glu_{27} in protein **15** (A, PDB 2QMT) and variant **16** (B, PDB 4KGR). C) Model of sidechain orientation for a β^2 residue built from the crystal structure of **16**.

Further support for the putative role of the Lys_{31} side chain is found by comparing the stability of a bacterially expressed wild-type version of GB1 (**WT**) with its **K$_{31}$A** mutant. Protein **WT** only differs from protein **15** by the first two residues MQ (**WT**) vs. DT (**15**). Loss of the Lys side-chain in mutant **K$_{31}$A** results in a destabilization of ~0.9 kcal mol^{-1}, accounting for the remaining energy difference between **15** and **20**.

From a design standpoint, the above results suggest that β^3- and β^2-residues are comparable in terms of fundamental folding propensity as components of heterogeneous-backbone α/β-peptide helices. Selection of the optimal regioisomer is context dependent and must take into account side-chain contacts important to folding and/or function. While the above examples show how $\beta^3 \rightarrow \beta^2$ substitution can be detrimental, it stands to reason that an identical shift in side-chain placement could be beneficial in other systems. Thus, while the commercial availability of protected β^3 amino acid building blocks make them a good choice for backbone modification, the more synthetically challenging β^2 analogues are likely to be valuable in some situations.

3.2 CALPHA-METHYLATED RESIDUES

As discussed in Section 3.1, restraining the conformational freedom of the backbone can improve the folded stability of proteins. This was observed in comparison of protein **16** and **15**, which differ in $\beta^3 \rightarrow \beta^{cyclic}$ substitution. The disadvantage of this design approach is the loss of a side-chain which can play a key role in tertiary interaction essential to folding (*vide supra*). Like β^{cyclic} residues, methylation of the alpha-carbon is known to conformationaly constrain the dihedral angles of amino acids and promote helix formation.[134]

Aminoisobutyric acid (Aib) is the C_α-methylated analogue of Ala and the simplest of this class of amino acids. Since it is achiral, Aib has the ability to form both left- and right-handed helices. When placed in the context of chiral amino acids, it will promote the helix handedness dictacted by the surrounding residues (i.e. a right-handed helix with L-residues).[194] Aib has also been found to increase folded stability in both small peptides[195] and protein contexts.[136] *In vitro* proteolytic degradation studies incorporating Aib also indicated that it imparts a greater degree of protection than β^3-residues.[120] Protein **21** was synthesized to determine whether Aib or β^{cyclic} residues are superior for stabilizing a helical fold– strategies which both have constrained conformational space and loss of the side-chain. Protein **21** is a variant of **16** where the outer β-residues are replaced with Aib.

A crystal structure of **21** was solved to 2.15 Å resolution and revealed successful incorporation of the Aib residues into the mixed α/β-helical fold (Figure 42).

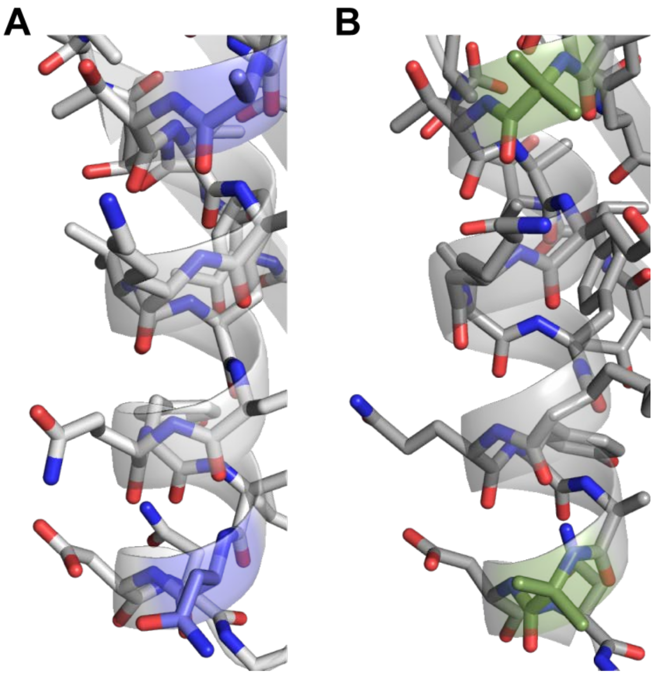


Figure 42. Crystal structure demonstrating successful incorporation of Aib residues into the tertiary fold comparing native sequence **15** (A, PDB 2QMT) with **21** (B, PDB 4HI1; Aib residues in green) .

The free energy of folding was stabilized by 1.7 kcal mol^{-1} over protein **16** (2β^3→2Aib) and 1.3 kcal mol^{-1} over protein **17** (2β^{cyclic}→2Aib, Figure 43). Surprisingly, this folded stability arose from enthalpic ($\Delta\Delta H_{fold}°$ = - 4.3 kcal mol^{-1} vs. **17** and -4.4 kcal mol^{-1} vs. **16**) contributions which were partially offset by entropic destabilization ($-T\Delta\Delta S_{fold}°$ = 3.0 kcal mol^{-1} vs. **17** and 2.7 kcal mol^{-1} vs. **16**). The enthalpic stabilization likely arises from a more native-like fold of the helix. It was hypothesized that the entropic destabilization was due to the ability of Aib to adopt left-handed dihedral angles in the unfolded state. To test this hypothesis a variant using the chiral C_α-methylated residue, C_α-methyl-Val, in place of Aib (protein **22**) was synthesized.

3.2.1 The Impact of Chirality

C_α-methyl-Val, a regioisomer of the β^{cyclic} residue ACPC, highly prefers right-handed helices over left-handed ones.[144] It is therefore an intermediate backbone rigidifier between the less constrained Aib and the highly constrained ACPC. All three residues contain aliphatic side-chains and exhibit thermodynamic trends with overall folded stability decreasing with backbone rigidity: ACPC<C_α-Me-Val<Aib. As backbone rigidity increases enthalpic stabilization decreases while entropic stabilization increases (Figure 43).

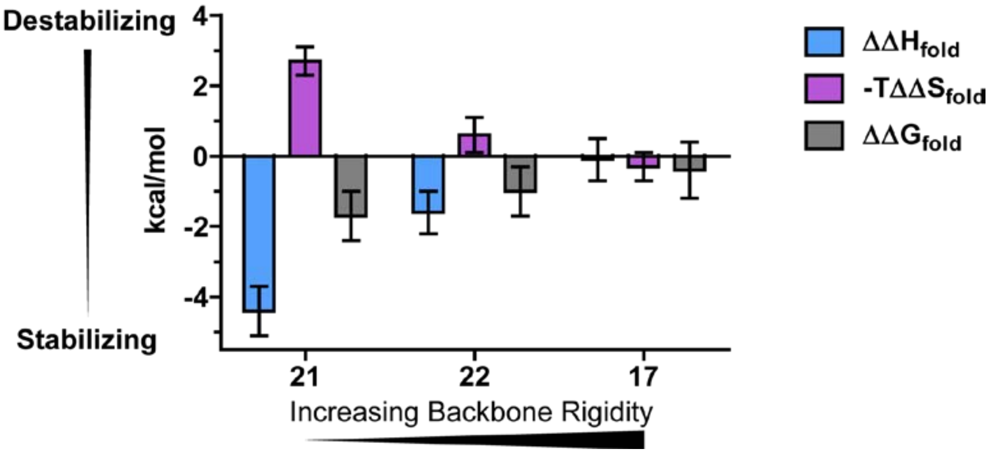


Figure 43. Effect of β^3 to more rigid backbone substitutions on the thermodynamics of folding: **21** ($\beta^3 \rightarrow$ Aib), **22** ($\beta^3 \rightarrow C_\alpha$-Me-Val), and **17** ($\beta^3 \rightarrow \beta^{cyclic}$). All proteins are referenced to the thermodynamic values for protein **16** (all four sites β^3).

Similar to protein **21**, protein **22** is more stable than both **16** ($\Delta\Delta G_{fold}°$ = -1.0 kcal mol^{-1}) and **17** ($\Delta\Delta G_{fold}°$ = -0.6 kcal mol^{-1}). As predicted, protein **22** has less entropic destabilization ($-T\Delta\Delta S_{fold}°$ = -2.1 kcal kcal mol^{-1} vs. **21**) to the fold than protein **21**, which is attributed to the difference in chirality between Aib and C_α-methyl-Val. This improvement in entropy of folding is not coupled with the same degree of enthalpic stabilization found with Aib leading to an overall less stable free energy of folding. The differences in enthalpy between and Aib and C_α-methyl-Val are unclear, but may be related to the increased sensitivity of protein **22** to chemical denaturation. Crystallization trials with protein **22** failed to produce diffraction-quality crystals, so structural detail beyond the CD measurements is unavailable.

Encouraged by the success of proteins **21** and **22** we synthesized protein **23**, a variant of **21** retaining Aib at positions 24 and 35 with C_α-methyl-Lys replacing the β^3-residues at positions 28 and 31. Due to the synthetic difficulty of incorporating so many sterically hindered amino acids, insufficient material was obtained for full thermodynamic and structural analysis. However, a simple CD thermal melt (Figure 44A) revealed the highest melting temperature (T_m) of any GB1 variant (79.6 ± 0.5 °C): 1.7 °C greater than protein **21** and only 2.5°C less than wild-type protein **15** (Table 6, Section 3.4.4).

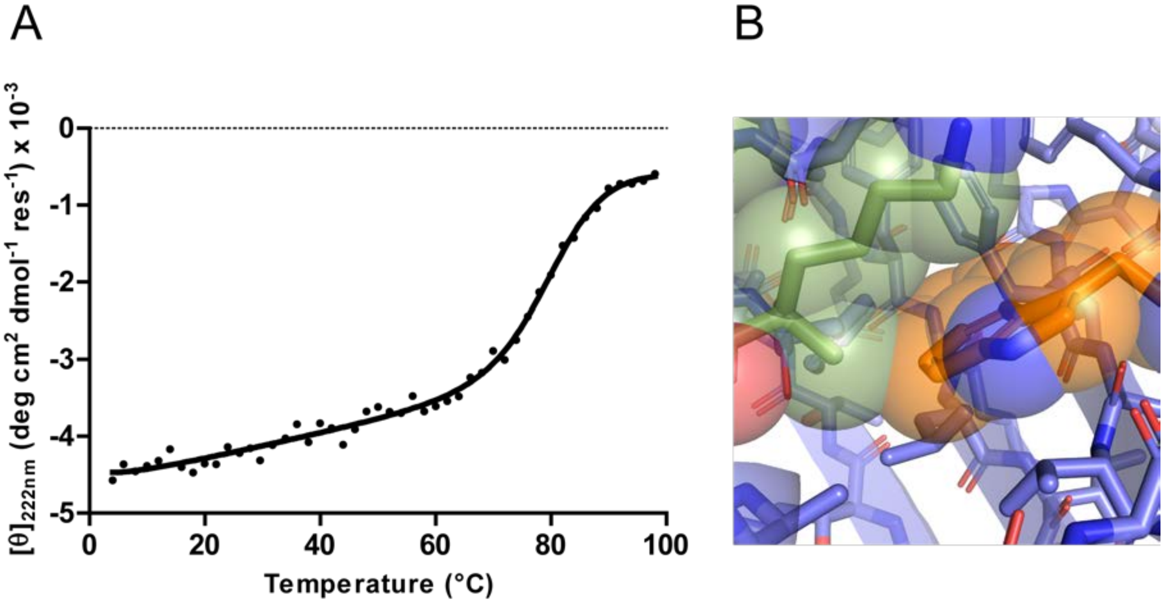


Figure 44. Stability of protein **23**.

A) CD thermal melt of protein **23** at 8 ± 3 µM in 20 mM phosphate buffer pH 7. B) Model of putative steric clash in protein **23** between Lys_{31} and Trp_{43} derived from a crystal structure of wild-type GB1 **15**.

The lower T_m than wild-type **15** may arise from a steric clash between the alpha-methyl carbon of Lys_{31} and Trp_{43} (Figure 44B). Another hypothesis is that the increased backbone rigidity decreased the folded stability in a manner similar to the trend seen between Aib, C_α-Me-Val, and ACPC. These results highlight the ability of C_α-methylation as an important design tool to stabilize helices in a tertiary fold with heterogeneous backbones.

3.3 CONCLUSIONS AND FUTURE DIRECTIONS

The results of the experiments described in this chapter provide insight into general design rules for constructing helices with a heterogeneous backbone in tertiary folds. It is clear that effects of substitutions of unnatural residues for native ones are context dependent. While β^3- and β^2-residues are largely thermodynamically equivalent, side-chain display is important in cases where there are intramolecular contacts. Positions where side-chain interactions are unimportant allow for the use of conformationaly constrained β^{cyclic} or C_α-methylated residues. Within this class, the order of stabilization was found to be Aib > C_α-methyl-Val > ACPC. C_α-methylated residues were better at stabilizing helices over β-residues in all cases, but synthetic difficulties may limit their utility somewhat.

It is intriguing to note that neither protein **16** (4 β^3- residues) nor protein **23** (4 C_α-methylated residues) resulted in higher thermal stability than native α-residues. This may imply that α-residues exist at a thermodynamic sweet spot between flexibility and rigidity to impart maximal thermodynamic stability to natural proteins. While the GB1 system is only a single test case for this hypothesis it suggests a possible explanation for the evolutionary selection of α-amino acids over other backbones as the building blocks of life. This is an especially interesting consideration in relation to the diversity of non-canonical amino acids isolated from meteorites and experiments mimicking pre-biotic environments.[196-197]

Recent advances modifying β-sheets with γ-[188] or dialkylated-[187] residues have resulted in near native-like stability. Combining these advances with those discussed above for helix modification suggest that it should be possible to design tertiary folds with a significant degree of backbone heterogeneity that exhibit native-like structure and thermodynamic stability. Targets that are currently being pursued to understand the effect of modifying helices in other contexts include zinc fingers[198] and disulfide-rich peptides along with large proteins such as ubiquitin. Furthering our understanding of these design principles will lead to peptide therapeutics with enhanced proteolytic stability.

3.4 EXPERIMENTAL

3.4.1 General Information

Solvents and all other reagents were purchased from Aldrich, Baker, EMD, or Fisher and used without further purification. HOBt was purchased from Anaspec Inc. HCTU, NovaPEG Rink Amide Resin, and Fmoc-protected α-amino acids were purchased from Novabiochem. Fmoc-protected β-amino acids were purchased from Aapptec. Fmoc-β^2-Ala-OH and Fmoc-β^2-Lys(Boc)$_2$-OH were synthesized according to published routes.[199] Fmoc-C_α-Me-Val-OH was purchased from Aspira. Fmoc-C_α-Me-Lys(Boc)-OH was purchased from W & J Pharma Chem Inc. Fmoc-β^2-Asn(Dcmp)-OH was synthesized by a route developed by Dr. Zach E. Reinert and Brian D. Griffith.[138]

3.4.2 Protein Synthesis and Expression

Proteins **15**, **16**, and **17** were synthesized as described previously.[131, 142] Proteins **18-20** were synthesized at room temperature (*Method 1*), while **21-23** were prepared by a combination of room temperature and microwave-assisted reactions (*Method 2*). All syntheses were carried out on a 70 µmol scale using NovaPEG Rink Amide resin.

Method 1. Room temperature reactions were performed on a PTI Tribute automated synthesizer. In a standard coupling reaction, 2.5 mL of a solution composed of 0.2 M HCTU, 0.4 M N-methylmorpholine in DMF was added to 7 equivalents of Fmoc-amino acid relative to resin. After a 2 minute pre-activation, the solution was added to the resin and vortexed for 45 minutes. Deprotections were performed by two treatments with 3 mL of 20% v/v 4-methylpiperidine in DMF for 4 minutes each. The resin was washed four times with 3 mL of DMF after each coupling and deprotection step. After the final Fmoc deprotection, the resin was rinsed three times with 3 mL each DMF, DCM, and MeOH. The resin was dried in a vacuum desiccator for 20 minutes prior to TFA cleavage.

Method 2. Microwave reactions were performed using a CEM Microwave-Assisted Reaction System (MARS). Coupling cycles consisted of a 1.5 minute ramp to 90°C followed by a 2 minute hold, while deprotection cycles consisted of a 1.5 minute ramp to 90°C followed by a 1 minute hold.[200] Coupling solutions included protected amino acid (7 equiv), HATU (6.9 equiv), and DIEA (10.5 equiv) in DMF, and were preactivated for 2 minutes prior to addition to resin. Fmoc deprotections made use of 20% v/v 4-methylpiperidine in DMF.

Unless otherwise indicated, synthesis was carried out by *Method 1*. Following are the modifications to the standard automated method made for specific proteins. For proteins **18-23**, Glu_{56} was double-coupled, and pseudoproline dipeptides were used for residues $Ala_{48}Thr_{49}$ and $Glu_{15}Thr_{16}$. Pseudoproline dipeptides were coupled with PyAOP or HATU for 90 minutes at room temperature. For protein **21**, Aib residues were coupled with PyAOP for 90 minutes and the Ala residues immediately following each were double-coupled with PyAOP for 90 minutes each. For protein **22**, αMeVal residues were coupled with HATU for 90 minutes, and the resin was capped with DMF/DIEA/Ac_2O (8/2/1, v/v/v) for 10 minutes prior to deprotection. The Ala residues following each were double-coupled in the microwave (*Method 2*), with capping. For protein **23**, Aib_{35} was coupled with HATU for 90 minutes. Ala_{34} was double-coupled with HATU for 90 minutes each, followed by capping. For residues 23-31, synthesis proceeded in the microwave (*Method 2*). Residues Phe_{30}, Glu_{27} and Ala_{23} were double-coupled. Capping was done after $\alpha MeLys_{31}$, Phe_{30}, $\alpha MeLys_{28}$, Glu_{27}, and Ala_{23}.

All proteins were cleaved from the resin in a cocktail of TFA/H_2O/EDT/TIS (92.5/3/3/1.5 by volume) for approximately 3.5 hours on a rocker. Crude protein was precipitated from the TFA solution by addition of cold Et_2O, and the mixture was centrifuged and decanted. The crude pellet was suspended in ~7 mL of 6M guanidinium, 0.025 M phosphate pH 7. The organic and aqueous layers of the resulting suspension were separated by centrifuge prior to purification.

Proteins (**WT** and **K31A**) were expressed in *E. coli* by Dr. Zach E. Reinert following published protocols[151] using plasmids graciously provided by Dr. Timothy F. Cunningham and Dr. Sunil K. Saxena (University of Pittsburgh).

3.4.3 Peptide Purification and Characterization

All proteins were purified by preparative (300 Å pores, 10 µm beads) C18 reverse-phase HPLC using gradients composed of 0.1% TFA in water/acetonitrile. Protein identities were confirmed using mass spectrometry on a Voyager DE Pro MALDI-TOF instrument (Table 5).

Table 5. Calculated and observed average masses of proteins **16-23**, **WT**, and **K31A**.

#	**$[M+H]^+$ *m/z* (average)**	
	Calculated	**Observed**
16	6235.8	6232.8
18	6235.8	6233.4
19	6235.8	6234.8
20	6235.8	6234.6
17	6244.8	6241.0
21	6191.7	6191.9
22	6247.8	6247.5
23	6191.7	6188.2
K31A	6164.7	6164.7
WT	6223.8	6223.0

Following HPLC, proteins were further purified by anion-exchange chromatography on a MonoQ 5/50GL column (GE Healthcare) using 0.02 M Tris pH 8 buffer eluted with increasing concentrations of NaCl. Following ion-exchange, protein **23** was further purified using semi-preparative (300 Å pores, 5 µm beads) C18 reverse-phase HPLC using gradients as described above. Final purity of each protein was ≥95% by analytical RP-HPLC (Figure 45).

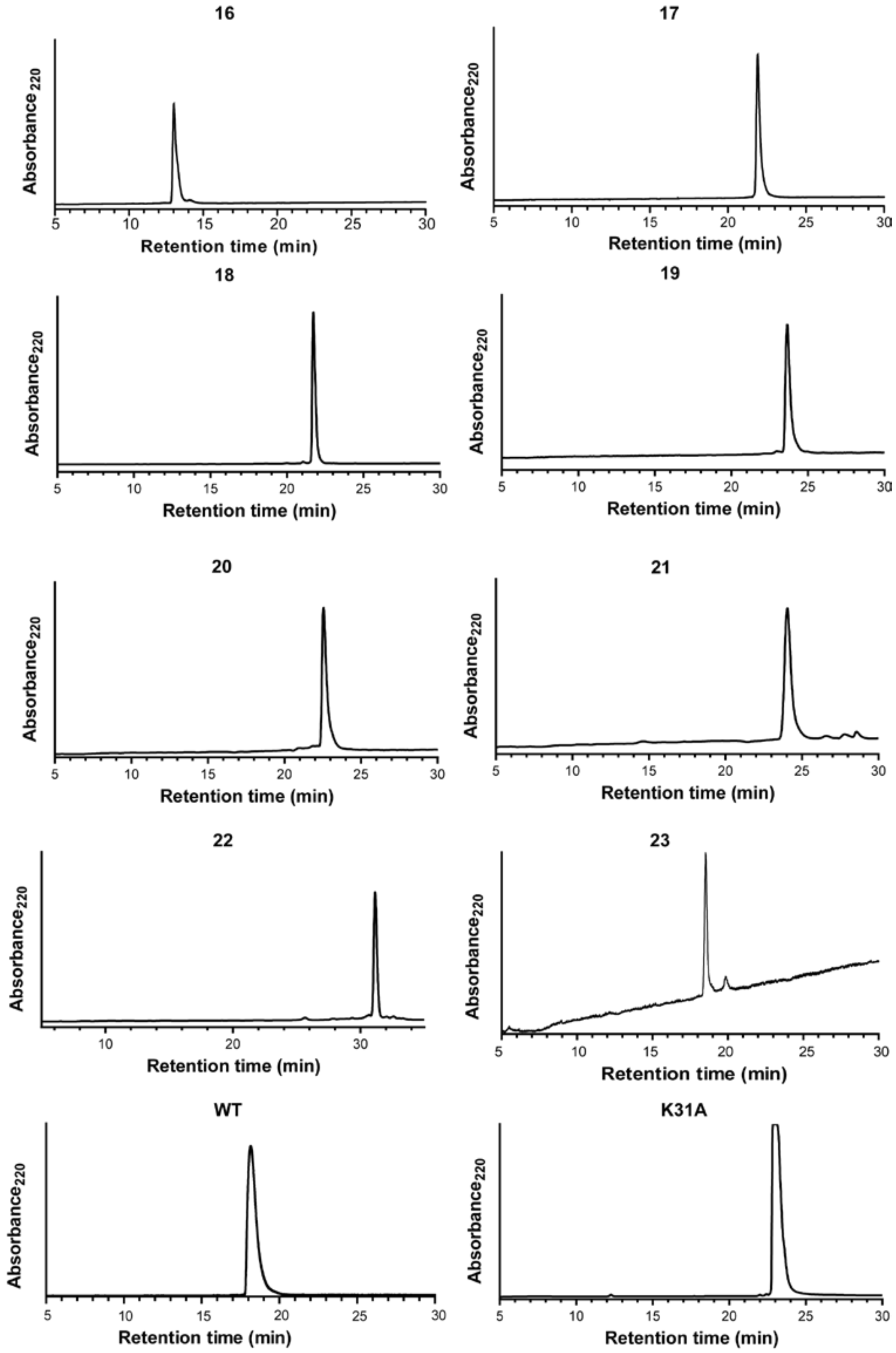


Figure 45. Analytical reverse phase HPLC chromatograms of purified proteins **16-23**, **WT**¸ and **K31A**. HPLC experiments were run on a C18 column (10 µm particle, 300 Å pore size) on gradients between water and acetonitrile with 0.1% TFA.

3.4.4 Circular Dichroism Measurements and Thermodynamics of Folding

Circular dichroism measurements and data analysis were performed as described previously.[131] Briefly, thermal melts of proteins in aqueous phosphate buffered (20 mM, pH 7) with varying concentrations of guanidinium chloride (0-6 M) were collected at 2 °C increments (2 min. equilibration) from 2-98 °C at 220 nm and 2 nm bandwidth with 5 second integration times. Concentration for all proteins was 40 µM except for protein **23** which was 8 µM due to poor synthetic yield. All measurements were buffer subtracted and baseline corrected. Raw data was then fit (Figure 46) to a series of multivariate equations (Equations 1-4)[131] to extract thermodynamic parameters using Mathematica (Wolfram).

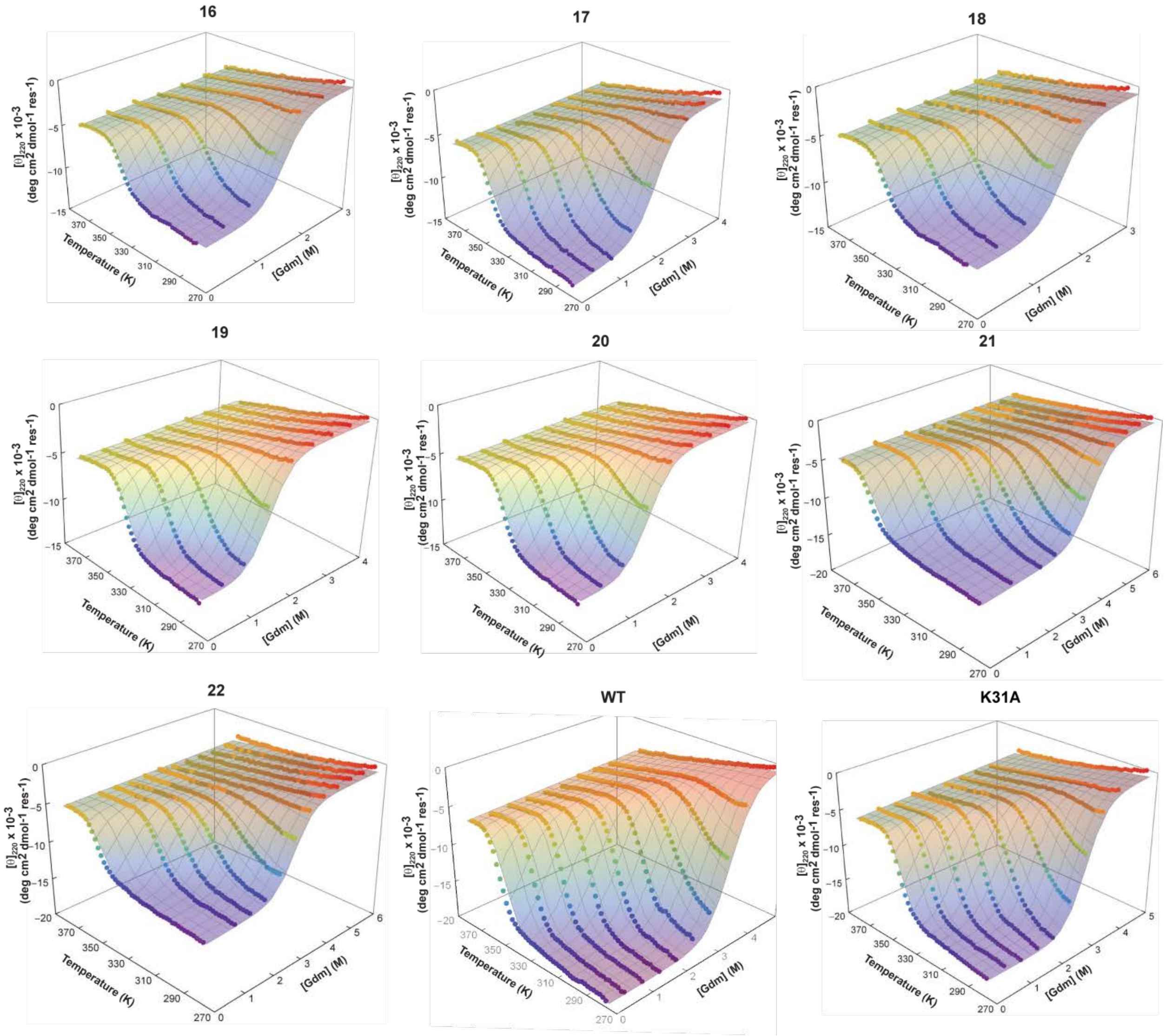


Figure 46. Tandem CD thermal melt and chemical denaturation data at 220 nm for proteins **16-22**, **WT**, and **K31A**. Raw data was fit (points) was fit (surface) to Equations **1-4** to extract thermodynamic parameters for protein folding.

The free energy of protein folding (ΔG) with respect to temperature (T) and a concentration of chemical denaturant (here guanidinium, denoted [Gdm] in M) can be described by Equation 1:

Equation 1

$$\Delta G = \Delta H^\circ - T\Delta S^\circ + \Delta C_p \times (T - T_0 + T \times ln\left(\frac{T_0}{T}\right) - m \times [Gdm]$$

where ΔH° (kcal mol^{-1}) and ΔS° (kcal mol^{-1} K^{-1}) are the folding enthalpy and entropy at a reference temperature (T_0, here T_0 = 20 °C). ΔC_p (kcal mol^{-1} K^{-1}) is the change in heat capacity and *m* (kcal mol^{-1} M^{-1}) is the susceptibility to unfolding by a chemical denaturant, here [Gdm]. The observed ellipticity (θ_{obs}) is related to the free energy and temperature at a constant [Gdm] by Equation 2:

Equation 2

$$\theta_{obs} = \frac{\left(\theta_n + \theta_u \times e^{\left(\frac{-\Delta G^{\circ}}{RT}\right)}\right)}{\left(1 + e^{\left(\frac{-\Delta G^{\circ}}{RT}\right)}\right)}$$

where θ_n is the fully folded ellipticity and θ_u is the fully unfolded ellipticity. Based on literature precedent[201] and the observed cooperativity of the melt data, it is was assumed that GB1 and its analogues followed a two-state folded/unfolded equilibrium. Folded and unfolded ellipticity values were assumed to vary linearly with T and [Gdm];[190, 202] raw data was fit by allowing ΔH°, ΔS°, ΔC_p, *m*, and the six parameters (*a*-*f*) in Equations 3-4:

Equation 3

$$\theta_u = a + bT + c[Gdm]$$

Equation 4

$$\theta_n = d + eT + f[Gdm]$$

to vary until the best fit solution to the non-linear model composed of Equations 1-4 was found. Thermodynamic parameters calculated from the fits are given below in Table 6.

Table 6. Thermodynamic parameters for the unfolding of synthetic proteins **15-23** and expressed proteins **WT** and **K31A**.[a]

Protein	ΔH° (kcal mol^{-1})	TΔS° (kcal mol^{-1})	ΔG° (kcal mol^{-1})	ΔC$_p$ (kcal mol^{-1} K^{-1})	*m* (kcal mol^{-1} M^{-1})	T_m (°C)[b]
15	22.0 ± 0.6	16.1 ± 0.3	5.9 ± 0.7	0.60 ± 0.02	1.80 ± 0.04	82.1 ± 0.1
16	18.3 ± 0.4	15.2 ± 0.3	3.1 ± 0.5	0.53 ± 0.02	2.48 ± 0.05	61.6 ± 0.1
17	18.4 ± 0.5	14.9 ± 0.3	3.5 ± 0.6	0.48 ± 0.02	2.00 ± 0.05	70.8 ± 0.2
18	18.8 ± 0.6	15.8 ± 0.3	3.0 ± 0.7	0.45 ± 0.03	2.25 ± 0.07	64.0 ± 0.2
19	18.7 ± 0.4	16.1 ± 0.3	2.6 ± 0.5	0.48 ± 0.02	2.36 ± 0.04	61.4 ± 0.4
20	17.0 ± 0.3	15.5 ± 0.3	1.5 ± 0.4	0.48 ± 0.02	2.45 ± 0.04	47.7 ± 0.2
21	22.7 ± 0.6	17.9 ± 0.5	4.8 ± 0.7	0.44 ± 0.02	1.57 ± 0.04	77.9 ± 0.1
22	19.9 ± 0.4	15.8 ± 0.3	4.1 ± 0.5	0.46 ± 0.02	1.73 ± 0.04	76.2 ± 0.1
23[c]	-	-	-	-	-	79.6 ± 0.5
WT	23.0 ± 0.5	17.9 ± 0.3	5.1 ± 0.6	0.57 ± 0.02	1.71 ± 0.04	78.5 ± 0.3
K31A	22.7 ± 0.4	18.5 ± 0.3	4.2 ± 0.5	0.67 ± 0.02	1.77 ± 0.03	67.6 ± 0.1

[a] Thermodynamic values are at 298 K and reported errors are from parameter uncertainties in the fit. [b] Midpoint of the CD thermal unfolding transition in the absence of chemical denaturant. [c] Protein **23** measured at 8 μM (0 M denaturant only) instead of 40 μM due to poor synthetic yield.

3.4.5 Crystallography

Crystals were grown using the hanging drop vapor diffusion method. Stock solutions (18 mg/mL for **18**, 20 mg/mL for **19** and **21**) were mixed (0.7 μL + 0.7 μL) with crystallization buffer (Table 7) and allowed to equilibrate at room temperature over a well of that buffer.

Table 7. Crystallization buffer formulations for proteins **18**, **19**, and **21**.

Protein	*Buffer*
18	200 mM NaOAc pH 4.6, 20% PEG 4000
19	100 mM Na Cacodylate pH 6.5, 0.15 M $Mg(OAc)_2$, 20% PEG 4000
21	100 mM NaOAc pH 4.5, 0.2 M $(NH_4)_2SO_4$, 20% PEG 4000

Harvested crystals were cryoprotected with 30% v/v glycerol in the mother liquor then flash frozen in liquid nitrogen. X-ray diffraction data were collected using Cu/K$_\alpha$ radiation on a Rigaku/MSC diffractometer (FR-E generator, VariMax optics) with a Saturn 944 CCD detector for **18** and **21** or a RAXIS HTC image plate detector for **19**. Crystals were maintained at 100 K during diffraction experiments with an X-Stream 2000 low-temperature system.

Raw diffraction data were processed with d*TREK. Structures were solved by molecular replacement with a published structure of the expressed wild-type GB1 (PDB 4QMT) as a search model. Model refinement was performed with Phenix,[171] and manual real-space model building was accomplished using Coot.[172] Final data collection and refinement statistics may be found in Table 8. Protein **19** exhibited signs of twinning as determined by Xtriage in Phenix, and the final structure was refined with a twin fraction of 0.27 and twin law of (H, -L, -K).

Table 8. X-ray data collection and refinement statistics for proteins **18**, **19**, and **21**.

	18	**19**	**21**
Data Collection			
Unit cell dimensions (Å, °)	*a* = 92.8, *b* = 22.4, *c* = 65.3 α = γ = 90, β = 134.1	*a* = *b* = 51.9, *c* = 96.4 α = β = γ = 90	*a* = 74.4, *b* = 73.4, *c* = 79.4 α = γ = 90, β=99.3
Space group	C2	$P4_1$	C2
Resolution (Å)	23.44-1.95 (2.02–1.95)	51.95-1.80 (1.86-1.80)	41.15-2.15 (2.23-2.15)
Total observations	34,719	297,770	260,925
Unique observations	7,188	22,477	22,880
Redundancy	4.8 (3.2)	13.25 (13.18)	11.40 (3.04)
Completeness (%)	97.8 (90.4)	95.0 (90.5)	99.2 (93.2)
I/σ	18.5 (3.6)	25.3 (4.9)	15.9 (4.2)
R_{merge} (%)	6.9 (15.2)	6.6 (40.5)	13.7 (23.9)
Refinement			
Resolution (Å)	23.44–1.95	51.95-1.80	41.15-2.15
R (%)	19.9	19.9	21.71
R_{free} (%)	23.0	21.7	25.23
Avg. B factor (Å^2)	22.1	27.3	24.33
RMSD			
Bonds (Å)	0.005	0.006	0.005
Angles (°)	1.04	1.13	1.018
Twinning			
Twin fraction	n/a	0.27	n/a
Twin Law	n/a	H, -K, -L	n/a
PDB ID	5HFY	5HG2	5HI1

4.0 DESIGN OF COILED-COIL BASED SUPRAMOLECULAR METALLOPOLYMERS

A portion of this work has been previously published as:

> Tavenor, N. A.; Murnin, M. J.; Horne, W. S. "Supramolecular Metal-Coordination Polymers, Nets, and Frameworks from Synthetic Coiled-Coil Peptides." *J. Am. Chem. Soc.* **2017**, *139*, 2212-2215.

Matthew J. Murnin synthesized peptides **33** and **34** and set-up crystal trays for both of these peptides. Synchrotron experiments were carried out by remote collection using the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22ID at the Advanced Photon Source (APS), Argonne National Laboratory. A special thanks to Dr. William F. Furey for organizing the beamline usage. All EPR measurements were collected and analyzed by Matthew J. Lawless (Saxena Lab, University of Pittsburgh).

Expanding upon the work described in Chapter 2, we endeavored to design coiled coil–based supramolecular materials based on a more facile synthetic approach. One of the disadvantages of the organic linkers employed in our prior systems was the necessity of a second round of purification after cross-linking. This increases the time and materials required and results in a lower final yield. Moreover, disassembly of these materials was only possible by unfolding the protein quaternary structure through thermal or chemical denaturation.

