Distribution Agreement

Signature:

____________________________ ____________________

Ordy Manuela Gnewou T Date

Cryo-EM as a tool to understand cross-α and β nanotubes peptide systems

By

Ordy Manuela Gnewou T
Doctor of Philosophy
Chemistry

Dr. Vincent P. Conticello
Advisor

Dr. David Lynn
Committee Member

Dr. Bill Wuest
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Cryo-EM as a tool to understand cross-α and β nanotubes peptide systems

By

Ordy Manuela Gnewou T.
B.A., Lehman College, City University of New York, 2018

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

2022

Abstract

Cryo-EM as a tool to understand cross-α and β nanotubes peptide systems

By Ordy Manuela Gnewou T

Chemical self-assembly is a necessary tool in the creation of *de novo* systems that would rival the level of complexity observed in nature. Within the scope of peptides nanomaterials, rational design is a vital means to help push the current boundaries within the protein design space. Cryo-EM, on the other hand, has become the standard rather than the exception in helping the scientific community bridge our knowledge gap of the structural interactions that govern diverse assembly systems, both natural and synthetic. To this end, several research projects will be presented here with the goal of contributing to the establishment of a set of design rules that would help us rationalize the relationship between sequence and the resulting structure of peptides nanotubes in ways that are both predictable and reproducible. These projects include the structural analysis of cross-α helical nanotubes as a means to gain insight into the designability of peptides materials, the structural and mechanistic study of cross-β sheet nanotubes from an amphipathic oligopeptide as well, as an overview of cryo-EM as it applies to helical polymers. In addition to electron microscopy in all its forms, we employed a wide spectrum of biophysical analyses to study higher-order structure. These projects contributed to widening the structural and mechanistic "toolbox" for self-assembling peptide materials while providing a better understanding of the key principles in peptide assembly.

Cryo-EM as a tool to understand cross-α and β nanotubes peptide systems

By

Ordy Manuela Gnewou T.
B.A., Lehman College, City University of New York, 2018

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

2022

Acknowledgments

I would like to sincerely thank my advisor Professor Vincent Conticello for his support, patience, and guidance over the course of my Ph.D. path. I could not have asked for a better advisor. He has led by example, setting high standards for academic research while taking into consideration my personal and professional growth. I am grateful to have had him as my mentor for four years and will carry forth the lessons he has taught me wherever life takes me. I would also like to express gratitude to those who have served as my committee members, Professor David Lynn and Professor Bill Wuest. They have elevated the quality of my research significantly with their wisdom, insight, and direction throughout my Ph.D. trajectory and have never failed to support my decisions as I navigate the world of graduate school. I must also thank the members of the Conticello lab who have helped me immensely: Jessalyn Miller, Duong Nguyen, Breana Laguera, and Abhinaba Das. These people have significantly improved both my research and my everyday life, providing a wonderful work environment as well as assistance with experimentation, insightful discussions, and encouragement. I am also grateful for former lab members as well, Dr. Spencer Hughes, Dr. Shengyuan Wang, Dr. Andrea Merg, Avi Tuachi, Gavin Touponse, George Cheng, and Claire Ogilvie. They have provided great assistance and encouragement along the way. I would also like to acknowledge a graduate student, not in the Conticello lab, but who has been critical to my success: my chemistry community, including the folks in NOBCChE particularly, the other student groups I have been a part of, the chemistry administrative staff and faculties who have held me together over the past few years (Antonio, Kira, Tracy, Kate, Todd). Thank you for helping me bring wisdom to my approach to not only research but also life in grad school.

A special thanks go to all the members of Robert. P. Apkarian Integrated Electron Microscopy Core, Dr. Ted, Joseph Whitworth, Dr. Ravi Koripella, Ms. Hong Yi, Ms. Jeannette Taylor, and Mr. Art McCanna, who has left us. A special thank you to Dr. Ricardo Guerrero C. and Dr. Puneet Juneja. Their mentorship helped me learn so much in the field of electron microscopy. They gave me the opportunity to understand all the aspects of data processing and analysis in a healthy environment with all the patience in the world. Doing so substantially improved my research and knowledge of the field.

I would like to express my gratitude for all of our wonderful collaborators over the past four years. Professor Ed Egelman (University of Virginia) and Dr. Fengbin Wang have been instrumental in helping us solve the multiple structures of some of the most complex peptide systems. Dr. Xiaobing Zuo (Argonne National Lab) has provided constant support in the field of small-angle X-ray scattering, and without his assistance, much of our work would not be possible. Professor Gevorg Grigoryan (Dartmouth University) provided the inspiration and sequences for many of our computationally based projects. Dr. Fred Strobel has been a constant source of assistance with mass spectrometry over the years.

My time in graduate school was supplemented by the pursuit of the Chateaubriand Fellowship in France in the Spring of 2021. There I worked with incredible people. Dr. Maité Paternostre and Dr. Human Rezaei. They provided me with the tools I needed to refine my approach to problem-solving and science presentation, constantly giving me room to grow, all of which made me a better scientist.

I must also thank those who inspired my scientific career and decision to pursue a Ph.D. Dr. Donna McGregor, Dr. Benjamin Burton-Pye, Professor Gustavo Lopez, Professor Pamella

Mills, and Michelle Alice Augustine are all from my Undergraduate Institution, Lehman College. These people were monumental in my decision to pursue a career in science. I thank them for being such positive role models. Without the help of all of these people, my research and Ph.D. career would be profoundly lacking.

Lastly, I acknowledge my family, who has provided incredible and unwavering support throughout my life. My sister for being my greatest source of support and eternally brings a smile to my face during difficult times. My brothers for always being there whenever I needed them during my time here at Emory, and most important of all, my mother has worked tirelessly to help put me into positions and schools where I could succeed and grow. She did the best she could with what she had. Without her love and support, I would not be where I am today.

Table of Content

List of Figures

List of tables

Chapter 1: Introduction

1.1- The importance of peptides self-assemblies

Self-assembly describes the process by which the components of a system, be it molecules or macroscopic particles, associate into a highly ordered pattern or structure. The hierarchical packing in self-assembly processes is often governed by thermodynamic principles[1], leading to specific local interactions among the individual parts. Such processes are observable across multiple multiple length-scales from the nanometer to the micrometer range. For instance, individual molecules organize in cellular nanostructure that further associate into cells, cells into tissues, tissues into organisms, etc. The observable molecular architectures can be obtained from a diverse set of building blocks such as phospholipids, peptides, and nucleic acids[2]. This thesis will focus entirely on the self-assemblies of peptides materials at the nanoscale level. Peptides can be constructed from a vast library of natural and synthetic amino acids. Structurally defined nanomaterials with sequence specificity such as theirs have physicochemical properties that grant them thermal, electrical, and mechanical properties, allowing them to be used in various fields. In recent years, peptides nanomaterials have been widely studied and used in biological and non-biological systems.

1.1.1- Peptides self-assemblies in biological systems

Evolutionary processes have resulted in an impressive set of biological machinery that are able to operate with high efficiency. It allows its organisms to constantly build, break down and recycle simple monomers into supramolecular assemblies. These engines are mostly found in living systems where they can accomplish various tasks cooperatively at speeds sometimes approaching the diffusion limit. Nature utilized a narrow set of elementary motifs such as α-helix and β-sheets secondary structure, tuning their interactions to optimize function. These biological self-associations can be found in many organisms, the most studied ones being archaea and bacteria, because of our ability to extract and purify them for high-resolution structural analyses[3]. To illustrate how much work has been done in this domain by the scientific community, it had been reported that as of December of 2021, 185,541 biological macromolecular structures were deposited in the Protein Database Bank (PDB; http://www.rcsb.org/pdb/) including 157,709 naturally occurring molecular structures, which include proteins, peptides, and viruses that display a critical roles in biological systems. Deep dives into their networks and modes of action have allowed us to understand the fundamental principle that governs native proteins' self-assembly processes and their involvement in activities such as cell migration, endocytic vesicle movements, and membrane-bound organelle as well as intercellular transport of small molecules[4,5,6]. While most protein sequences counting several hundreds of amino acids as building blocks are often encountered, substantially shorter sequences can also exhibit a high level of structural complexity in the way they assemble and are utilized[7,8]. In that respect, the intersection of fields such as chemistry, molecular biology, and materials science and engineering has allowed us to prepare synthetic materials with a level of complexity approaching their natural product counterpart[9,10,11]. Some of the most notable applications of these bio-inspired shorter peptides-based systems include the production of new antimicrobial agents[11], drug delivery generation of 3D hydrogel scaffolds for tissue engineering[12], cancer diagnostics and therapy[13], and much more.

1.1.2- Peptides self-assemblies in non-biological systems

A better understanding of the factors influencing the self-assembly process has allowed the development of peptide-based materials with utilities beyond biological systems. These factors primarily include foundational elements such as the properties of the side chains but also the number and sequence of amino acids within the given sequences. Second, to those, we count environmental factors such as temperature, pH, solvents, and even ultrasounds. In addition to the more common properties of peptides nanomaterials such as thermal and mechanical stability, they have been found to have electrical field control capabilities, conductivity, and semi-conductivity[13], piezoelectricity, ferroelectric[14], and optical properties[15]. These properties can be utilized to make peptide-based electrodes that can be used to build supercapacitors, peptides-based biosensors devices[4], as well as the scalable synthesis of melanin materials for cosmetic use[16]. The supramolecular structure formed can be classified into categories often based on their structures or functions. Moving forward, we will focus on their ability to self-assemble into helical filaments. To capitalize on the diversity in shape, size, properties, and functions, we need to be to interrogate their structure across multiple length scales.

1.2- Structural characterization of helical assemblies

Over the years, several experimental methods have been developed to characterize assemblies' structures. These methods can be divided into two main categories: Low and high-resolution methods (see **Table** 0.1) Low-resolution methods are often used as means to rapidly screen the assemblies to identify the propensity to form ordered assemblies and provide initial structural insight. While we can count multiple of those techniques, there are only a handful of methods able to give information at high resolution. The high-resolution methods can be further broken down into subcategories: near-atomic resolution (NAR) and true atomic resolution. True atomic resolution refers to the ability of a model to distinguish between separated atoms in an electron density map[5] and is agreed to be at ca. 1.2 Å. The NAR level, however, has been subjected to controversy but is usually employed to describe the resolution of density maps that can serve as the basis for the construction of reliable atomic models with a resolution limit claimed to be at ca. 4Å. Throughout this thesis, we make use of most of the low-resolution methods presented in this section. High-resolution structural characterization of our assemblies is, however, grounded in the use of cryo-EM for helical filaments and tubes as it presents multiple advantages over the other methods.

Table 0.1: Most common experimental methods for structural characterization of helical biomolecules.

Resolution	Method	Abbreviation	Usual Function
Low	Circular Dichroism	CD	Provide evidence of the formation of secondary and tertiary structures
	Fourier-transform infrared	FTIR	
	Raman Spectroscopy	Raman	
Low to Medium	Small angle X-Rau scattering	SAXS	Provide information on the structural features of the assemblies in solution[17]
	Small angle neutron scattering	SANS	
Medium	Small angle neutron scattering	SANS	Can allow assignment of the helical hand of the assemblies for cases in which the chirality of the surface features can be resolved[18]
	Atomic force microscopy	AFM	Can reveal the presence of supramolecular structure and provide insight into the morphological features of the resultant assemblies at medium resolution
	Transmission electron microscopy	TEM	
High	Single crystal X-ray diffraction	Crystallography	Structural determination of assemblies at NAR or true atomic resolution
	Fiber diffraction		
	Microcrystal electron diffraction	MicroED	
	Solid state nuclear magnetic resonance	ssNMR	
	Cryo-Electron Microscopy	Cryo-EM	

Among the available high-resolution methods, single-crystal X-ray diffraction has played a key role in our ability to understand short amyloidogenic peptide sequences. It requires the growth of suitable single crystals and the imposition of a crystallographic symmetry which represent challenges in the application of the method to helical assemblies for two main reasons. The first is that helical assemblies are difficult to crystalize because of their limited solubility, length variability, and structural polymorphism. They will often require the introduction of mutations to achieve the conditions suitable for crystallization. Second is the fact that helical arrays of chiral rod-like molecules must incur an energy penalty associated with elastic distortion of the protomers

