Distribution Agreement

Signature:

Chunfu Xu

Date

Programmed Self-assembly of Coiled-coil Peptides

By

Chunfu Xu
Doctor of Philosophy

Chemistry

Dr. Vincent P. Conticello
Advisor

Dr. Brian Dyer
Committee Member

Dr. Justin Gallivan
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Programmed Self-assembly of Coiled-coil Peptides

By

Chunfu Xu

B.S., Fudan University, 2008

Advisor: Vincent P. Conticello, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2013

Abstract

Programmed Self-assembly of Coiled-coil Peptides

By Chunfu Xu

Self-assembly is an extension of the central dogma of molecular biology, bridging the realm of linear information and the realm of protein assemblies. However, due to our limited understanding of principles of chemical self-assembly, the level of complexity of synthetic self-assembling systems pales in comparison to what nature exhibits all around us. With their relatively straightforward intermolecular interfaces, coiled coils could be a model system for studying rules of protein self-assembly and rationally designing complex peptide supramolecular assemblies. Surrounding the ubiquitous structural motif, coiled coil, different research topics will be presented including controlling peptide self-assembly through metal-induced registry shift, recoding heptameric coiled coil for nanotube formation, and designing large-diameter helical nanotubes. During these studies, various biophysical measurements conducted in solution and the solid state over multiple length scales of structural hierarchy were employed to verify our hypotheses. These studies provided design strategies for dynamically reconfigurable nanoscale materials and nano-porous protein-based materials, and investigated the primary sequence determinants which underlie the self-assembly processes of coiled coils. The results of these studies described in this volume are of both theoretical and practical significance in understanding primary rules of peptide/protein self-assembly and rational design of functional protein-based materials.

Programmed Self-assembly of Coiled-coil Peptides

By

Chunfu Xu

B.S., Fudan University, 2008

Advisor: Vincent P. Conticello, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2013

Acknowledgements

I would like to express my deepest gratitude to my advisor Prof. Vincent Conticello for his support and guidance for the past five years. He is a role model to me in many aspect both in academic research and in everyday life. His enthusiasm for research and high standard towards experimental data impressed me very much and will continue to have great impact on my future research. He was, is and will always be my mentor. Many thanks also to my committee members Prof. Brian Dyer and Prof. Justin Gallivan for sharing their knowledge through their classes and constructive criticisms through the course of my research. I would also like to thank the members of the Conticello lab, Dr. Melissa Patterson, Dr. Weilin Peng, Yunyun Pei, Dr. I-Lin Wu, Paolo Anzini, Tao Jiang, Elizabeth Magnotti, Becky Bartlett, Charlie Modlin, Spencer Hughes for making out lab like a big family.

I am very grateful to all members of Robert P. Apkarian Integrated Electron Microscopy Core, Prof. Elizabeth Wright, Ms. Hong Yi, Ms. Jeannette Taylor, and Mr. Art McCanna. I learned a lot about electron microscopy over the past four years which has been really helpful for my own research. I also want to thank Dr. Ricardo Guerrero-Ferreira and Dr. Gabriella Kiss for their help in my research and their friendship.

Special thanks also to the Lynn group for their advice and comradery in our shared areas of interest, especially Dr. Anil Mehta for his advice and his great work on solid-state NMR and computational modeling.

I would like to show my appreciation to our great collaborators, Prof. Borries Demeler (The University of Texas Health Science Center), Prof. Lars Hemmingsen

(University of Copenhagen, Denmark), Prof. Elizabeth Wright (Emory University), Dr. Anil Mehta (Emory University), Dr. Xiaobing Zuo (Argonne National Lab), Dr. Joseph Wall (Brookhaven National Lab), Prof. Dr. Louise Serpell (University of Sussex, UK), Dr. Stanislaw Dunin-Horkawicz (The International Institute of Molecular and Cell Biology, Germany), Prof. Ed Egelman (University of Virginia), Prof. Frank DiMaio (University of Washington at Seattle), and Prof. David Baker (University of Washington at Seattle). Without their great help and insights, our research projects would not have gone as well.

I am indebted to Prof. Vince Conticello, Prof. Elizabeth Wright, Prof. Ed Egelman, and Dr. Xiaobing Zuo for their recommendation in my process of searching for a postdoc position. Their support means a great deal to my future career. I am deeply grateful.

Finally, I would like to acknowledge my family. I always felt regretful for not being able to spend more time with my parents in the past several years. No words could capture my appreciation for them. I miss them. Similarly, my deep gratitude goes to my sister who has been taking care of our parents while I was away and always supported me. At last, I want to thank Emory for bringing my fiancé Sylvie to me. That is the most beautiful thing that happened to me in my life. We shared happiness, encouraged each other, traveled together, and will continue our life journey together forever.

Table of Contents

List of Figures

List of Table

Chapter I: Introduction

1.1 The importance of chemical self-assembly

"How far can we push chemical self-assembly" has been identified as one of the top 25 big questions facing science over the next quarter-century [1] together with other compelling scientific mysteries, e.g. " what is the universe made of" and "how and where did life on earth arise". The importance of chemical self-assembly has gradually been recognized, which has led to its rapid development over the past decades or so. Scientists study the rules of chemical self-assembly in order to fabricate functional nanometer length-scale structures, to understand and cure diseases, and to tracing the origins of life.

Chemical self-assembly is the science of molecules that put themselves together into ordered structures driven by energy minimization in a closed system. Self-assembly as a mechanism is very powerful and operates far beyond atomic dimensions to create structures of immense size and complexity exemplified by ribosomes and viruses (Figure 1.1).

Ribosomes are one of the wonders of the cellular world. 2009 Nobel Prize in Chemistry was awarded to three structural biologists Venkatraman Ramakrishnan, Thomas A. Steitz, and Ada E. Yonath, for their studies in structure and function of ribosomes. Ribosomes are composed of two subunits: a large subunit (PDB ID 1ffk), shown on the right, and a small subunit (PDB ID 1fka), shown on the left [2]. Both subunits are composed of long strands of RNA (orange and yellow) dotted with protein chains (blue). When synthesizing a new protein, the two subunits lock together with a messenger RNA trapped in the space between.

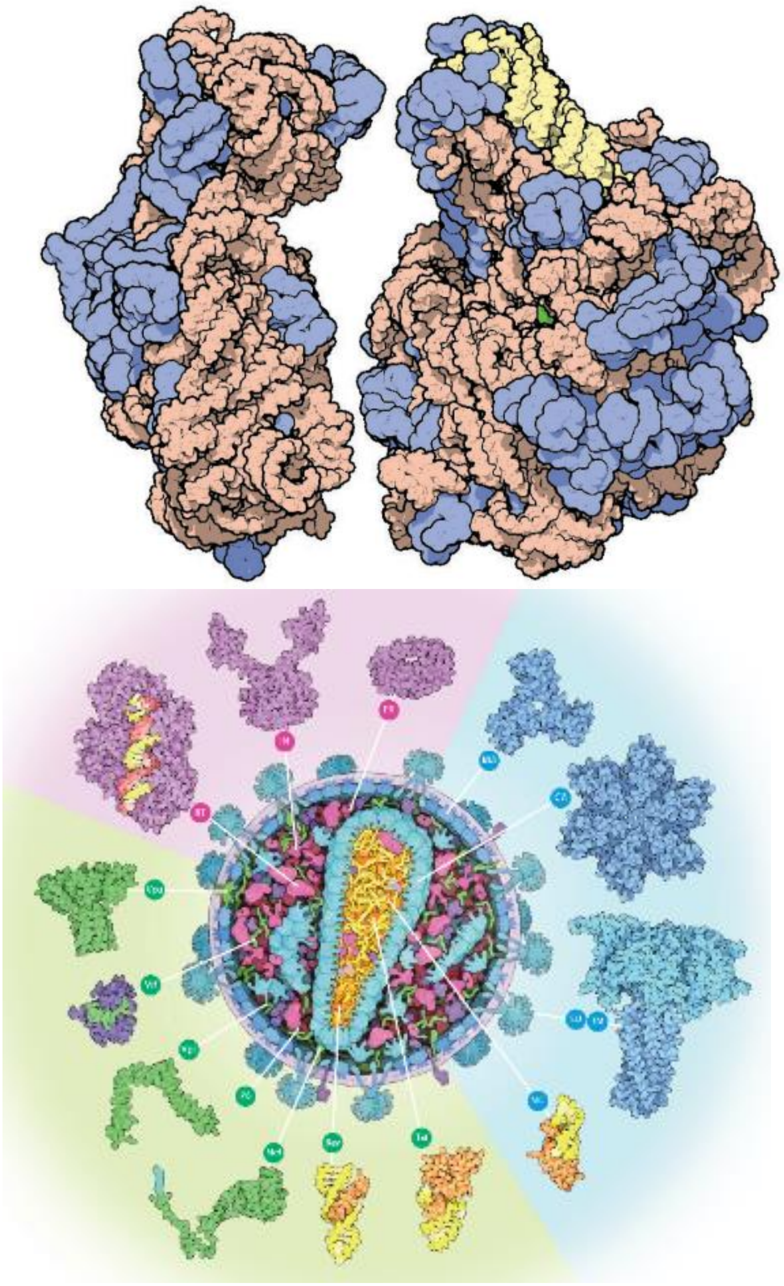


Figure 1.1. Ribosomes (top) and viruses (bottom) are masterpieces of molecular self-assembly. Ribosomes are composed of two subunits. Both consist of long strands of RNA (orange and yellow) dotted with protein chains (blue) [2]. HIV is composed of two strands of RNA, 15 types of viral proteins, and a few proteins from the last host cell it infected, all surrounded by a lipid bilayer membrane [3].

The ribosome then walks down the messenger RNA three nucleotides at a time, building a new protein piece-by-piece. The structural complexity and functional accuracy and efficiency of ribosomes exhibits how amazing macromolecular folding and self-assembly are. Similarly, viruses are another masterpiece of self-assembly. A virus is a single highly ordered structure that is "uniquely determined by size, number of component, geometry, and strength of interaction" [4]. Taking HIV (human immunodeficiency virus) as an example, an HIV virion is composed of two strands of RNA, 15 types of viral proteins, and a few proteins from the last host cell it infected, all surrounded by a lipid bilayer membrane [3]. Together, these molecules allow the virus to infect cells of the immune system and force them to build new copies of the virus. Each molecule in the virus plays a role in this process. They cooperate with each other from the first steps of viral attachment to the final process of budding.

1.1.1 Self-assembly creates the complex hierarchy of life.

The basis to understand supramolecular structure is the doctrine of self-assembly, an important segment in biological information flow (Figure 1.2). Self-assembly is an extension of the central dogma of molecular biology, bridging the realm of linear information and the realm of protein assemblies [5]. It helps us understand how each gene product functionally interact with other gene products, and forms the basis for us to comprehend higher level organization and information transfer in biological systems [6]. Self-assembly creates the complex hierarchy of life. The illustration below shows (Figure 1.3) the beautiful work done by chemical self-assembly.

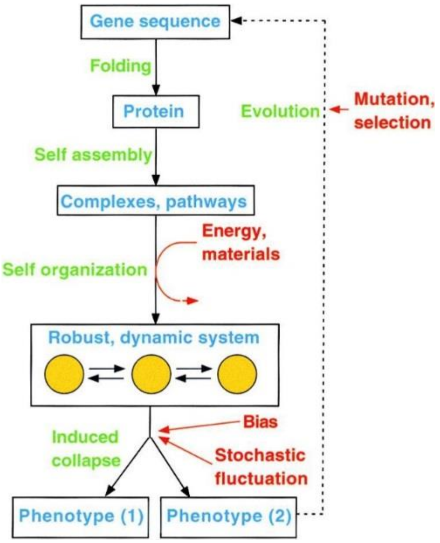


Figure 1.2. Information flow in biological organization. [5] The diagram depicted the information flow from the genotype to the phenotype. In this process, self-assembly is an extension of the central dogma of molecular biology.

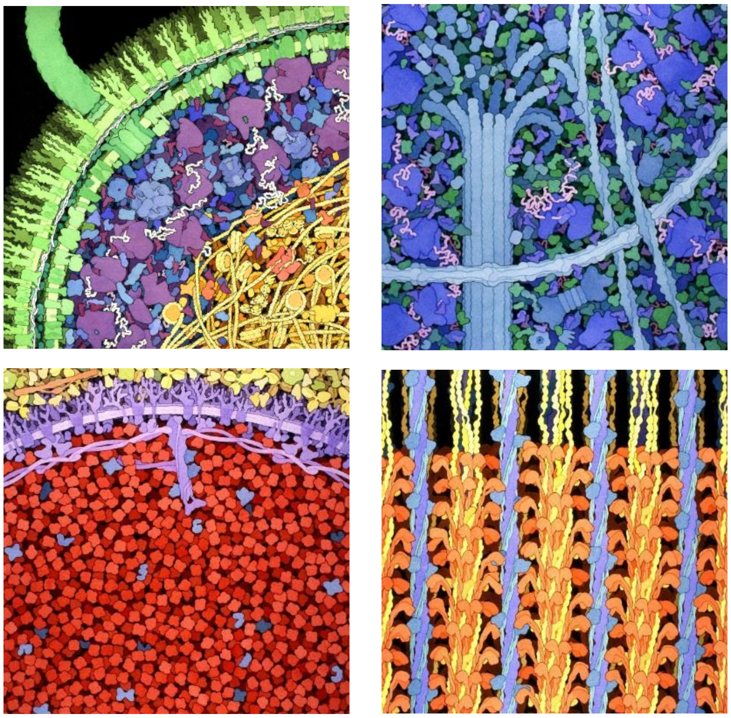


Figure 1.3. Beautiful artwork by chemical self-assembly. Illustration by David S. Goodsell, the Scripps Research Institute.[7] Top Left: A cross-section of a small

portion of an Escherichia coli cell; Top Right: A small portion of cytoplasm including three types of filaments that make up the cytoskeleton; Bottom Left: A portion of a red blood cell with hemoglobin in red; Bottom Right: Part of a muscle sarcomere with actin filaments in blue and myosin filaments in red.

1.1.2 Self-assembly is also the only practical approach for making a variety of nanostructures.

For the past several decades, scientists have made key strides in learning the fundamental rules of non-covalent bonding and have utilized them to design self-assembling systems with a modest degree of complexity. The most extraordinary achievements were accomplished in the field of DNA nanotechnology. Taking advantage of the remarkable specificity of the interactions between complementary nucleotides and the programmability of the nucleotide sequences, scientists have been able to build complex three-dimensional shapes and functional nanomachines [8-12]. Meanwhile, a great deal of efforts has been made by scientists toward designing complex peptide/protein supramolecular assemblies. Over the past decade or so, a certain degree of success has been demonstrated towards the construction of spherical[13-15], filamentous [16-19], tubular [20-23] and planar (our unpublished data) peptide/protein nanoarchitectures. More intriguingly, self-assembling protein cages [24] have been computationally designed with atomic level accuracy recently. The involvement of advanced computational modeling could potentially generate structurally defined protein nanomaterials that rival the complexity of DNA nanostructures. Besides, the folding and assembling of other sequence-specific polymers have also experienced rapid developments in recent years which was exemplified by an ultrathin two dimensional crystals formed by peptoid polymers [25].

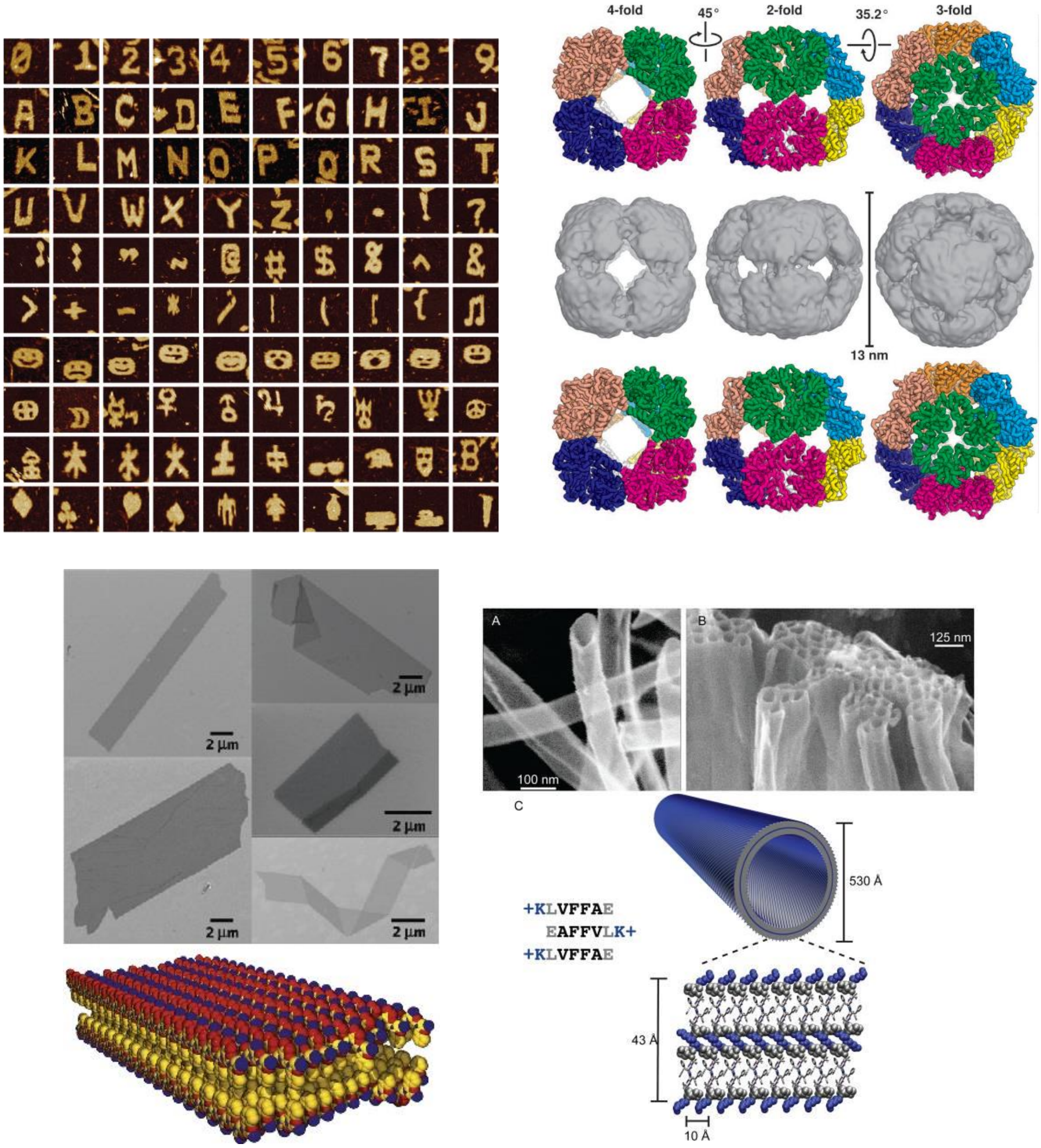


Figure 1.4. Various self-assembling nanostructures. Top Left: Complex shapes from single-stranded DNA tiles [9]; Top Right: Computational design of protein cages with atomic level accuracy [24]; Bottom Left: Ultrathin two-dimensional peptoid nanosheet [25]; Bottom Right [22]: Peptide nanotubes self-assembled from short amyloid sequence.

Impressive as it is, this level of complexity pales in comparison to what nature flaunts all around us. The fundamental rules of chemical self-assembly are yet to be appreciated and learned in order to replicate what nature can do, at least, in part, and to understand how biological information is transferred from genotype to phenotype and how life on the earth evolves.

1.2 Introduction to coiled coils

Inspired by structural DNA nanotechnology where the simple base pairing rules lead to people's proficiency in utilizing DNA "tiles" to construct sophisticated assemblies, coiled coils have been of great interests for protein engineering and have been considered as a model system for studying protein self-assembly due to their relatively straightforward sequence characteristics and well-understood sequence-structure relationships.

So what are coiled coils? Coiled coils are bundles of α-helices that are interwound into superhelical structures. Most commonly, they consist of two, three, or four helices, running in the same (parallel) or in opposite (antiparallel) directions, but structures with five and more helices have been determined. Coiled coils are ubiquitous and versatile structural motifs in biology involved in important biological functions mainly because coiled coils are ideal mediators of oligomerization. They assume this role in a multitude of proteins such as transcription factors (leucine zippers), molecular motors (myosin, kinesin), receptors (macrophage scavenger receptor, chemotaxis receptors), and signaling molecules (G protein βγ) [26]. Shown below in Figure 1.5 are some examples of coiled coil domains in native proteins.

Coiled coils have characteristic heptad sequence repeats of hydrophobic (*H*) and polar (*P*) residues, $(HPPHPPP)_n$, designated as $(abcdefg)_n$ (Figure 1.6). The driving force for the folding and assembling of two or more amphipathic α-helices into oligomers is the burial of hydrophobic residues [26, 27]. In undistorted α-helices, there are 3.6 residues per turn, while the periodicity of hydrophobic residues in a canonical coiled coil is 3.5

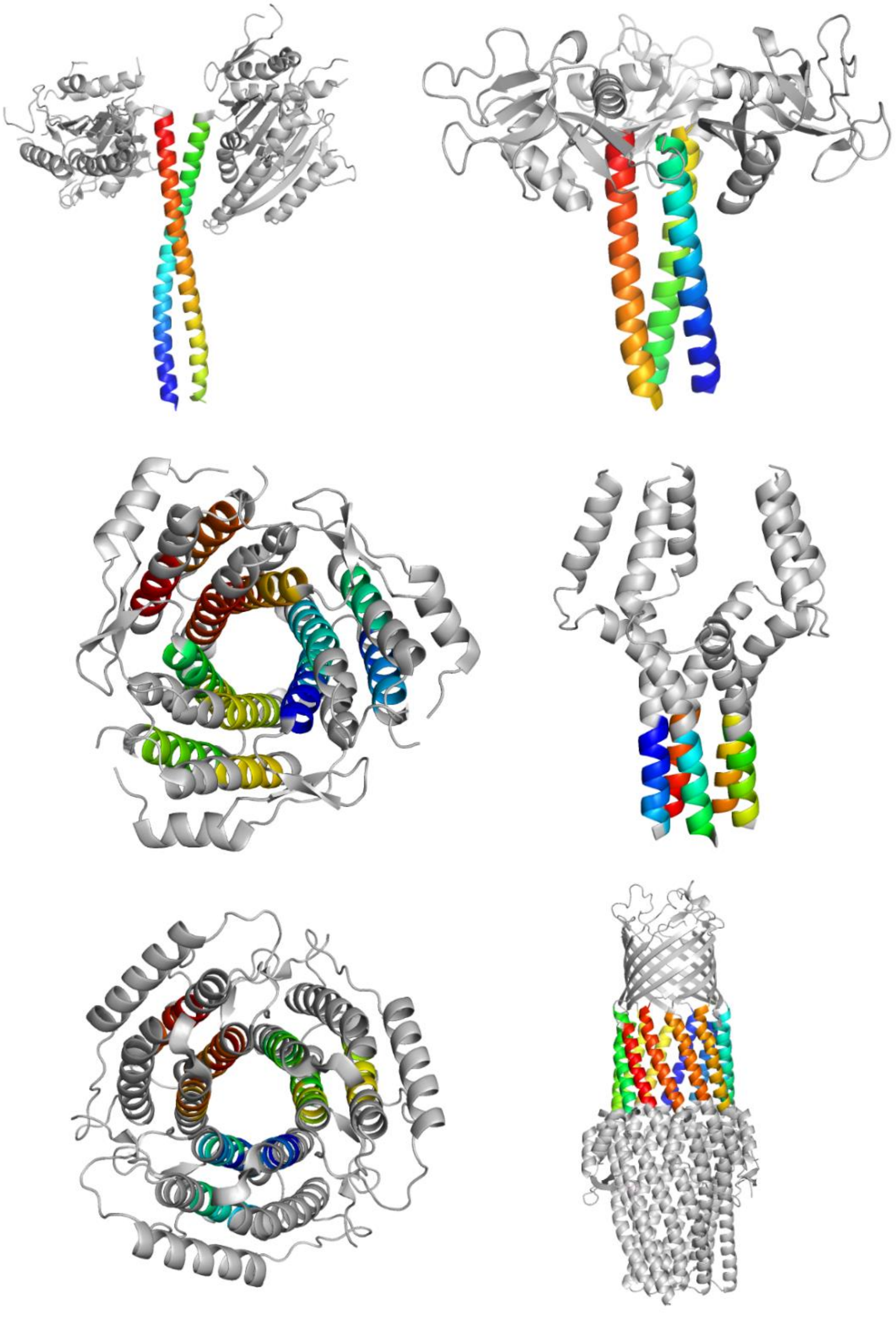


Figure 1.5. Examples of coiled-coil domains (colored segments) in native proteins with different oligomeric states. PDB IDs: 1N6M, 1RTM (top row); 1NIG, 2HYN (middle row); 2NT8, 1EK9 (bottom row).

(7/2, two hydrophobic residues in a heptad sequence repeat). So the hydrophobic seams winds around the axes of α-helices in a left-handed pattern. In order to maximize burial of hydrophobic side chains, helices in coiled coils need to wrap each other to form a left-handed superhelix (Figure 1.6). The interactions of hydrophobic side chains in the core of the bundle is called Knobs-into-Holes (KIH) packing proposed in 1950s by Francis Crick [28, 29] which is the hallmark of coiled coils. In a KIH packing, a side chain referred as a knob from one helix packs into a diamond shaped hole formed by four side chains from the partnering helix (Figure 1.7).

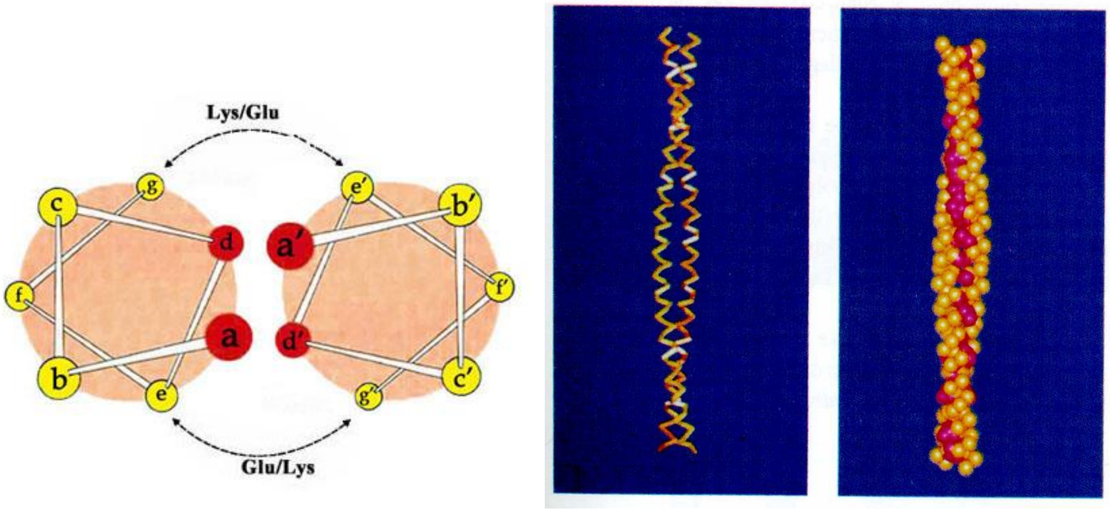


Figure 1.6. Left: Helical wheel representation of heptad sequence repeats of coiled coils. Right: Backbone model and space-filling model of a two-stranded coiled coil from a portion of sequence of tropomyosin. [27] a and d heptad positions are typically occupied by hydrophobic residues, while e and g positions are often charged residues. Burial of hydrophobic seams causes the superhelical twist.

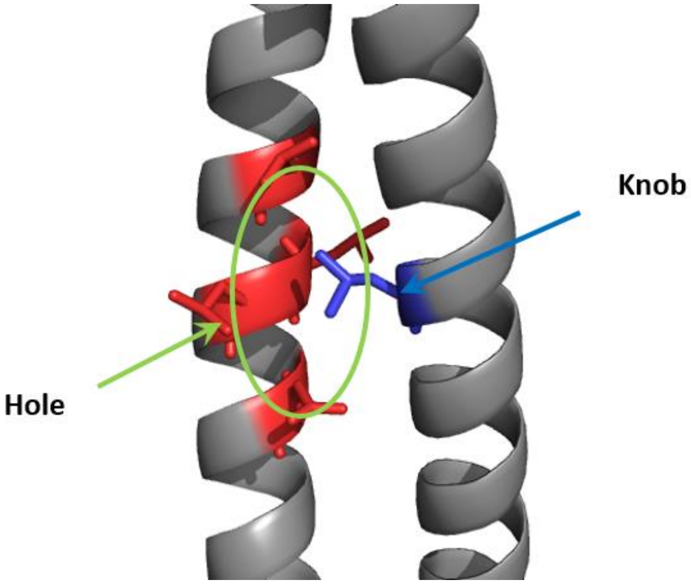


Figure 1.7. The hallmark Knobs-into-Holes parking of coiled coils (PDB ID: 2ZTA). A side chain referred as knob (blue) from one helix packs into a diamond shaped hole formed by four side chains (red) from the partnering helix.

1.3 Extended KIH packing and higher-order coiled coils

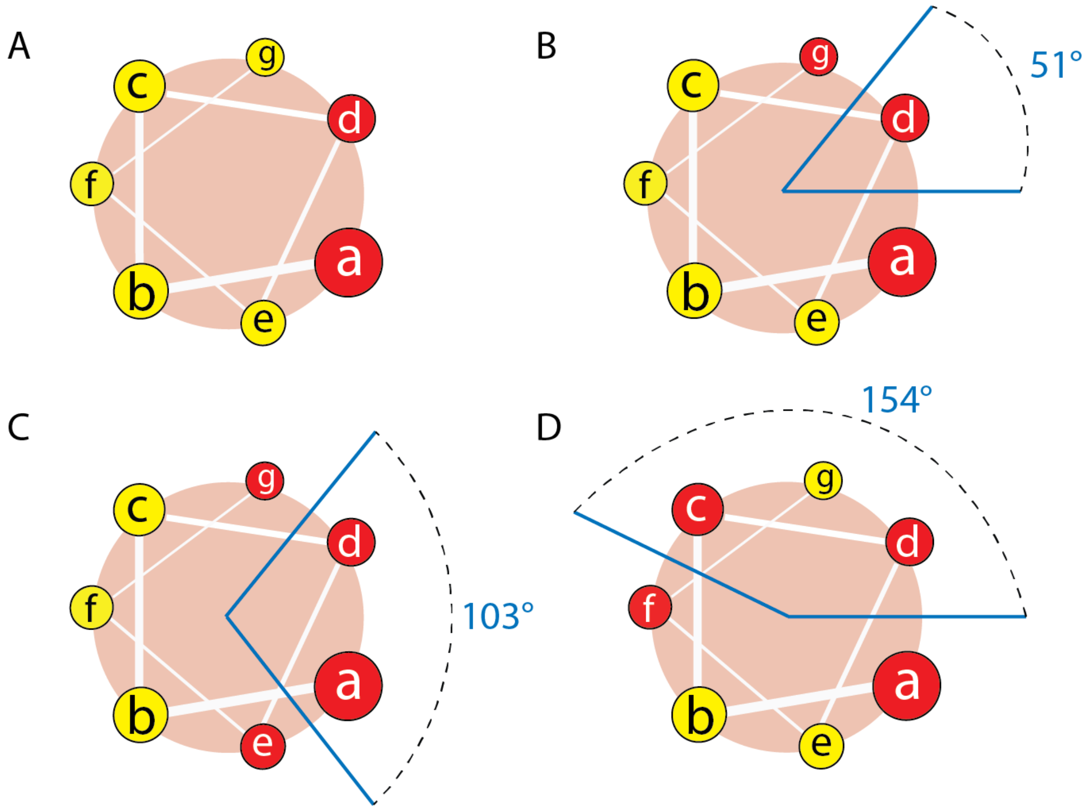


Figure 1.8. The hydrophobic interfaces are extended along with the increase of oligomeric state of coiled coils. A. Traditional heptad repeat: *HxxHxxx*. Only a and d positions are hydrophobic. B. Helical repeat: *HxxHxxH*, This pattern is seen in trimers and tetramers where side chains at g position are progressively buried. C. Helical repeat: *HxxHHxH*, seen in pentamers-heptamers. The two hydrophobic interfaces, ea and dg, are separated by 103°. D. Helical repeat: *HxHHxHx*, anticipated above heptamers. The ad and cf interfaces are separated by 154°.

In a canonical coiled-coil dimer, only the a and d positions are involved in knobs-into-holes packing. However, along with increasing oligomeric state, residues other than those at a and d positions become more involved in the helix-helix interfaces. Specifically, the e and g sites become progressively buried [30]. Actually in pentameric coiled coils and above, some of the sequences approximate to *HxxHHxH* repeats, rather than traditional *HxxHxxx* heptad repeats. The above *HxxHHxH* pattern can be split up to give two heptad repeats, *HxxxHxx* plus *xxxHxxH*, both of which are 3,4-repeats of H-type residues. The two heptad repeats generate two hydrophobic interfaces separated by 103°. The sequence of the seven helix bundle [31] (PDB ID: 2HY6), the largest freely-standing coiled coil thus far, agrees with the *HxxHHxH* pattern very well. The extended KIH packing contributes to the stabilization of the higher oligomeric state. There is another way of overlapping two heptad repeats which creates *HxHHxHx* sequence pattern. In this pattern, the two hydrophobic interfaces are further separated by 154°. It has been predicted that this sequence repeat could lead to the formation of open (α-sheet) and closed (α-cylinder) multihelical assemblies [32]. In nature, the bacterial transmembrane protein TolC provided the first example of α-cylinder [33], a 12-helix α-barrel-like domain contributed by three monomers, where heptad registers were unambiguously determined revealing knobs at relative a, c, d, and f positions. Interestingly, recent X-ray and cryo-EM studies [34-37] of the viral genome-delivery portal from bacteriophage P22 reveal a long 12-helix barrel, proposed to be responsible genome ejection into host cells. At this stage the resolution of this part of the structure is insufficient. Nonetheless, it would appear that the 12-helix barrel of TolC is not unique.

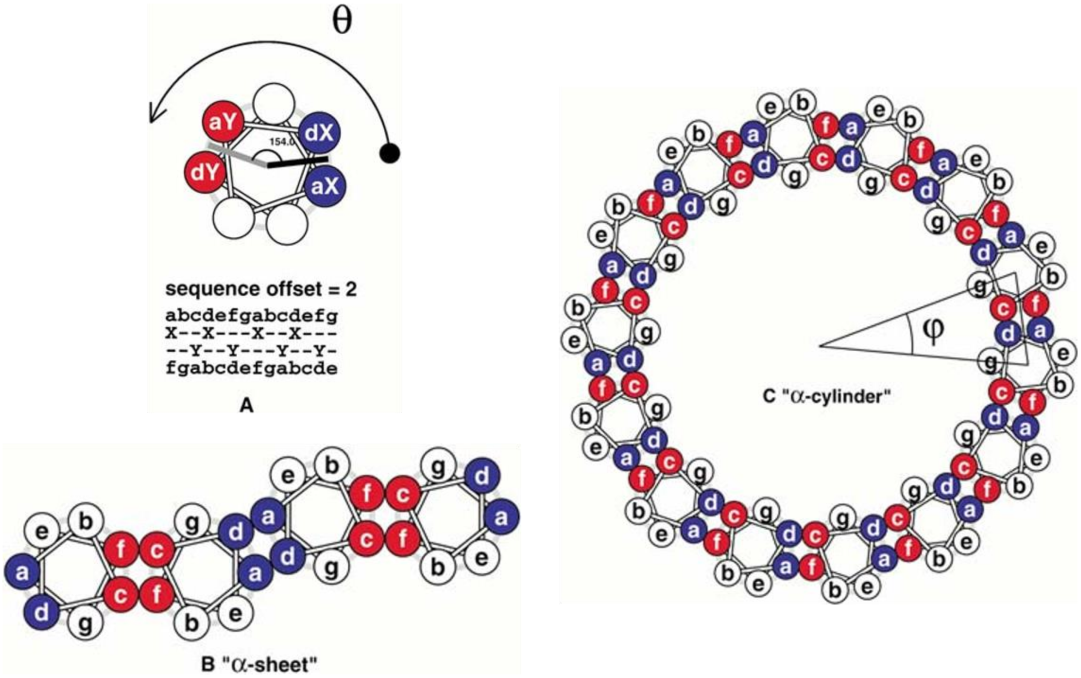


Figure 1.9. Open (“α-sheet”) (B) and closed (“α-cylinder”) (C) high-order coiled coil were anticipated to form from bifaceted α-helices (A). [32]

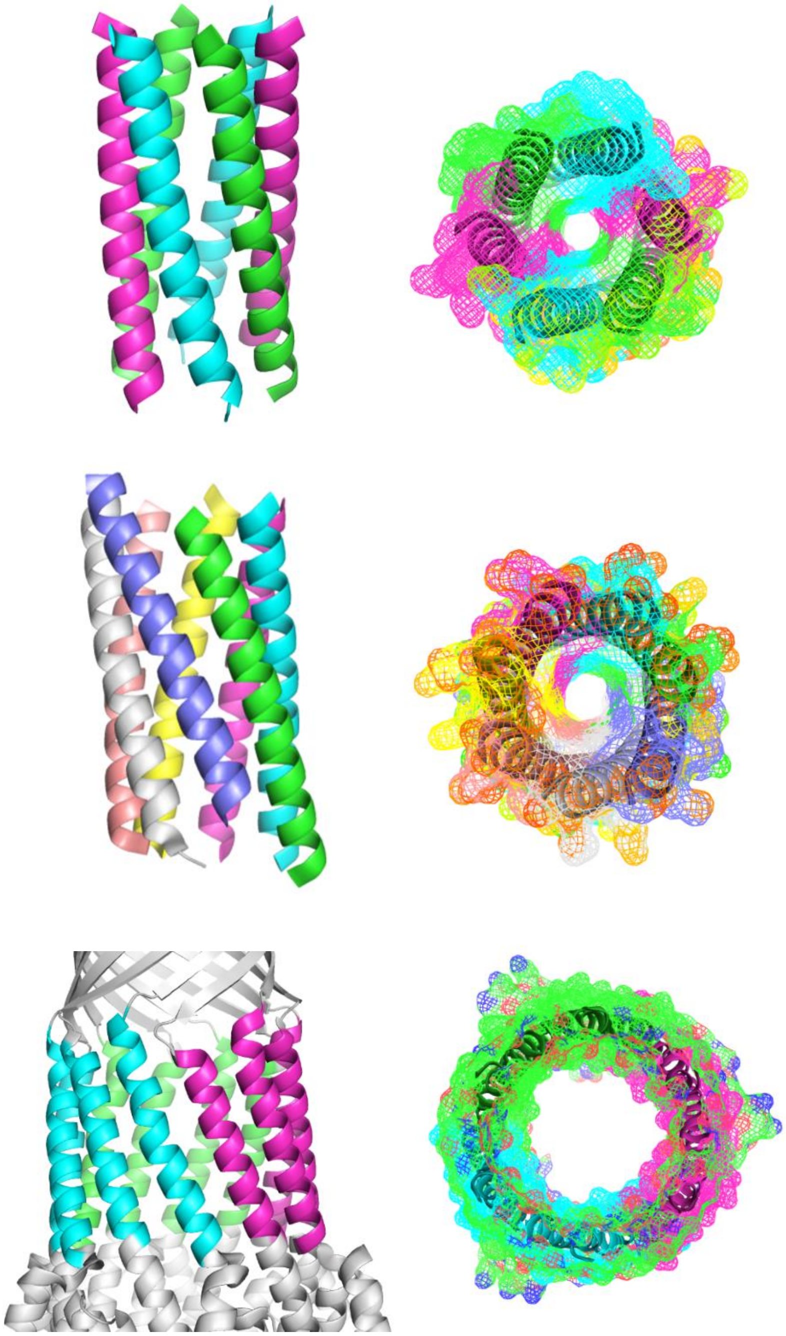


Figure 1.10. Complex coiled coils define a central channel with tailorable dimensions and chemistry. Top: Top and side view of a hexameric coiled coil (PDB ID 3R3K) with a central channel of 5-6 Å. Middle: Heptamer (PDB ID 2HY6), 7 Å. Bottom: dodecamer (PDB ID 1EK9), 25 Å.

Altought in nature, dimer, trimer, and tetramer are the most common oligomeric states of coiled coil domains, the potential to form larger helical bundle has long been recognized. Coiled coil complexes above pentamer are particularly interesting because they define a central channel with well-defined dimensions and chemistry, which could have an impact in the design of ion-channels, binding proteins, sensors and even enzymes. Shown in Figure 1.10 are de novo designed coiled-coil hexamer [38], heptamer [31], as well as a dodecamer [33] helical bundle domain of E. coli transmembrane protein TolC which define a central channel of 5-6 Å, 7Å, and 25Å, respectively[30].

1.4 Coiled-coil supramolecular assemblies

The first investigations into the structure of coiled coils were made by William Astbury at the University of Leeds in the 1930s using X-ray diffraction to study the structure of protein fibers [26]. Astbury named the class of proteins showing a strong meridional arc at 5.15 Å in their diffraction pattern 'k-m-e-f', for keratin, myosin, epidermin, and fibrinogen. The 5.15 Å spacing is the axial rise per turn of coiled coils which is unknown at that time. Coiled coil supramolecular assemblies constitute a major form of protein fibers due to the fact, at least in part, that coiled coils are usually long, rigid oligomers of helices with regular packing interactions and extended exposed surfaces. This enables them to assemble into large, mechanically rigid structures such as hair, horn, feathers (keratin), and blood clots (fibrinogen); extracellular matrices (laminin) and cytoskeletal networks (intermediate filaments); and a broad array of filaments (flagellins, pilins, phage coat proteins) [26].

Over the last decade and a half, a variety of synthetic coiled-coil supramolecular assemblies have been successfully constructed taking advantages of the relatively well-understood sequence-structure relationships of coiled coils. In the pursuit of structurally defined protein-based supramolecular materials, our group has previously designed and characterized self-assembling peptide fibers based on dimeric and trimeric coiled-coil motifs, and incorporated controlling mechanism into the self-assembly in response to environmental cues. In 2004, Zimenkov et al[17] designed a peptide **YZ1** with 42 amino acid residues which favors the formation of parallel dimeric coiled coil by selectively placing isoleucine at a positions and leucine at d positions. Moreover, **YZ1** prefers a staggered orientation between adjacent peptides that corresponded to an axial displacement of three heptads utilizing peripheral electrostatic interactions. The "sticky ends" enable the formation of high-aspect-ratio supramolecular fibers.

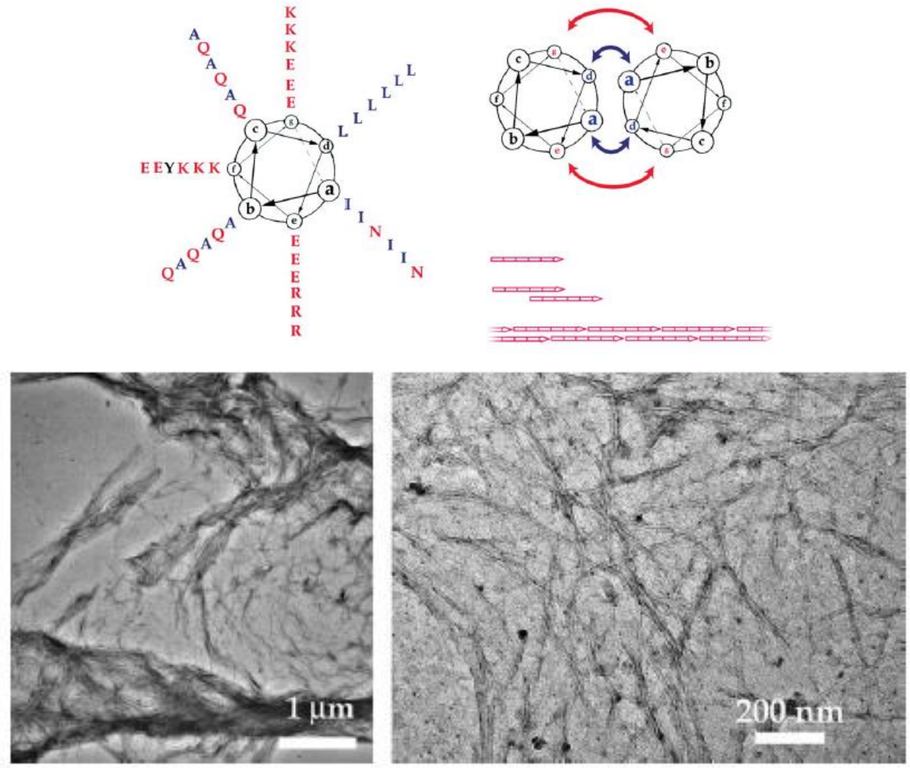


Figure 1.11. Design of peptide sequence of YZ1. [17] YZ1 favors the formation of staggered dimeric coiled coils which then associate further to form high-aspect-ratio fibers.

In other studies [18, 19], peptide TZ1H was designed and characterized based on a trimeric coiled-coil motif that reversibly assembles into long-aspect-ratio helical fibrils as a consequence of a pH and metal-induced conformational transition, respectively. These processes resemble the self-assembly mechanism of native protein fibrils in which supramolecular structures could be created or destabilized in response to environmental stimuli [39].

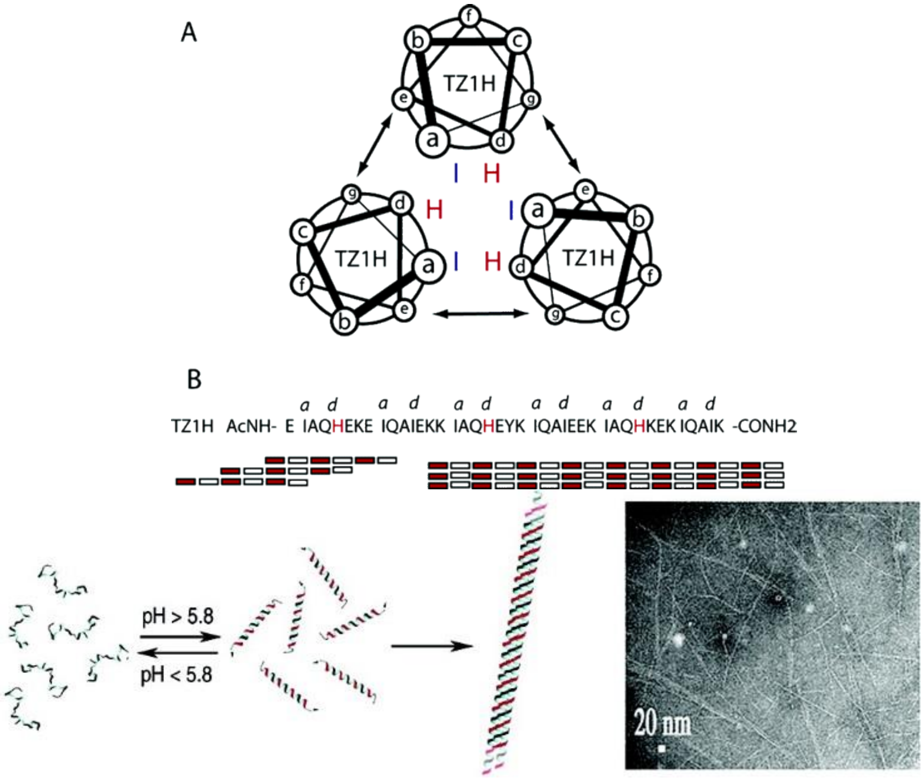


Figure 1.12. The design of peptide TZ1H. In the sequence of TZ1H, three histidine residues are in the core position which act as sensitive switch in response to environmental pH changes which regulates the self-assembly process [18]. Metal coordination was shown to play a similar role with pH changes [19].

Meanwhile, researchers from other groups have also made tremendous efforts towards fabrication of complex, structurally defined supramolecular assemblies based on coiled coil motif. The self-assembling fiber system, SAFs [40-42], designed by Woolfson and colleagues, is the most representative example which signified a breakthrough in this area. Paracrystalline fibers were created using a heterodimeric, sticky ended, coiled-coil building block, similar with the strategy used in designing **YZ1**. Recently, Woolfson group reported another interesting work on building self-assembling cages from coiled-coil modules [43]. The design comprises two, noncovalent, heterodimeric and homotrimeric coiled-coil bundles. These are joined back to back to render two complementary hubs, which when mixed form hexagonal networks that close to form cages approximately 100 nanometers in diameter. The design strategy provides routes to new compartmental materials for the delivery of drugs and biologicals.

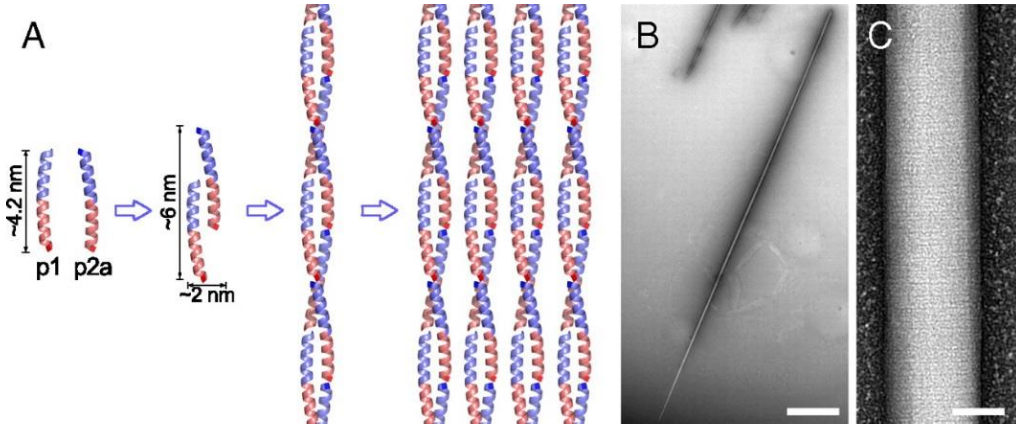


Figure 1.13. Structure and hierarchy of SAFs. [40] A: SAF-p1 and SAF-p2a form heterodimeric sticky-ended unit that associate further to form coiled-coil fibrils. B: Fibrils associate laterally to form fibers that display striations (C) under negative-stain TEM. Scale bars, 2 µm (B) and 50 nm (C).