Motivated to overcome the above limitations, we envisioned an alternative way to promote supramolecular assembly through metal coordination as a means of orthogonal control over the assembly to protein folding. This strategy has been used extensively in the supramolecular polymer and metal organic framework fields, which provided inspiration for this work.[140, 203-204] Incorporation of a metal chelating group into the coiled-coil peptide should allow for the formation of supramolecular materials by the simple addition of a metal ion at an appropriate stoichiometric ratio. This assembly process can then be easily tuned by the number of equivalents and identity of the metal (Figure 47).

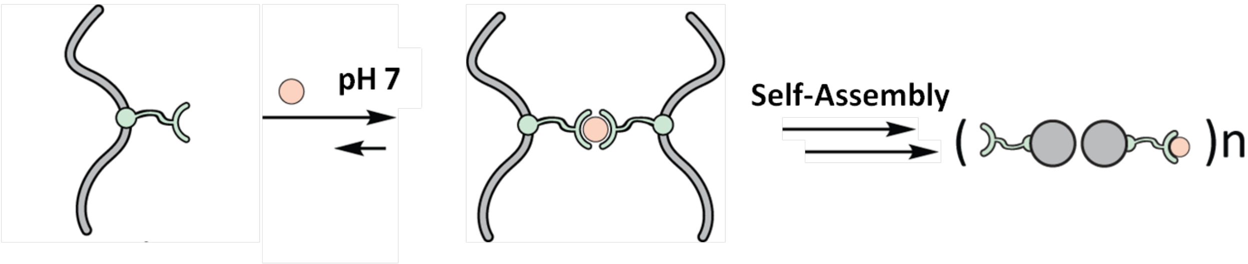


Figure 47. Cartoon representation of the chelation and subsequent supramolecular assembly of a dimeric coiled-coil peptide with a metal coordinating group.

Control over the size and structure of the resulting material can be realized through judicious placement of the metal coordinating group and the oligomerization state of the coiled coil.

4.1 CHOICE OF METAL BINDING SIDE CHAIN

While several natural amino acid residues can chelate metals (His, Asp, Glu, Cys), we sought to apply an unnatural one that would be orthogonal to side-chain functional groups found among the canonical amino acids. The properties desired for a metal chelating group were: 1) a small number of coordination modes, 2) ability to tightly chelate a range of metal ions, and 3) easy attachment to the peptide. Terpyridine (Tpy), a tridentate aromatic ligand that binds divalent transition metal ions with high affinity in either a bis- or mono- complex (Figure 48), was identified as an ideal candidate.

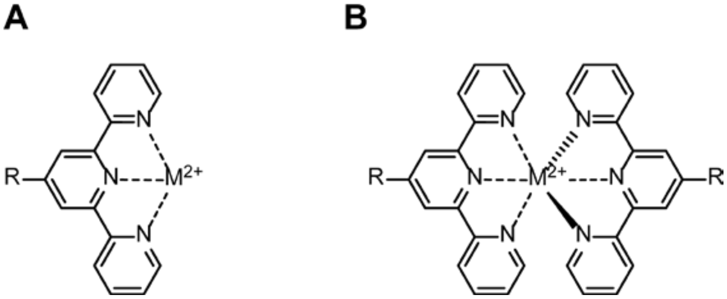


Figure 48. Generic mono (A) and bis-Tpy (B) complexation where M^{2+} is a divalent transition metal ion.

Tpy has long been used in supramolecular chemistry to form supramolecular polymers.[205-207] Synthesis of derivatives is well-established,[208] and 4'-carboxy-2,2':6',2"-terpyridine (where R = COOH) is accessible in a single step allowing for easy attachment to a lysine residue on resin. We hypothesized that supramolecular polymerization of the Tpy-functionalized peptide ligands could be controlled by the stoichiometry of the divalent transition metal ion by favoring either the mono- or bis-Tpy complex (Figure 48) thereby promoting either disassembly or polymerization, respectively. Supramolecular assembly of the peptides was investigated with the transition metal ions Zn^{2+}, Cu^{2+}, Co^{2+}, Ni^{2+}, Cd^{2+}, and Ru^{2+}.

Cu^{2+} proved uniquely well-suited among this series in promoting formation of highly ordered supramolecular architectures (see Section 4.2), albeit by a coordination mechanism different than the initial design. The utility of Cu^{2+} as an EPR spectroscopic probe for complex systems also allowed for the characterization of solution phase assemblies formed from terpyridine-modified peptides (see Section 4.3).

4.2 DESIGN OF CRYSTALLINE METALLOPEPTIDE ASSEMBLIES

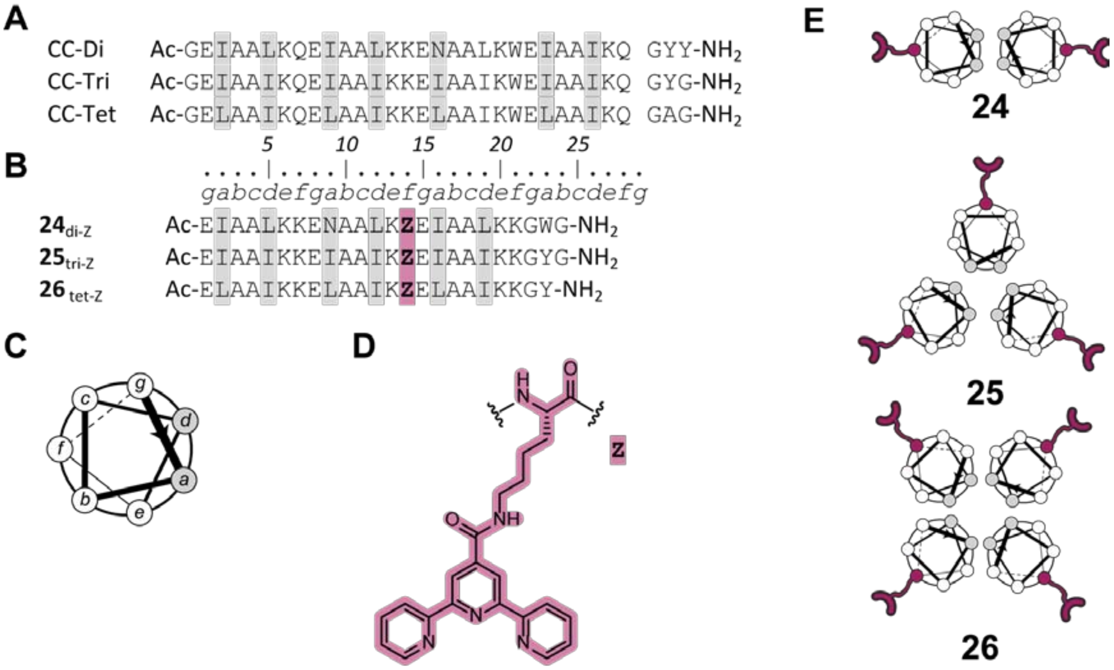


Figure 49. A) Sequences of peptides used for design inspiration[111] and B) terpyridine-modified peptides **24-38**. Lowercase letters denote heptad positions, and gray shaded residues occupy the hydrophobic core of the folded coiled-coil quaternary structure; these residues control oligomerization state of the expected favored assembly (dimer, trimer, or tetramer), which is indicated for each peptide as subscripts along with the type and number of Tpy residues. (C) Helical wheel diagram with a map of heptad positions looking down the axis of an α-helix. (D) structures of unnatural residue **Z** (magenta). (E) Schematic of expected oligomerization state of each peptide with location of terpyridine side chains (magenta shading) in relation to the hydrophobic core (gray shading).

We began our efforts to generate ordered supramolecular lattices with the series of peptides **24-26** (Figure 49). These peptides were intended to form dimeric, trimeric, and tetrameric quaternary structures, and their sequences were based on well-established design rules for α-helical coiled coils.[56, 85, 98] Positions along each individual chain that make up a coiled-coil fold are denoted *abcdefg*; this lettering is based on the ~7 residue repeat of an α-helix (Figure 49C). Positions *a* and *d* in each heptad make up the buried interface between helices in the quaternary structure, and the identity of these residues is the primary determinant of oligomerization state. Residues at *e* and *g* sites often participate in inter-helix electrostatic contacts that help to stabilize the assembly. Heptad positions *b*, *c*, and *f* are solvent exposed and have little impact on folding.

Peptide **24** bears Ile at *a* positions, Leu at *d* positions, and a single *a*-position Asn; all these features should work in concert to promote a dimeric coiled coil fold.[56, 85, 98] The hydrophobic core of peptide **25** (Ile at both *a* and *d*) was intended to specify a trimeric assembly, while that of peptide **26** (Leu at *a*, Ile at *d*) a tetrameric bundle.[56, 85, 98] It was hypothesized that differences in the oligomerization state could predictably tune the morphology of the supramolecular assembly leading to linear polymers, hexagonal nets, or tetrameric nets from dimer (**24**), trimer (**25**), and tetramer (**26**) coiled coils, respectively (Figure 50).

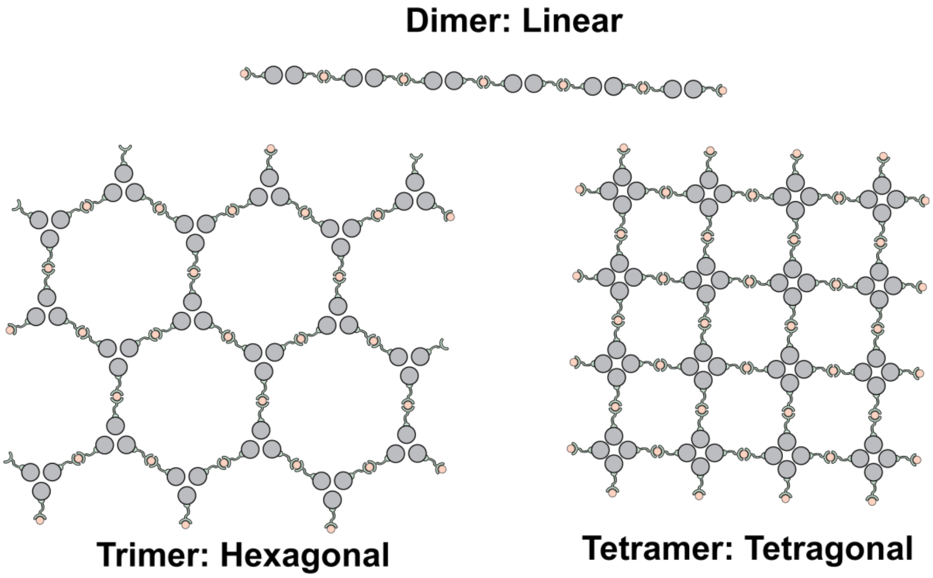


Figure 50. Hypothesized supramolecular lattices formed by a dimer, trimer and tetramer.

In the series **24-26**, the Tpy functionality was attached to each peptide via a Lys side chain (residue **Z**) at a solvent-exposed *f*-position near the midpoint of the sequence. Based on the symmetry of the coiled-coil quaternary structure expected from each sequence (Figure 49E), we hypothesized the addition of divalent metal would lead to either linear supramolecular polymers (for **24**), hexagonal nets (for **25**), or tetragonal nets (for **26**). Crystallization was attempted in the presence of a range of divalent transition metals (Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}) along with Mg^{2+} and Ca^{2+} as controls. Crystals were obtained, but only with Ca^{2+} and Mg^{2+}. From this, we concluded the terpyridine side chain was likely not participating in the lattice formation and therefore did not pursue this series further for crystalline assembly.

We hypothesized that one factor that might be hindering an ordered metal-directed assembly process in **24-26** was the flexibility of the side chain linking the terpyridine to the backbone. Therefore, we next probed the consequence of shortening this side chain by replacing the **Z** residue with an analogue (**X**) derived from diaminobutanoic acid instead of Lys. Making this substitution in peptide **26** led to peptide **27** (Figure 51).

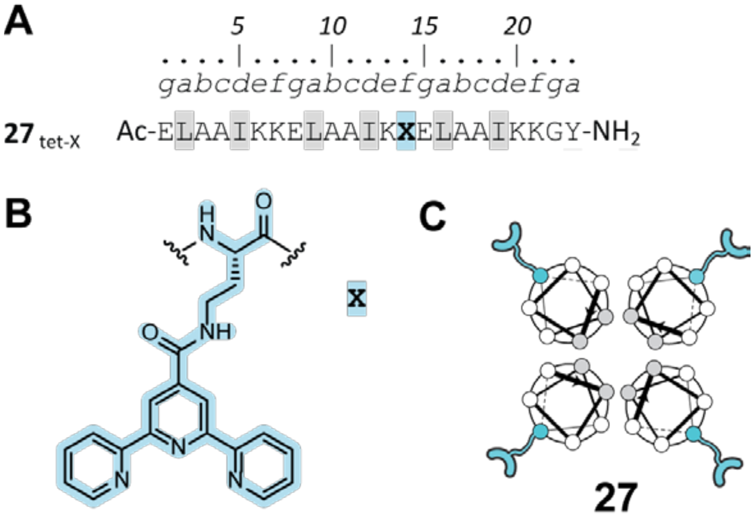


Figure 51. Design of peptide **27**.

(A) Sequence of peptide **27** with the Tpy **X** side-chain highlighted in blue and hydrophobic core residues in gray. (B) Structure of the Tpy **X** residue and (C) desired peptide oligomerization state.

Crystals of peptide **27** were successfully grown in the presence of a variety of divalent metals (Mg^{2+}, Ca^{2+}, Cu^{2+}, or Zn^{2+}). One of these crystals produced a strong diffraction pattern. The structure of this crystal was solved (data not shown) leading to two key observations: (1) the peptide had assembled to form a pentameric helix bundle instead of the expected tetramer and (2) the Tpy side chains were not involved in the formation of the lattice. With respect to the former point, we reasoned that extending the sequence by an additional heptad would lead to greater specificity for the desired tetrameric oligomerization state.

Thus, we prepared peptide **28**, a longer variant of **27** but with the same tetramer-promoting hydrophobic core composition (Ile at *a* positions, Leu at *d* positions). We performed crystallization trials with **28** and a range of divalent transition metals (Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}). Metal:peptide stoichiometries were varied from 0.3-6 equivalents in the presence of bis-TRIS with added formate, PEG 4000, or no precipitating agent. These experiments yielded microcrystals as well as amorphous aggregates but no diffraction-quality crystals. The observation of assembly to form microcrystals under dilute aqueous conditions lacking any precipitating agent led us to hypothesize that the metal ions and Tpy side chains were playing a key role in lattice formation.

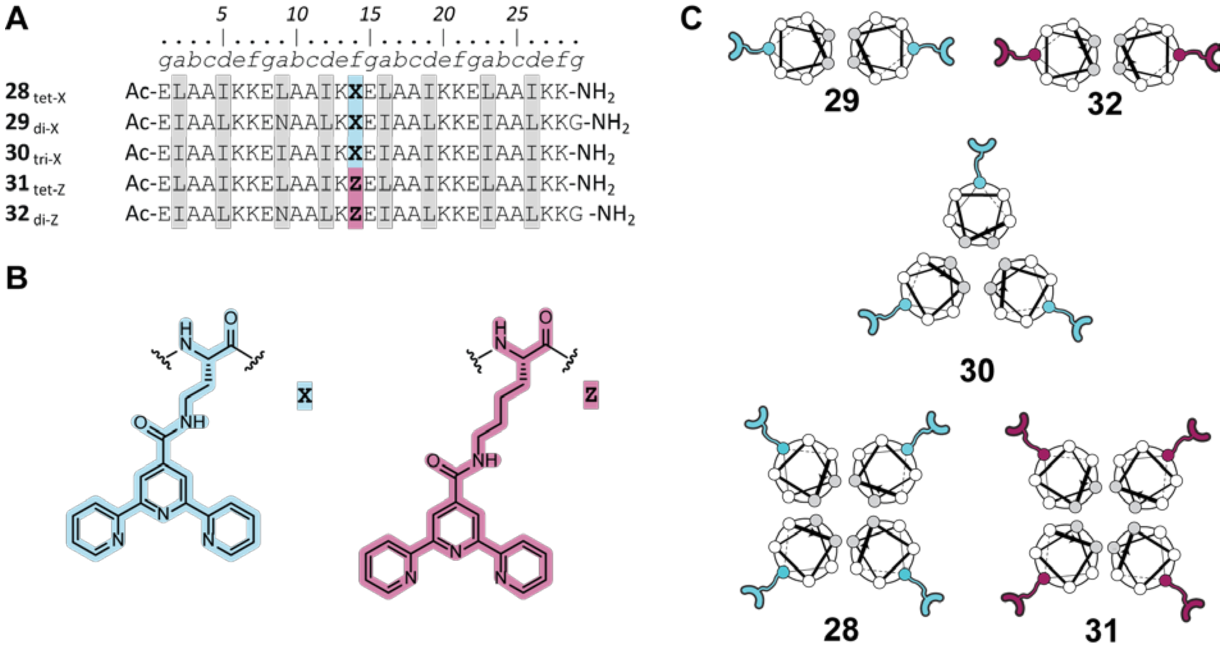


Figure 52. Design of four heptad sequences with a single terpyridine.

(A) Sequences of peptide **28-32** with the Tpy **X** side-chain highlighted in blue and hydrophobic core residues in gray. (B) Structures of the Tpy **X** and **Z** residues and (C) peptide oligomerization state.

Peptides **29** and **30** were designed as dimeric and trimeric variants of **28**, respectively (Figure 52). Both peptides contain the shorter **X** sidechain to promote a more ordered assembly than the **24-26** series. Variants of the tetramer (**31**) and dimer (**32**) were also synthesized with the longer **Z** sidechain to examine the effect of linker length. Although crystalline materials were not generated from this peptide series, the dimers (**29** and **32**) and tetramers (**28** and **31**) were used to study the solution-phase assembly of well-ordered metallopolymers (see Section 4.3).

The results obtained for peptide **28** were the most promising for crystallization among experiments carried out to that point, and we noted that the expected fold consisted of four Tpy groups per coiled coil (one in each chain of a C_4 symmetric tetramer). An alternate way to achieve a coiled-coil assembly bearing four metal-chelating side chains would be to incorporate two Tpy in a dimer forming sequence. We considered that such a peptide might lead to a more ordered lattice and further hypothesized introducing the Tpy groups near the helix termini would minimize the number of favorable packing arrangements among dimers in the crystal. Based on the above reasoning, we prepared peptides **33** and **34** (Figure 53).

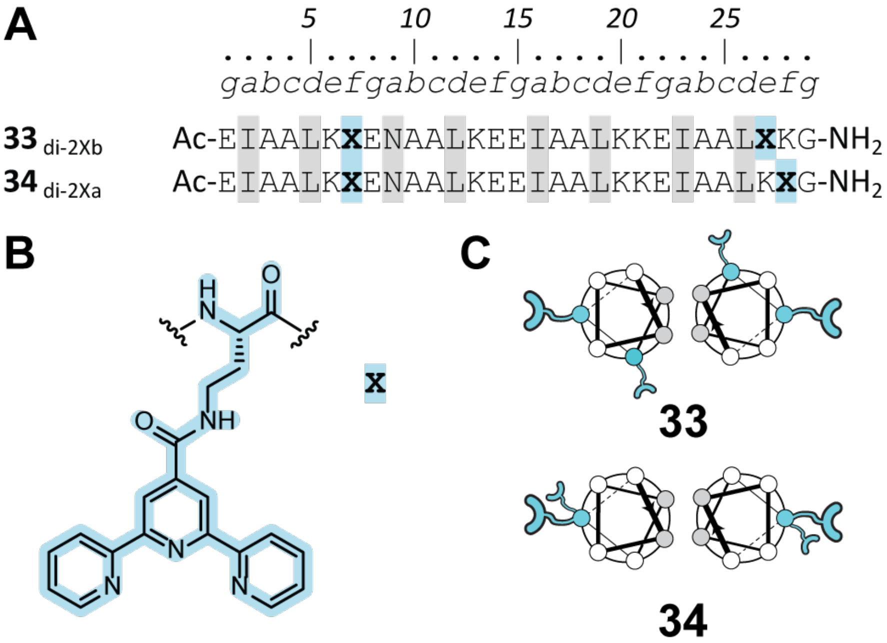


Figure 53. Design of coiled coil dimers with two Tpy residues.

(A) Sequence of peptides **33-34** with the Tpy **X** side-chain highlighted in blue and hydrophobic core residues in gray. (B) Structure of the Tpy **X** residue and (C) peptide oligomerization state with Tpy side chain display.

The design of peptide **33** was based on two primary considerations: (1) residues at *a*/*d* positions were selected to enforce a dimeric fold and (2) two Tpy residues were incorporated at solvent-exposed postitions—one near the N-terminus in the first heptad at position *f* and the other near the C-terminus in the fourth heptad at position *e*. The reasoning behind offsetting the heptad positions was to promote coordination in more than just a linear chain. Peptide **34** differs from **33** only in having both Tpy residues at the *f* positions of the first and second heptad. This was done to promote a ribbon-like assembly.

Although crystallization experiments of the peptides were attempted with an array of transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, and Cd^{2+}) known to form bis-terpyridine complexes,[209-211] only Cu^{2+} produced diffraction-quality crystals. This result was surprising since Cu^{2+} has the lowest propensity among the series to form the desired Tpy-M^{2+}-Tpy coordination motif preferring to form ternary complexes instead.[212]

Initial crystallization experiments of peptide **33** and **34** utilized Cu^{2+} and Zn^{2+} and a range of precipitating agents with bis-TRIS buffer—guided by its prior use as a weak competitor for Zn^{2+} in the growth of crystalline protein arrays.[76] These initial efforts yielded intergrown crystals. Optimization efforts focused on varying the buffer, additive, metal ion and precipitating agent. Crystals were only able to be grown in the presence of a weakly chelating buffer and Cu^{2+}. These efforts finally yielded a diffraction quality single crystal of peptide **33** (crystal form *a*) in the presence of citrate (weak chelating buffer), $CuCl_2$ (metal ion salt) and PEG 4000 (precipitating agent). The structure of this crystal was solved to 2.2 Å resolution (Figure 54, PDB ID 5U59). We also obtained small crystals of peptide **34** in the presence of Cu^{2+}, but the best diffraction of these crystals was weak (processed to 5.4 Å) and we were unable to find a molecular replacement solution.

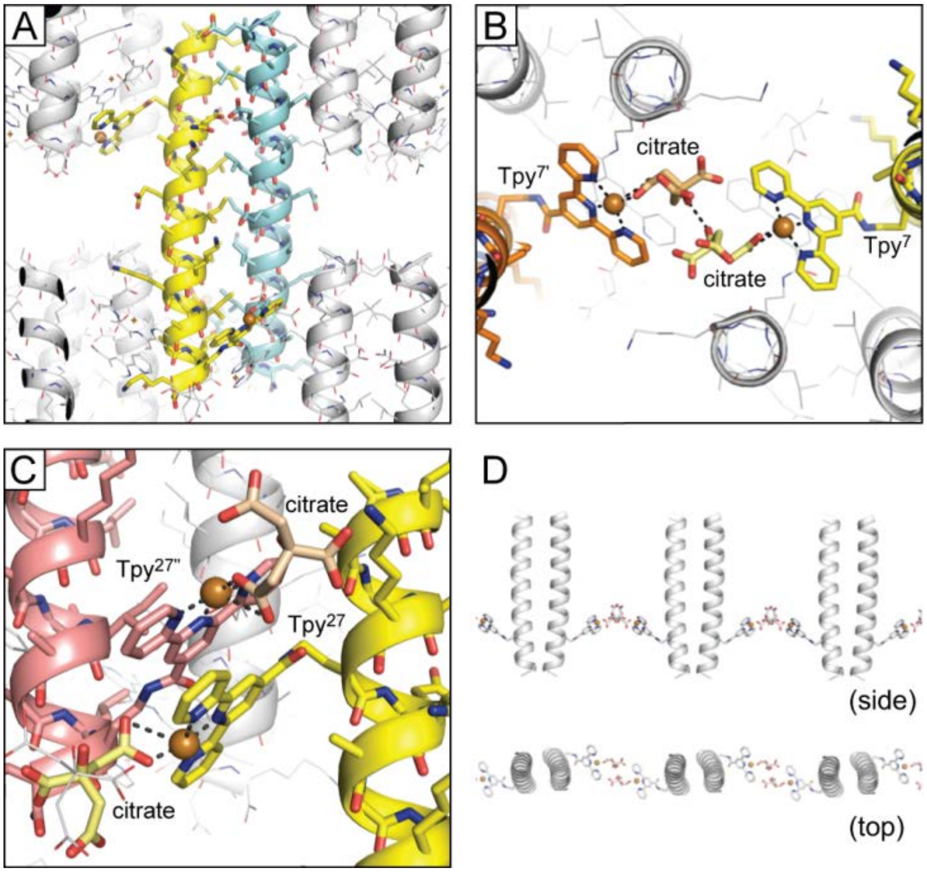


Figure 54. Structure of peptide **33** form *a*.

A) Crystal structure of peptide **33** (form *a*, PDB ID 5U59) grown in the presence of citrate; B,C) Tpy-Cu^{2+}-citrate coordination between peptides; and D) linear supramolecular polymer generated through the interface in panel B (Adapted with permission from Tavenor, N. A., et al. *J. Amer. Chem. Soc.* **2017**, *139* (6), pp 2212–2215. Copyright 2017 American Chemical Society).

Supporting our central hypothesis, the noncovalent interactions that make up the crystal lattice consist entirely of (1) coiled-coil hydrophobic interfaces and (2) inter-chain contacts involving Tpy residues and Cu^{2+} ions. The asymmetric unit consists of a single α-helix, and the dimeric coiled-coil is created by a crystallographic 2-fold symmetry axis propagated through the lattice by the C2 spacegroup. The dimer quaternary structure (Figure 54A) is virtually identical to that of a variant lacking the Tpy residues.[111] The remaining two crystallographically independent contacts between chains involve proximal Tpy residues coordinated to Cu^{2+} (Figure 54B,C). However, neither showed the simultaneous coordination of Cu^{2+} by two Tpy side chains (Figure 48B). Rather, a carboxylate from citrate in the crystallization buffer replaced one of the terpyridine moieties (Figure 54B,C) at both sites.

Unique among the first row transition metals is the propensity of Cu^{2+} to form CuAB type complexes over the corresponding CuA_2 or CuB_2 complex.[212] When bound to multi-dentate aromatic nitrogen ligands like terpyridine and its derivatives, there is a very strong propensity to bind an anionic oxygen ligand to form a ternary complex.[212-213] This motif was observed in all solved crystal structures (*vide infra*). Such terpyridine-Cu^{2+}-carboxylate complexes (Figure 55) are known and have been shown to be stable in aqueous solution.[213-214]

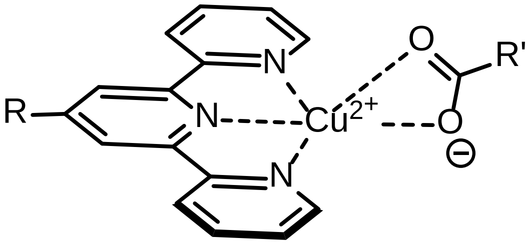


Figure 55. Structure of a terpyridine-Cu^{2+}-carboxylate complex with variable substituents.

This terpyridine-Cu^{2+}-carboxylate motif formed the basis for the supramolecular organization of the crystal lattice by a different mode at each Tpy site. The site at Tpy_7 forms a linear supramolecular polymer which propagates through the lattice (Figure 54D). This is held together by hydrogen bonds between the citrate molecules coordinated to the Tpy-Cu^{2+} complex (Figure 54B). The second site at Tpy_{27} supports the lattice through hydrophobic stacking of the Tpy aromatic groups; a motif which was found to be common in all structures (*vide infra*). This interaction allows the crystal to form along another vector as predicted by the design strategy (*vide supra*).

The above result led us to reason that more rigid carboxylate ligands than citrate may be able to bridge Cu^{2+}-bound Tpy residues to generate a MOF-like assembly. To test this hypothesis, we refocused optimization efforts with an aim to obtain crystals of peptide **33** replacing citrate with a bis-carboxylate linker: *cis*- and *trans*-1,4-cyclohexanedicarboxylic acid, malonate, tartrate, tricine, or terephthalate. It had also been previously observed that the combination of bis-TRIS buffer and Cu^{2+} most readily produced crystals under a wide range of conditions. Combining this observation with the effort to replace citrate with an alternative carboxylate linker yielded a new crystal form (*b*) of peptide **33** from a bis-TRIS buffer containing terephthalate, and the structure of this crystal was solved to 3.2 Å resolution (Figure 56, PDB ID 5U5A).

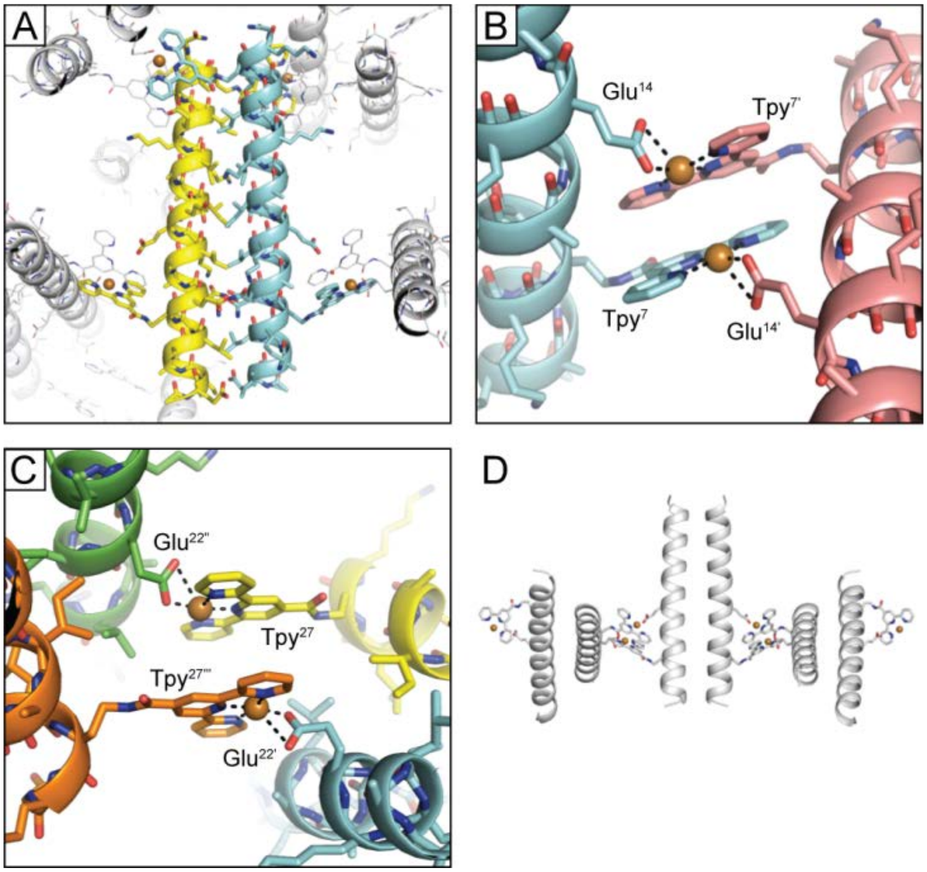


Figure 56. Structure of peptide **33** form *b*.

A) Crystal structure of peptide **33** (form *b*, PDB ID 5U5A) grown in the presence of terephthalate; B,C) Tpy-Cu^{2+}-Glu coordination between peptides; and D) linear polymer generated by the coordination in panel C (Adapted with permission from Tavenor, N. A., et al. *J. Amer. Chem. Soc.* **2017**, *139* (6), pp 2212–2215. Copyright 2017 American Chemical Society).

Like crystal form *a*, the asymmetric unit of crystal form *b* consists of a single α-helix, and the coiled-coil dimer is created by a crystallographic symmetry axis. The peptide folds are virtually identical in the two crystals, and the contacts that make up the lattice are composed entirely of coiled-coil hydrophobic interfaces and Cu^{2+}-bound Tpy residues in both. Despite these similarities, the lattices differ fundamentally in the packing between coiled coil units. In crystal form *b*, every carboxylate participating in a Tpy-Cu^{2+}-carboxylate motif comes from a Glu side chain on a neighboring peptide (Figure 56B,C). Two types of such interfaces exist in the lattice: one between two neighboring coiled coils (Figure 56B) and the other linking four chains in two coiled coils (Figure 56C). These Tpy-Cu^{2+}-Glu complexes generate a highly symmetrical lattice in the $P6_122$ spacegroup. From its absence in the lattice, we hypothesize the terephthalate may facilitate crystallization by acting as a stabilizer and leaving group for bound Cu^{2+} at the edge of the growing crystal since it played a key role in obtaining diffraction-quality single crystals.

Embedded in the lattice is an extended supramolecular polymer consisting of alternating peptide-peptide and metal-coordination interfaces (Figure 56D). This is generated through the complexes formed by two Tpy_7-Cu^{2+}-$Glu_{14'}$ creating a strong double cross-linked brace between coiled coils. This prevents rotation around the linker stabilizing the supramolecular polymer to form a lattice in contrast to the supramolecular polymers described in Chapter 2 which are singly cross-linked. Coordination site two (Tpy_{27}-Cu^{2+}-$Glu_{22''}$) connects the chain to a different symmetry-related copy than the first site generating the six-fold screw axis (Figure 57A).

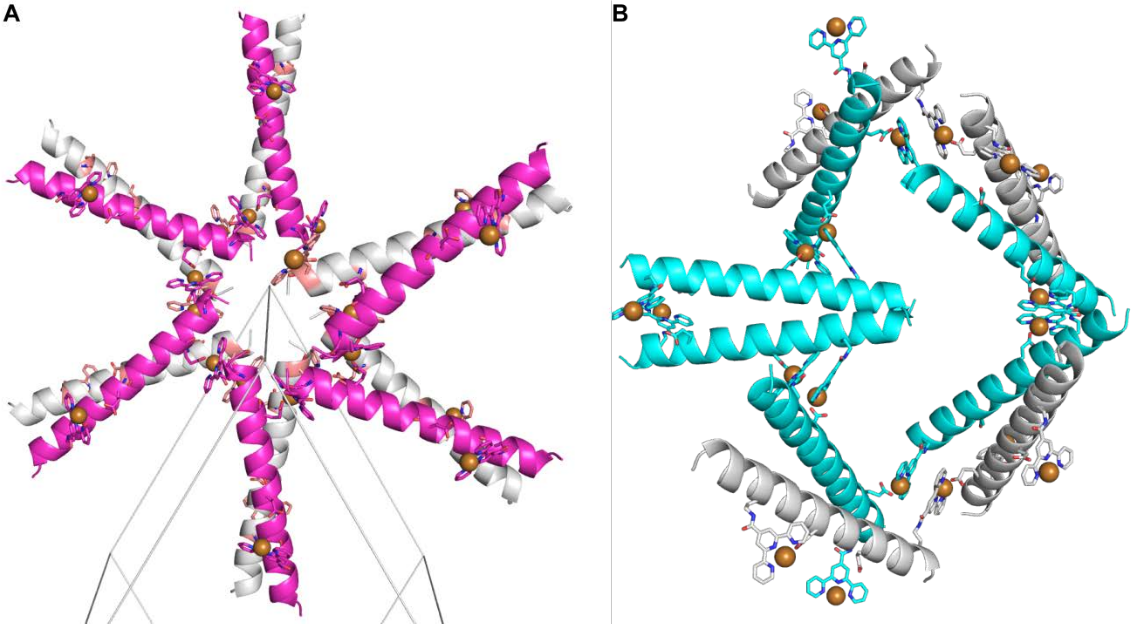


Figure 57. Lattice organization of peptide **33** crystal form *b*.

A) Six-fold screw axis (unit cell shown as gray box) of the asymmetric unit (magenta) generated by coordination site two (Tpy$_{27}$-Cu^{2+}-Glu$_{22''}$) and B) path through coordination bonds from one peptide chain back to its complementary partner (light blue). Cu^{2+} ions are display as brown spheres.

This leads to an interesting feature of crystal form *b*—every coiled coil is connected to every other chain through Tpy-Cu^{2+}-Glu coordination. This feature can be visualized by mapping a path through coordination bonds from one peptide chain back to its complementary partner in the dimer (Figure 57B). Since the asymmetric unit is a single chain, this implies the infinite connectivity of the framework analogous to MOFs derived from small molecule ligands.

The serendipitous result that Tpy-Cu^{2+}-Glu coordination could lead to the formation of highly ordered supramolecular assemblies of metallopeptides led to the design of peptides **35-38**. Based on the observations for dimeric coiled-coil peptide **33**, we sought to construct related assemblies from trimer and tetramer forming sequences using this Tpy-Cu^{2+}-Glu coordination motif. The four peptides **35-38** were created to form lattices using coiled coils with three-fold (**35** and **36**) or four-fold (**37** and **38**) symmetry (Figure 58).