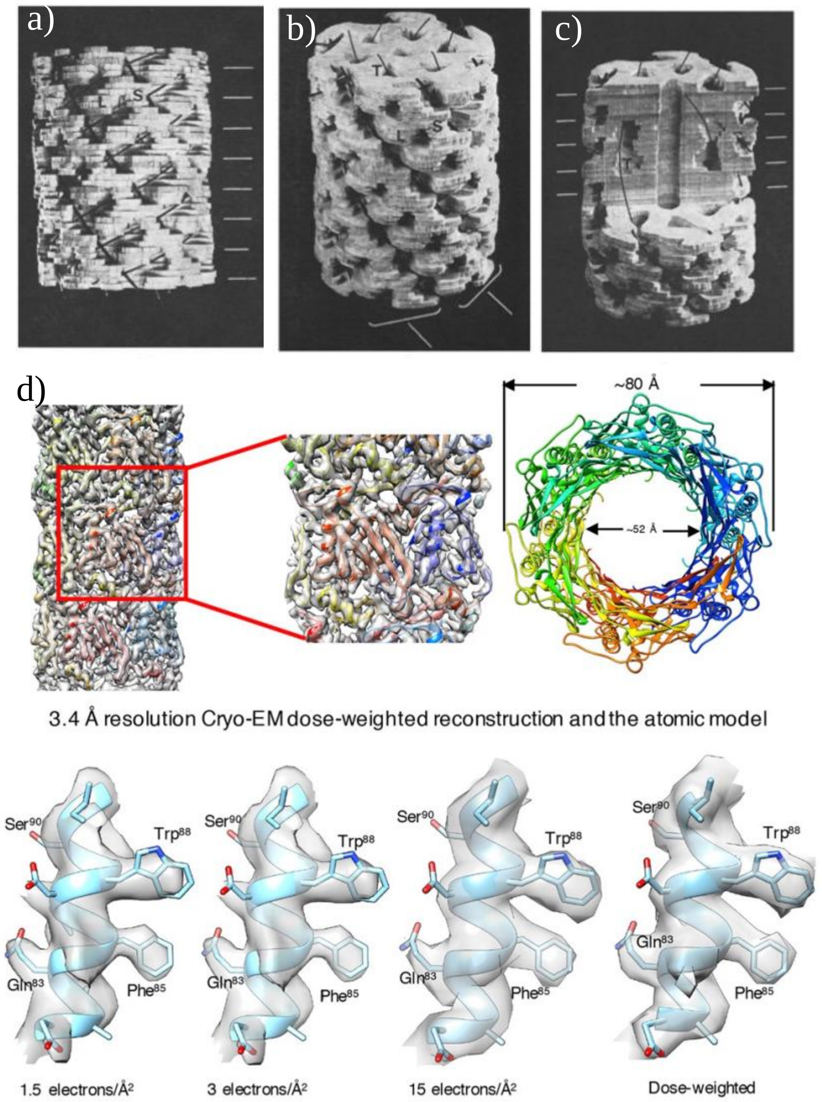


Figure 1-0-1 The top of the image is a model of the phage tail in the extended state at 35Å resolution from 1968 (De Rosier, 1968). a) the side view, b) model tilted forward, c) same as b but with part of the front half removed- Reprinted by permission from Reconstruction of three dimensional structures from electron micrographs, D.J. De Rosier, A. Klug , Copyright (1968), with permission from Elsevier. The bottom is the 3.4Å resolution map of the same bacteriophage tail tube solved in 2017 - Reprinted by permission from Refined Cryo-EM Structure of the T4 Tail Tube: Exploring the Lowest Dose Limit, Weili Zheng, Fengbin Wang, Nicholas M.I. Taylor, Ricardo Guerrero-Ferreira, Petr G. Leiman, Edward H. Egelman, Copyright (2017), with permission from Elsevier

molecules held[19] to be able to easily accommodate translational symmetry, which can significantly limit the dimension of the crystal. Finally, Peptides assemblies have a wide range of helical symmetries. They can go beyond the set of available crystallographic screw symmetries-restricted to two-, three-, four-, and six-fold rotations. Single state NMR, on the other hand, has been associated with scanning transmission electron microscopy (STEM) -mass per length measurements or cryo-EM[20] to address the issues of accurate determination of the interatomic distance. The structural polymorphism observed in most helical assemblies can significantly increase the effective number of resonances in NMR experiments, thus complicating the analysis. As a result, the use of cryo-EM has revolutionized our ability to get more accurate structural information from helical filaments. Our ability to preserve isolated and dispersed filaments as a thin film in hydrated states[21] through cryo vitrification reduces the probability of introducing structural artifacts like in conventional TEM. Advances in technology have allowed us to increase information content by ~ 1000-fold from the ∼35 Å resolution structure of Tobacco Mosaic Virus (TMV) that was obtained using a negatively stained sample in 1968 to a 3.4 Å resolution of the same bacteriophage T4 tail tube using cryo-EM helical reconstruction field[22] (**Figure 1-0-1**). While helical reconstruction of cryo-EM images can also present some difficulties, such as the correct assignment of the helical symmetry, the helical hand, and the ability to get optimal ice conditions depending on the sample, it still has many more advantages over the others. It has a high tolerance for the presence of impurities in the sample[23], as long as the targeted assemblies can clearly be distinguished on the micrograph. It can be applied to a small sample size (0.1-1mg/ml solution) and does not require heavy atom or isotopic labeling. As a result, the development of cryo-microscopy with direct electron detection capabilities has made cryo-EM the method of choice for structural characterization of helical polymer.

1.3- Self-assembly of helical biomolecules

In both prokaryotic and eukaryotic cells, most protein (by mass) can form helical assembliess. Well-known examples are actin, tubulin[24], flagellin, and pili[25]. These supramolecular assemblies can be constructed from peptides[26, 27] or non-peptide-based foldamers[28-30]. While the latter is not to be neglected and is distinguished itself from the former through differences in the chemical identity and stereochemistry of the backbone, this thesis will focus on the self-assembly of helical assemblies derived solely from peptides building blocks. The self-assembly behavior of the peptide building blocks can be influenced through the incorporation of hydrophilic and hydrophobic domains and/or the addition of charged amino acids in ways that will favor the formation of α-helix or β-sheets. Before we dive into the field of rational design of helical nanotubes, it is essential to have a good understanding of systems that already exist in nature as they represent a source of inspiration for *de novo-designed* ones.

1.3.1- Native helical assemblies

Native biological filaments display a structural hierarchy that ranges across length scales from the nanoscale to the mesoscale size regime. This degree of structural control is highly desirable and probably necessary for the evolution of complex functions, such as those observed in native filamentous assemblies, e.g., filaments that form extracellular proteins and viruses. However, it remains difficult, if not impossible, to achieve this degree of structural and functional control thus far for synthetic assemblies. As introduced above, one of the most studied self-assembly structures bearing a helical motif in natural systems is the TMV. It became widely popular due to its virulence against humans. It was first crystallized in 1935[31] and has since then been used in a variety of studies. The TMV assembles into a 300 nm long nanotube with an 18 nm thickness and a hollow 4 nm lumen in diameter (**Figure 1-0-2**). Its assembly is made of single-stranded RNA inserted between successive turns of right-handed coat protein superhelix. While multiple modes of assemblies have been stipulated, it was shown to form by "disk aggregation" of about 34 coat proteins subunits (roughly 2 super helical turns[32]) rather than the single subunit addition model as it better accounted for the rapid self-assembly observed. TMV is currently used as a mass-per-length standard in negative-stain STEM[22]. Because of its well know mass-per-length and mass-per-area, assemblies of interactions are incubated with TMV. The use of contrast between the TMV filament and the stained background allows the correlation of the brightness with mass measurements. The TMV virus has also been investigated for multiple applications, among which its use in antigen displays because of its well-ordered structure. Some have also used it as a conjugate with an antigen to show that the complex conjugate had a significantly more immunogenic character than the free antigen. Finally, it has been found to be a potential candidate structure for increasing surface area in micro-batteries for microelectronic machines (MEMs[33]).

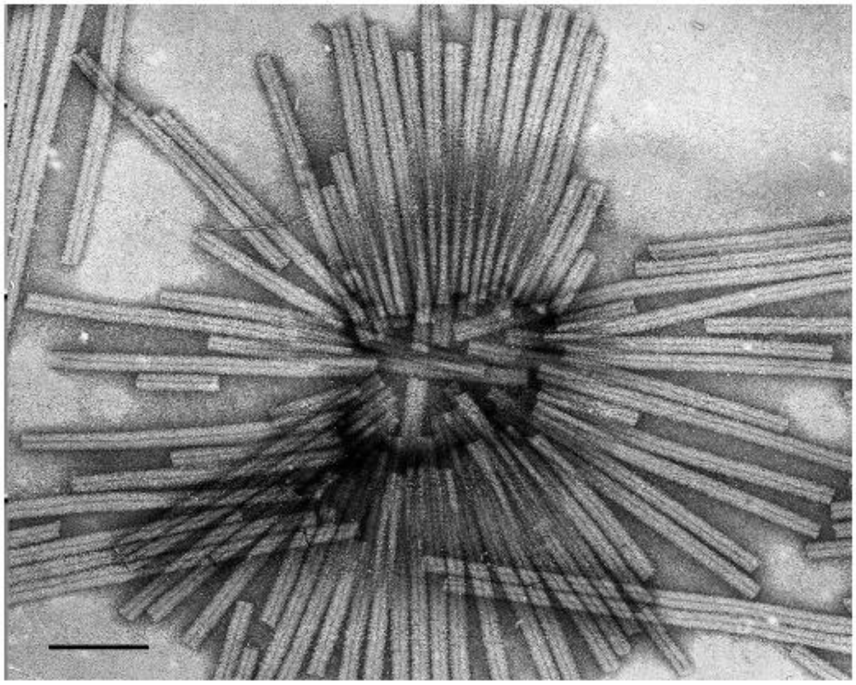


Figure 1-0-2: Negative-stained TEM of fully assembled TMV particles. Scale bar is 100 nm. Reprinted from "Archive of Biochemical Biophysics, 581, Harris, J. R., Transmission electron microscopy in molecular structural biology: A historical survey, 3-18." Copyright (2015) with permission from Elsevier.

Another great example of nature's helical assembly is the chaperone usher (CU) pili. They are comprised of two main components: a short tip fibrillin with three to four subunits and a superhelical rod comprising ~1000 copies of a single pilin protein, which can extend 1-2 microns from the bacterial cell surface[33]. CU pili play a critical role in biofilm formation as well as host cell recognition and adhesion. Their mode of action involves a coiling and uncoiling motion in response to shear stresses encountered during fluid flow *in vivo* (**Figure 1-0-3**). Resolved by cryo-EM, their structure reveals their general functions as well as their assembly mode[34]. It was found that the polar interactions holding the quaternary structure of the helical rod together would progressively break down under high shear forces. This does not, however, impact the integrity of the pilus due to the strong hydrophobic interactions driving the strand exchange. Further

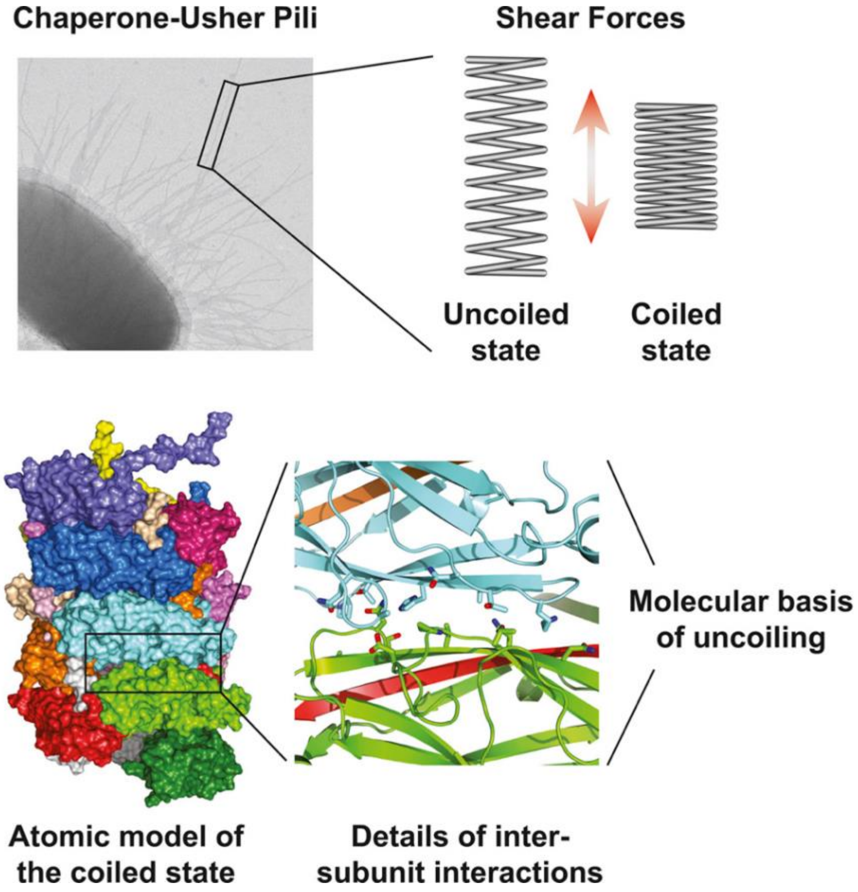


Figure 1-0-3: Atomic model of the P pilus generated from a 3.8 Å resolution cryo-electron microscopy reconstruction from Cell vol. 164,1-2 (2016): 269-278. doi:10.1016/j.cell.2015.11.049

experimentation is still required to fully understand the relationship between the formation of the quaternary structure and the extrusion process. These types of pilis can function as virulence factors for bacterial pathogens. Multiple approaches have been taken for therapeutic targeting the CU pili using other families as models. Vaccination and small-molecule inhibitor approaches generate protection against their virulence[35]. While multiple other native helical assemblies are still being discovered and studied, nonnative ones will be the focus of this thesis.