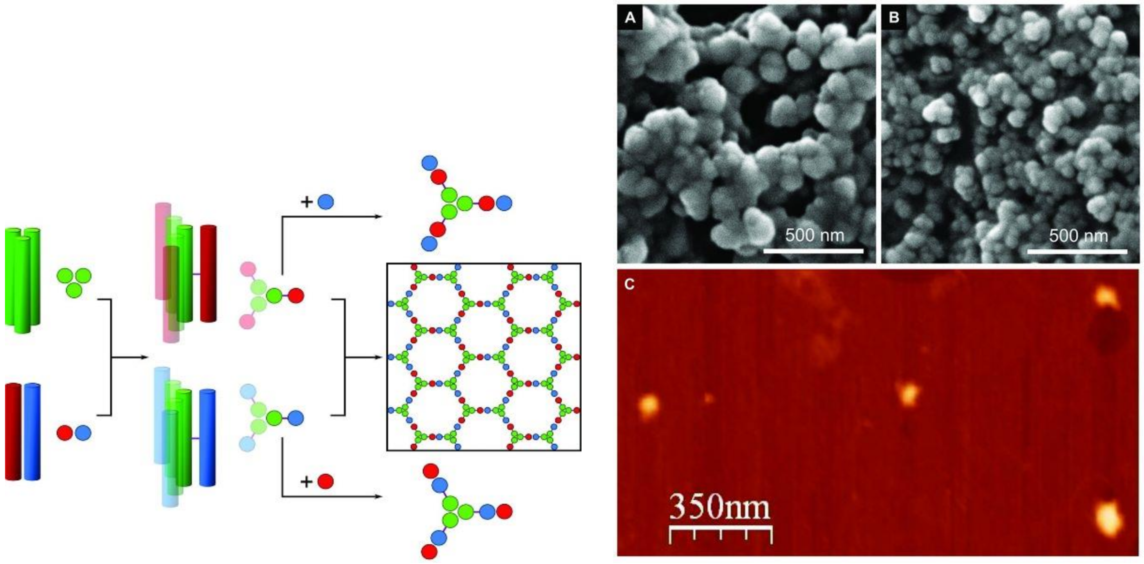


Figure 1.14. Self-assembling cages from coiled-coil modules. [43] Left: Schematics for the design and self-assembly of peptidebased cages. Dimeric coiled coil and trimeric coiled coil are used as building blocks. Mixing the subunits first produces a hexagonal network which closes to form cages. Right: SEM and AFM images of peptide cages.

Ever since the emergence of “DNA origami”, protein chemists have not stopped being jealous of DNA architects who have the ability of folding the molecule into minuscule smiley faces, boxes, letters of the alphabet and dozens of other intricate shapes. However, a piece of great work in the field of programmed protein self-assembly published in 2013 by Jerala and coworkers shed light on “protein origami”. They have designed and demonstrated the formation of a tetrahedron that self-assembles from a single polypeptide chain comprising 12 concatenated orthogonal coiled-coil segments separated by flexible peptide hinges [44].

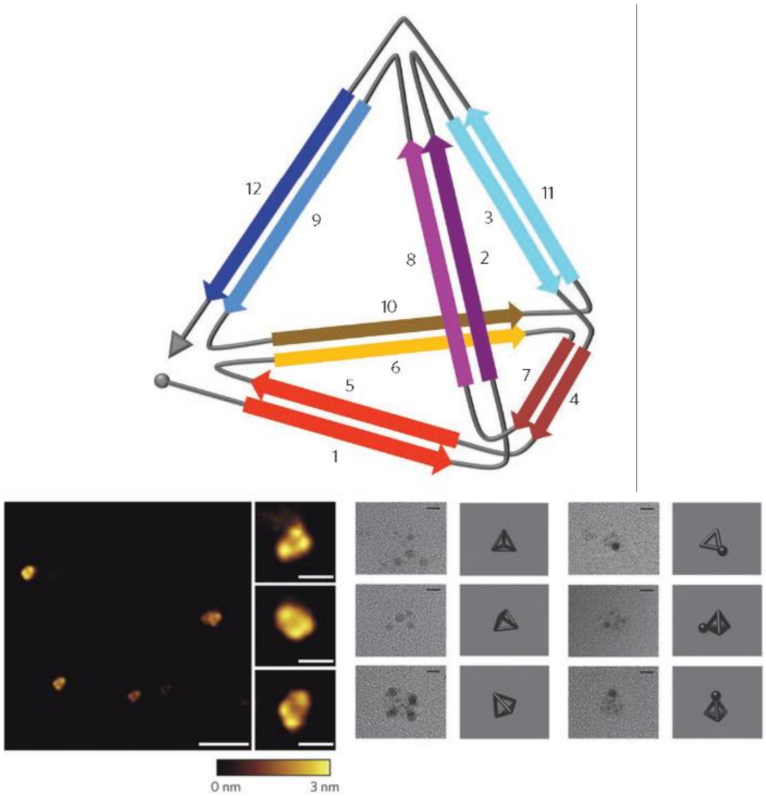


Figure 1.15. A single-chain polypeptide tetrahedron assembled from coiled-coil segments. Top: Schematic representation of the designed polypeptide tetrahedron. Six orthogonal dimeric coiled-coil pairs guide the formation of the tetrahedron. Bottom: AFM and TEM images showing the formation of tetrahedrons.

1.5 Summary

Chemical self-assembly represents the most practical approach for the design of structurally-defined functional materials which have been historically the most challenging to rationally construct and the most difficult to structurally analyze. Due to the relatively straightforward sequence characteristics and the well-understood sequence-structure relationship, coiled-coil motifs show great potential as a model system for fabricating structurally defined functional materials on the nano- to micrometer scale. Over the past decade or so, coiled-coil supramolecular assemblies, such as fibers, cages, and tetrahedrons described in this chapter, have been successfully designed and characterized. Information gained from the structural characterization of these coiled-coil supramolecular assemblies, can be employed to design more complex architectures with versatile functionality.

Previously reported assemblies primarily used dimeric and trimeric coiled coils as building blocks. However, coiled coils that are composed five or more α-helices are of particular interests because they define a central channel with tailorable dimensions and chemistry. They are ideal candidates for the construction of structurally-defined tubular and porous functional materials which have promising applications in directional transport, controlled release, and catalysis.

The crystal structures of de novo designed hexameric and heptameric coiled coils which were reported in recent years, improved our understanding of higher-order multihelical oligomers and largely inspired our studies. In the following chapters in this volume, different research topics will be presented including controlling peptide self-assembly through metal-induced registry shift, recoding heptameric coiled coil for nanotube formation, and designing large-diameter helical nanotubes.

1.6 References:

1. Service, R.F., *How far can we push chemical self-assembly.* Science, 2005. **309**(5731): p. 95-95.

2. *Ribosomal Subunits.* *http://www.rcsb.org/pdb/101/motm.do?momID=10*. [cited 2013 November 16th].

3. *The Structural Biology of HIV.* *http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/educational_resources/hiv-animation.html*. [cited 2013 November 16th].

4. Gerhart, J., M. Kirschner, and E.S. Moderbacher, *Cells, embryos, and evolution: Toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability.* 1997.

5. Kirschner, M., J. Gerhart, and T. Mitchison, *Molecular "vitalism".* Cell, 2000. **100**(1): p. 79-88.

6. Frederickson, R.M., *Macromolecular matchmaking: advances in two-hybrid and related technologies.* Current Opinion in Biotechnology, 1998. **9**(1): p. 90-96.

7. Goodsell, D.S. *Illustrations for Public Use.* *http://mgl.scripps.edu/people/goodsell/illustration/public*. [cited 2013 November 16th].

8. Linko, V. and H. Dietz, *The enabled state of DNA nanotechnology.* Curr Opin Biotechnol, 2013. **24**(4): p. 555-61.

9. Wei, B., M. Dai, and P. Yin, *Complex shapes self-assembled from single-stranded DNA tiles.* Nature, 2012. **485**(7400): p. 623-6.

10. Ke, Y., et al., *Three-dimensional structures self-assembled from DNA bricks.* Science, 2012. **338**(6111): p. 1177-83.

11. Pinheiro, A.V., et al., *Challenges and opportunities for structural DNA nanotechnology.* Nat Nanotechnol, 2011. **6**(12): p. 763-72.

12. Andersen, E.S., et al., *Self-assembly of a nanoscale DNA box with a controllable lid.* Nature, 2009. **459**(7243): p. 73-6.

13. Dreher, M.R., et al., *Temperature triggered self-assembly of polypeptides into multivalent spherical micelles.* Journal of the American Chemical Society, 2008. **130**(2): p. 687-694.

14. Herrero-Vanrell, R., et al., *Self-assembled particles of an elastin-like polymer as vehicles for controlled drug release.* Journal of Controlled Release, 2005. **102**(1): p. 113-122.

15. Lee, T.A.T., et al., *Thermo-reversible self-assembly of nanoparticles derived from elastin-mimetic polypeptides.* Advanced Materials, 2000. **12**(15): p. 1105-+.

16. Rele, S., et al., *D-periodic collagen-mimetic microfibers.* Journal of the American Chemical Society, 2007. **129**(47): p. 14780-7.

17. Zimenkov, Y., et al., *Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif.* Tetrahedron, 2004. **60**(34): p. 7237-7246.

18. Zimenkov, Y., et al., *Rational design of a reversible pH-responsive switch for peptide self-assembly.* Journal of the American Chemical Society, 2006. **128**(21): p. 6770-6771.

19. Dublin, S.N. and V.P. Conticello, *Design of a selective metal ion switch for self-assembly of peptide-based fibrils.* Journal of the American Chemical Society, 2008. **130**(1): p. 49-+.

20. Ballister, E.R., et al., *In vitro self-assembly of tailorable nanotubes from a simple protein building block.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3733-8.

21. Petrov, A., S. Lombardo, and G.F. Audette, *Fibril-mediated oligomerization of pilin-derived protein nanotubes.* Journal of Nanobiotechnology, 2013. **11**.

22. Childers, W.S., et al., *Peptide membranes in chemical evolution.* Curr Opin Chem Biol, 2009. **13**(5-6): p. 652-9.

23. Childers, W.S., et al., *Peptides Organized as Bilayer Membranes.* Angewandte Chemie-International Edition, 2010. **49**(24): p. 4104-4107.

24. King, N.P., et al., *Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy.* Science, 2012. **336**(6085): p. 1171-1174.

25. Nam, K.T., et al., *Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers.* Nature Materials, 2010. **9**(5): p. 454-460.

26. Lupas, A.N. and M. Gruber, *The structure of alpha-helical coiled coils.* Fibrous Proteins: Coiled-Coils, Collagen and Elastomers, 2005. **70**: p. 37-+.

27. Cohen, C. and D.A. Parry, *Alpha-helical coiled coils and bundles: how to design an alpha-helical protein.* Proteins, 1990. **7**(1): p. 1-15.

28. Crick, F.H.C., *Is Alpha-Keratin a Coiled Coil.* Nature, 1952. **170**(4334): p. 882-883.

29. Crick, F.H.C., *The Packing of Alpha-Helices - Simple Coiled-Coils.* Acta Crystallographica, 1953. **6**(8-9): p. 689-697.

30. Woolfson, D.N., et al., *New currency for old rope: from coiled-coil assemblies to alpha-helical barrels.* Curr Opin Struct Biol, 2012. **22**(4): p. 432-41.

31. Liu, J., et al., *A seven-helix coiled coil.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(42): p. 15457-15462.

32. Walshaw, J. and D.N. Woolfson, *Open-and-shut cases in coiled-coil assembly: alpha-sheets and alpha-cylinders.* Protein Science, 2001. **10**(3): p. 668-673.

33. Koronakis, V., et al., *Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export.* Nature, 2000. **405**(6789): p. 914-9.

34. Olia, A.S., et al., *Three-dimensional structure of a viral genome-delivery portal vertex.* Nature Structural & Molecular Biology, 2011. **18**(5): p. 597-U107.

35. Tang, J.H., et al., *Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22.* Structure, 2011. **19**(4): p. 496-502.

36. Parent, K.N., et al., *Cryo-reconstructions of P22 polyheads suggest that phage assembly is nucleated by trimeric interactions among coat proteins.* Phys Biol, 2010. **7**(4): p. 045004.

37. Chang, J., et al., *Cryo-EM asymmetric reconstruction of bacteriophage P22 reveals organization of its DNA packaging and infecting machinery.* Structure, 2006. **14**(6): p. 1073-1082.

38. Zaccai, N.R., et al., *A de novo peptide hexamer with a mutable channel.* Nature Chemical Biology, 2011. **7**(12): p. 935-941.

39. Conticello, V.P., S. Dublin, and Y. Zimenkov, *Engineering responsive mechanisms to control the assembly of peptide-based nanostructures.* Biochemical Society Transactions, 2009. **37**: p. 653-659.

40. Sharp, T.H., et al., *Cryo-transmission electron microscopy structure of a gigadalton peptide fiber of de novo design.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(33): p. 13266-13271.

41. Pandya, M.J., et al., *Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis.* Biochemistry, 2000. **39**(30): p. 8728-8734.

42. Ryadnov, M.G. and D.N. Woolfson, *Engineering the morphology of a selfassembling protein fibre.* Nature Materials, 2003. **2**(5): p. 329-332.

43. Fletcher, J.M., et al., *Self-assembling cages from coiled-coil peptide modules.* Science, 2013. **340**(6132): p. 595-9.

44. Gradisar, H., et al., *Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments.* Nature Chemical Biology, 2013. **9**(6): p. 362-+.

Chapter II: Controlling Self-assembly of a Peptide-Based Material via Metal-Ion Induced Registry Shift

2.1 Introduction and Design

Sequence-specific biomolecules, i.e., proteins and nucleic acids, confer significant advantages as substrates for the construction of structurally complex supramolecular materials [1-3]. Foremost, the sequence-structure correlations that have been elucidated from physical analysis of native biomolecular structures provide a context for the rational design of novel materials [4-6]. However, self-assembly of most designed peptide-based materials occurs commensurately with protein folding. In contrast, native protein assemblies undergo reversible self-association due to subtle shifts in conformation that are propagated hierarchically. We describe herein a mechanism to control supramolecular assembly of a synthetic peptide-based material through registry selection. This process recapitulates the native mechanism of fibril assembly in which a ligand binding event gates a reversible conformational transition between alternate forms of a folded peptide structure.

Peptide **TZ1C2**, a 41-residue sequence comprising six heptad repeats (Figure 1), was derived from modification of a previously reported trimeric coiled-coil **TZ1**. [7, 8] Sequence variants of peptide **TZ1** have been shown to self-assemble into high aspect-ratio helical fibrils, in which lateral registration between adjacent helical protomers is enforced through Coulombic interactions between charged residues at the *e*- and *g*-positions of the heptad repeats. Peptide **TZ1C2** differs from **TZ1** in that two cysteine residues were introduced in place of isoleucines at core *a*-positions within the second and fifth heptad sequences. The sulfhydryl groups of the cysteine residues have the latent capacity to serve as ligands for the coordination of metal ions.[9] However, in order to create effective binding sites, the cysteine residues should be oriented proximally across the helical interface, which requires an in-register alignment of helices within a trimeric coiled-coil structure.

Peptide **TZ1C2** can populate two distinct orientations within this structural context. The preferred orientation of helices depends on the presence of a metal ion, such as Cd(II), that can form a stable complex with the thiolate ligands of **TZ1C2**. Metal ion coordination may provide sufficient thermodynamic driving force to overcome electrostatic repulsion between residues at the *e*/*g*-positions and drive re-alignment of the structure (Figure 1). In the absence of metal coordination, electrostatic interactions and steric complementation[10] at core positions should favor the staggered alignment.[7, 8] Thus, the balance of intermolecular forces can be manipulated through changes in environmental conditions to shift the helices reversibly from a staggered (out-of-register) [7, 8, 11, 12] to aligned (in-register) orientation with a concomitant transition from a fibril to a discrete helical bundle.

The scheme for dis-assembly of the **TZ1C2** fibrils requires the presence of a metal ion to induce a registry shift that orients the cysteine residues within the same layer of the triple helical coiled-coil (Figure 2.1). Previous investigations have demonstrated that the chalcophilic cadmium(II) ion can bind to structurally related coiled-coil trimers in which cysteines have been substituted into core *a*- or *d*-positions within the heptad repeats. [13-15] Cysteine residues were placed at the *a*-positions of the second and fifth heptad repeats of **TZ1C2**. Prior research had indicated a preference for Cd(II) complex formation at the trigonal binding sites at the *a*-postions of coiled-coil trimers. [16] In addition, cadmium has the advantage that complexation can be monitored using ^{113}Cd NMR spectroscopy [17] and ^{111m}Cd perturbed angular correlation (PAC) spectroscopy, [18] which are sensitive tools for analysis of ligation environment and coordination geometry of metal ion-binding sites in proteins. [19]

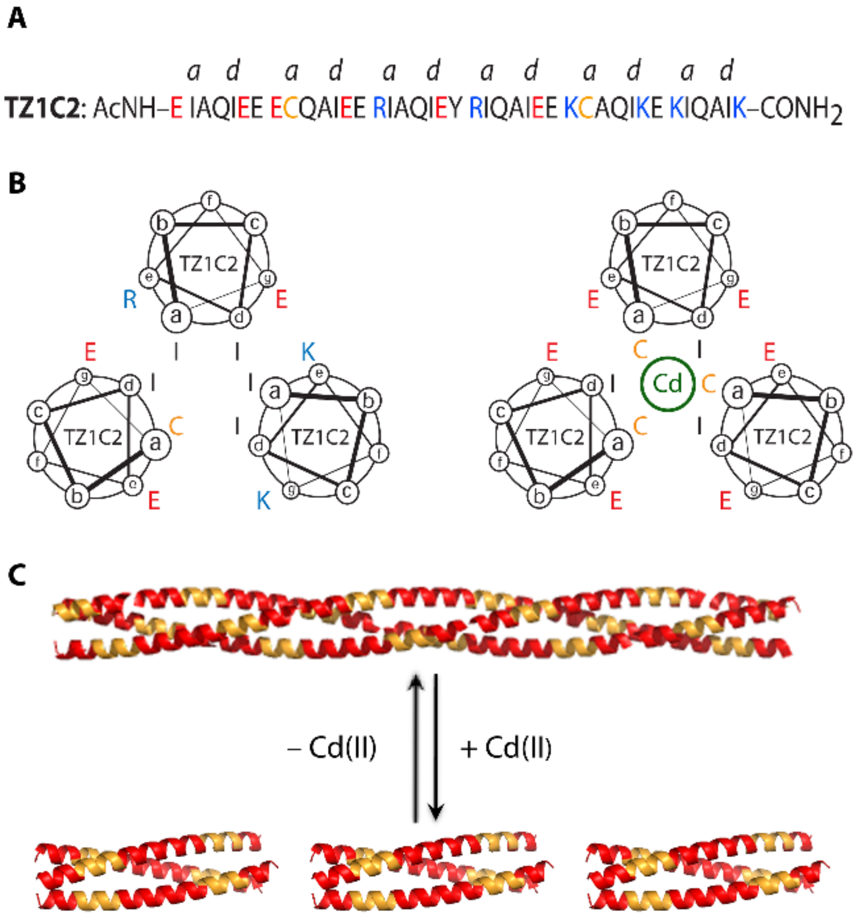


Figure 2.1. Design of TZ1C2 peptide. (A) Sequence of peptide TZ1C2. (B) Helical wheel representation of core layer packing within the second heptad of TZ1C2 in the staggered (left) and aligned (right) orientations. (C) Schematic representation of the registry shift that results from Cd(II) coordination to the cysteine residues of TZ1C2 (yellow: Cys-containing heptads and red: non-Cys-containing heptads).

2.2 Results and Discussion

2.2.1 Circular Dichroism

The circular dichroism (CD) spectrum of **TZ1C2** (100 μM) in TAPS buffer was consistent with the presence of an α-helical conformation throughout the pH range from 6 to 9 (Figure 2.2). CD spectropolarimetry of Cd(II)-containing solutions of peptide **TZ1C2** indicated that the α-helical conformation was retained in the complex, although the net helicity was reduced with respect to the corresponding value in the absence of cadmium ion (Figure 2.2).

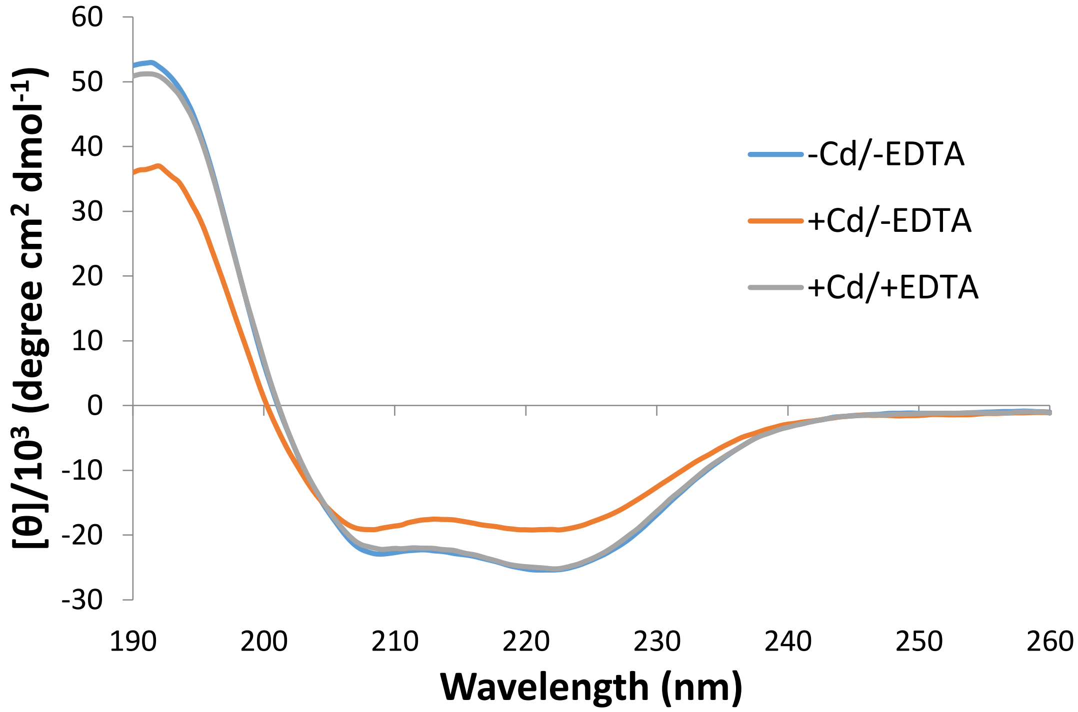


Figure 2.2. CD spectra of peptide TZ1C2 (100 μM) in TAPS buffer (10 mM, pH 8.5) and NaCl (100 mM). The existence of Cd(II) in peptide solution reduced the helicity.

CD titration experiments were performed to investigate the binding stoichiometry of Cd(II) to peptide. Change in CD spectropolarimetric signal at 222 nm ($[\theta]_{222}$) was monitored as a function of the number of equivalent of Cd(II) ion. The binding sites saturate at circa 0.65 equivalents of Cd(II), which closely corresponds to the expected value of 0.67 equivalents Cd(II) per peptide for binding of two cadmium ions to a coiled-coil trimer.

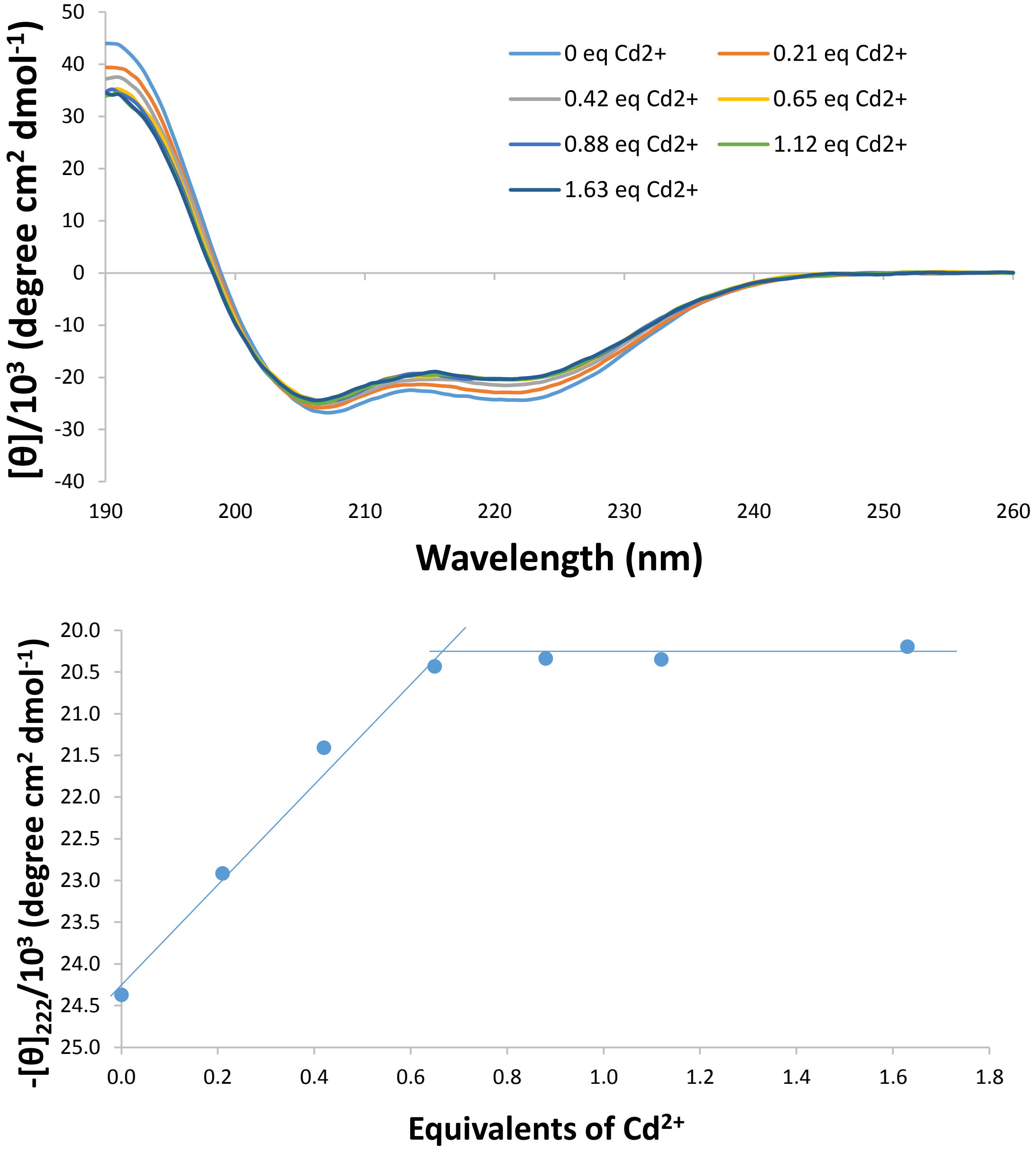


Figure 2.3. CD spectropolarimetric titration of TZ1C2 (75 μM) in TAPS buffer (10 mM, pH 8.5, 100 mM NaCl) with increasing concentrations of Cd(II)Cl2. Change in

CD spectropolarimetric signal at 222 nm ([θ]222) is shown as a function of the number of equivalents of Cd(II) ion (bottom).

2.2.2 Flow Linear Dichroism

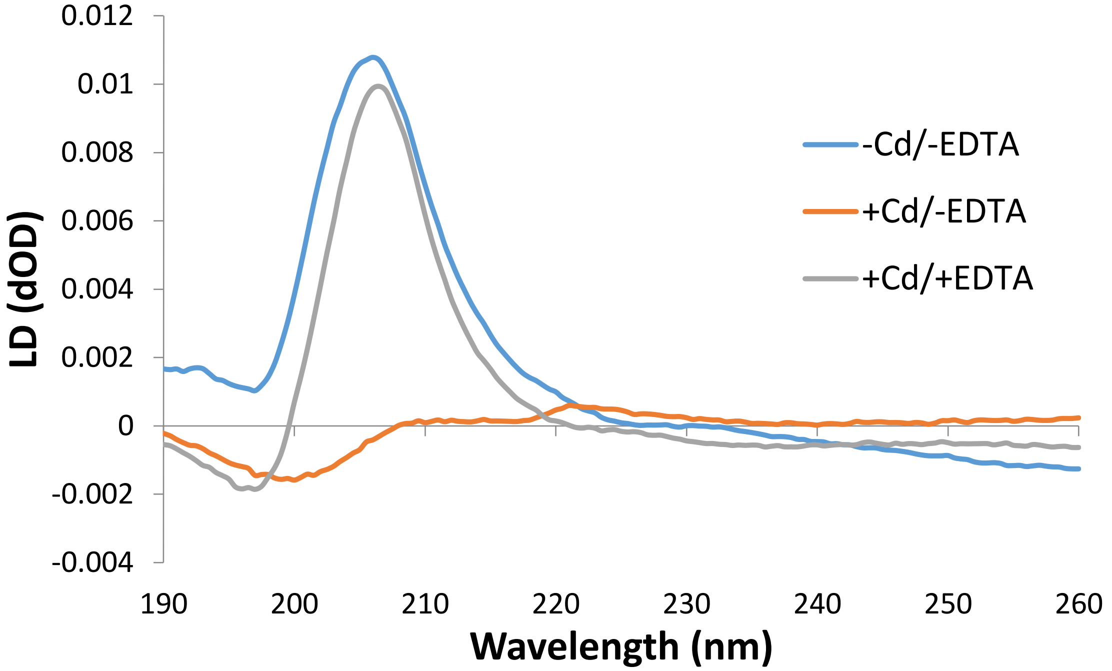


Figure 2.4. Flow LD spectra of peptide TZ1C2 (100 µM) in TAPS buffer (10 mM, pH 8.5) and NaCl (100 mM).

Flow linear dichroism (LD) of an aqueous solution of **TZ1C2** (100 µM) in TAPS buffer (10 mM, pH 8.5) displays a strong positive signal at 206 nm under a Couette flow of 3,000 rpm (Figure 2.4). The observed spectroscopic response indicated the presence of extended helical assemblies of **TZ1C2** that are stable to flow alignment. The flow LD data suggested a conformational arrangement of **TZ1C2** in which the amide bonds, and consequently, the α-helices, are oriented parallel to the flow direction. Similar behavior has been observed in the flow LD spectra of synthetic coiled-coil fibers and tropomyosin under flow alignment.[20]

The flow linear dichroism spectrum of **TZ1C2** was significantly reduced in the presence of cadmium(II). The strong positive signal at 206 nm disappeared and was replaced with a weak negative signal at 200 nm (Figure 2.4). The spectroscopic behavior was consistent with disassembly of **TZ1C2** from extended fibrillar structures into smaller species that could not be aligned under Couette flow.

2.2.3 Transmission Electron Microscopy

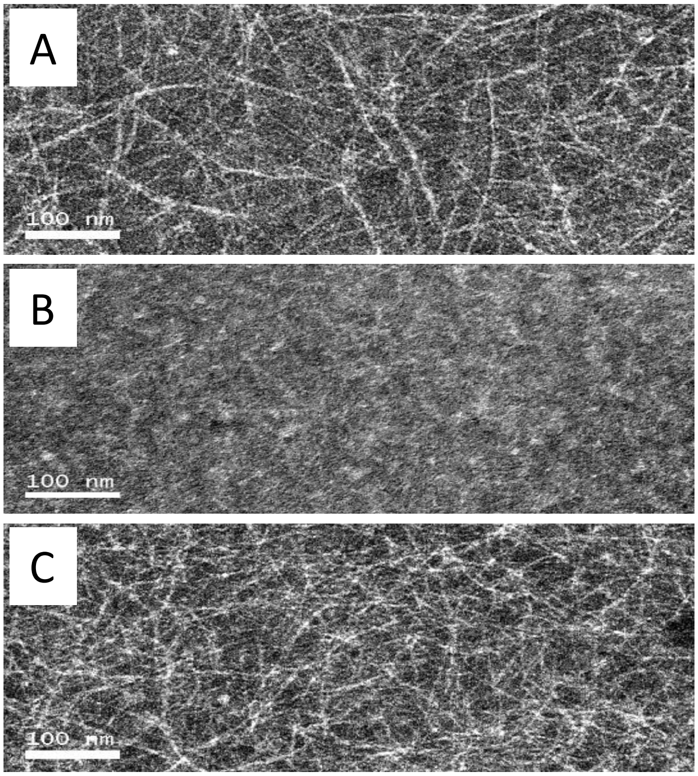


Figure 2.5. Negative-stain TEM images of TZ1C2 fibrils (500 µM in TAPS buffer (10 mM, pH 8.5)) and NaCl (100 mM) before Cd(II) addition (A) and after Cd(II) addition (B) followed treatment with excess EDTA (C).

TEM analysis of assemblies derived from **TZ1C2** confirmed the flow linear dichroism analysis. High aspect-ratio fibrils are observed in negatively stained specimens of **TZ1C2** at concentrations ≥ 100 μM (Figure 2.5). The combined data support the hypothesis that **TZ1C2** self-assembles into α-helical fibrils in which the peptides adopt a parallel, out-of-register orientation as observed previously for other **TZ1** peptide derivatives [7, 8].

Similarly, TEM analysis did not detect the formation of fibrils in aqueous solutions of **TZ1C2** in the presence of cadmium(II). The dis-assembly was readily reversed in the presence of excess EDTA (200 μM), which can tightly bind to Cd(II) ion.[21] The addition of EDTA restored the original flow LD signal (Figure 2.4) and induced the formation of fibrils that were observed in the TEM analysis (Figure 2.5).

2.2.4 Analytical Ultracentrifugation

The oligomeric state of **TZ1C2** in solution was investigated in the presence and absence of Cd(II) ion by sedimentation velocity analytical ultracentrifugation. The effect of Cd(II) ion was further examined as a function of **TZ1C2** concentration. When **TZ1C2** was measured under low loading concentration (160 μM) in the presence of Cd(II), a slightly heterogeneous sedimentation distribution with a weight-average s-value of as 1.68S was observed. Further analysis by genetic algorithm-Monte Carlo[22] methods indicated the presence of a small amount of monomeric **TZ1C2** species, as well as a major trimeric species (42% of absorption), and a small amount of a larger oligomer consistent with a hexameric species. Under higher loading concentration (480 μM), the weight average s-value shifted slightly to 1.89S, and the relative amount of trimer increased to 62% of absorption at the expense of the other species. This observation is consistent with a reversible mass action effect. In the absence of Cd(II), the sedimentation distributions changed drastically, and 38% of the absorbance corresponded to a monomeric species at low loading concentration (140 μM), while the remaining absorbance sedimented as a heterogeneous mixture with s-values ranging between 3-7S, corresponding to highly anisotropic aggregates ranging between 0.4-1 million Da in size (Figure 2.6). The latter species are consistent with the fibrils observed in the TEM images of **TZ1C2**. From these data it can be concluded that Cd(II) facilitates the formation of reversibly self-associating trimer, while, in the absence of Cd(II), **TZ1C2** is partially monomeric and partially forms much larger, irreversible aggregates. In addition, non-denaturing nano-electrospray ionization mass spectrometry confirmed the identity of the **TZ1C2** trimer as the di-cadmium adduct (Figure 2.9).

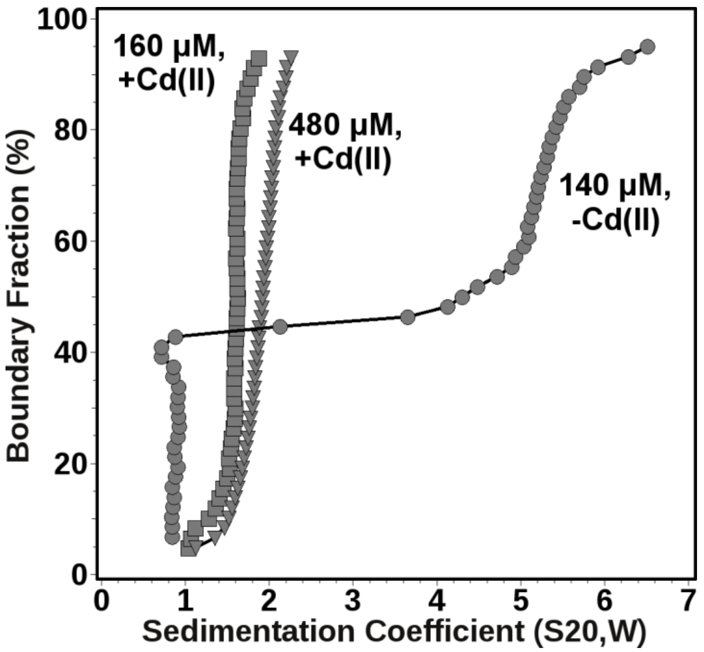


Figure 2.6. Diffusion-corrected integral sedimentation coefficient distributions obtained from a van Holde-Weischet analysis. [23] Shown are results for TZ1C2 in

the presence of excess Cd(II) at 160 µM (squares) and 480 µM (triangles) and in the absence of Cd(II) at 140 µM (circles) loading concentration.

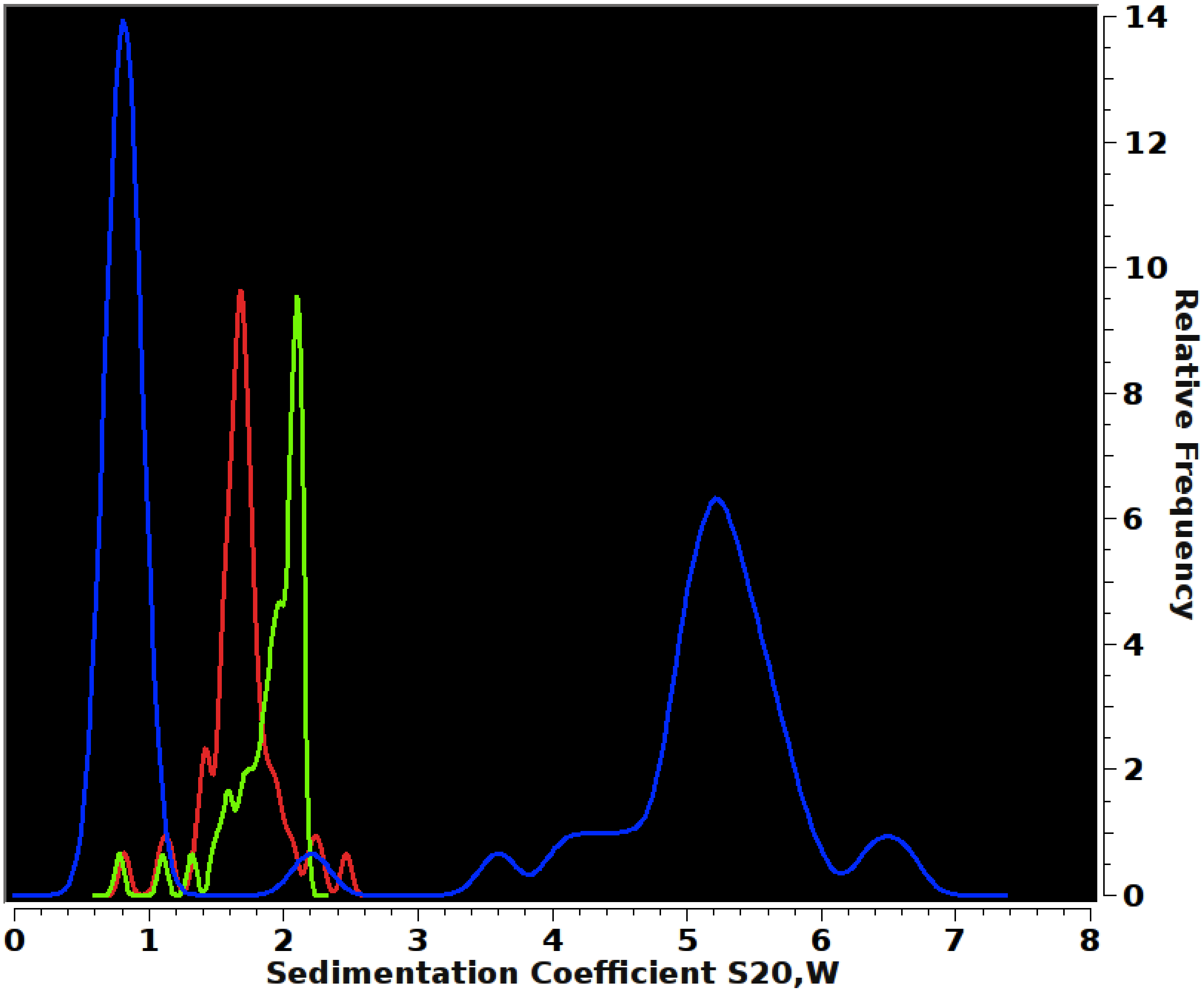


Figure 2.7. The g(s) distributions for TZ1C2 in the presence of excess Cd(II) at 160 µM (red) and 480 µM (green) and in the absence of Cd(II) at 140 µM (blue) loading concentration.

2.2.5 ^{113}Cd NMR spectroscopy and ^{111m}Cd PAC spectroscopy

The structural role of the Cd(II) ion in the **TZ1C2** assemblies was interrogated using a combination of ^{113}Cd NMR spectroscopy and ^{111m}Cd PAC spectroscopy (Figure 2.10) performed on isotopically enriched specimens. A single resonance was observed at 619 ppm in the ^{113}Cd NMR spectrum of an aqueous solution of peptide **TZ1C2** that had been treated with 2 equivalents of $^{113}CdCl_2$. The chemical shift was consistent with nearly identical CdS_3X coordination environments for the two structurally similar metal ion binding sites in **TZ1C2**. The ^{113}Cd NMR spectroscopic data were similar to that reported for Cd-(*S*-cysteinyl)$_3$ complexes of structurally analogous coiled-coil trimers within the TRI series.[13, 19] These spectroscopic results were interpreted in terms of a dynamic equilibrium on the NMR time-scale for a Cd(II) complex that involved three cysteinyl thiolate groups and a reversibly bound water molecule (X).

However, the ^{111m}Cd PAC spectroscopic data appear to refute a similar scenario for the cadmium(II) complex of **TZ1C2**. The ^{111m}Cd PAC spectrum of fully complexed **TZ1C2** gives a single very well defined nuclear quadrupole interaction (NQI), indicating that all Cd(II) ions are found in practically the same coordination geometry, i.e., the two Cd(II) binding sites are highly similar (Figure 2.10). The NQI, and thus the local structure, is different from those observed for Cd(II) complexes of the TRI peptide family. Both the frequency (ω_0 about 0.399 rad/ns) and the asymmetry parameter (η about 0.61) are relatively high for the **TZ1C2** complex. The high asymmetry parameter indicates the absence of axial symmetry, i.e., idealized trigonal planar CdS_3, and tetrahedral CdS_3X structures can be ruled out. The combined spectroscopic data suggest the presence of a single type of Cd(II) binding site, most likely of the CdS_3X type, in which the X-ligand is not in dynamic exchange. The deviation from an axially symmetric coordination environment may arise from structural distortions due to electrostatic repulsion between similarly charged *e*/*g*-residues at the Cd(II) coordination sites within the in-register coiled-coil trimer (Figure 2.1).

Table 2.1. Parameters fitted to PAC-data. The numbers in parenthesis are the standard deviations of the fitted parameters.

Peptide	t	pH	ω_0	η	$\Delta\omega_0/\omega_0$	$1/\tau_c$	A	χ_r^2
	℃	(@ 1 ℃)	(rad/ns)		×100	s^{-1}	×100	
TZ1C2	1	9.1	0.3993(5)	0.614(2)	1.5(2)	4.7(8)	7.4(3)	1.35

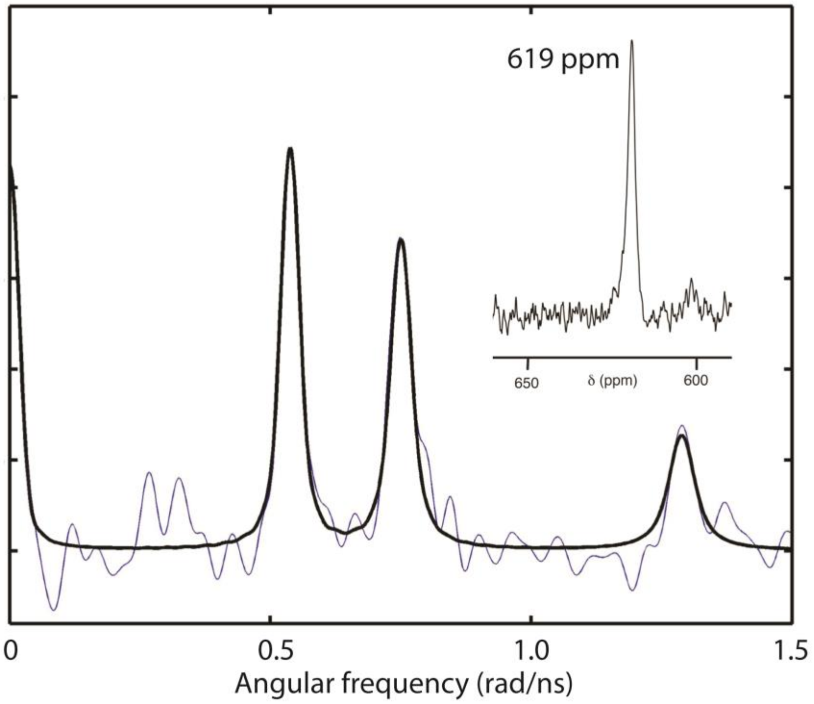


Figure 2.8. Fourier transformed 111mCd PAC spectroscopic data (blue: experimental data; black: fit) for the Cd(II) complex of TZ1C2. Inset: 113Cd NMR spectrum of the Cd(II) complex of TZ1C2 trimer.

2.2.6 Non-denaturing mass spectrometry

Non-denaturing mass spectrometry (NDMS) was also employed to verify the binding stoichiometry of Cd(II) to **TZ1C2** peptide. The largest population of species corresponds to the adduct with two Cd(II) ions per trimer, consistent with CD titration experiments.

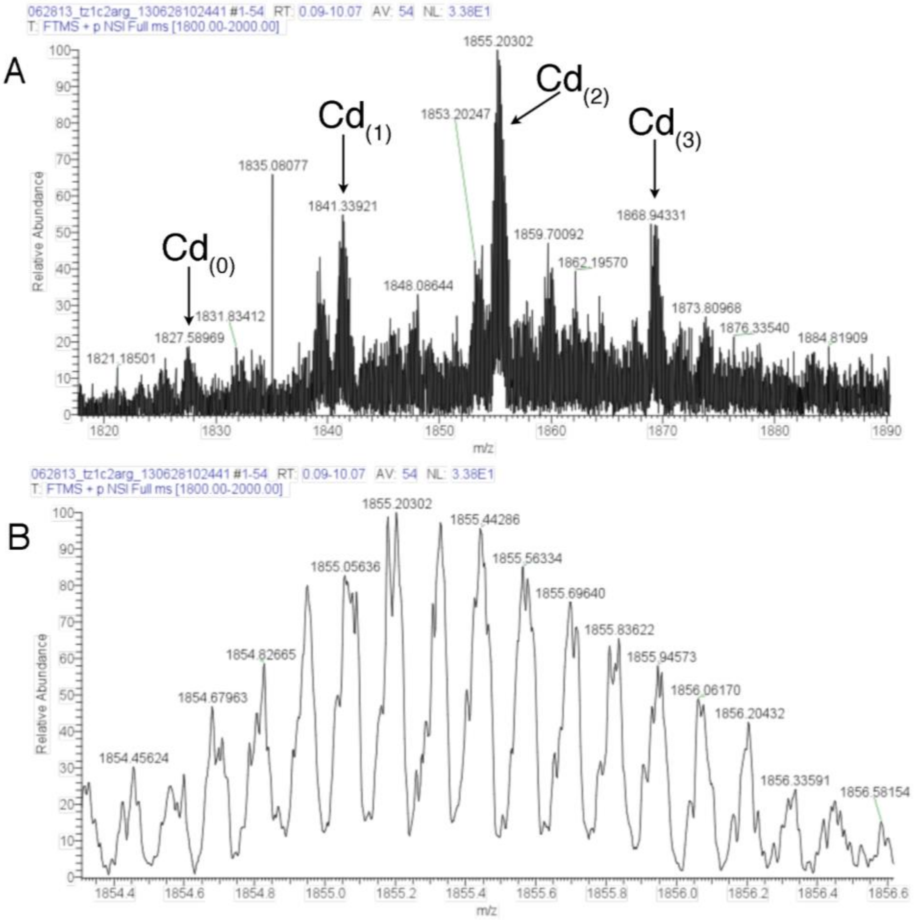


Figure 2.9. NDMS of TZ1C2 complex with Cd(II). A: MS of the TZ1C2 trimer in the +8 charge state. Arrows indicate the positions of adducts of the TZ1C2 trimer with 0, 1, 2, or 3 ^{113}Cd(II) ions. The largest population of species corresponds to the adduct with two Cd(II) ions per trimer. B. Expansion of the region of the mass spectrum corresponding to the di-cadmium adduct of the TZ1C2 trimer in the +8 charge state.