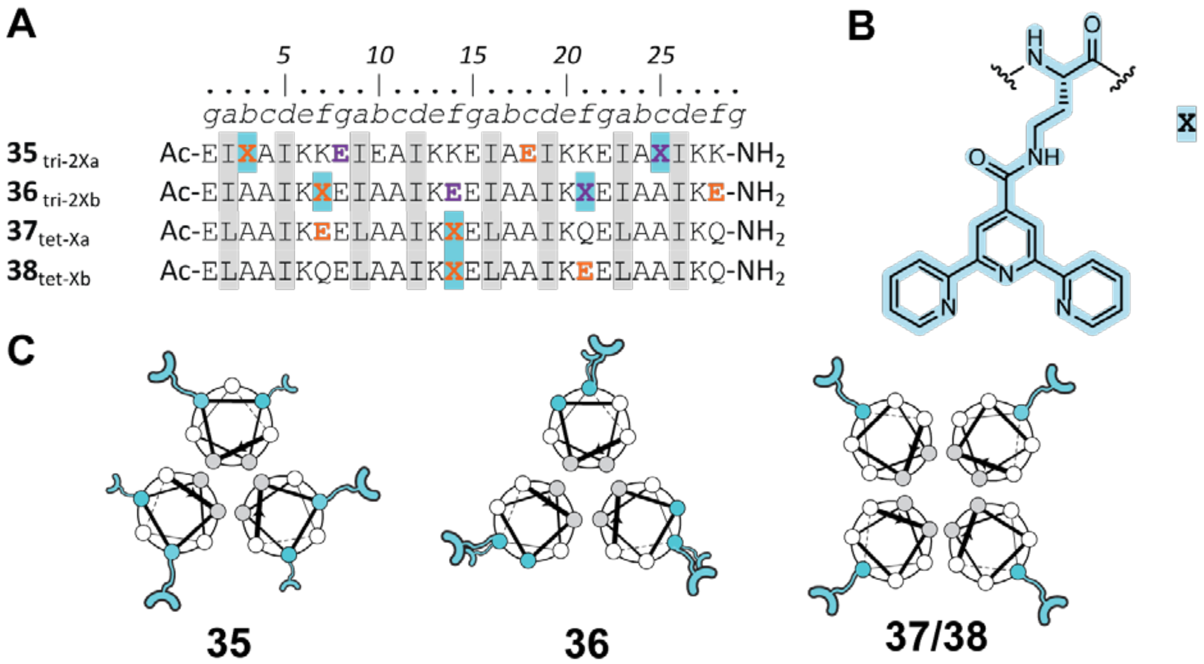


Figure 58. Design of tetrameric and trimeric coiled coils with engineered Tpy-Cu^{2+}-Glu interactions. A) Sequence of peptides **35-38** with the Tpy **X** side-chain highlighted in blue and hydrophobic core residues in gray; residues designed to chelate one another are in orange and purple. B) Structure of the Tpy **X** residue and C) peptide oligomerization state with Tpy side chain display.

All four peptides formed crystals, and three of these four resulted in structures that were solvable by molecular replacement. However, only two (**35** and **37**) were of high enough resolution for full refinement and structural elucidation.

For peptide **35**, the design hypothesis was that at *b*/*c* positions Tpy and Glu would interact to form a trigonal crystal lattice with trimers parallel and offset to one another. Thus, we incorporated Tpy at residue 3 (*b* position) and residue 25 (*c* position); additional Glu residues were introduced at positions 10 (*b* position) and 18 (*c* position) in an effort to provide the complementary coordinating group to complete each Tpy-Cu^{2+}-Glu motif. Summarizing the design of peptide **35**, the intended coordination mode was: Tpy$_{3}$(*b*)→Glu$_{18}$'(*c*) and Tpy$_{25}$(*c*)→Glu$_{10}$(*b*)' (heptad register of the residue in parentheses). Although the predominant interaction observed in the crystal, Tpy$_{25}$(*c*)→Glu$_{22}$(*g*)' differed from this design, it was nevertheless able to give rise to an ordered lattice.

Crystals of trimeric coiled-coil peptide **35** were grown in the presence of bis-Tris and Cu^{2+} and the structure was solved to 2.4 Å resolution (Figure 59, PDB ID 5U5B).

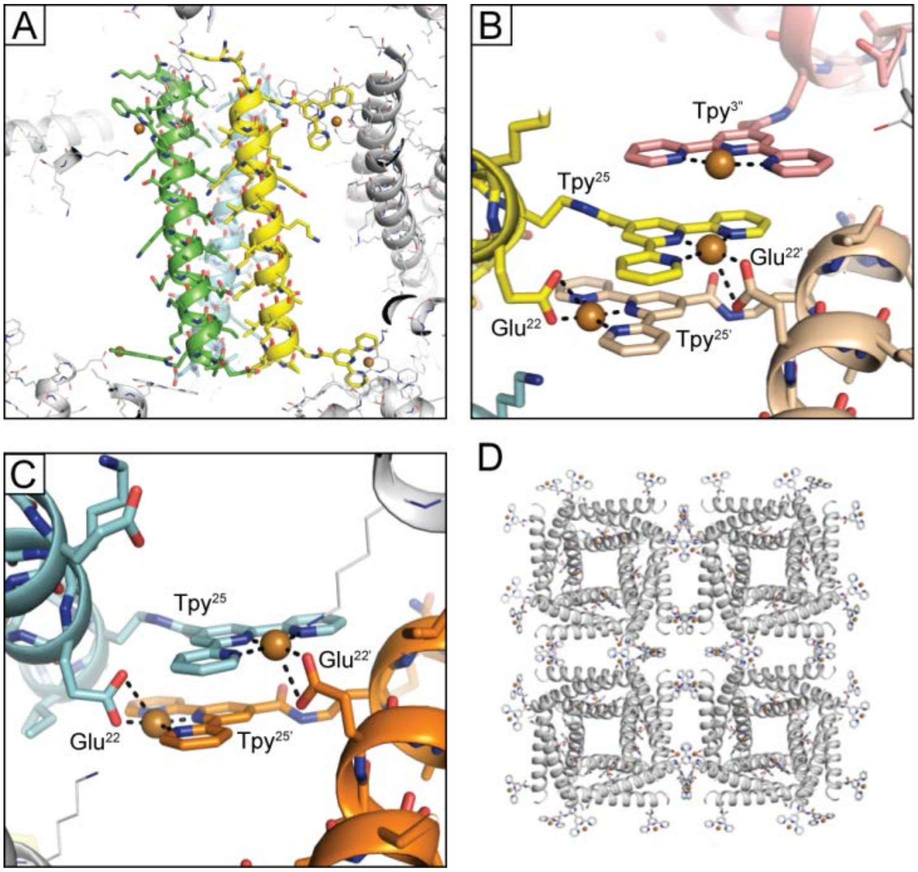


Figure 59. Structure of peptide **35**.

A) Crystal structure of peptide **35** (PDB ID 5U5B); B,C) Tpy-Cu^{2+}-Glu coordination between peptides; and D) 3-dimensional framework generated by this coordination (Adapted with permission from Tavenor, N. A., et al. *J. Amer. Chem. Soc.* **2017**, *139* (6), pp 2212–2215. Copyright 2017 American Chemical Society).

The asymmetric unit consists of the expected trimeric fold (Figure 59A), and the coiled coils are held together in the lattice entirely by Tpy side chains. Among four crystallographically independent interfaces, the same Tpy-Cu^{2+}-Glu coordination motif seen for peptide **33** form *b* (Figure 59C,B) is seen at most sites, alongside hydrophobic stacking of the aromatic systems minus Glu coordination (Figure 59B). These contacts generate an extended framework (Figure 59D) consisting of four-fold symmetric helical fibers (oriented along the 4_1 screw-axis).

Peptide **36** also contains a trimer-forming hydrophobic core and was designed to have alternating Tpy/Glu residues along the *f*-positions in order to facilitate a more open lattice than peptide **35**. It was expected that adjacent coiled coils would coordinate either 1) anti-parallel to one another to create sheets that would interface flush to one antother or 2) parallel, but offset, to create a Tpy/Glu zipper lattice where adjacent sheets would be coordinated together. We grew crystals of peptide **36**, however the best crystal (spots visible at ~3.4 Å resolution) degraded too quickly under synchrotron X-ray radiation to obtain a complete dataset (processable to ~6.5 Å resolution). Optimization has yet to yield a satisfactory crystal. If the design hypothesis is correct, it may be that the more open framework of the lattice is less stable and more easily perturbed by mechanical handling or environmental factors.

Peptide **37** was intended to form sheets through $Tpy_{14}(f) \rightarrow Glu_{7}(f)$ coordination with the coiled-coil tetramers anti-parallel and offset relative to one another. To decrease the likelihood of charge repulsion hindering self-assembly, the two other *f* position Lys residues in the starting sequence were replaced with Gln.

Crystals of peptide **37** were successfully grown in the presence of bis-Tris and Cu^{2+} with citrate and 1,6-hexanediol as additives, and the structure was solved to 2.1 Å resolution (Figure 60, PDB ID 5U5C).

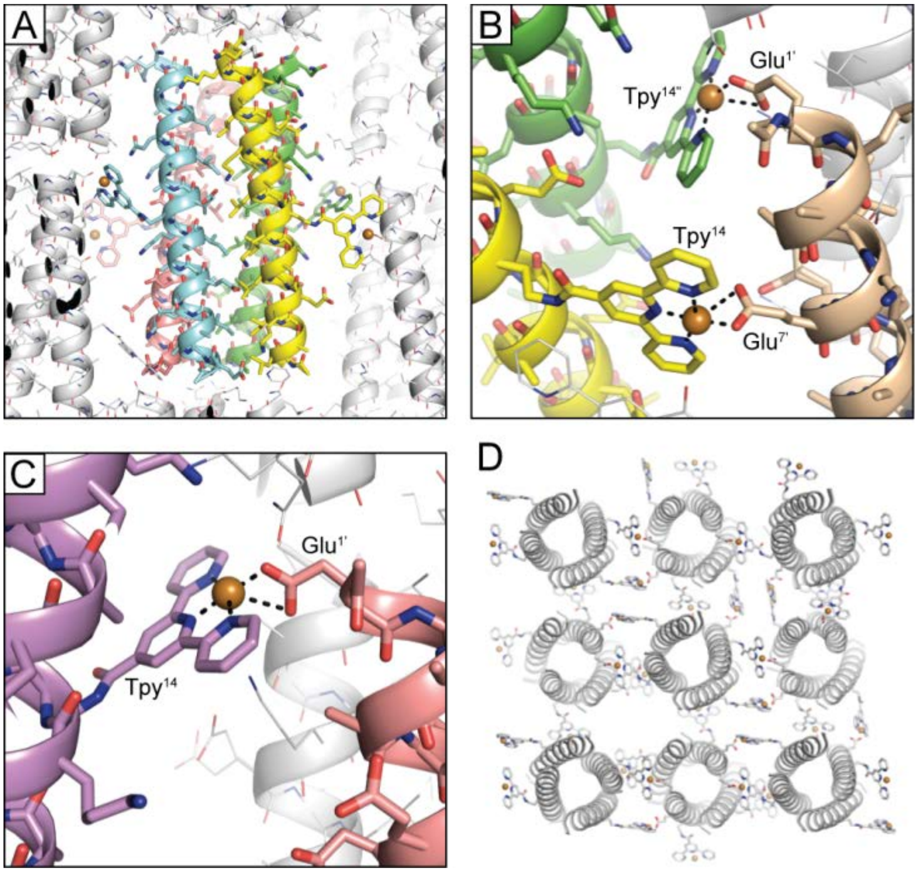


Figure 60. Structure of peptide **37**.

A) Crystal structure of peptide **37** (PDB ID 5U5C); B,C) Tpy-Cu^{2+}-Glu coordination between peptides; D) 2-dimensional net generated by the lattice (Adapted with permission from Tavenor, N. A., et al. *J. Amer. Chem. Soc.* **2017**, *139* (6), pp 2212–2215. Copyright 2017 American Chemical Society).

The asymmetric unit consists of six chains: four in a parallel coiled-coil tetramer, and the other two creating a second tetrameric assembly through a 2-fold symmetry axis. The tetramer fold is virtually identical to that of the variant without Tpy residues,[215] indicating the unnatural side chains direct assembly behavior independently of peptide folding. The single Tpy residue in each chain participates in a Tpy-Cu^{2+}-Glu interface with a neighboring tetramer in the lattice. The expected Tpy$_{14}$(*f*)→Glu$_{7}$(*f*) interaction was observed in two of six chains in the asymmetric unit; in the other four chains, Tpy$_{14}$(*f*) coordinated to Glu$_{1}$(*g*).

Although, the exact interacting residues differed from the initial design, the lattice did contain the expected overall assembly, consisting of coiled-coil tetramers antiparallel to closest neighbors and an infinitely propagating 2-dimensional tetragonal net (Figure 60D). These nets stack in layers via contacts between complementary (all N to C) peptide termini to form the third dimension of the lattice. Layer interfaces alternate pegs and holes facilitating a highly ordered lattice (Figure 61A).

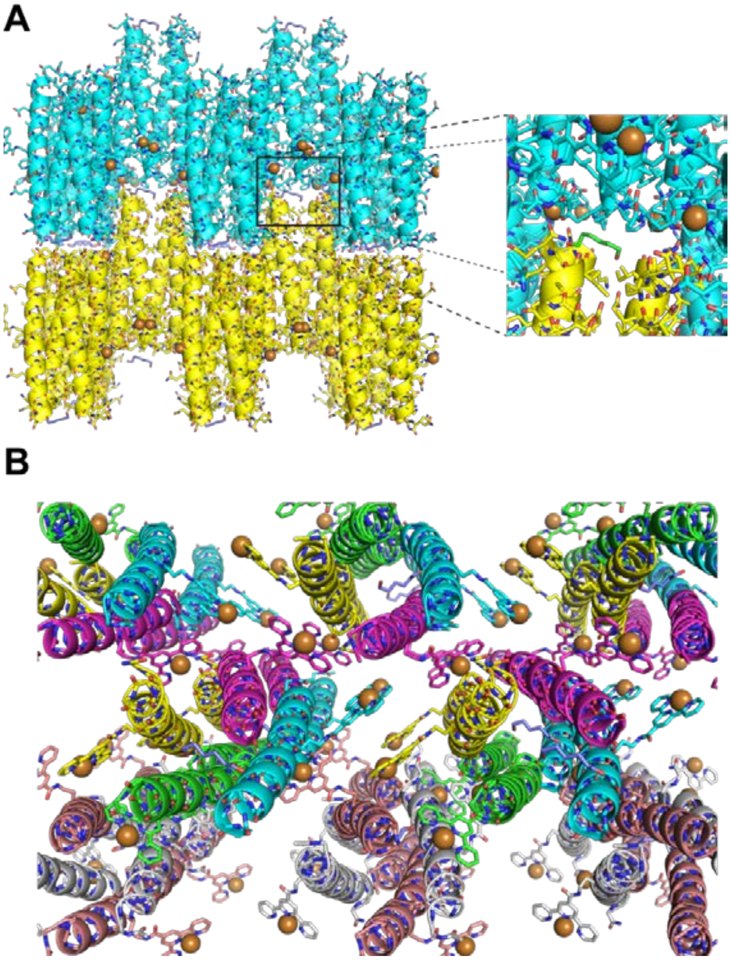


Figure 61. Lattice organization of peptide **37**.

A) Knobs into holes layering of peptide **37** with the interface stabilized by 1,6-hexanediol (green, inset). B) Patterning of the tetramers within each layer; the six crystallographically unique peptide chains are colored differently. Cu^{2+} atoms are displayed as brown spheres.

The interface between layers is stabilized by an ordered 1,6-hexanediol molecule (Figure 61A-inset). Within each layer, the two crystallographically unique tetramers pack in an AAB pattern of rows (Figure 61B). This patterning generates two different types of channels passing through the lattice. The alignment of the Cu^{2+} centers through these channels could be useful for studying charge transfer through the material. Combined, these interactions propagate the six chains of the asymmetric unit into the $P2_12_12$ lattice.

In contrast to peptide **37**, which was designed to form layer interfaces with alternating pegs and holes, peptide **38** was designed to form 2-dimensional nets with flush interfaces. This was to be achieved through coordination of $Tpy_{14} \rightarrow Glu_{21}$ of adjacent coiled coils to form a tetragonal lattice. It was anticipated that this design would be more compact than peptide **37** and contain a lower solvent content.

Although the crystal structure of tetrameric peptide **38** was unable to be fully refined, a molecular replacement solution (in space group $P4_222$) provided some insight (Figure 62). As intended, 2-dimensional sheets stack to form the third dimension of the crystal lattice by packing identical termini against one another (N to N and C to C, Figure 62B).

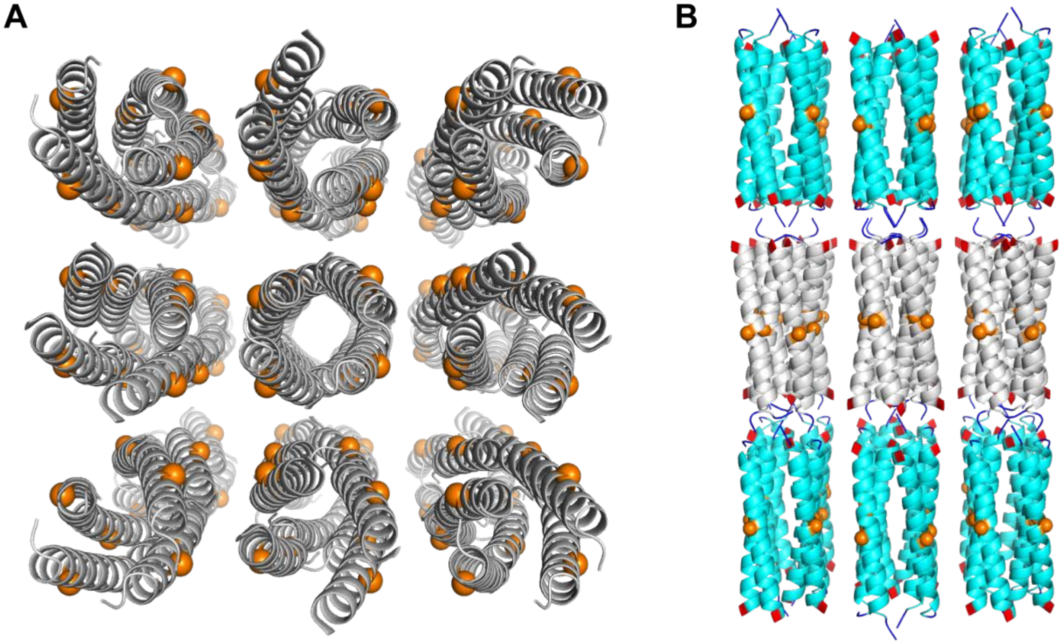


Figure 62. Crystal structure of peptide **38** with orange spheres representing the C_α of the Tpy residue: A) view down the axis of the tetramer and B) 2D parallel sheets which stack to form the third dimension of the crystal; N-termini are colored in dark blue while C-termini are colored in red.

Due to the low resolution of the diffraction (3.65 Å resolution), placement and coordination environment of the Tpy residues were too ambiguous to determine, and full structural refinement was not possible. Optimization of the crystallization conditions may yet yield a complete picture of this peptide's assembly architecture. It is likely that the flush layers of the crystal packing can more easily slide past one another giving rise to more disorder and the lower diffraction resolution. It may be possible to stabilize the lattice with an appropriate additive similar to the role 1,6-hexanediol plays at the interface between nets of peptide **37**; efforts along this vein are ongoing.

Although crystals were successfully grown with peptides **33-38**, only a handful of these were able to be fully refined to elucidate the structure. CD scans and thermal melts of peptides with fully refined structures indicate that peptides were fully folded under the conditions of the crystallization experiment (Figure 63).

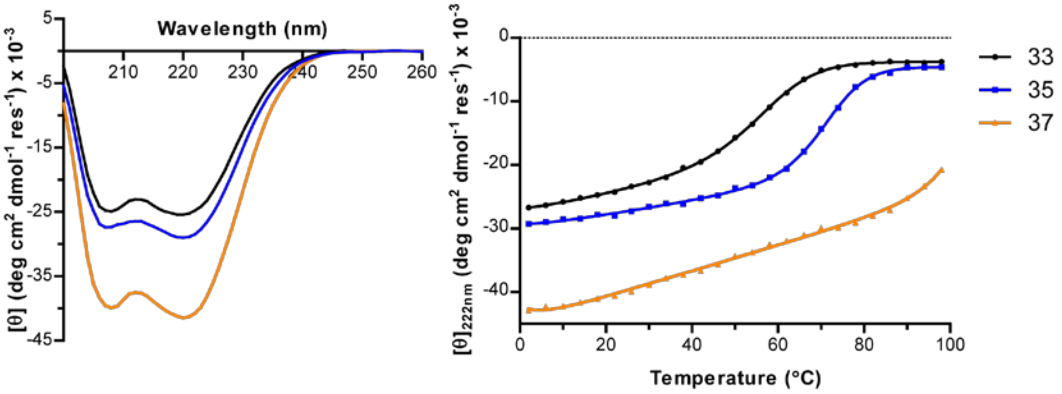


Figure 63. CD scans (left) and melts (right) of peptides **33**, **35**, and **37** at 25 µM in aqueous HEPES buffer (10 mM, pH 7) with 100 µM $CuCl_2$. (Adapted with permission from Tavenor, N. A., et al. *J. Amer. Chem. Soc.* **2017**, *139* (6), pp 2212–2215. Copyright 2017 American Chemical Society).

Collectively, these results provide a glimpse into a new class of highly ordered supramolecular architectures that can be generated through the combination of coiled-coil interfaces and terpyridine-Cu^{2+} coordination. Crystallization experiments of previous peptide designs were used to inform later designs finally leading to a series of highly-ordered crystalline assemblies. Iterating this design process further should enable the design of highly ordered metallopeptides with predictable supramolecular architectures. Due to the difficulty of crystallizing many of the sequences; solution-phase studies of a subset of the peptides was also undertaken in an attempt to probe the assembly of ordered soluble architectures.

4.3 SOLUTION-PHASE ASSEMBLY

Solution-phase experiments were undertaken to understand the role of Cu^{2+} in controlling the assembly of the coiled-coil based supramolecular metallopolymers. The peptide series **28-32** (*vide supra*, Figure 52) was designed with this goal in mind. These peptides all contain a single Tpy residue at position 14 (helix register *f* in the second heptad), a solvent-exposed site with a projection approximately normal to the coiled-coil axis. The peptides differ only in their hydrophobic cores giving rise to different oligomerization states (dimer, trimer or tetramer) and the length of the Tpy side-chain (**X** < **Z**). Although not observed in the crystalline assemblies (Section 4.2), it was hypothesized that the bis-Tpy complex may still form in solution in the presence of 0.5 equiv of Cu^{2+}.[210] Experiments were designed to determine how peptide sequence and Cu^{2+} stoichiometry affect the location of metal junctions within the assemblies and the preferred assembly modes of the peptides.

These structural properties were probed by a combination of biophysical techniques, similar to the approach used to characterize the supramolecular polymers in Chapter 2. As the simplest system, characterization initially focused on the dimeric peptides **29** (with shorter **Z** side-chain) and **32** (with longer **X** side-chain) and then expanded to include the more complex oligomerization states (peptides **28**, **30**, and **31**). The effect of Cu^{2+} on assembly properties was studied using both sub-stoichiometric and excess amounts (ratio of 0.5 and 1.5 equiv of Cu^{2+} relative to peptide, respectively). These ratios were determined based on the change in coordination environment observed in an EPR titration of Cu^{2+} with a side-chain analogue (ligand **40**, *vide infra* Figure 66).

Experimental evidence (*vide infra*) obtained for the supramolecular peptide assemblies support the hypothesis that a mixture of assembly modes exists in solution. A simplified model of this equilibrium for the dimer peptide is presented below (Figure 64).

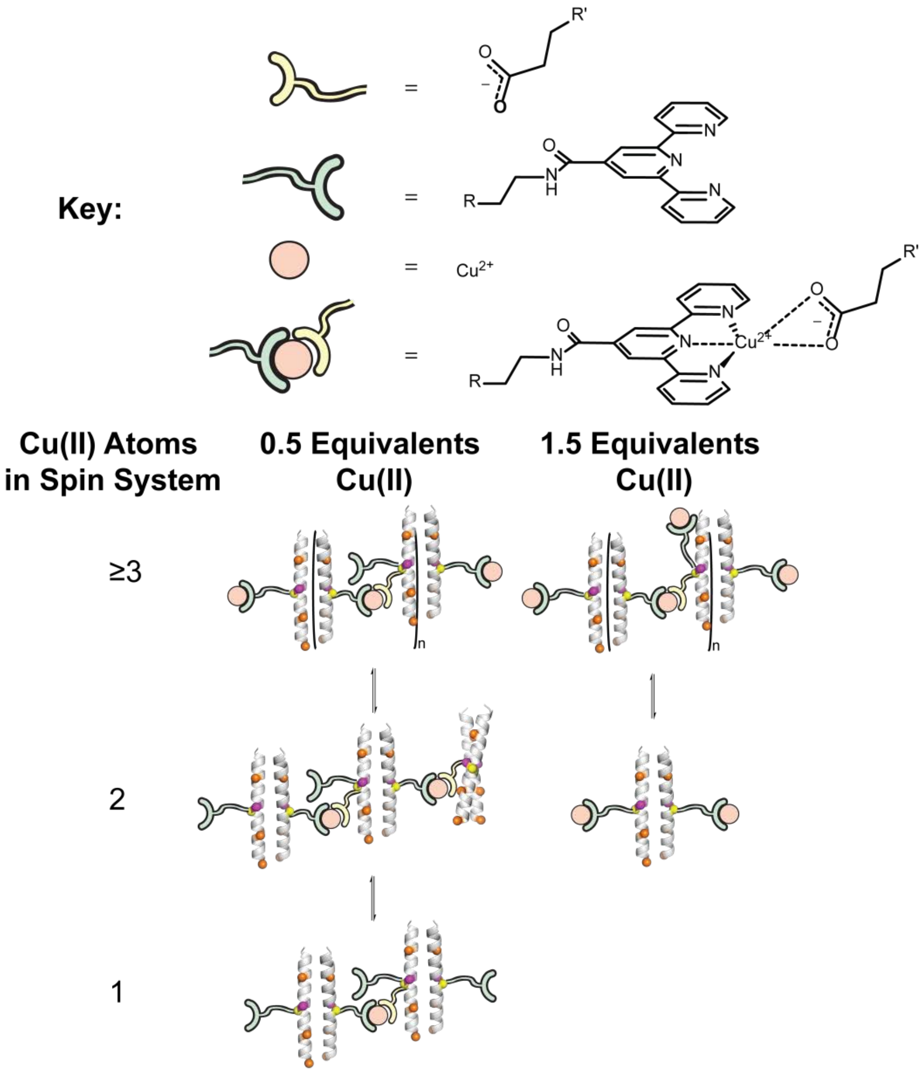


Figure 64. Schematic of dimeric peptide assembly with 0.5 or 1.5 equivalents of Cu^{2+}.

C$_{\alpha}$ atoms of important residues are highlighted as colored spheres: Tpy$_{14}$ (yellow), Glu (orange), and Glu$_{15}$ (purple).

This model assumes that the Cu^{2+} coordination is either in the form of Tpy-Cu^{2+}-Glu or Tpy-Cu^{2+}-(H$_2$O)$_n$ based on the EPR data (*vide infra*, Section 4.3.1). The assemblies shown in Figure 64 do not represent the full array of possibilities since there are four different Glu residues that can coordinate with Tpy-Cu^{2+} to form a ternary complex; a limitation of the current peptide designs. At 1.5 equivalents Cu^{2+}, it is assumed that every Tpy site is occupied based on the strength of the binding association between Tpy and Cu^{2+}.[210]

A combination of analytical methods was used to determine which assembly mode is likely to predominate. Dynamic light scattering was used to measure relative particle sizes. CD scans and melts probe the helicity and thermal stability of the coiled coil quaternary structure. Continuous wave (CW) and 2D pulsed Hyperfine Sub-level CORrElation (HYSCORE) EPR measurements probed the coordination environment around the Cu^{2+} ions. Finally, DEER spectroscopy was used to measure distances between Cu^{2+} centers in supramolecular systems and determine the relative local concentration of Cu^{2+} ions.

4.3.1 Cu^{2+} Coordination Environment Determined by EPR

We considered four possible coordination modes for the interaction of Cu^{2+} with peptides containing both Tpy and Glu residues (Figure 65).

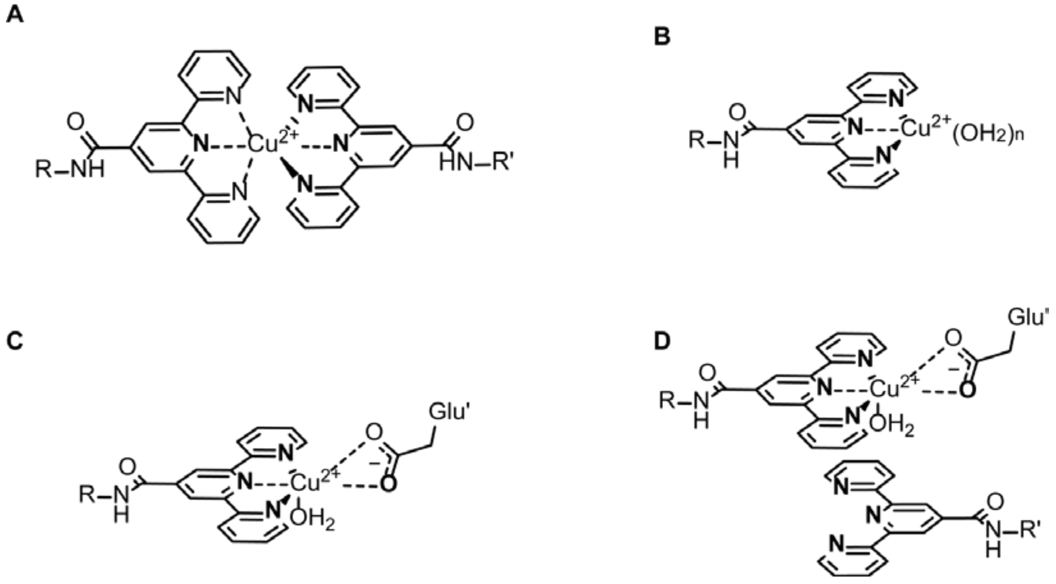


Figure 65. Possible coordination states of the Tpy-Cu^{2+} complex:

A) bis-Tpy-Cu^{2+}, B) hydrated mono-Tpy-Cu^{2+}, C) Tpy-Cu^{2+}-Glu', and D) Tpy-Cu^{2+}-Glu' with a distal non-chelating Tpy' (primes indicate another peptide chain).

Consistent with the crystalline assemblies, EPR titration of ligand **40** with Cu^{2+} and HYSCORE measurements suggests the bis-Tpy complex is not populated in the peptides (*vide infra*). This leaves the mono-complex where the empty ligand sites are occupied by water or Glu—the Cu^{2+}-bound Tpy may π-stack with other Tpy residues, as seen in the crystalline assemblies. Note that the immediate coordination environment of states C and D are identical; D merely contains a nearby Tpy residue which hydrophobically stacks with the one within the immediate coordination environment. A combination of continuous wave and pulsed EPR techniques were used to distinguish between the four possiblities.

Continuous wave (CW) EPR (Figure 66) indicates a shift in Tpy-Cu^{2+} binding environments in high vs. low equivalent samples.

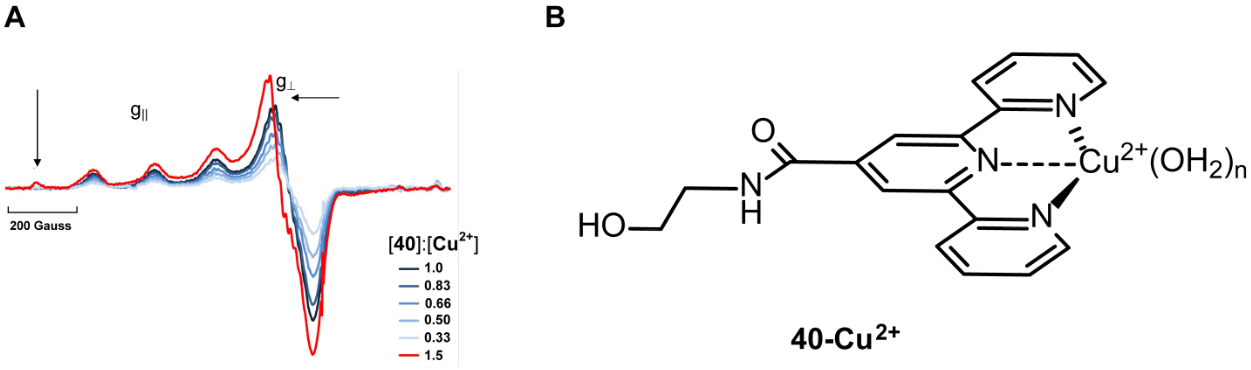


Figure 66. Coordination of Cu^{2+} with ligand **40**.

A) CW-EPR titration of ligand **40** (300 μM) with $CuCl_2$ in aqueous NEM buffer (40 mM pH 7.4); arrows indicate the shift in the spectrum upon reaching 1.5 equiv Cu^{2+}. B) Proposed coordination complex of Cu^{2+} and ligand **40** from 0.33-1.0 equiv Cu^{2+}.

The hyperfine and g-tensor parameters of the 80 K Cu^{2+} CW spectrum (extracted via simulation) are very sensitive to the immediate coordination environment. The spectra obtained for all samples of complex **40** are consistent with the lineshape of a Type II copper center (square planar, 4 equatorial ligands, 2 axial).[216] In type II Cu^{2+} centers, the unpaired electron only interacts with the planar $d_{x^2-y^2}$ orbital and contributions from the axial ligands are negligible.[216] From 0.33-1.0 equiv of Cu^{2+} with ligand **40**, the hyperfine and g-tensors of the EPR spectrum are consistent with a mono-coordinated terpyridine complex with 1 equatorial oxygen.[217] When 1.5 equiv of Cu^{2+} is reached, a new spectral component appears and the $g_{\perp}$ peak shifts downfield along with the presence of a new peak in the $g_{||}$ region. These features both imply a decrease in N coordination.[218] The broadening of the $g_{\perp}$ peak is indicative of a superposition of the two components' spectra; indicating the presence of two Cu^{2+} coordination rather than just one.

This second coordination state is consistent with Cu^{2+} coordinated to four equatorial oxygens (water). The presence of free Cu^{2+} was at first ruled out due to the presence of the NEM buffer which is well-known to precipitate out unbound Cu^{2+} ions as $[Cu(OH)_2]_n$ rendering them EPR inactive.[219] However, spectral simulations of a two component system containing the Tpy-Cu^{2+} monocomplex and free $CuCl_2$ in water are able to reproduce the spectra seen at 1.5 equiv Cu^{2+} (Figure 67).

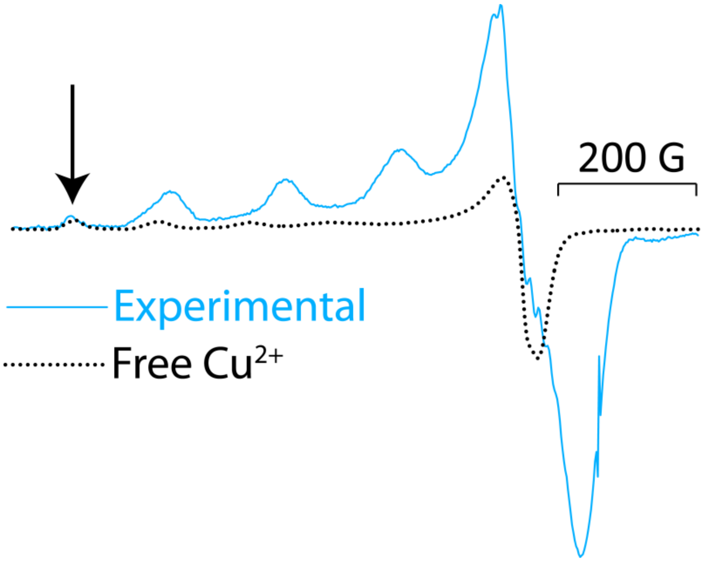


Figure 67. Spectrum of free Cu^{2+} compared to experimental spectrum of ligand **40** with 1.5 equivalents of Cu^{2+}.