1.3.2- Cross-α peptides nanotubes

Non-native helical assemblies can be engineered from synthetic alpha-helical peptides systems. The native α-helix is characterized by a 3.6 amino acid residue per turn geometry. Depending on the nature of the amino acid (*L* or *D*), the helix can either be right or left-handed. It is formed because of intermolecular hydrogen bonds between C=O of residue i and N-H of residue i+4[36,37] with a bond distance of 2.72Å and rise per turn 5.4 Å. The termini of the helix are stabilized by the presence of charged residues, negative at the *N*-terminus and positive at the *C*-terminus. and represent the building block of a cross-α peptides nanotubes. These α-helices are usually associated by coiling with each other into super-helical assemblies known as coiled-coils[38]. Coiled coils are formed based on repeating patterns of a heptad motif (*a-b-c-d-e-f-g*). The amino acids at the *a*- and *d*-positions are usually hydrophobic, allowing the self-association through the 'knob into holes' packing style first proposed by Crick in 1953[39]. Rather than the 3.6 residues per turn of the regular helix, the coiled coils pack into 3.5 residues per turn, allowing the helices to supercoil into a left-handed super-helical sense. Multiple strategies can be employed to promote the formation of helical filaments from these bundles. One of the earliest methods employed involves a variation of the helix-helix registry within the assembly to promote staggered self-association[40]. It occurs by introducing electrostatic interactions between the helices at the e and g positions while burying the polar interactions at either the a and d positions, allowing for the control of helix-helix alignment within the assembly. Another strategy involves the stacking of discrete oligomeric bundles to form continuous superhelices along the highest order rotational axis[41]. This method

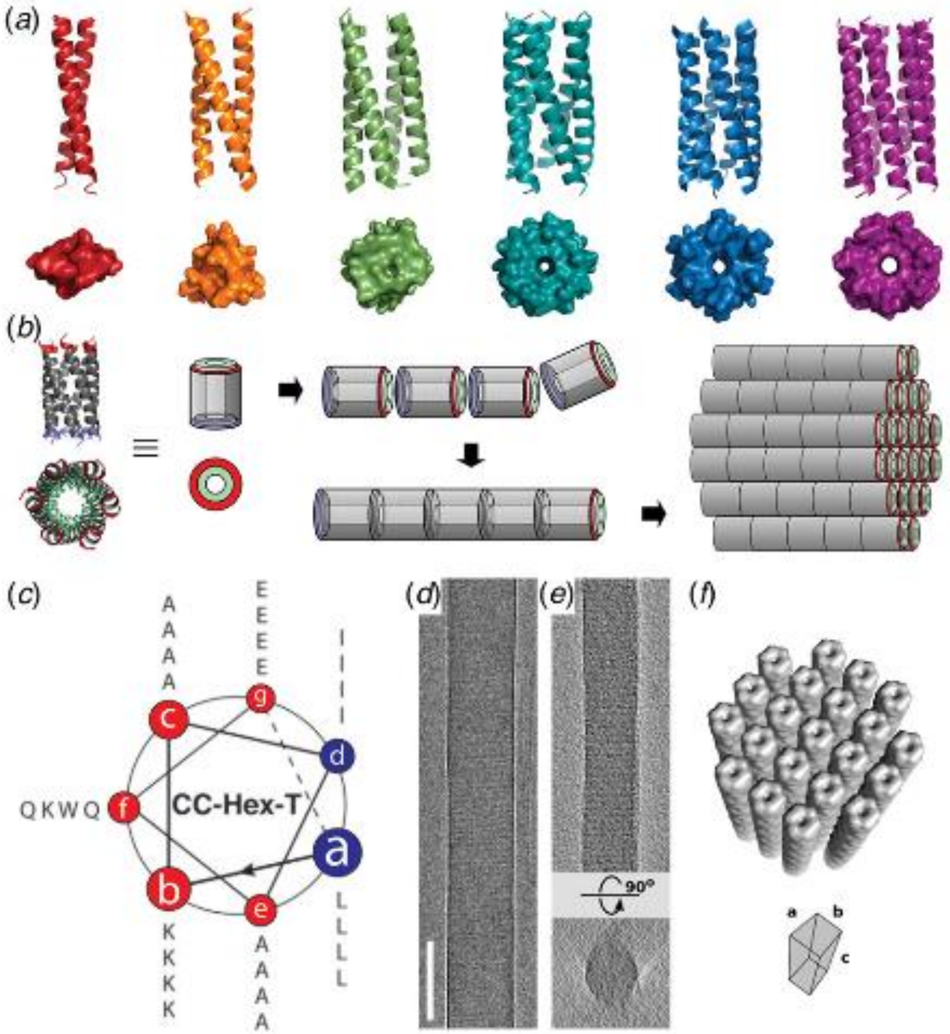


Figure 1-0-4: Blunt-ended coiled-coil oligomers self-assemble into peptide nanotubes. **(a)** Ribbon diagrams and orthogonal views of space-filling models for designed coiled-coil oligomers; CC-Di (red; PDB: 4DZM), CC-Tri (orange; 4DZL), CC-Tet (green; 3R4A), CC-Pent (turquoise; 4PN8), CC-Hex (blue; 3R3K), and CC-Hept (purple; 4PNA). **(b)** Proposed model for self-assembly of CC-Hex-T involving axial stacking of hexameric barrels into a filament**. (c)** Helical wheel diagram depicting the heptad register of the CC-Hex-T sequence. **(d)** Representative cryo-EM image of thermally annealed CC-Hex-T assemblies (scale bar = 100 nm). **(e)** Tomographic slice of a CC-Hex-T assembly from cryo-ET analysis depicting an approximately cylindrical cross section. **(f)** Tetragonal lattice model for packing of CC-Hex-T filaments in a para-crystalline array. Reprinted with permission from Burgess et al. (2015), Journal of the American Chemical Society, 137(33), 10554–10562. Copyright 2021 American Chemical Society. Please proof the figure captions and fix the borders

allows the formation of thermodynamically stable filaments by modifying the termini of the peptide sequence to promote axial interactions[42,43], especially if coupled with non-covalent interactions. Wolfson and coworkers utilized this method and extended the range of oligomerization states from trimers to heptamers (**Figure 1-0-4**). While these examples primarily involve individual helices aligning nearly parallel to the superhelical axis, other styles of packing involving helices stacking in bilayer arrays with individual peptides oriented perpendicular to the long axis of the filament have been observed.

One example is a short cytolytic peptide secreted from a virulent strain of staphylococcus aureus called phenol-soluble modulin (PSMs). Crystallographic analysis has shown that peptides form cross-α super-helical arrays through stacking of parallel dimers in an alternating antiparallel orientation. Though these filaments displayed limited interfaces over which KIH packing was observed between protomers[33,], their cytolytic membrane activity and their self-assembly behaviors have been attributed to amphipathic character[33,34] i.e. alternating patterns of hydrophobic and hydrophilic amino acids. The Conticello group and the Egelman group have explored multiple of those amphipathic peptide assemblies. Their most recent work used cryo-EM to analyze the self-assembly behavior of two distinct classes of PSMs, PSMα3 and PSMβ2, in aqueous solution[35]. These two classes are differentiated by their length, typically ~20 to 25 amino acids long for PSMα and ~44 to 45 amino acids long for PSMβ. They demonstrated that these PSM nanotubes displayed polymorphism in terms of global morphology (twisted filament, helical ribbons, and nanotube) and the number of protofilaments within a given peptide assembly[36] (**Figure 1-0-5**) . Analysis of the PSMα3 assemblies showed two distinct populations of nanotubes with diameters of ∼370 Å versus ∼410 Å. Their cryo-EM structures generated reconstructions at resolutions (map versus model) of 3.9 Å and a C_1 symmetry for the larger tube and 4.1 Å and a C_7

symmetry for the smaller one. PSMβ2, on the other hand, provided a volume at 4.3-Å with a C_6 symmetry containing 12 protofilaments.

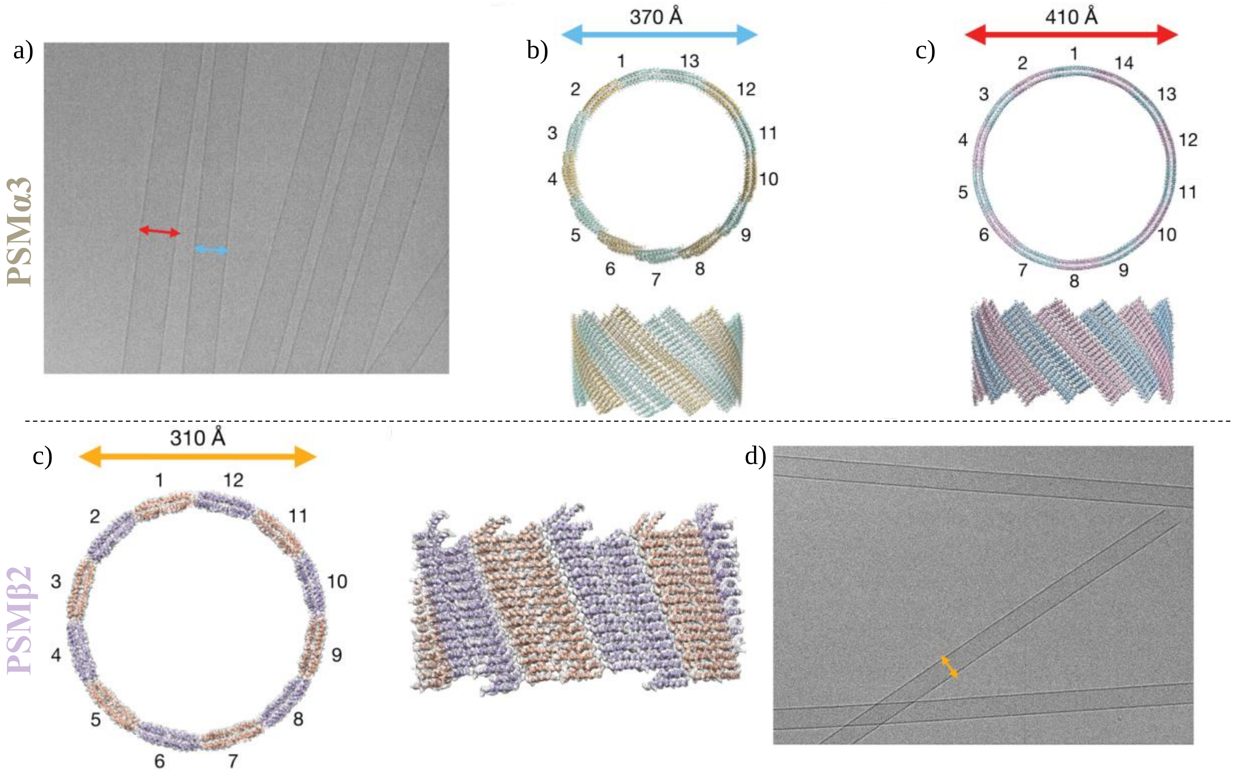


Figure 1-0-5: Cryo-EM structure of the PSMα3 and PSMβ2 nanotubes. (a) Cryo-electron micrograph of PSMα3 nanotubes of different diameters, the two most common diameters being near 370 Å (blue double arrow) and 410 Å (red double arrow). (b) Model of the smaller diameter tube fit into its density map (light gray). The top image shows a view through the lumen of the tube, while the bottom image shows a side view. The protofilaments are distinguished by alternating gold and green strands. (c) Model of the larger diameter tube fit into its density map (light gray). Top and bottom images are the same views as in B. The protofilaments are indicated as alternating pink and blue strands. (a) Top and side view of the PSMβ2 nanotube model fit into its density map Protofilaments are colored in alternating orange and purple. (d) Cryo-electron micrograph of the PSMβ2 nanotubes. The orange double arrow indicates a 310-Å diameter. Images adapted from Phenol-soluble modulins PSMα3 and PSMβ2 form nanotubes that are cross-α amyloids. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.2121586119. Epub 2022 May 9. PMID: 35533283; PMCID: PMC9171771.