2.3 Summary

These data demonstrate that metal ion coordination can reversibly control chain registry and, consequently, the assembly state of a designed nanomaterial. A similar phenomenon has been observed adventitiously for β-sheet assemblies, [24-26] in which pH-dependent changes in strand registry can trigger reversible transitions between fibril and nanotube structures.[24, 25] In addition, experimental evidence suggests that a registry shift between helices in an anti-parallel coiled-coil dimer within the cytoskeletal motor protein dynein may underlie the mechanism of locomotion.[27] Thus, control of chain alignment represents an attractive strategy for the design of dynamically reconfigurable nano-scale materials.

2.4 Materials and methods

Materials. Chemical reagents were purchased from Fisher Scientific, Inc. (Pittsburgh, PA) or Sigma Chemical Co. (St. Louis, MO), unless otherwise noted. Isotopically enriched $^{113}CdCl_2$ (enrichment level: 93.35%) was obtained from Cambridge Isotope Laboratories, Inc. (Andover, MA). Fmoc-amino acid derivatives for solid-phase synthesis were purchased from AnaSpec Inc. (Fremont, CA).

Peptide Synthesis. Peptide synthesis was performed using microwave-assistance on a CEM Liberty solid-phase peptide synthesis instrument. Peptides were synthesized on a 4-(hydroxymethyl)phenoxyacetamido-methyl]-PEG-PS resin from Applied Biosystems, Inc. (Foster City, CA). Standard Fmoc protection chemistry was employed with coupling cycles based on HBTU/DIEA-mediated activation protocols and base-induced deprotection (20% piperidine in DMF with 0.1 M HOBt) of the Fmoc group. The peptides were purified via RP-HPLC on a C18 column with a gradient of water–acetonitrile (0.1% trifluoroacetic acid). The purity was assessed to be above 95% by analytical HPLC. Peptide mass was confirmed using electro-spray ionization mass spectrometry. The peptides were lyophilized, sealed, and stored at -20 ℃.

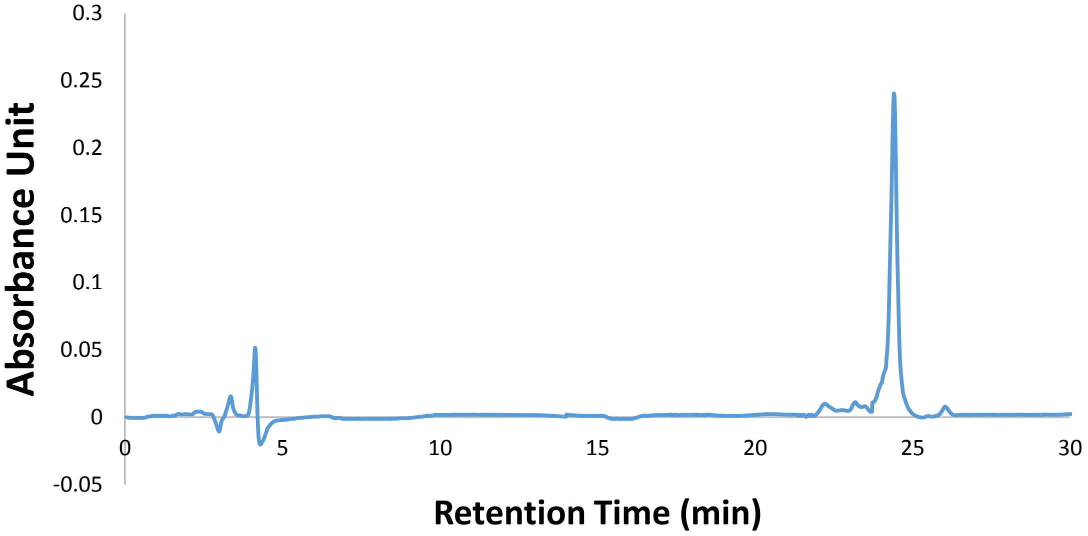
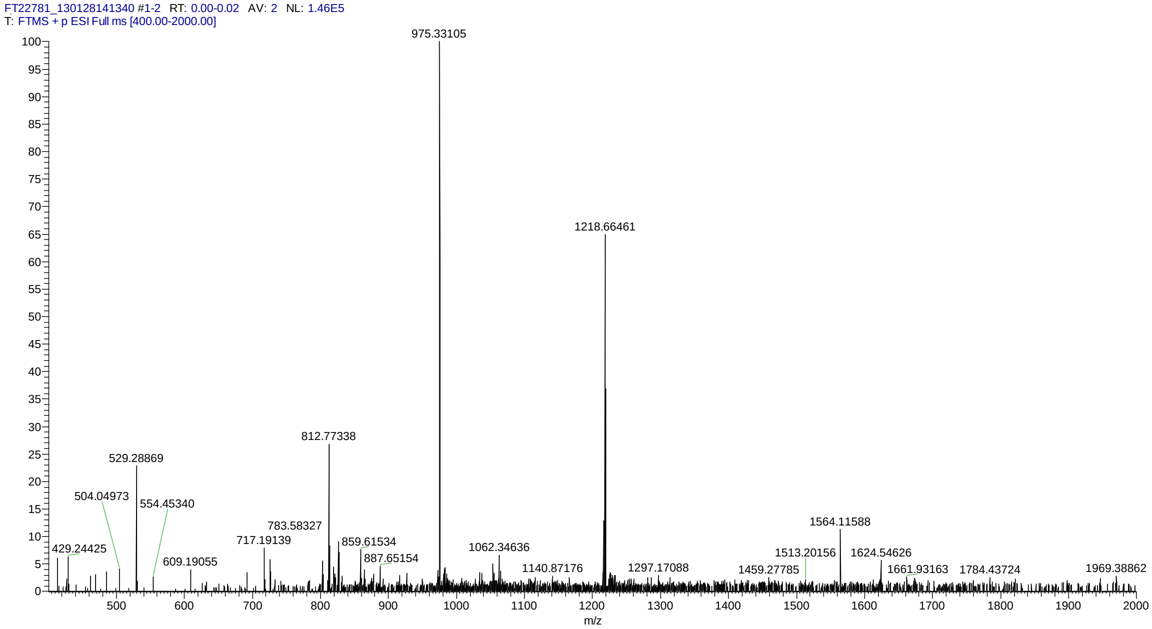


Figure 2.10. Analytical HPLC trace showing the purity of peptide TZ1C2.



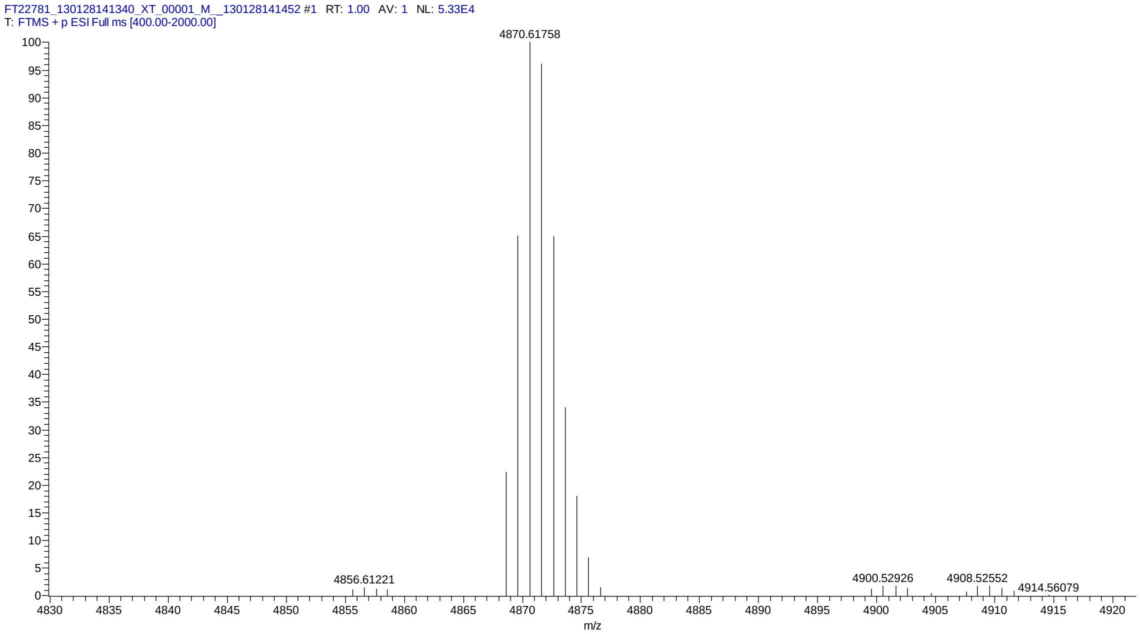


Figure 2.11. ESI-mass spectrum (upper) and the corresponding spectral deconvolution for purified peptide TZ1C2.

Circular Dichroism Spectropolarimetry. CD spectra were recorded on a Jasco J-810 CD spectropolarimeter in 0.10 mm quartz cells (Hellma Analytics) at a concentration of 100 μM of **TZ1C2** in 10 mM TAPS buffer pH 8.5 with 100 mM NaCl. Spectra were recorded from 190 to 260 nm at a scanning rate of 100 nm/min and a resolution of 0.5 nm. The peptide concentration was determined spectroscopically by measuring the absorbance at 280 nm (A_{280}).[1] For peptides containing Tyr, Trp or Cys residues, the peptide concentration can be calculated from the following equation:

$$\mathrm{MW}\times {}^{A_{280}}/_{c} = 1280n_Y + 5690n_W + 120n_C$$

in which c is the concentration of peptide in mg/mL, and n_Y, n_W, and n_C are the numbers of tyrosine, tryptophan and cystine residues, respectively, in the peptide sequence. As

TZ1C2 contains only a single tyrosine residue per molecule, then c = MW×A_{280}/1280. To eliminate error in determination of absorbance that could arise as a result of UV light scattering due to peptide self-assembly, aqueous solutions of peptide were mixed with 6 M guanidinium chloride in 1:9 v/v ratio to completely denature the sample prior to performing the absorbance measurements.

Linear Dichroism Spectroscopy. Linear Dichroism was recorded on JASCO J-810 Circular Dichroism Spectropolarimeter using a microvolume cuvette with a path length of 50 µm[2] at a concentration of 100 µM of **TZ1C2** in 10 mM TAPS buffer pH 8.5 with 100 mM NaCl. The background scattering was obtained by measuring the spectrum at 0 rpm rotation speed. The LD spectrum was measured after 15 minutes rotation at a speed of 3000 rpm to establish Couette flow.

Electron Microscopy. TEM specimens were prepared from aqueous **TZ1C2** solution in 10 mM TAPS buffer with 100 mM NaCl pH 8.5. The samples were deposited onto 200 mesh carbon coated copper grids (Electron Microscopy Sciences). After a 30 seconds incubation period, excess liquid was wicked away and the specimens were stained with 1% methylamine tungstate (Ted Pella, Inc.). Excess stain was wicked away after incubation on the grid for 1 min. The sample grids were dried under vacuum and stored in a desiccator. TEM measurements were acquired on a Hitachi H-7500 transmission electron microscope at an accelerating voltage of 75 kV. The micrographs were recorded at a magnification of 200,000x using a Gatan CCD digital camera.

Analytical Ultracentrifugation. In order to identify the oligomerization state of **TZ1C2** and to characterize mass and shape distributions of the oligomerization products, sedimentation velocity (SV) experiments were performed. SV experiments characterize the solution behavior of macromolecules and can identify dynamic processes such as mass-action driven reversible associations by observing the sedimentation and diffusion behavior of all species in the system simultaneously. The experimental data are analyzed with the van Holde-Weischet analysis,[4] or modeled with solutions of the Lamm equation,[5, 6] which are optimized by 2-dimensional spectrum analysis (2DSA). [7, 8] 2DSA solutions were further refined using genetic algorithm (GA) analysis[9] to obtain a parsimoniously regularized solution.[10] The GA solution is then analyzed by Monte Carlo analysis to determine confidence limits for the determined parameters.[11] The calculations are computationally intensive and are carried out on high-performance computing platforms.[12] The van Holde-Weischet approach provides diffusion-corrected sedimentation coefficient distributions from which an assessment of heterogeneity and reversible self-association can be obtained. The 2DSA and GA analyses further provide molecular weight and anisotropy for any species found in the mixture.

All sedimentation experiments were performed with a Beckman Optima XL-I at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center at San Antonio. Sedimentation velocity data were analyzed with the UltraScan-III software[20] version 2.0.[13] All calculations were performed on the Lonestar and Stampede clusters at the Texas Advanced Computing Center at the University of Texas at Austin, or on the Alamo cluster at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio. **TZ1C2** was measured in the presence of Cd(II) at a loading concentration of 0.24 OD and 0.73 OD at 280 nm, and at 0.22 OD 280 nm in the absence of Cd(II). All measurements were made in a buffer containing 10 mM TAPS and 100 mM NaCl. The experimental data were collected in intensity mode at 20 ℃, and at 60 krpm, using standard epon 2-channel centerpieces. Hydrodynamic corrections for buffer density, viscosity and partial specific volume were made according to methods outlined in Laue, et al.,[14] and as implemented in UltraScan.[20] The partial specific volume of **TZ1C2** was determined to be 0.743 cm^3/g. All data were first analyzed by 2-dimensional spectrum analysis (2DSA)[7, 8] with simultaneous removal of time- and radial-invariant noise and fitting of the meniscus position, followed by van Holde-Weischet analysis and genetic algorithm (GA) refinement,[10] followed by Monte Carlo analysis (MC).[11] We acknowledge the support of the San Antonio Cancer Institute grant P30 CA054174 for the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center at San Antonio.

^{113}Cd NMR Spectroscopy. Lyophilized **TZ1C2** peptide (20 mg) was dissolved in 400 μL of 15% D_2O/H_2O purged with N_2 in advance to remove dissolved oxygen. The pH was adjusted to 8.5 using 0.1 M NaOH in 15% D_2O/H_2O. The peptide concentration was determined spectrophotometrically by measuring the absorbance at 280 nm (see above for detail). After determination of peptide concentration, two equivalents of isotopically enriched $^{113}CdCl_2$ were added based on peptide concentration. The ^{113}Cd NMR spectrum was collected at room temperature on a Varian Inova 600 spectrometer (113.08 MHz for ^{113}Cd) equipped with a 5 mm broadband probe. A spectral width of 844 ppm (112,360 Hz) was sampled using a 6.0 μs 90° pulse and 0.2 second acquisition time with 0.05 second delay between scans. Chemical shifts were recorded in parts per million relative to 0.1 M $Cd(ClO_4)_2$ in D_2O.

Perturbed Angular Correlation Spectroscopy. Methods for sample preparation and data acquisition followed the general procedures described in Matzapetakis, et al.[16] Perturbed angular correlation (PAC) experiments were performed with a setup using six detectors at a temperature of 1 ± 2 ℃ that was controlled using a Peltier element. The radioactive cadmium was produced on the day of the experiment at the University Hospital cyclotron in Copenhagen. The ^{111m}Cd solution (10−40 μL) was mixed with nonradioactive cadmium acetate and TRIS buffer. The **TZ1C2** peptide was then added (dissolved in 10 mM TAPS buffer, pH 8.5, 100 mM NaCl), and the sample was left to equilibrate for 10 min to allow for metal ion binding. All buffers were purged with Ar and treated so as to lower metal contamination. Finally, sucrose was added to produce in order to further reduce the Brownian tumbling of the molecules. Final sample conditions were 300 μM peptide, 1.9/3 eq. Cd(II), 10 mM TAPS, 100 mM NaCl, (final pH value of 9.1 at 1 ℃), and 55 % w/w sucrose. The experiments were carried out at 1 ℃. The time resolution was 0.860 ns, and time per channel was 0.562 ns. All fits were carried out with 300 data points, disregarding the first 5 points due to systematic errors. The PAC spectroscopic results are presented in Table 2.1.

Non-denaturing Mass Spectrometry. A solution (500 μL) of peptide **TZ1C2** (100 μM) was incubated with tris(2-carboxyethyl)phosphine (TCEP, 500 μM) for 1 h. The sample was dialyzed for two days against 1200 mL of 500 mM ammonium acetate, pH 8.5. The sample was removed from dialysis and incubated with a one to one molar ratio of $^{113}CdCl_2$ for 12 h at 4 ℃. Mass spectrometry experiments were performed on a Thermo LTQ-FTMS with a 7 T magnet. The ionization was performed using a nanospray source with PicoTip Emitters (New Objective). The voltage on the PicoTip emitter was maintained at 1.5 kV. The capillary temperature was approximately 40 ℃. The capillary voltage was 10 V and the tube lens was 100 V. The automatic gain control (AGC) was set to 1.00 x 10^6. The maximum inject time was 5000 µs. The spectrum was taken at a resolution of 500,000, and the mass range was set to 1800 to 2000 m/z. The spectrum was signal averaged for ten minutes (54 scans).

2.5 References:

1. Hauser, C.A. and S. Zhang, *Designer self-assembling peptide nanofiber biological materials.* Chemical Society Reviews, 2010. **39**(8): p. 2780-2790.

2. Woolfson, D.N. and Z.N. Mahmoud, *More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials.* Chemical Society Reviews, 2010. **39**(9): p. 3464-3479.

3. Tørring, T., et al., *DNA origami: a quantum leap for self-assembly of complex structures.* Chemical Society Reviews, 2011. **40**(12): p. 5636-5646.

4. King, N.P., et al., *Computational design of self-assembling protein nanomaterials with atomic level accuracy.* Science, 2012. **336**(6085): p. 1171-1174.

5. Lanci, C.J., et al., *Computational design of a protein crystal.* Proceedings of the National Academy of Sciences, 2012. **109**(19): p. 7304-7309.

6. Brodin, J.D., et al., *Metal-directed, chemically tunable assembly of one-, two-and three-dimensional crystalline protein arrays.* Nature chemistry, 2012. **4**(5): p. 375-382.

7. Zimenkov, Y., et al., *Rational design of a reversible pH-responsive switch for peptide self-assembly.* J Am Chem Soc, 2006. **128**(21): p. 6770-1.

8. Dublin, S.N. and V.P. Conticello, *Design of a selective metal ion switch for self-assembly of peptide-based fibrils.* J Am Chem Soc, 2008. **130**(1): p. 49-51.

9. Peacock, A.F., O. Iranzo, and V.L. Pecoraro, *Harnessing natures ability to control metal ion coordination geometry using de novo designed peptides.* Dalton Transactions, 2009(13): p. 2271-2280.

10. Schnarr, N.A. and A.J. Kennan, *Peptide tic-tac-toe: heterotrimeric coiled-coil specificity from steric matching of multiple hydrophobic side chains.* J Am Chem Soc, 2002. **124**(33): p. 9779-83.

11. Potekhin, S., et al., *De novo design of fibrils made of short α-helical coiled coil peptides.* Chemistry & biology, 2001. **8**(11): p. 1025-1032.

12. Papapostolou, D., et al., *Engineering nanoscale order into a designed protein fiber.* Proceedings of the National Academy of Sciences, 2007. **104**(26): p. 10853-10858.

13. Matzapetakis, M., et al., *Comparison of the binding of cadmium(II), mercury(II), and arsenic(III) to the de novo designed peptides TRI L12C and TRI L16C.* J Am Chem Soc, 2002. **124**(27): p. 8042-54.

14. Li, X., et al., *Soft metal ions, Cd (II) and Hg (II), induce triple‐stranded α‐helical assembly and folding of a de novo designed peptide in their trigonal geometries.* Protein Science, 2000. **9**(7): p. 1327-1333.

15. Kharenko, O.A. and M.Y. Ogawa, *Metal-induced folding of a designed metalloprotein.* Journal of inorganic biochemistry, 2004. **98**(11): p. 1971-1974.

16. Matzapetakis, M. and V.L. Pecoraro, *Site-selective metal binding by designed alpha-helical peptides.* J Am Chem Soc, 2005. **127**(51): p. 18229-33.

17. Hemmingsen, L., et al., *First principle calculations of (113)Cd chemical shifts for proteins and model systems.* J Biol Inorg Chem, 2004. **9**(5): p. 591-9.

18. Hemmingsen, L., K.N. Sas, and E. Danielsen, *Biological applications of perturbed angular correlations of gamma-ray spectroscopy.* Chem Rev, 2004. **104**(9): p. 4027-62.

19. Iranzo, O., et al., *The correlation of 113Cd NMR and 111mCd PAC spectroscopies provides a powerful approach for the characterization of the structure of Cd(II)-substituted Zn(II) proteins.* Chemistry, 2009. **15**(15): p. 3761-72.

20. Bulheller, B.M., et al., *Flow linear dichroism of some prototypical proteins.* Journal of the American Chemical Society, 2009. **131**(37): p. 13305-13314.

21. Andersen, O., *Chelation of cadmium.* Environmental health perspectives, 1984. **54**: p. 249.

22. Demeler, B. and E. Brookes, *Monte Carlo analysis of sedimentation experiments.* Colloid and Polymer Science, 2008. **286**(2): p. 129-137.

23. Demeler, B. and K.E. van Holde, *Sedimentation velocity analysis of highly heterogeneous systems.* Anal Biochem, 2004. **335**(2): p. 279-88.

24. Liang, Y., et al., *Cross-strand pairing and amyloid assembly.* Biochemistry, 2008. **47**(38): p. 10018-10026.

25. Mehta, A.K., et al., *Facial symmetry in protein self-assembly.* J Am Chem Soc, 2008. **130**(30): p. 9829-35.

26. Petkova, A.T., et al., *Solid state NMR reveals a pH-dependent antiparallel beta-sheet registry in fibrils formed by a beta-amyloid peptide.* J Mol Biol, 2004. **335**(1): p. 247-60.

27. Kon, T., et al., *Helix sliding in the stalk coiled coil of dynein couples ATPase and microtubule binding.* Nature structural & molecular biology, 2009. **16**(3): p. 325-333.

Chapter III: Rational Design of Helical Nanotubes from Self-assembly of Coiled-coil Lock Washers

3.1 Introduction

Structurally defined materials on the nanometer length-scale have been historically the most challenging to rationally construct and the most difficult to structurally analyze [1, 2]. Sequence-defined polypeptides represent attractive design elements for construction of these types of nano-scale materials. Sequence-structure correlations from native proteins can be employed for the design of ordered assemblies in which functional properties can be controlled through the progression of structural hierarchy encoded at the molecular level. Furthermore, proteins are amenable to preparative scale synthesis and display a rich portfolio of structure and function in the native state. However, the diversity of sequence space and the current limitations of theoretical approaches to reliably define the relationship between sequence and supramolecular structure present a significant challenge to the *de novo* design of novel materials architectures.

Simple protein motifs, such as α-helical coiled-coils [3-15], β-strands [16-25], β-hairpins [26-29], and collagen triple helices [30-39], have been employed as structural elements for the *de novo* design of fibrillar protein-based assemblies with a notable degree of success. However, thus far, it has proven difficult to reliably predict higher order structure and almost impossible to specify function to a similar extent as observed for native protein assemblies. This phenomenon may be attributed, at least in part, to differences in the self-assembly mechanism between synthetic and native protein assemblies. For the synthetic systems, self-assembly usually occurs commensurately with folding of the peptide sequence into the target structure. This strategy requires the ability to define structure simultaneously over multiple length-scales. In contrast, the structural subunits of native protein fibrils usually correspond to folded protein domains or structurally defined oligomeric assemblies that self-associate through the recognition between structurally complementary interfaces, usually reversibly and controllably, in response to environmental cues.

An alternative strategy for the creation of synthetic protein assemblies may be envisioned that hews more closely to the native mechanism of self-assembly [40-42]. Since the structures of folded protein domains can be determined to high resolution using single crystal X-ray diffraction or multi-dimensional NMR spectroscopic analyses, one can potentially introduce structurally informed mutations at surface-exposed positions within these sequences to promote specific modes of self-association that would permit retention of native function and/or the creation of novel function within the self-assembled material. We employ this recoding strategy to redesign the sequence of the peptide **GCN4-pAA** [43] to promote specific recognition between interfaces such that linear propagation occurs to form uniaxially oriented tubular assemblies [44-51].

The crystal structure of **GCN4-pAA** (PDB ID: 2HY6), a *de novo* designed peptide derived from the leucine zipper region of the *S. cerevisiae* transcription factor GCN4,[52] displays a discrete 7-helix bundle structure that comprises the largest, freely standing coiled-coil oligomer that has been structurally characterized thus far (Figure 3.1). The heptameric assembly defines a continuous central channel with an internal diameter of approximately 7 Å.[43] Computational analysis using the program CASTp [53] indicated

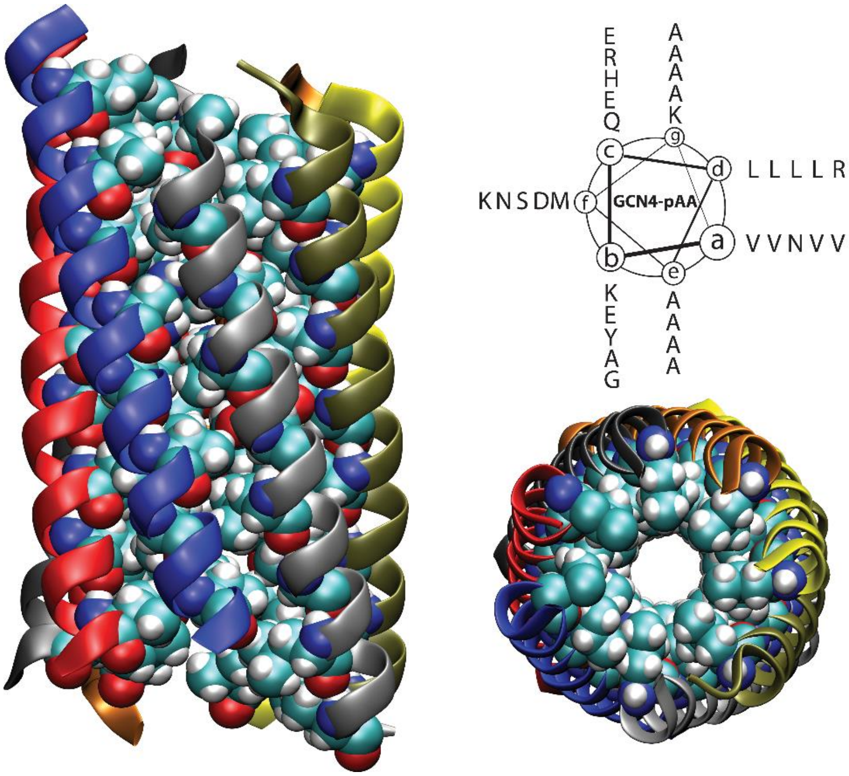


Figure 3.1. Crystal structure (PDB ID: 2HY6) of the 7-helix bundle resulting from self-association of the peptide GCN4-pAA. Full-length and top-down views are shown on the left and in the lower right, highlighting side-chain atoms of residues pointing towards the tube interior. Individual helices (*A*-*G*) are colored separately. The displaced edge of the structure occurs at the interface between the first (*A*, blue) and seventh (*G*, gray) helices. The helical wheel projection of the amino acid sequence of GCN4-pAA in 7_2-supercoil helix space (upper right).

the presence of an internal void volume of 1,880 Å^3 associated with the central channel. The presence of several hexane-1, 6-diol molecules in the central channel of the heptamer crystal structure indicates that it is capable of accommodating appropriately shaped small-molecules within the cavity. In addition, the **GCN4-pAA** bundle structure displays a single residue shift in registry between adjacent helices, which resulted in an overall displacement of seven residues (i.e., one coiled-coil heptad) at the interface between helices *A* and *G* of the bundle structure (Figure 3.1). The 7-helix bundle of **GCN4-pAA** resembles a lock washer, in which the displaced edges at the interface between helices *A* and *G* interface provide additional surface area for complementary interactions between the coiled-coil protomers. In contrast, most coiled-coil structures do not display a commensurate shift in helix registry.

The program MSMS [54] was employed to calculate the solvent-accessible surface area (SASA) associated with the exposed edges at the *N*-terminal and *C*-terminal surfaces of the **GCN4-pAA** seven-helix bundle. SASA values of 2,411 Å^2 and 1,310 Å^2 were determined for the upper (*N*-terminal) surface and exposed edge of helix *A* and for the lower (*C*-terminal) surface and exposed edge of helix *G*, respectively. The SASA differences between the upper and lower surfaces may be attributed to differences in amino acid composition and are, therefore, amenable to modification. The surface area that would be buried through formation of a helical interface between stacked bundles was estimated by comparison to the remnant solvent accessible surface area calculated for structurally homologous positions at internal sites within the structure of **GCN4-pAA** (see *Methods*). Subtracting these values from the total SASA of the exposed upper and lower interfaces of **GCN4-pAA**, one would expect approximately 1,100-1,200 Å^2 and

900 Å^2 of surface area, respectively, to be buried upon self-association between structurally complementary surfaces that establishes a continuous coiled-coil structure at the interface. Bennett, et al.,[55] proposed a minimum interaction surface area of 856 Å^2 per interface for the formation of stable fibrillar assemblies, based on the structural analysis of interfaces in protein crystals postulated by Ponstingl, et al.[56] The interfaces of **GCN4-pAA** seem more than sufficient to meet the minimal criteria for the formation of stable end-to-end association, which would represent a necessary pre-requisite for the formation of structurally defined assemblies.

Despite these observations, heptameric coiled-coil assemblies of **GCN4-pAA** do not appear to interact significantly in solution or the crystalline state.[43] No axial interactions were observed between 7-helix bundles at cut-off distances $\leq$ 9.0 Å in the crystal structure of **GCN4-pAA**. This distance is equivalent to the axial rise of six amino acid residues within an α-helical conformation, which implies that a significant gap exists between co-axially oriented 7-helix bundles in the crystal structure of **GCN4-pAA**. Moreover, sedimentation equilibrium analytical ultracentrifugation of solutions of **GCN4-pAA** indicated the clean formation of a heptamer within the experimentally measured concentration range (30-300 μM) in ABS buffer (50 mM acetate, pH 5.2, 150 mM NaCl) at 20 ℃. [43]

Several features of the **GCN4-pAA** sequence may act to frustrate interfacial interactions between helical bundles and, therefore, may be subject to improvement through rational design. Notably, an arginine residue occurs at a *d*-position within the *C*-terminal heptad of **GCN4-pAA** (Figure 3.1), which breaks the hydrophobic periodicity of the repeat sequence. Charged gatekeeper residues at protein-protein interfaces have been demonstrated to act as negative design elements that prevent self-association in β-sheet proteins.[57-60] Similarly, the positive charge of the arginine residue should have a direct influence on the Coulombic attraction between the uncapped *N*- and *C*-termini of **GCN4-pAA**. Thus, the disruption of the hydrophobic core and the introduction of repulsive interfacial interactions would be expected to inhibit self-association between heptameric assemblies derived from **GCN4-pAA**. We hypothesized that if the sequence of **GCN4-pAA** were modified to remove these inhibitory structural features, then end-to-end association should occur between complementary surfaces of the lock washer structures that would result in formation of a high aspect-ratio fibril with a continuous channel throughout the assembly.

3.2 Design of 7HSAP1 Sequence

Peptide **7HSAP1** was designed to test the hypothesis described above (Figure 3.2). The design of **7HSAP1** reflects several critical structural considerations. The thermodynamic driving force for self-assembly of the nanotubes should be a combination of electrostatic attraction between the oppositely charged surfaces and burial of solvent accessible surface area at structurally complementary interfaces between lock washer subunits (i.e., 7-helix bundles). The hydrophobic *a*-*d*-*e*-*g* interface [43] of the original **GCN4-pAA** peptide was maintained as it was reasoned that this feature was necessary to ensure formation of the seven-helix bundle. However, the single arginine residue at a *d*-position within the sequence of **GCN4-pAA** was replaced with a canonical leucine residue in order to maintain a continuous hydrophobic interface and remove repulsive electrostatic interactions that might arise between bundles. The *N*- and *C*-termini of the peptide were left uncapped to promote head-to-tail stacking interactions through electrostatic attraction between oppositely charged interfaces of the lock washer structures. In addition, electrostatic interactions were maximized between the spatially proximal *b*- and *c*-positions on adjacent helices (Figure 3.2) with glutamate and lysine, respectively, in order to strengthen association between individual helices within the 7-helix bundle. The peptide sequence was formally split between the *b*- and *c*-positions of the heptad repeat in order to minimally perturb the *a*-*d*-*e*-*g* hydrophobic core. To preserve continuity of the coiled-coil structural periodicity upon axial stacking of the lock washer assemblies, the sequence of **7HSAP1** was based on an integral number of five heptad repeats. Finally, two arginine residues were placed at *f*-positions within the **7HSAP1** sequence. The resultant positive charge of (+2) per peptide should hinder lateral association between bundles and promote dispersion of the assemblies.

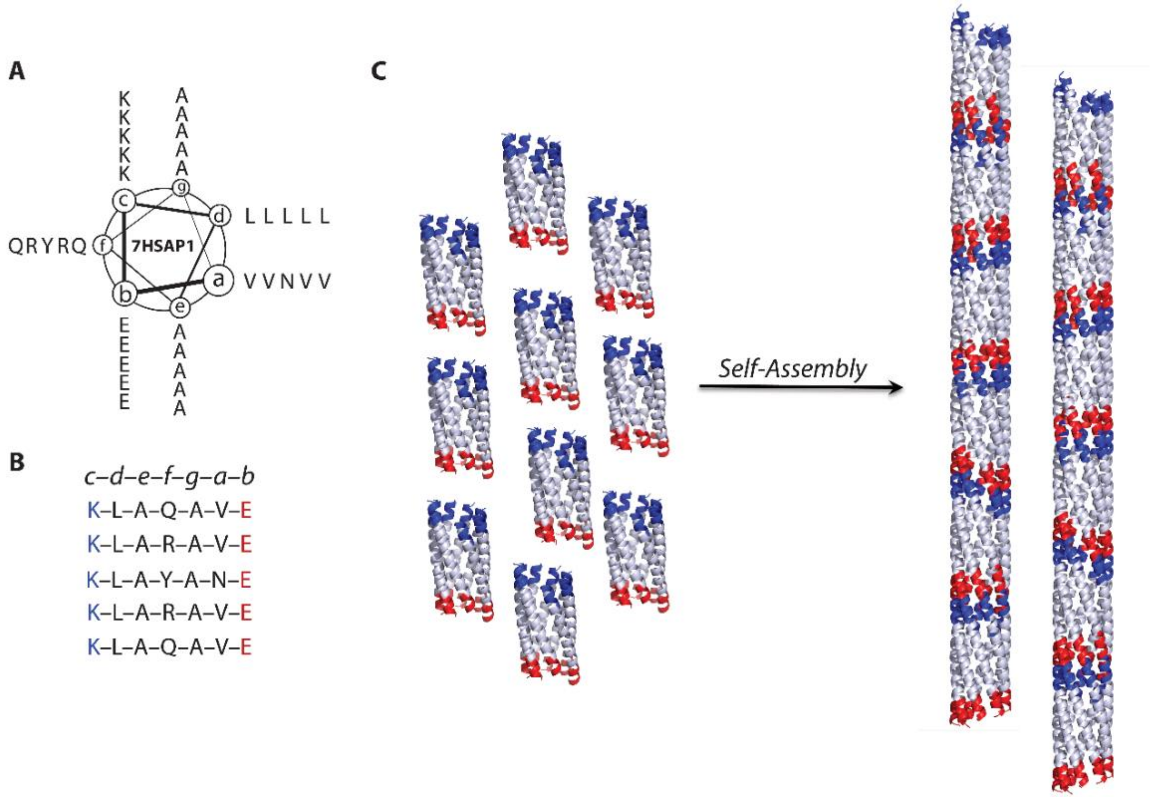


Figure 3.2. Left: Helical wheel (A) and linear depiction (B) of the sequence of peptide 7HSAP1. Right: Schematic representation of the proposed model for self-assembly (C) of lock washer structures derived from the 7-helix bundle of peptide 7HSAP1 into helical nanotubes. The blue and red surfaces represent the positively charged (*N*-terminal) heptads and negatively charged (*C*-terminal) heptads at the interfaces between 7-helix bundle subunits of 7HSAP1 assemblies.

A range of analytical methods was employed to evaluate the structure of **7HSAP1** over multiple length-scales in solution and the solid-state.

3.3 Results and Discussion

3.3.1 Circular Dichroism and Flow Linear Dichroism

The conformation of **7HSAP1** peptide was initially interrogated with circular dichroism (CD) spectropolarimetry. The CD signal is strongly α-helical with MRE values that exceed those of the control peptide **GCN4-pAA** [43] under identical conditions of buffer concentration and pH (100 μM peptide in 10 mM MES buffer pH 6.0) and approach 100% helicity (Figure 3.3). Under these conditions, peptide **7HSAP1** did not display a melting transition between 4 ℃ and 95 ℃, which is consistent with the presence of an extended hydrophobic core and is commensurate with earlier CD melting studies of **GCN4-pAA** under similar conditions (Figure 3.4).

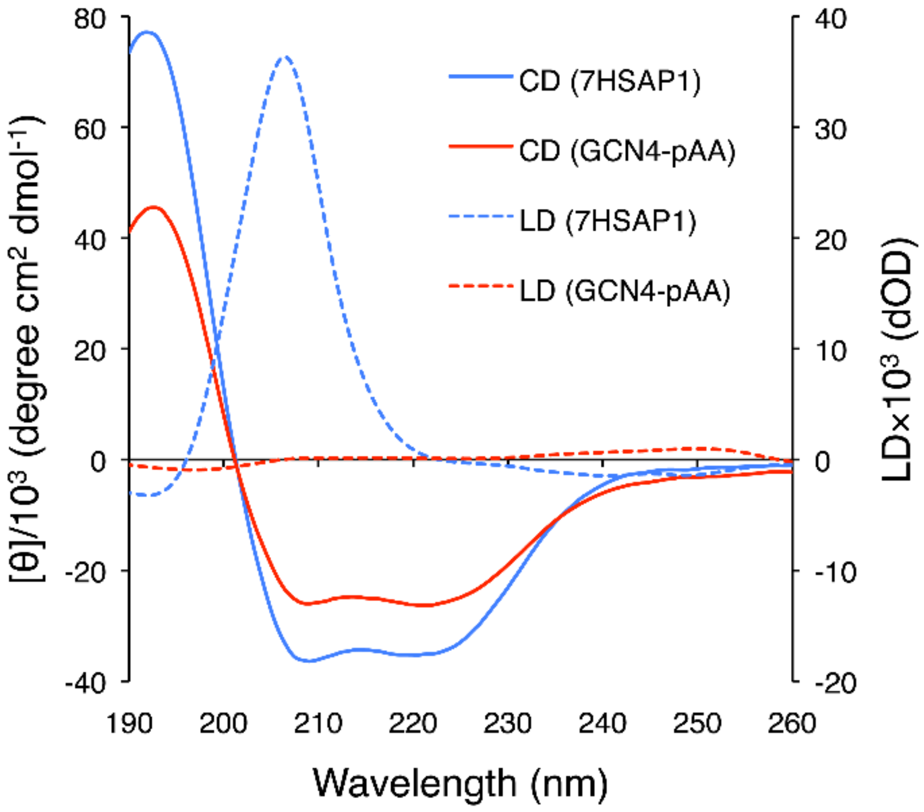


Figure 3.3. Circular dichroism and flow linear dichroism spectra of peptides GCN4-pAA and 7HSAP1 (100 µM) in 10 mM MES buffer, pH 6.0. Flow linear dichroism spectra were acquired under a Couette flow of 3,000 rpm.

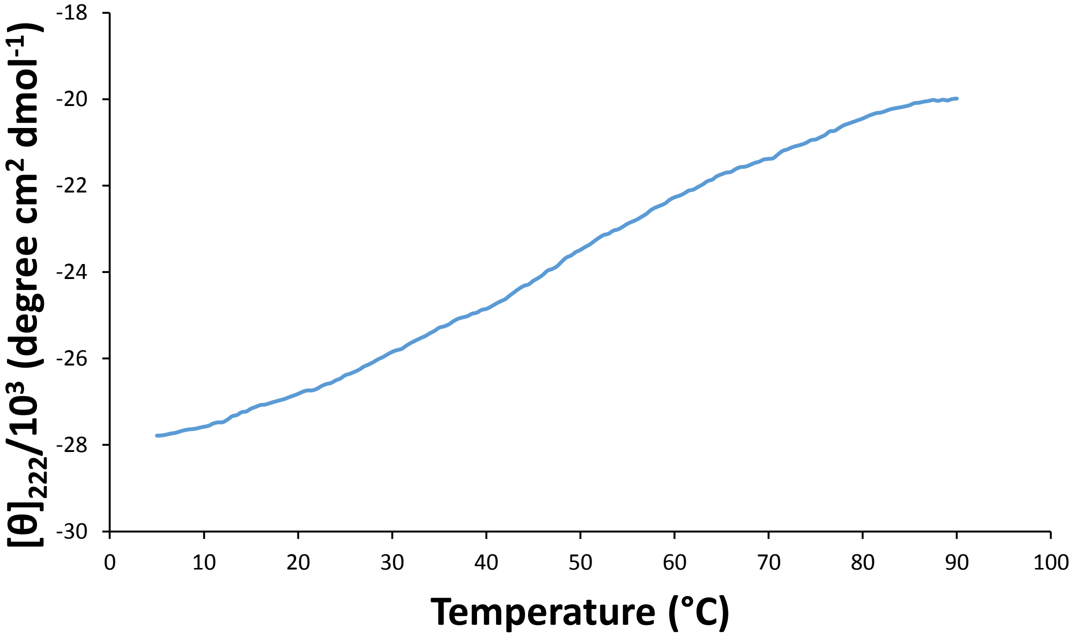


Figure 3.4. CD melting curve of 7HSAP1 (93 µM) in 10 mM MES buffer, pH 6.0

Flow linear dichroism [61, 62] was employed under conditions of similar concentration to assess the degree of anisotropy that developed within the peptide samples under a flow alignment (3,000 rpm) in a Couette cell (Figure 3.3). Solutions of **7HSAP1** displayed an LD spectrum with a strong positive signal at 207 nm, which provides evidence that the amide bond chromophores are aligned in parallel to the flow direction and therefore along the long-axis of the assembly. In addition, the negative feature at 190 nm would also be expected for a helix oriented in the direction of flow, since it arises from a transition perpendicular to the helical axis. Similar LD spectroscopic results have been observed for solutions of a synthetic coiled-coil fiber and the dimeric coiled-coil protein tropomyosin that had been aligned under Couette flow [63].

The LD and CD results are consistent with the hypothesis that **7HSAP1** assembles in solution as an α-helical bundle to form extended arrays that stack along the direction of the super-helical axis (Figure 3.2). In contrast, solutions of **GCN4-pAA** did not display a flow LD spectrum that differed significantly from the baseline (Figure 3.3). This observation suggests that solutions of **GCN4-pAA** could not be aligned in the Couette cell, presumably as a consequence of their inability to form stable, extended arrays. Changes in solution conditions, such as lowered pH (50 mM acetate, pH 5.2) or the addition of co-solvents (0.1% dimethylformamide), resulted in loss of the flow LD signal of **7HSAP1**. These LD results suggest that self-association between the structurally complementary interfaces of **7HSAP1** were subject to disruption under shear flow, even though the CD spectroscopic signature and fibrillar morphology (*vide infra*) were not perturbed under identical static conditions. The malleable nature of the interactions between subunits is consistent with the action of non-covalent forces (i.e., electrostatic, van der Waals, and hydrogen bonding interactions); the strength of which can be manipulated through changes in environmental conditions.

3.3.2 Electron Microscopy

TEM analysis (Figure 3.5A) of **7HSAP1** in MES buffer (10 mM, pH 6.0) confirmed the formation of fibrillar structures over a wide range of peptide concentration (at least 25 μM to 4.0 mM). Under conditions of conventional TEM analysis, the fibrils were observed to bundle together, occasionally affording highly ordered para-crystalline assemblies (Figure 3.5B) at higher concentrations (≥ 2 mM). In contrast, fibrillar structures were not observed in the TEM analysis of solutions of peptide **GCN4-pAA**, even at the higher concentration limit. Capping of **7HSAP1** as the *N*-acetyl and *C*-amide derivative significantly inhibits but does not abolish fibrillization of the peptide, as has been observed for other coiled-coil fibrils [4-8]. This observation suggests that **7HSAP1** has a stronger intrinsic propensity for self-assembly than **GCN4-pAA**, which may result from the introduction of energetically stabilizing interactions, and the removal of destabilizing interactions, in the initial peptide design (*vide supra*).

Cryo-TEM (Figure 3.5D) of solutions of **7HSAP1** indicated the presence of fibrils of circa 3 nm in diameter. STEM analysis of negatively stained specimens of **7HSAP1** confirmed the fibril diameter (3.0 ± 0.1 nm). A lesser population (≤ 10%) of thinner fibrils was observed with a diameter of circa 2.0 nm. The observed diameter of the major population of **7HSAP1** fibrils compared well with the diameter of 3.1 nm observed for the 7-helix bundle in the crystal structure of **GCN4-pAA**.[43] In addition, the lateral striations in the negatively stained, para-crystalline assemblies of **7HSAP1** were determined to have a spacing of approximately 5 nm (Figure 3.5B). This distance coincides with the length of the **7HSAP1** peptide projected onto the super-helical structural parameters determined from the seven-helix bundle structure of **GCN4-pAA** (super-helical rise/residue of 0.143 nm x 35 residues = 5.01 nm) [43]. Similar striations have been observed in the TEM analysis of negatively stained fibrils derived from self-assembly of coiled-coil dimers.[7, 8]

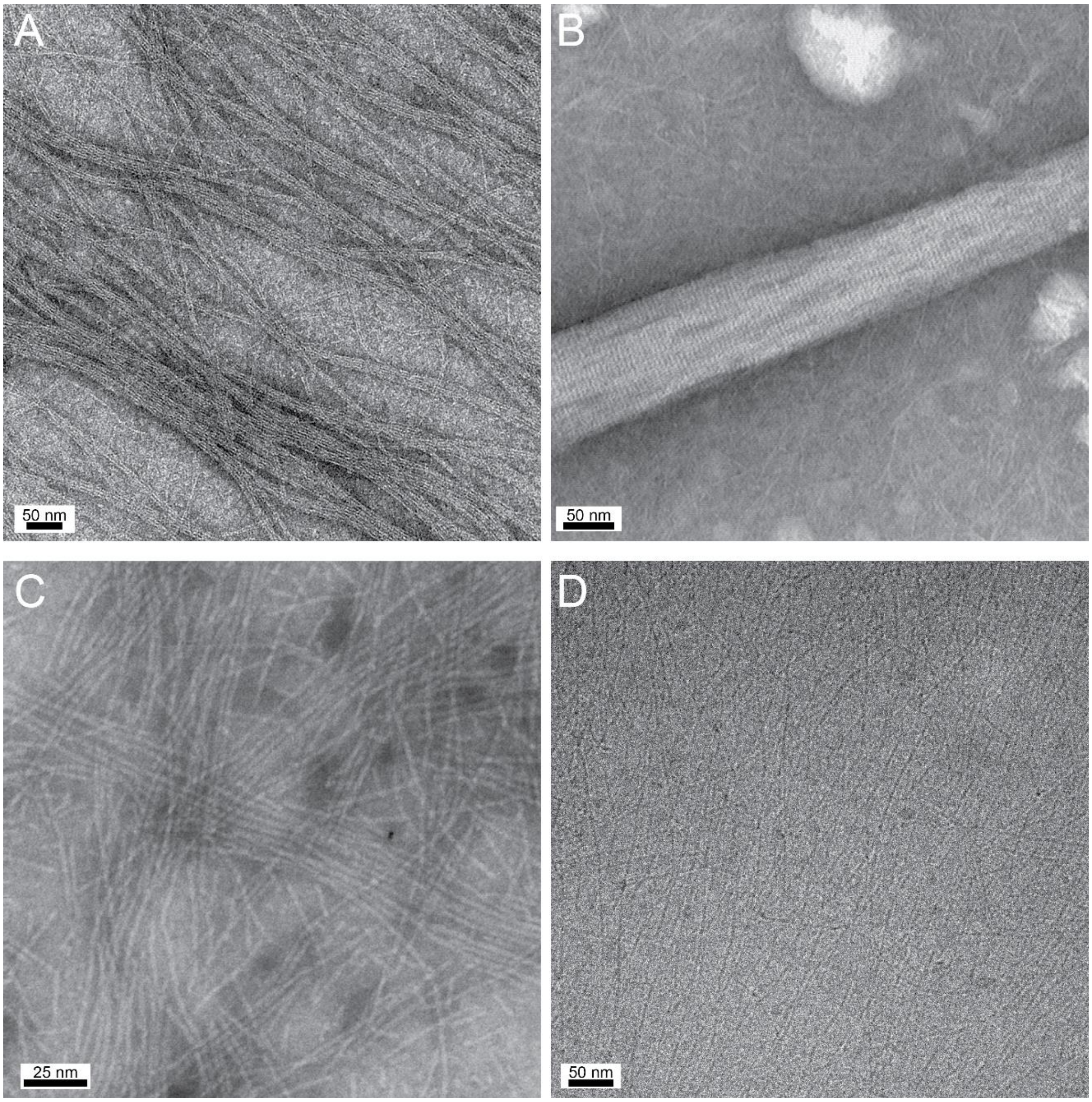


Figure 3.5. Electron microscopy analysis of 7HSAP1 assemblies. A. Conventional TEM of 7HSAP1 (2.2 mM) in MES buffer (10 mM, pH 6.0) stained with 2% mixture of 1:1 methylamine vanadate and methylamine tungstate. B. Conventional TEM of large para-crystalline assemblies derived from 7HSAP1 (2 mM) in MES buffer (10 mM, pH 6.0) stained with 1% uranyl acetate. C. STEM of 7HSAP1 (80 µM) in MES buffer (10 mM, pH 6.0) stained with 2% methylamine vanadate. D. Cryo-TEM of 7HSAP1 (130 µM) in MES buffer (10 mM, pH 6.0).