4.3.1.1 CW EPR of Peptides

The spectra obtained for samples of all four peptides (**28**, **29**, **31**, and **32**) with 0.5 and 1.5 equivalent of Cu^{2+} are consistent with the lineshape of a Type II copper center (4 equatorial ligands, 2 axial).[216] The hyperfine and g-tensor values for the Cu^{2+} spectrum of the 0.5 equivalents sample for all peptides are consistent with 3 nitrogens and 1 oxygen equatorially bound (Figure 68).[216]

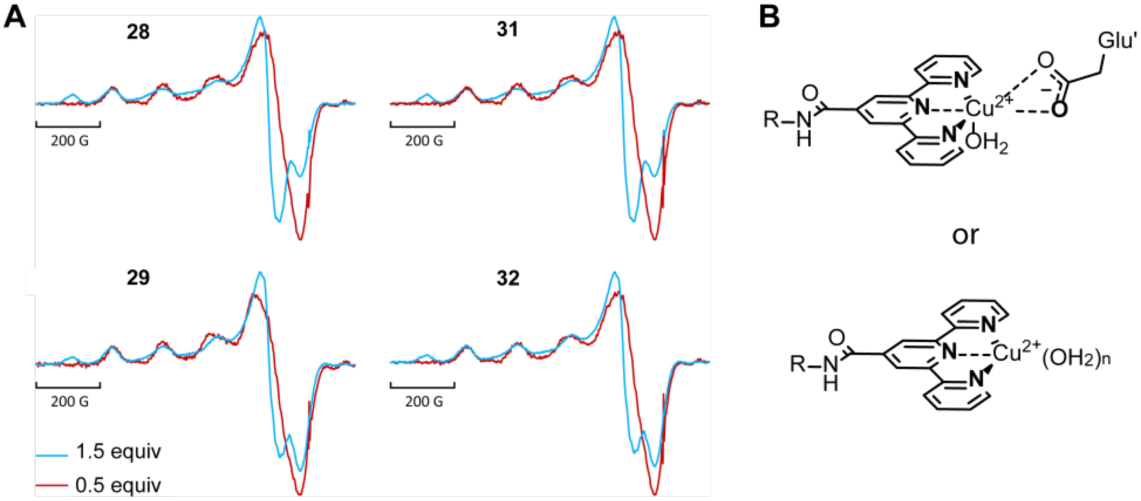


Figure 68. Proximal coordination of Cu^{2+} in Tpy-modified peptides.

A) CW-EPR spectra of peptides **28**, **29**, **31**, and **32** with either 0.5 (red) or 1.5 (blue) equiv of Cu^{2+}. B) Proximal coordination states supported by the data.

This is consistent with Tpy-Cu^{2+}-Glu or Tpy-Cu^{2+}-$(H_2O)_n$ coordination;[217] both of which would be expected to result in similar spectra. In all samples, the same hyperfine and g-tensor values also fit the Cu^{2+} CW spectrum of the Tpy ligand alone (Figure 66) with up to a 1:1 ratio of Cu^{2+}:Tpy and formation of the bis-Tpy-Cu^{2+} is inconsistent with the observed CW data.[217]

Upon addition of excess Cu^{2+}, a second component corresponding to free Cu^{2+} is observed, as seen for the model ligand (Figure 69).

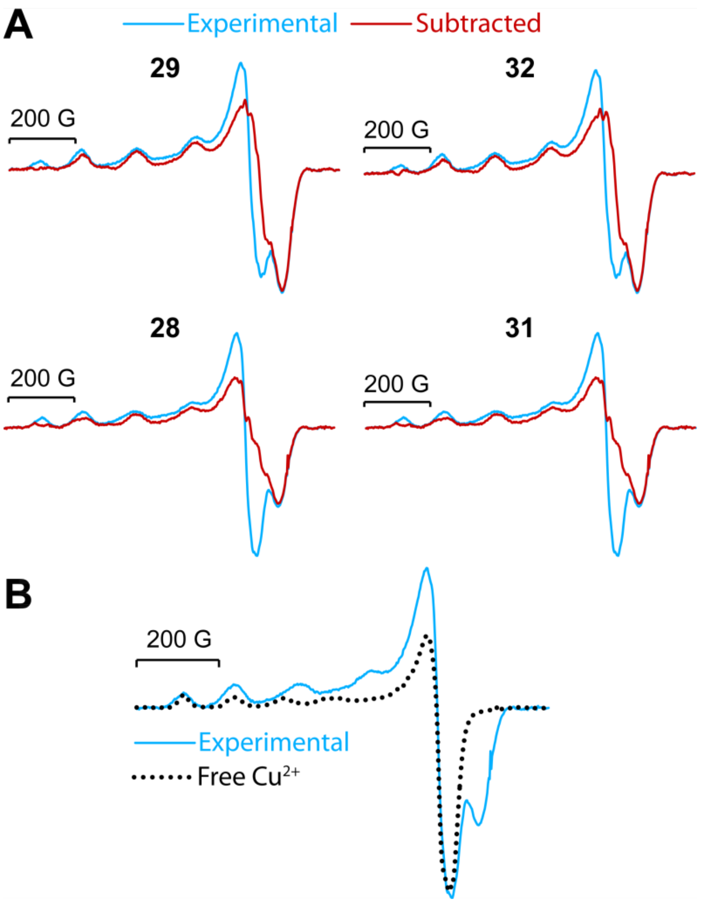


Figure 69. Spectral simulations of excess Cu^{2+} peptide samples compared to experimental data. A) Simulated CW spectra of peptides **28**, **29**, **31**, and **32** with free Cu^{2+} subtracted out compared to experimental data ; B) Free Cu^{2+} spectrum compared to experimental spectrum of peptide **28** with 1.5 equivalents of Cu^{2+}.

Overall, results are consistent among all four peptides with the only variation being between the ratio of the two components at 1.5 equivalents Cu^{2+} when altering the oligomerization state; the dimers have a higher proportion of the second component than the tetramers (Figure 68A). Based only on the CW data, any of the supramolecular assembly models are possible since it is not possible to distinguish between Tpy-Cu^{2+}-Glu or Tpy-Cu^{2+}-$(H_2O)_n$ coordination (Figure 68B).

4.3.1.2 HYSCORE EPR of Peptides

2D pulsed Hyperfine Sub-level CORrElation (HYSCORE) EPR spectroscopy reveals the atoms coordinated distally to the Cu^{2+} center. The FFT HYSCORE spectra of all samples shows strong cross-peaks about 14 MHz (Figure 70), indicating strongly coordinating distal solvent hydrogens.[220]

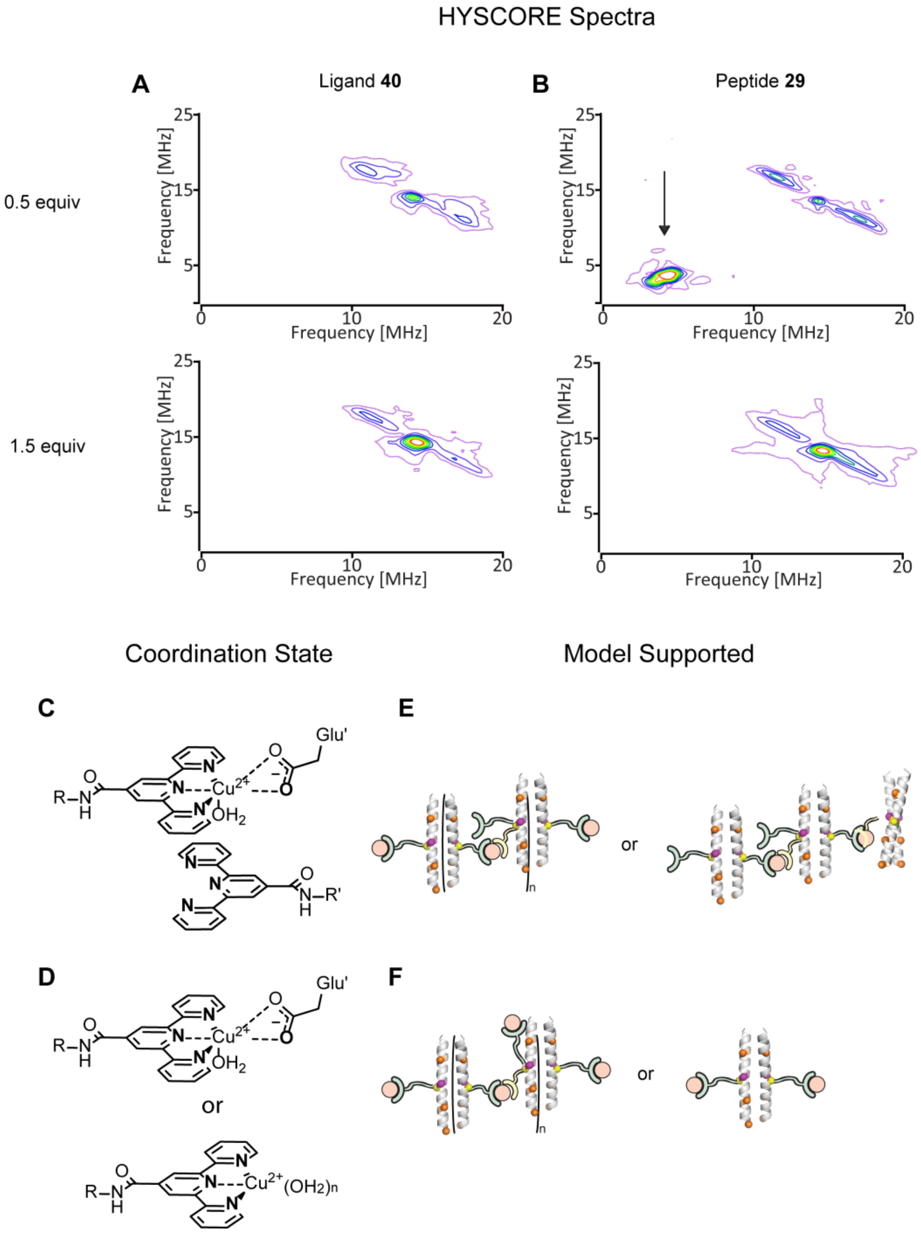


Figure 70. Distal coordination of Cu^{2+} with Tpy-modified peptide ligands.

HYSCORE spectra of ligand **40** (A) and peptide **29** (B) with 0.5 (top) and 1.5 (bottom) equiv of Cu^{2+}; spectra of peptides **28**, **31**, and **32** are identical to that of **29**. The arrow indicates a the signal for the presence of a distal nitrogen. Coordination states (C, D) and plausible supramolecular assembly model supported by the HYSCORE spectrum for 0.5 (C, D, E) and 1.5 (D, F) equivalents of Cu^{2+}.

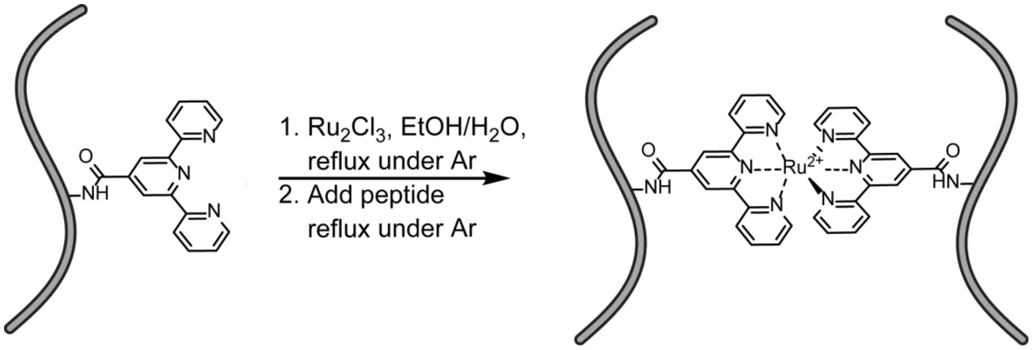
This strong H signal would not be expected for a bis-Tpy-Cu^{2+} complex lending support to the hypothesis that only mono-Tpy (Tpy-Cu^{2+}-$(H_2O)_n$ or Tpy-Cu^{2+}(H_2O)-Glu) coordination is present in all samples. Both mono-Tpy coordination modes would give rise to this signal due to the presence of coordinating water ligands and are likely in equilibrium with one another; the crystal structure of peptide **35** contained both coordination states (*vide infra,* Section 4.2) supporting this hypothesis.

In addition to features expected from distal hydrogen, peaks stemming from distal nitrogen are present for the peptides with 0.5 equivalents Cu^{2+}, but disappear upon addition of 1.5 equivalents Cu^{2+}. This peak is not present in the corresponding sample of the small molecule analogue, ligand **40**, implying this is a unique feature of the peptides. As evident from crystal structures (Section 4.2) of similar peptides, it is possible that the Tpy-Cu^{2+}(H_2O)-Glu coordination is coupled with another Tpy residue interacting through hydrophobic stacking in the 0.5 equivalent case. The close proximity of the two stacked Tpy could give rise to the distal nitrogen signal. Due to the ambiguity of distinguishing between coordination Glu or H_2O this is the strongest evidence that the peptides are forming supramolecular polymers.

The absence of the nitrogen signal with the addition of 1.5 equiv of Cu^{2+} suggests that hydrophobic stacking of a Tpy ligand to a Tpy-Cu^{2+}(H_2O)-Glu coordination is no longer the preferred mode of assembly, possibly due to electrostatic repulsion between the positively charged Cu^{2+} centers. This could either be a result of a change in the morphology of the supramolecular assembly or a lower degree of supramolecular polymerization.

4.3.2 Particle Sizing with Dynamic Light Scattering

DLS can be used to measure the relative size of particles in solution. Larger particles exhibit scattering events that have a longer correlation time, while smaller particles exhibit shorter correlations.[221] Mathematical models can be used to derive the particle size from the resulting correlation data. However, current theoretical models break down when the materials are heterogeneous and non-spherical (two characteristics of the system examined here).[222-224] With this caveat in mind, we have presented the raw correlation data as a qualitative measure of supramolecular assembly properties. Apparent hydrodynamic diameters are included only as a means of relative comparison. Peptide **29** was also cross-linked with Ru^{2+} to form a bis-Tpy-Ru^{2+} complex resulting in subunit **43** (Scheme 5).



28 or **29**

42 or **43**

Scheme 5. Cross-linking reaction of peptide **28** or **29** with Ru^{2+} to produce subunit **42** or **43**, respectively.

Unlike the copper complexes, bis-Tpy Ru(II) complexes do not readily exchange and are even stable under HPLC.[225] Thus, this subunit acts as the reference point for the formation of the supramolecular polymers from an inert coordination complex.

Dynamic light scattering (DLS) measurements of the dimeric peptides **29** and **32** suggest that supramolecular assemblies are created in small quantities upon addition of Cu^{2+} (Figure 71).

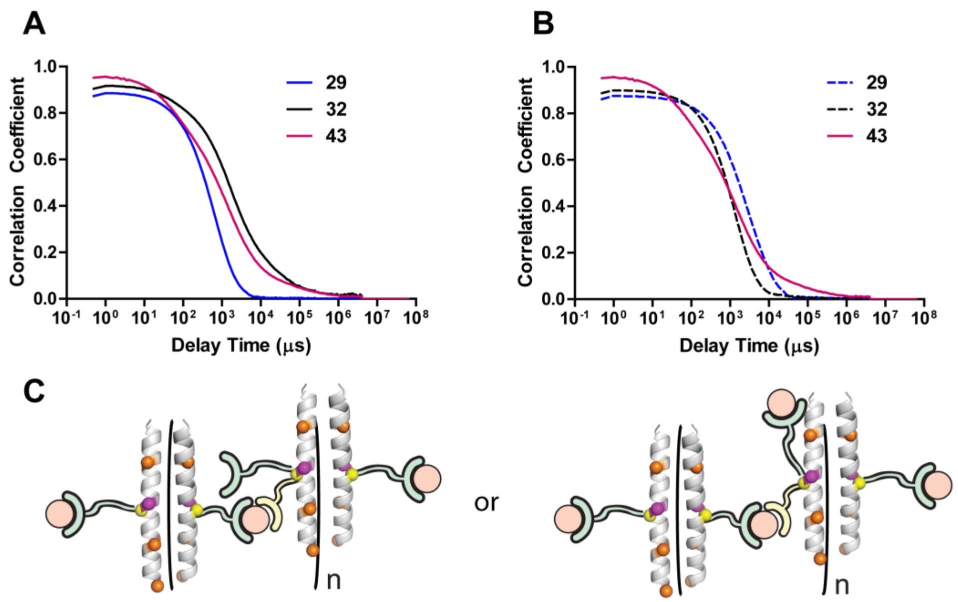


Figure 71. DLS correlation curves of dimeric peptides **29** (blue) and **32** (black) in the presence of 0.5 (solid, A) or 1.5 equiv (dashed, B) Cu^{2+}. Assembly of Ru^{2+} cross-linked peptide **43** is shown as a reference point for both cases. All measurements were performed with 300 μM (peptides **29** and **32**) or 150 μM (subunit **43**) in aqueous 10 mM HEPES buffer (pH 7) with 0.1 M NaCl. C) Cartoon models of the supramolecular polymer formed by the peptides in the presence of Cu^{2+}.

Peptides **29** and **32** were measured in the presence of 0.5 and 1.5 equivalents of Cu^{2+}. Apparent average hydrodynamic diameters were 186±7 nm (**29**) and 600±100 nm (**32**) with 0.5 equivalents of Cu^{2+}. Peptide **29** increased to 830±60 nm with 1.5 equivalents of Cu^{2+} while peptide **32** decreased to 390±20 nm. Peptides **29** and **32** in the absence of copper resulted in no correlation signal. The apparent size of assemblies generated by subunit **43** was 300±60 nm.

Consistent with the DEER data (*vide infra*, Section 4.3.3), in the presence of 0.5 equivalents of Cu^{2+} peptide **32** has a larger particle size than peptide **29**. This difference in particle size is due to three possible effects related to peptide stability and linker length. First, **32** contains the longer **Z** side chain; increasing the side-chain length intuitively increases assembly size by increasing the space between peptide monomers. Second, even though the side chain is longer, with 0.5 equivalents of Cu^{2+} peptide **32** has less conformational freedom than peptide **29** (see Section 4.3.3); an effect which is known to increase particle size in an analogous system (see Chapter 2).[158] Third, although the length of the Tpy side-chain itself does not affect the coiled-coil fold, it does influence the extent of destabilization due to complexation with Cu^{2+} (Figure 72).

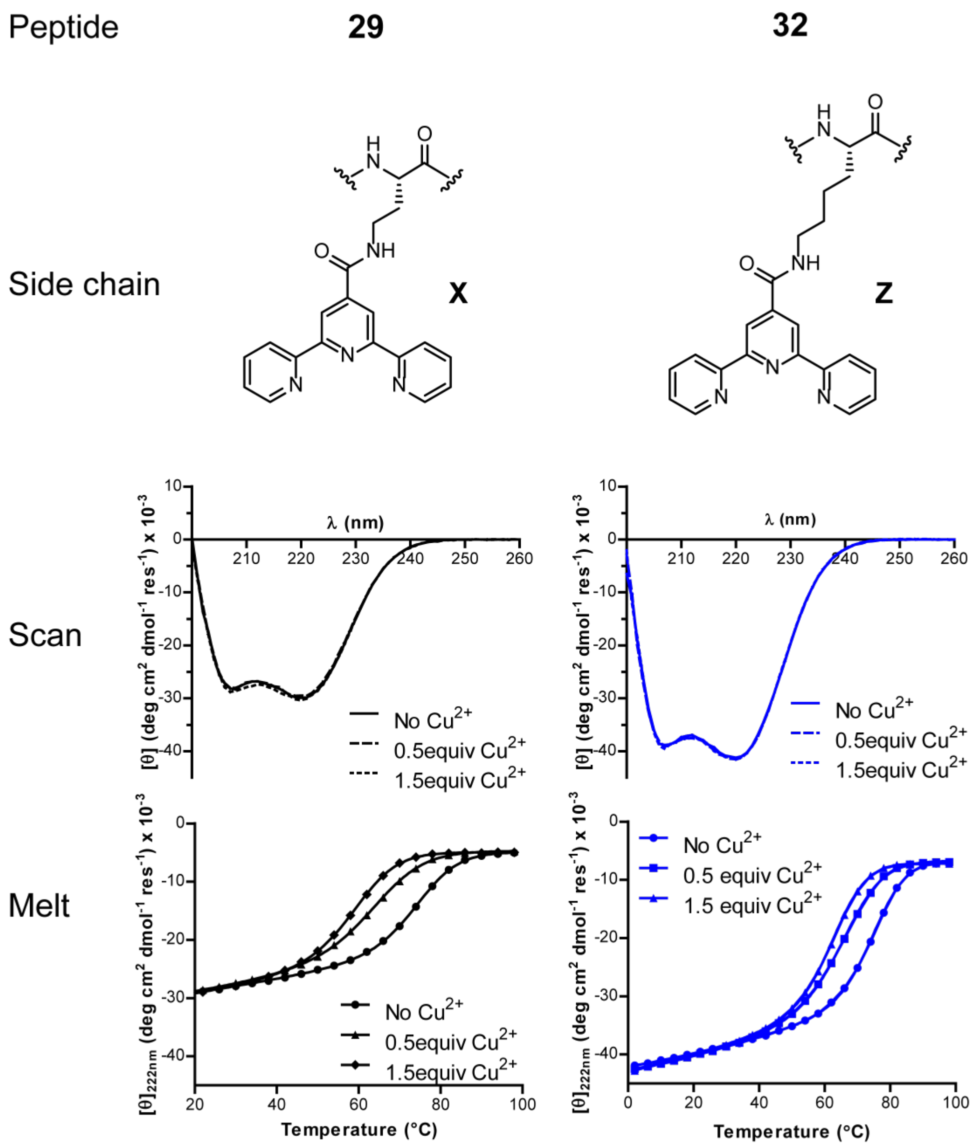


Figure 72. CD scans and melts for dimeric peptides **29** (black) and **32** (blue) at 100 µM in aqueous HEPES buffer (10 mM, pH 7) with 0, 0.5 or 1.5 equivalents of $CuCl_2$.

In the absence of Cu^{2+}, both peptides have a similar midpoint in the thermal melting curve (T_m): 74.4 ± 0.1 °C (**29**) and 74.7 ± 0.1 °C (**32**). Upon addition of 0.5 equivalents of Cu^{2+}, the midpoint of the thermal melting curve decreases for both peptides compared to corresponding samples lacking Cu^{2+} (ΔT_m = -9.0 ± 0.2 °C and -10.6 ± 0.1 °C for **32** and **29**, respectively, Figure 72). The T_ms of peptide **32** in the presence of Cu^{2+} are higher at 65.7 ± 0.2 °C (0.5 equiv Cu^{2+}) and 62.5 ±

0.1 °C (1.5 equiv Cu^{2+}); whereas for peptide **29** the T_ms are 63.8 ± 0.2 °C (0.5 equiv Cu^{2+}) and 58.9 ± 0.1 °C (1.5 equiv Cu^{2+}). The higher folded stability of peptide **32** in the presence of Cu^{2+}, though slight, could contribute to the increase in the assembly size of peptide **32** compared to peptide **29**.[161]

Upon addition of excess Cu^{2+}, meant to favor monovalent Tpy coordination, assemblies of peptide **29** increased in size while those of peptide **32** decreased in size (Figure 71). This can likely be explained by the difference in conformational flexibilty between the two linkers. The greater degree of flexibility of the Lys allows peptide **32** to collapse into a more compact material similar to the effect seen with the EDA linker in Chapter 2.[158]

Replacing the Cu^{2+} ion with Ru^{2+} to create a covalently bound bis-Tpy cross-linked peptide results in a supramolecular polymer larger than that with Cu^{2+} (Figure 71B). Surprisingly, the T_m of this cross-linked Ru^{2+}subunit is much lower than the analogous Cu^{2+} cross-link (47.5 °C vs. 63.8°C, Figure 73).

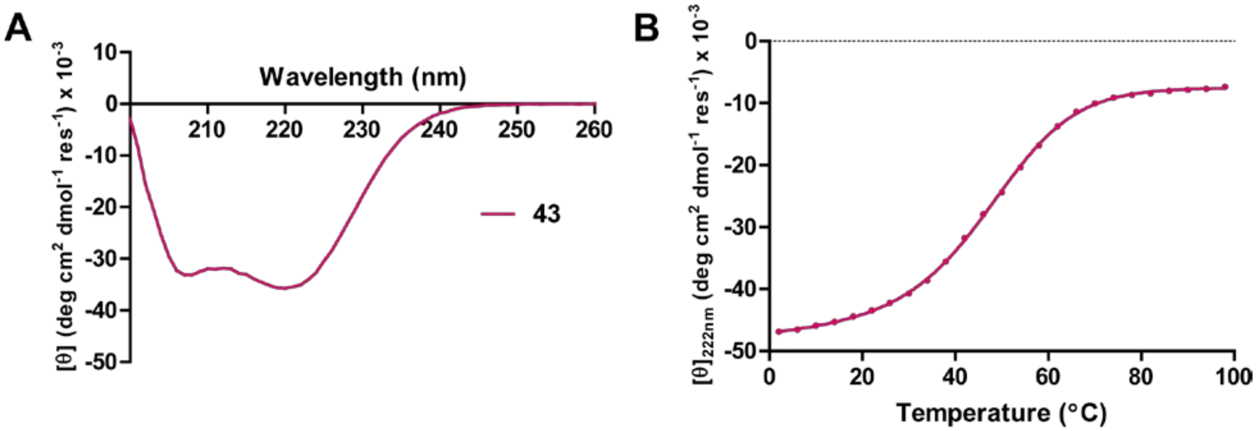


Figure 73. CD scan (A) and melt (B) of subunit **43**

at 50 µM in aqueous HEPES buffer (10 mM pH 7).

This is likely either due to the fact that 100% of the peptides are cross-linked in the case of Ru^{2+} or hints that the coordination environments of the two metals are not isomorphous as would be expected if Cu^{2+} formed solely bis-Tpy complexes, a conclusion also supported by the EPR data. The high degree of coiled-coil destabilization may also explain why assemblies based on the Ru^{2+} cross-link are smaller than those formed by **32** in the presence of Cu^{2+}. The lower molar ellipticity observed in the CD scan (measured pre-melt at 20 °C) vs. the melt at the same temperature may be due to kinetic trapping which is stabilized upon cooling to 2°C. This is supported by the presence of visible aggregates post-melt.

4.3.3 Assembly Morphology

Double Electron-Electron Resonance (DEER) measurements obtained from the materials also demonstrate formation of a supramolecular metallopolymer (Figure 74A, B).

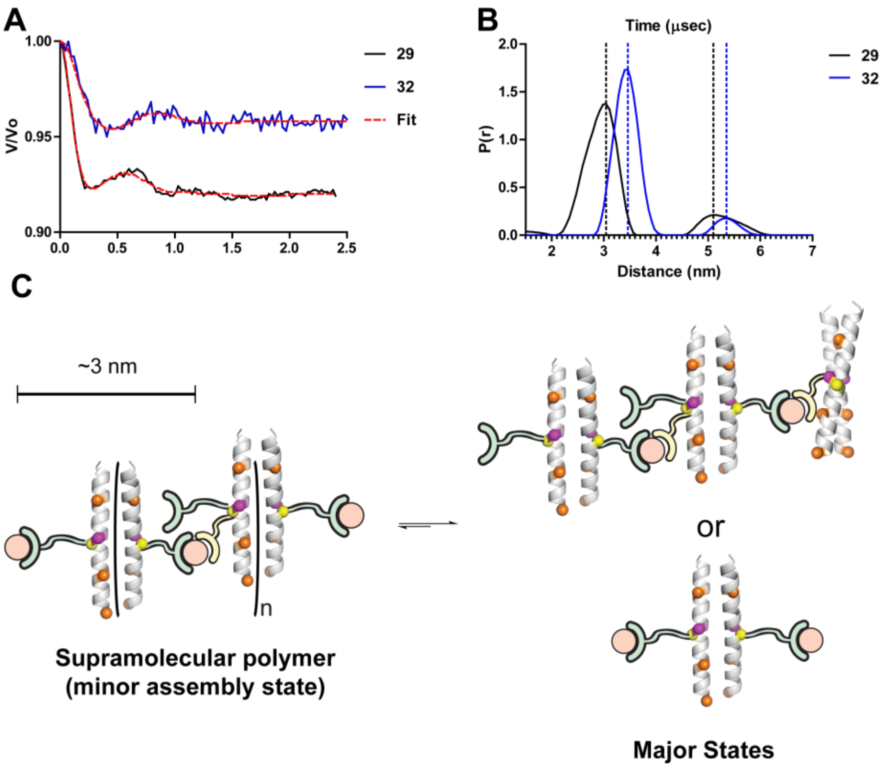


Figure 74. DEER measurements of dimer peptides in the presence of 0.5 equiv Cu^{2+}.

A) Time domain spectrum and resulting distance distribution (B) for dimeric peptides **29** (black) and **32** (blue); dashed vertical mark the most probable distance for each peak. C) Equilibrium between assembly modes based on the DEER data.

Upon addition of 0.5 equiv Cu^{2+}, peptide **29** assembles into supramolecular oligomers with a most probable distance between the linkers of 3.0 nm with a second peak at 5.1 nm (Figure 74B). Since it is not possible for the copper ions on the same coiled coil to be this far apart, we infer that there must be at minimum of three copper ions in the system to produce this signal. The shorter peak likely results from copper ions chelated one subunit away while the longer peak must result from coppers two subunits away along the supramolecular polymer.

Note that the ratio of population between the two peaks (0.2) is not what would be expected for an infinite assembly (1) or even for two coiled coils joining together. This implies that there is only a small number of coiled coils assembled into longer oligomers and that most of them exist in structures containing at 2-3 copper ions (Figure 74C). While the DLS measurements on these samples showed large assemblies, DLS is biased to the observation of large particles and cannot be considered quantitative for determining the ratio between polymerized and unpolymerized material.[222]

Increasing the length of the side chain by two methylene units (peptide **32**) results in a increase in the most probable distance between copper ions of 0.4 nm. There is also a small shift of 0.2 nm in the longer-distance peak. Interestingly, the longer side chain, which is expected to be more flexible, gives rise to a narrower distance distribution implying it has *less* conformational freedom than the shorter side chain. It is not clear why this might be the case, but we hypothesize that the longer side chain may be able to more easily find an energy minimum due to its ability to search a larger conformational energy landscape.

Addition of excess Cu^{2+} alters the equilibrium of the Tpy residue coordination states resulting in the disappearance of peaks around 5 nm (Figure 75).

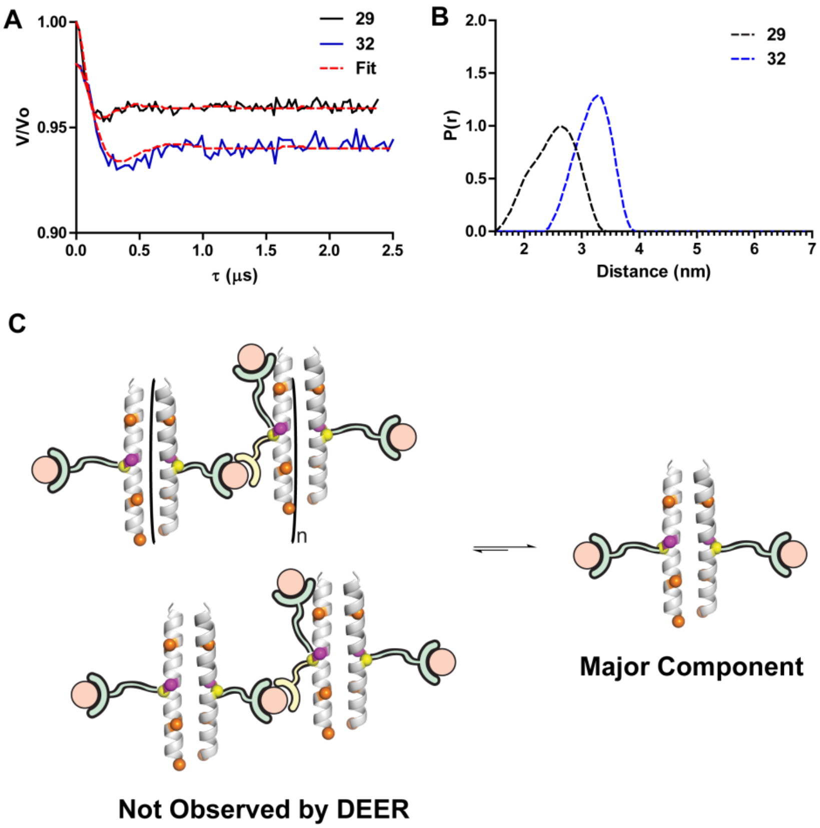


Figure 75. DEER measurements of dimer peptides in the presence of 1.5 equiv Cu^{2+}. A) Time domain spectra and B) distance distribution of peptides **29** and **32**. C) Assembly model based on the DEER data.

This absence implies that large assemblies containing more than two Cu^{2+} ions are not prevalent enough to be observed and that the majority of coiled coils exist independently of one another. Since the only way to form the supramolecular assemblies is through formation of Tpy-Cu^{2+}-Glu linkages, this implies that the population of this linkage has also decreased. This is the strongest evidence that there is a decrease supramolecular polymer formation caused by increasing the Cu^{2+} ion concentration. Even though supramolecular assembly formation is decreased considerably it does not completely disappear as evident from the DLS data (*vide supra*, Section 4.3.2).

4.3.3.1 Altering Coiled-Coil Oligomerization State

Peptides **28** (tetramer) and **30** (trimer) were designed to explore the effect of altering the coiled-coil oligomerization state on the assembly and structure of the materials. The only evidence that these peptides differ in their assembly morphology comes from measuring the distances between Cu^{2+} ions. CW and HYSCORE spectra are consistent with the dimer peptides. DEER spectroscopy of the peptides in the presence of 0.5 equivalents Cu^{2+} yields more complex spectra than the dimers indicating the formation of higher dimensional materials (Figure 76).

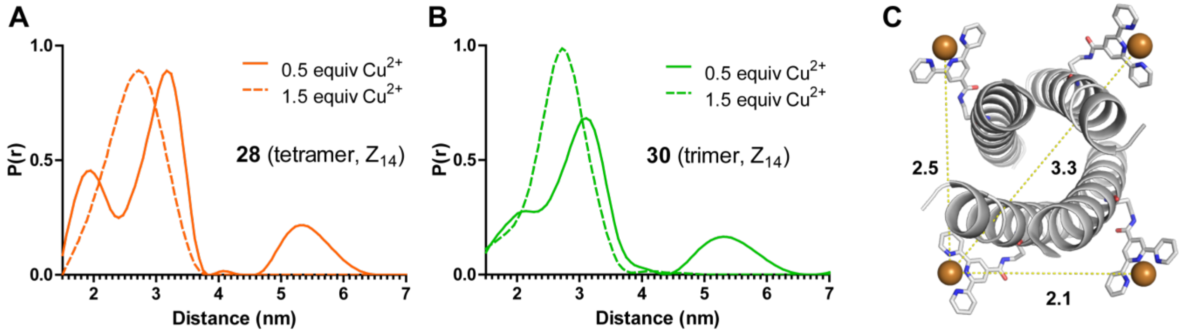


Figure 76. DEER measurements of peptides with higher order oligomerization in the presence of Cu^{2+}. Distance distribution for tetrameric peptide **28** (A, orange) and trimeric peptide **30** (B, green) in the presence of 0.5 (solid) or 1.5 (dashed) equiv Cu^{2+}. C) Cu^{2+}- Cu^{2+} distances (nm) observed in the crystal structure of peptide **37** (PDB ID 5U5B).

Both the tetramer (Figure 76A) and trimer (Figure 76B) distance distributions exhibit a set of three peaks. Two of these overlap and are centered at 5.3 and 3.1 nm. As seen previously with the dimer peptides, the longer peak at 5.3 nm indicates the formation of supramolecular polymers. The shortest distance peaks differ only slightly in most probable distance 1.9 (**28**, tetramer) and 2.0 nm (**30**, trimer). The peaks in the tetramer sequence are much narrower, indicating more ordered assemblies. The double peaks at ~3 and ~2 nm likely arise from the close proximity of the four Cu^{2+} ions in the tetramer; each Cu^{2+} ion can resonate with three other Cu^{2+} ions creating the observed distribution. These distances are very close to what was observed in the crystal structure of peptide **37** (Figure 76C); in both peptides the Tpy **X** residue is located at position 14. For the trimer the origin of both peaks is less apparent, but may be due to an asymmetry between the three Cu^{2+} sites.

Upon addition of excess Cu^{2+} the longer distance vanishes for both peptides (Figure 76B,C) and is replaced by a single broad peak at 2.7. This is to be expected if the majority of Tpy groups are now mono-coordinated allowing for an increased flexibility of the side-chain to create a single distance distribution similar to the model proposed for the dimer system. The absence of the longer peak at ~5.3 nm indicates that supramolecular assembly formation has significantly decreased.

Due to the higher degree of order apparent in its assembly, the tetramer was chosen to examine the role of increasing the side-chain length to compare with the dimer series. Peptide **31** is a variant of peptide **28** containing the longer **Z** side chain in place of the shorter **X** side chain. Upon addition of 0.5 equiv Cu^{2+}, peptide **31** forms supramolecular assemblies as indicated by the peak at 5.1 nm; distances for the other two peaks are at 1.7 and 3.0 nm (Figure 77).