Together, these structures can be used to demonstrate the importance of the cross-α fold in protein structures and self-assembled peptides. Advancements in the near-atomic resolution structures de novo designs from biologically derives peptides [37] open ups questions related to their designability and fold robustness[38]. Finally, despite the differences in their cohesive interactions and local conformations, their supramolecular architectures appear superficially similar to cross-β folds[39], which are the basis for amyloid fibrils structures[40,22,] as well as many designed β-sheet peptides.

1.3.3- Cross-β peptides nanotubes

Much like alpha-helical peptides assemblies, β-sheet assemblies are a widespread phenomenon in nature. They can be broken down into pathogenic and nonpathogenic amyloid structures. One nonpathogenic example includes Curli fibrils. They are extracellular amyloid protrusions of *E. coli* that enable colonization and adhesion to surfaces[44]. Pathogenic ones, however, have been associated with Alzheimer's disease, Parkinson's disease, spongiform encephalopathies, and many more[45]. The cross-β assemblies derive from individual peptides that have adopted extended β-strand conformation. The strands can be parallel or antiparallel, depending on their orientation. These β-sheets are stabilized by multiple interactions, which include directional hydrogen bonding between neighboring amide backbone groups as well as hydrophobic, aromatic, and coulombic interactions between adjacent side chain groups. They can then self-associate into cross-β filaments or nanotubes. The first synthetic assemblies of cross-β filaments were shown to display a facial amphiphilicity that resulted in amphipathic β-sheets[46, 47]. The introduction of complementary electrostatic and hydrogen bonding interaction between the side chains played a big role in making oligopeptides into stable β-sheet filaments[22]. While early designs assumed that the strands would be organized in an antiparallel fashion resulting in cross strands pairing between opposite charged residues within the fibrils, it was later shown through ssNMR analysis that some of these peptides can assemble into bilayer filaments from parallel β-sheet protofilements[4]. A different mode of self-association of these fibers includes metal coordination as a selective mechanism to help drive the assemblies but also as a method to introduce structural specificity within the peptide filaments[48, 49].

In some cases, these cross-β filaments can further self-associates to form cross-β nanotubes. Up-to-date synthetic design of these assemblies has either derived from amyloidogenic peptide sequences or from more complex β-peptides building blocks. Early structural characterization of β-sheet nanotubes from biologically derived to bio-inspired assemblies was performed by Lynn and coworking. They reported the self-association of an amyloidogenic truncated sequence from an Aβ peptide sequence Ac-^{16}KLVFFA^{22}E-NH$_2$ (Aβ$_{16-22}$). Their assembly resulted in nanotubes under acidic conditions- pH2^{50} and twisted filaments from a pair of protofilaments at under neutral conditions. (**Figure 1-0-6**).

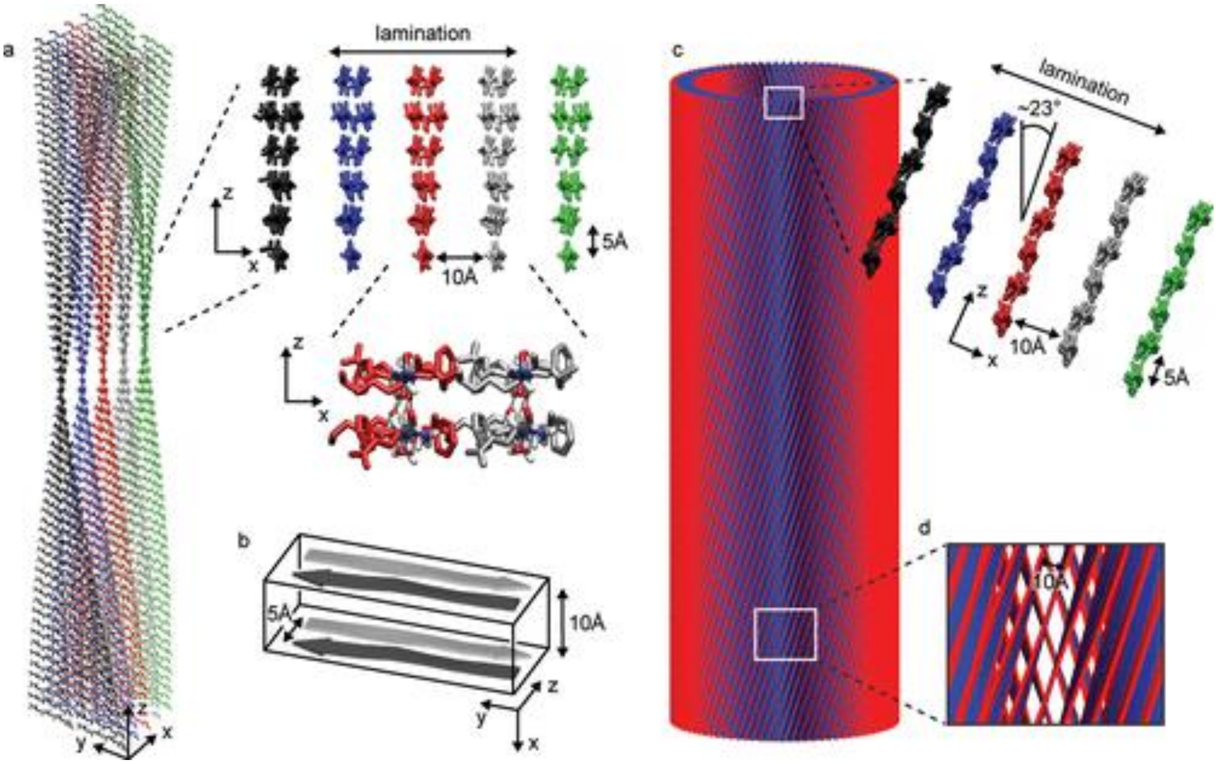


Figure 1-0-6: Morphological variants of Aβ16-22. (a) Atomic model of the laminated filament of Aβ16-22 observed at neutral pH. (b) Proposed unit cell describing the packing of peptides in the laminated filament of Aβ16-22. (c) Atomic model of the bilayer nanotube of Aβ16-22 observed at acidic pH. (d) Expanded representation of the tilted packing of β-strands within a bilayer nanotube of Aβ16-22. Reprinted with permission from Mehta et al. (2008), Journal of the American Chemical Society, 130(30), 9829–9835. Copyright 2021 American Chemical Society.

A similar oligopeptide-derived sequence from an Aβ, α-synuclein (α-syn) was able to assemble into helical ribbons. The eight-residue truncated sequence NH_2-^{37}VLYVGSK^{44}T-COOH revealed an assembly evolution from helical ribbons to close nanotubes over time. Despite having different sequences, both the α-Sβ1 and $A\beta_{16\text{-}22}$ peptides demonstrated similar self-association behavior. Their different amphiphilic patterns ie facially amphiphilic for $A\beta_{16\text{-}22}$ vs conventional amphiphilic for α-Sβ1 accounted for the difference in packing. Further studies of the same peptide showed that their amphiphilic character allows them to achieve long-range order while associated with straight chain fatty acids[51]. On the other hand, more complex peptide sequences have been used to promote nanotube self-assemblies. Peptide surfactant-based bolamphiphile architectures, for instance, have demonstrated abilities to form a cross-β nanotubes. They are designed with a hydrophobic core and polar residues at their terminal[20,22]. They were shown to form wide diameter thin wall nanotubes than can turn into ribbons upon switching the solvent from water to Ceric Ammonium Nitrate (CAN), leading to the weakening of the hydrogen bonding interactions. Another great example is the recently published Lanreotide acetate by Paternostre and coworkers (**Figure 1-0-7**). They assembled an 8-residue disulfide-linked cyclic peptide hormone. The nanotube formation was related to a hexagonal columnar mesophase[20,52]. The self-assembly formed uniform diameter tubes with negligible polymorphism. The cryo-EM structure of the nanotubes was recently solved at a 2.5 A resolution[53].[52]. It was found to have a right-handed 1-start helix with asymmetric units based on eight peptides arranged in a monolayer shell. The peptide backbone from that monolayer extends outward from the central axis of the assembly. The structure resolved represents the highest resolution achieved thus far for synthetic peptides

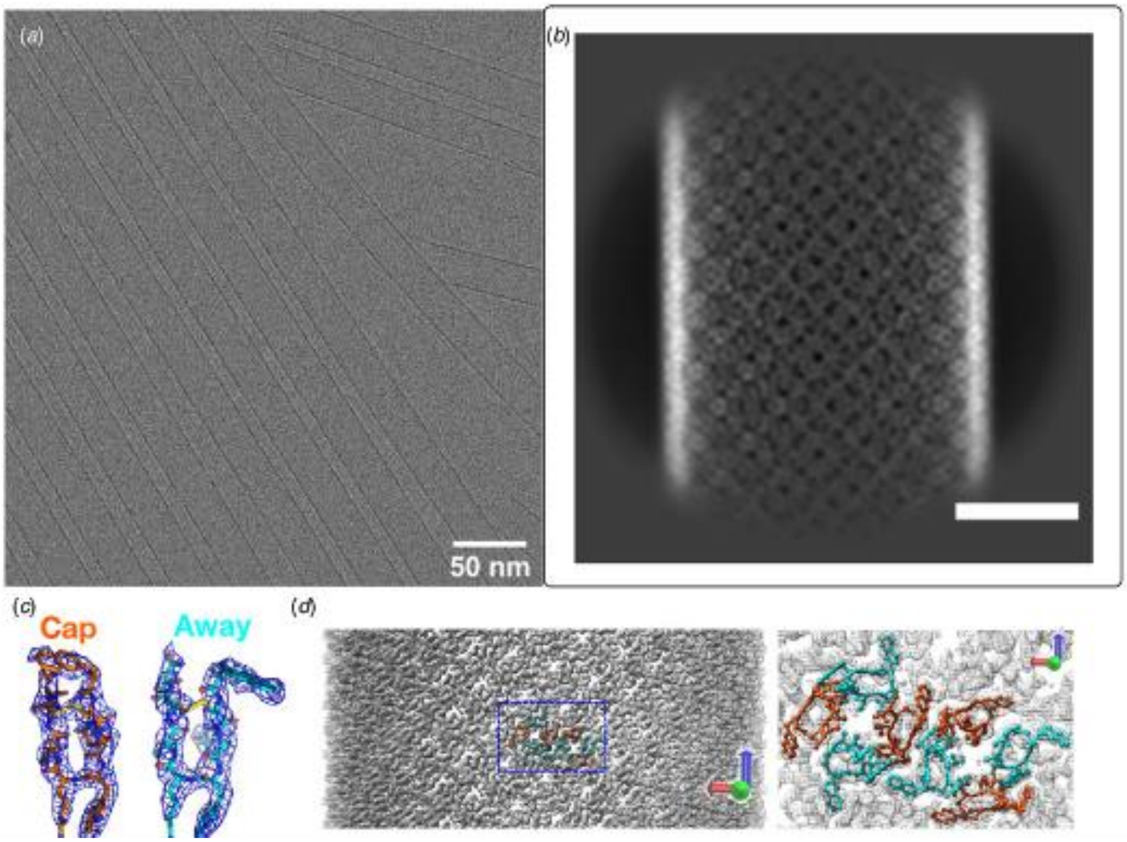


Figure 1-0-7: Cryo-EM structure of lanreotide nanotubes. (a) Representative cryo-EM image of peptide nanotubes derived from self-assembly of lanreotide acetate. (b) Representative 2D class average derived from cryo-EM analysis of lanreotide nanotubes. (c) Density maps for the two main conformations, capped versus away, of the lanreotide peptide within the nanotubes. (d) Density map for the lanreotide nanotube in which the eight peptides in an asymmetric unit are highlighted within the assembly. An expansion of the asymmetric unit is depicted on the right in which the capped (orange) and away (cyan) conformations are indicated.

nanotubes. Despite the fact that many more cross-β sheet base nanotubes have been investigated (macrocyclic β-peptides, synthetic surfactant-like peptides, and amphiphilic block copolypeptides among others), it is difficult to establish a set of rules to help with the controlled design and understanding of the mechanism of formation of these nanotubes assemblies since there is limited structural information available at near-atomic resolution despite these recent advances.