While the combined data suggested that the major structural feature was fibrils derived from self-assembly of 7-helix bundle structures, the STEM measurements indicated a smaller population of thinner fibrils (≤ 10%) within the specimens assembled at the lower concentration conditions (≤ 100 μM) employed for imaging (Figure 3.6A). Mass action effects could tilt the equilibrium toward fibril disassembly under these circumstances. The STEM measurements indicated the presence of breaks within the filaments at the more dilute concentration range (< 80 μM), which was consistent with this hypothesis. Solid-state ^{13}C CP/MAS NMR spectroscopy of fibers derived from the isotopically labeled peptide variant **7HSAP1*** (*vide infra*) indicated the absence of β-sheet conformation even under conditions of assembly at high peptide concentration (Figure 3.9).

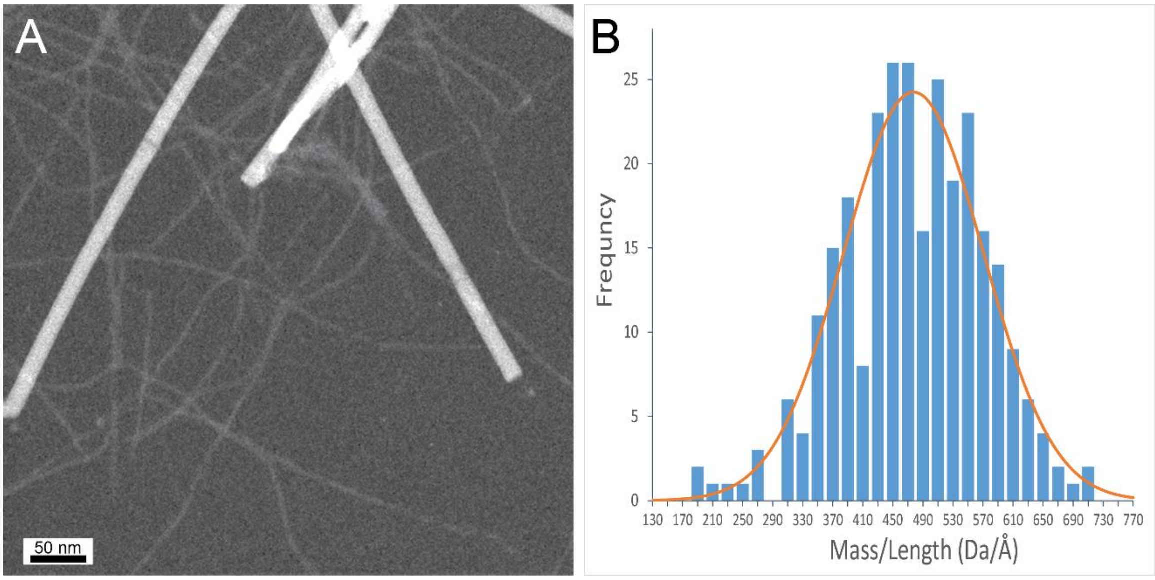


Figure 3.6. STEM mass-per-length measurements of 7HSAP1. A. Dark-field STEM of freeze-dried, unstained specimens of 7HSAP1 assemblies. The larger assemblies

correspond to the TMV calibrant. B. Histogram of mass-per-unit-length (MPL) measurements of the 7HSAP1 specimen (after normalization to the TMV standard).

Mass-per-length (MPL) measurements were performed on **7HSAP1** filaments (8 μM in 10 mM MES buffer pH 6.0) imaged using dark-field STEM of unstained, freeze-dried specimens (Figure 3.6A). Tobacco mosaic virus (TMV) particles were employed as an internal mass standard. The majority of the specimen corresponded to filaments of approximately 3 nm in diameter, as observed previously. A MPL value of 479 ± 93 Da/Å was determined for the 3 nm filaments, after normalization based on MPL measurements of the TMV mass standard (Figure 3.6B). A smaller population (circa 10% of the specimen) corresponded to thinner filaments with approximately half of the MPL of the major component. A relatively large error (± 19%) was observed on the MPL measurements in comparison to the values usually obtained (circa ± 9%) in measurements of similar peptide and protein filaments [64-68].

The observed MPL of the **7HSAP1** fibrils can be compared to that calculated for nanotube assemblies derived from stacking of 7-helix bundles analogous to that of **GCN4-pAA**. We used the crystal structure of **GCN4-pAA** to generate a model for the **7HSAP1** fibril (*vide infra*), in which a bundle length of circa 52 Å was calculated based on the super-helical rise/residue of 1.45 Å and an average spacing between bundles of 1.6 Å. Using this value and the molar mass of **7HSAP1** (3823 Da), a MPL value of 511 Da/Å was calculated for a stacked assembly of 7-helix bundles. This value is within experimental error of the MPL value observed from dark-field STEM measurements. However, due to the large variance in the MPL measurements, a filament based on a 6-helix bundle structure could not be ruled out as the structural subunit. The design and structural characterization of individual (i.e., non-fibrillar) 6-helix bundles based on coiled-coil motifs has been described recently [69, 70]. However, the 6-helix bundles differ significantly in core sequence from those of **7HSAP1** and **GCN4-pAA**, and form in-register cylinders rather than edge displaced lock-washer structures.

3.3.3 Solution X-ray scattering measurements

Small- and wide-angle X-ray scattering (SAXS/WAXS) data were collected on aqueous solutions of **7HSAP1** in MES buffer (10 mM, pH 6.0) to interrogate the structural hierarchy of the resultant assemblies at length scales over two decades. The small-angle range, i.e., the scattering momentum transfer (q), less than 0.2 $Å^{-1}$ reflects the global shape of the molecule, while the middle q range, ca. $0.2 < q < 1.1$ $Å^{-1}$, reflects the tertiary fold of the assembly.

In the small-angle region, the experimental scattering intensities at $q < 0.1$ $Å^{-1}$ roughly follow the q^{-1} power law (Figure 3.7A), indicating the assembly in solution has rod- or cylinder-like form. The pair distance distribution function (PDDF) [71] in Figure 3.7B derived from the SAXS data has a characteristic long tail, which further confirms the cylindrical shape. A cross-section PDDF derived from $q*I(q)$ exhibits a bimodal form, indicating the cylindrical molecular assembly is hollow. The largest dimension of the cross-section is about 32 Å. The average wall thickness and the diameter of the middle of the cylindrical shell can be estimated from the two maxima of the PDDF as ca. 10 and 20 Å, respectively, and the diameter of the hollow channel is about 10 Å. The value of R_c, the radius of gyration of the rod cross-section, was obtained as 12.4 Å through the fitting of the modified Guinier equation, [72] $I(q) = \pi q I(q=0)\exp(-0.5R_c^2q^2)$. The discrepancy between the data and rod-like q^{-1} power law at $q < 0.01$ Å^{-1} reflects the inhomogeneity of the assembly length in solution.

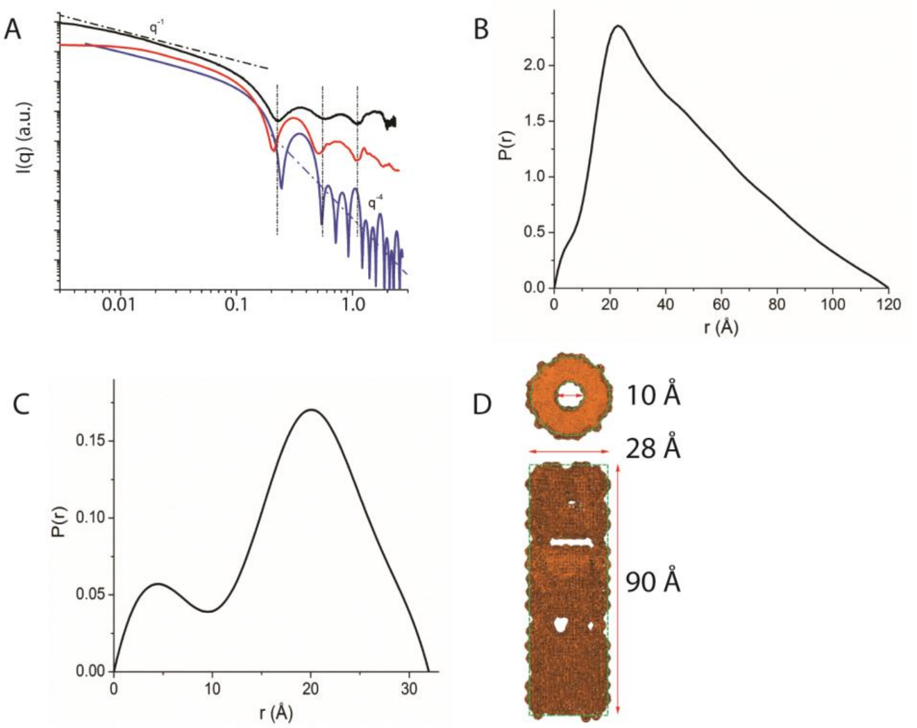


Figure 3.7. SAXS/WAXS studies of 7HSAP1. A. Experimental SAXS/WAXS scattering profile (black curve) for 7HSAP1 (1 mM) in MES buffer (10 mM, pH 6.0), along with simulated data for a hollow cylindrical model (blue curve) and a molecular model (red curve) based on five stacked seven-helix bundles of 7HSAP1. B. The pair distance distribution function (PDDF) of 7HSAP1, calculated from the SAXS (*I(q)* versus *q*) data using program GNOM [71]. C. The cross section PDDF calculated from SAXS data (*q*I(q)* versus *q*) using GNOM. D. The top and side view of the SAXS molecular envelope derived from a nanotube search space with R_{in} of 5 Å and R_{out} of 14 Å using the scattering data from 7HSAP1 and the program

DAMMIN [73]. Scattering data within the *q* range of 0.03-0.58 Å^{-1} were used in calculations of B, C and D. The largest molecular dimension (r_{max} in B) and the length of the SAXS molecular envelope may be under-estimated due to the limited available low *q* data.

To further interpret the scattering data at higher-angle regions, a molecular model was generated that corresponded to five stacked seven-helix bundles of **7HSAP1** (*vide infra*). Simulation of the X-ray scattering profile for this molecular model approximately reproduces the two scattering peaks centered at 0.35 and 0.80 Å^{-1} in the experimental data in terms of the peak and valley positions (Figure 3.7A). Simulations on the hollow cylindrical object model showed that the inner and outer radii determined the valley position. Therefore, the cross-section dimensions of the molecular model are very close to those in the sample. The more damped peak-valley ratio in the experimental data is partially due to higher thermal dynamics/disorder in the sample and the limitations inherent in the molecular model. At high angle, the scattering peak centered at 1.5 Å^{-1} arises from the repeatable occurrence of the atom pairs with 3-6 Å distances within the α-helices, which were also partially reproduced in the simulated scattering for the molecular model.

The full-range scattering data could not be fit using a structural model based on a hollow cylindrical geometric object of uniform electron density. The impact of the electron density fluctuation within the molecule becomes significant at *q* values of ≥ 0.3 Å^{-1}, which influences the intensity and the position of the peaks and valleys (Figure 3.7A). However, an *ab initio* three-dimensional envelope reconstruction from SAXS data was performed and the derived SAXS molecular envelope exhibits a hollow cylindrical shape with inner and outer radii of circa 5 and 14 Å, respectively. The SAXS envelope yields a R_c of 10.5 Å under the uniform electron density assumption where $R_c = \sqrt{\frac{R_{in}^2+R_{out}^2}{2}}$. The R_c value of the envelope agrees fairly well with the previous R_c from the modified Guinier fitting, considering the density fluctuation in the SAXS envelope model. Taken together, the SAXS/WAXS data support the presence of hollow nanotube assemblies in solutions of **7HSAP1** with lateral dimensions that approximate those observed in the crystal structure of the 7-helix bundle of **GCN4-pAA** (R_{out} = 15.5 Å, R_{in} = 3.5 Å).

3.3.4 X-ray Fiber Diffraction

X-ray fiber diffraction data collected from air-dried, partially aligned bundles of fibers (Figure 3.8) show a sharp signal on the meridian at 10.1 Å as well as weaker, more diffuse reflections at 5.1 Å and 4.3 Å. On the equator, reflections appear at around 26 Å, 14.5 Å and 8 Å. The meridional reflections observed at 5.1 and 10.1 Å are consistent with the regular repeat that arises from α-helices and the position on the meridian supports the view that the α -helices are aligned parallel to the fiber axes, which is consistent with the flow LD spectral data in solution. Analyses of the equatorial spacings are commensurate with a unit cell of similar dimensions to the diameter of the model structure (30 Å). An X-ray fiber diffraction pattern was calculated from structural coordinates of a molecular model corresponding to five stacked seven helix bundles of **7HSAP1** (*vide infra*). The model was placed into a unit cell with hexagonal packing (*a*=30 Å, *b*=30 Å, *c*=52 Å, α=β=90°, γ=120°) using the program Clearer.[74] The calculated pattern is shown in comparison to the experimental diffraction data in Figure 3.8. Calculated signals match well with positions of signals in the experimental data. Discrepancies between relative intensity are observed and this is likely to arise from differences in side chain rotamers and minor backbone conformational differences. However, the results are supportive of the proposed model structure for assemblies of **7HSAP1** that correspond to stacked seven-helix bundle structures (Figure 3.2). Alternatively, 6-helix bundle structures [69, 70] were examined as models, however the 7-helix bundle structure afforded a better fit to the unit cell and was more consistent with the experimental fiber diffraction.

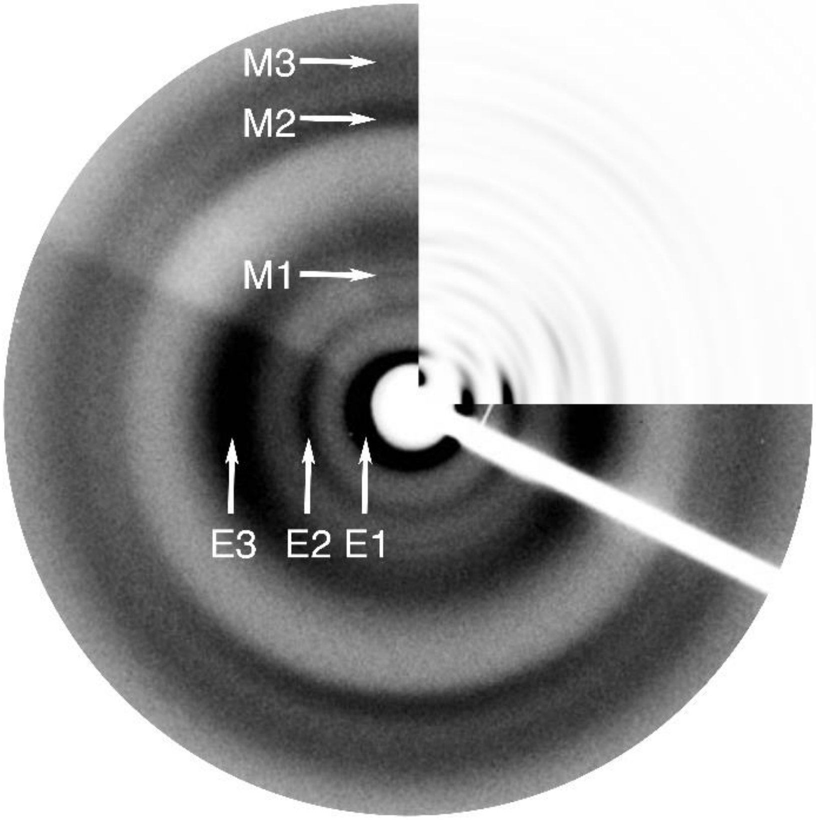


Figure 3.8. X-ray fiber diffraction of 7HSAP1 obtained from dried, partially oriented fibrillar assemblies. Equatorial reflections are noted at 26 Å (E1), 14.5 Å (E2), and 8 Å (E3), while meridional reflections are noted at 10.2 Å (M1), 5.1 Å (M2)

and 4.3 Å (M3). The upper right quadrant corresponds to the calculated diffraction pattern of stacked 7-helix bundle assemblies of 7HSAP1 based on a unit cell with hexagonal packing (*a*=30 Å, *b*=30 Å, *c*=52 Å, α=β=90°, γ=120°). The M1 reflection is too weak to observe in the simulation, but can be detected through the off-meridionals that are part of that layer line.

3.3.5 Solid-state NMR measurements

In order to obtain proof of the registry shift within the fibrillar assemblies of **7HSAP1**, we used distance measurements from solid-state NMR. A variant peptide, **7HSAP1***, was synthesized in which the Ala12 CO and Ala24 CH_3 groups were enriched with ^{13}C and the Leu16 NH group was enriched with ^{15}N. This labeling scheme, with a circa 160 ppm chemical shift difference between ^{13}CO and $^{13}CH_3$, allows for independent determination of both the intramolecular backbone Ala12 ^{13}CO - Leu16 ^{15}N distance and the intermolecular Ala24 CH_3 - Leu16 amide ^{15}N distance within a single peptide sample. The [1-^{13}C]Ala12 - [3-$^{13}CH_3$]Ala24 intramolecular distance of circa 17 Å minimizes any homonuclear ^{13}C-^{13}C dipolar coupling. The Ala12 ^{13}CO–Leu16 amide ^{15}N distance corresponds to an intramolecular *i,i+4* hydrogen-bonded pair of residues and can be correlated to the fraction of the **7HSAP1*** peptide that has folded into an α-helical conformation.

The Ala24 CH_3 - Leu16 amide ^{15}N distance corresponds to an intermolecular contact between adjacent helical interfaces within the 7-helix bundle structure. However, as solid-state NMR ^{13}C-^{15}N distance measurements are limited to less than 7 Å, this distance can only be determined at the interface between helices *A* and *G* of the suppositious 7-

helix bundles of **7HSAP1** assemblies. Using the crystal structure of **GCN4-pAA** as a guide (Figure 3.10E) [43], the distance between labeled atoms at structurally homologous positions in **7HSAP1*** could be calculated. The intramolecular Ala24 $^{13}CH_3$ - Leu16 amide ^{15}N distance in **7HSAP1** were determined to be circa 13 Å, while the closest intermolecular distance between structurally adjacent helices within the 7-helix bundle is circa 11 Å at the non-*A*/*G* helical interfaces. Both of these distances correspond to ^{13}C-^{15}N dipolar couplings that are too weak to be measured with solid-state NMR. However, the registry shift between adjacent helices places the Leu16 amide of helix *G* and the Ala24 CH_3 of helix *A* in close contact; 3.5 Å within the crystal structure of **GCN4-pAA**. Since this intermolecular contact defines the displaced edge of the 7-helix bundle, the corresponding distance can be correlated with peptide registry within the assembly and retention of the screw symmetry observed in the parent **GCN4-pAA** system (Figure 3.10E).

The unbiased fit of the Ala12 ^{13}CO $^{13}C[^{15}N]$REDOR [75] dephasing data, using a Boltzmann maximum-entropy approach (BS-REDOR) [76], as isolated ^{13}C–^{15}N spin-pairs (Figure 3.10A and B) indicates that 85% of Ala12 ^{13}CO are directly H-bonded to Leu16 amide ^{15}NH with the remaining 15% of the Ala12 ^{13}CO having ^{13}C-^{15}N distances great than 7 Å. This observation suggests that in the NMR sample preparation approximately 85% of the peptide is folded into an α-helix. The ^{13}C CP/MAS NMR spectrum of **7HSAP1** (Figure 3.9) did not exhibit a distinct upfield β-sheet resonance for the Ala12 ^{13}CO chemical shift, consistent with the hypothesis that the 15 % of the Ala12 ^{13}CO with no ^{15}N REDOR dephasing corresponded to peptide in a disordered conformation. The disordered fraction of the sample could result from denaturation during sample preparation and/or incomplete self-assembly. However, the REDOR data suggest that the helical portion of the sample is structurally defined with a narrow distribution of ^{13}C-^{15}N distances centered at 4.1 Å, as expected for hydrogen-bonded contacts within α-helices.

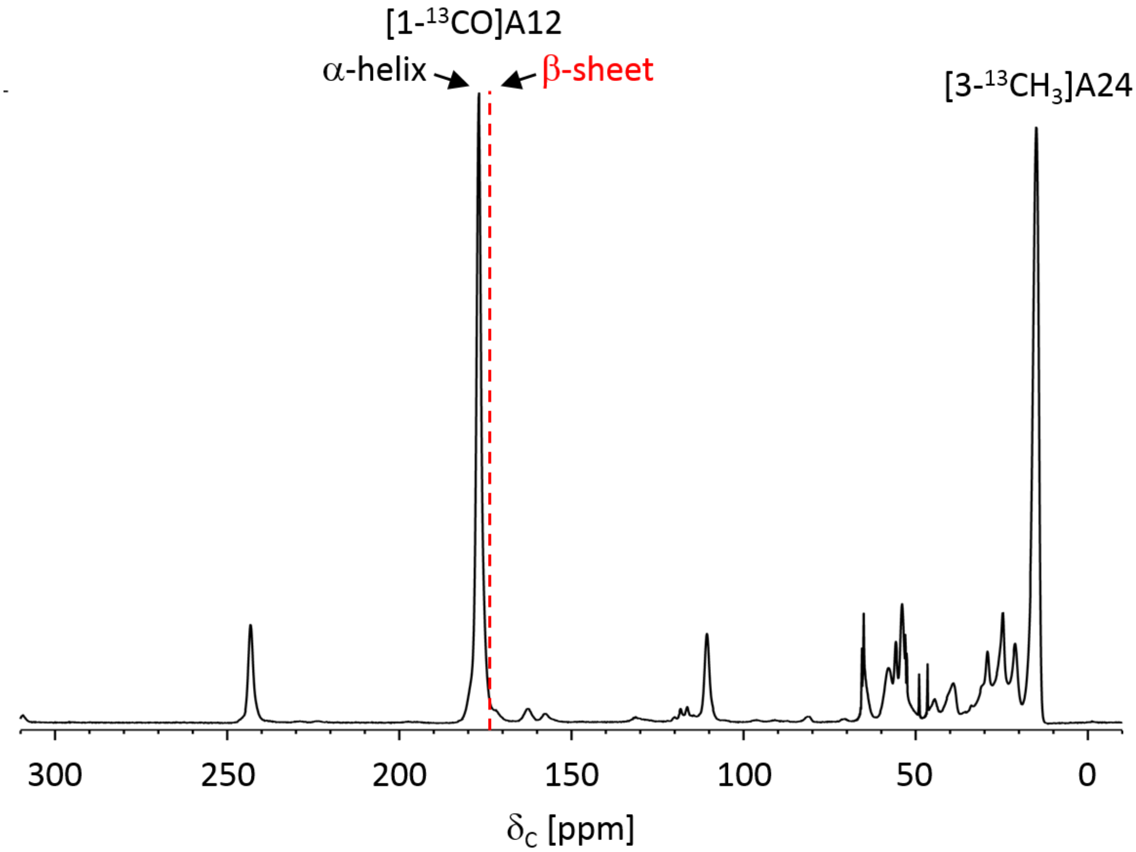


Figure 3.9. One-dimensional ^{13}C CP/MAS solid-state NMR spectrum of peptide 7HSAP1* indicating the chemical shift positions of the 1-^{13}C-alanine and 3-^{13}C-alanine labels.

The BS-REDOR fits of $^{13}C[^{15}N]$ REDOR dephasing of Ala24 $^{13}CH_3$ indicates that 8 ± 2% of the Ala24 methyl carbons are 3.9 ± 0.1 Å from Leu16 amide nitrogen and 8 % of

Ala24 $^{13}CH_3$ have a ^{13}C-^{15}N distance greater than 6 Å (Figure 3.10C and D). The 6 Å distance and its contribution to the 3.9 Å is ill-defined due to an experimental limit for maximum REDOR dephasing evolution time of circa 70 ms. The observed distance of 3.9 Å correlates reasonably well with the corresponding distance (3.5 Å) between helices *A* and *G* in the crystal structure of **GCN4-pAA**, which provides support that the single

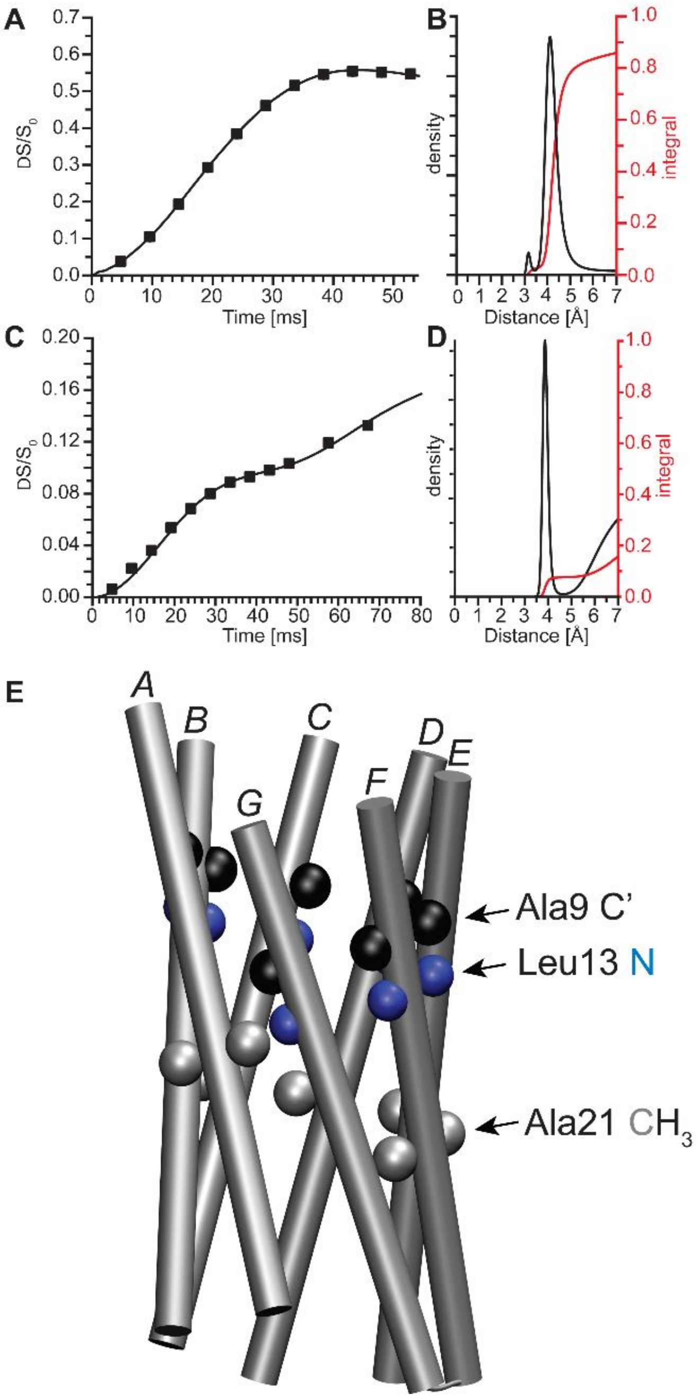


Figure 3.10. Solid-state NMR ^{13}C-^{15}N distance measurements of nanotubes derived from self-assembly of 7HSAP1* peptides. (A) ^{13}C[^{15}N]REDOR dephasing for [1-^{13}C]A12 [^{15}N]L16-7HSAP (7HSAP*). Solid curve is the calculated best fit REDOR curve with the distances from (B) BS-REDOR fits to experimental data. The BS-

REDOR best fit has circa 85% (red curve) of Ala12 ^{13}CO to Leu16 ^{15}N spin-pairs centered at 4.1 Å (black curve), which is consistent with Ala12 CO H-bonded to Leu16 amide NH. (C) ^{13}C[^{15}N]REDOR dephasing for [3-^{13}C]A24 [^{15}N]L16-7HSAP1 and (D) distance distribution from BS-REDOR fit to experimental data with 8% of Ala24 ^{13}CH$_3$ at 3.9 Å from Leu16 ^{15}N with a narrow distribution and a second ill-defined distribution with circa 8% of ^{13}CH$_3$-^{15}N distances > 6 Å. (E) Diagram of the 7-helix bundle structure of GCN4-pAA[43] in which helices *A* through *G* are drawn as cylindrical rods. The colored spheres indicate atoms within the 7-helix bundle of GCN4-pAA that occupy structural homologous positions with respect to the isotopically labeled sites in 7HSAP1*: carbonyl carbons (C') of Ala9 (black), amide nitrogens of Leu13 (blue) and the methyl carbons of Ala21 (gray). The closest contact between AlaCH3 – LeuN corresponds to a distance of 3.5 Å at the displaced interface between helices *A* and *G*.

heptad registry shift is retained within the structure of the **7HSAP1**-derived fibers. Given that the Ala12 ^{13}CO - Leu16 ^{15}N REDOR dephasing constrains 85% of the peptide to be folded into an α-helix, 1/7th of 85% (i.e., 12 %) of the Ala24 sidechain ^{13}CH$_3$ should be within 3.9 Å. Together this suggests that a fraction (2-4 %) of the 7-helix bundle structures in the **7HSAP1** assemblies is locally disordered. The presence of the 6 Å distance may be related to disorder in the helical sample. However, the observed distance is too short to correspond to other intramolecular or intermolecular interactions within the 7-helix bundles (*vide infra*). One possible scenario is that the longer distance may result from local disorder at the ends of the fibrillar assemblies of **7HSAP1**. Modeling studies of **7HSAP1** assemblies (*vide infra*) provide support for this hypothesis in that the Ala24 CH_3 to Leu16 amide ^{15}N distance increases at the ends of stacked assemblies and become less defined, especially in comparison to the internal positions of the nanotube fibril (Figure 3.12).

3.3.6 Computational Modeling

While the solid-state NMR measurements on the **7HSAP1*** fibrils were consistent with an assembly based on 7-helix bundles with displaced edges, the measured REDOR distance (3.9 Å) between the 3-^{13}C-Ala(24) on the helix *A* and the ^{15}N-Leu(16) on the helix *G* differed from the corresponding distance (3.5 Å) measured between structurally homologous positions within the crystal structure of **GCN4-pAA**. To better understand this discrepancy between the helix-helix contact distances at the displaced edge, an analysis of the 7-helix bundle structure of **GCN4-pAA** was undertaken. Prior analysis of the side-chain packing within this structure was performed using the program SOCKET [77] and were available in CC+, the coiled-coil structural database [78, 79]. The SOCKET analysis indicated a complex structure, in which knob-into-holes packing occurred over an extended hydrophobic interface involving the *a*-*d*-*e*-*g* residues within multiple helices.

The coiled-coil analysis program samCC [80, 81] was modified to determine the structural parameters associated with individual helices within the structure of **GCN4-pAA**. This algorithm was originally developed to detect deviations in the Crick parameters for 4-helix bundle structures through comparison to an idealized tetrameric coiled-coil model displaying canonical knob-into-holes interactions. Similarly, an idealized 7-helix bundle structure based on central residues 4-31 of the **GCN4-pAA** sequence was constructed using the program BeamMotifCC (see *Methods*) [81]. This structural model

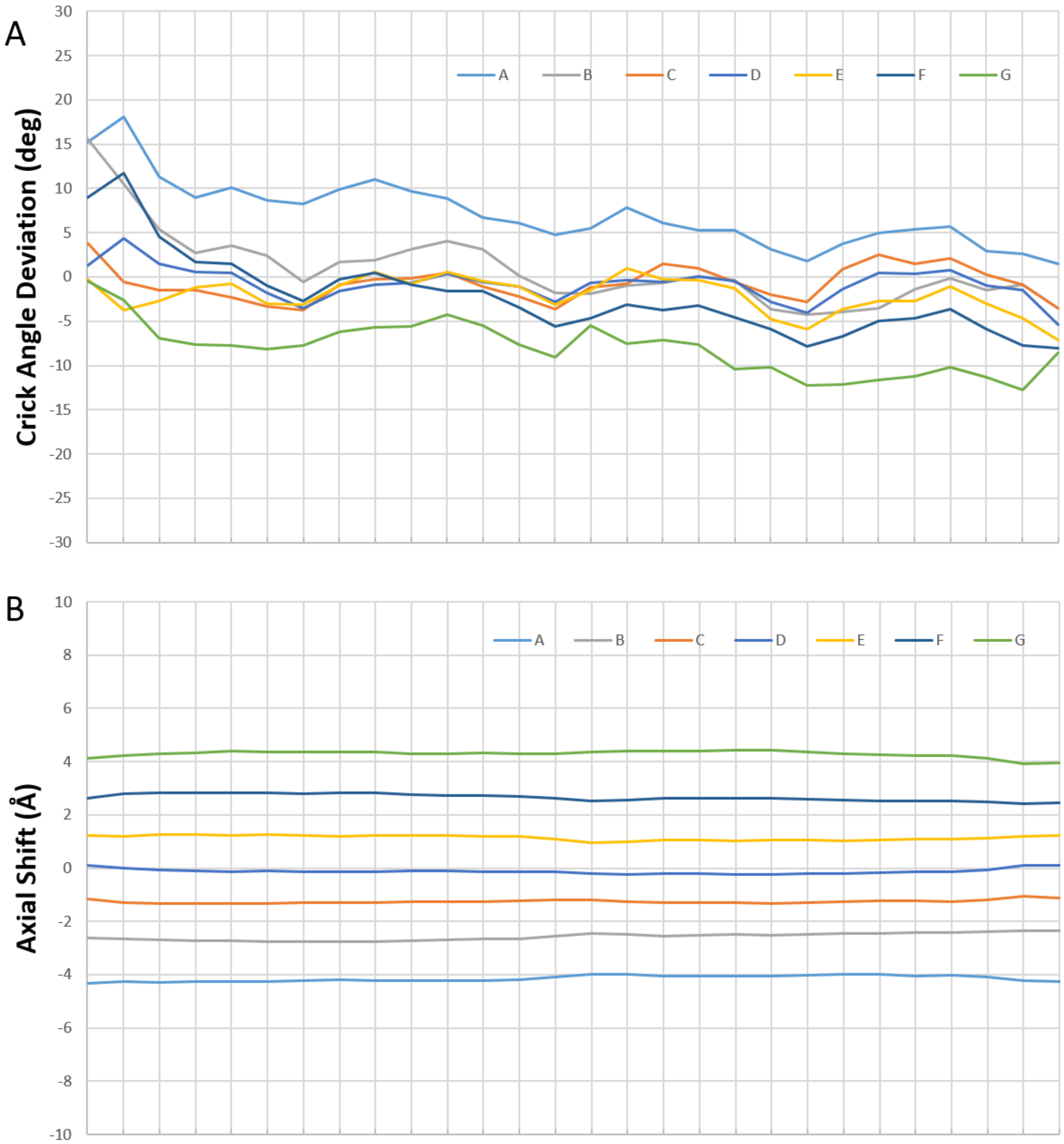


Figure 3.11. Crick angle deviations (A) and axial shifts (B) of GCN4-pAA, calculated using samCC, as a function of residue position (Lys4-Val31) for helices A-G in the 7-helix bundle structure (PDB ID: 2HY6) of GCN4-pAA.

(2HY6.ideal.pdb) was employed as a benchmark for the analysis of the crystal structure of **GCN4-pAA**.

Two significant differences were observed between the actual **GCN4-pAA** structure and the reference model calculated based on Crick's equations (Figure 3.11) [81-83]. The first difference involved the axial displacement between adjacent helices in the bundle, which corresponded to an average value of 1.41 Å. This displacement can be compared to the average axial rise/residue of 1.51 Å that is associated with individual amino acid residues within an α-helix. This feature was noted previously in the crystal structure analysis of **GCN4-pAA** [40-42] and underlies the screw symmetry of the lock washer structure.

The second structural difference observed for **GCN4-pAA** vis-à-vis the idealized model structure involved deviations of the minor helical phases, ϕ_1, from the ideal Crick angle values for helices *A* and *G*. The axial rotations corresponded to approximately +7 ° and –7 ° for helices *A* and *G,* respectively (Figure 3.11). Helices *B*-*F* maintained near-ideal Crick angle values for residues in the heptad repeat units. The net result of the Crick angle deviations in the structure of **GCN4-pAA** was an outward rotation of helices *A* and *G* with respect to the other helices in the bundle. This rotation altered the hydrophobic packing at the displaced edge away from classical knobs-into-holes (*a*,*d*-layer) packing towards *x*,*da*-layer packing [81], *a*-*d*-*g* and *a*-*d*-*e* packing for helices *A* and *G*, respectively.

To understand the influence of the Crick angle deviation on the inter-helix separation, a simple model structure for **7HSAP1** was constructed that incorporated the Crick parameters of the idealized 7-helix bundle structure in conjunction with the axial translation observed in the crystal structure of **GCN4-pAA**. The structure fitter module of the Crick coiled-coil parameterization (CCCP) program [84, 85] was employed to computationally generate Crick parameters for the 7-helix bundles of the actual and ideal **GCN4-pAA** structures from the corresponding protein database files. The structure generator of the CCCP program was then used to create an idealized 7-helix bundle structure in which an axial translation equivalent to a single amino acid occurred between adjacent helices. The Crick parameters of the idealized in-register model were used as input data, with the exception that the $\Delta Z_{offset,}$, i.e., axial displacement, values were derived from **GCN4-pAA**. The resultant structural model was derived from a 7-helix bundle based on a polypeptide consisting of thirty-five alanine residues that displayed the axial translation and the displaced edge of the **GCN4-pAA** structure (Ala35_7.pdb).

The distances between methyl carbon atoms on chain *A* and amide nitrogen atoms on chain *G* were measured at structurally homologous positions at the displaced edge within this $[(\mathrm{Ala})_{35}]_7$ structural model. The corresponding N-Ala(16)^{G}-3-C-Ala(24)^{A} distance (4.8 Å) within the model structure was significantly longer than the distances of 3.9 Å and 3.5 Å derived from REDOR NMR measurements on **7HSAP1*** and the crystal structure of **GCN4-pAA**, respectively. The longer ^{15}N-Leu(16)^{G}-3-^{13}C-Ala(24)^{A} distance of 3.9 Å for **7HSAP1*** may imply a lesser deviation of the Crick angle values from those observed for the corresponding helices in **GCN4-pAA**. Given that the two sequences are not identical and were prepared under different conditions, one might expect that subtle structural differences may occur even though the global conformation of the 7-helix bundle is retained.

Molecular modeling with molecular dynamics (MD) simulation was employed for the analysis of a stacked assembly of five 7-helix bundle structures of **7HSAP1** in the presence of explicit water. A simulation for 1.2 ns indicated convergence to a model structure, (**7HSAP1_**assembly_MD.pdb), in which non-covalent interactions persisted between helical bundles (Figure 3.12). An average distance of 1.6 Å was observed for the gap between successive helical bundles in the center of the stack. Significant differences were observed in the Crick parameters between internal and terminal bundles within the stacked assembly. CCCP analysis of the central helical bundle of the assembly (**7HSAP1_**centralbundle.pdb) indicated that its Crick parameters compared quite well to those determined for the 7-helix bundle structure of **GCN4-pAA**, including the deviations in the Crick angles for helices *A* and *G* at the displaced edge. In addition, the ^{15}N-Leu(16)^{G}-3-^{13}C-Ala(24)^{A} distance of 3.5 Å for the central 7-helix bundle model matched the corresponding value for the **GCN4-pAA** structure. These results suggest that assembly into nanotube should not cause significant distortion of the 7-helix bundle structure at internal positions within the assembly. However, the REDOR measurement indicates that some degree of structural distortion must occur, as judged on the basis of the larger ^{15}N-Leu(16)^{G}-3-^{13}C-Ala(24)^{A} distance. A slight rotation of the helices at the displaced edge would result in a change in the Crick angle vis-à-vis the **GCN4-pAA** structure, which would account for the increased helix-helix distance in **7HSAP1** assemblies. This situation may result from the formation of lateral interactions between fibrils and/or removal of waters of solvation; either of which could occur during sample preparation for solid-state NMR experiments.

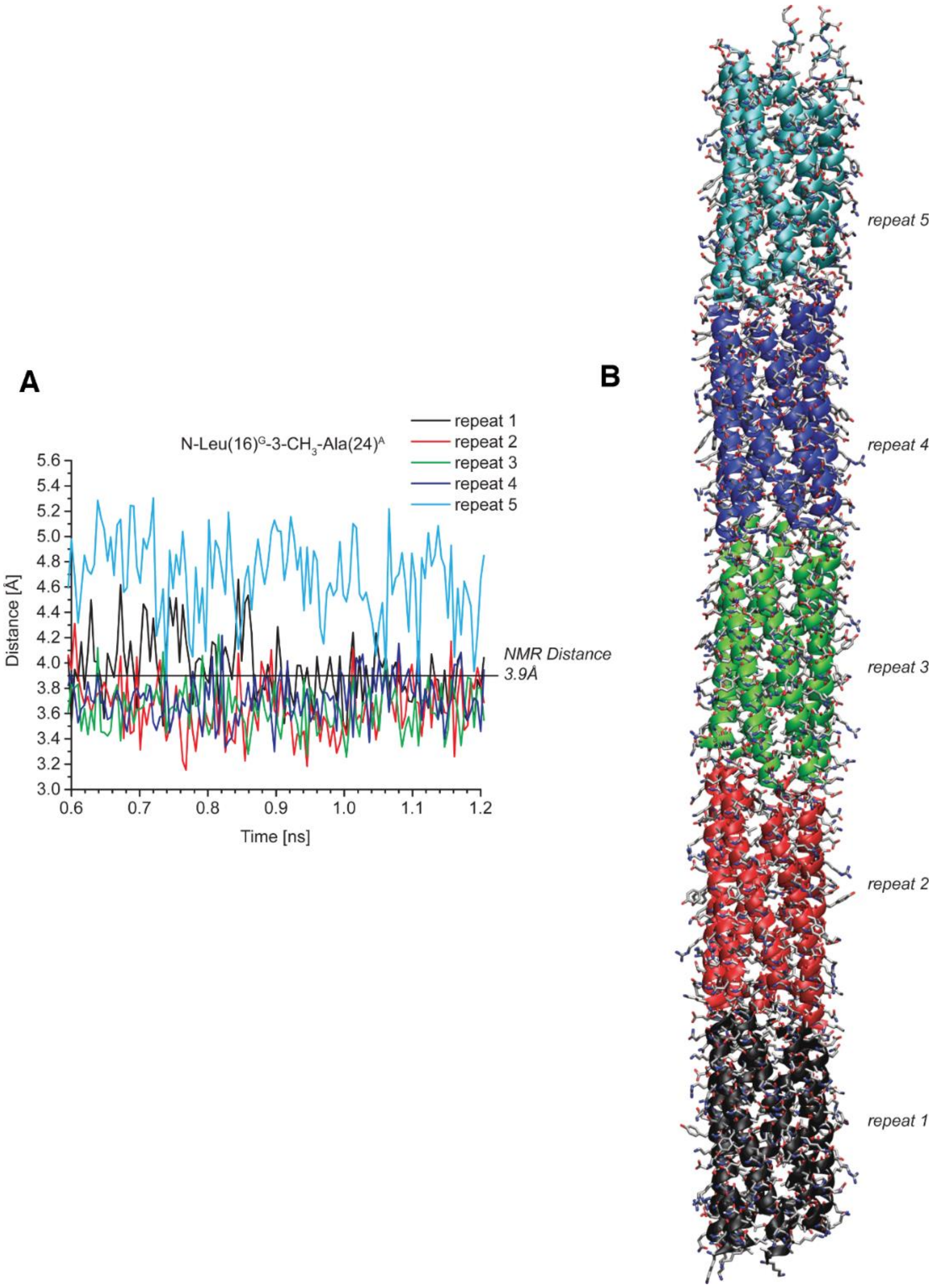


Figure 3.12. Molecular dynamics simulation of 7HSAP1 bundle. A. Plot of ^{15}N-Leu(16)^G-3-^{13}C-Ala(24)^A distances at the displaced edge. B. Molecular model of the corresponding stacked assembly after 1.2 ns of molecular dynamics simulation.

Individual bundles (repeats) are color-coded in the model and the corresponding plot.

3.3.7 Fluorescence Spectroscopy

In analogy to **GCN4-pAA**, the 7-helix bundle sub-unit of **7HSAP1**-derived fibrils should define a central channel with an internal diameter of 7 Å and a length of circa 50 Å. The diameter of this channel compares in dimension to the internal cavity of γ-cyclodextrin (γ -CD) [86, 87] and, therefore, the **7HSAP1** assemblies should be able to accommodate substrates of similar size and shape as molecules that form inclusion complexes with γ-CD. The solvatochromic fluorophore PRODAN,[88, 89] 6-propionyl-2-(*N*,*N*-dimethylamino)-naphthalene, was employed as a structural probe in substrate binding assays as it has been shown to bind in a concentration dependent manner within the internal cavity of γ-CD[90] (Figure 3.12).

The fluorescent emission spectrum was monitored as a function of **7HSAP1** concentration at a fixed concentration of PRODAN (1 µM in 10 mM MES buffer, pH 6.0). Upon addition of the peptide, the λ_{max} for the PRODAN emission shifted from 520 nm to 430 and gained strongly in intensity. The observed spectroscopic behavior is commensurate with an increase in hydrophobicity in the environment of the fluorophore. Similar wavelength displacements have been observed for binding of PRODAN within the hydrophobic internal cavity of γ-CD. In contrast, we performed a similar titration with peptide **YZ1**, which forms a fibrillar structure derived from a coiled-coil dimer.[11] The fluorescence of PRODAN did not shift in position from the initial λ_{max} value of 520 nm, nor did the intensity of the emission increase significantly over the course of the experiment (Figure 3.14). These results suggest that **YZ1** does not bind PRODAN, which supports the hypothesis that PRODAN binds selectively to **7HSAP1** but not generally to coiled-coil assemblies.

In the case of the **7HSAP1**-derived fibrils, the fluorescence data indicate that PRODAN binds within the hydrophobic channel of the 7-helix bundle structures. Fluorescence titration of PRODAN with **GCN4-pAA** under identical conditions afforded a similar spectroscopic response (Figure 3.15), although the binding response did not saturate under the experimental conditions. Since the 7-helix bundle structures of **GCN4-pAA** do not self-associate further under these conditions,[43] the fluorescence response must be associated solely with binding of PRODAN within the central hydrophobic cavity of the **GCN4-pAA** assembly and, by analogy, the 7-helix bundle sub-units of the **7HSAP1** nanotubes. We cannot rule out the possibility that PRODAN binding occurs only within bundles that have dissociated from the fibril. However, PRODAN does not affect the global self-assembly of **7HSAP1** as judged from TEM analysis, in which fibrils were observed in the presence of the fluorophore that were identical is structure to those observed in its absence.

To estimate the binding affinity of PRODAN to the **7HSAP1**-derived fibrils, the binding isotherm was determined as a function of peptide concentration (Figure 3.12, inset). The fluorescence level reached a plateau at a peptide concentration of 75 µM (circa 10 µM in 7-helix bundle), which is approximately two orders of magnitude lower in concentration of the host at ligand saturation than the corresponding values for PRODAN binding to either β- or γ-cyclodextrin.[85] These data suggest that PRODAN binds tightly within the central channel of the **7HSAP1**-derived assemblies. The relatively tight binding response of the peptide-based assembly for PRODAN suggests the potential for the encapsulation of shape-appropriate substrates within the nanotube assemblies for controlled release applications.

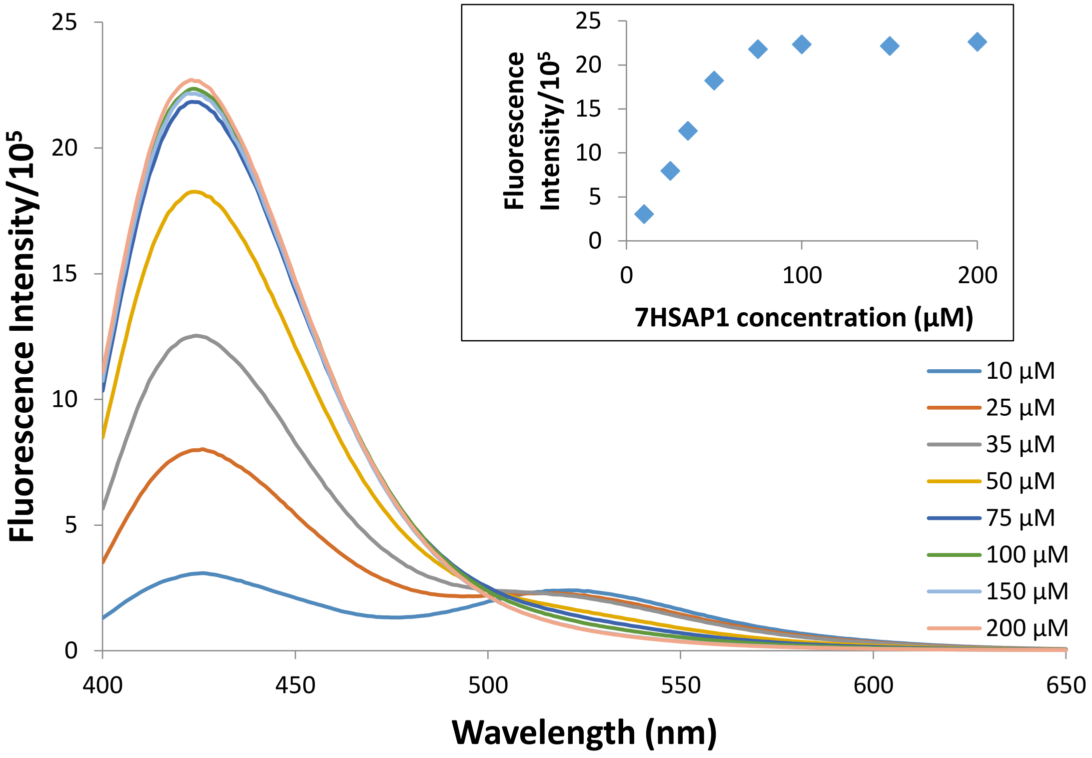


Figure 3.13. Fluorescence titration of 7HSAP1 into PRODAN (1 µM) in MES buffer (10 mM, pH 6.0). Inset: Dependence of fluorescence emission intensity on 7HSAP1 concentration.