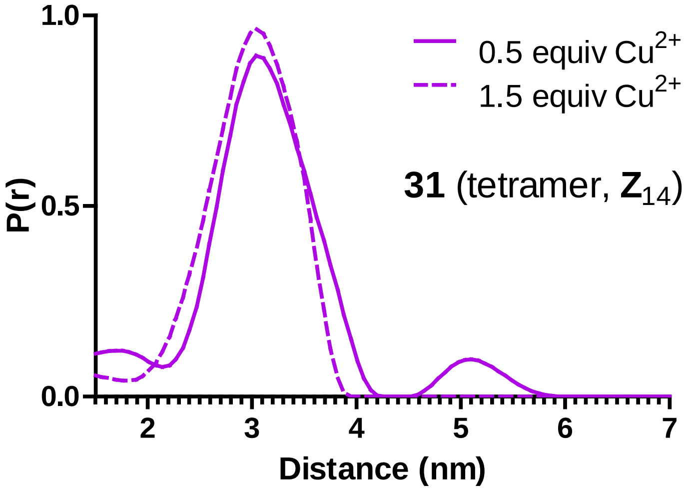


Figure 77. DEER distance distribution for tetrameric peptide **31** (purple) in the presence of 0.5 (solid) or 1.5 (dashed) equiv Cu^{2+}.

Interestingly, all of the major peaks for this peptide are shorter than in peptide **28** even though the side-chain is longer. Although the origin of this phenomenon is not clear it may be that the more flexible side-chain is able to form a more compact structure resulting in the shorter distances. When 1.5 equiv of Cu^{2+} are present the distance distribution converges to a single peak at 3.0 nm, 0.3 nm longer than the shorter **X** side-chain. Since the disappearance of the longer distance occurs for all peptides in the presence of 1.5 equiv Cu^{2+}, this further supports the hypothesis that the majority of coiled coils are in an isolated state under these conditions.

Although the subunits of the assembly generated from peptide **31** with 1.5 equivalents of Cu^{2+} are closer in proximity than those in peptide **28** (5.1 vs. 5.3 nm, respectively), the overall assembly size by DLS is larger for peptide **31** (Figure 78).

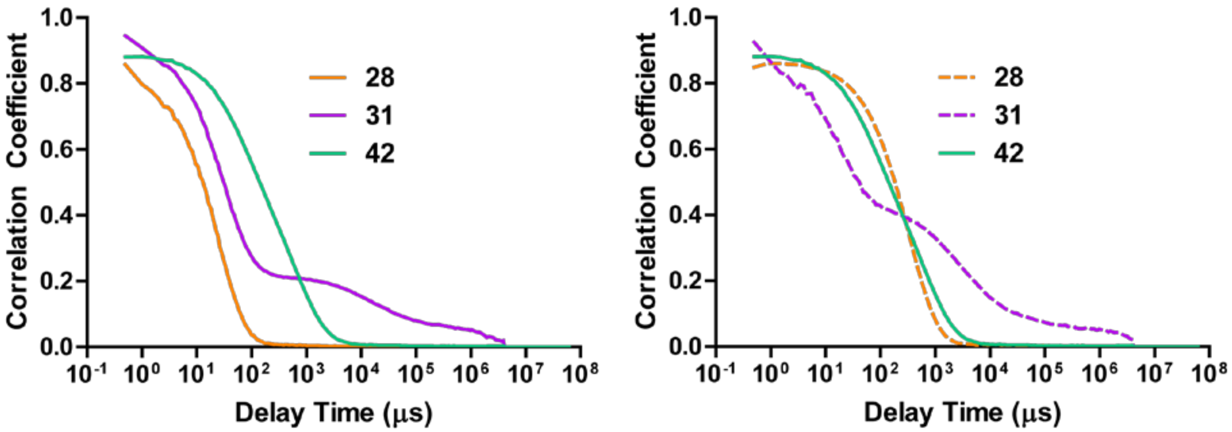


Figure 78. DLS correlation curves of dimeric peptides **28** (orange) and **31** (purple) in the presence of 0.5 (solid, A) or 1.5 equiv (dashed, B) Cu^{2+}. Assembly of Ru^{2+} cross-linked subunit **42** is shown as a reference point for both cases. All measurements were performed with 300 µM (peptides **28** and **31**) or 150 µM (subunit **42**) in aqueous 10 mM HEPES buffer (pH 7) with 0.1 M NaCl.

Consistent with the DEER data, the DLS suggests that peptide **31** generates more heterogeneous assemblies as evident by the tailing of the correlation curve. The larger assembly size of **31** vs. **28** can be explained by the lower helicity of **28** compared to **31** in the presence of 0.5 equivalents Cu^{2+} (Figure 79).

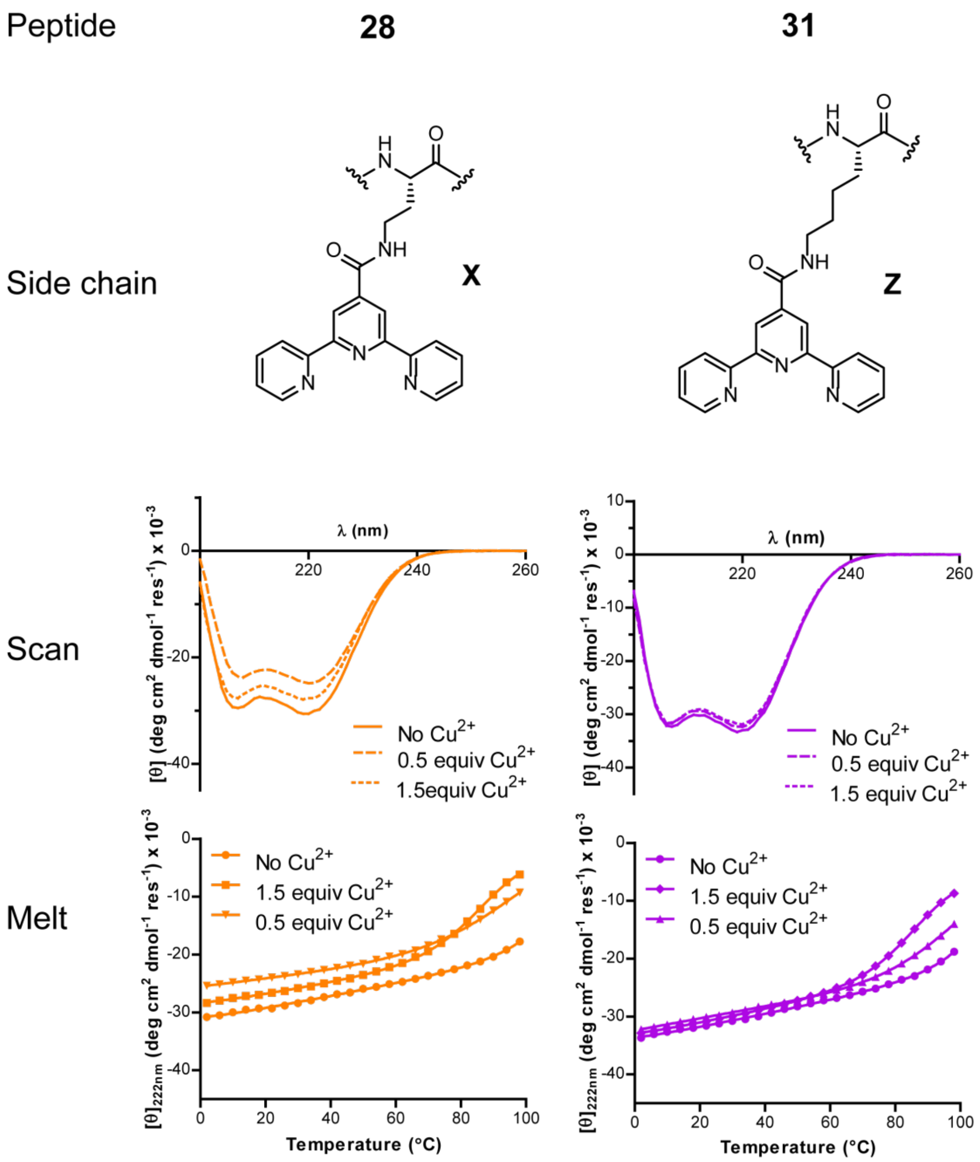


Figure 79. CD scans and melts for tetrameric peptides **28** (orange) and **32** (purple) at 100 µM in aqueous HEPES buffer (10 mM, pH 7) with 0, 0.5 or 1.5 equivalents of $CuCl_2$.

Both uncoordinated peptides have T_m values > 100 °C. Unlike the dimer peptides, both tetramers, while still destabilized by addition of Cu^{2+}, have similar stabilities in the presence of the metal ion. With 0.5 equivalents of Cu^{2+}, they still retain T_m values > 100 °C; unlike with the dimer peptides. However, there is a noticeable decrease in drop peptide **31**'s helicity when coordinating Cu^{2+}; this may explain the smaller assembly sizes observed in the DLS. In the presence of 1.5 equivalents of Cu^{2+} the T_m values are equivalent at 86.1 ± 0.4 °C (**31**) and 86.2 ± 0.4°C (**28**).

Cross-linking peptide **28** with Ru^{2+} to create subunit **42** once again results in larger assemblies compared to the parent peptide. This is accompanied by a decrease in thermal stability (T_m = 86.4 ± 0.5 °C for peptide **42**, Figure 80).

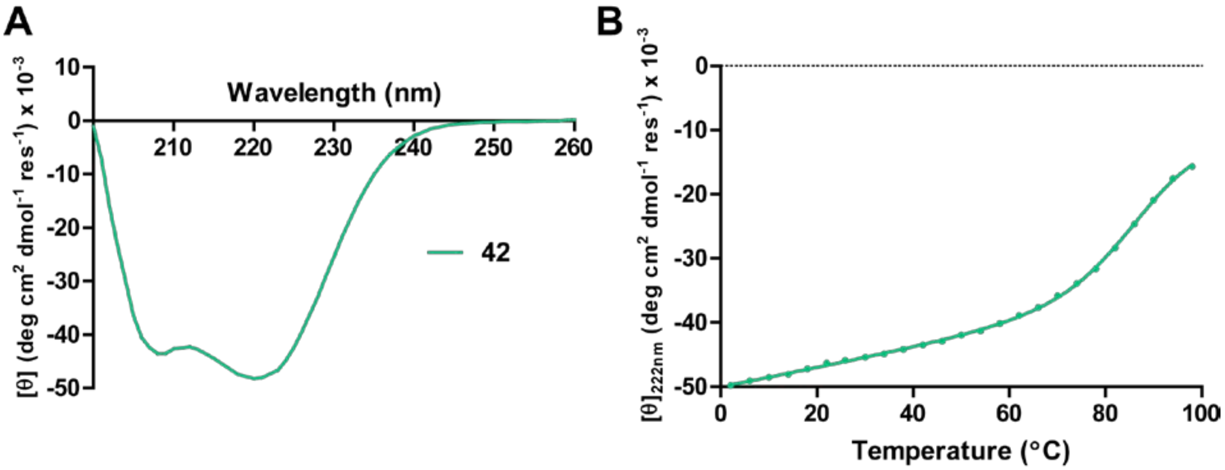


Figure 80. CD scan (A) and melt (B) of Ru^{2+} cross-linked subunit **42** at 50 µM in aqueous HEPES buffer (10 mM pH 7).

This result is consistent with that seen for the dimer peptides, except that no visible aggregation was observed post-melt. It is interesting to note that peptide **28** with 1.5 equivalents of Cu^{2+} and subunit **42** have identical T_m values along with nearly identical DLS correlation curves.

4.3.3.2 DEER Modulation Depth and Background Slope Analysis

Besides the distance distribution, DEER was also used to determine the number of Cu^{2+} ions in the spin system and relative local Cu^{2+} concentration. The decrease in the modulation depth in samples from 0.5 to 1.5 equivalents of Cu^{2+} (Figure 81A) is indicative of fewer Cu^{2+} ions in the spin system.

This data further supports the hypothesis that large assemblies are disfavored when the Cu^{2+} is increased.

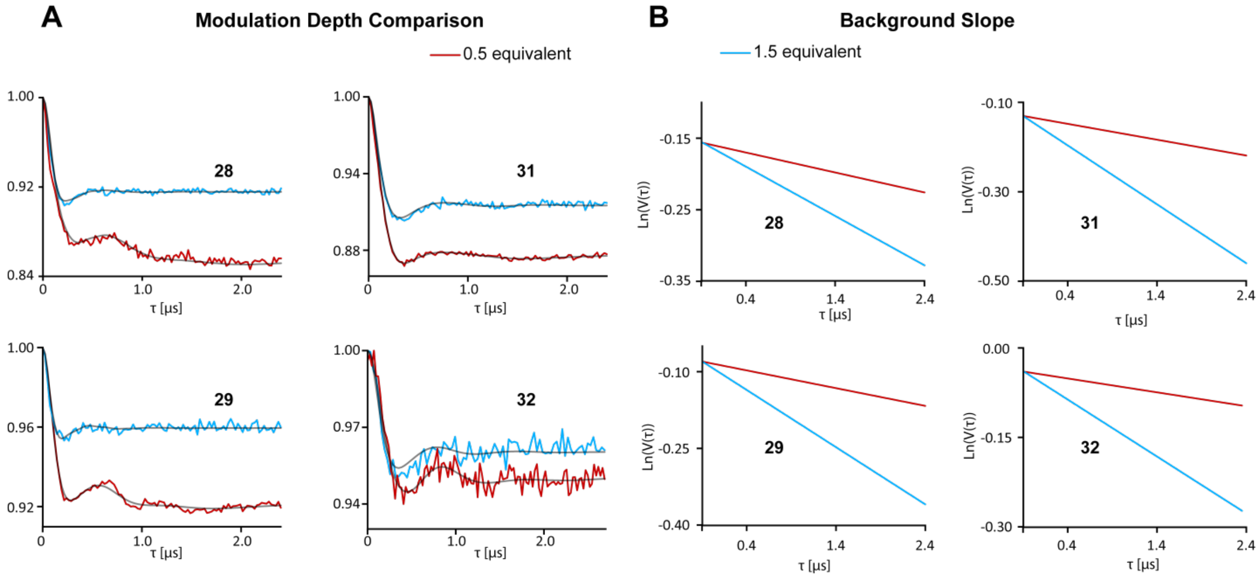


Figure 81. Relative amount of Cu^{2+} ions within the spin system and local concentrations. A) Background subtracted time domain DEER signal showing the decrease in modulation depth upon addition of excess Cu^{2+}. B) Logarithmic scale of the DEER background decay signal. The slope is proportional to the local concentration of bound Cu^{2+}.

This decrease in modulation depth from 0.5 equiv to 1.5 equiv was consistent across all peptides. Furthermore, the raw DEER time domain signal can be analyzed to extract the local concentration of bound Cu^{2+}. Addition of excess Cu^{2+} results in a larger local concentration (Figure 81B) indicative of more bound Cu^{2+} per peptide. The increase in local concentration, but decrease in modulation depth strongly supports the hypothesis that the supramolecular structure is disrupted upon addition of excess Cu^{2+}. Coupled with the change in the HYSCORE spectrum with the loss of the nitrogen signal, it is likely that most of the peptides exist in a disassembled state. However, this disassembly is incomplete as there is also evidence for some degree of larger assemblies from

DLS measurements (*vide supra*, Figure 71B). Another possibility is that the local concentration increases due to binding of the excess Cu^{2+} by glutamates.

Combining all of the available data (CD, EPR and DLS) allows a picture of the supramolecular assembly to be established (Figure 82); although only the dimer peptide is depicted for ease, this model can be extrapolated to the other oligomerization states as well.

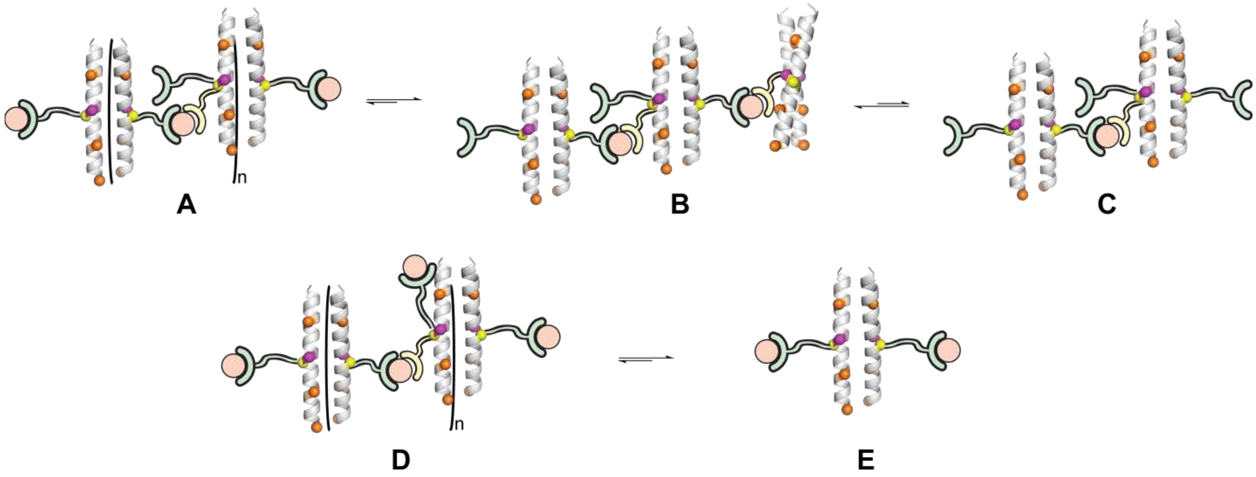


Figure 82. Supramolecular polymer assembly model of dimer peptides.

In the presence of 0.5 equiv of Cu^{2+}, the predominant assembly mode is likely state B with a significant presence of state A and an unknown amount of state C. State A is likely lower in population since the longer distance peak observed in the DEER has a much smaller area than the shorter distance. DLS data imply that some some large assemblies do exist though. States B and C are more likely than state E due to the presence of the nitrogen signal observed in the HYSCORE data. In the presence of 1.5 equiv of Cu^{2+}, the predominant assembly state is state E with a small amount of state D. Evidence for this is most readily seen from the DEER data by the absence of any longer distances. De-polymerization likely results from charge-charge repulsion between the Tpy-Cu^{2+} moieties, so that water coordination becomes favored over glutamate coordination. The destabilization of the peptide fold observed by CD melts with addition of excess Cu^{2+} may be evidence of this hypothesis. However, larger assemblies must still exist as evidenced by the particle sizes seen by DLS. It is likely though that these assemblies are formed through non-specific aggregation since EPR samples which were not immediately flash-frozen showed signs of aggregation.

4.4 CONCLUSIONS AND FUTURE DIRECTIONS

The results presented in this chapter represent the first steps toward a new class of protein-based materials that can be generated by combining the metal chelating power of terpyridine with the self-assembly of coiled-coil peptides. Control over the coordinating residues should be possible by matching the number of glutamate and terpyridine groups in each sequence. Assemblies formed from two different sequences should also be possible if one contains glutamate and no terpyridine residues while the other contains terpyridine and no glutamate residues. Derivatization of the terpyridine moiety should enable the exploration of an array of electronic and catalytic functionality,[226-229] enabling the development of new nanomaterials. Future designs of crystalline assemblies will take these ideas into consideration.

Although a tentative model has for solution-phase supramolecular assembly has been presented here, there are still some unanswered questions concerning these systems that need to be addressed. Determination of the identity of Glu coordination may be determined through synthesizing a series of peptides identical those described here each with a different Glu isotopically labeled with ^{18}O then looking for the presence of Cu^{2+}-^{18}O coupling with HYSCORE. EPR measurements on these samples should disambiguate both the coordination mode and identity of Glu binding. Likewise, peptides with only a single Glu residue could be designed to improve the order of the assemblies by only providing a single coordination site.

In an effort to study the originally proposed bis-Tpy-M^{2+} complex by EPR, Co^{2+} could be used in place of Cu^{2+} as the spectroscopic probe. Co^{2+} also has a higher affinity for bis-Tpy complexes[230] than Cu^{2+} and should therefore promote the formation of larger, more specific assemblies. The complication of using Co^{2+} is its coordination is less studied than that of Cu^{2+} and contains two possible spin states (Cu^{2+} has only one) , which can even vary for the same ligand system.[231] DEER measurements on Co^{2+}-Co^{2+} systems are also unprecedented in the literature although there is an example of a DEER measurement between a Co^{2+} and nitroxide center.[232] Solubility may also be a concern using cobalt since all crystallization screens involving Co^{2+} precipitated into large aggregates. Pioneering the use of Co^{2+}-Co^{2+} DEER measurements could also be useful in other contexts like the structure of zinc fingers where Co^{2+} long been used in place of Zn^{2+} to monitor binding by UV-Vis spectroscopy.[233]

4.5 EXPERIMENTAL

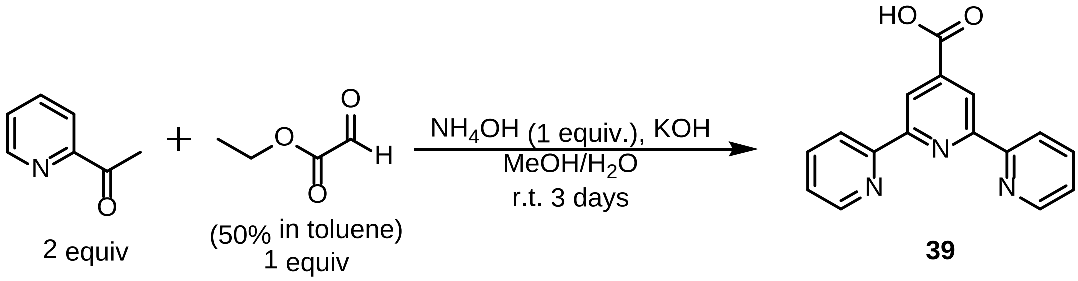
4.5.1 General Information

HATU, NovaPEG Rink Amide Resin, and Fmoc-protected α-amino acids were purchased from Novabiochem. HCTU was purchased from Aapptec. Fmoc-L-Dab(Alloc)-OH was purchased from ChemImpex. Solvents and all other reagents were purchased from AcrosOrganics, Aldrich, AlfaAesar, Fisher, or Hampton Research.

4.5.2 Terpyridine Side-chain Analogues

4.5.2.1 Synthesis

Synthesis of Compound 39.

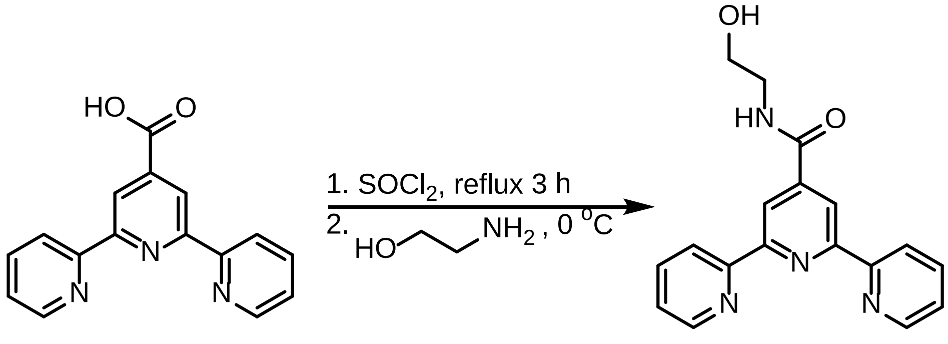


39

Scheme 6. Synthesis of 4'-carboxy-2,2':6',2"-terpyridine (compound **39**).

4'-Carboxy-2,2':6',2"-terpyridine (**39**) was synthesized as described previously (Scheme 6).[208] **^{1}HNMR** and high-resolution **MS** data matched published values.

Synthesis of Ligand 40.



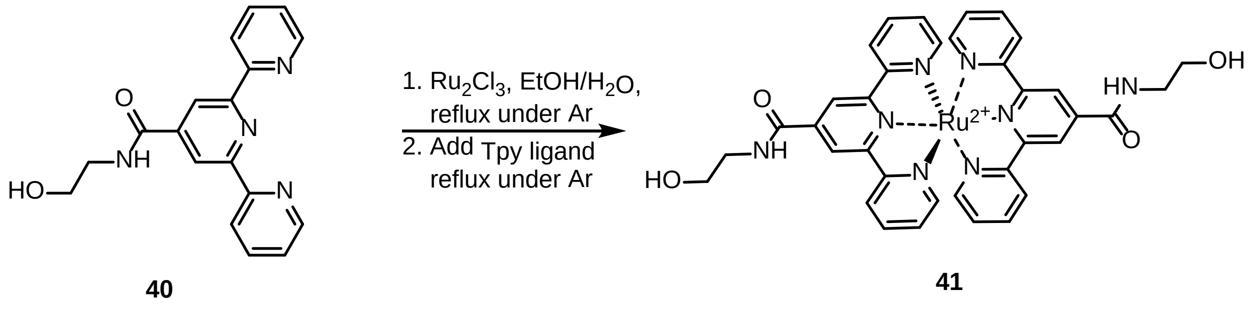
39

40

Scheme 7. Synthesis of ligand **40** from compound **39**.

Synthesis of **40** (Scheme 7**)** was adapted from the literature.[234] A solution of **39** (279.7 mg, 1.009 mmol) in excess thionyl chloride (18 mL) was refluxed for 3 hours. Solvent was removed to dryness by vacuum distillation to afford the acid chloride intermediate, which was used without further purification. To the acid chloride, a solution of excess ethanolamine (7 mL) in dry dichloromethane (7 mL) was added dropwise at 0 °C under nitrogen resulting in a translucent red solution. The reaction was stirred overnight under nitrogen at room temperature resulting in a cloudy purple suspension. The reaction was precipitated into water (50 mL) and extracted three times with dichloromethane (50 mL each). Solvent was removed via vacuum distillation. The resulting solid was re-suspended in minimal ethanol and precipitated with water (100 mL). The precipitate was filtered and dried under high vacuum to obtain the product as a white solid (219.6 mg, 68% yield). **^{1}HNMR** (400 MHz, DMSO-d_6) δ 9.06 (t, *J* = 5.6 Hz, 1H), 8.84 (s, 2H), 8.77 (dq, *J* = 4.8, 0.8 Hz, 2H), 8.66 (d, *J* =8 Hz, 2H), 8.05 (td, *J* = 7.6, 1.6 Hz, 2H), 7.54 (ddd, *J* =7.6, 4.8, 1.2 Hz, 2H), 4.80 (t, *J* = 5.6 Hz, 1H), 3.58 (q, *J* = 6 Hz, 2H) , 3.41 (q, *J* = 6 Hz, 2H) ppm; **^{13}CNMR** (100 MHz, DMSO-d_6) δ 164.8, 155.7, 154.6, 149.4, 144.2, 137.6, 124.7, 121.0, 118.4, 59.5, 42.4 ppm; **HRMS** (ESI) *m/z* calculated for $C_{18}H_{17}N_4O_2$ $[M+H]^+$: 321.1352; found 321.1351.

Synthesis of Complex 41.



Scheme 8. Synthesis of complex **41** from ligand **40**.

Synthesis of complex **41** (Scheme 8) was adapted from the literature.[235] A solution of $RuCl_3$•H_2O (31 mg, 0.15 mmol, 1 equiv) in EtOH/H_2O (1:1 v/v, 12 mL) was bubbled with Ar for 20 min. The de-oxygenated solution was heated and stirred under Ar at ~90 °C for 4 hours. A solution of ligand **40** (96 mg, 0.30 mmol, 2 equiv) in EtOH/H_2O (1:1 v/v, 20 mL) was bubbled with Ar for 30 min. then added to the ruthenium solution. The solution was left to stir at ~90 °C overnight. The product was purified by reverse phase HPLC using a gradient between water and acetonitrile with 0.1% TFA. Pure fractions were combined and lyophilized yielding the red TFA salt. **^{1}HNMR** (400 MHz, MeCN-d_3) δ 9.34 (d, *J* = 2.4 Hz, 4H); 8.91 (s, broad, NH); 8.71 (d, *J* = 8 Hz, 4 H), 7.92 (td, *J*= 8, 1.2 Hz, 4 H), 7.37 (d, *J* = 5.2 Hz, 4H); 7.15 (td, *J* = 6.6, 1.2 Hz, 4H); 3.85 (t, *J* = 5.2 Hz, 4H); 3.71 (q, *J* = 5.2 Hz, 4H); 3.58 (s, broad, OH) ppm; **^{13}CNMR** (100 MHz, 0.1M NaPhos pH 7 in 10% D_2O/90% H_2O) δ 167.4, 157.2, 155.6, 152.0, 140.5, 138.2, 127.5, 125.6, 121.1, 60.1, 42.7; **^{13}CNMR** (100 MHz, MeCN-d_3) δ 166.2, 164.8, 158.7, 156.4, 153.5, 139.2, 128.6, 125.8, 122.7, 118.3, 44.5 ppm;* **HRMS** (ESI) *m/z* calculated for $C_{36}H_{32}N_8O_4Ru$ $[M]^{+2}$: 371.0790; found 371.07640 (*z* = 2). *(second aliphatic peak not visible)

4.5.2.2 Molar Absorptivity Determination

A stock solution of ligand **40** (2.9 mg in 10 mL, 905 µM) was prepared in H_2O/DMSO (80%/20%, v/v). Samples were prepared from the stock solution by dilution with H_2O at concentrations of 7.5, 15, 30, 60, and 90 µM. UV-Vis spectra of each sample was measured from 200-600nm and referenced to H_2O. The molar absorptivity at 310 nm was determined using Beer's law to be 8400±200 M^{-1} cm^{-1} (Figure 83).

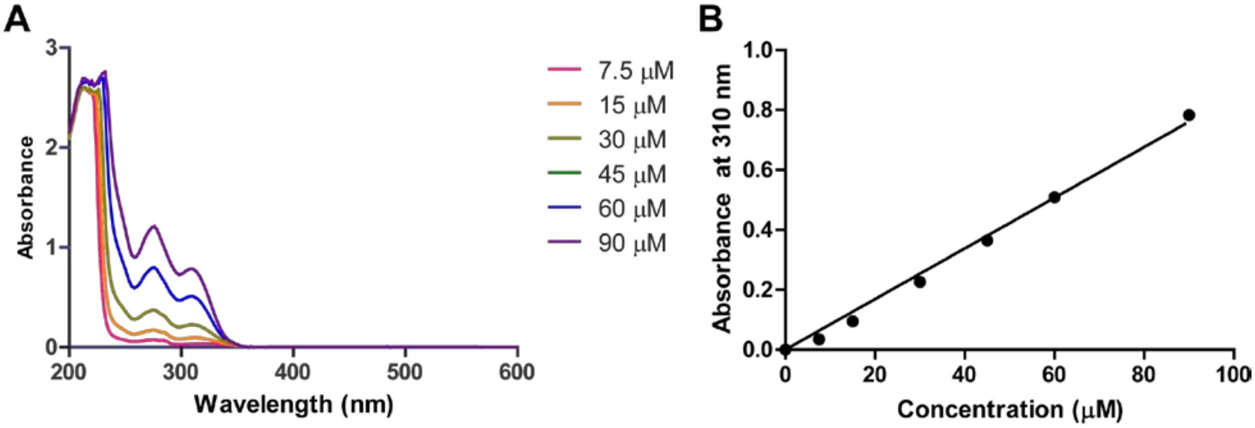


Figure 83. Molar absorptivity measurement of ligand **40**.

A) Absorbance spectra of ligand **40** from 7.5-90 μM and B) determination of its molar absorptivity at 310 nm.

Stock solutions of complex **41** were prepared in triplicate at a concentration of 10.0 mM in 0.2 M sodium phosphate pH 7. Samples were prepared in triplicate from the three stock solutions by dilution with 0.2 M sodium phosphate pH 7 to 2.5, 5, 10, 20, 40, and 80 μM and measured from 200-800nm referenced to a solution of 0.2 M sodium phosphate pH 7. A molar absorptivity of 21,900 ± 100 M^{-1} cm^{-1} at 486 nm was found using Beer's law (Figure 84).

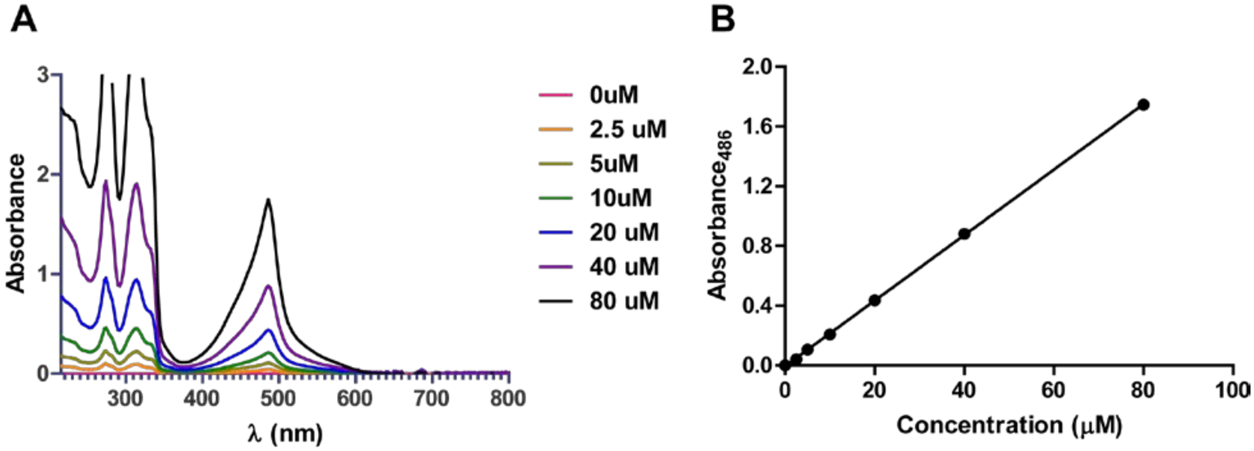


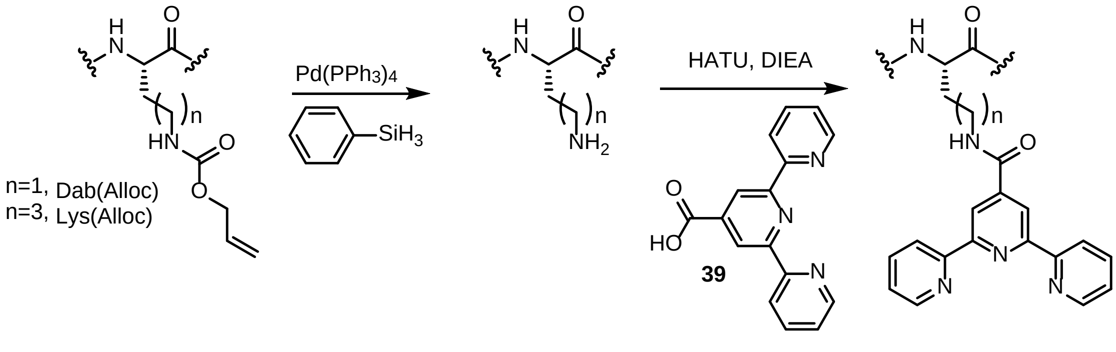
Figure 84. Molar absorptivity measurement of complex **41.**

A) Absorbance spectra of complex **41** from 2.5-80 μM and B) determination of its molar absorptivity at 486 nm.

4.5.3 Peptide Synthesis

Peptides **24-38** were synthesized by manual microwave-assisted Fmoc solid phase methods using a CEM MARS microwave and NovaPEG Rink Amide resin. Typical coupling reactions were performed with a 1.5 min ramp to 90 °C followed by a 2 min hold at that temperature, and Fmoc deprotections performed with a 1.5 min ramp to 90 °C followed by a 1 min hold at that temperature.[200] Coupling solutions were composed of Fmoc-protected amino acid (5 equiv relative to resin), HCTU (4.9 equiv), and DIEA (7.5 equiv) in NMP. Fmoc deprotections used 20% v/v 4-methylpiperidine in DMF. The resin was washed three times with DMF between each reaction. Full-length peptides were acetyl capped by treatment with DMF/DIEA/Ac_2O (8/2/1, v/v/v) at room temperature for 20 mins.

Terpyridine-modified side chains were introduced by coupling Fmoc-Lys(Alloc)-OH (to produce residue **Z**) or Fmoc-Dab(Alloc)-OH (to produce residue **X**) at the desired sites during synthesis of the full-length chain and subsequent on-resin conversion as detailed below (Scheme 9).



Scheme 9. Alloc deprotection and Tpy attachment to the produce side-chain **Z** (n=3) and **X** (n=1).

Alloc was removed by treatment of resin with $Pd(PPh_3)_4$ (0.25 equiv relative to Dab) and phenylsilane (24 equiv) in anhydrous CH_2Cl_2 under Ar for 0.5 h.[236] Resin was washed three times each with DCM, THF, DMF, 0.02M sodium diethyldithiocarbamate in DMF, and DMF.[237] Carboxylic acid **39**[208] (5 equiv relative to Dab) was then coupled to the resulting amines with HATU (4.9 equiv) and DIEA (7.5 equiv) in NMP according to the microwave heating protocol detailed above. Resin was washed three times each with DMF, CH_2Cl_2, and MeOH and dried in a vacuum desiccator for 20 min prior to cleavage.

Peptides **24-36** were cleaved from the resin by treatment with TFA/H_2O/EDT/TIS (92.5/3/3/1.5 by volume) for ~4 h. Peptide **37** and **38** were cleaved from the resin TFA/H_2O/EDT/Anisole/TIS (90/3/3/3/1 by volume) for ~3.5 h. After filtration, crude peptides were precipitated from the TFA cleavage solution with cold Et_2O, and the suspension was centrifuged and decanted. Pellets were re-dissolved in 6 M guanidinium, 25 mM phosphate pH 7 and stored in the freezer until purification.