1.4- Aims and scope of the dissertation

Despite being widely used in electronics, biology, and medicine, the structure peptide self-assemblies have also proven difficult to reliably predict and rationally control. As a result, the general theme of this dissertation focuses on utilizing the development of electron cryo-microscopy (cryo-EM) with direct electron detection to understand the structure of peptides nanotubes from derived from self-assembly of α-helical and β-sheet motifs in order to ultimately establish the structural basis of a supramolecular peptide folding code.

Chapter 2 focuses on *de novo* predictive design of α-helical assemblies as it represents a significant challenge due to the conformational plasticity of quaternary structure in sequence-space and the lability of helical symmetry (*vide supra*). We have demonstrated that cryo-EM can be used to generate reliable atomic models for synthetic helical assemblies. We have employed the information from these studies to understand the structurally critical interactions that underlie filament formation and to predict the effect of mutations within the structure.

Chapter 3 explores the design space for synthetic β-sheet-based filaments derived from the self-assembly of the well-studied octapeptide KFE8 (FKFEFKFE). We focused on a series of modifications XKXEXKXE (X = Val, Ile, Phe), which had been previously studied using low-resolution analytical methods. We then make use of the KFE8 peptide to illustrate the challenge associated with de novo design of synthetic helical assemblies. In particular, the presence of structural polymorphism, the lability of helical symmetry, and the presence of rarely observed structural interfaces.

Chapter 4 expands on the structural information from chapter 2 to gain more insight into the designability of a family of α-helical nanotubes using computational design as it has proven effective to optimize interfacial interactions between protomers within known structures. We compare computational models to original high-resolution experimental structures and ultimately show that the computationally redesigned sequences from targeted structures can maintain the integrity of the inter-protomer interfaces.

Lastly, Chapter 5 details the molecular self-assembly mechanism of the previously characterized cross-β-sheet nanotube from chapter 3. We use complementary experimental techniques such as static light scattering and FT-IR to investigate the successive steps involved in the nanotube formation process and gain more insight into the supramolecular organization mechanism of the FKE8 assemblies.

1.5- References`

(1)Chiti, F.; Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat Chem Biol* **2009**, *5* (1), 15-22. DOI: 10.1038/nchembio.131 From NLM.

(2)Bowerman, C. J.; Nilsson, B. L. Self-assembly of amphipathic β-sheet peptides: insights and applications. *Biopolymers* **2012**, *98* (3), 169-184. DOI: 10.1002/bip.22058 From NLM.

(3)Zhang, S.; Lockshin, C.; Cook, R.; Rich, A. Unusually stable beta-sheet formation in an ionic self-complementary oligopeptide. *Biopolymers* **1994**, *34* (5), 663-672. DOI: 10.1002/bip.360340508 From NLM.

(4)Cormier, A. R.; Pang, X.; Zimmerman, M. I.; Zhou, H. X.; Paravastu, A. K. Molecular structure of RADA16-I designer self-assembling peptide nanofibers. *ACS Nano* **2013**, *7* (9), 7562-7572. DOI: 10.1021/nn401562f From NLM.

(5)Lee, M.; Wang, T.; Makhlynets, O. V.; Wu, Y.; Polizzi, N. F.; Wu, H.; Gosavi, P. M.; Stöhr, J.; Korendovych, I. V.; DeGrado, W. F.; et al. Zinc-binding structure of a catalytic amyloid from solid-state NMR. *Proc Natl Acad Sci U S A* **2017**, *114* (24), 6191-6196. DOI: 10.1073/pnas.1706179114 From NLM.

(6)Morris, R. J.; Bricogne, G. Sheldrick's 1.2 A rule and beyond. *Acta Crystallogr D Biol Crystallogr* **2003**, *59* (Pt 3), 615-617. DOI: 10.1107/s090744490300163x From NLM.

(7)Zhao, Y.; Wang, J.; Deng, L.; Zhou, P.; Wang, S.; Wang, Y.; Xu, H.; Lu, J. R. Tuning the self-assembly of short peptides via sequence variations. *Langmuir* **2013**, *29* (44), 13457-13464. DOI: 10.1021/la402441w From NLM.

(8)Hamley, I. W.; Burholt, S.; Hutchinson, J.; Castelletto, V.; da Silva, E. R.; Alves, W.; Gutfreund, P.; Porcar, L.; Dattani, R.; Hermida-Merino, D.; et al. Shear Alignment of Bola-

Amphiphilic Arginine-Coated Peptide Nanotubes. *Biomacromolecules* **2017**, *18* (1), 141-149. DOI: 10.1021/acs.biomac.6b01425 From NLM.

(9)Valéry, C.; Artzner, F.; Robert, B.; Gulick, T.; Keller, G.; Grabielle-Madelmont, C.; Torres, M. L.; Cherif-Cheikh, R.; Paternostre, M. Self-association process of a peptide in solution: from beta-sheet filaments to large embedded nanotubes. *Biophys J* **2004**, *86* (4), 2484-2501. DOI: 10.1016/s0006-3495(04)74304-0 From NLM.

(10)Valéry, C.; Paternostre, M.; Robert, B.; Gulik-Krzywicki, T.; Narayanan, T.; Dedieu, J. C.; Keller, G.; Torres, M. L.; Cherif-Cheikh, R.; Calvo, P.; et al. Biomimetic organization: Octapeptide self-assembly into nanotubes of viral capsid-like dimension. *Proc Natl Acad Sci U S A* **2003**, *100* (18), 10258-10262. DOI: 10.1073/pnas.1730609100 From NLM.

(11)Pieri, L.; Wang, F.; Arteni, A. A.; Vos, M.; Winter, J. M.; Le Du, M. H.; Artzner, F.; Gobeaux, F.; Legrand, P.; Boulard, Y.; et al. Atomic structure of Lanreotide nanotubes revealed by cryo-EM. *Proc Natl Acad Sci U S A* **2022**, *119* (4). DOI: 10.1073/pnas.2120346119 From NLM.

(12)Miller, J. G.; Hughes, S. A.; Modlin, C.; Conticello, V. P. Structures of synthetic helical filaments and tubes based on peptide and peptido-mimetic polymers. *Q Rev Biophys* **2022**, 1-103. DOI: 10.1017/s0033583522000014 From NLM.

(13)Acar, H.; Srivastava, S.; Chung, E. J.; Schnorenberg, M. R.; Barrett, J. C.; LaBelle, J. L.; Tirrell, M. Self-assembling peptide-based building blocks in medical applications. *Adv Drug Deliv Rev* **2017**, *110-111*, 65-79. DOI: 10.1016/j.addr.2016.08.006 From NLM Medline.

(14)Ekiz, M. S.; Cinar, G.; Khalily, M. A.; Guler, M. O. Self-assembled peptide nanostructures for functional materials. *Nanotechnology* **2016**, *27* (40), 402002. DOI: 10.1088/0957-4484/27/40/402002 From NLM Medline.

(15)Ramakers, B. E.; van Hest, J. C.; Lowik, D. W. Molecular tools for the construction of peptide-based materials. *Chem Soc Rev* **2014**, *43* (8), 2743-2756. DOI: 10.1039/c3cs60362h From NLM Medline.

(16)Nagy-Smith, K.; Moore, E.; Schneider, J.; Tycko, R. Molecular structure of monomorphic peptide fibrils within a kinetically trapped hydrogel network. *Proc Natl Acad Sci U S A* **2015**, *112* (32), 9816-9821. DOI: 10.1073/pnas.1509313112.

(17)Zhang, S. Q.; Huang, H.; Yang, J.; Kratochvil, H. T.; Lolicato, M.; Liu, Y.; Shu, X.; Liu, L.; DeGrado, W. F. Designed peptides that assemble into cross-alpha amyloid-like structures. *Nat Chem Biol* **2018**, *14* (9), 870-875. DOI: 10.1038/s41589-018-0105-5.

(18)Chen, K. H.; Corro, K. A.; Le, S. P.; Nowick, J. S. X-ray Crystallographic Structure of a Giant Double-Walled Peptide Nanotube Formed by a Macrocyclic beta-Sheet Containing Abeta16-22. *J Am Chem Soc* **2017**, *139* (24), 8102-8105. DOI: 10.1021/jacs.7b03890.

(19)Kajander, T.; Cortajarena, A. L.; Mochrie, S.; Regan, L. Structure and stability of designed TPR protein superhelices: unusual crystal packing and implications for natural TPR proteins. *Acta Crystallogr D Biol Crystallogr* **2007**, *63* (Pt 7), 800-811. DOI: 10.1107/S0907444907024353.

(20)Shen, H.; Fallas, J. A.; Lynch, E.; Sheffler, W.; Parry, B.; Jannetty, N.; Decarreau, J.; Wagenbach, M.; Vicente, J. J.; Chen, J.; et al. De novo design of self-assembling helical protein filaments. *Science* **2018**, *362* (6415), 705-709. DOI: 10.1126/science.aau3775.

(21)Hughes, S. A.; Wang, F.; Wang, S.; Kreutzberger, M. A. B.; Osinski, T.; Orlova, A.; Wall, J. S.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Ambidextrous helical nanotubes from self-assembly of designed helical hairpin motifs. *Proc Natl Acad Sci U S A* **2019**, *116* (29), 14456-14464. DOI: 10.1073/pnas.1903910116 From NLM Medline.

(22)Egelman, E. H.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.; Baker, D.; Conticello, V. P. Structural plasticity of helical nanotubes based on coiled-coil assemblies. *Structure* **2015**, *23* (2), 280-289. DOI: 10.1016/j.str.2014.12.008.

(23)Nagy-Smith, K.; Beltramo, P. J.; Moore, E.; Tycko, R.; Furst, E. M.; Schneider, J. P. Molecular, Local, and Network-Level Basis for the Enhanced Stiffness of Hydrogel Networks Formed from Coassembled Racemic Peptides: Predictions from Pauling and Corey. *ACS Cent Sci* **2017**, *3* (6), 586-597. DOI: 10.1021/acscentsci.7b00115.

(24)Fujii, T.; Cheung, M.; Blanco, A.; Kato, T.; Blocker, A. J.; Namba, K. Structure of a type III secretion needle at 7-Å resolution provides insights into its assembly and signaling mechanisms. *Proceedings of the National Academy of Sciences* **2012**, *109* (12), 4461-4466. DOI: 10.1073/pnas.1116126109.

(25)Loquet, A.; Sgourakis, N. G.; Gupta, R.; Giller, K.; Riedel, D.; Goosmann, C.; Griesinger, C.; Kolbe, M.; Baker, D.; Becker, S.; et al. Atomic model of the type III secretion system needle. *Nature* **2012**, *486* (7402), 276-279. DOI: 10.1038/nature11079.

(26)Cuccia, L. A.; Ruiz, E.; Lehn, J. M.; Homo, J. C.; Schmutz, M. Helical self-organization and hierarchical self-assembly of an oligoheterocyclic pyridine-pyridazine strand into extended supramolecular fibers. *Chemistry* **2002**, *8* (15), 3448-3457. DOI: 10.1002/1521-3765(20020802)8:15<3448::Aid-chem3448>3.0.Co;2-# From NLM.

(27)Cuccia, L. A.; Lehn, J. M.; Homo, J. C.; Schmutz, M. Encoded Helical Self-Organization and Self-Assembly into Helical Fibers of an Oligoheterocyclic Pyridine - Pyridazine Molecular Strand. *Angew Chem Int Ed Engl* **2000**, *39* (1), 233-237. DOI: 10.1002/(sici)1521-3773(20000103)39:1<233::aid-anie233>3.0.co;2-r From NLM.

(28)Lombardo, C. M.; Collie, G. W.; Pulka-Ziach, K.; Rosu, F.; Gabelica, V.; Mackereth, C. D.; Guichard, G. Anatomy of an Oligourea Six-Helix Bundle. *J Am Chem Soc* **2016**, *138* (33), 10522-10530. DOI: 10.1021/jacs.6b05063 From NLM.

(29)Collie, G. W.; Bailly, R.; Pulka-Ziach, K.; Lombardo, C. M.; Mauran, L.; Taib-Maamar, N.; Dessolin, J.; Mackereth, C. D.; Guichard, G. Molecular Recognition within the Cavity of a Foldamer Helix Bundle: Encapsulation of Primary Alcohols in Aqueous Conditions. *J Am Chem Soc* **2017**, *139* (17), 6128-6137. DOI: 10.1021/jacs.7b00181 From NLM.