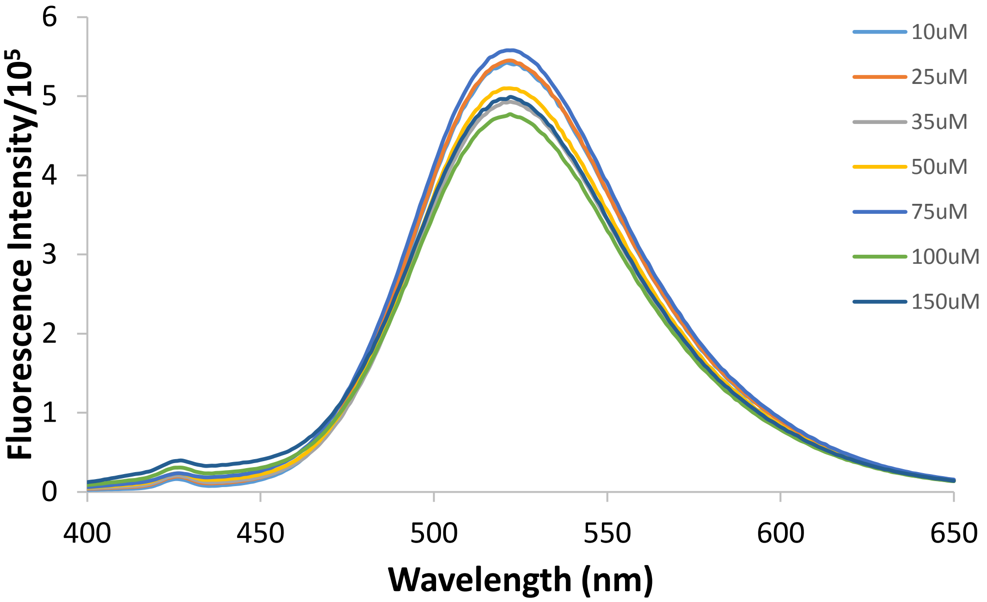


Figure 3.14. Fluorescence titration of peptide YZ1 into a solution of PRODAN (1 µM) in MES buffer (10mM, pH6.0). Peptide YZ1 forms bundled fibrils based on a dimeric coiled-coil structural motif and, as such, should not be able to bind PRODAN as it lacks an internal cavity that can accommodate the fluorophore.

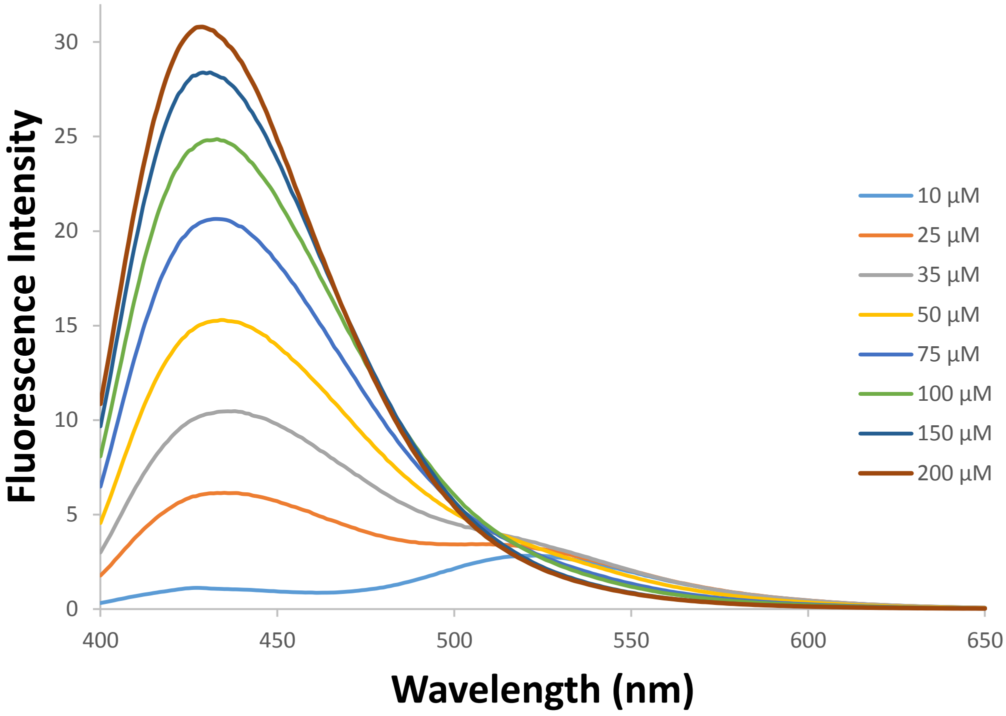


Figure 3.15. Fluorescence titration of peptide GCN4-pAA into a solution of PRODAN (1µM) in MES buffer (10mM, pH6.0). Peptide GCN4-pAA forms a soluble 7-helix boiled coil under these conditions and does not fibrillize to any detectable extent.

3.3 Summary

Protein-based nanotubes represent attractive candidates for the construction of functional nano-porous materials. In native biological systems, tubular structures derived from self-assembly of folded protein domains occur frequently in functional roles that include directional transport [91], controlled release [92], and catalysis [93], Nanotubes derived from native protein assemblies, including tubulin [94], flagellin [95], and filamentous viruses [96], have been explored as structural elements for the construction of nanoscale devices for diverse applications [97, 98]. However, a far greater repertoire of protein structures is available that display defined internal cavities but do not self-associate to form stable supramolecular assemblies. The functional diversity of nanotube assemblies could be greatly expanded if these non-assembling protein motifs could be recoded at the sequence level to drive self-association into structurally defined assemblies.

This approach has been employed previously with a degree of success to create synthetic protein nanotubes from structurally modified ring-like protein assemblies. Ballister, *et al*., [99] modified the hexameric protein Hcp1 from *Pseudomonas aeruginosa* through introduction of cysteine residues at structurally informed positions at the protein surface. Under oxidizing conditions, the Hcp1 mutant formed nanotubes up to 100 nm in length that consisted of circa 25 hexameric sub-units. Similarly, Miranda, *et al*., [100] introduced cysteine as well as several other mutations at surface positions within the undecameric assembly of TRAP (trp RNA binding attenuation protein) from *Bacillus stearothermophilus*. Surprisingly, the mutant TRAP assembled into nanotubes only in the presence of a bifunctional thiol, which presumably acted indirectly as a cross-linking agent between surface cysteines, albeit at a low-level of substitution. Individual nanotubes were observed to be several hundred nanometers in length and could bundle into larger assemblies. This strategy relied mainly on covalent capture through disulfide bond formation to drive nanotube formation, although some degree of surface interaction was implied as a stabilizing influence on nanotube assembly.

In contrast, the self-assembly of **7HSAP1** relies exclusively on selective recognition between structurally complementary interfaces to drive self-assembly. These interactions are reinforced through Coulombic attraction between oppositely charged interfaces and, potentially, through hydrogen bond formation (either directly between peptides or mediated by waters of solvation). Minor alterations introduced into the sequence of **GCN4-pAA** were sufficient to alter its assembly behavior from an isolated heptameric bundle to a nanotube of defined diameter with lengths in the multi-micron range. Moreover, these high aspect-ratio assemblies formed at concentrations of **7HSAP1** as low as 100 μM. However, at lower concentrations, the STEM data indicated nanotubes contained imperfections including breaks in continuity and/or the presence of lower mass species. While the structural data are consistent with nanotubes derived from stacking of 7-helix bundles, the interfaces have not been optimized energetically to promote the formation of stable contacts.

Tezcan, *et al*. have recently reported the stable formation of a complex range of one-, two-, and three-dimensional, including nanotubes, via metal-mediated self-assembly of cytochrome *cb-562* mutants [101, 102]. The program Rosetta [103] was employed to computationally optimize interactions at the protein-protein interfaces, which enhanced the stability of the resultant assemblies. While the mode of assembly of the cytochrome *cb-562* mutants is different from the lock washer stacking of **7HSAP1**, computational optimization of the complementary interfaces represents an attractive strategy to further refine the formation of nanotube assemblies.

Nonetheless, we have demonstrated that simple modification of structurally complementary interfaces within a folded protein structure can drive self-assembly of high aspect-ratio nanotubes. Moreover, structural analysis across multiple length-scales has provided strongly supportive evidence that the nanotube assemblies derived from **7HSAP1** comprise sub-units that retain the 7-helix bundle of **GCN4-pAA** with minor structural modifications. In addition, these nanotubes can bind shape-appropriate small-molecules with an affinity that exceeds common host molecules of similar internal dimensions such as cyclodextrins. The external and internal surfaces of the nanotubes are structurally distinct and display the potential for modification through mutagenesis of the peptide sequence. Although the structural plasticity of **7HSAP1** has not been explored, the utility of the corresponding nanotubes would be greatly enhanced if the chemistry of the internal channel could be controlled to facilitate applications in transport and catalysis. Woolfson, *et al.*, [63] recently reported that internal channel of 6-helix bundles was amenable to mutagenesis with retention of the overall fold. While the structure of the 7-helix bundle of **GCN4-pAA** is more complex than the classical knobs-into-holes packing of coiled-coils, [78, 79] preliminary studies have indicated that the structure is compatible with limited mutagenesis at core residue positions within the peptide sequence [43]. We envision that chemically modified nanotubes derived from **7HSAP1** may offer considerable promise for the development of functional nano-porous materials.

3.4 Materials and methods

Peptide Synthesis. Peptides **GCN4-pAA** and **7HSAP1** were either purchased from GenScript USA, Inc. (Piscataway, N.J.) or synthesized in-house. In the latter case, peptides were prepared using microwave-assisted synthesis on a CEM Liberty solid-phase peptide synthesis instrument using a 4-(hydroxymethyl)phenoxyacetamido-methyl]-PEG-PS resin from Applied Biosystems, Inc. (Foster City, CA). Standard Fmoc protection chemistry was employed with coupling cycles based on HBTU/DIEA-mediated activation protocols and base-induced deprotection (20% piperidine in DMF with 0.1 M HOBt) of the Fmoc group. The peptides were purified via RP-HPLC on a C18 column with a gradient of water–acetonitrile (0.1% trifluoroacetic acid). The purity was assessed to be above 95% by analytical HPLC. Peptide mass was confirmed using electro-spray ionization mass spectrometry. The peptides were lyophilized, sealed, and stored at -20 ℃. Samples for analytical studies were prepared by dissolving the peptide at the appropriate concentration in aqueous MES buffer (10 mM 2-(*N*-morpholino)ethanesulfonic acid, pH 6.0). Peptide solutions were dialyzed against buffer solution to remove the remaining trifluoroacetic acid (MWCO = 2,000 Da). The isotopically labeled peptide, 1-^{13}C-Ala(12),^{15}N-Leu(16),3-^{13}C-Ala(24)-**7HSAP1** (**7HSAP1***), was synthesized, purified, and assembled using an identical procedure. The protected amino acids Fmoc-1-^{13}C-alanine, Fmoc-3-^{13}C-alanine, and Fmoc-^{15}N-leucine were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA). Labeled amino acids were introduced manually during microwave peptide synthesis. Mass shifts due to the incorporation of the labels were confirmed using ESI-mass spectrometry.

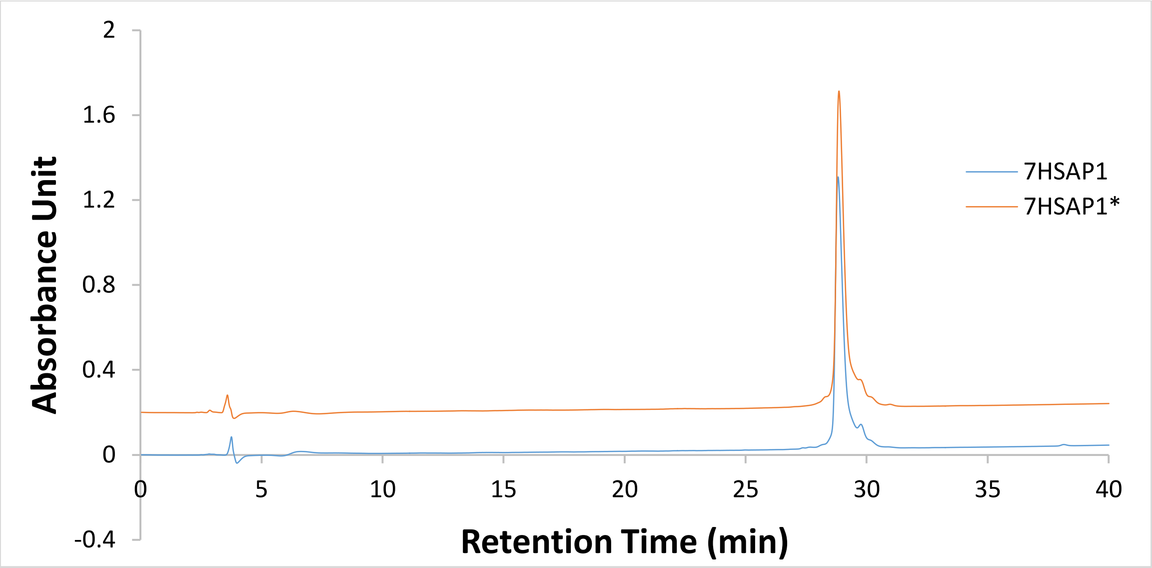
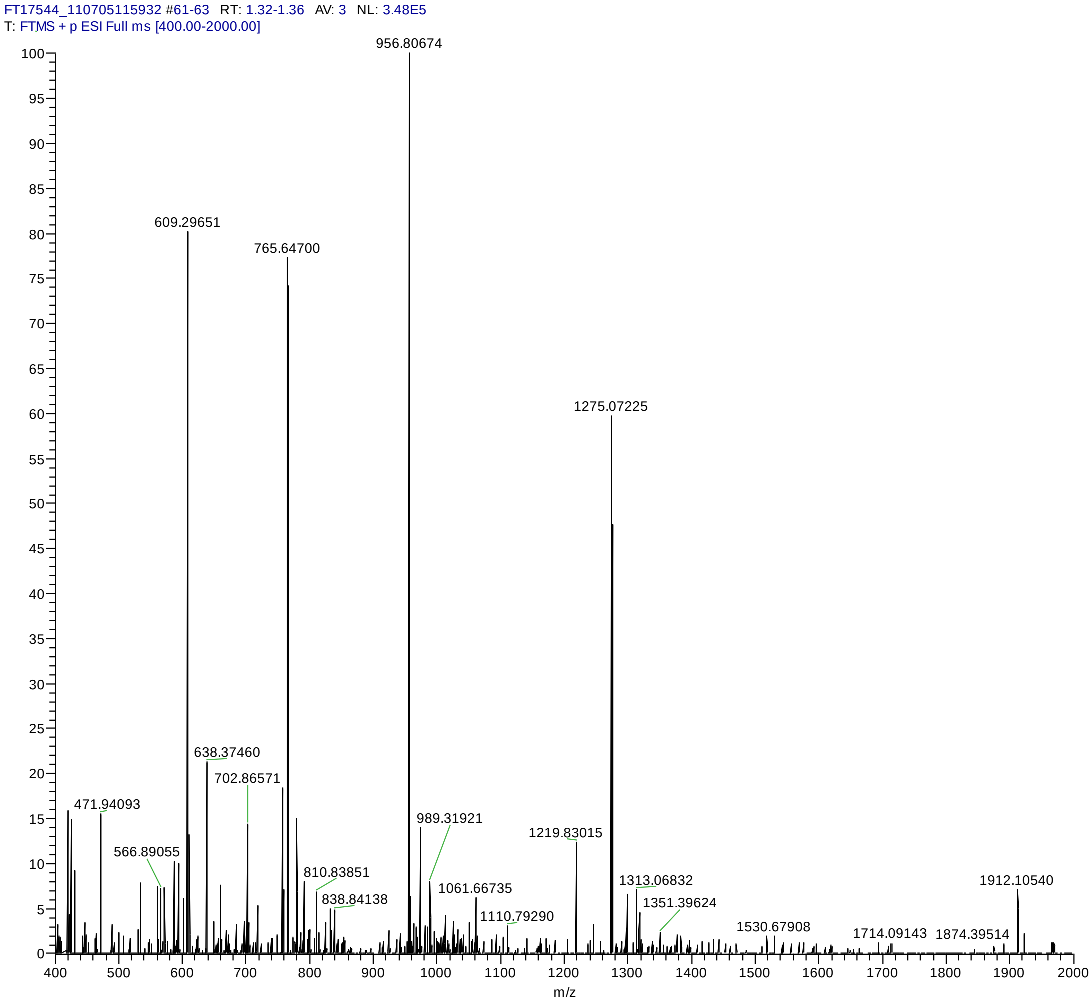


Figure 3.16. Analytical HPLC traces of 7HSAP1 and 7HSAP1*.



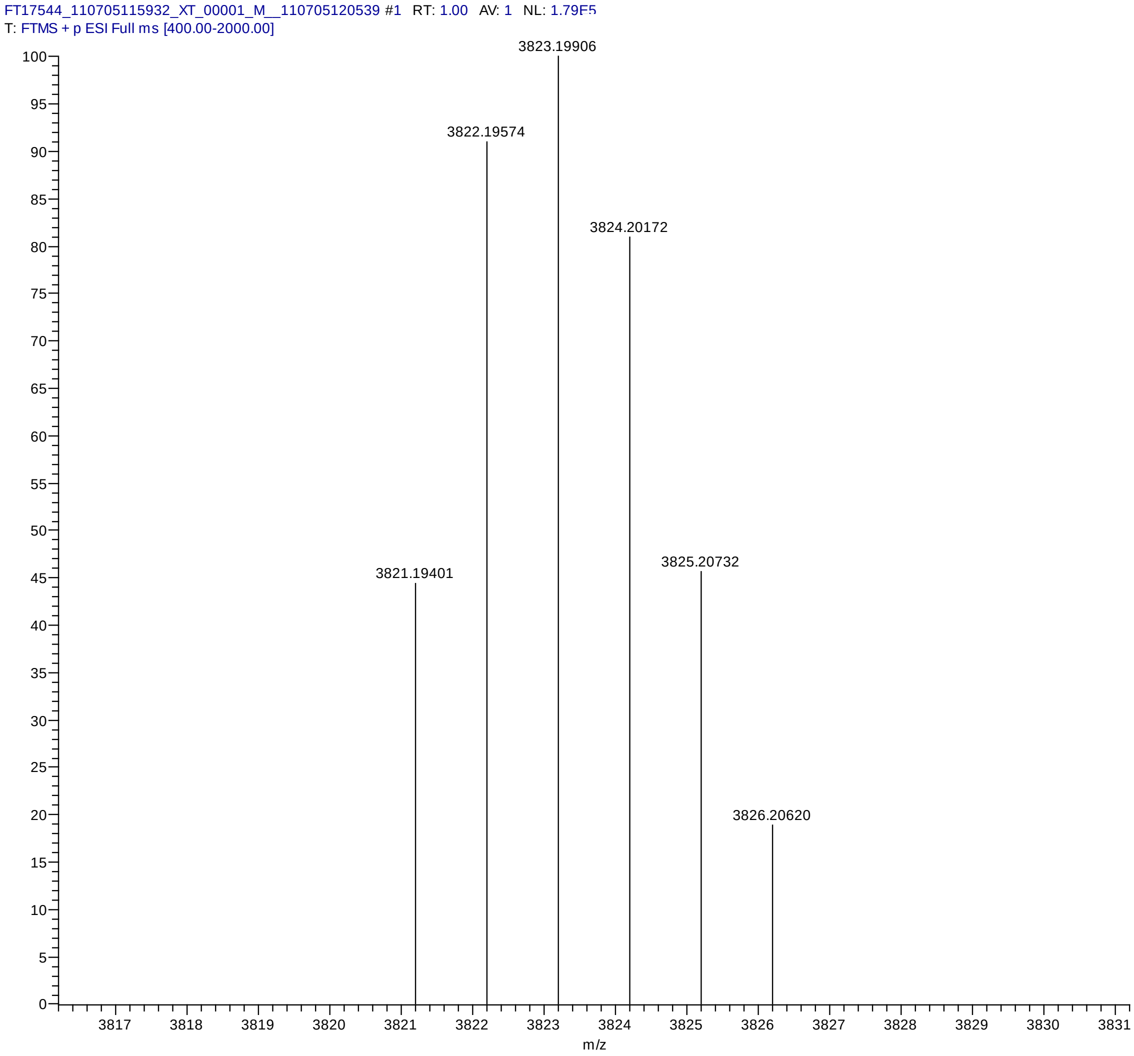
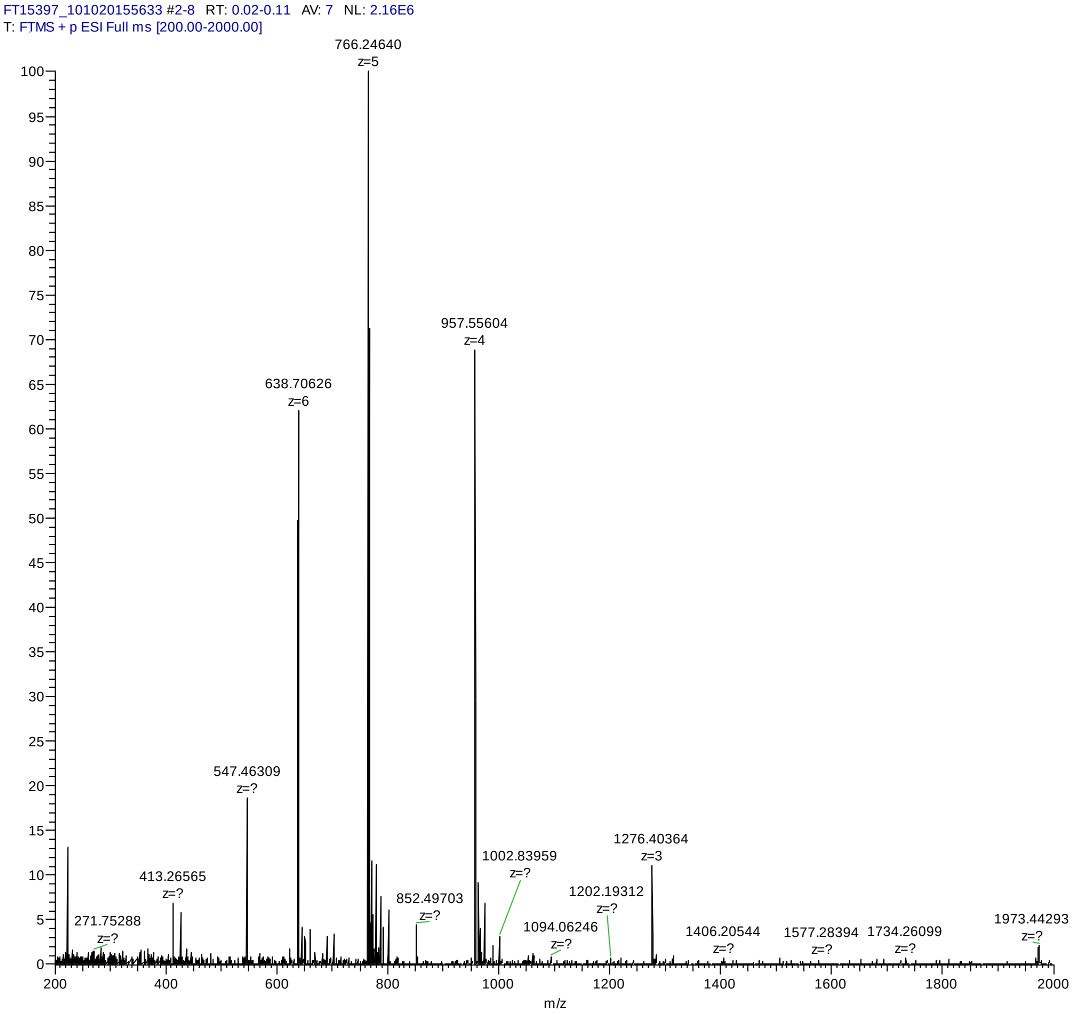


Figure 3.17. ESI-mass spectrometry of purified 7HSAP1.


FT15397_101020155633 #2-8 RT: 0.02-0.11 AV: 7 NL: 2.16E6
T: FTMS + p ESI Full ms [200.00-2000.00]
766.24640
z=5
957.55604
z=4
638.70626
z=6
547.46309
z=?
413.26565
z=?
271.75288
z=?
852.49703
z=?
1002.83959
z=?
1094.06246
z=?
1202.19312
z=?
1276.40364
z=3
1406.20544
z=?
1577.28394
z=?
1734.26099
z=?
1973.44293
z=?
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
200
400
600
800
1000
1200
1400
1600
1800
2000
m/z

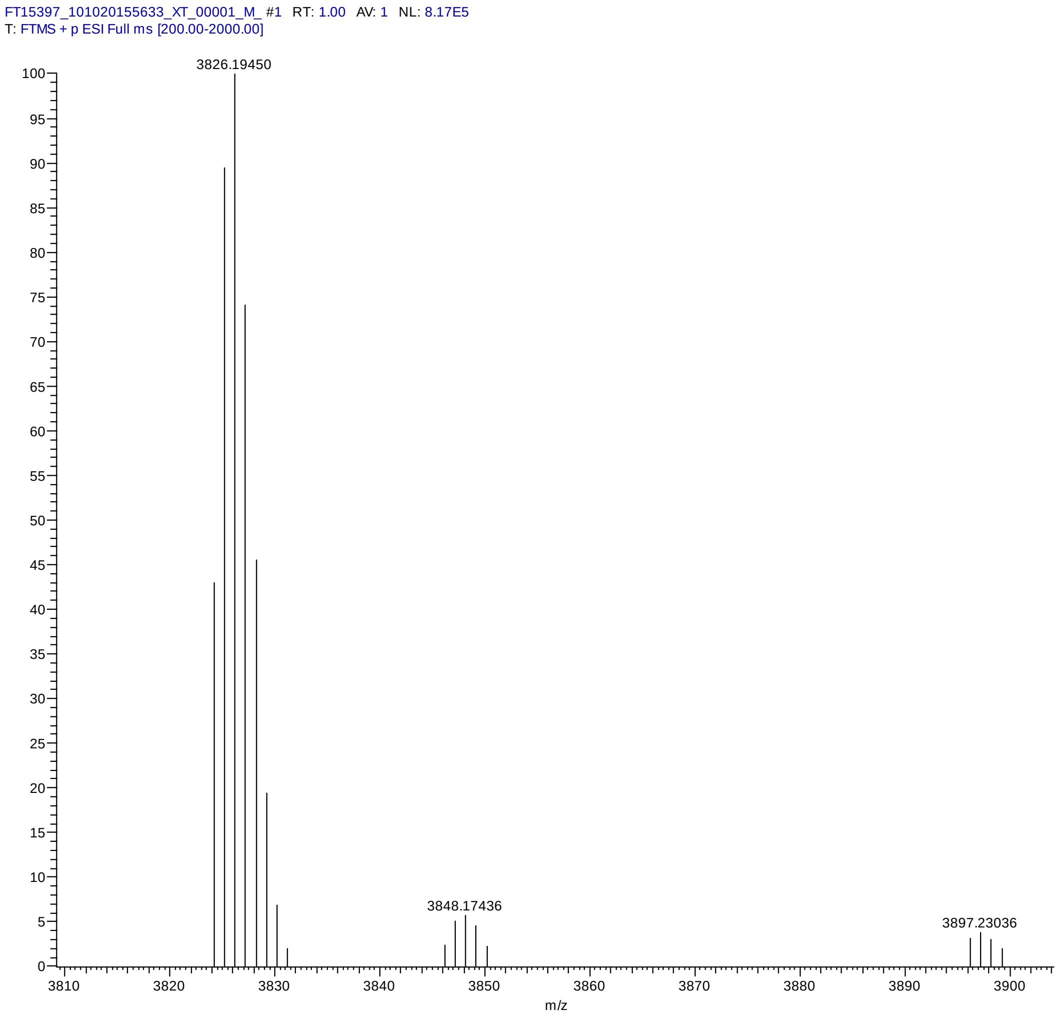


Figure 3.18. ESI-mass spectrometry of purified 7HSAP1*.

Circular Dichroism Spectropolarimetry. CD spectra were recorded on a Jasco J-810 CD spectropolarimeter in 0.10 mm quartz cells at a concentration of 100 μM in MES buffer (10 mM, pH 6.0). Spectra were recorded from 260 to 190 nm at a scanning rate of 100 nm/min and a resolution of 0.5 nm. Peptide concentration was determined spectrophotometrically from measurement of the absorbance at 280 nm (A_{280}). For peptides containing Tyr, Trp or Cys residues, the peptide concentration can be calculated from equation (1):

$$\mathrm{MW}\times {}^{A_{280}}/_{c} = 1280n_Y + 5690n_W + 120n_C \qquad (1)$$

in which c is the concentration of peptide in mg/mL, and n_Y, n_W, and n_C are the numbers of tyrosine, tryptophan and cysteine residues, respectively, in the peptide sequence [104]. As **7HSAP1** and **GCN4-pAA** and its derivatives contain only a single tyrosine residue per molecule, then $c = \mathrm{MW} \times A_{280}/1280$. To eliminate error in determination of absorbance that could arise as a result of UV light scattering due to peptide self-assembly, aqueous solutions of peptide were mixed with 6 M guanidinium chloride in 1:9 v/v ratio and heated at 100 ℃ for 15 min to completely denature the sample prior to performing the absorbance measurements.

Flow Linear Dichroism Spectroscopy. Flow linear dichroism spectra were recorded on a JASCO J-810 circular dichroism spectropolarimeter using a microvolume cuvette with a path length of 50 µm and a rotation speed of 3000 rpm to establish Couette flow.[105] The background scattering for each sample was obtained from the LD spectra of samples at 0 rpm. The LD spectra were measured after 15 min of rotation.

Transmission Electron Microscopy. TEM specimens were prepared from aqueous **7HSAP1** solution in MES buffer (10 mM, pH 6.0). The samples were deposited onto 200 mesh carbon coated copper grids from Electron Microscopy Sciences (Hatfield, PA). After a 30-second incubation period, excess liquid was wicked away and the specimens were stained with either a 1:1 mixture (v/v) of the negative stains NanoVan (methylamine vanadate, 2%) and Nano-W (methylamine tungstate, 2%) from Nanoprobes, Inc. (Yaphank, NY) or a solution of uranyl acetate (2%). Excess stain was wicked away after incubation on the grid for 1 min. The sample grids were dried under vacuum and stored in a desiccator. TEM measurements were acquired on a Hitachi H-7500 or H-7100 transmission electron microscope at an accelerating voltage of 75 kV or 100 kV, respectively. The micrographs were recorded at magnifications of 200,000 or 40,000 using a Gatan CCD digital camera.

Cryo-TEM. Aliquots of the protein fiber solution were applied onto glow-discharged, 200 mesh, Quantifoil grids and plunge frozen into liquid ethane using a Mark III Vitrobot (FEI, Hillsboro, OR). Images were collected on a JEOL JEM-2200FS 200 kV field emission transmission electron microscope with an in-column Omega energy filter. The images were energy filtered with a slit width of 20 eV. Images were collected under low dose conditions with the sample maintained at a temperature of circa –177 ℃. Images were captured on a high sensitivity Gatan US4000 4k x 4k Ultrascan CCD camera with minimal defocus applied.

Scanning Transmission Electron Microscopy. The STEM data were acquired at Brookhaven National Laboratory (BNL). The STEM instrument operates at 40 keV with a scanning probe of < 0.3 nm diameter produced from a cold field-emission source. Every electron emerging from the specimen is detected by one of the scintillator-photomultiplier detectors collecting 0-15 mRadian (bright field), 15-40 MRadian (small-angle dark field) and 40-200 mRadian (large-angle dark field). The large-angle signal is proportional to the mass of atoms in the path of the beam. Specimen quality and mass calibration are checked by detailed comparison of the image to the known structure of tobacco mosaic virus (TMV).

Specimens are deposited on thin carbon (circa 2 nm thick) supported on a thicker holey carbon film mounted on a titanium grid using the wet-film, hanging-drop method http://www.bnl.gov/biology/stem/SpecPrepDetails.asp. TMV is added to the grid first as an internal control, followed by injection buffer, then specimen solution (in 10 mM MES buffer, pH 6.0) for 1 min, then 10 washes of 20 mM ammonium acetate pH 7.0. Excess solution is wicked from the edge with filter paper between each injection. After the last wash the grid is wicked to a thin layer (ca. 1 µm), fast frozen by plunging into liquid nitrogen shush and stored under liquid nitrogen. Grids are freeze-dried overnight in an ion-pumped chamber with an efficient cold trap and transferred under vacuum to the STEM cold stage (-160 ℃). Imaging typically uses a dose of 20 el/Å^2 (causing < 5% mass loss, corrected by comparison to TMV). Mass measurements are performed off-line with customized software (PCMass, available at ftp.stem.bnl.gov). The program masks out objects significantly above background and computes the value for the thin carbon in the remaining areas, which is subtracted, and pixels within the contour of filaments are summed and divided by length to give mass-per-unit-length. Accuracy is determined by cleanliness of the background between objects and by counting statistics of the scattered electrons. At 10 el/Å^2, TMV SD (standard deviation) is ideally circa 1 % and SD of filaments of 6 kDa/nm is circa 20% for a single segment 10 nm long. For TMV the program provides automatic searching and measurement but for the thin filaments the low S/N requires manual selection whereupon the software "locks on" to a segment giving angle, offset and mass per unit length. PCMass also provides statistics in a database for individual images or groups of images.

Solid-state NMR Measurements. The labeled peptide, [1-^{13}C]A12-, [3-^{13}CH$_3$]A24-, [^{15}N]L16-**7HSAP1** (**7HSAP1***), was assembled either in aqueous solution at neutral pH or in the presence of 25 mM MES adjusted to neutral pH. The assembled peptide was pelleted by centrifugation and dried *in vacuo*. Each NMR sample (25-50 mg) was packed into a 4 mm solid-state NMR rotor and centered using boron nitride spacers. NMR spectra were collected with a Bruker (Billerica, MA) Avance 600 spectrometer using a 4 mm HCN BioSolids magic-angle spinning (MAS) probe. MAS frequency was actively controlled at 10,000 ± 2 Hz with cooling and spinning air exit temperature maintained below -1 ℃ to ensure MAS and radio-frequency (RF) heating did not denature the samples. ^{13}C (150.8 MHz) CP-MAS spectra before and after REDOR experiments confirmed that the samples did not change during the experiment. The pulse sequence for ^{13}C[^{15}N] rotational-echo double-resonance (REDOR) [75] consists of two parts, an S sequence with both ^{13}C and ^{15}N pulses, and the S_0 sequence that is identical but does not contain any ^{15}N dephasing pulses. Applying pulses to the dephasing ^{15}N spins interferes with the averaging due to magic-angle spinning and reintroduces the dipolar coupling which is observed in the REDOR S spectrum, where the signal decays according to both T_2 (spin-spin relaxation) and the heteronuclear ^{13}C-^{15}N dipolar coupling. Maximum dephasing occurs when the spacing between π-pulses is equal to ½ of the rotor-period (Tr). The sequence without ^{15}N dephasing π -pulses gives the REDOR full-echo or S_0 spectra, where the magnetization decays according to only T_2. The difference between the REDOR S and S_0 signal (ΔS) is directly proportional to the dipolar coupling, hence the distance between the two spins.

All REDOR data were collected with the ^{13}C of interest placed on-resonance with xy8 ^{13}C[^{15}N]REDOR pulse sequence[106, 107] with dephasing ^{15}N 8 μs π-pulses centered at Tr/2 and refocusing ^{13}C 4 μs π-pulses centered at Tr and EXORCYCLE phase cycling[108, 109] of the final Hahn-echo 4 μs ^{13}C pulse. The xy8 phase cycling compensated for pulse imperfections in both the ^{13}C and ^{15}N REDOR π-pulses. The initial ^{1}H 90 ° pulse was 1.9 μs, ^{1}H CP RF fields were ramped from 50 to 70 kHz and the ^{13}C cross-polarization (CP) RF field was kept constant at 50 kHz. A 128 kHz Spinal64 ^{1}H (600.3 MHz) decoupling[110] was applied during REDOR evolution and acquisition. SPINAL64 pulse-widths and ^{1}H decoupling resonance frequencies were optimized by comparing the peak heights of the CH and CH_2 resonances of fumaric acid monoethyl ester.

The long REDOR evolution times required to measure the weak dipolar couplings between the Ala $^{13}CH_3$ carbons and the peptide backbone ^{15}N's are susceptible to RF homogeneity,[109, 111] which can lead to a lower REDOR dephasing plateau than predicted with stronger dipolar couplings. Therefore, to determine proper level for ^{13}C π-pulses the power level was arrayed in REDOR S_0 pulse sequence at long REDOR evolution times (>50 ms typically corresponding to ca. 500 4 μs ^{13}C π-pulses) and choosing the power level that corresponded to the maximum signal intensity.[112] Similarly, ^{15}N 180 ° pulses were determined by arraying the ^{15}N power level using the REDOR S experiment at REDOR evolution times corresponding to a $\Delta S/S_0$ between 0.3 and 0.5.[112]

REDOR data points are the integrated sum of center- and sideband peaks. Error bars were calculated using the noise of each spectrum as the maximum peak height deviation. To normalize for the decay due to T2, individual REDOR curves are plotted as $\Delta S/S_0$. The ideal scaling factor was determined with ^{13}C[^{15}N]REDOR spectra of [1-^{13}C,^{15}N]alanine diluted 10:1 with natural abundance alanine. This dilution ratio does not create a perfectly isolated spin-pair, due to the probability of having an enriched 13C/15N alanine as a next nearest neighbor. Correcting for this, the probe REDOR scaling factor was determined to be 93%. BS-REDOR[76] fits account for the effects of natural abundance ^{13}C contribution to ^{13}CO and $^{13}CH_3$ REDOR S_0 signal and natural abundance ^{15}N dephasing contribution to REDOR S signal intensity.

X-ray Fiber Diffraction and Alignment. Fiber samples were aligned by suspending a droplet of fibril suspension between two wax-tipped glass capillaries (0.7 mm) and air-dried at room temperature. Aligned fiber samples were mounted on a goniometer head and data was collected on a home source Pilatus 6M detector with a Saturn 944+ CCD. Exposure times were 1-15 sec and the specimen-to-detector distance was 100 mm. Diffraction data were examined using Clearer[74] and X-ray signals measured. Calculated diffraction patterns were generated from input model coordinates (**7HSAP1_**assembly_MD.pdb or **7HSAP1_**centralbundle.pdb) that had first been aligned such the helices ran approximately parallel to the vertical axis of the fiber. The coordinates were placed within a specified unit cell and calculations were performed using default values[74] with a crystallite size of 400 nm^3 and a sampling size of 1 pixel. Diffraction patterns were compared to experimental data both by visual comparison of the patterns and radial scans were imported into Microsoft Excel and compared as graphical traces.

Small- and Wide-Angle X-ray Scattering Measurements. Synchrotron SAXS/WAXS measurements were performed at the 12-ID-B beamline of Advanced Photon Source at Argonne National Laboratory. A SAXS/WAXS simultaneous setup was utilized and the sample-to-detector distances were set such that the overall scattering momentum transfer q range was achieved from 0.003-2.4 $Å^{-1}$, where $q = 4\pi \sin \theta/\lambda$, 2θ denoting the scattering angle and λ the x-ray wavelength. The wavelength was set at 1.033 Å during the measurements. Scattered X-ray intensities were measured using a Pilatus 2M (DECTRIS Ltd) detector for SAXS and Pilatus 300K for WAXS. SAXS/WAXS measurements were performed on aqueous solutions of peptide **7HSAP1** at concentrations of 0.5 or 1 mM in MES buffer (10 mM, pH 6.0) at 25 ℃. A flow cell equipped with a quartz capillary (1.5 mm diameter) was used to prevent the radiation damage. Twenty images were collected for each sample and buffer. The 2-D scattering images were converted to 1-D SAXS curves through azimuthally averaging after solid angle correction and then normalizing with the intensity of the transmitted x-ray beam, using the software package at beamline 12ID-B. The 1-D curves of the samples were averaged and subtracted with the background measured from the corresponding buffers. The R_c analysis was done in Igor Pro software (WaveMatrics, Inc.) using the following modified Guinier equation[72]: $\ln[q * I(q)] = -R_c^2 q^2/2$, where I(q) is the scattering intensity at momentum transfer q. The scattering simulations on hollow cylindrical geometry models were performed in Matlab (MathWorks, Inc.), details described elsewhere.[49] The scattering simulation on the molecular model (**7HSAP1**_assembly_MD.pdb) was performed using program CRYSOL.[113] The SAXS molecular envelope calculations were performed using program DAMMIN.[73] To reduce the impact of the length heterogeneity and the electron density fluctuation with the molecule, SAXS data with q range of 0.03-0.56 Å^{-1} was used in the DAMMIN calculations. Due to the intrinsic degeneracy in SAXS data, a tight hollow cylindrical search space with R_{in} of 3, R_{out} of 15 and height of 100 Å was employed for all calculations. Twenty independent individual calculations were performed and the 3-D shape results were further averaged to yield the final envelope.

Fluorescence Spectroscopy.

PRODAN (6-propionyl-2-(*N*,*N*-dimethylamino)naphthalene) was purchased from Sigma-Aldrich (St. Louis, MO). Fluorescence spectra were recorded on a Fluoromax-3 spectrophotometer (Horiba Scientific, Inc.) using a quartz cell with a 1-cm path-length (Starna Cells, Inc) at ambient temperature (ca. 22 °C). Emission and excitation slit widths were set to 5 nm. The excitation wavelength was set to 371 nm. Before each measurement, an aliquot (1 μL) of a PRODAN stock solution (1 mM in ethanol) was mixed with 1 mL peptide solution at the appropriate concentration to obtain a PRODAN concentration of 1 μM. The mixtures were equilibrated at ambient temperature for 6 h before taking fluorescence measurements. Fluorescence spectra were recorded over the wavelength range of 400 nm to 650 nm.

Computational Analyses. The programs MSMS[54] and CASTp[53] were employed to analyze the solvent-accessible surface area and internal void volume, respectively, of the 7-helix bundle structure of **GCN4-pAA** (PDB ID: 2HY6). The solvent-accessible surface area calculation was performed using MSMS on the NIH server.[114] CASTp calculations were performed using the web-based resource at the University of Illinois, Champaign-Urbana.[115] A modified version of the samCC [77] program was used to quantify the deviation of **GCN4-pAA** crystal structure from a reference 7-helix bundle. The reference structure was modeled as canonical coiled-coil adopting knobs-into-holes packing using the BeamMotifCC program.[81] SamCC not only provided the detailed information on the local variations in **GCN4-pAA** but also allowed the generation of an idealized model that is in agreement with the reference structure (2HY6.ideal.pdb).

The Crick Coiled-Coil Parameterization (CCCP) program[84] was employed through the web-based server at Dartmouth University.[85] The pdb files for **GCN4-pAA** and its ideal, in-register structure were uploaded into the structure fitter module of CCCP. Residues 4-31 of the corresponding structure of **GCN4-pAA** (PDB ID: 2HY6) and the corresponding residues of its idealized counterpart (2HY6.ideal.pdb) were fit using this program. The structural parameters generated from analysis of 2HY6.ideal.pdb were loaded into the structure generator of the CCCP program with the exception that ΔZ_{offset} values were derived from the wild-type structure of **GCN4-pAA**. The resulting model (Ala35_7.pdb) consisted of a 7-helix bundle comprising 35-residue peptides based on a polyalanine backbone. The same parameters were employed in CCCP to create a 7-helix bundle structure consisting exclusively of 175-residue (i.e., five repeats of 35-residues) polyalanine sequences. This structure was employed as the input for subsequent molecular modeling studies

Molecular Models. The structural model consisting of the 7-helix bundle polyalanine sequence of 175 residues was split into peptides of 35 residues in length, which were arranged as five separate 7-helix bundles. Side chains were mutated to match the **7HSAP1** sequence and the resultant 7-helix bundle structure was energy minimized in Macromodel (Schrodinger, Inc.) with 200 steps of steepest decent. The resulting assembly was placed in a simulation box with explicit waters and neutralizing chloride anions. Molecular dynamics simulation was performed for 1.2 ns using the Desmond simulation package[116] and the OPLS-AA/SPC force field.[117] PDB files of the stacked assembly (**7HSAP1_**assembly_MD.pdb) and the central 7-helix bundle (**7HSAP1_**centralbundle.pdb) provided models for determination of distance measurements at the displaced edge between helices *A* and *G*, for comparison to REDOR

NMR measurements on **7HSAP***, and for simulations of solution SAXS/WAXS measurements and X-ray fiber diffraction data.

3.5 References

1. Loquet, A., B. Habenstein, and A. Lange, *Structural Investigations of Molecular Machines by Solid-State NMR.* Accounts of chemical research, 2013.

2. Zhou, Z.H., *Atomic resolution cryo electron microscopy of macromolecular complexes.* Adv Protein Chem Struct Biol, 2011. **82**: p. 1-35.

3. Kojima, S., et al., *Fibril formation by an amphipathic alpha-helix-forming polypeptide produced by gene engineering.* Proceedings of the Japan Academy Series B-Physical and Biological Sciences, 1997. **73**(1): p. 7-11.

4. Pandya, M.J., et al., *Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis.* Biochemistry, 2000. **39**(30): p. 8728-8734.

5. Ryadnov, M.G. and D.N. Woolfson, *Engineering the morphology of a selfassembling protein fibre.* Nature Materials, 2003. **2**(5): p. 329-332.

6. Ryadnov, M.G. and D.N. Woolfson, *Introducing branches into a self-assembling peptide fiber.* Angewandte Chemie-International Edition, 2003. **42**(26): p. 3021-3023.

7. Papapostolou, D., et al., *Engineering nanoscale order into a designed protein fiber.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(26): p. 10853-10858.

8. Sharp, T.H., et al., *Cryo-transmission electron microscopy structure of a gigadalton peptide fiber of de novo design.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(33): p. 13266-13271.

9. Ogihara, N.L., et al., *Design of three-dimensional domain-swapped dimers and fibrous oligomers.* Proceedings of the National Academy of Sciences, 2001. **98**(4): p. 1404-1409.

10. Potekhin, S., et al., *De novo design of fibrils made of short α-helical coiled coil peptides.* Chemistry & biology, 2001. **8**(11): p. 1025-1032.

11. Zimenkov, Y., et al., *Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif.* Tetrahedron, 2004. **60**(34): p. 7237-7246.

12. Zimenkov, Y., et al., *Rational design of a reversible pH-responsive switch for peptide self-assembly.* J Am Chem Soc, 2006. **128**(21): p. 6770-1.

13. Dublin, S.N. and V.P. Conticello, *Design of a selective metal ion switch for self-assembly of peptide-based fibrils.* Journal of the American Chemical Society, 2008. **130**(1): p. 49-+.

14. Wagner, D.E., et al., *Toward the development of peptide nanofilaments and nanoropes as smart materials.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(36): p. 12656-12661.

15. Dong, H., S.E. Paramonov, and J.D. Hartgerink, *Self-assembly of alpha-helical coiled coil nanofibers.* Journal of the American Chemical Society, 2008. **130**(41): p. 13691-13695.

16. Aggeli, A., et al., *Responsive gels formed by the spontaneous self-assembly of peptides into polymeric beta-sheet tapes.* Nature, 1997. **386**(6622): p. 259-62.

17. Aggeli, A., et al., *pH as a trigger of peptide beta-sheet self-assembly and reversible switching between nematic and isotropic phases.* J Am Chem Soc, 2003. **125**(32): p. 9619-28.

18. Aggeli, A., et al., *Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta -sheet tapes, ribbons, fibrils, and fibers.* Proc Natl Acad Sci U S A, 2001. **98**(21): p. 11857-62.

19. Zhang, S., et al., *Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3334-8.

20. Marini, D.M., et al., *Left-handed helical ribbon intermediates in the self-assembly of a beta-sheet peptide.* Nano Letters, 2002. **2**(4): p. 295-299.

21. Janek, K., et al., *Water-soluble beta-sheet models which self-assemble into fibrillar structures.* Biochemistry, 1999. **38**(26): p. 8246-52.

22. Matsumura, S., S. Uemura, and H. Mihara, *Fabrication of nanofibers with uniform morphology by self-assembly of designed peptides.* Chemistry, 2004. **10**(11): p. 2789-94.

23. Dong, H., et al., *Self-assembly of multidomain peptides: balancing molecular frustration controls conformation and nanostructure.* J Am Chem Soc, 2007. **129**(41): p. 12468-72.

24. Bowerman, C.J., et al., *Tuning β-sheet peptide self-assembly and hydrogelation behavior by modification of sequence hydrophobicity and aromaticity.* Biomacromolecules, 2011. **12**(7): p. 2735-2745.

25. Swanekamp, R.J., et al., *Coassembly of Enantiomeric Amphipathic Peptides into Amyloid-Inspired Rippled β-Sheet Fibrils.* Journal of the American Chemical Society, 2012. **134**(12): p. 5556-5559.