Peptides were purified via preparative reverse phase HPLC on C18 (150 Å pore size, 10 µm particle size) columns using gradients between 0.1% TFA in water and 0.1% TFA in acetonitrile. The purity and identity of each peptide was confirmed by analytical HPLC (Figure 85) and MALDI-TOF mass spectrometry (Table 9), respectively.

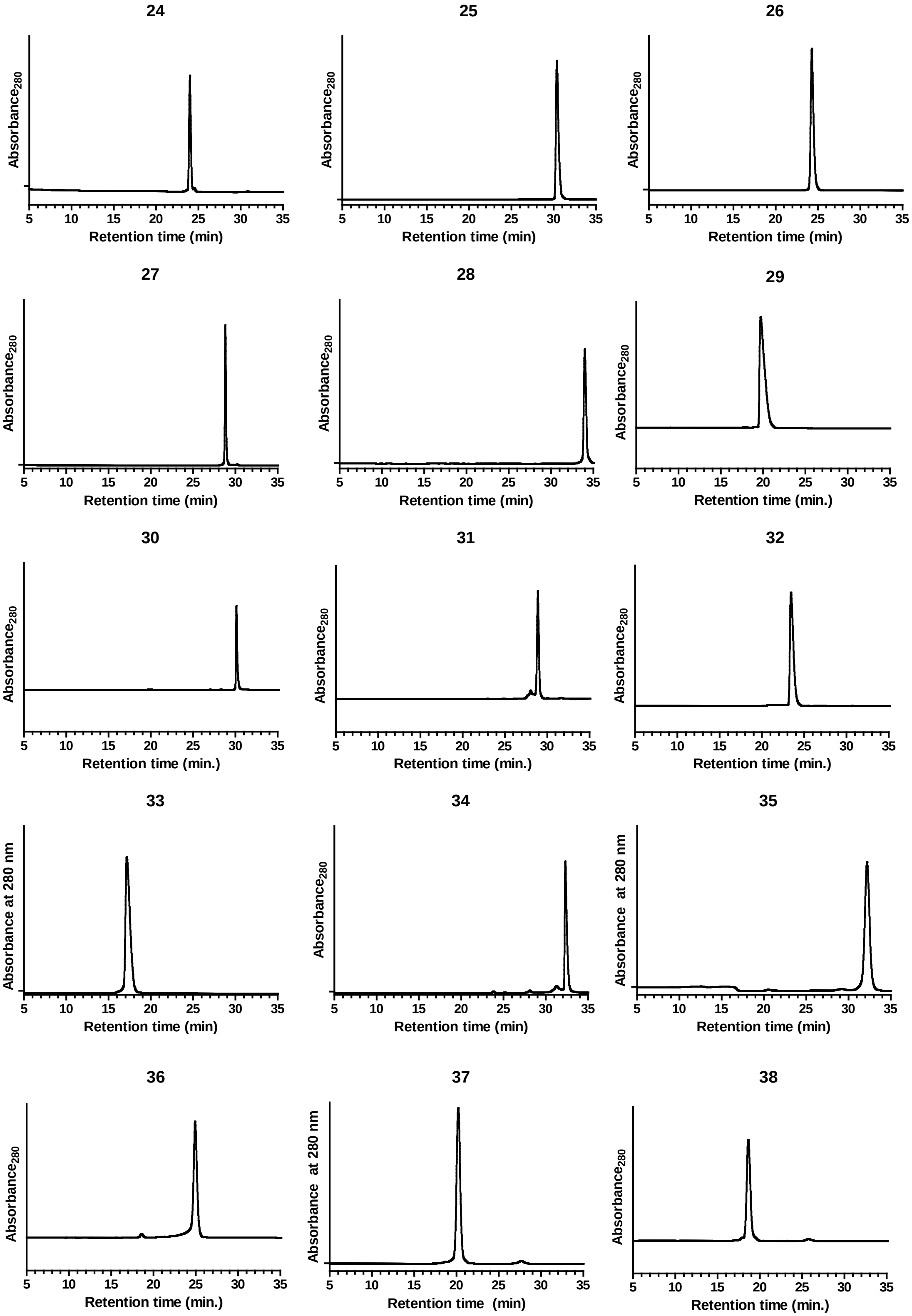


Figure 85. Analytical HPLC chromatograms of purified peptides **24-38**.

Table 9. Calculated and observed average masses of proteins **24-38**.

#	$[M+H]^+$ *m/z* (average)	
	Calculated	**Observed**
24	2882.4	2881.7
25	2858.4	2858.5
26	2801.3	2799.8
27	2773.3	2771.8
28	3306.8	3306.0
29	3365.0	3364.5
30	3363.8	3363.6
31	3334.8	3333.1
32	3393.1	3391.0
33	3595.8	3595.5
34	3595.8	3595.0
35	3768.5	3768.7
36	3540.1	3539.2
37	3307.8	3307.9
38	3307.8	3306.7

4.5.3.1 Peptide Ru^{2+} Cross-Linking

Ruthenium(II) cross-linked subunits **42** and **43** were synthesized and purified from peptides **28** and **29**, respectively, following a similar procedure as that of complex **41** with the peptides substituted for ligand **40**. Identity of the products was confirmed by MALDI-TOF spectrometry (Table 10) and purity by analytical HPLC (Figure 86).

Table 10. Calculated and observed average masses of subunits **42** and **43**.

#	$[M-H_2O]^+$ *m/z* (average)		$[M]^{2+}$ *m/z* (average)	
	Calculated	**Observed**	**Calculated**	**Observed**
42	6694.6	6694.8	3356.3	3355.6
43	6811.0	6809.9	3414.5	3416.6

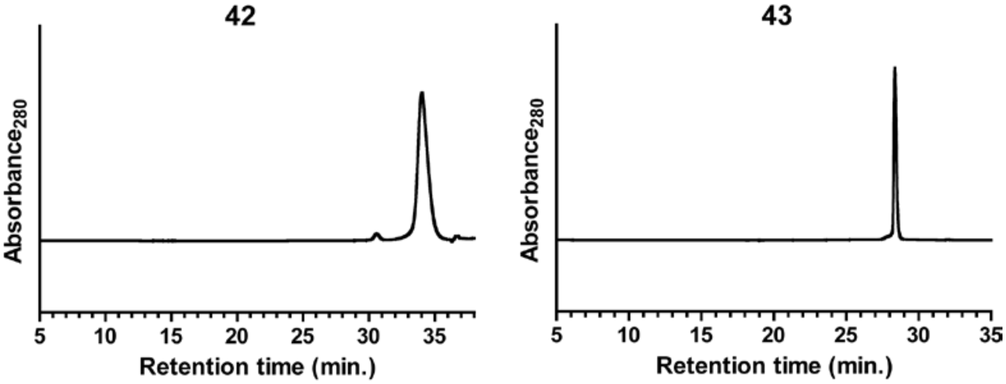


Figure 86. Analytical HPLC chromatograms of purified subunits **42** and **43**.

4.5.4 Circular Dichroism

CD measurements were performed in 1 mm quartz cuvettes on an Olis DSM 17 CD spectrometer. All buffer components were syringed filtered through a 0.22 µm pore filter. Scans were measured from 200-260 nm at 1 nm increments with a 2 nm bandwidth and 5 sec integration time at each wavelength. Melts were monitored at 222 nm from 2-98 °C with 4 °C temperature steps and 2 min equilibration at each temperature. All scans and melts were baseline corrected against buffer blanks measured in the same cuvette. Raw scan data were smoothed and melts fit to a two-state unfolding model[169] using GraphPad Prism.

4.5.5 Crystallography

Crystals of peptides **33, 35, 37,** and **38** were grown by hanging drop vapor diffusion. Stock solutions (15 mg/mL for **33** and **38**, and 7.5mg/mL for **35** and **37**) in water were mixed (0.7 µL + 0.7 µL) with the crystallization buffer indicated in Table 11 and equilibrated at room temperature over a well of the same buffer (700 µL). Two different crystal forms of peptide **33**, denoted *a* and *b*, were obtained under different growth conditions. It is worth noting that all peptides gave rise to crystals in the presence of only BIS-TRIS buffer and Cu^{2+} salt, but these were usually not single (with the exception of peptide **35**).

Table 11. Crystallization buffer conditions foe peptides **33** (forms *a* and *b*), **35**, **37**, and **38**.

Peptide	**Buffer**
33 (crystal form *a*)	0.1 M sodium citrate pH 6.0, 3 mM $CuCl_2$, 15% w/v PEG 4000
33 (crystal form *b*)	0.1 M BIS-TRIS with 2 mM terephthalic acid pH 6.5, 10 mM $CuCl_2$, 10% w/v PEG 4000
35	0.1 M BIS-TRIS pH 6.5, 10 mM $CuCl_2$
37	0.095 M BIS-TRIS pH 6.5, 9.5 mM $CuCl_2$, 5 mM sodium citrate pH 5.6, 0.125 M 1,6-hexanediol
38	0.1 M BIS-TRIS with 1 mM terephthalic acid pH 6.5, 10 mM $CuCl_2$, 10% w/v PEG 4000

Crystals were harvested, cryoprotected with solutions of 30% v/v glycerol in mother liquor, and then flash frozen in liquid nitrogen. X-ray diffraction data for peptide **33**, crystal form *a*, were collected using Cu/K_α radiation on a Rigaku diffractometer (FR-E generator, VariMax optics) with a Saturn 944 CCD detector. Data for peptide **33**, crystal form *b*, along with peptides **35**, **37**, and **38** were collected at the Advanced Photon Source at Argonne National Laboratory on beamline 22-ID operated at a wavelength of 1.0 Å.

Raw diffraction data were processed using d*TREK. Structures were solved by molecular replacement using Phaser.[170] Truncated versions of the following published structures were used as search models: peptide **33** (PDB 4DZM, residues 2-30), peptide **35** (PDB 4DZL, residues 2-29), peptides **37** and **38** (PDB 3R4A, residues 2-29). Refinement was performed using Phenix[171] in tandem with Coot[172] for manual real-space model building. TLS parameters were used in the refinement of crystal form *b* of peptide **33**. Geometric restraints for the terpyridine-modified Dab residue were generated using the REFMAC[238] monomer library sketcher in the CCP4 software suite.[239] Geometric restraints for the coordination of Cu(II) by the terpyridine residue was based on a crystal structure of bis(2,2':6',2"-Terpyridyl)-copper(II) hexafluorophosphate (CSD entry BEJPUB),[240] and restraints for the binding between carboxylates and terpyridine-Cu(II) residues from the crystal structure of μ_2-succinato-bis[aqua(2,2':6',2"-terpyridine)copper(II)] dinitrate dihydrate (CSD entry PUHWIZ).[241] Final data collection and refinement statistics are listed in Table 12.

Table 12. Data collection and refinement statistics for peptides **33** (forms *a* and *b*)**, 35,** and **37**.

Peptide	**33** (crystal form *a*)	**33** (crystal form *b*)	**35**	**37**
Data Collection				
Unit Cell Dimensions (Å, °)	*a*=36.3, *b*=52.2, *c*=25.6 α=γ=90, β=104.0	*a*=*b*=57.1, *c*=80.2 α=β=90, γ=104.0	*a*=*b*=77.6, *c*=55.1 α=β=γ=90	*a*=114.3, *b*=38.6, *c*=44.2 α=β=γ=90
Space Group	C2	$P6_{1}22$	$P4_{1}2_{1}2$	$P2_{1}2_{1}2$
Mosaicity	1.40	1.05	0.71	2.05
Resolution	29.21-2.20 (2.28-2.20)	31.15-3.23 (3.35-3.23)	38.81-2.40 (2.49-2.40)	29.09-2.10 (2.18-2.10)
Total Reflections	20805	10970	48001	40191
Unique Reflections	2391	1430	6926	11924
Redundancy	8.70 (4.42)	7.67 (8.15)	6.93 (7.19)	3.37 (3.42)
% Completeness	99.6 (97.6)	99.6 (100.0)	99.3 (100.0)	99.3 (100.0)
Rmerge %	5.3 (21.2)	11.7 (29.8)	9.6 (39.7)	11.0 (38.8)
I/σ (averaged)	29.0 (3.8)	11.0 (3.7)	10.8 (3.8)	6.0 (2.3)
I/σ (unaveraged)	9.6 (1.9)	4.3 (1.5)	4.5 (1.8)	3.6 (1.6)
Refinement				
Resolution (Å)	26.09-2.20	31.15-3.23	38.88-2.40	29.09-2.10
R (%)	22.9	30.0	25.6	26.5
R_{free} (%)	25.7	31.3	28.2	27.6
Avg. B factor (Å^2)	35	89	55	35
RMSD				
Bonds (Å)	0.003	0.005	0.003	0.006
Angles (°)	1.068	1.288	1.040	0.881

Although full structural refinement was not possible for peptide **38**, a molecular replacement solution was found in space group $P4_{2}22$ with two chains of the tetramer from PDB ID 3R4A as a search model; the asymmetric unit also consists of only two chains (half the tetramer). Data collection statistics are shown below in Table 13.

Table 13. Data collection statistics for peptide **38**.

Data Collection	
Unit Cell Dimensions (Å, °)	*a*=*b*=37.3, *c*=94.33 α= β=γ=90
Space Group	$P4_222$
Mosaicity	1.21
Resolution	37.33-3.65 (3.78-3.65)
Total Reflections	5319
Unique Reflections	885
Redundancy	6.01 (6.14)
% Completeness	97.9 (100.0)
Rmerge %	0.070 (0.322)
I/σ (averaged)	10.7 (3.2)
I/σ (unaveraged)	4.6 (1.6)

4.5.6 Dynamic Light Scattering

All DLS measurements were made in a 1 cm quartz cuvette on a Malvern Zetasizer Nano ZS90 with a 632.8 nm laser and an angle of 90° at a temperature of 20°C. Peptide assembly samples were prepared containing 600 µM peptide with 300 µM (0.5 equiv) $CuCl_2$ in 10 mM HEPES pH 7 (**28-29** and **31-32**) or 150 µM peptide in 0.1 M NaCl, 10 mM HEPES pH 7 (subunits **33-34**) then equilibrated overnight at room temperature. Each sample was measured 5-9 times and the averaged correlation curve reported.

4.5.7 EPR Measurements

All EPR measurements were performed and analyzed by Matthew J. Lawless (Saxena Lab, University of Pittsburgh).

4.5.7.1 Sample Preparation

Samples used in EPR experiments were 150 µL containing 20% v/v glycerol as a cryoprotectant. Concentration of ligand **40** and peptides **28-32** was kept at 300 µM while $^{63}Cu^{2+}$ concentration was either 150 µM or 450 µM for the 0.5 equivalent and 1.5 equivalent samples respectively. Solvent pH was controlled with 40 mM NEM at pH 7.4. Upon addition Cu^{2+} addition samples were vortexed to mix. Under one minute after mixing samples were placed in a quartz tube (3 mm inner diameter by 4 mm outer diameter), flash frozen in liquefied MAPP gas, and inserted into a sample cavity pre-cooled to 80 K. Experiments were performed at either 80K or 20K using an Oxford ITC503 temperature controller and an Oxford CF935 dynamic continuous flow cryostat connected to an Oxford LLT 650 low-loss transfer tube.

4.5.7.2 Continuous Wave Measurements

CW experiments were performed on a Bruker ElexSys E580 CW/FT X-band spectrometer with a Bruker ER4118X-MD5 resonator. The CW spectra were collected consisting of 1024 data points using a center field of 3100 with a 2000 G sweep width. A modulation depth of 4 G, a modulation frequency of 100 kHz, a conversion time of 20.48 ms and a time constant of 10.24 ms were used. 25 scans were collected for all spectra. Spectral simulations were performed with the Bruker Simfonia software.

4.5.7.3 HYSCORE Measurements

The two dimensional, four-pulse HYSCORE experiment was performed at X-band frequencies on either a Bruker ElexSys E580 CW/FT X-band spectrometer with a Bruker ER4118X-MD5 resonator or a Bruker ElexSys E680 CW/FT X-band spectrometer equipped with a Bruker EN4118X-MD4 resonator. The pulse sequence used was as follows: $\pi/2 - \tau - \pi/2 - \tau 1 - \pi - \tau 2 - \pi/2$ – echo, where the $\pi/2$ and π pulse lengths were 16 ns and 32 ns respectively. $\tau 1$ and $\tau 2$ were both stepped out by 16 ns. Spectra were all collected at the field corresponding with maximum echo intensity detected by the echo-detected field swept spectrum. Data acquisition was 16 hours. Raw data was baseline corrected, zero filled and fast Fourier transformed. The fast Fourier transformation was reported as a contour plot.

4.5.7.4 DEER Measurements

The dead time free four pulse DEER experiment was performed at X-band frequencies on a Bruker ElexSys E680 CW/FT X-band spectrometer equipped with a Bruker EN4118X-MD4 resonator. All DEER experiments were performed at 20 K. The pulse sequence used was as follows: $(\pi/2)\nu 1 - \tau 1 - (\pi)\nu 1 - T - (\pi)\nu 2 - \tau 2 - (\pi)\nu 1 - \tau 2$ – echo.[242] The observer pulse lengths, $(\pi/2)\nu 1$ and $(\pi)\nu 1$, were 16 ns and 32 ns respectively. The pump pulse length, $(\pi/2)\nu 2$, was 16 ns. The delay, T was incremented by a step size of 20 ns for 128 points. The pump frequency, $\nu 2$, was positioned at the maximum of the echo detected field swept Cu^{2+} spectrum. The observer frequency, $\nu 1$, was offset 150 MHz downfield from the pump pulse. The raw time domain DEER data were analyzed via DeerAnalysis2013[243] using Tikhonov regularization. Data acquisition ranged from 6 to 12 hours.

5.0 CONCLUSION AND FUTURE PERSPECTIVES

Proteins are excellent scaffolds for the design of novel materials and structures. Structural control at the sequence level allows for a great degree of tunability and modularity in their properties. By incorporating synthetic cross-links and unnatural building blocks we can expand this structural and functional diversity even further.

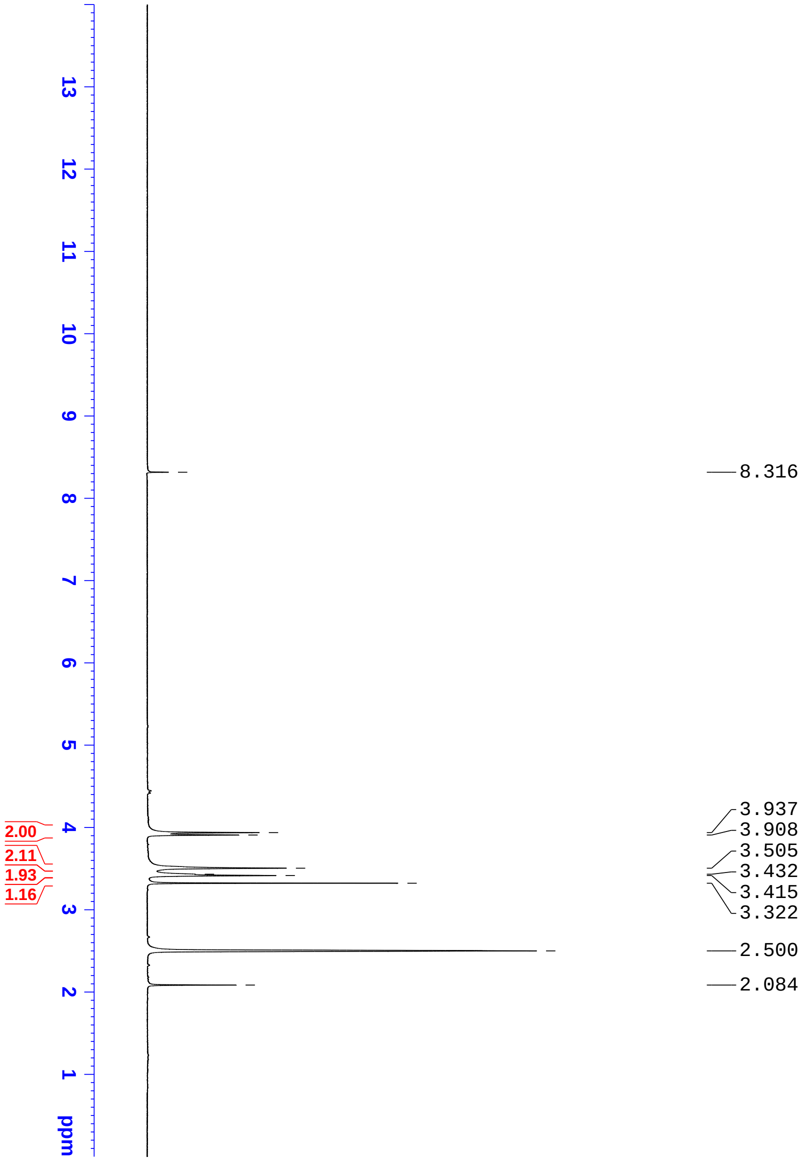
The work described in this thesis highlights three projects that have increase our understanding of helical design strategies for the creation of supramolecular polymers (Chapter 2), crystalline metal-coordinated architectures (Chapter 3), and stabilization of tertiary fold mimics (Chapter 4). Macroscopic differences between similar linkers were shown to arise from the presence of two sub-populations within the conformations of the more rigid cross-linked subunit. Through replacement of the covalent linker with metal-coordination in tandem with rationally designed oligomeric helical bundles, a new class of metallopeptides which give rise to crystalline supramolecular architectures was developed. Cα-methylation of amino acids incorporated into a helix were found to impart a greater thermodynamic stabilization to mixed-backbone tertiary fold mimics than β-residues.

The knowledge of protein design described within this thesis will assist in the future design of novel protein mimics and protein-based materials. By incorporating Cα-methylated residues into helices within other tertiary fold contexts, it can be determined if the observation of the GB1 system was general or an anomaly. The Tpy-Cu^{2+}-Glu motif described in Chapter 4 may prove useful in designing assemblies comprised of multiple oligomerization states. The Tpy-modified residue may also provide a scaffold to develop Co^{2+}-Co^{2+} DEER spectroscopy enabling structural analysis of materials formed through bis-terpyridyl coordination. It will be exciting to see what novel materials and applications may arise based on the design motifs described herein.