(30)Collie, G. W.; Pulka-Ziach, K.; Lombardo, C. M.; Fremaux, J.; Rosu, F.; Decossas, M.; Mauran, L.; Lambert, O.; Gabelica, V.; Mackereth, C. D.; et al. Shaping quaternary assemblies of water-soluble non-peptide helical foldamers by sequence manipulation. *Nat Chem* **2015**, *7* (11), 871-878. DOI: 10.1038/nchem.2353 From NLM.

(31)Stanley, W. M. Chemical studies on the virus tobacco mosaic. *Phytopathology* **1935**, *25*, 475-492.

(32)Zhang, J.; Zheng, F.; Grigoryan, G. Design and designability of protein-based assemblies. *Current Opinion in Structural Biology* **2014**, *27*, 79-86. DOI: https://doi.org/10.1016/j.sbi.2014.05.009.

(33)Adamcik, J.; Mezzenga, R. Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape. *Angew Chem Int Ed Engl* **2018**, *57* (28), 8370-8382. DOI: 10.1002/anie.201713416.

(34)Galkin, V. E.; Orlova, A.; Schroder, G. F.; Egelman, E. H. Structural polymorphism in F-actin. *Nat Struct Mol Biol* **2010**, *17* (11), 1318-1323. DOI: 10.1038/nsmb.1930.

(35)Galkin, V. E.; Kolappan, S.; Ng, D.; Zong, Z.; Li, J.; Yu, X.; Egelman, E. H.; Craig, L. The structure of the CS1 pilus of enterotoxigenic Escherichia coli reveals structural polymorphism. *J Bacteriol* **2013**, *195* (7), 1360-1370. DOI: 10.1128/JB.01989-12.

(36)Egelman, E. H. A tale of two polymers: new insights into helical filaments. *Nat Rev Mol Cell Biol* **2003**, *4* (8), 621-630. DOI: 10.1038/nrm1176.

(37)Lu, A.; Li, Y.; Yin, Q.; Ruan, J.; Yu, X.; Egelman, E.; Wu, H. Plasticity in PYD assembly revealed by cryo-EM structure of the PYD filament of AIM2. *Cell Discov* **2015**, *1*. DOI: 10.1038/celldisc.2015.13 From NLM.

(38)Guenther, E. L.; Ge, P.; Trinh, H.; Sawaya, M. R.; Cascio, D.; Boyer, D. R.; Gonen, T.; Zhou, Z. H.; Eisenberg, D. S. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nat Struct Mol Biol* **2018**, *25* (4), 311-319. DOI: 10.1038/s41594-018-0045-5.

(39)Close, W.; Neumann, M.; Schmidt, A.; Hora, M.; Annamalai, K.; Schmidt, M.; Reif, B.; Schmidt, V.; Grigorieff, N.; Fandrich, M. Physical basis of amyloid fibril polymorphism. *Nat Commun* **2018**, *9* (1), 699. DOI: 10.1038/s41467-018-03164-5.

(40)Zhang, W.; Falcon, B.; Murzin, A. G.; Fan, J.; Crowther, R. A.; Goedert, M.; Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **2019**, *8*. DOI: 10.7554/eLife.43584 From NLM.

(41)Wu, Y.; Norberg, P. K.; Reap, E. A.; Congdon, K. L.; Fries, C. N.; Kelly, S. H.; Sampson, J. H.; Conticello, V. P.; Collier, J. H. A Supramolecular Vaccine Platform Based on alpha-Helical Peptide Nanofibers. *ACS Biomater Sci Eng* **2017**, *3* (12), 3128-3132. DOI: 10.1021/acsbiomaterials.7b00561.

(42)Wu, D.; Sinha, N.; Lee, J.; Sutherland, B. P.; Halaszynski, N. I.; Tian, Y.; Caplan, J.; Zhang, H. V.; Saven, J. G.; Kloxin, C. J.; et al. Polymers with controlled assembly and rigidity made with click-functional peptide bundles. *Nature* **2019**, *574* (7780), 658-662. DOI: 10.1038/s41586-019-1683-4.

(43)Fries, C. N.; Wu, Y.; Kelly, S. H.; Wolf, M.; Votaw, N. L.; Zauscher, S.; Collier, J. H. Controlled Lengthwise Assembly of Helical Peptide Nanofibers to Modulate CD8+ T-Cell Responses. *Advanced Materials n/a* (n/a), 2003310. DOI: 10.1002/adma.202003310.

(44)DiMaio, F.; Song, Y.; Li, X.; Brunner, M. J.; Xu, C.; Conticello, V.; Egelman, E.; Marlovits, T.; Cheng, Y.; Baker, D. Atomic-accuracy models from 4.5-A cryo-electron microscopy data with density-guided iterative local refinement. *Nat Methods* **2015**, *12* (4), 361-365. DOI: 10.1038/nmeth.3286.

(45)Ramakrishnan, V. The Ribosome Emerges from a Black Box. *Cell* **2014**, *159* (5), 979-984. DOI: 10.1016/j.cell.2014.10.052 (acccessed 2020/09/21).

(46)Burkoth, T. S.; Benzinger, T. L. S.; Urban, V.; Morgan, D. M.; Gregory, D. M.; Thiyagarajan, P.; Botto, R. E.; Meredith, S. C.; Lynn, D. G. Structure of the β-Amyloid(10-35) Fibril. *Journal of the American Chemical Society* **2000**, *122* (33), 7883-7889. DOI: 10.1021/ja000645z.

(47)Benzinger, T. L.; Gregory, D. M.; Burkoth, T. S.; Miller-Auer, H.; Lynn, D. G.; Botto, R. E.; Meredith, S. C. Propagating structure of Alzheimer's beta-amyloid(10-35) is parallel beta-sheet with residues in exact register. *Proc Natl Acad Sci U S A* **1998**, *95* (23), 13407-13412. DOI: 10.1073/pnas.95.23.13407 From NLM.

(48)Morgan, D. M.; Dong, J.; Jacob, J.; Lu, K.; Apkarian, R. P.; Thiyagarajan, P.; Lynn, D. G. Metal switch for amyloid formation: insight into the structure of the nucleus. *J Am Chem Soc* **2002**, *124* (43), 12644-12645. DOI: 10.1021/ja0273086 From NLM.

(49)Dong, J.; Canfield, J. M.; Mehta, A. K.; Shokes, J. E.; Tian, B.; Childers, W. S.; Simmons, J. A.; Mao, Z.; Scott, R. A.; Warncke, K.; et al. Engineering metal ion coordination to regulate amyloid fibril assembly and toxicity. *Proc Natl Acad Sci U S A* **2007**, *104* (33), 13313-13318. DOI: 10.1073/pnas.0702669104 From NLM.

(50)Feng, Z.; Wang, H.; Wang, F.; Oh, Y.; Berciu, C.; Cui, Q.; Egelman, E. H.; Xu, B. Artificial Intracellular Filaments. *Cell Reports Physical Science* **2020**, *1* (7), 100085. DOI: https://doi.org/10.1016/j.xcrp.2020.100085.

(51)Ni, R.; Childers, W. S.; Hardcastle, K. I.; Mehta, A. K.; Lynn, D. G. Remodeling cross-β nanotube surfaces with peptide/lipid chimeras. *Angew Chem Int Ed Engl* **2012**, *51* (27), 6635-6638. DOI: 10.1002/anie.201201173 From NLM.

(52)Tayeb-Fligelman, E.; Tabachnikov, O.; Moshe, A.; Goldshmidt-Tran, O.; Sawaya, M. R.; Coquelle, N.; Colletier, J. P.; Landau, M. The cytotoxic Staphylococcus aureus PSMalpha3 reveals a cross-alpha amyloid-like fibril. *Science* **2017**, *355* (6327), 831-833. DOI: 10.1126/science.aaf4901.

(53)Walshaw, J.; Woolfson, D. N. Open-and-shut cases in coiled-coil assembly: alpha-sheets and alpha-cylinders. *Protein Sci* **2001**, *10* (3), 668-673. DOI: 10.1110/ps.36901 From NLM Medline.

Chapter 2: Structural analysis of a family of cross-helical nanotubes

This chapter has been adapted with permission from Wang, F., Gnewou, O., Modlin, C. *et al.* Structural analysis of cross α-helical nanotubes provides insight into the designability of filamentous peptide nanomaterials. *Nat Commun* 12**,** 407 (2021). https://doi.org/10.1038/s41467-020-20689-w **Copyright © 2022 Nature Communication (https://rdcu.be/cSHsg)**

2.1- Introduction

Peptide-based filamentous assemblies have been employed with great success over the past several decades for the fabrication of structurally ordered materials within the nano-scale size regime for applications in biomedicine and nanotechnology.[1-4]The molecular design of these materials has relied thus far on relatively simple sequence-structure correlations derived from structural informatics analysis of native protein folds. The mode of assembly involves self-association of structural subunits (i.e., protomers) into long non-covalent polymers that display helical symmetry. While de novo design has been employed frequently to identify suitable candidate peptides for construction of synthetic helical assemblies, the number of structures that have been characterized at near-atomic resolution is relatively limited.[5-10]As recent high-resolution structural analyses have demonstrated for both native and synthetic protein filaments, atomic models for helical assemblies based on low-resolution data are incomplete and, often in error.[11-14] In the infrequent cases in which high resolution analysis has been performed on designed peptide filaments, the observed structures can differ significantly from the models that were employed as the basis for the design. These results suggest that our current knowledge of the design principles that underlie the formation of helical peptide and protein assemblies remains limited in scope, even for the simple structural motifs that have been employed as substrates in these studies. In addition, it raises the question of whether quaternary structure, i.e, interactions at the interfaces between protomers, is robust in sequence space, that is, designable,[15]and can be employed for reliable and predictable design of synthetic peptide filaments.[16-24]

Given the difficulties involved in correctly predicting quaternary structure, the de novo design of synthetic peptide assemblies requires validation of the predicted model through structural determination at near-atomic level resolution.

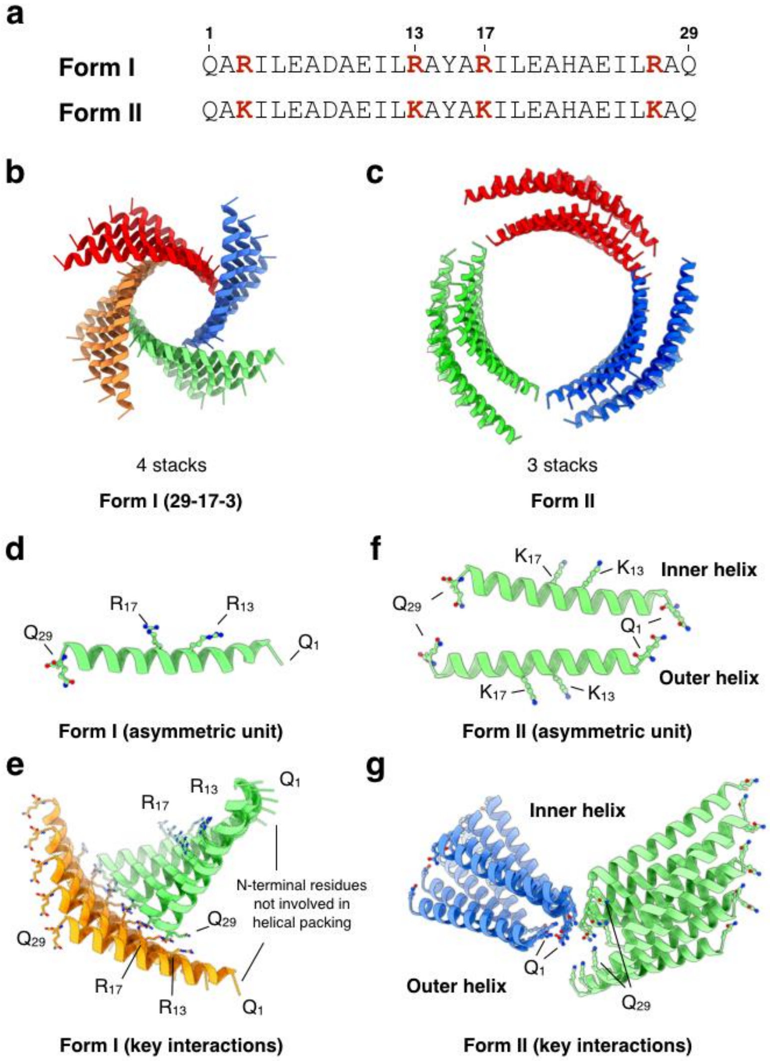


Figure 2-1: Structures and critical interactions in the Form peptide filaments. **a,** The 29-residue sequences of Form I and Form II peptides. **b-c,** Cryo-EM structure of Form I (b) and Form II (c) filaments. **d,** The building block (asymmetric unit) of Form I peptide with key residues shown in sticks **e,** Essential interactions between adjacent helix stacks in Form I filaments that maintain the helical packing**. f,** The building block (asymmetric unit) of the re-examined Form II peptide with key residues shown in sticks **g,** Essential interactions between adjacent helix stacks in Form II filaments that maintain the helical packing.