26. Choo, D.W., et al., *Nucleated antiparallel β-sheet that folds and undergoes self-assembly: a template promoted folding strategy toward controlled molecular architectures.* Macromolecules, 1996. **29**(1): p. 355-366.

27. Schneider, J.P., et al., *Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide.* Journal of the American Chemical Society, 2002. **124**(50): p. 15030-15037.

28. Pochan, D.J., et al., *Thermally reversible hydrogels via intramolecular folding and consequent self-assembly of a de novo designed peptide.* Journal of the American Chemical Society, 2003. **125**(39): p. 11802-11803.

29. Nagarkar, R.P., et al., *De novo design of strand-swapped β-hairpin hydrogels.* Journal of the American Chemical Society, 2008. **130**(13): p. 4466-4474.

30. O'Leary, L.E., et al., *Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel.* Nature chemistry, 2011. **3**(10): p. 821-828.

31. Xu, F., et al., *Compositional control of higher order assembly using synthetic collagen peptides.* Journal of the American Chemical Society, 2011. **134**(1): p. 47-50.

32. Kar, K., et al., *Aromatic interactions promote self-association of collagen triple-helical peptides to higher-order structures.* Biochemistry, 2009. **48**(33): p. 7959-68.

33. Rele, S., et al., *D-periodic collagen-mimetic microfibers.* Journal of the American Chemical Society, 2007. **129**(47): p. 14780-7.

34. Przybyla, D.E. and J. Chmielewski, *Metal-triggered radial self-assembly of collagen peptide fibers.* Journal of the American Chemical Society, 2008. **130**(38): p. 12610-12611.

35. Cejas, M.A., et al., *Collagen-related peptides: Self-assembly of short, single strands into a functional biomaterial of micrometer scale.* Journal of the American Chemical Society, 2007. **129**(8): p. 2202-2203.

36. Cejas, M.A., et al., *Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions.* Proceedings of the National Academy of Sciences, 2008. **105**(25): p. 8513-8518.

37. Yamazaki, C.M., et al., *Artificial collagen gels via self-assembly of de novo designed peptides.* Peptide Science, 2008. **90**(6): p. 816-823.

38. Koide, T., et al., *Self-complementary peptides for the formation of collagen-like triple helical supramolecules.* Bioorganic & medicinal chemistry letters, 2005. **15**(23): p. 5230-5233.

39. Kotch, F.W. and R.T. Raines, *Self-assembly of synthetic collagen triple helices.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(9): p. 3028-3033.

40. Bruning, M., et al., *Bipartite Design of a Self-Fibrillating Protein Copolymer with Nanopatterned Peptide Display Capabilities.* Nano Letters, 2010. **10**(11): p. 4533-4537.

41. King, N.P., et al., *Computational design of self-assembling protein nanomaterials with atomic level accuracy.* Science, 2012. **336**(6085): p. 1171-1174.

42. Fletcher, J.M., et al., *Self-assembling cages from coiled-coil peptide modules.* Science, 2013. **340**(6132): p. 595-9.

43. Liu, J., et al., *A seven-helix coiled coil.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(42): p. 15457-15462.

44. Valéry, C., F. Artzner, and M. Paternostre, *Peptide nanotubes: molecular organisations, self-assembly mechanisms and applications.* Soft Matter, 2011. **7**(20): p. 9583-9594.

45. Bong, D.T., et al., *Self-assembling organic nanotubes.* Angewandte Chemie-International Edition, 2001. **40**(6): p. 988-1011.

46. Ghadiri, M.R., et al., *Self-Assembling Organic Nanotubes Based on a Cyclic Peptide Architecture.* Nature, 1993. **366**(6453): p. 324-327.

47. Vauthey, S., et al., *Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles.* Proceedings of the National Academy of Sciences, 2002. **99**(8): p. 5355-5360.

48. Reches, M. and E. Gazit, *Casting metal nanowires within discrete self-assembled peptide nanotubes.* Science, 2003. **300**(5619): p. 625-627.

49. Lu, K., et al., *Exploiting amyloid fibril lamination for nanotube self-assembly.* Journal of the American Chemical Society, 2003. **125**(21): p. 6391-6393.

50. Valéry, C., et al., *Biomimetic organization: Octapeptide self-assembly into nanotubes of viral capsid-like dimension.* Proceedings of the National Academy of Sciences, 2003. **100**(18): p. 10258-10262.

51. Morris, K.L., et al., *The structure of cross-beta tapes and tubes formed by an octapeptide, alphaSbeta1.* Angew Chem Int Ed Engl, 2013. **52**(8): p. 2279-83.

52. O'Shea, E.K., et al., *X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil.* Science, 1991. **254**(5031): p. 539-544.

53. Binkowski, T.A., S. Naghibzadeh, and J. Liang, *CASTp: computed atlas of surface topography of proteins.* Nucleic Acids Research, 2003. **31**(13): p. 3352-3355.

54. Sanner, M.F., A.J. Olson, and J.C. Spehner, *Reduced surface: an efficient way to compute molecular surfaces.* Biopolymers, 1996. **38**(3): p. 305-20.

55. Bennett, M.J., M.R. Sawaya, and D. Eisenberg, *Deposition diseases and 3D domain swapping.* Structure, 2006. **14**(5): p. 811-824.

56. Ponstingl, H., K. Henrick, and J.M. Thornton, *Discriminating between homodimeric and monomeric proteins in the crystalline state.* Proteins: Structure, Function, and Bioinformatics, 2000. **41**(1): p. 47-57.

57. Richardson, J.S. and D.C. Richardson, *Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation.* Proceedings of the National Academy of Sciences, 2002. **99**(5): p. 2754-2759.

58. Pedersen, J.S., G. Christensen, and D.E. Otzen, *Modulation of S6 fibrillation by unfolding rates and gatekeeper residues.* Journal of molecular biology, 2004. **341**(2): p. 575-588.

59. Rousseau, F., L. Serrano, and J.W. Schymkowitz, *How evolutionary pressure against protein aggregation shaped chaperone specificity.* Journal of molecular biology, 2006. **355**(5): p. 1037-1047.

60. Monsellier, E. and F. Chiti, *Prevention of amyloid-like aggregation as a driving force of protein evolution.* EMBO reports, 2007. **8**(8): p. 737-742.

61. Marrington, R., M. Seymour, and A. Rodger, *A new method for fibrous protein analysis illustrated by application to tubulin microtubule polymerisation and depolymerisation.* Chirality, 2006. **18**(9): p. 680-90.

62. Rodger, A., et al., *Looking at long molecules in solution: what happens when they are subjected to Couette flow?* Physical Chemistry Chemical Physics, 2006. **8**(27): p. 3161-3171.

63. Bulheller, B.M., et al., *Flow linear dichroism of some prototypical proteins.* Journal of the American Chemical Society, 2009. **131**(37): p. 13305-13314.

64. Baxa, U., et al., *In Sup35p filaments (the [PSI+] prion), the globular C-terminal domains are widely offset from the amyloid fibril backbone.* Molecular microbiology, 2011. **79**(2): p. 523-532.

65. Goldsbury, C., et al., *Amyloid structure and assembly: insights from scanning transmission electron microscopy.* Journal of structural biology, 2011. **173**(1): p. 1-13.

66. Sen, A., et al., *Mass analysis by scanning transmission electron microscopy and electron diffraction validate predictions of stacked β-solenoid model of HET-s prion fibrils.* Journal of Biological Chemistry, 2007. **282**(8): p. 5545-5550.

67. Hodgkinson, J., et al., *E. 563 P. Morris, JS Wall, SM Lea, and AJ Blocker. 2009. Three-dimensional 564 reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry 565 and novel features throughout.* Nat Struct Mol Biol. **16**: p. 477-85.

68. Baxa, U., et al., *Architecture of Ure2p Prion Filaments THE N-TERMINAL DOMAINS FORM A CENTRAL CORE FIBER.* Journal of Biological Chemistry, 2003. **278**(44): p. 43717-43727.

69. Zaccai, N.R., et al., *A de novo peptide hexamer with a mutable channel.* Nature Chemical Biology, 2011. **7**(12): p. 935-941.

70. Burton, A.J., et al., *Accessibility, Reactivity, and Selectivity of Side Chains within a Channel of de Novo Peptide Assembly.* Journal of the American Chemical Society, 2013. **135**(34): p. 12524-12527.

71. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria.* Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.

72. Glatter, O. and O. Kratky, *Small-angle X-ray Scattering*. 1982, London: Academic Press.

73. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophysical Journal, 1999. **76**(6): p. 2879-2886.

74. Sumner Makin, O., P. Sikorski, and L.C. Serpell, *CLEARER: a new tool for the analysis of X-ray fibre diffraction patterns and diffraction simulation from atomic structural models.* Journal of Applied Crystallography, 2007. **40**(5): p. 966-972.

75. Gullion, T. and J. Schaefer, *Rotational-Echo Double-Resonance Nmr.* Journal of Magnetic Resonance, 1989. **81**(1): p. 196-200.

76. Gehman, J.D., et al., *Boltzmann statistics rotational-echo double-resonance analysis.* Journal of Physical Chemistry B, 2007. **111**(27): p. 7802-7811.

77. Walshaw, J. and D.N. Woolfson, *SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures.* Journal of Molecular Biology, 2001. **307**(5): p. 1427-1450.

78. Testa, O.D., E. Moutevelis, and D.N. Woolfson, *CC plus : a relational database of coiled-coil structures.* Nucleic Acids Research, 2009. **37**: p. D315-D322.

79. *The CC+ Database*. Available from: http://coiledcoils.chm.bris.ac.uk/ccplus/search/.

80. Dunin-Horkawicz, S. and A.N. Lupas, *Measuring the conformational space of square four-helical bundles with the program samCC.* Journal of Structural Biology, 2010. **170**(2): p. 226-35.

81. Offer, G., M.R. Hicks, and D.N. Woolfson, *Generalized crick equations for modeling noncanonical coiled coils.* Journal of Structural Biology, 2002. **137**(1-2): p. 41-53.

82. Crick, F.H., *The Fourier transform of a coiled-coil.* Acta Crystallographica, 1953. **6**(8-9): p. 685-689.

83. Crick, F.H.C., *The Packing of Alpha-Helices - Simple Coiled-Coils.* Acta Crystallographica, 1953. **6**(8-9): p. 689-697.

84. Grigoryan, G. and W.F. DeGrado, *Probing designability via a generalized model of helical bundle geometry.* Journal of molecular biology, 2011. **405**(4): p. 1079-1100.

85. *CCCP (Coiled-coil Crick Parameterization).* Available from: http://arteni.cs.dartmouth.edu/cccp/.

86. Aree, T., et al., *Crystal structure of octakis (2, 3, 6-tri-< i> O</i>-methyl)-γ-cyclodextrin ·4.5 H< sub> 2</sub> O: evidence for conformational flexibility of permethylated cyclodextrins.* Carbohydrate Research, 2000. **328**(3): p. 399-407.

87. Lindner, K. and W. Saenger, *Crystal structure of the γ-cyclodextrin n-propanol inclusion complex; Correlation of α-, β-, γ-cyclodextrin geometries.* Biochemical and Biophysical Research Communications, 1980. **92**(3): p. 933-938.

88. Weber, G. and F.J. Farris, *Synthesis and spectral properties of a hydrophobic fluorescent probe: 6-propionyl-2-(dimethylamino) naphthalene.* Biochemistry, 1979. **18**(14): p. 3075-3078.

89. Sun, S., et al., *6-Propionyl-2-(N, N-dimethylamino) naphthalene (PRODAN) revisited.* Applied spectroscopy, 1997. **51**(9): p. 1316-1322.

90. Crane, N.J., et al., *Cyclodextrin inclusion complexes with a solvatochromic fluorescent probe.* Journal of Chemical Education, 2002. **79**(10): p. 1261-1263.

91. Korten, T., A. Månsson, and S. Diez, *Towards the application of cytoskeletal motor proteins in molecular detection and diagnostic devices.* Current opinion in biotechnology, 2010. **21**(4): p. 477-488.

92. Tang, J.H., et al., *Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22.* Structure, 2011. **19**(4): p. 496-502.

93. Shnyrova, A.V., et al., *Geometric catalysis of membrane fission driven by flexible dynamin rings.* Science, 2013. **339**(6126): p. 1433-6.

94. Dinu, C.Z., et al., *Tubulin encapsulation of carbon nanotubes into functional hybrid assemblies.* Small, 2009. **5**(3): p. 310-315.

95. Kumara, M.T., et al., *Bioengineered flagella protein nanotubes with cysteine loops: Self-assembly and manipulation in an optical trap.* Nano Letters, 2006. **6**(9): p. 2121-2129.

96. Hou, C., et al., *Construction of GPx active centers on natural protein nanodisk/nanotube: A new way to develop artificial nanoenzyme.* ACS nano, 2012. **6**(10): p. 8692-8701.

97. Lee, B.Y., et al., *Virus-based piezoelectric energy generation.* Nat Nanotechnol, 2012. **7**(6): p. 351-6.

98. Miller, R.A., A.D. Presley, and M.B. Francis, *Self-assembling light-harvesting systems from synthetically modified tobacco mosaic virus coat proteins.* J Am Chem Soc, 2007. **129**(11): p. 3104-9.

99. Ballister, E.R., et al., *In vitro self-assembly of tailorable nanotubes from a simple protein building block.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3733-8.

100. Miranda, F.F., et al., *A self-assembled protein nanotube with high aspect ratio.* Small, 2009. **5**(18): p. 2077-84.

101. Salgado, E.N., et al., *Metal templated design of protein interfaces.* Proceedings of the National Academy of Sciences, 2010. **107**(5): p. 1827-1832.

102. Brodin, J.D., et al., *Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays.* Nat Chem, 2012. **4**(5): p. 375-82.

103. Liu, Y. and B. Kuhlman, *RosettaDesign server for protein design.* Nucleic acids research, 2006. **34**(suppl 2): p. W235-W238.

104. Gill, S.C. and P.H. Von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data.* Analytical biochemistry, 1989. **182**(2): p. 319-326.

105. Marrington, R., et al., *Validation of new microvolume Couette flow linear dichroism cells.* Analyst, 2005. **130**(12): p. 1608-1616.

106. Gullion, T., D.B. Baker, and M.S. Conradi, *New, compensated carr-purcell sequences.* Journal of Magnetic Resonance (1969), 1990. **89**(3): p. 479-484.

107. Christensen, A.M. and J. Schaefer, *Solid-state NMR determination of intra-and intermolecular phoshorus-31-carbon-13 distances for shikimate 3-phosphate and [1-13C] glyphosate bound to enolpyruvylshikimate-3-phosphate synthase.* Biochemistry, 1993. **32**(11): p. 2868-2873.

108. Rance, M. and R.A. Byrd, *Obtaining high-fidelity spin-12 powder spectra in anisotropic media: Phase-cycled Hahn echo spectroscopy.* Journal of Magnetic Resonance (1969), 1983. **52**(2): p. 221-240.

109. Sinha, N., K. Schmidt-Rohr, and M. Hong, *Compensation for pulse imperfections in rotational-echo double-resonance NMR by composite pulses and EXORCYCLE.* Journal of Magnetic Resonance, 2004. **168**(2): p. 358-365.

110. Fung, B.M., A.K. Khitrin, and K. Ermolaev, *An Improved Broadband Decoupling Sequence for Liquid Crystals and Solids.* Journal of Magnetic Resonance, 2000. **142**(1): p. 97-101.

111. Weldeghiorghis, T.K. and J. Schaefer, *Compensating for pulse imperfections in REDOR.* Journal of Magnetic Resonance, 2003. **165**(2): p. 230-236.

112. Ni, R., et al., *Remodeling Cross-beta Nanotube Surfaces with Peptide/Lipid Chimeras.* Angewandte Chemie-International Edition, 2012. **51**(27): p. 6635-6638.

113. Svergun, D., C. Barberato, and M. Koch, *CRYSOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates.* Journal of Applied Crystallography, 1995. **28**(6): p. 768-773.

114. *StrucTools - PDB coordinate calculations*. Available from: http://helixweb.nih.gov/structbio/basic.html.

115. *Computed Atlas of Surface Topography of proteins. CASTp calculation*. Available from: http://sts.bioengr.uic.edu/castp/calculation.php.

116. Bowers, K.J., et al. *Scalable algorithms for molecular dynamics simulations on commodity clusters*. in *SC 2006 Conference, Proceedings of the ACM/IEEE*. 2006. IEEE.

117. Kaminski, G.A., et al., *Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides.* The Journal of Physical Chemistry B, 2001. **105**(28): p. 6474-6487.

Chapter IV: *De Novo* Design of Helical Nanotubes from Coiled-coil Peptides

4.1 Introduction

Protein-based nanotubes represent attractive candidates for the construction of functional nano-porous materials with medically relevant applications in controlled release [1], directional transport [2], and catalysis [3]. In native biological systems, tubular structures derived from self-assembly of protein and peptide structural motifs occur frequently in functional roles that include ion channels [4, 5], small-molecule pumps [6], DNA/RNA delivery vehicles (i.e., viruses) [2, 7], and other directional transport systems.

Compared with their carbon or inorganically derived nanotubes, protein-based nanotubes embrace several advantages such as significantly less cytotoxicity and intrinsic biocompatibility. Furthermore, protein-based nanotubes hold the promise of ease of functionalization, modular molecular recognition [8], and reversible assembly/disassembly in response to physiological environmental cues. Additionally, wormlike, filamentous nanoparticles are shown to be better than spherical ones at avoiding immune responses allowing for longer circulation times due to the difficulty of macrophages have adjusting tertiary and/or quaternary structure to engulf such elongated particles [9, 10].

A diverse group of protein-based nanotubes have emerged in recent years using viral capsids [11, 12], amyloid peptide [13, 14], actin [15], tubulin [16], pili [17], and cyclic D/L peptides [18], and some interesting applications have been investigated, e.g. the engineering of tobacco mosaic virus to form an artificial light harvesting apparatus [19], and use of peptide nanotube and flagellin protein FliC nanotubes as template for synthesis of metal nanowires [20, 21].

However, the possibility of employing another ubiquitous structural motif, α-helical barrels, to fabricate tubular nano-systems has yet to be explored. α-Helical barrels are a class of membrane-spanning or membrane-associated protein domains that defines various pore and channel sizes ranging from angstrom to nanometer length scale [22] (Figure 4.1). They represent an intriguing prototype for the design of a series of protein-based nanotubes with tailorable channel dimensions and chemistry which hold promise for versatile applications. The significance of native α-helical barrels to de novo protein-based materials design have long been recognized, and the structure TolC has been analyzed theoretically to understand its sequence-structure relationships[23]. Unfortunately, these efforts have not led to successful de novo design of synthetic counterparts of α-helical barrels which have been predicted designable over a decade ago [24].

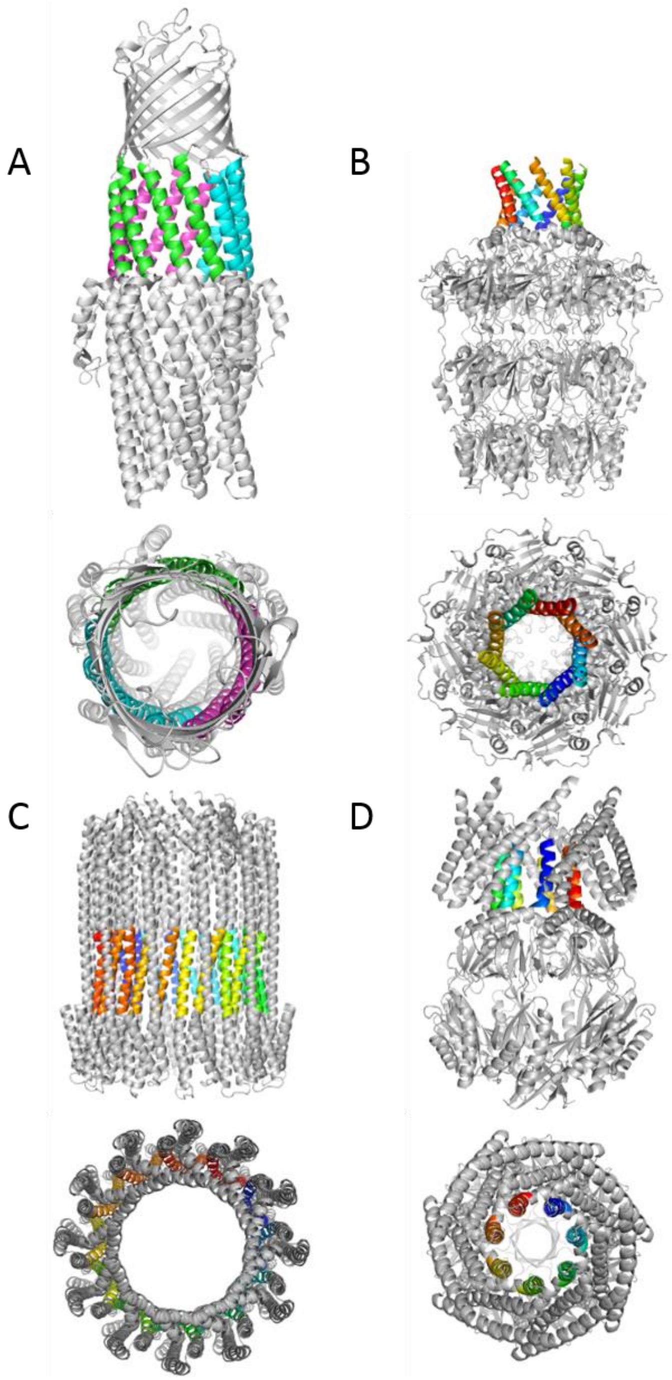


Figure 4.1. The α-helical barrel domains in native proteins defining various pore sizes. A: TolC bacterial transmembrane protein, PDB ID 1EK9. B: Wza the translocon for E. coli capsular polysaccharides, PDB ID 2J58. C: The cytotoxin ClyA, PDB ID 2WCD. D: E. coli mechanosensitive channel Msc5, PDB ID 2VV5.

4.2 Barrel3CLys, a bilayered helical nanotube

4.2.1 Sequence Design

The design **Barrel3CLys** peptide was based on our current understanding of the sequence characteristics of higher-order coiled coils. As described in Chapter 1, extended KIH packing contributes to the stabilization of the higher oligomeric state. It has been envisioned that α-helices with two hydrophobic interfaces which are separated by 154° could possibly form open (α-sheet) and closed (α-cylinder) multihelical assemblies [24](Figure 1.9). Based on these considerations and inspired by **GCN4-pAA** sequence in which alanines are employed to extend the hydrophobic interfaces [25], the sequence of **Barrel3CLys** was designed as depicted in Figure 4.2.

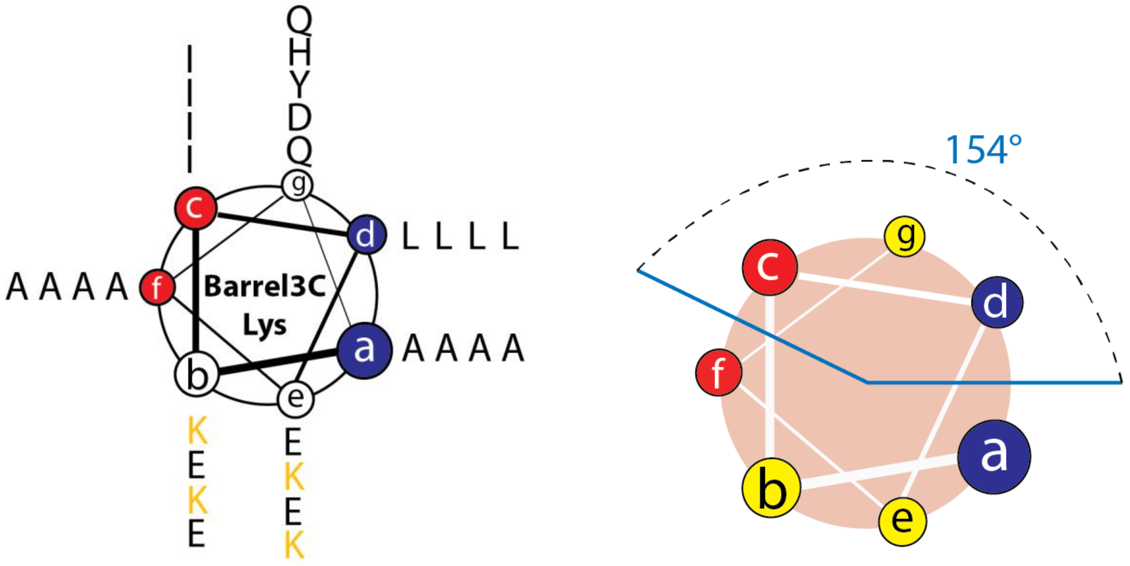


Figure 4.2. Left: The designed sequence of Barrel3CLys prejected on a helical wheel diagram. Right: The two hydrophobic interfaces are separated by 154°.

In the sequence, hydrophobic leucine and isoleucine residues are placed at c and d positions, respectively, in the heptad repeat, and alanine residues flank leucine and isoleucine at a and f positions. So both a, d and c, f seams are hydrophobic. The b, e positions are inhabited by charged residues where the pattern of charged residues is critical which requires explicit negative design against β-sheet formation. The pattern we use here is KEEKKEEK. There is more freedom in filling up g positions.

4.2.2 Results and Discussion

4.2.2.1 Circular Dichroism and Flow Linear Dichroism

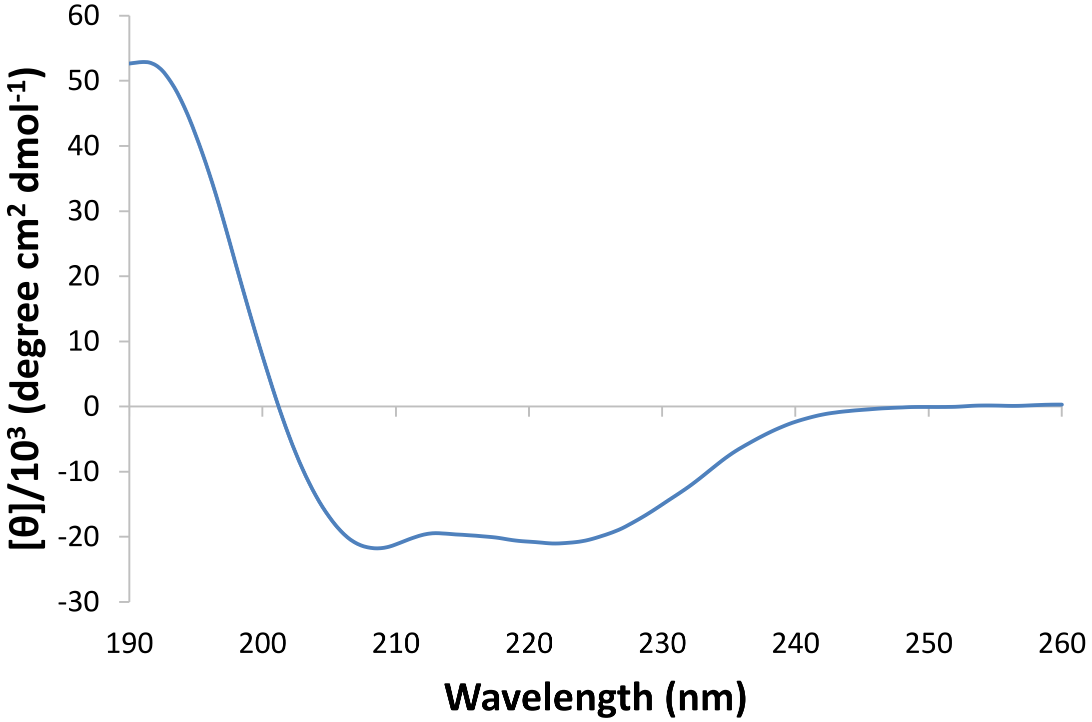


Figure 4.3. CD spectrum of peptide Barrel3CLys (100 µM) in acetate buffer (10 mM, pH 4.0). MRE is relatively weak indicating not full-length of peptide is in an α-helical conformation.

The conformation of **Barrel3CLys** peptide was initially interrogated with circular dichroism (CD) spectropolarimetry. The CD spectrum indicates the formation of α-helical secondary structure. However, the MRE (Mean Residue Ellipticity) values are relatively weak suggesting that part of sequence is melted under experimental conditions (100 μM peptide in 10 mM acetate buffer pH 4.0).

After observation of formation of tubular structures under TEM, flow linear dichroism [26-29] was employed under conditions of similar concentration to assess the degree of anisotropy that developed within the peptide samples under a flow alignment (5,000 rpm) in a Couette cell. Solutions of **Barrel3CLys** displayed an LD spectrum with a strong positive signal at 192 nm, and a weaker, broader positive signal at 220nm, which suggest that these electronic transitions are aligned parallel with the flow direction, hence the helices perpendicular with the long axis of peptide assemblies, since it arises from 222nm nπ* and 190nm ππ* transitions that are polarized perpendicular to the helical axis.

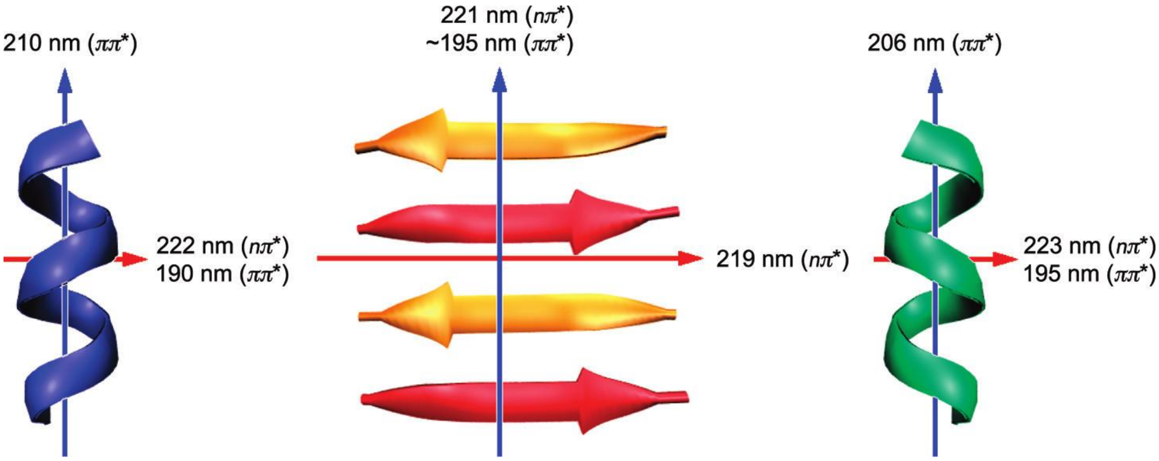


Figure 4.4. Polarizations of transitions in (left) an α-helix, (middle) a β-sheet, and (right) a PII-helix. [26] Note that 222nm nπ* and 190nm ππ* transitions are polarized perpendicular to the helical axis.

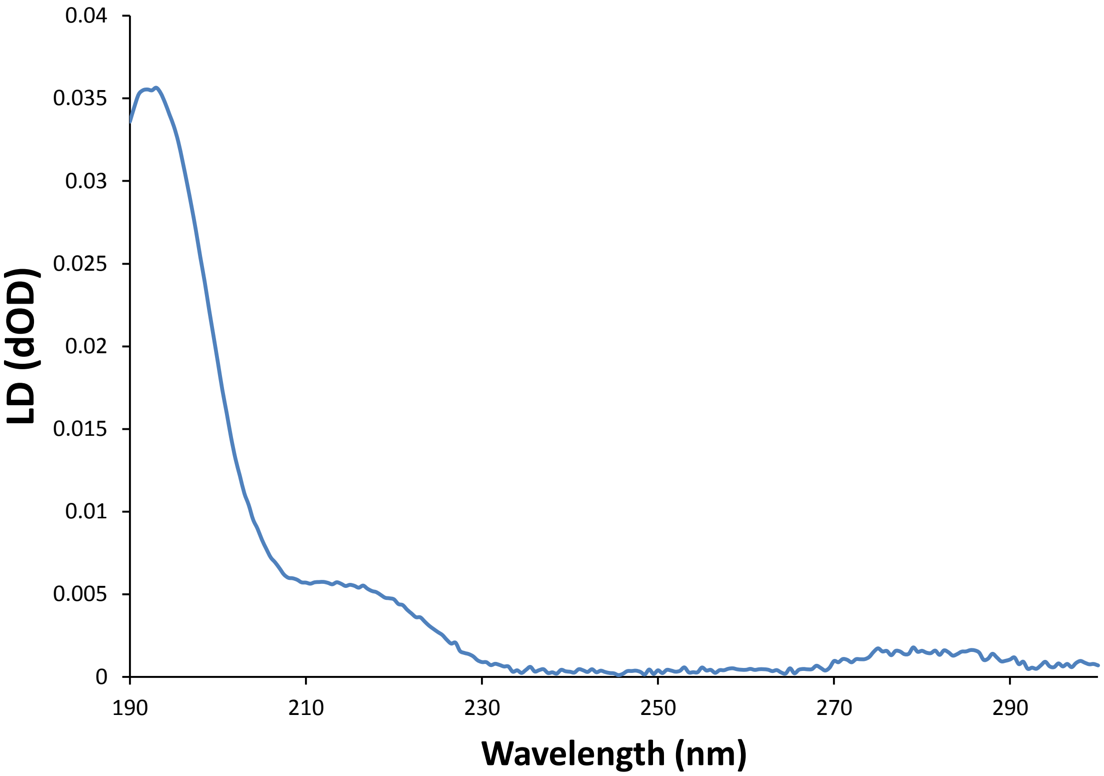


Figure 4.5. Flow LD spectrum of peptide Barrel3CLys (50 µM) in acetate buffer (10 mM, pH 4.0) showing that 222 nm and 190 nm transitions are aligned with the flow direction, hence peptides oriented perpendicular with flow direction.

4.2.2.2 Negative-stain Electron Microscopy and Mass-per-Length Measurements

Transmission Electron Microscopy (TEM) analysis (Figure 4.6) of **Barrel3CLys** in acetate buffer (10 mM, pH 4.0) revealed the formation of tubular structures with a diameter of 10 nm along coexisting with thinner fibrils (4 nm in diameter). The large nanotubes could be isolated from fibrils using preparative ultracentrifuge equipped with a Beckman 70.1 Ti rotor with a speed of 32,000 rpm for 45 minutes.

When observed using higher-resolution negative-stain Scanning Transmission Electron Microscopy (STEM), the termini of the tubular assemblies unwound into fibrils with a diameter about 4 nm. This observation provided very important macroscopic information regarding the structure of tubular assemblies which are formed by two or more fibrils wrapping around each other. However, at this stage, no information about molecular-level packing within the supramolecular assemblies was obtained.

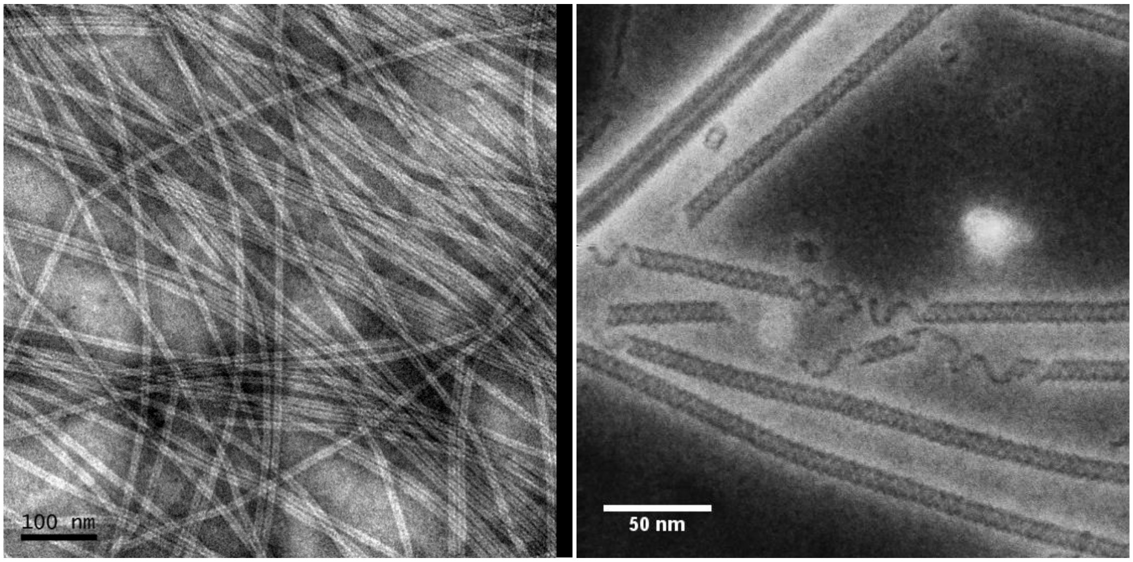
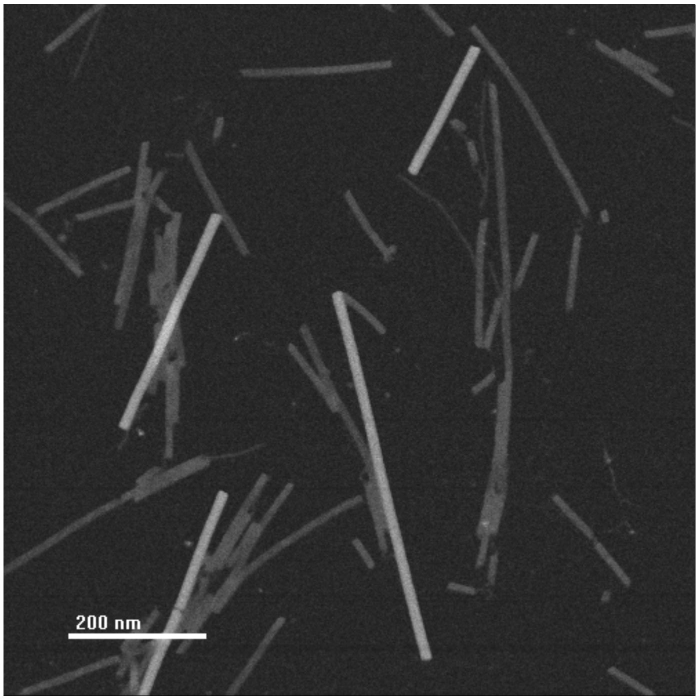


Figure 4.6. Negative-stain TEM (left) and STEM (right) of Barrel3CLys nanotubes. The diameter of these nanotubes is 10nm. Higher-resolution STEM image shows that the ends of the nanotubes unwind into 2 or more fibrils (circa 4 nm).

STEM mass-per-length (MPL) measurements [30, 31] were performed on **Barrel3CLys** nanotubes and fibrils imaged using dark-field STEM of unstained, freeze-dried specimens. Tobacco mosaic virus (TMV) particles were employed as an internal mass standard. A MPL value of ca. 3300 Da/Å was determined for the 10 nm nanotubes, and ca. 700 Da/Å for 4 nm fibrils, after normalization based on MPL measurements of the TMV mass standard. The MPL value for the tubes corresponds to ca. 1 Å axial rise



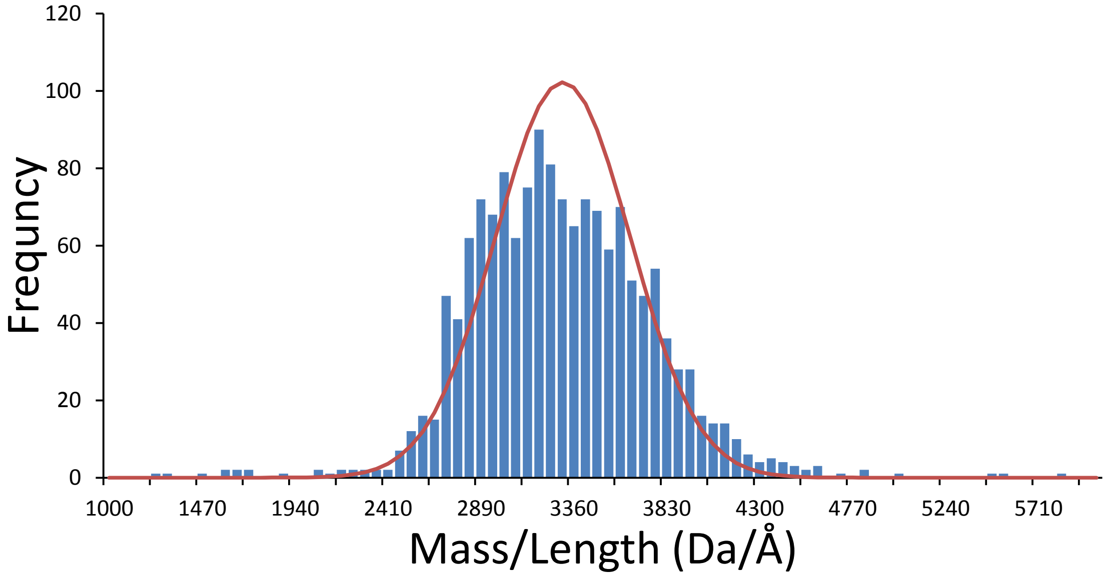


Figure 4.7. STEM mass-per-length measurements of Barrel3CLys nanotubes. Top: A typical STEM image used for mass-per-length (MPL) measurements. Tobacco Mosaic Viruses are used as internal reference for mass measurements. The majority of other filaments are Barrel3CLys nanotubes with thinner fibrils. Bottom: The MPL histogram. The averaged MPL for Barrel3CLys nanotubes is 3300 Da/Å

per peptide which was further confirmed by helical symmetry determination by cryo-EM. The MPL value for thin fibrils corresponds to a helical bilayer confirmation very well in which helical axes are oriented perpendicular to fibrils.

4.2.2.3 Small Angle X-ray Scattering

Small- and wide-angle X-ray scattering (SAXS/WAXS) data were collected on aqueous solutions of **Barrel3CLys** nanotubes in acetate buffer (10 mM, pH 4.0) to interrogate the structural hierarchy of the resultant assemblies at length scales over two decades. The small-angle range, $q < 0.2$ Å^{-1} reflects the global shape of the molecule, while the middle q range, ca. $0.2 < q < 1.1$ Å^{-1}, reflects the tertiary fold of the assembly [32, 33].

In the small-angle region, the experimental scattering intensities at $q < 0.1$ Å^{-1} roughly follow the q^{-1} power law (Figure 4.8), indicating the assembly in solution has rod- or cylinder-like form. The pair distance distribution function (PDDF) [34] in Figure 4.9 derived from the SAXS data has a characteristic long tail, which further confirms the cylindrical shape. A cross-section PDDF derived from $q*I(q)$ exhibits a bimodal form, indicating the cylindrical molecular assembly is hollow. The largest dimension of the cross-section is about 105 Å. The average wall thickness and the diameter of the middle of the cylindrical shell can be estimated from the two maxima of the PDDF as ca. 30 and 65 Å, respectively, and the diameter of the hollow channel is about 35 Å. The value of R_c, the radius of gyration of the rod cross-section, was obtained as 40.7 Å through the fitting of the modified Guinier equation [33], $I(q) = \pi q I(q = 0)\exp(-0.5 R_c^2 q^2)$. The discrepancy between the data and rod-like q^{-1} power law at $q < 0.01$ Å^{-1} reflects the inhomogeneity of the assembly length in solution.

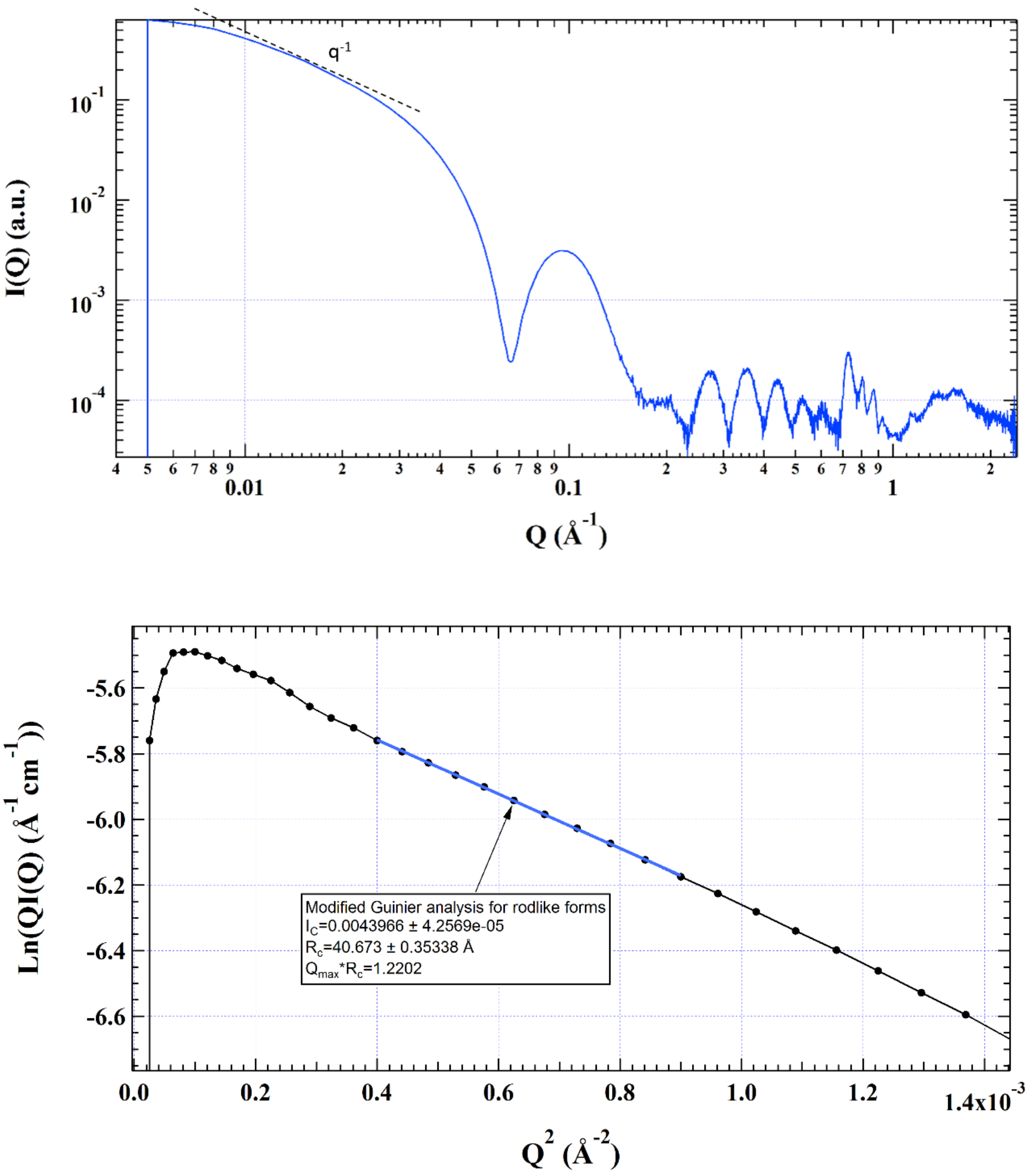
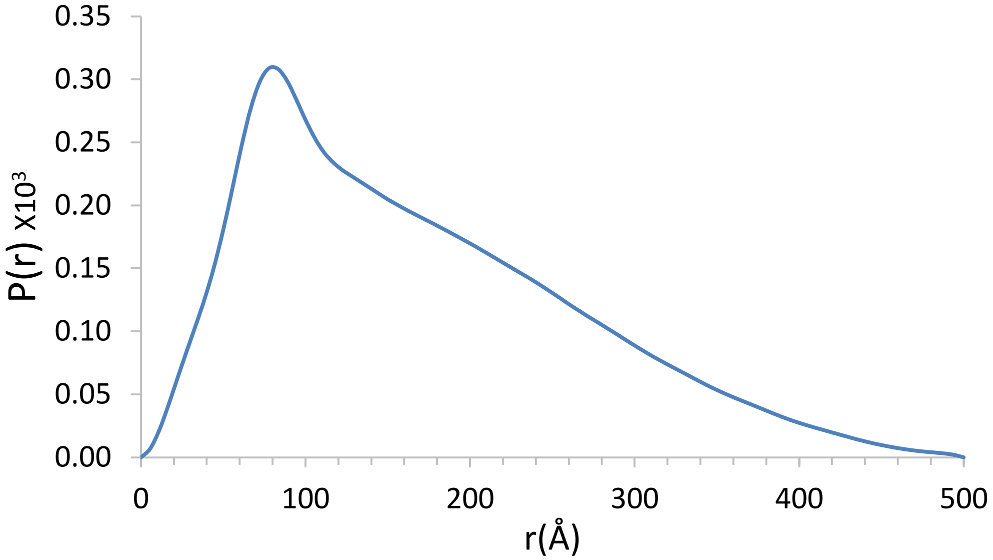


Figure 4.8. Top: SAXS and WAXS profile of Barrel3CLys nanotubes. The scattering intensity at $q < 0.1$ Å^{-1} roughly follows the q^{-1} power law indicating the assembly in solution has rod-like form. Guinier Analysis for rod-like forms (bottom) gives a cross sectional radius of gyration of 40.7 Å.