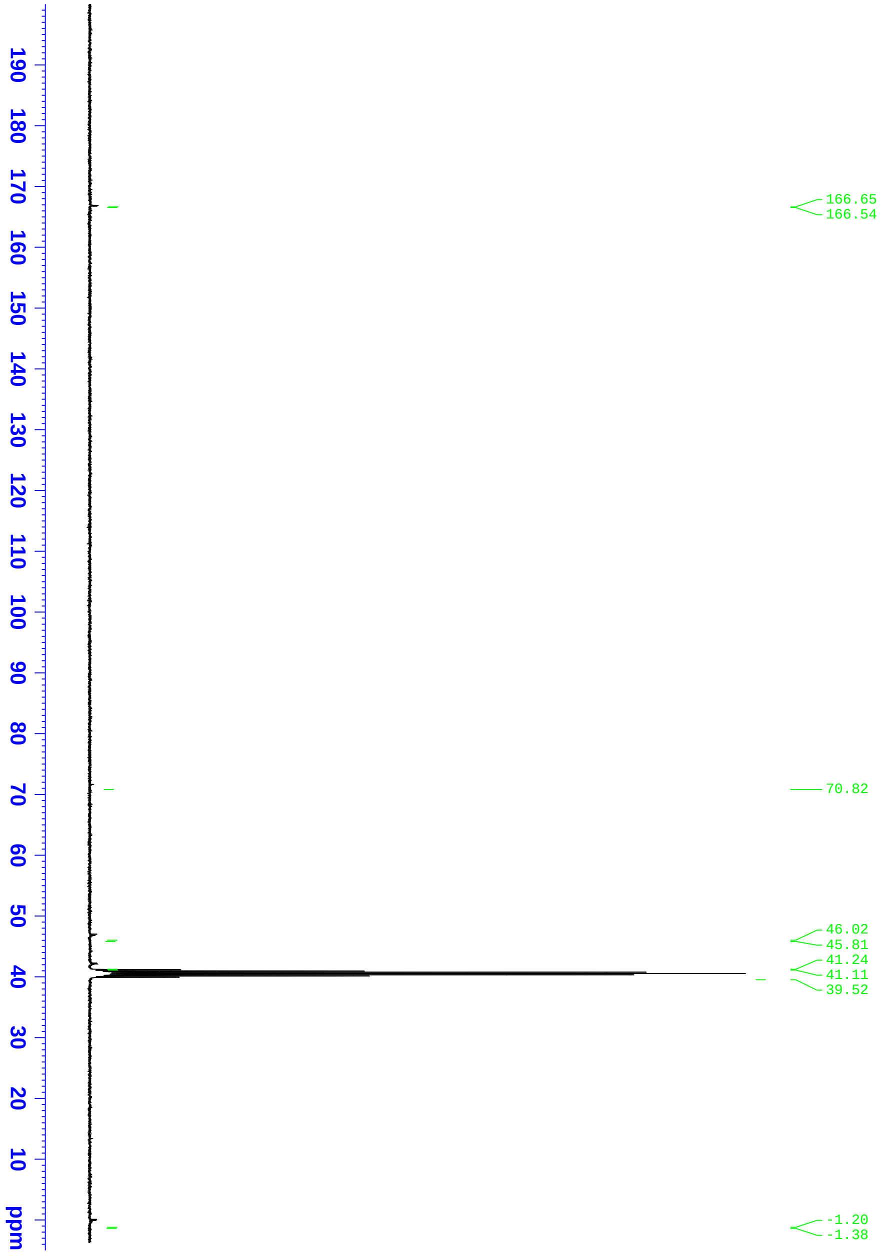
APPENDIX A

NMR SPECTRA

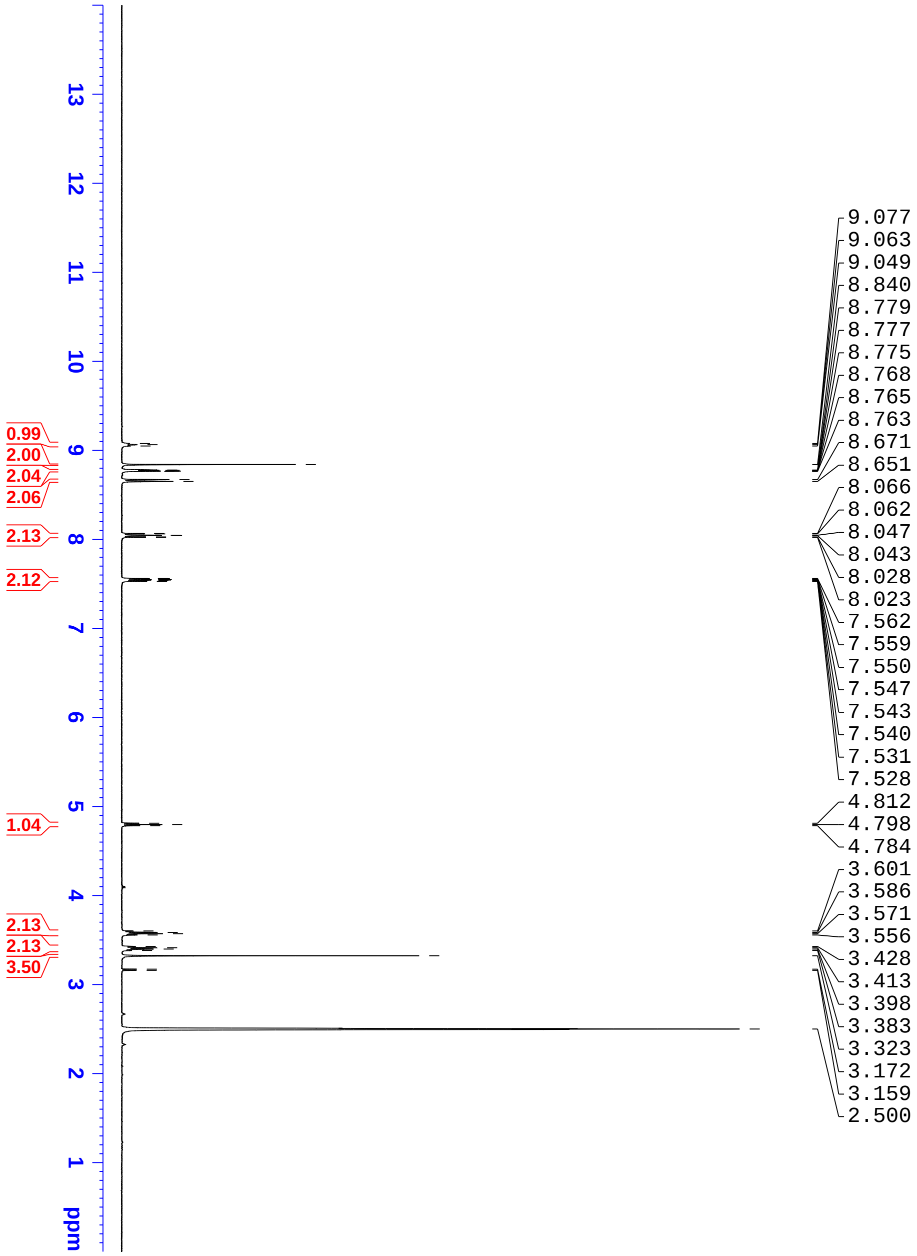
^{1}**HNMR** Linker **7**



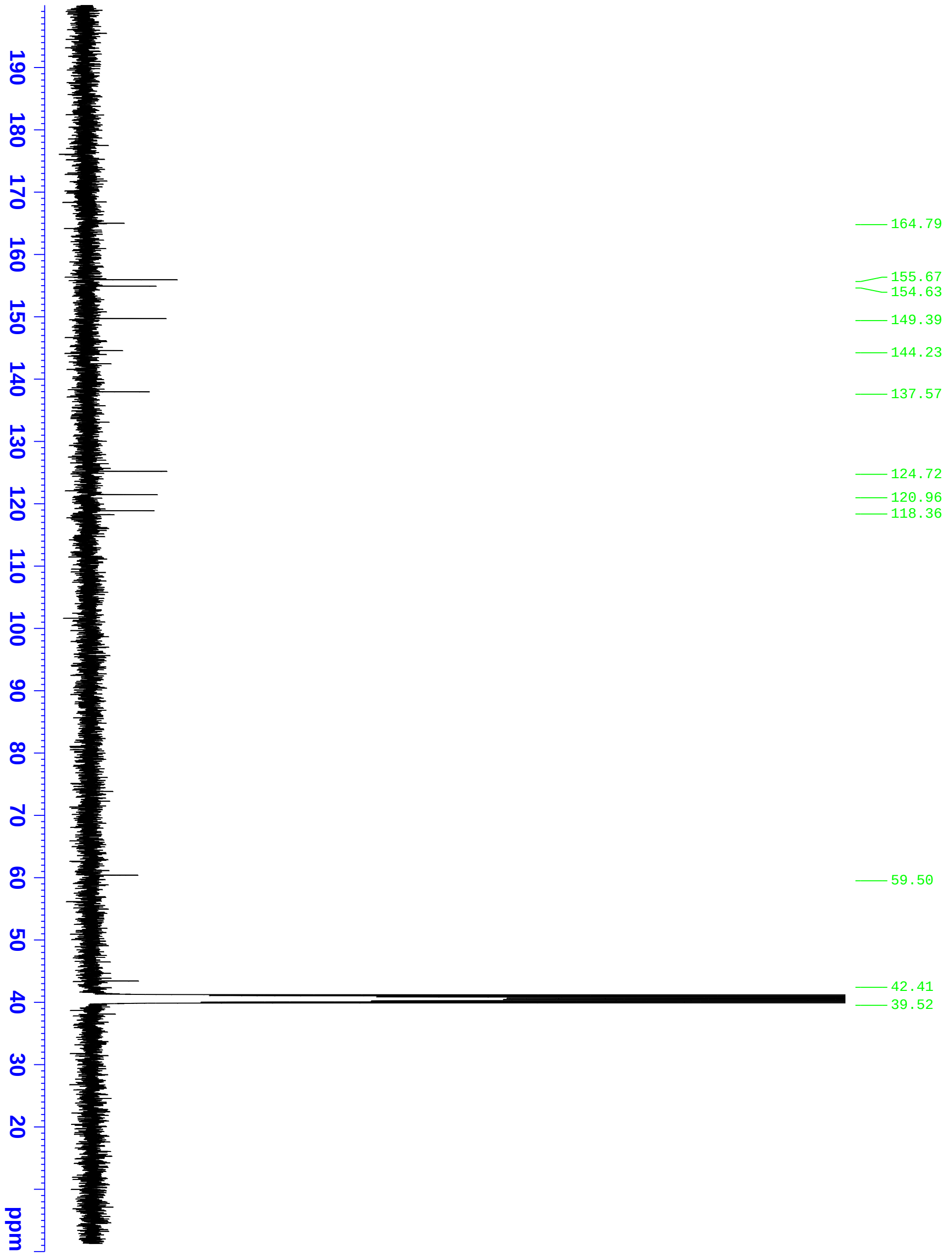
^{13}**CNMR** Linker **7**



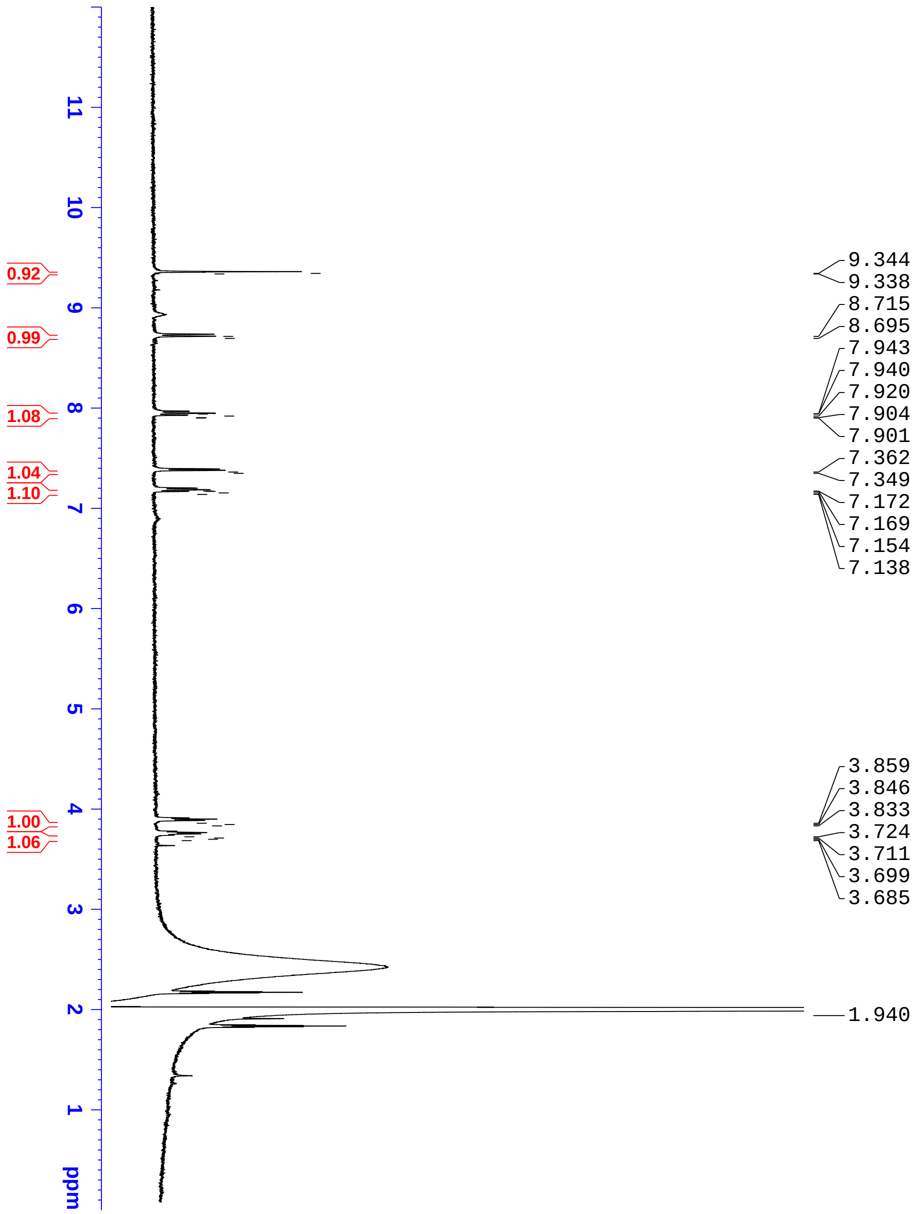
^{1}HNMR Ligand **40**



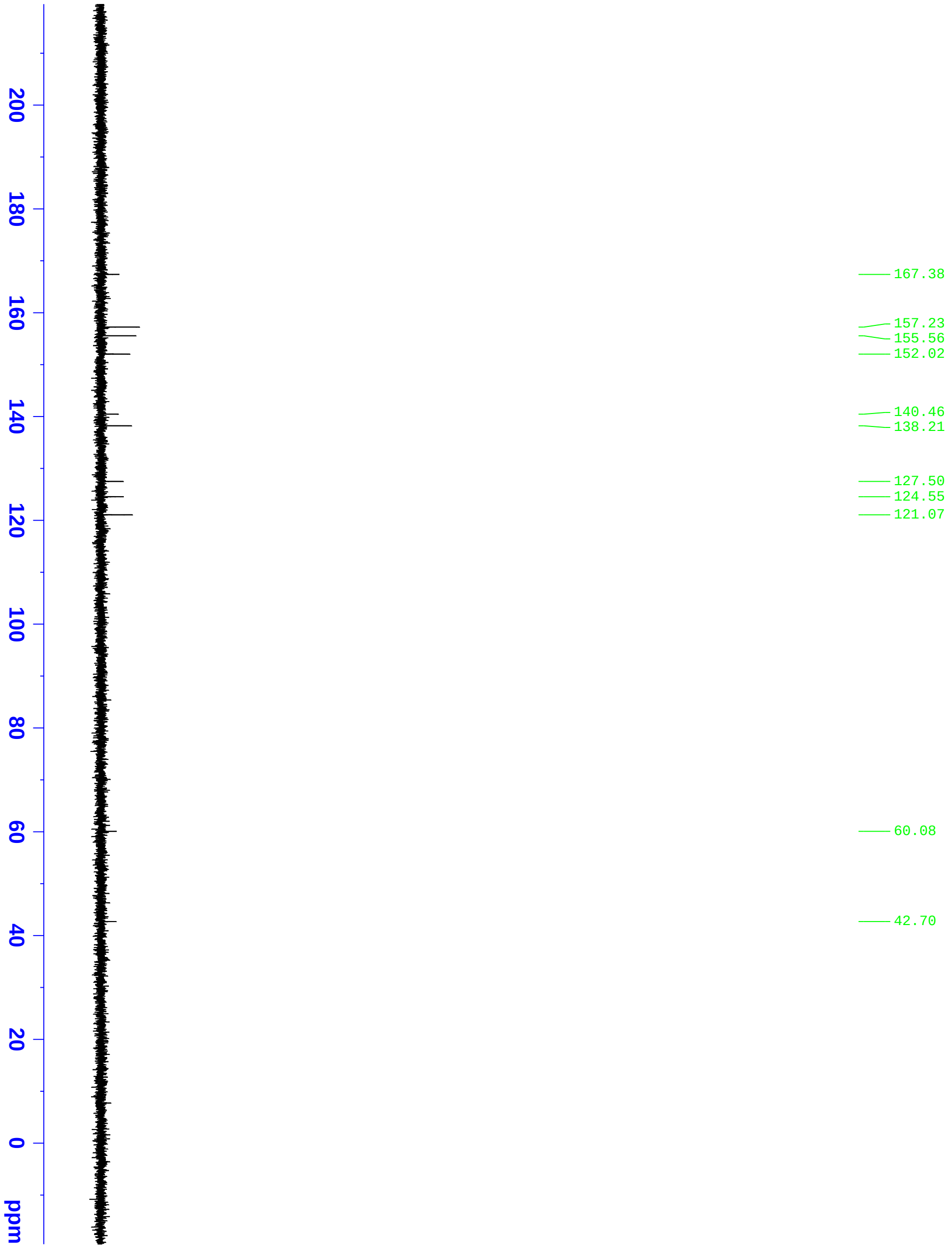
^{13}**CNMR** Ligand **40**



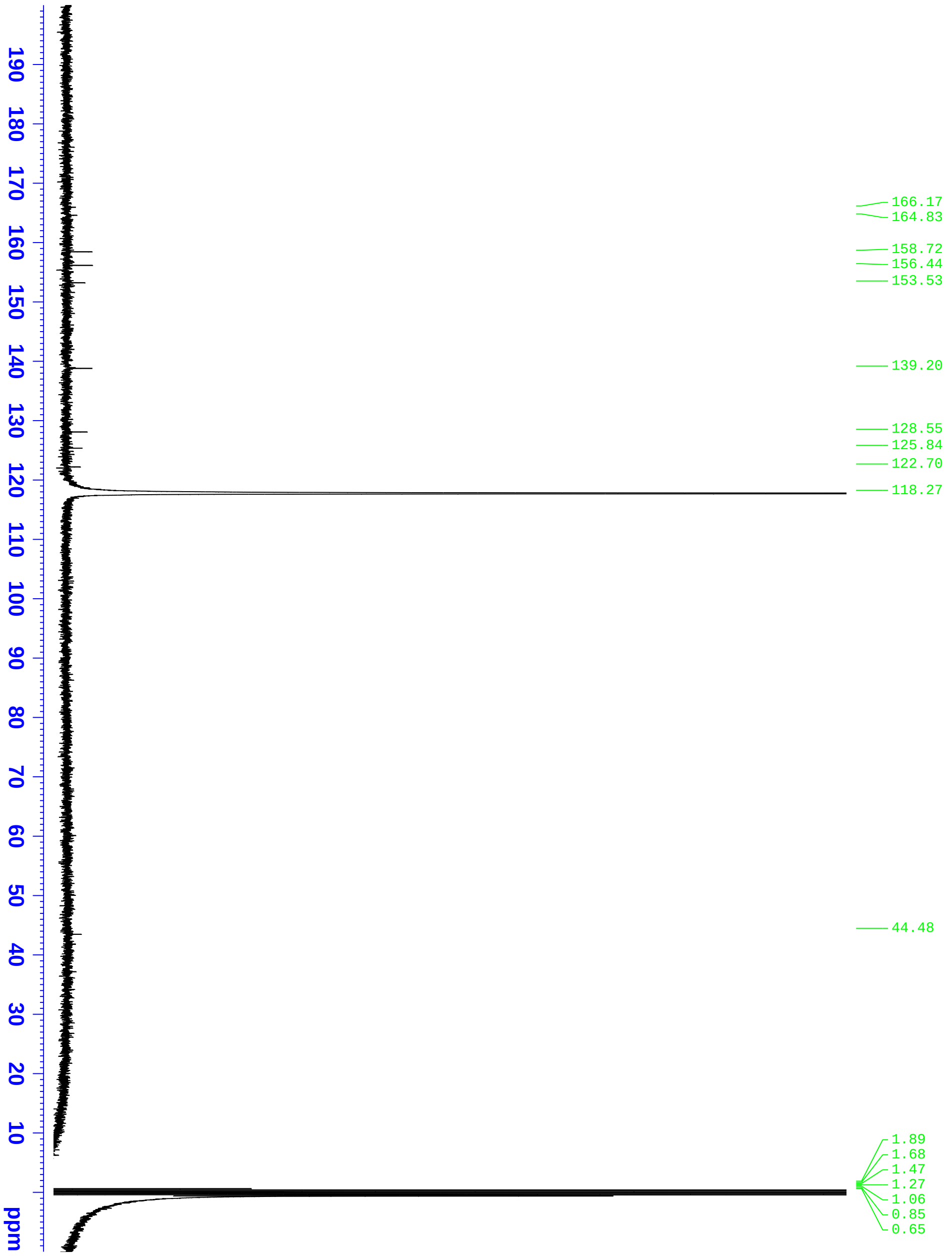
^{1}**HNMR** Complex **41**



^{13}**CNMR** Complex **41**
in 10% D_2O/90% H_2O, 0.1 M
phosphate buffer, pH 7



^{13}**CNMR** Complex **41**
in MeCN-d_3



APPENDIX B

MOLECULAR DYNAMICS-AIDED MODELLING SCRIPTS

Below are the scripts used to model and curate the results of the molecular dynamics simulations of the linkers described in Section 2.6.8. All scripts are written in Python and were run in Pymol on the Center for Simulation and Modeling's supercomputer. Inputs for each script are the output of the previous script. Script "ClashTest.py" (B.1) appends two coiled-coil dimers to each structure within the library of linker conformations generated by MD simulations then tests for steric clashes within each structure; only those without steric clashes are retained. Script "DblRngFilter.py" (B.2) removes any conformations inconsistent with the experimental DEER data. Script "Get Frames.py" (B.3) reduces the library of final conformations to a manageable number for visualization.

B.1 CLASHTEST.PY

```
#ClashTest3 changes:  added in clash checks for outer helix with inner helix terminii
#ClashTest4 changes: changes helix clash checking to a function and then runs using the two major rotamers of Cys
#ClashTest5 changes: now checks for linker atom clashes with adjacent helix
#ClashTest6 changes: now checks for outer to outer helix N and C terminii clashes
#ClashTest6 changes: Redefines CA-CA clashes as its own function
import pymol
from pymol import cmd
import sys,getopt

## AAD 2014-02-06
# Process command line arguments. There is no longer a need for separate py files
try:
```

```
 opts,args = getopt.getopt(sys.argv[1:],"hl:i:")
except getopt.GetoptError:
 print "Arguments are required"
 print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
 cmd.quit()
for opt,arg in opts:
 if opt == '-h':
  print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
  cmd.quit()
 elif opt == '-l':
  if arg == 'EDA':
   inputLinker = arg
  elif arg == 'PipC':
   inputLinker = 'PipC'
  elif arg == 'PipT':
   inputLinker = 'PipT'
  else:
   print "<linker> must be either EDA, PipC or PipT"
   print "  " + arg + " is not accepted"
   cmd.quit()
 elif opt == '-i':
  inputfile = arg


#Load Linker MD file
print("ClashTest6 started.")
print "  Linker is " + inputLinker
print "  Input file is " + inputfile

#DEFINE function LnkClshChk:
#Checks for linker atom clashes with the specified helix
def LnkClshChk (Helix, linker, ResChk, frame):
    clash = "no"
#Van der Waals radii obtained from  Table II of Seeliger, D. and de Groot, B. L. Proteins, 2007. 68: 595-601.
    HlxAtms = ['CB','CA','N','O',]
    #CB, CA, N and linker C's radii are the sum of the carbon/nitrogen's radius plus their hydrogen's radius
    HlxAtmRad = [3.08, 2.51, 2.42, 1.41]
    if (linker == 'EDA'):
        LnkrAtmRad = [3.08, 1.41, 2.42, 3.08, 3.08, 2.42, 1.41, 1.43]
            if (Helix == 'B'):
                CC = 'AB'
                LnkrAtms = ['CA1', 'O1', 'N1', 'C1', 'C2', 'O2', 'N2', 'CO1']
            elif (Helix == 'D'):
                CC = 'CD'
                LnkrAtms = ['CA2', 'O2', 'N2', 'C2', 'C1', 'O1', 'N2', 'CO2']
        else:
            print "Non-adjacent helix chosen."
            return
    elif (linker == 'PipC' or 'PipT'):
            LnkrAtmRad = [3.08, 1.41, 3.08, 3.08, 3.08, 3.08, 1.41, 1.43]
            if (Helix == 'B'):
                    CC = 'AB'
                    LnkrAtms = ['CA1', 'O1', 'CR1', 'CR2', 'CR3', 'CR4', 'O2', 'CO1']
            elif (Helix == 'D'):
```

```
                    CC = 'CD'
                    LnkrAtms = ['CA2', 'O2', 'CR3', 'CR4', 'CR1', 'CR2', 'O1', 'CO2']
        else:
            print "Non-adjacent helix chosen."
            return
        else:
        print "Linker undefined."
                return
    for resnum in ResChk:
        for hatms in HlxAtms:
            atomA = ("CC_" + CC + "_%05d and chain " + Helix + " and resi %d and name " + hatms)%(frame, resnum)
            for lnkatms in LnkrAtms:
                if frame >= 10000:
                    atomB = ("traj_%05d and chain Z and name "+lnkatms)%frame
                else:
                    atomB = ("traj_%04d and chain Z and name "+lnkatms)%frame
                #Distance cutoff for specific atom types from Table III of Seeliger, D. and de Groot, B. L. Proteins,
                 2007. 68: 595-601.
                #Default is sum of atom van der Waal's radii
                if (hatms == 'O' and lnkatms == 'N1' or 'N2') or (hatms == 'N' and lnkatms == 'O1' or 'O2'):
                    cutoff = 2.82
                elif (hatms == 'O' and lnkatms == 'O1' or 'O2'):
                    cutoff = 3.3
                elif (hatms == 'CA' and lnkatms == 'O1' or 'O2'):
                    cutoff = 3.18
                else:
                    cutoff =  sum (hatms + lnkatms)
                atmdist = cmd.dist("tmpdist", atomA, atomB)
                if (atmdist < cutoff and atmdist > 0):
                                clash = "yes"
                    print "Linker clash of %g Angstroms from a %g Angstroms cutoff."%(atmdist, cutoff)
                    break
                elif (atmdist >= cutoff):
                    clash = "no"
                else:
                    clash = "Error"
                    break
            if clash == "yes":
                break
        if clash == "yes":
            break
        return clash

#DEFINE function CaClshTst:
def CaClshTst(Helix1, Helix2, frame, CC1='AB', CC2='CD', TotResHlx1=30, TotResHlx2=30, cutoff=5.5):
    clash = "no"
    for resnumA in xrange(1, TotResHlx1+1):
            atomA = ("CC_" + CC1 + "_%05d and chain " + Helix1 + " and resi %d and name CA")%(frame, resnumA)
        for resnumB in xrange(1, TotResHlx2+1):
                atomB = ("CC_" + CC2 + "_%05d and chain " + Helix2 + " and resi %d and name CA")%(frame, resnumB)
                cadist = cmd.dist("tmpdist", atomA, atomB)
                cmd.delete("tmpdist")
            if cadist < cutoff and cadist > 0:
                clash = "yes"
```

```
                print "%c/%c clash of %g Angstroms."%(Helix1, Helix2, cadist)
                break
            elif cadist < 0:
                print cadist
                clash = "Error"
                print "Error in starting pdb file %c or %c chain."%(Helix1, Helix2)
                break
        if clash == "yes":
            break
    return clash

#DEFINE function ClshTst:
# create n identical objects for CoiledCoilAB and CoiledCoilCD [n must match number of states in traj.pdb]
# fit CA+CB+SG for each copy of CoiledCoilAB and CoiledCoilCD to the corresponding atoms in one of the states of
traj.pdb
#merge all CoiledCoilAB into single object and merge all CoilCoilCD into single object
def ClshTst (CC1, CC2, states, cutoff=5.5, totresB=30, totresD=30, clashnum=0, finsts=0):
    for idx in xrange(1,states+1):
            print "Frame: %d"%idx
        cmd.copy("CC_AB_%05d"%idx,CC1)
        cmd.copy("CC_CD_%05d"%idx,CC2)
        if idx >= 10000:
            cmd.pair_fit("CC_AB_%05d and resn CYS and name CA"%idx,"traj_%05d and name
CM1"%idx,"CC_AB_%05d and resn CYS and name SG"%idx,"traj_%05d and chain Z and name S1"%idx,"CC_AB_%05d
and resn CYS and name CB"%idx,"traj_%05d and chain Z and name CE1"%idx)
                cmd.pair_fit("CC_CD_%05d and resn CYS and name CA"%idx,"traj_%05d and name
CM2"%idx,"CC_CD_%05d and resn CYS and name SG"%idx,"traj_%05d and chain Z and name S2"%idx,"CC_CD_%05d
and resn CYS and name CB"%idx,"traj_%05d and chain Z and name CE2"%idx)
        else:
            cmd.pair_fit("CC_AB_%05d and resn CYS and name CA"%idx,"traj_%04d and name
CM1"%idx,"CC_AB_%05d and resn CYS and name SG"%idx,"traj_%04d and chain Z and name S1"%idx,"CC_AB_%05d
and resn CYS and name CB"%idx,"traj_%04d and chain Z and name CE1"%idx)
            cmd.pair_fit("CC_CD_%05d and resn CYS and name CA"%idx,"traj_%04d and name
CM2"%idx,"CC_CD_%05d and resn CYS and name SG"%idx,"traj_%04d and chain Z and name S2"%idx,"CC_CD_%05d
and resn CYS and name CB"%idx,"traj_%04d and chain Z and name CE2"%idx)
        clash = "no"
        #Check for linker clashes with adjacent helices
        ResChk = [10, 11, 13, 15, 17, 18]
        LnkrHlcs = ['B', 'D']
        print "Linker/helix clash test running."
        for Helix in LnkrHlcs:
            clash = LnkClshChk (Helix=Helix, linker=linker , ResChk = ResChk, frame=idx)
            if clash == "yes":
                break
        print "Linker/helix clash test finished."
        if clash == "no":
            CaHlcs = [('B','D'), ('B','C'),('A','D'),('A','C')]
            for Helix1, Helix2 in CaHlcs:
                    clash = CaClshTst(Helix1=Helix1, Helix2=Helix2, frame=idx)
                if clash == "yes":
                    break
        if clash == "yes":
            clashnum = clashnum + 1
            print "Clash %d found."%clashnum
```

```
            cmd.delete("CC_AB_%05d"%idx)
            cmd.delete("CC_CD_%05d"%idx)
            if idx >= 10000:
                cmd.delete("traj_%05d"%idx)
            else:
                    cmd.delete("traj_%04d"%idx)
            cmd.delete("site1")
            cmd.delete("tmpdist")
                continue
        elif clash == "no":
            print "No clashes."
                finsts = finsts + 1
                cmd.create("CC_AB_MD","CC_AB_%05d"%idx,0,finsts)
                cmd.create("CC_CD_MD","CC_CD_%05d"%idx,0,finsts)
                cmd.delete("CC_AB_%05d"%idx)
                cmd.delete("CC_CD_%05d"%idx)
            if idx >= 10000:
                cmd.create("Linker_MD","traj_%05d"%idx,0,finsts)
                cmd.delete("traj_%05d"%idx)
            else:
                cmd.create("Linker_MD","traj_%04d"%idx,0,finsts)
                    cmd.delete("traj_%04d"%idx)
                cmd.delete("site1")
            cmd.delete("tmpdist")
                continue
        else:
            print "Error! Program terminated early."
            break
    ClshRslt = [clashnum]
    return ClshRslt

#Define rotamer combinations to check
CysRotComb = [(1,1), (1,2), (2,2)]
comb = 1
nxtfrm = 1

#Initialize clashnum variable
clashnum = 0

#Check for clashes in each rotamer combination and output non-clashing conformations
for ABrot, CDrot in CysRotComb:
    cmd.load(inputfile,"traj")
    linker = inputLinker
    cmd.alter("traj", "chain='Z'")
    print "Linker trajectory file opened successfully."

    #Count number of linker conformations
    states = cmd.count_states("traj")
    cmd.split_states("traj")
    cmd.delete("traj")

    #load coiled-coil chain A and B
    cmd.load("PDB/CoiledCoilAB_r%d.pdb"%ABrot,"CoiledCoilAB")
```

```
    #load coiled-coil chain C and D
    cmd.load("PDB/CoiledCoilCD_r%d.pdb"%CDrot, "CoiledCoilCD")

    #Run Clash Test for rotamer combination
    ClshTstRslt = ClshTst(CC1="CoiledCoilAB", CC2="CoiledCoilCD", states=states, clashnum=clashnum)

    #Merge non-clashing coiled coils and linker traj into one trajectory object
    print "Merging linker and coiled coil trajectories..."
    cmd.create("CCSMP_TempMD%d"%comb, "CC_AB_MD or CC_CD_MD or Linker_MD")
    cmd.create("CCSMP_FinalMD", "CCSMP_TempMD%d"%comb,0,nxtfrm)
    nxtfrm = nxtfrm + cmd.count_states("CCSMP_TempMD%d"%comb)

    #Reset and clashnum
    clashnum = ClshTstRslt[0]

    #Delete old objects
    cmd.delete("CoiledCoilAB")
    cmd.delete("CoiledCoilCD")
    cmd.delete("CC_AB_MD")
    cmd.delete("CC_CD_MD")
    cmd.delete("Linker_MD")
    cmd.delete("tmpdist")
    comb = comb + 1

#Output final number of clashes
print "\nThere were %d frames with clashes out of %d frames checked."%(clashnum, states*(comb-1))
print "There were %d frames with no clashes."%(states*(comb-1)-clashnum)
#remove extraneous atoms
print "Removing extraneous or duplicitous atoms..."
cmd.remove("CCSMP_FinalMD and chain Z and name HM1+HM2+HM3+HM4+HM5+HM6+CM1+CM2+CE1+CE2")
cmd.remove("CCSMP_FinalMD and resn CYS and name SG+2HB+3HB+HG")

#Save Non-clashing frames
print "Saving final output..."
basename = inputfile.split(".")[0]

cmd.save (basename+"."+inputLinker+".Final.pdb", "CCSMP_FinalMD", 0)
cmd.save (basename+"."+inputLinker+".Finalr1r1.pdb","CCSMP_TempMD1",0)
cmd.save (basename+"."+inputLinker+".Finalr1r2.pdb","CCSMP_TempMD2",0)
cmd.save (basename+"."+inputLinker+".Finalr2r2.pdb","CCSMP_TempMD3",0)
print "Clash test finished."
```

B.2 DBLRNGFILTER.PY

```
import pymol
from pymol import cmd
import sys,getopt
```

```
## NAT 2014-03-02
# Process command line arguments. There is no longer a need for separate py files


try:
 opts,args = getopt.getopt(sys.argv[1:],"hl:i:")
except getopt.GetoptError:
 print "Arguments are required"
 print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
 cmd.quit()
for opt,arg in opts:
 if opt == '-h':
  print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
  cmd.quit()
 elif opt == '-l':
  if arg == 'EDA':
   inputLinker = arg
  elif arg == 'PipC':
   inputLinker = 'PipC'
  elif arg == 'PipT':
   inputLinker = 'PipT'
  else:
   print "<linker> must be either EDA, PipC or PipT"
   print "  " + arg + " is not accepted"
   cmd.quit()
 elif opt == '-i':
  inputfile = arg

#The following function, DistWin, filters conformations that fall into a given distance range between two atoms
def DistWin (resnum, states, start, end, objname, finsts=1, AtmNamA='SPN', AtmNamB='SPN', ChainA='A', ChainB='C'):
    startnm = str(start/10.0)
    endnm = str(end/10.0)
    finalname = objname + startnm + "-" + endnm + "nm"
    for frame in xrange(1,states+1):
            print "Frame: %d"%frame
        atomA = (objname + "%04d and chain " + ChainA + " and resi %d and name " + AtmNamA)%(frame, resnum)
        atomB = (objname + "%04d and chain " + ChainB + " and resi %d and name " + AtmNamB)%(frame, resnum)
        SpnDis = cmd.dist("tmpdist", atomA, atomB)
        cmd.delete("tmpdist")

        if SpnDis >= start and SpnDis <= end:
            cmd.create(finalname,objname + "%04d"%frame,0,finsts)
            finsts = finsts + 1
            print "This one's on the team!"
            print "Team members: %d"%finsts
        else:
            print ("You don't make the cut!")
            cmd.delete(objname + "%04d"%frame)
    return finsts, finalname
def RngFilt (resnum, Ranges, outname, inname):
#Set distance range selections here in Angstroms:
    for start, end in Ranges:
        #Count number of linker conformations
        states = cmd.count_states(inname)
```

```
        cmd.split_states(inname)
        cmd.delete(inname)
        FinConf, finalname= DistWin(resnum=resnum, states=states, start=start, end=end, objname=(inname + "_"))
        startnm = start/10.0
        endnm = end/10.0
        print "There were %d conformers in the range %g to %g nm."%(FinConf, startnm, endnm)
        cmd.create(outname + finalname, finalname, 0)
        nextname = (outname + finalname)
    return nextname
cmd.load(inputfile, "ConfLib")
print "Conformer library opened successfully."
if inputLinker == 'EDA':
  Range1 = [(36, 45)]
  Range2 = [(32.7, 50.9)]
  nextname1 = RngFilt(resnum=11, Ranges = Range1, outname="Inner_", inname="ConfLib")
  print nextname1
  nextname2 = RngFilt(resnum=4, Ranges = Range2, outname="NTerm+", inname = nextname1)
  print nextname2
  cmd.save("DblFilt"+inputLinker+".pdb", nextname2, 0)
elif inputLinker == 'PipC' or 'PipT':
  Range1 = [(31.8,47.8)]
  Range2 = [(30.1,53.9)]
  Range3 = [(19.7,26.7)]
  nextname1 = RngFilt(resnum=11, Ranges = Range1, outname="Inner_", inname="ConfLib")
  cmd.save(InputLinker+"InnerFilt.pdb", nextname1)
  print nextname1
  nextname2 = RngFilt(resnum=4, Ranges = Range2, outname="NTermLong+", inname = nextname1)
  print nextname2
  cmd.save("DblFilt"+inputLinker+".pdb", nextname2, 0)
  cmd.load(InputLinker+"InnerFilt.pdb", InputLinker+"InnerFilt")
  nextname3 = RngFilt(resnum=4, Ranges = Range3, outname="NTermShort+", inname = InputLinker+"InnerFilt")
  print nextname3
  cmd.save("DblFilt"+inputLinker+".pdb", nextname3, 0)

import pymol
from pymol import cmd
import sys,getopt
```

B.3 GETFRAMES.PY

```
## NAT 2014-10-04
# This script outputs a smaller set of frames from a large trajectory file.

try:
  opts,args = getopt.getopt(sys.argv[1:],"hl:i:")
except getopt.GetoptError:
  print "Arguments are required"
```

```
  print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
  cmd.quit()
for opt,arg in opts:
  if opt == '-h':
    print "Usage: pymol -qrc ClashTest.py -- -l <linker> -i <input.pdb>"
    cmd.quit()
  elif opt == '-l':
    if arg == 'EDA':
      inputLinker = arg
    elif arg == 'PipC':
      inputLinker = 'PipC'
    elif arg == 'PipT':
      inputLinker = 'PipT'
    else:
      print "<linker> must be either EDA, PipC or PipT"
      print "  " + arg + " is not accepted"
      cmd.quit()
  elif opt == '-i':
    inputfile = arg

def GetFrames (frames_out, inname):
  states = cmd.count_states(inname)
  remainder = states % frames_out
  states = states - remainder
  step = states / frames_out
  cmd.split_states(inname)
  cmd.delete(inname)
  idx = 1
  for frames in xrange(1, states, step):
    cmd.create(inputfile+"_SmallSet", "ConfLib_%04d"%frames, 0, idx)
    idx = idx + 1
    print idx
  return
cmd.load(inputfile, "ConfLib")
print "Conformer library opened successfully."
if inputLinker == 'EDA':
  GetFrames(frames_out = 1000, inname = "ConfLib")
elif inputLinker == 'PipC' or 'PipT':
  GetFrames(frames_out = 500, inname = "ConfLib")
cmd.save("DblFilt"+inputLinker+"_SmallSet.pdb",inputfile+"_SmallSet", 0)
cmd.quit
```

APPENDIX C

A SET OF TOAC-LABELED COILED-COIL HOMODIMERS

A side project springing from collaboration with the Saxena lab sought to measure the dihedral angles between pairs of TOAC spin-labels using orientational selective DEER experiments within the context of a coiled coil dimer (GCN4p1). Three peptides (Figure 87) were synthesized with the spin-label in the second heptad at three different solvent-exposed sites (heptad positions *b*, *c*, and *f*).

....|....|....|....|....|....|...

gabcdefgabcdefgabcdefgabcdefgabcd

Ac-RMKQLEDKV**Σ**ELLSKNYHLENEVARLKKLVGER-NH_2 (**C1**)

Ac-RMKQLEDKVE**Σ**LLSKNYHLENEVARLKKLVGER-NH_2 (**C2**)

Ac-RMKQLEDKVEELL**Σ**KNYHLENEVARLKKLVGER-NH_2 (**C3**)

Figure 87. Sequences of peptides **C1-C3**. The spin-labeled amino acid TOAC (Σ) is denoted in bold.

Detailed below is the synthesis, purification and characterization of each peptide along with sample preparation for the EPR measurements.

C.1 TOAC-LABELED PEPTIDE SYNTHESIS AND PURIFICATION

C.1.1 General Information

Solvents and all other reagents were purchased from Acros Organics, Aldrich, Advanced Chem Tech, Baker, EMD, Fluka, or Fisher and used without further purification. NovaPEG Rink Amide Resin, Fmoc-protected amino acids, HATU, PyBOP, and PyBrOP were purchased from Novabiochem. HCTU was purchased from Aapptec or Novabiochem. PyAOP was purchased from AK Scientific. 2,2,6,6-Tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid (TOAC) was purchased from Santa Cruz Biotechnology or Toronto Research Chemicals. MALDI of peptides was collected on an AB Sciex Voyager DE Pro MALDI-TOF.

C.1.2 Peptide Synthesis

Peptides **C1** and **C2** were synthesized using solid-phase Fmoc methods on an automated synthesizer at room temperature for residues with residues 8-11 coupled manually with microwave-assisted heating (*Method 1*). Peptide **C3** was synthesized manually using microwave-assisted heating (*Method 2*). Automated couplings and deprotections were performed on a PTI Tribute automated synthesizer while microwave-assisted couplings were performed on a CEM Microwave-Assisted Reaction System (MARS).

Method 1. In a standard automated coupling reaction, 2.5 mL of a solution composed of 0.2 M HCTU, 0.4 M N-methylmorpholine in DMF was added to 7 equivalents of Fmoc-amino acid relative to resin. After a 2 minute pre-activation, the solution was added to the resin and vortexed for 45 minutes. Deprotections were performed by two treatments with 3 mL of 20% v/v

4-methylpiperidine in DMF for 4 minutes each. The resin was washed four times with 3 mL of DMF after each coupling and deprotection step. After residue 12, the synthesizer was paused and residues 8-11 were synthesized manually using micrwave-assisted heating. TOAC was coupled as the free amino acid with PyBOP (5 equiv TOAC, 4.9 equiv PyBOP, 12.5 equiv DIEA in NMP) in the microwave with a 2 minute ramp to 70 °C and an 8 minute hold. For peptide **C1**, residue 9 was double coupled using TFFH (5 equiv Fmoc-amino acid, 5equiv TFFH, 12.5 equiv DIEA). For peptide **C2**, residues 10 and 9 were double coupled using TFFH. For both peptides, residue 8 (Lys) was double coupled with PyAOP (5 equiv Fmoc-Lys-OH, 5 equiv PyAOP, 7.5 equiv DIEA). Fmoc deprotections were performed with one treatment of 20% v/v 4-methylpiperidine in DMF with a 2 minute ramp to 80° C and a 2 minute hold. Resin was washed 3 times with DMF between steps.

<u>*Method 2.*</u> Standard coupling solutions consisted of 5 equivalents of Fmoc- protected amino acid with 4.9 equivalents of HCTU and 7.5 equivalents of DIEA in NMP. Coupling solutions were pre-activated for 2 minutes prior to addition to the resin then subjected to a 1.5 minute ramp to 90° C with a 2 minute hold. Fmoc deprotections were performed with 20% 4-methylpiperidine in DMF with a 2 minute ramp to 80° C with a 2 minute hold. Resin was washed 3 times with DMF between steps. Residue 18 (His) was coupled by stirring at room temperature for 45 minutes. TOAC at position 14 was coupled the same way as in *Method 1*. Leu at position 13 was coupled as the activated acid fluoride; Fmoc-Leu-F was prepared by a known method.[166]

After the final Fmoc deprotection, the N-terminus of each peptide was acetylated with a solution of DMF/DIEA/Ac_2O (8:2:1 by volume) for 20 minutes at room temperature. The resin was rinsed three times with 3 mL each DMF, DCM, and MeOH. The resin was dried in a vacuum desiccator for 20 minutes prior to TFA cleavage. Cleavage conditions for all peptides was

TFA/H_2O/Anisole/TIS (91/3/3/3 v/v/v/v) for 4.5 hours. Peptides were then precipitated in ~40 mL of chilled Et_2O, centrifuged and decanted resulting in a crude solid. The nitroxide radical was regenerated by treatment with of 10% NH_4OH for ~4 hours.

Peptides were purified by reverse phase HPLC on a C_{18} column then lyophilized. Identity and purity was confirmed by MALDI-MS (Table 14) and analytical HPLC (Figure 88), respectively. The impurity prior to the main peak in peptide C1 is due to the a small population of the TOAC residue in the hydroxylamine form rather than nitroxide.

Table 14. Calculated and observed masses of peptides **C1-C3**.

#	$[M+H]^+$ *m/z* (average)	
	Calculated	**Observed**
C1	4106.8	4106.3
C2	4106.8	4105.3
C3*	4163.5 $[M+NH_4]^+$	4164.2

*Most abundant peak was $[M+NH_4]^+$ for peptide **C3** due to treatment with NH_4OH.

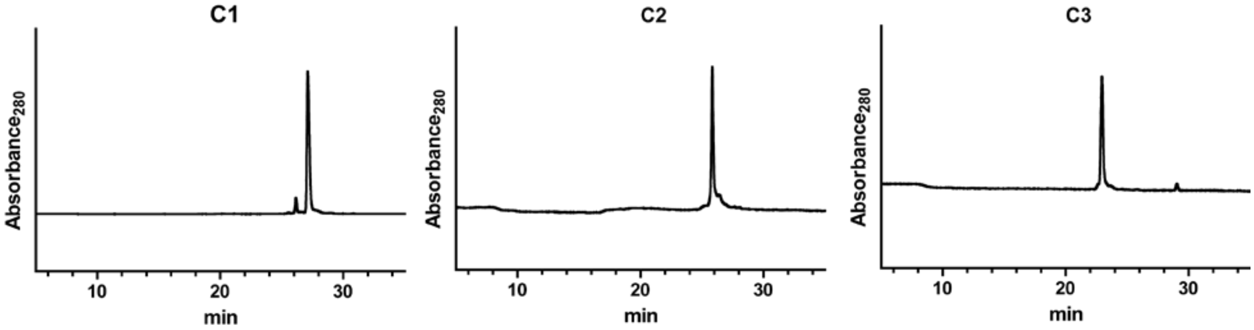


Figure 88. Analytical HPLC chromatograms of purified peptides **C1-C3**.

C.2 BIOPHYSICAL MEASUREMENTS

C.2.1 Circular Dichroism Melts and Scans

CD spectroscopy measurements were carried out in 1 mm quartz cuvettes on an Olis DSM 17 CD spectrometer. Peptide stock concentrations were determined by UV-Vis spectroscopy. Buffer components were syringe filtered through 0.22 µm pore filter prior to sample preparation. All measurements were made with a 2 nm bandwidth and 5 second integration time. CD scans were measured from 200-260 nm at 20 °C. Thermal melts were measured at 222 nm with a baseline of 260 nm from 2-98 °C with 4° increments. Samples were incubated at each temperature for 2 minutes. Thermal melt data was fit to a two-state unfolding model[169] using GraphPad Prism.

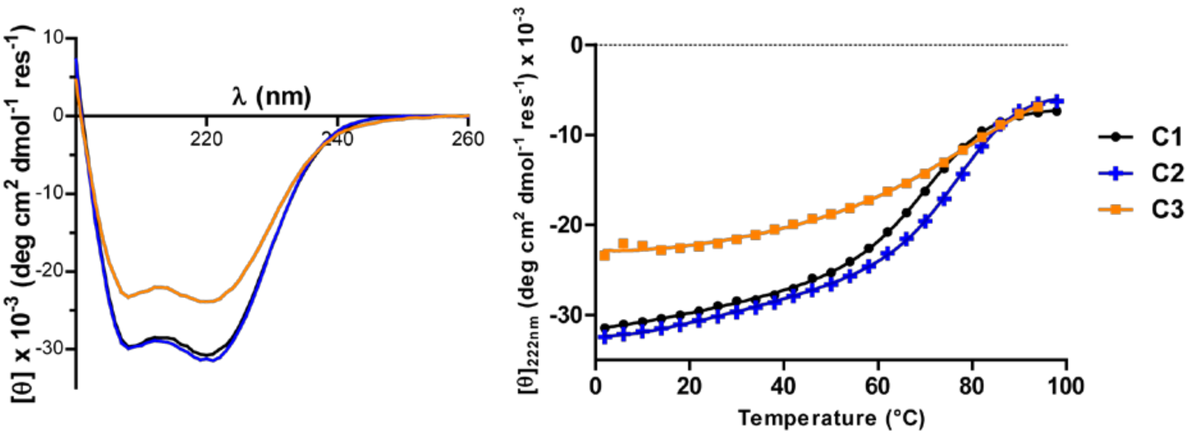


Figure 89. CD scans and melts of peptides **C1-C3**.

C.2.2 Sample Preparation for DEER

Peptide sample were prepared at a concentration of 100 µM in aqueous HEPES buffer (10 mM, pH 7) with 20% v/v glycerol as a cryoprotectant. Samples were then given to Dr. K. Ishara Silva for orientational selective DEER measurements.