This structural information is essential for understanding the inter-subunit contacts that stabilize the interfaces within filamentous peptide assemblies, especially as these interactions are critically important for functional properties such as mechanical response and bioactivity.[24-26]Structural polymorphism is frequently observed for designed helical peptide assemblies, as well as for biologically derived protein filaments that have been assembled in vitro.[9, 21, 26-29]Diverse populations of structurally distinct filaments can arise from the same peptide or protein sequence. Selection for monomorphic variants of defined and predictable structure represents a significant challenge to de novo design of peptide-based materials. Slight variations in preparative conditions can result in the formation of distinctly different structural variants in a manner that has yet to be understood at the level of atomic interactions.[30-32]Therefore, structural information that results from near-atomic resolution analyses of helical assemblies, yielding reliable atomic models, is essential for interrogating the limits of our knowledge of peptide design as well as for reverse engineering of structure to generate assemblies reliably and predictably.

We recently described the structural analysis of helical filaments derived from two related synthetic coiled-coil peptides, Form I and Form II (**Figure 2-2-1**).[33] Atomic models were generated using Rosetta modeling of the peptides into the density map derived from electron cryomicroscopy (cryo-EM) with direct electron detection(**Figure 2-1**)[8]. Although cryo-EM has come to dominate structural biology within the past several years and has emerged as the main technique for determining the atomic structure of macromolecular complexes,[11, 20, 34]applications of this method in chemistry and materials science have been relatively sparse.[35-39]However, the resolution revolution in cryo-EM has enabled direct structural determination of designed helical assemblies at near-atomic resolution that were previously inaccessible.

While the sequences of Form I and Form II peptides differed solely in the substitution of arginine residues in the former sequence for lysine residues in the latter sequence (**Figure 2-1-a**), the corresponding atomic models for the filaments displayed significant differences in helical symmetry and the nature of the cohesive interactions between protomers. These differences were not anticipated in that the corresponding sequence substitutions occurred at positions that were not expected to participate directly in interfacial interactions (**Figure 2-2-1)**[40, 41] Moreover, site-directed mutagenesis of one or two residues within the respective sequences of Form I and Form II resulted in interconversion between the two alternative helical structures (vide infra). While the structural transition could be rationalized on the basis of the differences between the atomic models of the respective filaments, neither structure was predicted a priori, nor could the structures be considered as robust in sequence space, due to the significant impact of limited mutagenesis. Indeed, the dramatic change in quaternary structure due to the semi-conservative mutagenesis of one or two amino acids is suggestive of nearly chaotic behavior.

Here we report seven high-resolution cryo-EM structures of filamentous peptide nanotubes based on cross-α supramolecular architectures that provide insight into the designability of helical peptide filaments. In addition, emergent structural behavior is observed among the corresponding helical filaments. A structural interaction is identified based on an appropriately placed pair of arginine residues (RxxxR) within the corresponding peptide sequences. This local interaction motif, designated as an arginine clasp, is robust within the context of these cross-α assemblies and can guide the coalescence of helical protofilaments into structurally defined cross-α nanotubes. Despite this observation, the arginine clasp motif does not appear to have been evolutionarily sampled as a mechanism to mediate similar helix-helix interactions in naturally occurring protein structures and, therefore, cannot be considered as natively designable.

2.2- Result and discussion

2.2.1- Structural differences between Form I and Form II Filaments β, α

The respective assemblies of Form I and Form II epitomized an unusual packing motif, recently designated as a cross-α structure[42] (alternatively, an α-sheet structure[43]). In contrast to most structurally characterized filaments based on synthetic α-helical coiled-coils,[35, 44-51]the chain axes of the helical protomers are oriented in a plane nearly perpendicular to the long axis of the assembly. The initial reconstruction of Form II resulted in an atomic model of relatively low (circa 7 Å) resolution, in contrast to the higher (3.6-Å) resolution of the corresponding model for Form I (PDB ID: 3J89, [https://www.rcsb.org/structure/3j89]). In order to better compare the two models, cryo-EM data were collected on a fresh preparation of Form II filaments (**Figure 2-1-c**). Helical reconstruction afforded an atomic model at 4.2-Å resolution with a different helical symmetry (i.e., a rise of 1.93 Å and a rotation of 124.4°) than that proposed for the original model.[52] The higher resolution model of Form II is consistent with a parallel orientation of the inner-outer helix pair, which was suggested but could not be unambiguously determined in the previous analysis.[52]

A structural comparison of the two assemblies indicated that helical protomers were arranged in four (Form I) or three (Form II) cross-helical stacks (i.e., protofilaments) that self-associated laterally to form single-walled and double-walled nanotubes, respectively (**Figure 2-1**). The axial interactions along the respective stacking interfaces were conserved between the different filaments. These interfaces were stabilized through a series of heterotypic knobs-into-holes (KIH) interactions between hydrophobic residues at the a/d and c/f faces of axially adjacent helices within a protofilament (**Figure 2-2-1**).[52, 53] The main distinguishing feature between the Form I and Form II assemblies was the mode of lateral association between the respective cross

α-helical stacks. The Form I structure displayed a unique interaction involving a pair of arginine residues, R13 and R17, that formed a network of hydrogen bonds and electrostatic interactions with the C-terminal residues of helices in an adjacent protofilament (**Figure 2-1-e**). Mutagenesis of either of these residues to lysine resulted in conversion of the assembly to a Form II-like structure.[54] Conversely, a K13R, K17R double mutant of the Form II sequence resulted in conversion of the resultant assembly to a Form I-like structure, which suggested that the presence of both arginine residues was necessary for formation of the Form I structure. High-resolution cryo-EM structural analysis led to the identification of the critical interactions that were responsible for the structural differences between the Form I and Form II filaments, as well as the ability to directly interrogate the importance of these interactions through site-directed mutagenesis.

These results suggested that the RxxxR structural unit, in which the two arginine residues, R13 and R17, are arranged as i,i+4 within the Form I sequence, defined a side-to-end helical interaction motif within the assembly.[55]While interactions involving individual RxxxR motifs may be relatively weak, many such interactions occurred at the lateral interfaces between protofilaments along the contour length of the Form I structure. Consequently, the influence of these localized interactions on the stability of the helical assembly was magnified significantly. Numerous instances of the stabilizing effect of structural interactions between appropriately substituted i,i+4 residues have been described within the sequences of helical peptides.[9, 35, 56-62]In addition, specific interaction motifs such as the left-handed leucine zipper, i.e., LxxxLxx,[55, 63-65]and the right-handed glycine zipper,[30, 65-67]i.e., GxxxG, have been observed to promote and stabilize lateral (side-to-side) interactions in coiled-coils and transmembrane helical bundles, respectively.[68]The crossing angle, Ω, between laterally interacting pairs of helices in the Form I structure was circa 94°. In contrast, the leucine and glycine zipper motifs define side-to-side association between helices and display more acute crossing angles (Ω = -20° and +40°, respectively).[68-70] PISA[71] analysis indicated that the lateral interface between interacting helices in the Form I filament buried approximately 275 Å^2/peptide. For comparison, the coiled-coil

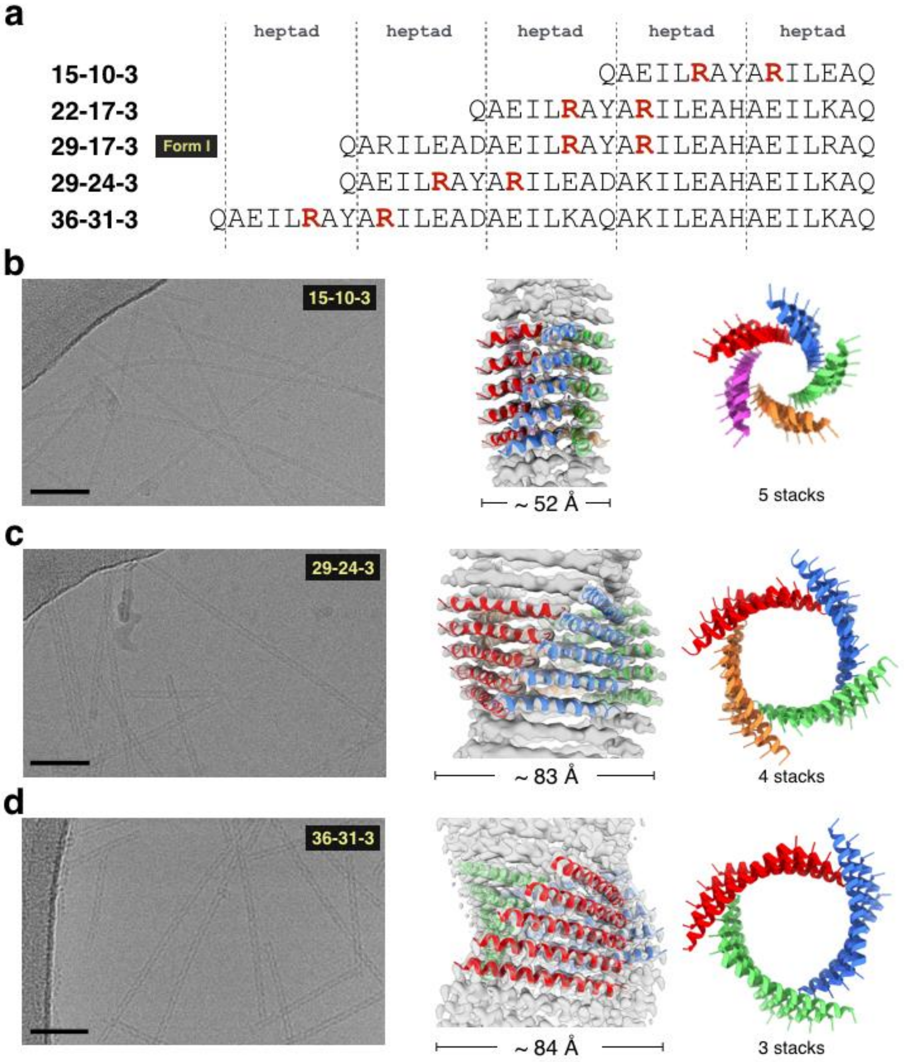


Figure 2-2: **Cryo-EM of arginine clasp peptide filaments based on Form I.**
a, Sequences of the Form I peptide and four designed peptides derived from Form I. **b-d**, Cryo-EM of peptides 15-10-3, 29-24-3, and 36-31-3. Representative raw micrographs are shown on the left (scale bar 500 Å) out of a total number of 84, 279, and 314 images recorded, respectively. The cryo-EM reconstructions are shown in the middle, with models fitted into the maps. The top views of the atomic models are shown on the right.