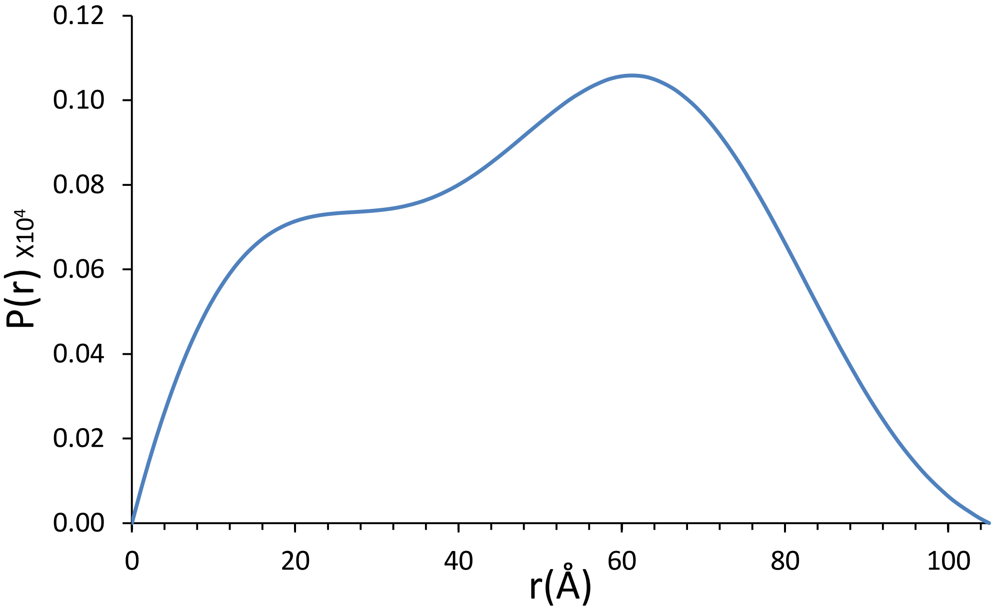


Figure 4.9. The pair distribution function (PDDF) of Barrel3CLys nanotubes. Top: PDDF calculated from the SAXS (*I(q)* versus *q*) data using program GNOM [34] further confirms the cylindrical shape. Bottom: The cross section PDDF calculated from SAXS data (*q*I(q)* versus *q*) using GNOM. The bimodal form indicates the cylindrical molecular assembly is hollow.

4.2.2.4 Cryo-EM Helical Reconstruction

In the past several decades, electron microscopy has played a central role in the elucidation of the structure of helical polymers which are usually not suitable for X-ray crystallography because rare filamentous polymers can pack with crystalline symmetry [35, 36]. Helical filaments were the first structures that to be reconstructed in three dimensions using electron microscopic images [37-40]. The development of cryogenic-temperature electron microscopy facilitates high-resolution structure determination and allows imaging macromolecular assemblies in near-native conditions. Additionally, the recent advancement in Direct Detection camera enables near-atomic-resolution structural determination.

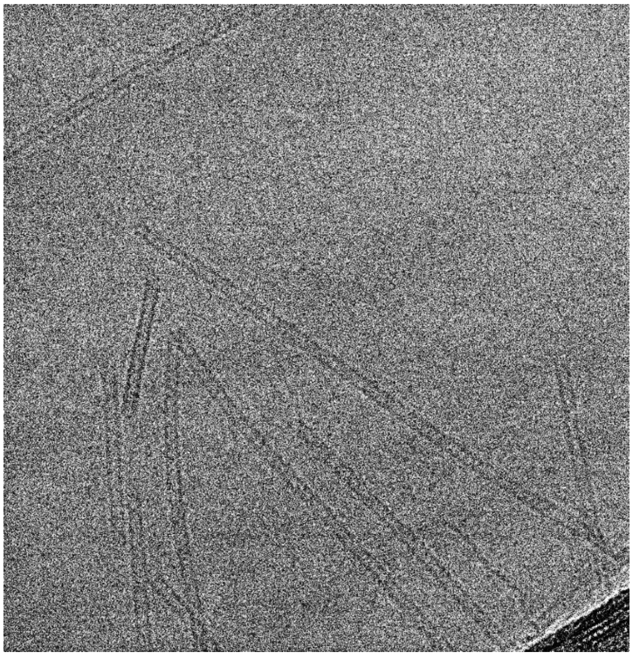


Figure 4.10. A typical cryo-EM image of Barrel3CLys nanotubes.

Cryo-EM images of **Barrel3CLys** nanotubes were collected on a high-resolution microscope (FEI Company) equipped with Direct Detection camera. 3D helical reconstruction was then performed on images of **Barrel3CLys** nanotubes using IHRSR algorithm [41] generating a density map at 5Å resolution. The results illustrated that the wall of tubes are composed of a bilayer of helices which confirmed our postulation according to STEM MPL measurements, and the helices are oriented roughly perpendicular to the tube axis, consistent with the LD measurements. The helical symmetry of the tube was shown to have a rotation of 115.5 °and an axial rise of 2 Å per asymmetric subunits which contains two peptides, one at inner radius and one at outer radius. The 1 Å axial rise per peptide matches the results of STEM MPL measurements very well.

Although side chains are not clearly visible at 5 Å resolution, some structural information could be speculated reasonably. Judging from the crossing angle between adjacent helices, it is very likely that both interactions within each layer of helices and between inner and outer walls are in a coiled-coil fashion, which agrees with the original design guideline. It was expected that α-helical barrels would form with the helices aligned along the long axis of the multihelical assembly. However, the smaller sized alanines that occupy the hydrophobic positions may cause the interface to open up and form coiled-coil ribbons instead of closed barrels. In the cryo-EM structure, we can clearly identify that three coiled-coiled ribbons, each of which is a bilayer of helices, wind with each other to form the tube and that the contact between ribbons are formed by end-to-end helix-helix interaction. Therefore, it is likely that the hydrogen bonding between the terminal glutamine residues holds the structure together. These speculations could potentially be verified by constructing and refining an atomic model into the cryo-EM density map, which is currently ongoing.

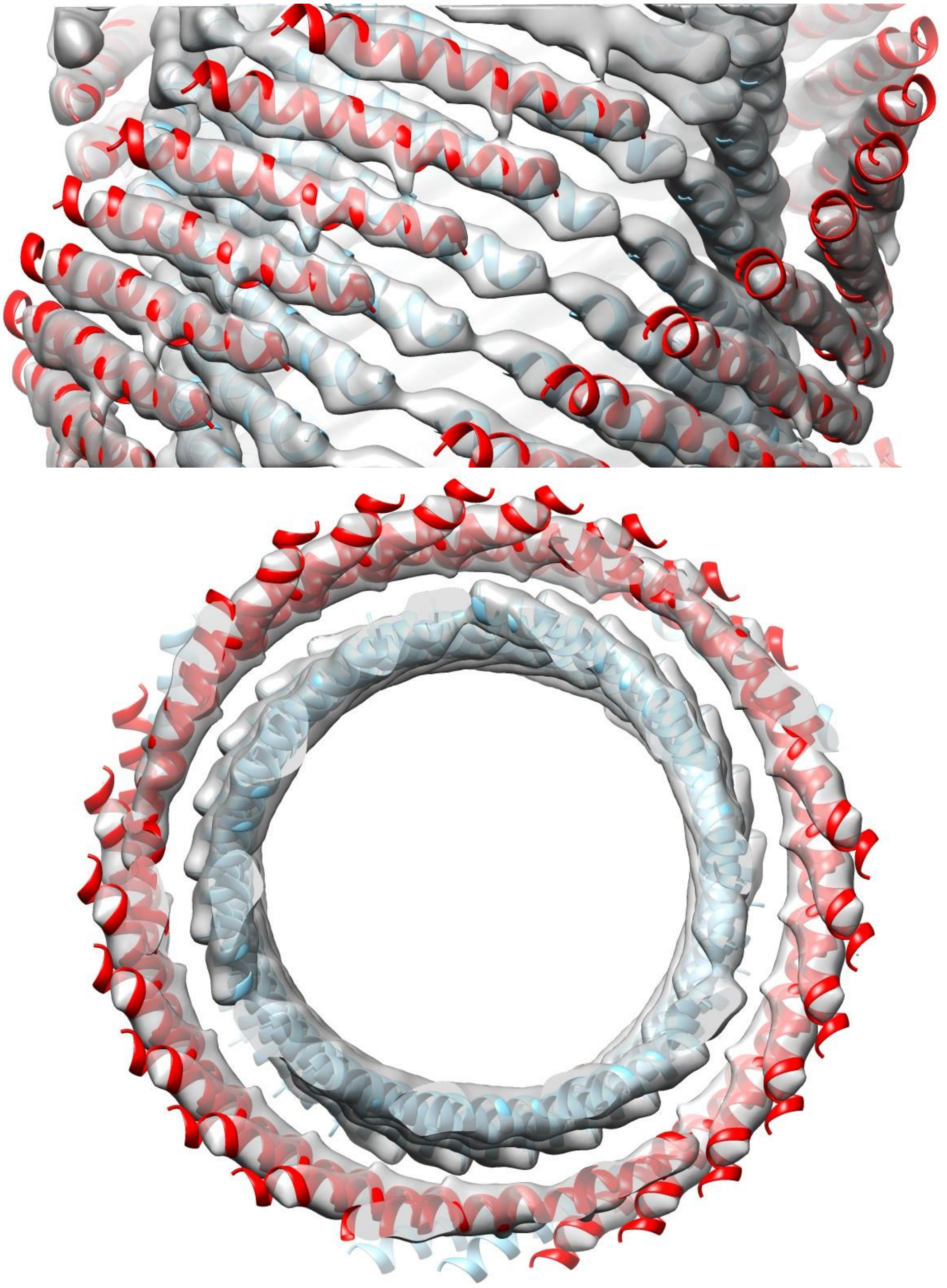


Figure 4.11. Side and top view of cryo-EM structure of unprecedented Barrel3CLys nanotube at circa 5 Å resolution. The density map shows that the wall of the nanotube is composed of a bilayer of helices.

To the best of our knowledge, the structure of **Barrel3CLys** nanotubes is unprecedented in both native and synthetic worlds. It is of great significance in studying the rules the peptide/protein self-assembly, especially helical supramolecular assemblies. Once the atomic model is obtained, the sequence-structure relationship will be better understood, which will guide the design of other forms of complex multihelical assemblies, for example, α-helical barrels which is the original target of this study. The structure of **Barrel3CLys** nanotubes also hold great practical promise in applications such as in controlled release, facilitated transport, and directional catalysis.

4.2.2.5 D-Barrel3CLys, a mirror image of Barrel3CLys

In order to investigate the role of handedness of α-helices in dictating the nanotube formation, **D-Barrel3CLys** is synthesized which is a D-peptide analogue (all D-amino acids) to the wild-type **Barrel3CLys**. The CD confirmed that **D-Barrel3CLys** formed left-handed helices opposite to **Barrel3CLys**. Negative-stain TEM also showed the formation of 10 nm nanotubes indistinguishable to the wild-type ones, and SAXS analysis of **D-Barrel3CLys** nanotubes showed identical characteristics with those of **Barrel3CLys**. All these data suggested that the handedness of helices is not a critical factor determining the structure. However, the helical symmetry of the nanotubes assembled from **D-Barrel3CLys** might be in an opposite sense.

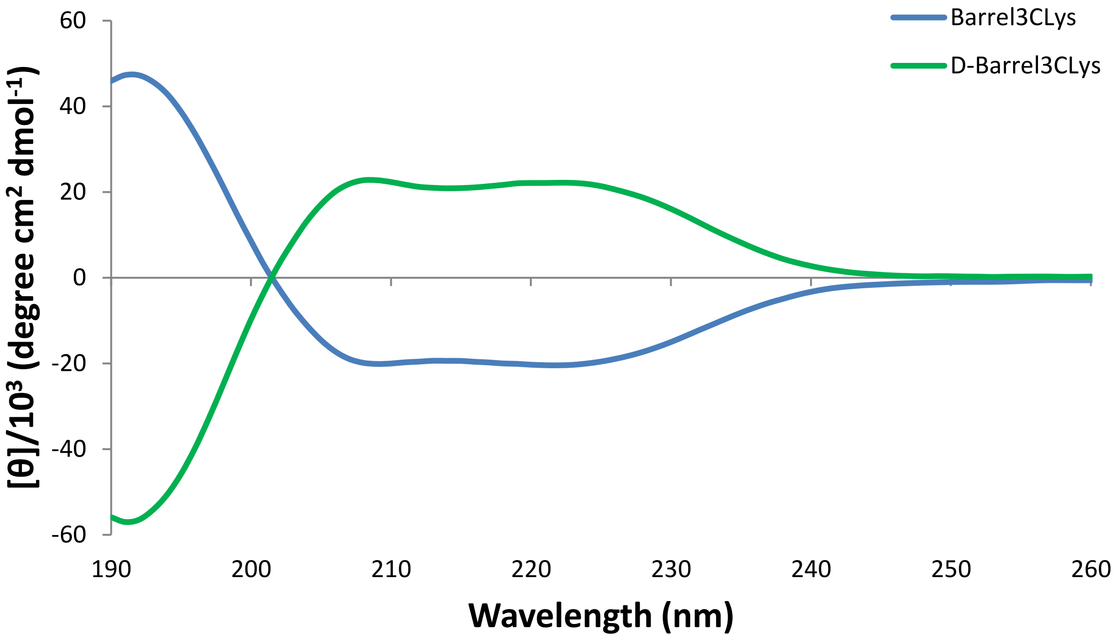


Figure 4.12. CD spectra of Barrel3CLys and D-Barrel3CLys (100 μM) in acetate buffer (10 mM, pH 4.0). Spectrum of D-Barrel3CLys is a mirror image of that of Barrel3CLys indicating D-Barrel3CLys forms left-handed helical conformation.

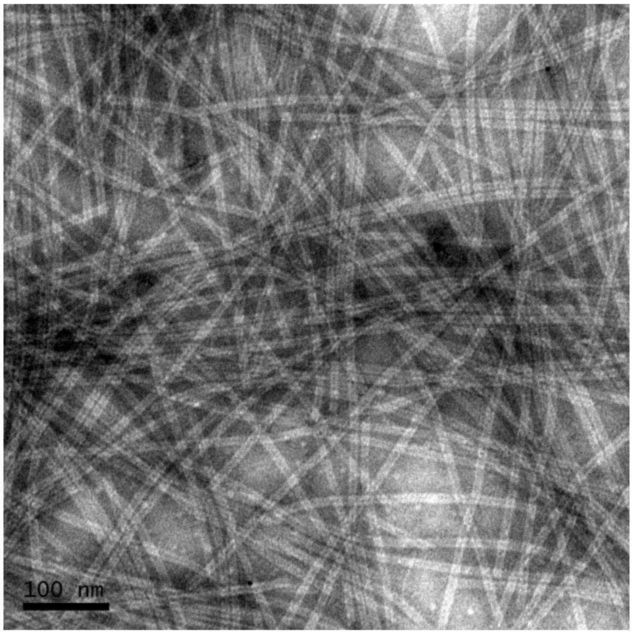


Figure 4.13. TEM image of D-Barrel3CLys nanotubes identical to Barrel3CLys nanotubes.

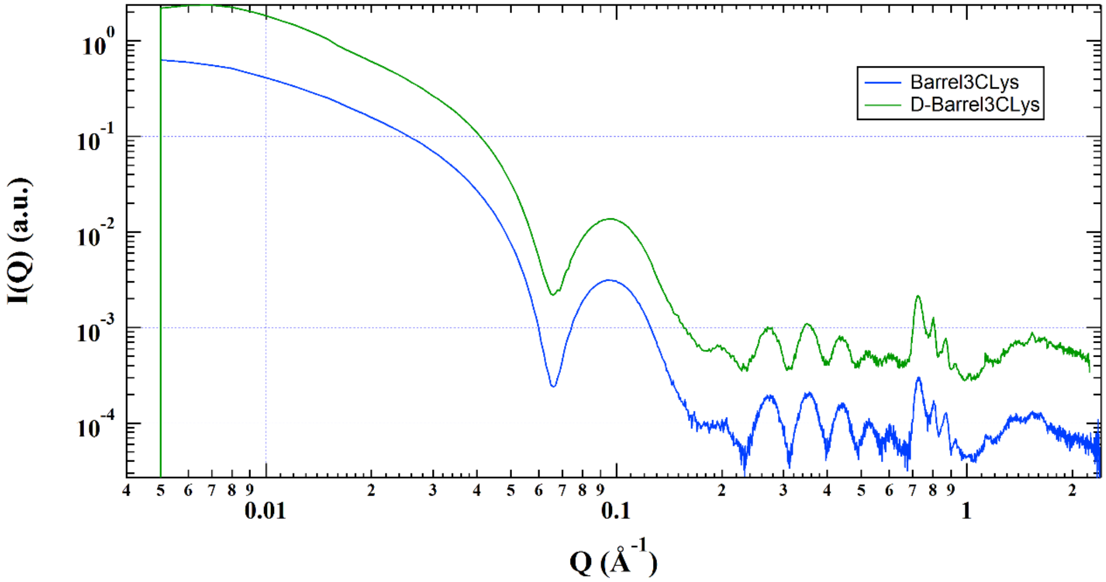


Figure 4.14. Comparison of SAXS and WAXS profiles of D-Barrel3CLys and Barrel3CLys nanotubes. The identical features suggest the shape, size, as well as molecular packing within nanotubes are all the same.

4.3 Barrel3CArg, a conservative mutant of Barrel3CLys

After the discovery of the unprecedented **Barrel3CLys** nanotubes, initial efforts toward the understanding of its sequence-structure relationship were attempted by introducing mutations to the original sequence. None of the mutant peptide was able to reproduce the tubular structure of **Barrel3CLys**. However, one of the filamentous assemblies formed by a mutant in which all lysine residues were replaced with arginine residues, named **Barrel3CArg**, attracted more attentions. **Barrel3CArg** formed filamentous assemblies in acetate buffer (10 mM, pH 4.0) with homogeneous diameter comparable to that of **Barrel3CLys**. So we speculated that **Barrel3CArg** might have also formed tubular structure.

4.3.1 Sequence of Barrel3CArg

As described in section 4.2.1, charged residues were put at b and e heptad positions of **Barrel3CLys** according to the design criteria. Lysine and glutamic acid were initially chosen because of their helix favoring property. However, other charged amino acids might potentially play the same role in determining the final supramolecular structure. Considering that Barrelr3CLys nanotubes were grown only in acidic condition close to the pKa of glutamic acid, the Glu residues were kept unchanged, while four lysine residues at b and e heptad positions were all replaced with arginine. The resultant sequence of **Barrel3CArg** is projected onto the helical wheel diagram (Figure 4.15).

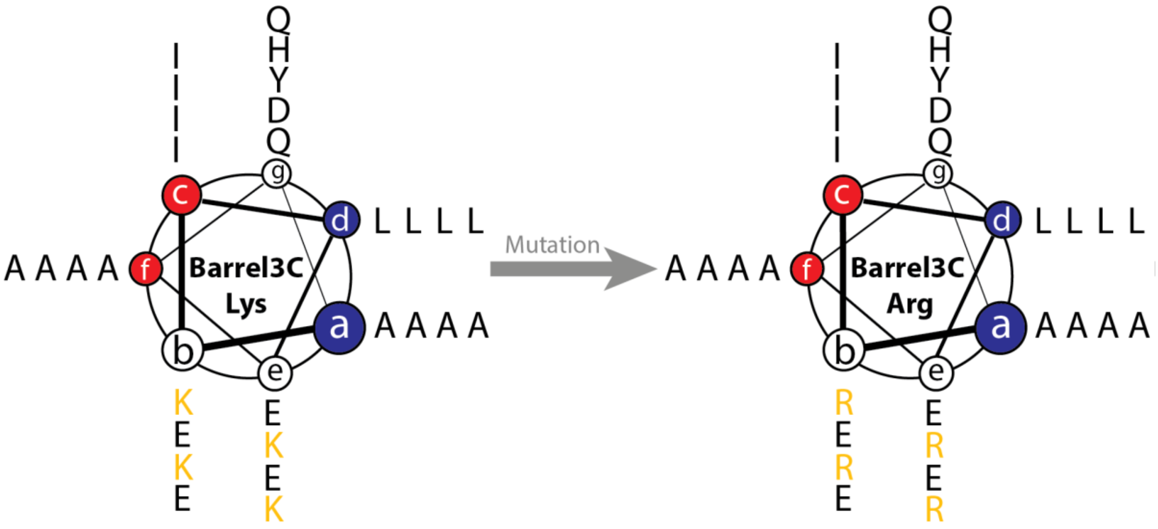


Figure 4.15. Sequence comparison between Barrel3CLys and Barrel3CArg. All four lysine residues at b and e heptad positions are mutated to arginine residues.

4.3.2 Results and Discussion

4.3.2.1 Circular Dichroism and Flow Linear Dichroism

The conformation of **Barrel3CArg** peptide was initially interrogated with circular dichroism (CD) spectropolarimetry. Similarly with its parent peptide **Barrel3CLys**, The CD spectrum indicates the formation of α-helical secondary structure with relatively weak MRE indicating the existence of unstructured segments under experimental conditions (100 μM peptide in 10 mM acetate buffer pH 4.0).

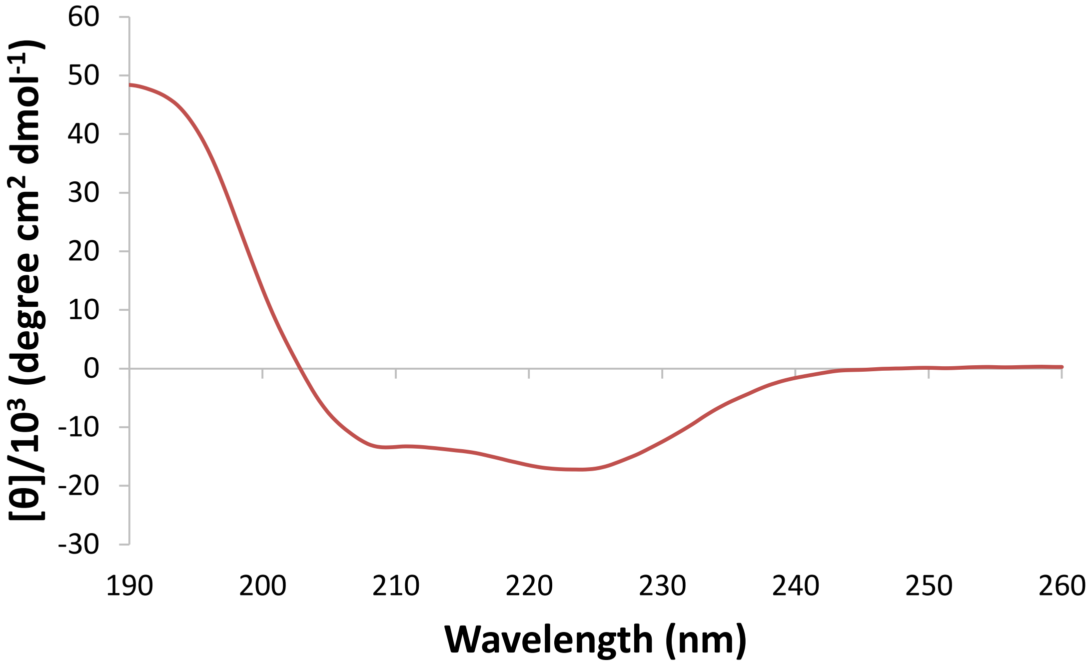


Figure 4.16. CD spectrum of peptide Barrel3CArg (100 µM) in acetate buffer (10 mM, pH 4.0). MRE is also weak indicating not full-length of peptide is in a α-helical conformation.

Flow linear dichroism was then employed to investigate the helix orientation within the filamentous assemblies of **Barrel3CArg**. The LD spectrum displayed a strong positive signal at 192 nm, and a weaker, broader positive signal at 220 nm, just like what was observed for **Barrel3CLys**, suggesting that α-helices are aligned perpendicular with the long axis of peptide assemblies. At the same experimental condition (50 µM peptide, 10 mM acetate buffer, pH 4.0, 5000 rpm rotation speed), the intensity of LD signal for **Barrel3CArg** was much stronger than the wild-type peptide. This distinction could be rationalized by TEM observation that **Barrel3CArg** formed much longer filamentous assemblies than **Barrel3CLys**, which is easier to get aligned by Couette flow.

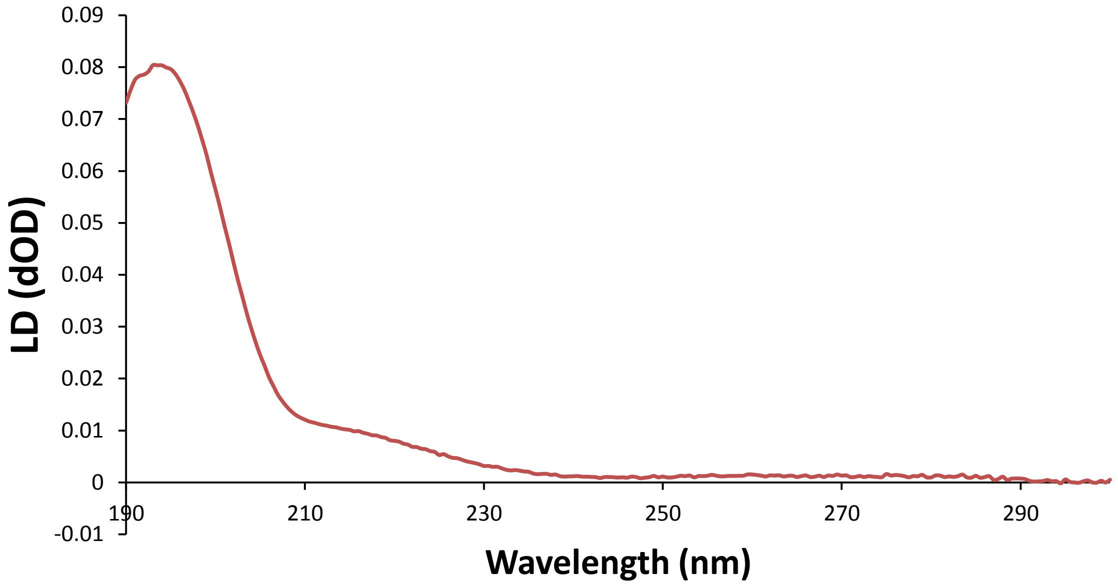


Figure 4.17. Flow LD spectra of peptide Barrel3CArg (50 µM) in acetate buffer (10 mM, pH 4.0) showing that 222 nm and 190 nm transitions are aligned with the flow direction, hence peptides oriented perpendicular with flow direction.

4.3.2.2 Negative-stain Electron Microscopy

TEM analysis (Figure 4.18) of **Barrel3CArg** in acetate buffer (10 mM, pH 4.0) revealed the formation of filamentous assemblies with a homogeneous diameter of ca. 9nm. The length of these assemblies could extend to several microns. Although we speculated that these filamentous assemblies were likely to be tubular, more evidence is required to make this judgment as the lumen was not clearly seen in negative-stain TEM images.

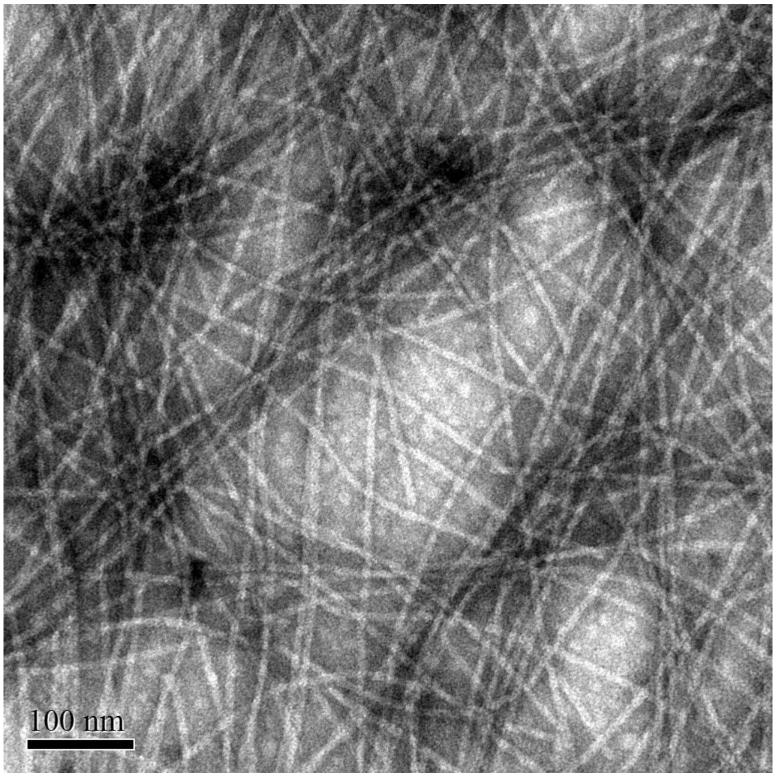


Figure 4.18. Negative-stain TEM of Barrel3CArg nanotubes. The diameter of these nanotubes is circa 9 nm. The central channels of the nanotubes are not clearly visible.

4.3.2.3 Solution X-ray Measurements

Small- and wide-angle X-ray scattering (SAXS/WAXS) data were collected on aqueous solutions of **Barrel3CArg** filaments in acetate buffer (10 mM, pH 4.0) to interrogate the structural hierarchy. The scattering profile showed enormous distinctions from **Barrel3CLys** scattering profile suggesting that the two structures differs in the overall size and shape as well as the internal packing interactions.

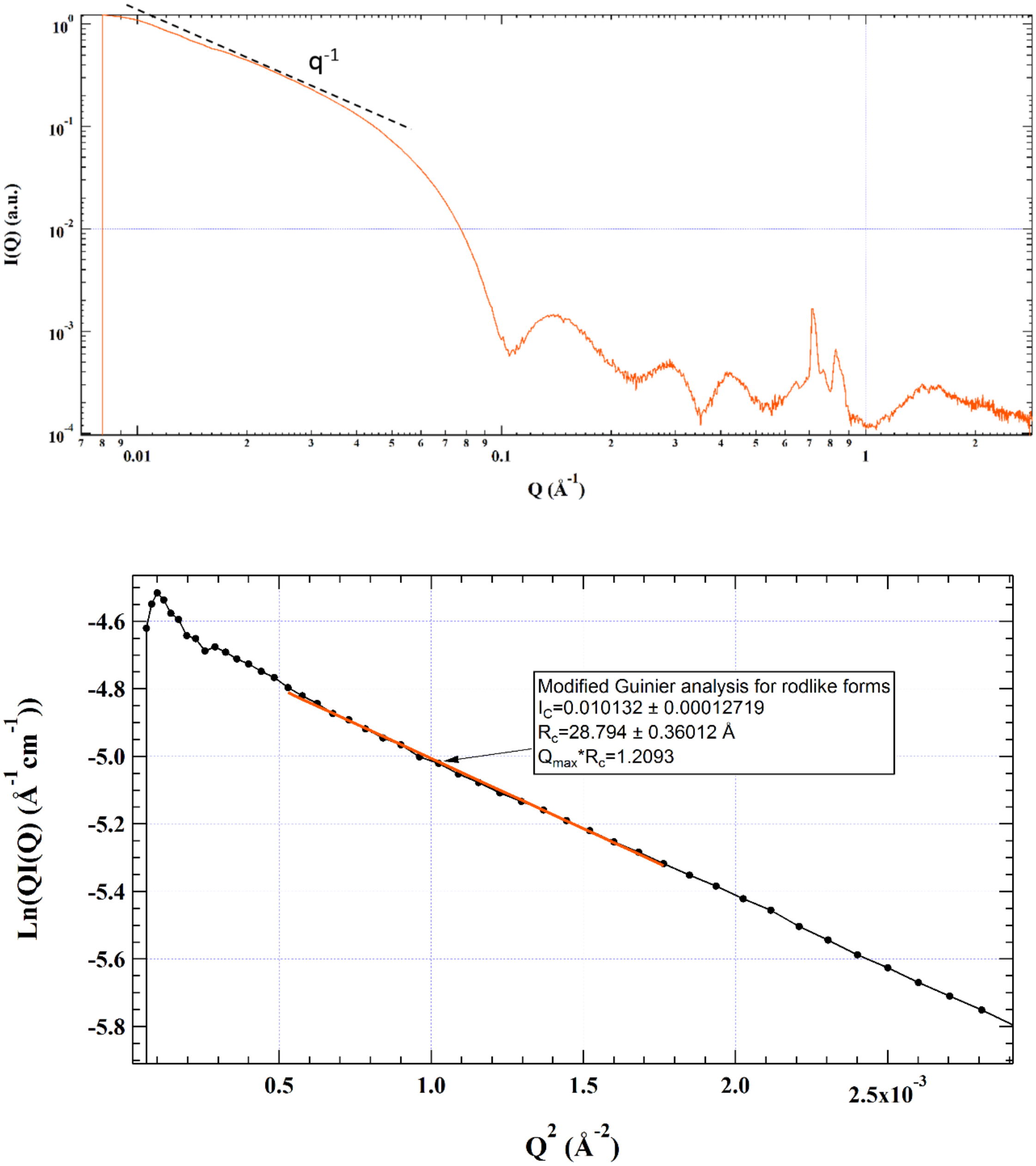
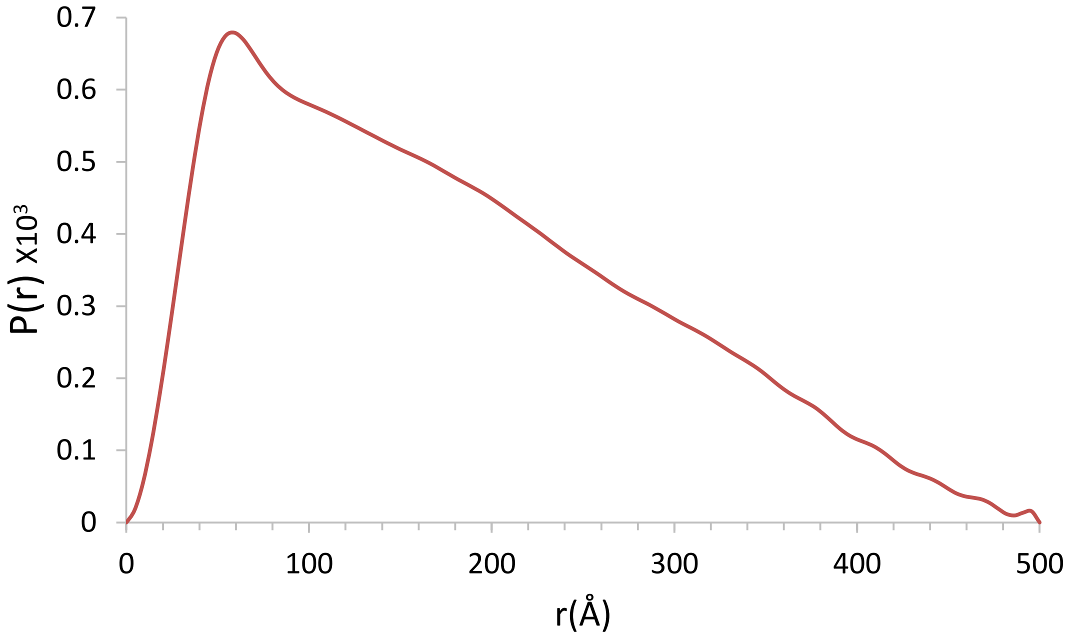


Figure 4.19. SAXS/WAXS studies of Barrel3CArg nanotubes. Top: SAXS and WAXS profile of Barrel3CArg nanotubes. The scattering intensity at $q < 0.1$ Å^{-1} roughly follows the q^{-1} power law indicating the assembly in solution has rod-like form. Guinier Analysis for rod-like forms (bottom) gives a cross sectional radius of gyration of 28.8 Å, smaller than Barrel3CLys nanotubes consistent with TEM observation.



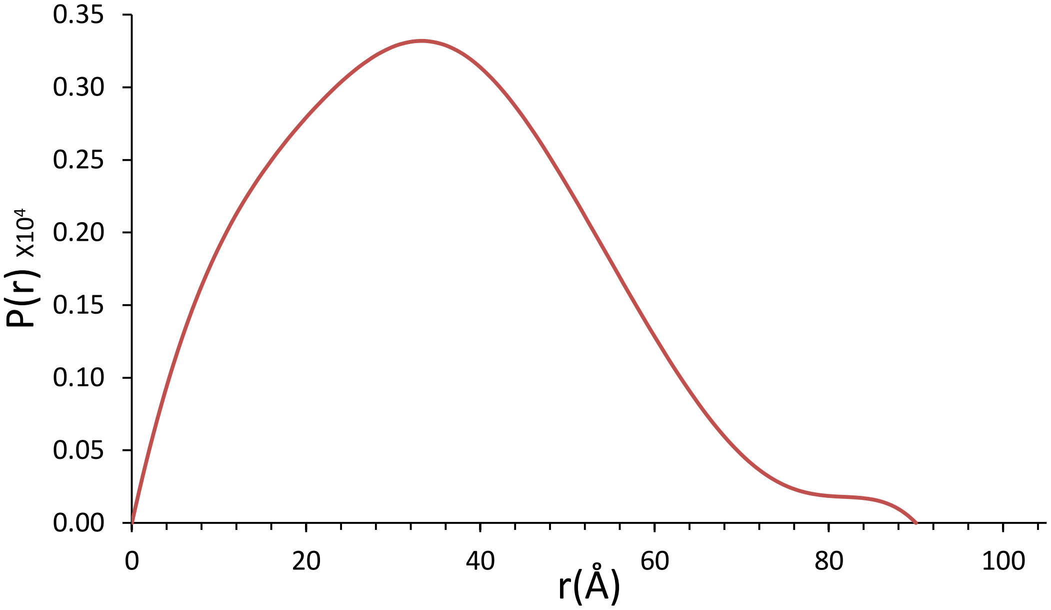


Figure 4.20. The pair distribution function (PDDF) of Barrel3CArg nanotubes. Top: PDDF of Barrel3CArg nanotubes, calculated from the SAXS (*I(q)* versus *q*) data using program GNOM further confirms the cylindrical shape. Bottom: The cross section PDDF calculated from SAXS data (*q*I(q)* versus *q*) using GNOM.

In the small-angle region, the experimental scattering intensities at $q < 0.1$ Å^{-1} roughly follow the q^{-1} power law (Figure 4.19), indicating the assembly in solution has rod- or cylinder-like form. The pair distance distribution function (PDDF) in Figure 4.20 derived from the SAXS data has a characteristic long tail, which further confirms the cylindrical shape. However, the cross-section PDDF derived from $q*I(q)$ does not exhibit a typical bimodal form characteristics of hollow cylindrical assemblies, consistent with TEM observation where the central channels of **Barrel3CArg** nanotubes are not clearly visible. The largest dimension of the cross-section is about 90 Å, consistent with the TEM observation. The value of R_c, the radius of gyration of the rod cross-section, was obtained as 28.8 Å through the fitting of the modified Guinier equation, $I(q) = \pi q I(q = 0)\exp(-0.5R_c^2 q^2)$, smaller than **Barrel3CLys** nanotubes.

4.3.2.4 Cryo-EM Helical Reconstruction

In order to get a high-resolution structure **Barrel3CArg** supramolecular assemblies, cryo-EM images were collected. The filaments evidently showed a central lumen which confirmed our speculation. A 3.5 Å resolution structure was obtained by performing 3D helical reconstruction using IHRSR algorithm [41].

The resultant structure surprised us by the significant distinctions that it exhibits from the supramolecular structure of the parent peptide **Barrel3CLys**. Firstly, instead of a helical bilayer, the wall of **Barrel3CArg** tubes is composed of a single layer of α-helices. Secondly, the top view of the tube shows a square cross section, while **Barrel3CLys** nanotube is formed by three helical ribbons. Thirdly, the interfaces that hold the wall together laterally are end-to-middle helix-helix interactions rather than end-to-end. These structural distinctions are astonishing considering that **Barrel3CArg** is an conservative mutant of **Barrel3CLys**.

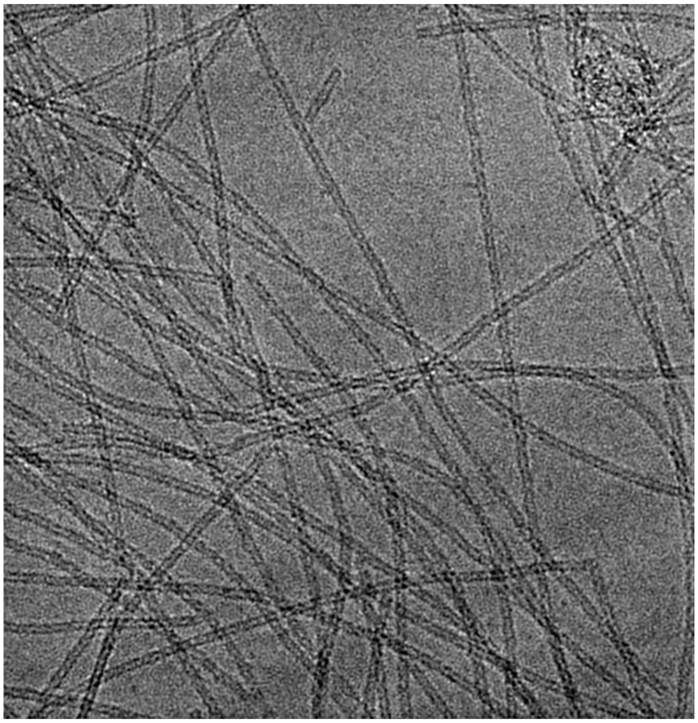


Figure 4.21. A typical cryo-EM image of Barrel3CLys nanotubes.

However, there are certain similarities between the wild-type and mutant supramolecular structures. In these two peptide assemblies, α-helices are both aligned perpendicular with the long axis of tubes, consistent with the LD results. It is also reasonable to conjecture that similar packing motif, most probably coiled-coil knobs-into-holes packing, guides the axial growth of these nanotubes. Both peptides display relatively low helicity by CD measurements, which can be explained by the existence of melted segments. In the **Barrel3CLys** assembly, the helices in the outer layer are partially unstructured, while, in this case, the lack of electron density at the termini of α-helices pointing outside indicates the existence of random coil conformation.

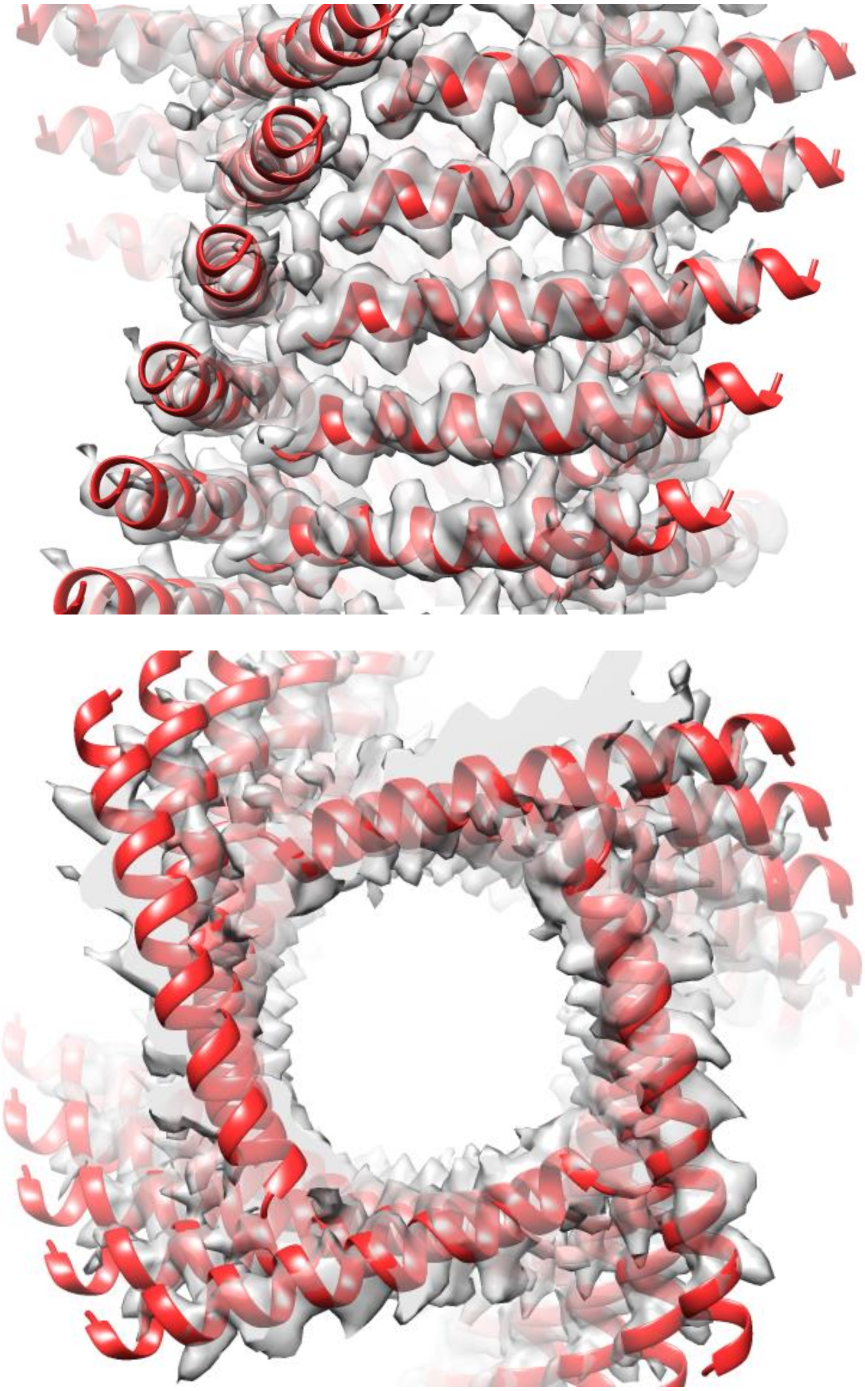


Figure 4.22. Side and top view of cryo-EM structure of unprecedented Barrel3CArg nanotube at circa 3.5 Å resolution. The density map shows that the wall of the nanotube is composed of a single layer of helices, distinct from the structure of Barrel3CLys nanotubes.

The cryo-EM structure also provides hints about the forces that drive the bilayered tubular structure into single layered tubular structure with totally different helical symmetry. At 3.5 Å resolution, large side chains are visible as shown in Figure 4.23. We can see that one of the arginine residues is in close proximity with the lateral interface suggesting that this arginine residue actively interacts with helices from the adjacent wall. These interactions potentially dictated the structure.

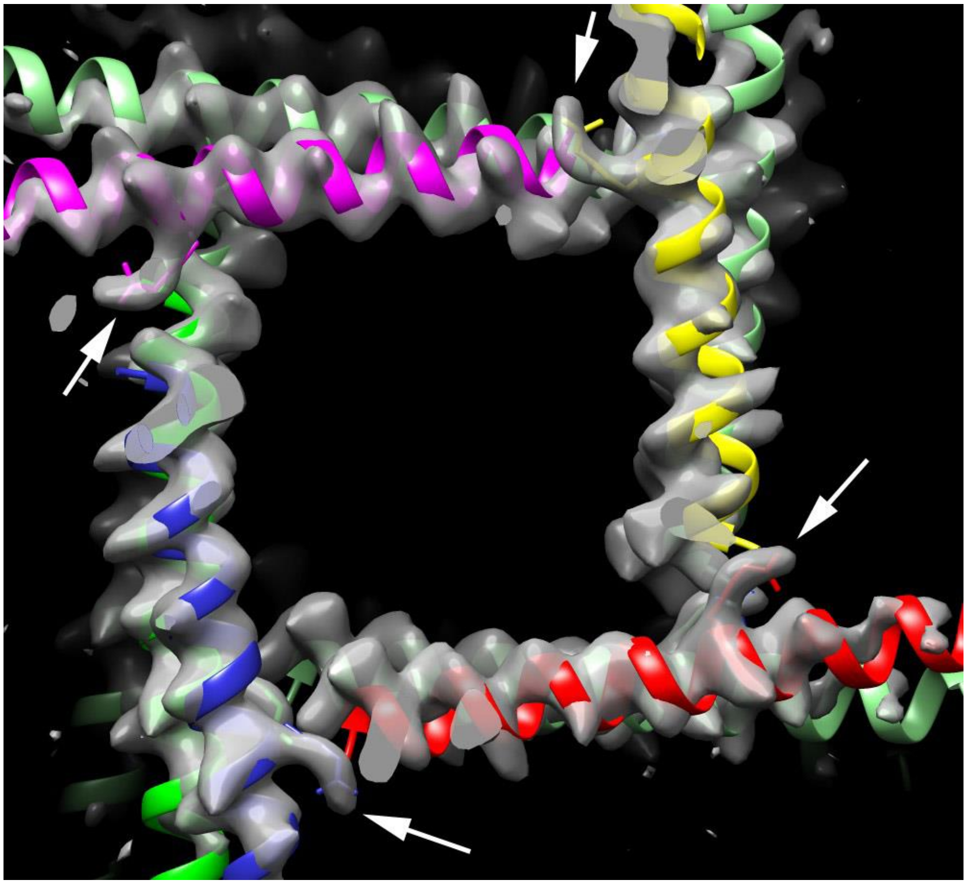
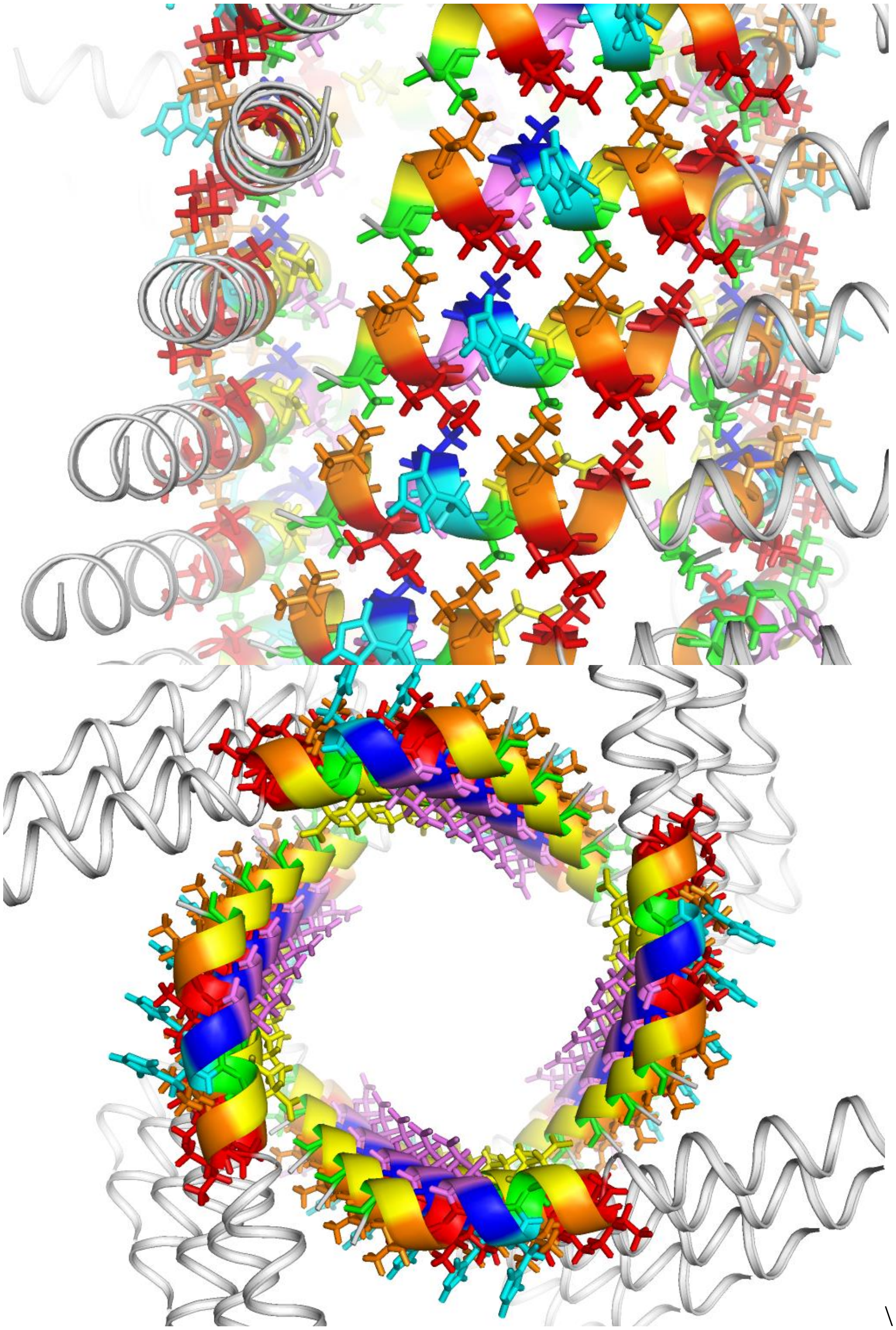


Figure 4.23. The reconstructed cryo-EM density map clearly shows that an arginine residue (pointed by arrows) inhabits near the ribbon-ribbon interface.