REFERENCES

1. Krieg, E.; Bastings, M. M. C.; Besenius, P.; Rybtchinski, B., *Chem. Rev.* **2016,** *116*, 2414-2477.
2. Yang, L.; Tan, X.; Wang, Z.; Zhang, X., *Chem. Rev.* **2015**.
3. Herbst, F.; Döhler, D.; Michael, P.; Binder, W. H., *Macromol. Rapid Commun.* **2013,** *34*, 203-220.
4. Wei, P.; Yan, X.; Huang, F., *Chem. Soc. Rev.* **2015,** *44*, 815-32.
5. Fouquey, C.; Lehn, J.-M.; Levelut, A.-M., *Adv. Mater.* **1990,** *2*, 254-257.
6. Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W., *Science* **1997,** *278*, 1601.
7. Coulibaly, S.; Heinzmann, C.; Beyer, F. L.; Balog, S.; Weder, C.; Fiore, G. L., *Macromolecules* **2014,** *47*, 8487-8496.
8. Cantekin, S.; de Greef, T. F. A.; Palmans, A. R. A., *Chem. Soc. Rev.* **2012,** *41*, 6125-6137.
9. Besenius, P.; Portale, G.; Bomans, P. H. H.; Janssen, H. M.; Palmans, A. R. A.; Meijer, E. W., *Proc. Natl. Acad. Sci. USA* **2010,** *107*, 17888-17893.
10. Albertazzi, L.; van der Zwaag, D.; Leenders, C. M. A.; Fitzner, R.; van der Hofstad, R. W.; Meijer, E. W., *Science* **2014,** *344*, 491-495.
11. Aida, T.; Meijer, E. W.; Stupp, S. I., *Science* **2012,** *335*, 813-817.
12. Weingarten, A. S.; Kazantsev, R. V.; Palmer, L. C.; Fairfield, D. J.; Koltonow, A. R.; Stupp, S. I., *J. Am. Chem. Soc.* **2015**.
13. Weingarten, A. S.; Kazantsev, R. V.; Palmer, L. C.; McClendon, M.; Koltonow, A. R.; SamuelAmanda, P. S.; Kiebala, D. J.; Wasielewski, M. R.; Stupp, S. I., *Nat. Chem.* **2014,** *6*, 964-970.
14. Pedersen, C. J., *J. Am. Chem. Soc.* **1967,** *89*, 7017-7036.
15. Cram, D. J.; Cram, J. M., *Science* **1974,** *183*, 803-9.
16. Yu, G.; Jie, K.; Huang, F., *Chem. Rev.* **2015,** *115*, 7240-7303.
17. Virovets, A. V.; Blatov, V. A.; Shevchenko, A. P., *Acta Crystallogr., Sect. B: Struct. Sci* **2004,** *60*, 350-357.
18. Harada, A., *Carbohydr. Polym.* **1997,** *34*, 183-188.
19. Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A., *Chem. Rev.* **2015,** *115*, 12320-12406.
20. Kondo, M.; Yoshitomi, T.; Matsuzaka, H.; Kitagawa, S.; Seki, K., *Angew. Chem. Int. Ed.* **1997,** *36*, 1725-1727.
21. Li, H.; Eddaoudi, M.; Groy, T. L.; Yaghi, O. M., *J. Am. Chem. Soc.* **1998,** *120*, 8571-8572.
22. Ferey, G., *Chem. Soc. Rev.* **2008,** *37*, 191-214.
23. Cook, T. R.; Zheng, Y. R.; Stang, P. J., *Chem. Rev.* **2013,** *113*, 734-77.
24. Hoskins, B. F.; Robson, R., *J. Am. Chem. Soc.* **1989,** *111*, 5962-5964.
25. Tian, Y.-K.; Chen, L.; Tian, Y.-J.; Wang, X.-Y.; Wang, F., *Polym. Chem.* **2013,** *4*, 453-457.
26. Yaghi, O. M.; Li, G.; Li, H., *Nature* **1995,** *378*, 703-706.
27. Seo, J. S.; Whang, D.; Lee, H.; Jun, S. I.; Oh, J.; Jeon, Y. J.; Kim, K., *Nature* **2000,** *404*, 982-986.

28. Rosi, N. L.; Eckert, J.; Eddaoudi, M.; Vodak, D. T.; Kim, J.; O'Keeffe, M.; Yaghi, O. M., *Science* **2003,** *300*, 1127-1129.
29. Lee, C. Y.; Farha, O. K.; Hong, B. J.; Sarjeant, A. A.; Nguyen, S. T.; Hupp, J. T., *J. Am. Chem. Soc.* **2011,** *133*, 15858-15861.
30. Ma, L.; Falkowski, J. M.; Abney, C.; Lin, W., *Nat. Chem.* **2010,** *2*, 838-846.
31. Bloch, E. D.; Queen, W. L.; Krishna, R.; Zadrozny, J. M.; Brown, C. M.; Long, J. R., *Science* **2012,** *335*, 1606-1610.
32. Stavila, V.; Schneider, C.; Mowry, C.; Zeitler, T. R.; Greathouse, J. A.; Robinson, A. L.; Denning, J. M.; Volponi, J.; Leong, K.; Quan, W.; Tu, M.; Fischer, R. A.; Allendorf, M. D., *Adv. Funct. Mater.* **2016,** *26*, 1699-1707.
33. Zhang, G.; Wei, G.; Liu, Z.; Oliver, S. R. J.; Fei, H., *Chem. Mater.* **2016,** *28*, 6276-6281.
34. Goswami, S.; Ma, L.; Martinson, A. B. F.; Wasielewski, M. R.; Farha, O. K.; Hupp, J. T., *ACS Appl. Mater. Interfaces* **2016,** *8*, 30863-30870.
35. Rothemund, P. W. K., *Nature* **2006,** *440*, 297-302.
36. Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H., *Nat. Methods* **2011,** *8*, 221-229.
37. Torring, T.; Voigt, N. V.; Nangreave, J.; Yan, H.; Gothelf, K. V., *Chem. Soc. Rev.* **2011,** *40*, 5636-5646.
38. van Meer, G.; Voelker, D. R.; Feigenson, G. W., *Nat. Rev. Mol. Cell Biol.* **2008,** *9*, 112-124.
39. Pieters, B. J. G. E.; van Eldijk, M. B.; Nolte, R. J. M.; Mecinovic, J., *Chem. Soc. Rev.* **2016,** *45*, 24-39.
40. Rambaran, R. N.; Serpell, L. C., *Prion* **2008,** *2*, 112-117.
41. Shoulders, M. D.; Raines, R. T., *Annu. Rev. Biochem* **2009,** *78*, 929-958.
42. Blanchoin, L.; Boujemaa-Paterski, R.; Sykes, C.; Plastino, J., *Physiol. Rev.* **2014,** *94*, 235-263.
43. Lee, C.-H.; Kim, M.-S.; Chung, B. M.; Leahy, D. J.; Coulombe, P. A., *Nat. Struct. Mol. Biol.* **2012,** *19*, 707-715.
44. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H., *Science* **1994,** *266*, 75-81.
45. Kneussel, M.; Wagner, W., *Nature Rev. Neurosci.* **2013,** *14*, 233-247.
46. Engel, A.; Gaub, H. E., *Annu. Rev. Biochem* **2008,** *77*, 127-148.
47. De Santis, E.; Ryadnov, M. G., *Chem. Soc. Rev.* **2015,** *44*, 8288-8300.
48. Shen, J.-R., *Annu. Rev. Plant Biol.* **2015,** *66*, 23-48.
49. Lenaz, G.; Genova, M. L., Respiratory Cytochrome Supercomplexes. In *Cytochrome Complexes: Evolution, Structures, Energy Transduction, and Signaling*, Cramer, W. A.; Kallas, T., Eds. Springer Netherlands: Dordrecht, 2016; pp 585-628.
50. Moreau, R. J.; Schubert, C. R.; Nasr, K. A.; Török, M.; Miller, J. S.; Kennedy, R. J.; Kemp, D. S., *J. Am. Chem. Soc.* **2009,** *131*, 13107-13116.
51. Marsh, J. A.; Teichmann, S. A., *Annu. Rev. Biochem* **2015,** *84*, 551-575.
52. Boyle, A. L.; Woolfson, D. N., Rational Design of Peptide-Based Biosupramolecular Systems. In *Supramolecular Chemistry: From Molecules to Nanomaterials*, Gale, P. A.; Steed, J. W., Eds. John Wiley and Sons: 2012; Vol. 8, pp 1639-1664.
53. Wang, J.; Liu, K.; Yan, L.; Wang, A.; Bai, S.; Yan, X., *ACS Nano* **2016,** *10*, 2138-2143.
54. Mantion, A.; Massüger, L.; Rabu, P.; Palivan, C.; McCusker, L. B.; Taubert, A., *J. Am. Chem. Soc.* **2008,** *130*, 2517-2526.
55. Salgado, E. N.; Radford, R. J.; Tezcan, F. A., *Acc. Chem. Res.* **2010,** *43*, 661-672.

56. Parry, D. A.; Fraser, R. D.; Squire, J. M., *J. Struct. Biol.* **2008,** *163*, 258-269.
57. Woolfson, D. N.; Ryadnov, M. G., *Curr. Opin. Chem. Biol.* **2006,** *10*, 559-67.
58. Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P., *Tetrahedron* **2004,** *60*, 7237-7246.
59. Ryadnov, M. G.; Woolfson, D. N., *Nat. Mater.* **2003,** *2*, 329-32.
60. Potekhin, S. A.; Melnik, T. N.; Popov, V.; Lanina, N. F.; Vazina, A. A.; Rigler, P.; Verdini, A. S.; Corradin, G.; Kajava, A. V., *Chem. Biol.* **2001,** *8*, 1025-1032.
61. Zaccai, N. R.; Chi, B.; Thomson, A. R.; Boyle, A. L.; Bartlett, G. J.; Bruning, M.; Linden, N.; Sessions, R. B.; Booth, P. J.; Brady, R. L.; Woolfson, D. N., *Nat. Chem. Biol.* **2011,** *7*, 935-41.
62. Staples, J. K.; Oshaben, K. M.; Horne, W. S., *Chem. Sci.* **2012,** *3*, 3387.
63. Zhou, M.; Bentley, D.; Ghosh, I., *J. Am. Chem. Soc.* **2004,** *126*, 734-5.
64. 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.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N., *Science* **2013,** *340*, 595-9.
65. Hartgerink, J. D.; Beniash, E.; Stupp, S. I., *Proc. Natl. Acad. Sci. USA* **2002,** *99*, 5133-5138.
66. Merg, A. D.; Boatz, J. C.; Mandal, A.; Zhao, G.; Mokashi-Punekar, S.; Liu, C.; Wang, X.; Zhang, P.; van der Wel, P. C. A.; Rosi, N. L., *J. Am. Chem. Soc.* **2016,** *138*, 13655-13663.
67. Korevaar, P. A.; Newcomb, C. J.; Meijer, E. W.; Stupp, S. I., *J. Am. Chem. Soc.* **2014,** *136*, 8540-8543.
68. Ghadiri, M. R.; Kobayashi, K.; Granja, J. R.; Chadha, R. K.; McRee, D. E., *Angew. Chem. Int. Ed.* **1995,** *34*, 93-95.
69. Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N., *Nature* **1993,** *366*, 324-327.
70. Ashkenasy, N.; Horne, W. S.; Ghadiri, M. R., *Small* **2006,** *2*, 99-102.
71. Chapman, R.; Danial, M.; Koh, M. L.; Jolliffe, K. A.; Perrier, S., *Chem. Soc. Rev.* **2012,** *41*, 6023-6041.
72. Brea, R. J.; Reiriz, C.; Granja, J. R., *Chem. Soc. Rev.* **2010,** *39*, 1448-1456.
73. Peacock, A. F. A., *Curr. Opin. Chem. Biol.* **2013,** *17*, 934-939.
74. Zou, R.; Wang, Q.; Wu, J.; Wu, J.; Schmuck, C.; Tian, H., *Chem. Soc. Rev.* **2015,** *44*, 5200-5219.
75. Radford, R. J.; Brodin, J. D.; Salgado, E. N.; Tezcan, F. A., *Coord. Chem. Rev.* **2011,** *255*, 790-803.
76. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., *Nat. Chem.* **2012,** *4*, 375-382.
77. Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A., *J. Am. Chem. Soc.* **2015,** *137*, 11598-11601.
78. Wei, P.; Yan, X.; Huang, F., *Chem. Soc. Rev.* **2015,** *44*, 815-832.
79. Przybyla, D. E.; Chmielewski, J., *J. Am. Chem. Soc.* **2008,** *130*, 12610-12611.
80. Przybyla, D. E.; Chmielewski, J., *J. Am. Chem. Soc.* **2010,** *132*, 7866-7867.
81. Przybyla, D. E.; Rubert Pérez, C. M.; Gleaton, J.; Nandwana, V.; Chmielewski, J., *J. Am. Chem. Soc.* **2013,** *135*, 3418-3422.
82. Vandermeulen, G. W. M.; Tziatzios, C.; Schubert, D.; Andres, P. R.; Alexeev, A.; Schubert, U. S.; Klok, H.-A., *Aust. J. Chem.* **2004,** *57*, 33-39.

83. Tangbunsuk, S.; Whittell, G. R.; Ryadnov, M. G.; Vandermeulen, G. W. M.; Woolfson, D. N.; Manners, I., *Chem. Eur. J.* **2012,** *18*, 2524-2535.
84. Nepal, M.; Sheedlo, M. J.; Das, C.; Chmielewski, J., *J. Am. Chem. Soc.* **2016,** *138*, 11051-11057.
85. Lupas, A. N.; Gruber, M., *Adv. Protein Chem.* **2005,** *70*, 37-78.
86. Parry, D. A.; Fraser, R. D.; Squire, J. M., *J. Struct. Biol.* **2008,** *163*, 258-69.
87. O'Shea, E. K.; Klemm, J.; Kim, P. S.; Alber, T., *Science* **1991,** *254*, 539-544.
88. Lupas, A., *Trends Biochem. Sci* **1996,** *21*, 375-382.
89. Crick, F., *Acta Crystallographica* **1953,** *6*, 689-697.
90. Oshaben, K. M.; Salari, R.; McCaslin, D. R.; Chong, L. T.; Horne, W. S., *Biochemistry* **2012,** *51*, 9581-91.
91. Su, J. Y.; Hodges, R. S.; Kay, C. M., *Biochemistry* **1994,** *33*, 15501-15510.
92. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T., *Science* **1993,** *262*, 1401-1407.
93. Wagschal, K.; Tripet, B.; Mant, C.; Hodges, R. S.; Lavigne, P., *Protein Sci.* **1999,** *8*, 2312-2329.
94. Tsang, B. P.; Bretscher, H. S.; Kokona, B.; Manning, R. S.; Fairman, R., *Biochemistry* **2011,** *50*, 8548-8558.
95. Liu, J.; Yong, W.; Deng, Y.; Kallenbach, N. R.; Lu, M., *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 16156-16161.
96. Liu, J.; Zheng, Q.; Deng, Y.; Cheng, C.-S.; Kallenbach, N. R.; Lu, M., *Proc. Natl. Acad. Sci. USA* **2006,** *103*, 15457-15462.
97. V, K.; Sharff, A.; Koronakis, E.; Luisi, B.; Hughes, C., *Nature* **2000,** *405*, 914-919.
98. Grigoryan, G.; Keating, A. E., *Curr. Opin. Struct. Biol.* **2008,** *18*, 477-483.
99. Knappenberger, J. A.; Smith, J. E.; Thorpe, S. H.; Zitzewitz, J. A.; Matthews, C. R., *J. Mol. Biol.* **2002,** *321*, 1-6.
100. Zeng, X.; Herndon, A. M.; Hu, J. C., *Proc. Natl. Acad. Sci. USA* **1997,** *94*, 3673-8.
101. Nautiyal, S.; Woolfson, D. N.; King, D. S.; Alber, T., *Biochemistry* **1995,** *34*, 11645-51.
102. Zhu, H.; Celinski, S. A.; Scholtz, J. M.; Hu, J. C., *Protein Sci.* **2001,** *10*, 24-33.
103. O'Shea, E.; Rutkowski, R.; Stafford, W.; Kim, P., *Science* **1989,** *245*, 646-648.
104. O'Shea, E. K.; Lumb, K. J.; Kim, P. S., *Curr. Biol.* **1993,** *3*, 658-657.
105. Root, B. C.; Pellegrino, L. D.; Crawford, E. D.; Kokona, B.; Fairman, R., *Protein Sci.* **2009,** *18*, 329-36.
106. Ryan, S. J.; Kennan, A. J., *J. Am. Chem. Soc.* **2007,** *129*, 10255-10260.
107. Schnarr, N. A.; Kennan, A. J., *J. Am. Chem. Soc.* **2001,** *123*, 11081-2.
108. Travisano, P.; Kennan, A. J., *Org. Lett.* **2004,** *6*, 4219-4222.
109. Dietz, H.; Bornschlögl, T.; Heym, R.; König, F.; Rief, M., *New Journal of Physics* **2007,** *9*, 424-424.
110. Grigoryan, G.; Degrado, W. F., *J. Mol. Biol.* **2011,** *405*, 1079-100.
111. Fletcher, J. M.; Boyle, A. L.; Bruning, M.; Bartlett, G. J.; Vincent, T. L.; Zaccai, N. R.; Armstrong, C. T.; Bromley, E. H. C.; Booth, P. J.; Brady, R. L.; Thomson, A. R.; Woolfson, D. N., *ACS Synth. Biol.* **2012,** *1*, 240-250.
112. Cummings, C. G.; Hamilton, A. D., *Curr. Opin. Chem. Biol.* **2010,** *14*, 341-346.
113. Davis, J. M.; Tsou, L. K.; Hamilton, A. D., *Chem. Soc. Rev.* **2007,** *36*, 326-334.
114. Jin, L.; Wang, W.; Fang, G., *Annu. Rev. Pharmacool. Toxicol.* **2014,** *54*, 435-456.
115. Tan, Y. S.; Lane, D. P.; Verma, C. S., *Drug Discov. Today* **2016,** *21*, 1642-1653.
116. Reinert, Z. E.; Horne, W. S., *Org. Biomol. Chem.* **2014,** *12*, 8796-8802.

117. Olsen, C. A., *ChemBioChem* **2010,** *11*, 152-160.
118. Henchey, L. K.; Jochim, A. L.; Arora, P. S., *Curr. Opin. Chem. Biol.* **2008,** *12*, 692-697.
119. Gellman, S. H., *Acc. Chem. Res.* **1998,** *31*, 173-180.
120. Werner, H. M.; Cabalteja, C. C.; Horne, W. S., *ChemBioChem* **2016,** *17*, 712-8.
121. Horne, W. S.; Yadav, M. K.; Stout, C. D.; Ghadiri, M. R., *J. Am. Chem. Soc.* **2004,** *126*, 15366-15367.
122. Sun, J.; Zuckermann, R. N., *ACS Nano* **2013,** *7*, 4715-4732.
123. Violette, A.; Fournel, S.; Lamour, K.; Chaloin, O.; Frisch, B.; Briand, J.-P.; Monteil, H.; Guichard, G., *Chem. Biol.* **2006,** *13*, 531-538.
124. Jiang, H.; Léger, J.-M.; Dolain, C.; Guionneau, P.; Huc, I., *Tetrahedron* **2003,** *59*, 8365-8374.
125. Dado, G. P.; Gellman, S. H., *J. Am. Chem. Soc.* **1994,** *116*, 1054-1062.
126. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H., *J. Am. Chem. Soc.* **1996,** *118*, 13071-13072.
127. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X.; Barchi, J. J.; Gellman, S. H., *Nature* **1997,** *387*, 381-384.
128. Seebach, D.; Schreiber, J. V.; Arvidsson, P. I.; Frackenpohl, J., *Helv. Chim. Acta* **2001,** *84*, 271-279.
129. Hart, S. A.; Bahadoor, A. B. F.; Matthews, E. E.; Qiu, X. J.; Schepartz, A., *J. Am. Chem. Soc.* **2003,** *125*, 4022-4023.
130. Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H., *J. Am. Chem. Soc.* **2007,** *129*, 4178-4180.
131. Reinert, Z. E.; Horne, W. S., *Chem. Sci.* **2014,** *5*, 3325-3330.
132. Szefczyk, M.; Weglarz-Tomczak, E.; Fortuna, P.; Krzyszton, A.; Rudzinska-Szostak, E.; Berlicki, L., *Angew. Chem. Int. Ed.* **2017,** *56*, 2087-2091.
133. Raap, J.; Erkelens, K.; Ogrel, A.; Skladnev, D. A.; Brückner, H., *J. Pept. Sci.* **2005,** *11*, 331-338.
134. Toniolo, C.; Bonora, G. M.; Bavoso, A.; Benedetti, E.; di Blasio, B.; Pavone, V.; Pedone, C., *Biopolymers* **1983,** *22*, 205-215.
135. Banerjee, R.; Basu, G., *ChemBioChem* **2002,** *3*, 1263-1266.
136. De Filippis, V.; De Antoni, F.; Frigo, M.; Polverino de Laureto, P.; Fontana, A., *Biochemistry* **1998,** *37*, 1686-1696.
137. Maffucci, I.; Pellegrino, S.; Clayden, J.; Contini, A., *J. Phys. Chem. B* **2014**.
138. Tavenor, N. A.; Reinert, Z. E.; Lengyel, G. A.; Griffith, B. D.; Horne, W. S., *Chem. Commun.* **2016,** *52*, 3789-3792.
139. Tangbunsuk, S.; Whittell, G. R.; Ryadnov, M. G.; Vandermeulen, G. W. M.; Woolfson, D. N.; Manners, I., *Chem. Eur. J.* **2012,** *18*, 2524-2535.
140. Whittell, G. R.; Hager, M. D.; Schubert, U. S.; Manners, I., *Nat. Mater.* **2011,** *10*, 176-88.
141. Stock, N.; Biswas, S., *Chem. Rev.* **2012,** *112*, 933-969.
142. Reinert, Z. E.; Lengyel, G. A.; Horne, W. S., *J. Am. Chem. Soc.* **2013,** *135*, 12528-12531.
143. Möhle, K.; Günther, R.; Thormann, M.; Sewald, N.; Hofmann, H.-J., *Biopolymers* **1999,** *50*, 167-184.
144. Crisma, M.; Toniolo, C., *Pept.Sci.* **2015,** *104*, 46-64.
145. Jeschke, G., *ChemPhysChem* **2002,** *3*, 927-932.
146. Jeschke, G., *Annu. Rev. Phys. Chem.* **2012,** *63*, 419-46.
147. Milov, A. D.; Ponomarev, A. B.; Tsvetkov, Y. D., *Chem. Phys. Lett.* **1984,** *110*, 67-72.

148. Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W., *J. Magn. Reson.* **2000,** *142*, 331-340.
149. Schmidt, T.; Wälti, M. A.; Baber, J. L.; Hustedt, E. J.; Clore, G. M., *Angew. Chem. Int. Ed.* **2016,** *55*, 15905-15909.
150. Martin, L.; Ivancich, A.; Vita, C.; Formaggio, F.; Toniolo, C., *J. Pept. Res.* **2001,** *58*, 424-32.
151. Cunningham, T. F.; McGoff, M. S.; Sengupta, I.; Jaroniec, C. P.; Horne, W. S.; Saxena, S., *Biochemistry* **2012,** *51*, 6350-6359.
152. Toniolo, C.; Valente, E.; Formaggio, F.; Crisma, M.; Pilloni, G.; Corvaja, C.; Toffoletti, A.; Martinez, G. V.; Hanson, M. P.; Millhauser, G. L.; et al., *J. Pept. Sci.* **1995,** *1*, 45-57.
153. Schreier, S.; Bozelli, J. C., Jr.; Marin, N.; Vieira, R. F.; Nakaie, C. R., *Biophys. Rev.* **2012,** *4*, 45-66.
154. Schiemann, O.; Prisner, T. F., *Q. Rev. Biophys.* **2007,** *40*, 1-53.
155. Marchetto, R.; Schreier, S.; Nakaie, C. R., *J. Am. Chem. Soc.* **1993,** *115*, 11042-11043.
156. Kaduk, C.; Wenschuh, H.; Beyermann, M.; Forner, K.; Carpino, L. A.; Bienert, M., *Lett. Pept. Sci.* **1996,** *2*, 285-288.
157. Karim, C. B.; Zhang, Z.; Thomas, D. D., *Nat. Protoc.* **2007,** *2*, 42-9.
158. Tavenor, N. A.; Silva, K. I.; Saxena, S.; Horne, W. S., *J. Phys. Chem. B* **2014,** *118*, 9881-9889.
159. Saviano, M.; Improta, R.; Benedetti, E.; Carrozzini, B.; Cascarano, G. L.; Didierjean, C.; Toniolo, C.; Crisma, M., *ChemBioChem* **2004,** *5*, 541-544.
160. Smythe, M. L.; Nakaie, C. R.; Marshall, G. R., *J. Am. Chem. Soc.* **1995,** *117*, 10555-10562.
161. Oshaben, K. M.; Horne, W. S., *Biomacromolecules* **2014,** *15*, 1436-1442.
162. Giorgino, T., *Comput. Phys. Commun.* **2014,** *185*, 1109-1114.
163. Marinelli, F.; Faraldo-Gómez, José D., *Biophys. J.* **2015,** *108*, 2779-2782.
164. Barducci, A.; Bonomi, M.; Parrinello, M., *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011,** *1*, 826-843.
165. Lescrinier, T.; Hendrix, C.; Kerremans, L.; Rozenski, J.; Link, A.; Samyn, P.; Van Aerschot, A.; Lescrinier, E.; Eritja, R.; Van Beeumen, J.; Herdewijn, P., *Chem. Eur. J.* **1998,** *4*, 425-433.
166. Kaduk, C.; Wenschuh, H.; Beyermann, M.; Forner, K.; Carpino, L.; Bienert, M., *Lett Pept Sci* **1996,** *2*, 285-288.
167. Gill, S. C.; von Hippel, P. H., *Anal. Biochem.* **1989,** *182*, 319-326.
168. Ma, M.; Bong, D., *Org. Biomol. Chem.* **2011,** *9*, 7296-7299.
169. Shortle, D.; Meeker, A. K.; Freire, E., *Biochemistry* **1988,** *27*, 4761-4768.
170. Bunkoczi, G.; Echols, N.; McCoy, A. J.; Oeffner, R. D.; Adams, P. D.; Read, R. J., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2013,** *69*, 2276-2286.
171. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2010,** *66*, 213-221.
172. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2010,** *66*, 486-501.
173. Martin, R. E.; Pannier, M.; Diederich, F.; Gramlich, V.; Hubrich, M.; Spiess, H. W., *Angew. Chem. Int. Ed.* **1998,** *37*, 2833-2837.

174. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H., *Appl. Magn. Reson.* **2006,** *30*, 473-498.
175. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., *Bioinformatics* **2013,** *29*, 845-54.
176. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E., *Proteins: Struct., Funct., Bioinf.* **2010,** *78*, 1950-1958.
177. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., *J. Mol. Graphics Modell.* **2006,** *25*, 247-260.
178. Sousa da Silva, A.; Vranken, W., *BMC Res. Notes* **2012,** *5*, 367.
179. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *J. Chem. Phys.* **1983,** *79*, 926-935.
180. Seeliger, D.; de Groot, B. L., *Proteins: Struct., Funct., Bioinf.* **2007,** *68*, 595-601.
181. Checco, J. W.; Lee, E. F.; Evangelista, M.; Sleebs, N. J.; Rogers, K.; Pettikiriarachchi, A.; Kershaw, N. J.; Eddinger, G. A.; Belair, D. G.; Wilson, J. L.; Eller, C. H.; Raines, R. T.; Murphy, W. L.; Smith, B. J.; Gellman, S. H.; Fairlie, W. D., *J. Am. Chem. Soc.* **2015,** *137*, 11365-11375.
182. Werner, H. M.; Horne, W. S., *Curr. Opin. Chem. Biol.* **2015,** *28*, 75-82.
183. Johnson, L. M.; Gellman, S. H., *Methods Enzymol.* **2013,** *523*, 407.
184. Cheloha, R. W.; Maeda, A.; Dean, T.; Gardella, T. J.; Gellman, S. H., *Nat. Biotechnol.* **2014,** *32*, 653.
185. Price, J. L.; Horne, W. S.; Gellman, S. H., *J. Am. Chem. Soc.* **2010,** *132*, 12378-12387.
186. Spencer, R. K.; Li, H.; Nowick, J. S., *J. Am. Chem. Soc.* **2014**.
187. Karnes, M. A.; Schettler, S. L.; Werner, H. M.; Kurz, A. F.; Horne, W. S.; Lengyel, G. A., *Org. Lett.* **2016,** *18*, 3902-3905.
188. Lengyel, G. A.; Reinert, Z. E.; Griffith, B. D.; Horne, W. S., *Org. Biomol. Chem.* **2014,** *12*, 5375-5381.
189. Lengyel, G. A.; Frank, R. C.; Horne, W. S., *J. Am. Chem. Soc.* **2011,** *133*, 4246-4249.
190. Kuhlman, B.; Raleigh, D. P., *Protein Sci.* **1998,** *7*, 2405-2412.
191. Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N., *Nature Chemical Biology* **2010,** *6*, 615-620.
192. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T., *J. Am. Chem. Soc.* **2013,** *135*, 7843-6.
193. Frericks Schmidt, H. L.; Sperling, L. J.; Gao, Y. G.; Wylie, B. J.; Boettcher, J. M.; Wilson, S. R.; Rienstra, C. M., *J. Phys. Chem. B* **2007,** *111*, 14362-14369.
194. Royo, S.; De Borggraeve, W. M.; Peggion, C.; Formaggio, F.; Crisma, M.; Jiménez, A. I.; Cativiela, C.; Toniolo, C., *J. Am. Chem. Soc.* **2005,** *127*, 2036-2037.
195. Maffucci, I.; Pellegrino, S.; Clayden, J.; Contini, A., *J. Phys. Chem. B* **2015,** *119*, 1350-1361.
196. Burton, A. S.; Stern, J. C.; Elsila, J. E.; Glavin, D. P.; Dworkin, J. P., *Chem. Soc. Rev.* **2012,** *41*, 5459-5472.
197. Bada, J. L., *Chem. Soc. Rev.* **2013,** *42*, 2186-2196.
198. George, K. L.; Horne, W. S., *J. Am. Chem. Soc.* **2017**.
199. Chi, Y.; English, E. P.; Pomerantz, W. C.; Horne, W. S.; Joyce, L. A.; Alexander, L. R.; Fleming, W. S.; Hopkins, E. A.; Gellman, S. H., *J. Am. Chem. Soc.* **2007,** *129*, 6050-6055.
200. Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S., *Org. Lett.* **2014,** *16*, 940-943.

201. Alexander, P.; Fahnestock, S.; Lee, T.; Orban, J.; Bryan, P., *Biochemistry* **1992,** *31*, 3597-3603.
202. Buer, B. C.; Levin, B. J.; Marsh, E. N. G., *J. Am. Chem. Soc.* **2012,** *134*, 13027-13034.
203. Cui, Y.; Li, B.; He, H.; Zhou, W.; Chen, B.; Qian, G., *Acc. Chem. Res.* **2016,** *49*, 483-493.
204. Guillerm, V.; Kim, D.; Eubank, J. F.; Luebke, R.; Liu, X.; Adil, K.; Lah, M. S.; Eddaoudi, M., *Chem. Soc. Rev.* **2014,** *43*, 6141-6172.
205. Bazzicalupi, C.; Bencini, A.; Bianchi, A.; Danesi, A.; Faggi, E.; Giorgi, C.; Santarelli, S.; Valtancoli, B., *Coord. Chem. Rev.* **2008,** *252*, 1052-1068.
206. Schubert, U. S.; Hofmeier, H.; Newkome, G. R., *Modern Terpyridine Chemistry*. 2006.
207. Hofmeier, H.; Schubert, U. S., *Chem. Soc. Rev.* **2004,** *33*, 373-399.
208. Stublla, A.; Potvin, P. G., *Eur. J. Inorg. Chem.* **2010,** *2010*, 3040-3050.
209. Holyer, R. H.; Hubbard, C. D.; Kettle, S. F. A.; Wilkins, R. G., *Inorg. Chem.* **1966,** *5*, 622-625.
210. Calì, R.; Rizzarelli, E.; Sammartano, S.; Siracusa, G., *Transition Met. Chem.* **1979,** *4*, 328-332.
211. van der Gucht, J.; Besseling, N. A. M.; van Leeuwen, H. P., *J. Phys. Chem. B* **2004,** *108*, 2531-2539.
212. Sigel, H., *Angew. Chem. Int. Ed.* **1975,** *14*, 394-402.
213. Wang, P.; Moorefield, C. N.; Panzer, M.; Newkome, G. R., *Chem. Commun.* **2005**, 4405-4407.
214. Zhou, W.; Wang, X.; Hu, M.; Guo, Z., *J. Inorg. Biochem.* **2013,** *121*, 114-120.
215. Zaccai, N. R.; Chi, B.; Thomson, A. R.; Boyle, A. L.; Bartlett, G. J.; Bruning, M.; Linden, N.; Sessions, R. B.; Booth, P. J.; Brady, R. L.; Woolfson, D. N., *Nat. Chem. Biol.* **2011,** *7*, 935-941.
216. Savelieff, M. G.; Wilson, T. D.; Elias, Y.; Nilges, M. J.; Garner, D. K.; Lu, Y., *Proc. Natl. Acad. Sci. USA* **2008,** *105*, 7919-7924.
217. Bonomo, R.; Riggi, F., *Transition Met. Chem.* **1984,** *9*, 308-311.
218. Peisach, J.; Blumberg, W. E., *Arch. Biochem. Biophys.* **1974,** *165*, 691-708.
219. Aronoff-Spencer, E.; Burns, C. S.; Avdievich, N. I.; Gerfen, G. J.; Peisach, J.; Antholine, W. E.; Ball, H. L.; Cohen, F. E.; Prusiner, S. B.; Millhauser, G. L., *Biochemistry* **2000,** *39*, 13760-13771.
220. Dikanov, S. A.; Samoilova, R. I.; Kolling, D. R. J.; Holland, J. T.; Crofts, A. R., *J. Biol. Chem.* **2004,** *279*, 15814-15823.
221. Koppel, D. E., *J. Chem. Phys.* **1972,** *57*, 4814-4820.
222. Baalousha, M.; Lead, J. R., *Environ. Sci. Technol.* **2012,** *46*, 6134-6142.
223. Zhu, X.; Shen, J.; Thomas, J. C., *Appl. Opt.* **2012,** *51*, 7537-7548.
224. Wang, Y. J.; Shen, J.; Liu, W.; Sun, X. M.; Dou, Z. H., *Laser Phys.* **2013,** *23*, 085701.
225. Wang, C.; Hao, X.-Q.; Wang, M.; Guo, C.; Xu, B.; Tan, E. N.; Zhang, Y.-Y.; Yu, Y.; Li, Z.-Y.; Yang, H.-B.; Song, M.-P.; Li, X., *Chem. Sci.* **2014**.
226. Zhang, G.; Liu, E.; Yang, C.; Li, L.; Golen, J. A.; Rheingold, A. L., *Eur. J. Inorg. Chem.* **2015,** *2015*, 939-947.
227. Elgrishi, N.; Chambers, M. B.; Artero, V.; Fontecave, M., *PCCP* **2014,** *16*, 13635-13644.
228. Henderson, I. M.; Hayward, R. C., *Polym. Chem.* **2012,** *3*, 1221-1230.
229. Wild, A.; Winter, A.; Schlutter, F.; Schubert, U. S., *Chem. Soc. Rev.* **2011,** *40*, 1459-1511.
230. Kim, K.-Y.; Nancollas, G. H., *J. Phys. Chem.* **1977,** *81*, 948-952.

231. Hayami, S.; Nakaya, M.; Ohmagari, H.; Alao, A. S.; Nakamura, M.; Ohtani, R.; Yamaguchi, R.; Kuroda-Sowa, T.; Clegg, J. K., *Dalton Trans.* **2015,** *44*, 9345-9348.
232. Spindler, P. E.; Glaser, S. J.; Skinner, T. E.; Prisner, T. F., *Angew. Chem. Int. Ed.* **2013,** *52*, 3425-3429.
233. Burke, C. J.; Sanyal, G.; Bruner, M. W.; Ryan, J. A.; LaFemina, R. L.; Robbins, H. L.; Zeft, A. S.; Middaugh, C. R.; Cordingley, M. G., *J. Biol. Chem.* **1992,** *267*, 9639-9644.
234. Chu, W.; Tu, Z.; McElveen, E.; Xu, J.; Taylor, M.; Luedtke, R. R.; Mach, R. H., *Biorg. Med. Chem.* **2005,** *13*, 77-87.
235. Hofmeier, H.; Andres, P. R.; Hoogenboom, R.; Herdtweck, E.; Schubert, U. S., *Aust. J. Chem.* **2004,** *57*, 419-426.
236. Grieco, P.; Gitu, P. M.; Hruby, V. J., *J. Pept. Res.* **2001,** *57*, 250-256.
237. Leelasvatanakij, L.; Aldrich, J. V., *J. Pept. Res.* **2000,** *56*, 80-87.
238. Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2004,** *60*, 2184-2195.
239. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2011,** *67*, 235-242.
240. Arriortua, M. I.; Rojo, T.; Amigo, J. M.; Germain, G.; Declercq, J. P., *Acta Crystallogr., Sect. B: Struct. Sci* **1982,** *38*, 1323-1324.
241. Mei, C.; Lei, Q.; Zhang, P., *Acta Crystallogr., Sect E* **2010,** *66*, m346-m347.
242. Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiees, H. W., *J. Magn. Reson.* **2000,** *142*, 331-340.
243. Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J., *Appl. Magn. Reson.* **2006,** *30*, 473-498.