4.3.2.5 Computational Modeling

Computational refinements are critical in analyzing, interpreting, and annotating structural information in cryo-EM density maps which can be used directly to facilitate the construction and evaluation of a protein structural model [42-44]. An atomic model was constructed and refined into the cryo-EM density map which provides more details regarding the packing mode within **Barrel3CArg** nanotubes and confirmed our speculations.

Hydrophobic effect is the main driving force for the axial growth of the **Barrel3CArg** peptide nanotubes. When the atomic model was run through SOCKET [45], a program that locate Knobs-into-Holes packing between α-helices in PDB structures, extended Knobs-into-Holes packing was identified (Figure 4.24) in the structured segments that surround the central channel, which agrees with our design. However, the extended Knobs-into-Holes packing leads to the formation of open multihelical assemblies, coiled-coil ribbons, rather than closed multihelical assemblies, α-helical barrels. Hydrogen bonding between arginine 13 and C-terminal glutamine hold the coiled-coil ribbons together to form nanotubes. Show in Figure 4.25 is the ribbon-ribbon interface where the C-terminal glutamine side chain forms hydrogen bonds with two arginine 13 side chains of adjacent helices from the other coiled-coil ribbon.



\

Figure 4.24. Knobs-into-Holes packing (colored) of Barrel3CArg nanotube. KIH packing surrounding the core of the tube is identified by SOCKET program [45] indicating the formation of coiled coils.

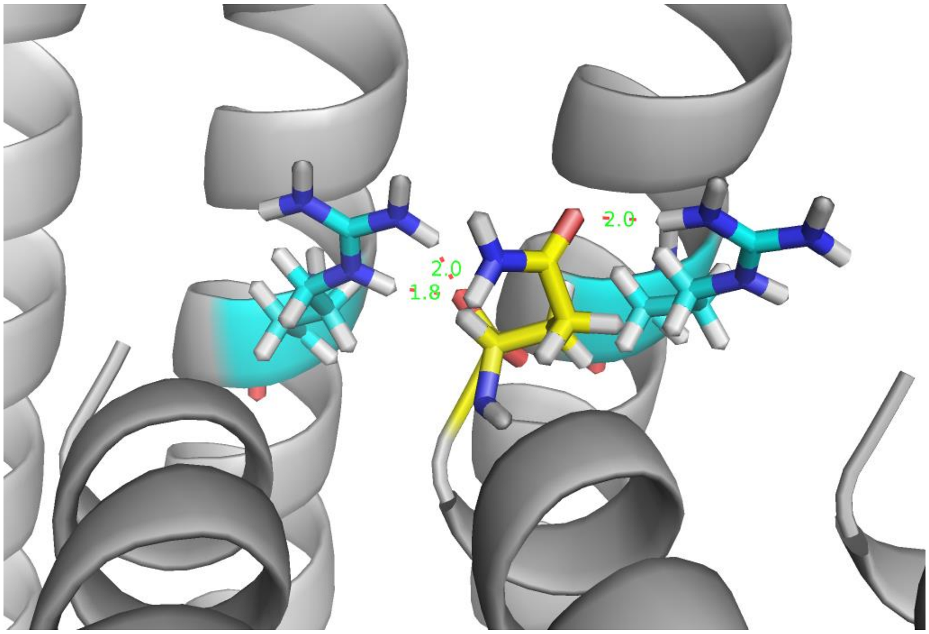


Figure 4.25. A close-up of the ribbon-ribbon interface of Barrel3CArg nanotube. The C-terminal glutamine (yellow) side chain forms hydrogen bonds with two arginine 13 (cyan) side chains of adjacent helices from the other coiled-coil ribbon.

The inner surface of the nanotube is primarily occupied by hydrophilic residues while the outer surface is highly hydrophobic. This will cause unpolymerized monomers to non-specifically bind to the outside of the tubes to reduce the hydrophobic solvent-accessible surface area, which explains the peripheral “dusts” around the tubes in cryo-EM density map.

Similarly with **Barrel3CLys** nanotubes, the structure of **Barrel3CArg** supramolecular assemblies is also unprecedented in both native and synthetic worlds, and holds great significance in both theoretical and practical studies on peptide/protein self-assembly. On one hand, thorough characterization and analysis of the supramolecular structure provide insights to the design of self-assembling protein-based materials, especially higher-order coiled-coil assemblies. On the other hand, new research questions are raised such as why arginine 13, why squre-shaped cross section, and so on. Finding answers to these questions will help scientists better understand the rules that govern the folding and self-assembly of peptides/proteins.

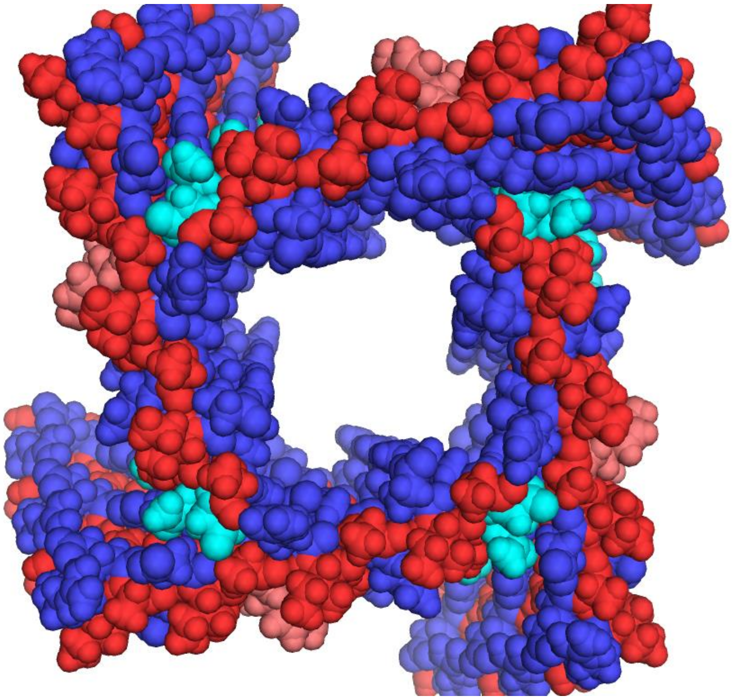


Figure 4.26. Space-filling model of Barrel3CArg nanotube. Hydrophobic residues are colored red and salmon. Hydrophilic residues are colored blue and cyan.

4.3.2.6 Arginine 13, the structural switch

As has been stated above, cryo-EM structure and the atomic model reveal that arginine 13 plays key role in determining the final structure of **Barrel3CArg** assemblies, which lead to an intuitive hypothesis that, if arginine 13 is mutated back to lysine, the self-assembly of mutant peptide will form **Barrel3CLys**-like nanotubes. To test this hypothesis, a mutant peptide, **Barrel3CArg R13K** was synthesized and preliminarily characterized.

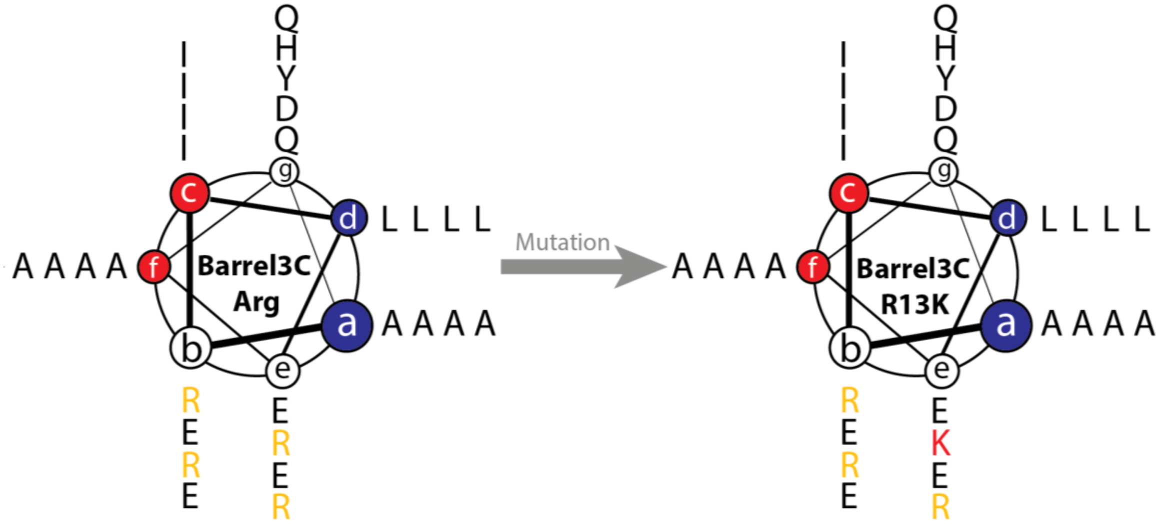


Figure 4.27. Sequence comparison between Barrel3CArg and Barrel3CArg R13K.

TEM analysis showed that the **Barrel3CArg R13K** formed tubular assemblies with a diameter of ca. 10 nm and clearly visible central lumen more similar with **Barrel3CLys** nanotubes. However, higher-resolution evidence is required to assert that **Barrel3CArg R13K** nanotubes resemble **Barrel3CLys** nanotubes in atomic details.

If the hypothesis here is confirmed by high-resolution techniques, e.g. cryo-EM helical reconstruction, this will be a classical case showing the liability of protein quaternary structures. One residue mutation completely alters energy landscape for a peptide supramolecular assembly. Knowledge learned from this case can be used to better understand protein sequence-structure relationship.

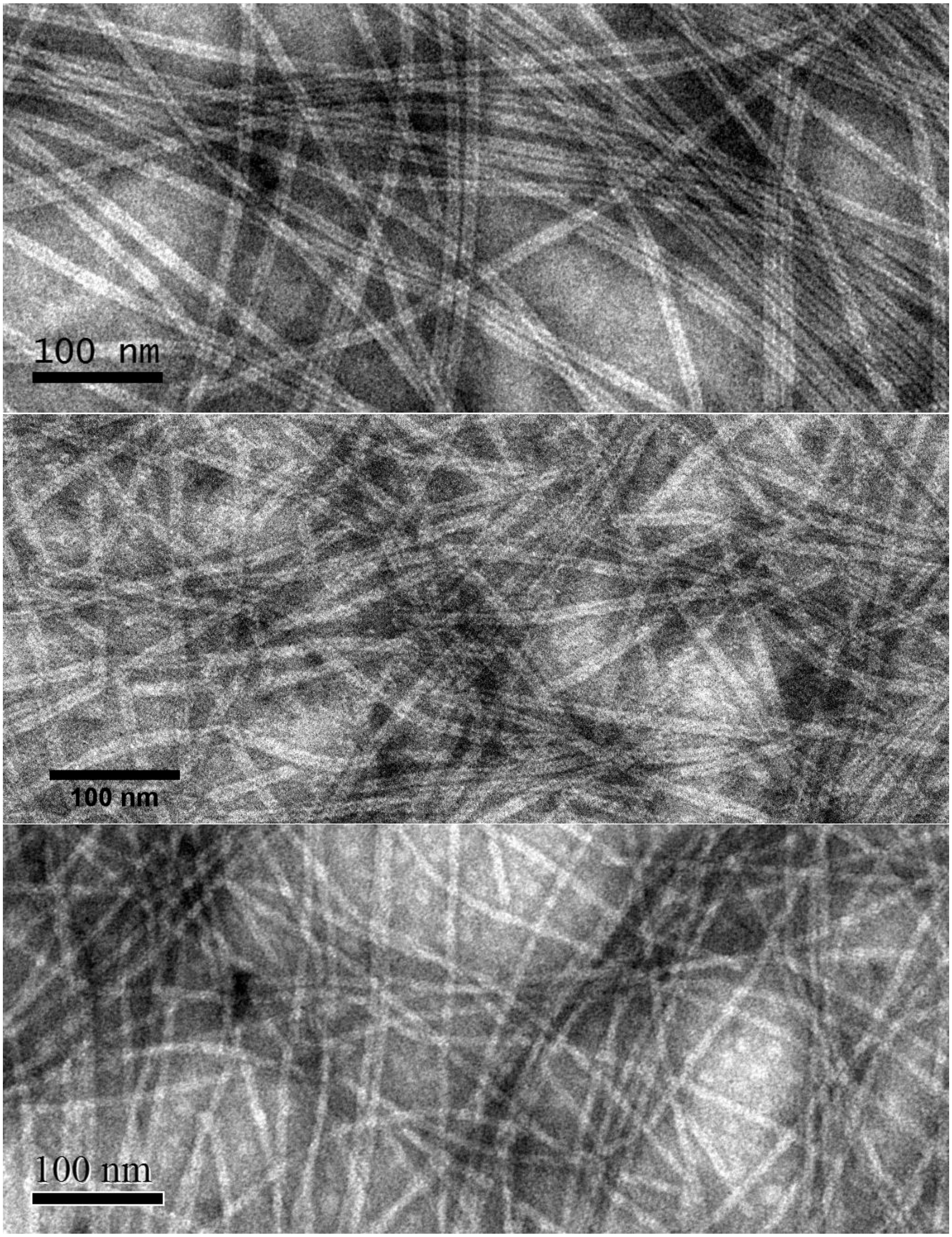


Figure 4.28. TEM images of Barrel3CLys nanotubes (top), Barrel3CArg R13K nanotubes (middle), and Barrel3CArg nanotubes (bottom). Barrel3CArg R13K nanotubes resemble those of Barrel3CLys.

4.4 Summary

Protein-based nanotubes represent attractive candidates for the construction of functional nano-porous materials with medically relevant applications in controlled release, directional transport, and catalysis. High-order coiled-coil assemblies are ideal building blocks for the design of tubular structure thanks to their tailorable channel dimension and chemistry as well as their relatively straightforward sequence characteristics. However, efforts toward the construction of α-helical barrels over the past decade have always been frustrated.

Here, our attempts were also unsuccessful in terms of failing to achieve our original targeted structures. However, the adventitiously discovered structures which are unique and unprecedented made our attempts even more valuable and significant. They not only provide chemists with more insights about protein sequence-structure relationship, but also offer a perfect platform for protein engineering. More specifically, thoroughly evaluate the physics of these structures could guide the design of high-order multihelical assemblies.

Furthermore, the liability of protein quaternary structures have been demonstrated, at least partially, through the example of **Barrel3CArg R13K** peptide where a single-residue mutation completely alters the final supramolecular structure compared with its parent peptide. To provide more support, **Barrel3CLys K13R** mutant will be of great interest to study, which is currently ongoing.

4.5 Materials and methods

Peptide Synthesis. Peptides **Barrel3CLys**, **D-Barrel3CLys**, **Barrel3CArg**, and **Barrel3CArg R13K** were either purchased from GenScript USA, Inc. (Piscataway, N.J.) or synthesized in-house. In the latter case, peptides were prepared using microwave-assisted synthesis on a CEM Liberty solid-phase peptide synthesis instrument using a 4-(hydroxymethyl)phenoxyacetamido-methyl]-PEG-PS resin from Applied Biosystems, Inc. (Foster City, CA). Standard Fmoc protection chemistry was employed with coupling cycles based on HBTU/DIEA-mediated activation protocols and base-induced deprotection (20% piperidine in DMF with 0.1 M HOBt) of the Fmoc group. The peptides were purified via RP-HPLC on a C18 column with a gradient of water–acetonitrile (0.1% trifluoroacetic acid). The purity was assessed to be above 95% by analytical HPLC and MALDI mass spectrometry. Peptide mass was confirmed using electro-spray ionization mass spectrometry. The peptides were lyophilized, sealed, and stored at -20 ℃. Samples for analytical studies were prepared by dissolving the peptide at the appropriate concentration in aqueous buffer. Peptide solutions were dialyzed against buffer solution to remove the remaining trifluoroacetic acid (MWCO = 2,000 Da).

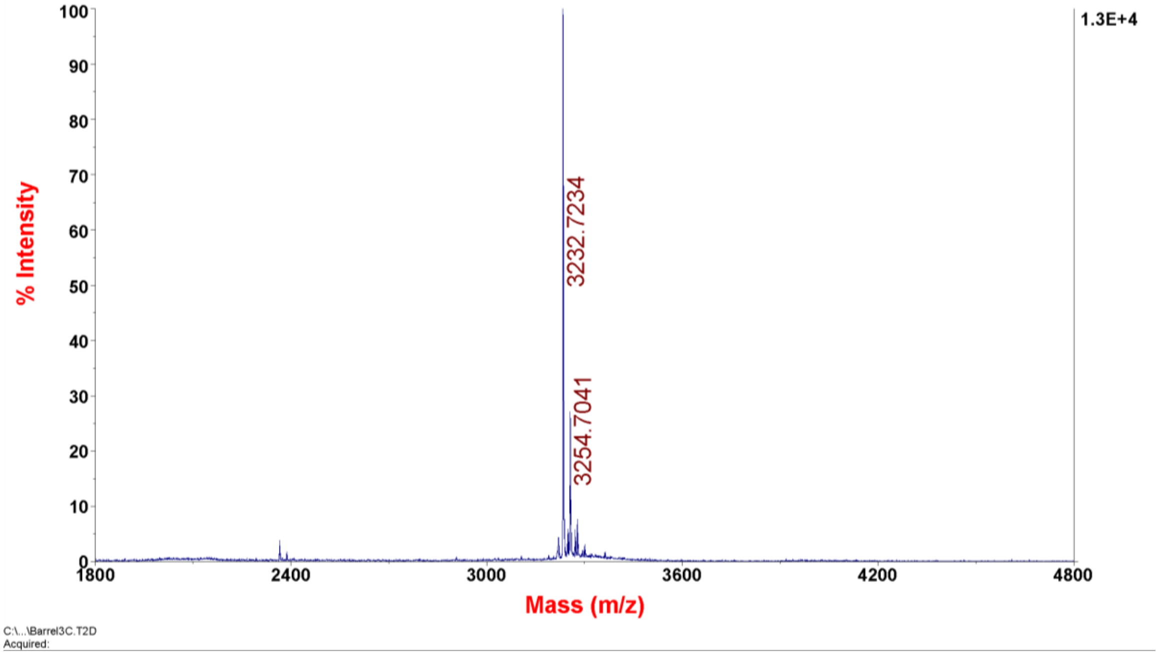


Figure 4.29. MALDI-mass spectrum for purified peptide Barrel3CLys.

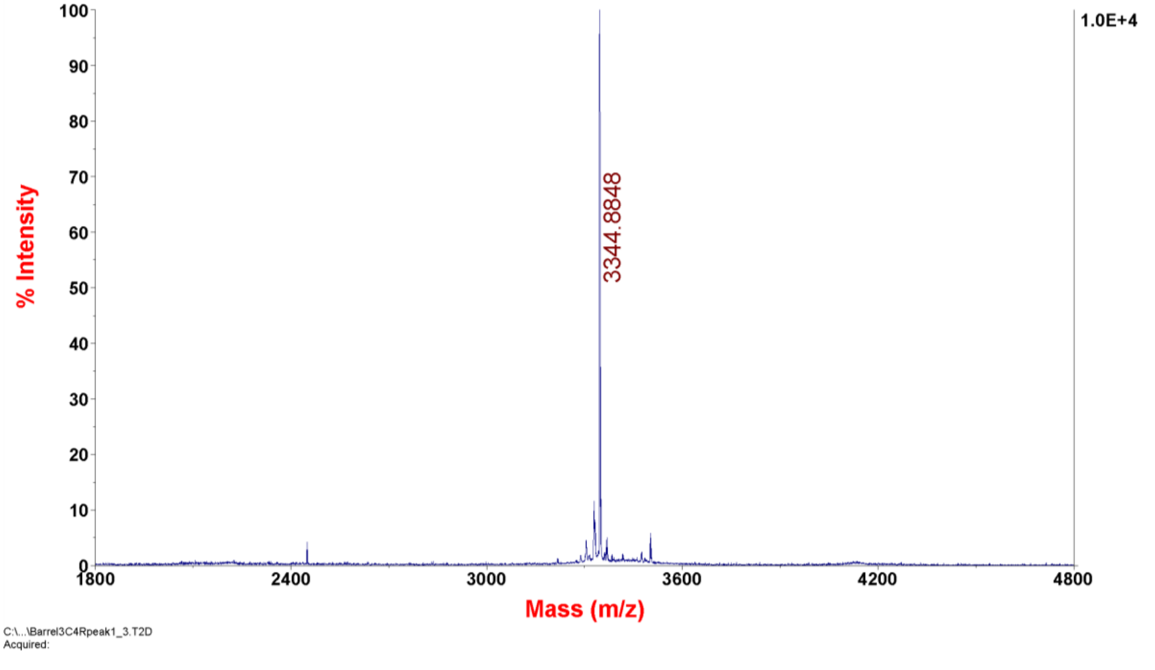


Figure 4.30. MALDI-mass spectrum for purified peptide Barrel3CArg.

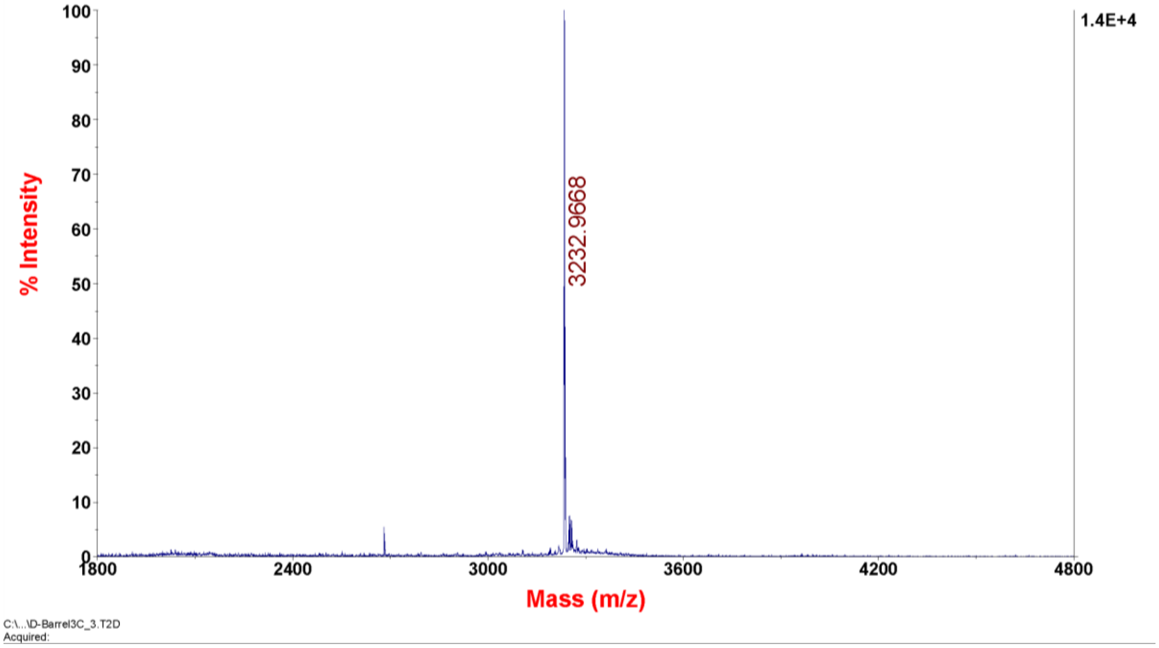


Figure 4.31. MALDI-mass spectrum for purified peptide D-Barrel3CLys.

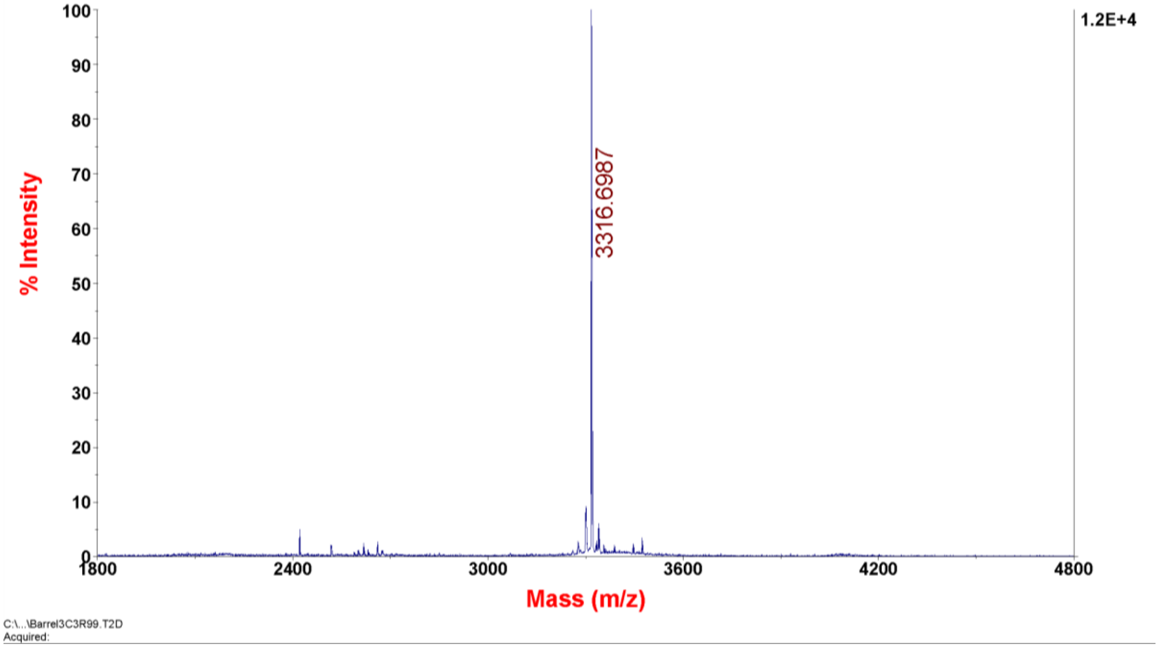


Figure 4.32. MALDI-mass spectrum for purified peptide Barrel3CArg R13K.

Circular Dichroism Spectropolarimetry. CD spectra were recorded on a Jasco J-810 CD spectropolarimeter in 0.10 mm quartz cells at a concentration of 100 µM in acetate buffer (10 mM, pH 4.0). Spectra were recorded from 260 to 190 nm at a scanning rate of 100 nm/min and a resolution of 0.5 nm. Peptide concentration was determined spectrophotometrically from measurement of the absorbance at 280 nm (A_{280}). For peptides containing Tyr, Trp or Cys residues, the peptide concentration can be calculated from equation (1):

$$\mathrm{MW}\times {}^{A_{280}}\!/_{c} = 1280n_Y + 5690n_W + 120n_C \qquad (1)$$

in which c is the concentration of peptide in mg/mL, and n_Y, n_W, and n_C are the numbers of tyrosine, tryptophan and cysteine residues, respectively, in the peptide sequence [46]. As **Barrel3CLys** and its derivatives contain only a single tyrosine residue per molecule, then $c = \mathrm{MW}\times A_{280}/1280$. To eliminate error in determination of absorbance that could arise as a result of UV light scattering due to peptide self-assembly, aqueous solutions of peptide were mixed with 6 M guanidinium chloride in 1:9 v/v ratio and put in room temperature for 1hr to completely denature the sample prior to performing the absorbance measurements.

Flow Linear Dichroism Spectroscopy. Flow linear dichroism spectra were recorded on a JASCO J-810 circular dichroism spectropolarimeter using a microvolume cuvette with a path length of 50 μm and a rotation speed of 5000 rpm to establish Couette flow. The background scattering for each sample was obtained from the LD spectra of samples at 0 rpm. The LD spectra were measured after 15 min of rotation.

Transmission Electron Microscopy. In order to get relatively long nanotubes without fibrils, 3mg/mL **Barrel3CLys** unbuffered solution with pH adjusted to 4.0 was incubated at room temperature for 10 days. And then the large tubes were pelleted down using preparative ultracentrifuge equipped with a Beckman 70.1 Ti rotor with a speed of 32,000 rpm for 45 minutes. The pellet was then resuspended into 10 mM acetate buffer pH4.0. TEM specimens were prepared from aqueous peptide solutions in acetate buffer (10 mM, pH 4.0). The sample samples were deposited onto 200 mesh carbon coated copper grids from Electron Microscopy Sciences (Hatfield, PA). After a 30-second incubation period, excess liquid was wicked away and the specimens were stained with 1% methylamine tungstate from Electron Microscopy Sciences (Hatfield, PA). Excess stain was wicked away after incubation on the grid for 1 min. The sample grids were dried under vacuum and stored in a desiccator. TEM measurements were acquired on a Hitachi H-7500 transmission electron microscope at an accelerating voltage of 75 kV.

Cryo-EM. Cryo-EM images were acquired at Z Hong Zhou lab in Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, California, USA. The detailed imaging methods could be found elsewhere [47]. The IHRSR algorithm [41] was used for the 3D helical reconstruction.

Computational Modeling. The fit of atomic model into cryo-EM density map was performed by Dr. Frank DiMaio at Department of Biochemistry, University of Washington at Seattle. Details can be found elsewhere [42].

Scanning Transmission Electron Microscopy. The STEM data were acquired at Brookhaven National Laboratory (BNL). The STEM instrument operates at 40 keV with a scanning probe of < 0.3 nm diameter produced from a cold field-emission source. Every electron emerging from the specimen is detected by one of the scintillator-photomultiplier detectors collecting 0-15 mRadian (bright field), 15-40 MRadian (small-angle dark field) and 40-200 mRadian (large-angle dark field). The large-angle signal is proportional to the mass of atoms in the path of the beam. Specimen quality and mass calibration are checked by detailed comparison of the image to the known structure of tobacco mosaic virus (TMV).

Specimens are deposited on thin carbon (circa 2 nm thick) supported on a thicker holey carbon film mounted on a titanium grid using the wet-film, hanging-drop method http://www.bnl.gov/biology/stem/SpecPrepDetails.asp. TMV is added to the grid first as an internal control, followed by injection buffer, then specimen solution (in 10 mM acetate buffer, pH 4.0) for 1 min, then 10 washes of 20 mM ammonium acetate pH 7.0. Excess solution is wicked from the edge with filter paper between each injection. After the last wash the grid is wicked to a thin layer (ca. 1 μm), fast frozen by plunging into liquid nitrogen shush and stored under liquid nitrogen. Grids are freeze-dried overnight in an ion-pumped chamber with an efficient cold trap and transferred under vacuum to the STEM cold stage (-160 ℃). Imaging typically uses a dose of 20 el/Å^2 (causing < 5% mass loss, corrected by comparison to TMV). Mass measurements are performed off-line with customized software (PCMass, available at ftp.stem.bnl.gov). The program masks out objects significantly above background and computes the value for the thin carbon in the remaining areas, which is subtracted, and pixels within the contour of filaments are summed and divided by length to give mass-per-unit-length. Accuracy is determined by cleanliness of the background between objects and by counting statistics of the scattered electrons. At 10 el/Å^2, TMV SD (standard deviation) is ideally circa 1 % and SD of filaments of 6 kDa/nm is circa 20% for a single segment 10 nm long. For TMV the program provides automatic searching and measurement but for the thin filaments the low S/N requires manual selection whereupon the software "locks on" to a segment giving angle, offset and mass per unit length. PCMass also provides statistics in a database for individual images or groups of images.

Small- and Wide-Angle X-ray Scattering Measurements. Synchrotron SAXS/WAXS measurements were performed at the 12-ID-B beamline of Advanced Photon Source at Argonne National Laboratory. A SAXS/WAXS simultaneous setup was utilized and the sample-to-detector distances were set such that the overall scattering momentum transfer q range was achieved from 0.005-2.3 $Å^{-1}$ for **Barrel3CLys** and **D-Barrel3CLys**, 0.005-2.3 $Å^{-1}$ for **Barrel3CArg**, where $q = 4\pi sin\theta/\lambda$, 2θ denoting the scattering angle and λ the x-ray wavelength. The wavelength was set at 1.033 Å during the measurements. Scattered X-ray intensities were measured using a Pilatus 2M (DECTRIS Ltd) detector for SAXS and Pilatus 300K for WAXS. SAXS/WAXS measurements were performed on aqueous solutions of peptide **Barrel3CLys**, **D-Barrel3CLys** and **Barrel3CArg** at concentrations of 1 mM in acetate buffer (10 mM, pH 4.0) at 25 ℃. A flow cell equipped with a quartz capillary (1.5 mm diameter) was used to prevent the radiation damage. Twenty images were collected for each sample and buffer. The 2-D scattering images were converted to 1-D SAXS curves through azimuthally averaging after solid angle correction and then normalizing with the intensity of the transmitted x-ray beam, using the software package at beamline 12ID-B. The 1-D curves of the samples were averaged and subtracted with the background measured from the corresponding buffers. The R_c analysis was done in Igor Pro software (WaveMatrics, Inc.) using the following modified

Guinier equation [33]: $\ln[q * I(q)] = -R_c^2 q^2/2$, where I(q) is the scattering intensity at momentum transfer q.

4.6 References

1. Korten, T., A. Månsson, and S. Diez, *Towards the application of cytoskeletal motor proteins in molecular detection and diagnostic devices.* Current opinion in biotechnology, 2010. **21**(4): p. 477-488.
2. Tang, J.H., et al., *Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22.* Structure, 2011. **19**(4): p. 496-502.
3. Shnyrova, A.V., et al., *Geometric catalysis of membrane fission driven by flexible dynamin rings.* Science, 2013. **339**(6126): p. 1433-6.
4. Bocquet, N., et al., *X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation.* Nature, 2009. **457**(7225): p. 111-4.
5. Wang, W.J., et al., *The structure of an open form of an E. coli mechanosensitive channel at 3.45 angstrom resolution.* Science, 2008. **321**(5893): p. 1179-1183.
6. Koronakis, V., et al., *Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export.* Nature, 2000. **405**(6789): p. 914-9.
7. Olia, A.S., et al., *Three-dimensional structure of a viral genome-delivery portal vertex.* Nature Structural & Molecular Biology, 2011. **18**(5): p. 597-U107.
8. Ballister, E.R., et al., *In vitro self-assembly of tailorable nanotubes from a simple protein building block.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3733-8.
9. Nishiyama, N., *Nanomedicine: nanocarriers shape up for long life.* Nat Nanotechnol, 2007. **2**(4): p. 203-4.
10. Geng, Y., et al., *Shape effects of filaments versus spherical particles in flow and drug delivery.* Nat Nanotechnol, 2007. **2**(4): p. 249-55.

11. Mao, C.B., et al., *Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires.* Science, 2004. **303**(5655): p. 213-217.

12. Lee, S.W., et al., *Ordering of quantum dots using genetically engineered viruses.* Science, 2002. **296**(5569): p. 892-895.

13. Childers, W.S., et al., *Peptide membranes in chemical evolution.* Curr Opin Chem Biol, 2009. **13**(5-6): p. 652-9.

14. Mehta, A.K., et al., *Facial symmetry in protein self-assembly.* J Am Chem Soc, 2008. **130**(30): p. 9829-35.

15. Wong, G.C., et al., *Hierarchical self-assembly of F-actin and cationic lipid complexes: stacked three-layer tubule networks.* Science, 2000. **288**(5473): p. 2035-9.

16. Raviv, U., et al., *Cationic liposome-microtubule complexes: pathways to the formation of two-state lipid-protein nanotubes with open or closed ends.* Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11167-72.

17. Petrov, A., S. Lombardo, and G.F. Audette, *Fibril-mediated oligomerization of pilin-derived protein nanotubes.* Journal of Nanobiotechnology, 2013. **11**.

18. Ghadiri, M.R., et al., *Self-Assembling Organic Nanotubes Based on a Cyclic Peptide Architecture.* Nature, 1993. **366**(6453): p. 324-327.

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

20. Reches, M. and E. Gazit, *Casting metal nanowires within discrete self-assembled peptide nanotubes.* Science, 2003. **300**(5619): p. 625-627.

21. Kumara, M.T., et al., *Bioengineered flagella protein nanotubes with cysteine loops: Self-assembly and manipulation in an optical trap.* Nano Letters, 2006. **6**(9): p. 2121-2129.

22. Woolfson, D.N., et al., *New currency for old rope: from coiled-coil assemblies to alpha-helical barrels.* Curr Opin Struct Biol, 2012. **22**(4): p. 432-41.

23. Calladine, C.R., A. Sharff, and B. Luisi, *How to untwist an alpha-helix: structural principles of an alpha-helical barrel.* J Mol Biol, 2001. **305**(3): p. 603-18.

24. Walshaw, J. and D.N. Woolfson, *Open-and-shut cases in coiled-coil assembly: alpha-sheets and alpha-cylinders.* Protein Science, 2001. **10**(3): p. 668-673.

25. Liu, J., et al., *A seven-helix coiled coil.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(42): p. 15457-15462.

26. Bulheller, B.M., et al., *Flow linear dichroism of some prototypical proteins.* Journal of the American Chemical Society, 2009. **131**(37): p. 13305-13314.

27. Dafforn, T.R. and A. Rodger, *Linear dichroism of biomolecules: which way is up?* Current Opinion in Structural Biology, 2004. **14**(5): p. 541-546.

28. Marrington, R., et al., *Micro-volume couette flow sample orientation for absorbance and fluorescence linear dichroism.* Biophysical Journal, 2004. **87**(3): p. 2002-12.

29. Dafforn, T.R., et al., *Protein fiber linear dichroism for structure determination and kinetics in a low-volume, low-wavelength couette flow cell.* Biophysical Journal, 2004. **86**(1): p. 404-410.

30. Hodgkinson, J.L., et al., *Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout.* Nat Struct Mol Biol, 2009. **16**(5): p. 477-85.

31. Wall, J.S., et al., *Mass mapping of large globin complexes by scanning transmission electron microscopy.* Methods Enzymol, 2008. **436**: p. 487-501.

32. Koch, M.H.J., P. Vachette, and D.I. Svergun, *Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution.* Quarterly Reviews of Biophysics, 2003. **36**(2): p. 147-227.

33. Glatter, O. and O. Kratky, *Small-angle X-ray Scattering*. 1982, London: Academic Press.

34. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria.* Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.

35. Egelman, E.H., *Single-particle reconstruction from EM images of helical filaments.* Curr Opin Struct Biol, 2007. **17**(5): p. 556-61.

36. Egelman, E.H., *Reconstruction of helical filaments and tubes.* Methods Enzymol, 2010. **482**: p. 167-83.

37. De Rosier, D.J. and A. Klug, *Reconstruction of three dimensional structures from electron micrographs.* Nature, 1968. **217**(5124): p. 130-4.

38. Moore, P.B., H.E. Huxley, and D.J. DeRosier, *Three-dimensional reconstruction of F-actin, thin filaments and decorated thin filaments.* J Mol Biol, 1970. **50**(2): p. 279-95.

39. DeRosier, D.J. and P.B. Moore, *Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry.* J Mol Biol, 1970. **52**(2): p. 355-69.

40. Wagenknecht, T., et al., *Three-dimensional reconstruction of the flagellar hook from Caulobacter crescentus.* J Mol Biol, 1981. **151**(3): p. 439-65.

41. Egelman, E.H., *A robust algorithm for the reconstruction of helical filaments using single-particle methods.* Ultramicroscopy, 2000. **85**(4): p. 225-234.

42. Baker, M.L., et al., *Analyses of subnanometer resolution cryo-EM density maps.* Methods Enzymol, 2010. **483**: p. 1-29.

43. DiMaio, F., et al., *Cryo-EM model validation using independent map reconstructions.* Protein Science, 2013. **22**(6): p. 865-868.

44. Zhang, J., et al., *Cryo-EM structure of a group II chaperonin in the prehydrolysis ATP-bound state leading to lid closure.* Structure, 2011. **19**(5): p. 633-9.

45. Walshaw, J. and D.N. Woolfson, *SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures.* Journal of Molecular Biology, 2001. **307**(5): p. 1427-1450.

46. Gill, S.C. and P.H. Von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data.* Analytical biochemistry, 1989. **182**(2): p. 319-326.

47. Zhang, X.K., et al., *Cryo-EM structure of the mature dengue virus at 3.5-angstrom resolution.* Nature Structural & Molecular Biology, 2013. **20**(1): p. 105-U133.

Chapter V: Conclusion and outlook

5.1 Conclusion

With their relatively straightforward intermolecular interfaces, coiled coils serve as a model system for studying rules of protein self-assembly and rationally designing complex peptide supramolecular assemblies. Surrounding this ubiquitous structural motif, the studies described in this volume span different research topics including controlling peptide self-assembly through metal-induced registry shift [1], recoding heptameric coiled coil for nanotube formation [2], and designing large-diameter helical nanotubes (manuscript in preparation). The overall purpose of these studies was to define the guiding principles for the rational design of structurally defined protein-based materials and to explore primary rules that govern the self-assembly of synthetic peptides, more specifically, coiled-coil peptides.

In the **TZ1C2** project, a mechanism to control supramolecular assembly of a synthetic peptide-based material through registry selection was described. This process recapitulates the native mechanism of fibril assembly in which a ligand binding event gates a reversible conformational transition between alternate forms of a folded peptide structure. A variety of characterization techniques were employed to verify our design. It was successfully demonstrated that metal ion coordination could reversibly control chain registry and, consequently, the assembly state of a designed nanomaterial. The control of chain alignment represents an attractive strategy for the design of dynamically reconfigurable nanoscale materials.

In the **7HSAP1** project, design of a structurally defined helical assembly is described that involves recoding of the amino acid sequence of peptide **GCN4-pAA**, a seven-helix coiled-coil bundle that resembles a supramolecular lock washer. Structurally informed mutagenesis of the sequence of **GCN4-pAA** afforded peptide **7HSAP1**, which undergoes self-association into a nanotube via noncovalent interactions between complementary interfaces of the coiled-coil lock-washer structures. Biophysical measurements conducted in solution and the solid state over multiple length scales of structural hierarchy are consistent with self-assembly of nanotube structures derived from 7-helix bundle subunits. Fluorescence studies of the interaction of **7HSAP1** with the solvatochromic fluorophore PRODAN indicated that the nanotubes could encapsulate shapeappropriate small molecules with high binding affinity.

In the **Barrel3C** project, the efforts toward higher-order coiled-coil assemblies were presented, which lead to the discovery of two unprecedented helical supramolecular assemblies which hold great significance in understanding the sequence-structure relationship of coiled coils and in designing functional tubular or porous protein-based materials. More interestingly, a single-residue mutagenesis completely alters the preference for two extinct supramolecular structures, which partially validate the idea about the liability of protein quaternary structures.

5.2 Outlook

The studies described in this volume as well as studies from other research groups [3, 4] have demonstrated notable degree of success in the design of self-assembling protein-based architectures. However, the complexity still pales in comparison to what nature exhibits all around us. I want to show the illustrative pictures here again. Every time I look at them, I am amazed by how crowded and ordered life is. How far can we push chemical self-assembly? [5] Can we think someday in a nature's way? These are big questions facing science.

The solution to these questions will involve cracking and formulating the complex inter- and intra-molecular interactions of biopolymers, especially proteins where forces that drive the self-assembly are far more complicated than other macromolecules. One way to do this is to apply computational advancements to our current understanding of folding and assembling of sequence-defined macromolecules and to develop more accurate forcefield to fully depict the energy landscape in atomic details. A recently reported study demonstrated the computational design of self-assembling protein nanomaterials with atomic level accuracy [4], which brings us hope for the design of more complex supramolecular structures. However, there is still a long way to go on the road of pursuing nature-level complexity. Although some synthetic viruses and designer viruses have been created in test tubes over the past decade by several research groups, they all utilized the existing generic information encoded in the original viral genome. Up until now, there is hardly any reports on the successful construction of de novo designed co-assembling systems with two or more type of biological macromolecules, e.g. protein and nucleic acid, the simplest form of life on early.

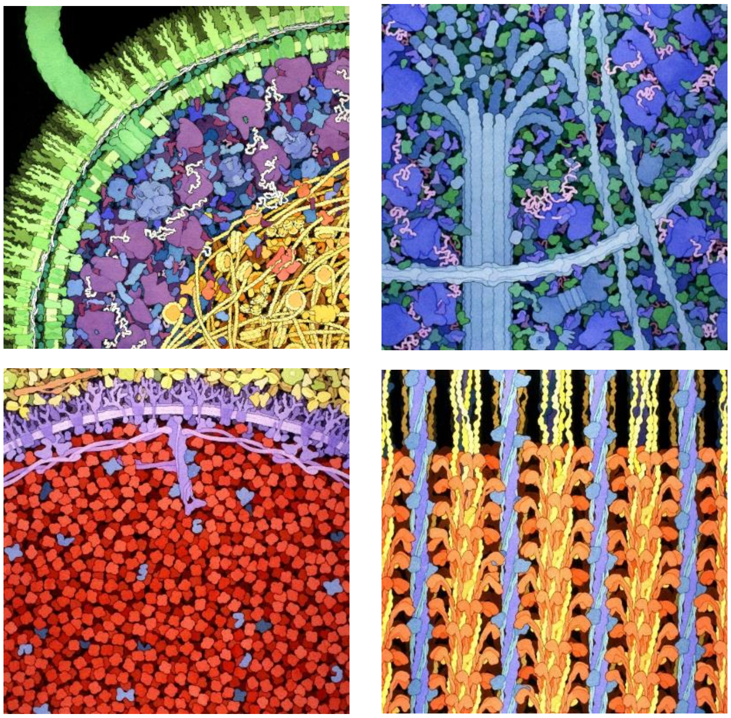


Figure 5.1. Beautiful artwork by chemical self-assembly. Illustration by David S. Goodsell, the Scripps Research Institute.[6] Top Left: A cross-section of a small portion of an Escherichia coli cell; Top Right: A small portion of cytoplasm including three types of filaments that make up the cytoskeleton; Bottom Left: A portion of a red blood cell with hemoglobin in red; Bottom Right: Part of a muscle sarcomere with actin filaments in blue and myosin filaments in red.

Another way of addressing these big questions is to employ an evolutionary approach in which the comprehension nature’s way of thinking is explored from the very beginning, a simpler world. Hopefully the study of chemical self-assembly and chemical evolution could fuse at certain point in the future to acquire a clear understanding of Mother Nature.

5.3 References

1. Anzini, P., et al., *Controlling self-assembly of a peptide-based material via metal-ion induced registry shift.* J Am Chem Soc, 2013. **135**(28): p. 10278-81.

2. Xu, C., et al., *Rational design of helical nanotubes from self-assembly of coiled-coil lock washers.* J Am Chem Soc, 2013. **135**(41): p. 15565-78.

3. Fletcher, J.M., et al., *Self-assembling cages from coiled-coil peptide modules.* Science, 2013. **340**(6132): p. 595-9.

4. King, N.P., et al., *Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy.* Science, 2012. **336**(6085): p. 1171-1174.

5. Service, R.F., *How far can we push chemical self-assembly.* Science, 2005. **309**(5731): p. 95-95.

6. Goodsell, D.S. *Illustrations for Public Use.* http://mgl.scripps.edu/people/goodsell/illustration/public. [cited 2013 November 16th].