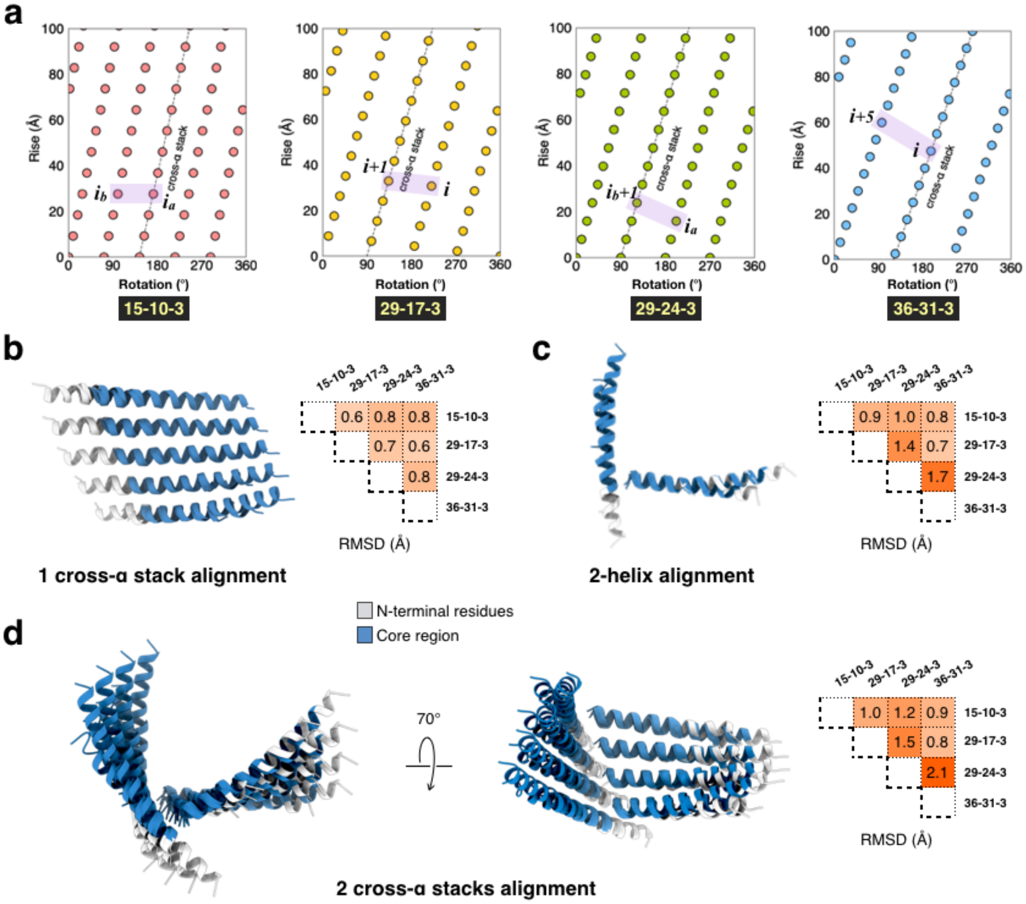
interaction along the axial direction of the Form I protofilaments defined an interface of approximately 700 Å^2 of surface area/peptide with a helix crossing angle of circa –11.5°. We designated this RxxxR interaction motif as an arginine clasp, since it mediated a distinctive local interaction between appropriately oriented protomers on structurally adjacent protofilaments.

2.2.2- Dependence of filaments structure and peptide length

. In order to test the utility of the arginine clasp in mediating interactions within a defined structural context, we designed a series of peptides derived from the Form I sequence (**Figure 2-2**). In the latter structure, the N-terminal heptad sequence, upstream of the RxxxR motif, was unstructured and did not contribute to the helix-helix interactions. We reasoned that the major contribution to the structural integrity of the helical filament was the region in Form I sequence space from the RxxxR motif to the C-terminus of the peptide. The corresponding peptide segment contained the critical structural elements responsible for the lateral and axial interactions. Therefore, the designed series of Form I-like peptides encompassed minimal peptide sequences in which the structurally determinative distance between the RxxxR motif and the C-terminus was systematically varied over the range from 2 to 5 heptad repeats (**Figure 2-2**). In each case, the most N-terminal heptad contained the RxxxR motif and, in order to prevent competition, lysine residues occupied all of the other basic amino acid sites at solvent-contacting b- and e-positions within the respective heptad sequences. Lysines contribute to stabilizing electrostatic interactions with glutamic acid residues that occur between peptides within a cross-α stack, but minimally to the cohesive interactions between stacks.[72]The nomenclature of the peptides was changed to one that reflected the structural elements that were considered to be critical to the design. All peptides were designated as (x-y-z), in which x refers to peptide length, y refers to the number of residues from the first arginine to the C-terminus, and z to the number of residues between the two arginine

(**Figure 2-2**). The original Form I sequence corresponded to peptide 29-17-3 under this system of nomenclature.

The designed peptide sequences were based on the assumption that the helical symmetry of the respective filaments and, consequently, the axial and lateral interfaces of the Form I filament, would be conserved in the resultant assemblies. In this scenario, the RxxxR motifs within a given protofilament would interact selectively with the C-terminal residues within adjacent helices oriented (i+1) along a left-handed 1-start helix oriented similarly to that of the Form I filament



	15-10-3	29-17-3	29-24-3	36-31-3
15-10-3		0.6	0.8	0.8
29-17-3			0.7	0.6
29-24-3				0.8
36-31-3				

	15-10-3	29-17-3	29-24-3	36-31-3
15-10-3		0.9	1.0	0.8
29-17-3			1.4	0.7
29-24-3				1.7
36-31-3				

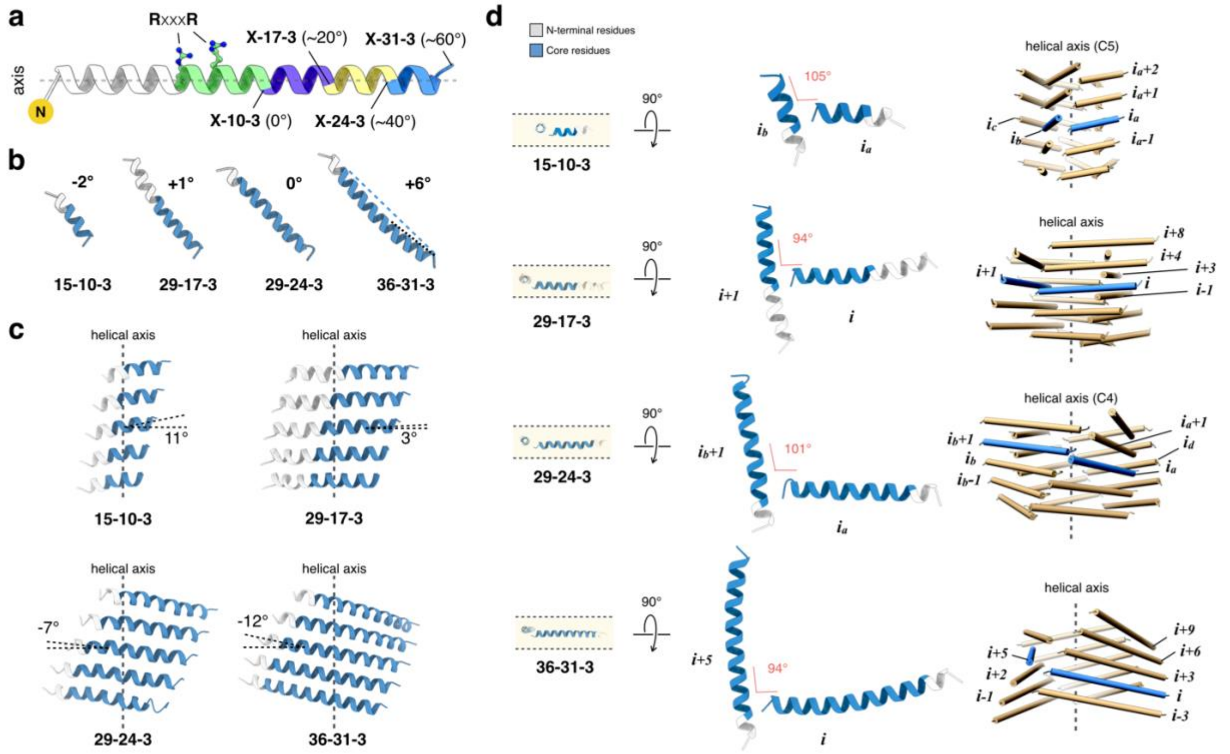
	15-10-3	29-17-3	29-24-3	36-31-3
15-10-3		1.0	1.2	0.9
29-17-3			1.5	0.8
29-24-3				2.1
36-31-3				

Figure 2-3: Structural conservations of the Form I based arginine clasp filaments
a, Helical nets for the peptide 15-10-3, 29-17-3 (Form I), 29-24-3, and 36-31-3 filaments. The helical nets show the unrolled surface lattice viewed from the outside of the filament. One of the right-handed n-start helices–associated with a cross-α protofilament within the respective assembly–is indicated with a straight line in the corresponding helical net diagram. Adjacent helices shown in a, with the conserved interactions are highlighted in a purple box. **b,** Structural alignment of a single cross-α stack containing five helices (left) and their Cα RMSD (right). **c,** Structural alignment of two adjacent helices (left) and their Cα RMSD (right). d, Structural alignment of two cross-α stacks (left) and their Cα RMSD (right).

(**Figure 2-3-a**). If a given peptide self-associated into a Form I-like structure, the diameter of the resultant nanotubes would depend on the length of the peptide segment from the first arginine to the C-terminus (i.e., the y-value). The Form I assembly displayed a nearly square cross-section, in which the nanotube diameter was determined by the seventeen amino acid length of the interacting peptide segments within a laterally adjacent pair of protofilaments within the assembly (**Figure 2-1**). If the Form I-like structure were retained, then the width of the assembly should be predictable from the position of the arginine clasp motif and the length of the peptide, as these parameters determine the cross-sectional dimensions of the corresponding assemblies. If the lengths of the interacting peptide segments were shorter than for Form I, a narrower nanotube would result and, if longer, the corresponding nanotube should display a wider diameter.

All four peptides in the (x-y-3) series adopted an α-helical conformation and formed thin filaments in aqueous buffers (**Figure 2-2-4** and **Figure 2-2-5)** Three filaments, derived from peptides 15-10-3, 29-24-3, and 36-31-3, were selected for further analysis using cryo-EM imaging with direct electron detection. The filament derived from 22-17-3 was not analyzed further, as it was expected that the structure would be similar to the previously described Form I/29-17-3, since it corresponded to a truncation of the unstructured N-terminal heptad. The iterative helical real space reconstruction (IHRSR) algorithm[40, 41, 73, 74]was employed to generate 3D reconstructions at final resolutions (Supplementary Table 2-1) of 4.2-Å, 4.1-Å, and 4.0-Å for the 15-10-3, the 29-24-3, and the 36-31-3 filaments, respectively (**Figure 2-2**-b). Surprisingly, the number of protofilaments (i.e., cross-α helical stacks) within the nanotubes was found to have an inverse relationship to peptide length. Five protofilaments interacted to form the 15-10-3 filament, four interacted in the 29-24-3 filament, and three interacted in the 36-31-3 filament. The Form I/29-17-3 structure fits the observed trend, in that the 17 amino acid length of the structurally determinative segment lay between the 10 amino acid and 24 amino acid segment length of 15-10-3 and 29-24-3, respectively.

High-resolution structural data enabled a detailed comparison between the arrangement of peptides within the respective atomic models (**Figure 2-3**). In each case, the α-helical protomers within a protofilament stacked along a right-handed n-start helix (**Figure 2-3**-a), in which the value of n corresponded to the number of protofilaments within a nanotube. For example, each 5-start helix of 15-10-3 passed through every fifth protomer and therefore corresponded to the axial stacking direction in a protofilament (**Figure 2-2**-b and **Figure 2-3**-a). Similarly, the 4-start helices



<u>Figure 2-4:</u> Three degrees of freedoms to maintain the conserved arginine clasp contacts. **a,** The building blocks of four types of peptide filaments. The N-terminal residues not involved in filament packing are colored white. The RxxxR motif is highlighted. The core region of X-10-3 peptide is colored in green. The core region of X-17-3, X-24-3, and X-31-3 have one (magenta), two (yellow) or three (blue) additional heptads, respectively. **b,** The helix tilt angle of the core region in four types of peptide filaments. **c,** The cross-α stack within the respective protofilaments, indicating tilt angle from the plane normal to the helical axis **d,** Two adjacent helices that make the key contacts in these peptide filaments. The adjacent helices are oriented to be in the same plane (left), and the angle between the two helices are indicated (middle). The relative positions of these two adjacent helices (blue) in the filaments are shown (right) of 29-17-3 and 29-24-3 and the 3-start helices of 36-31-3, passing through every fourth subunit and every third subunit, respectively, were associated with the axial stacking direction of the protofilaments within the corresponding assemblies. The arrangement of protofilaments within the respective assemblies can be clearly discerned in the corresponding helical net diagrams (**Figure 2-3**-a). The right-handed curvature of the protofilament derived from a two-residue progressive displacement between successive helical protomers along the rise of the n-start helices. This displacement was a consequence of a heterotypic coiled-coil interaction along the stacking interface between the two offset hydrophobic faces (a/d and c/f) of the protomers, as predicted by Walshaw and Woolfson for type III coiled-coil structures (**Figure 2-2-1**).[75]Structural overlays of helices within a protofilament (**Figure 2-3-b**) showed reasonably good alignment (root mean square standard deviation (RMSD) < 1 Å), which presumably resulted from the constraints of KIH packing at the coiled-coil interfaces despite differences that were observed at the lateral packing interfaces (**Figure 2-3-c** and **Figure 2-3-Figure 2-3-d**).

Structural comparison of the filaments suggested that three mechanisms could be distinguished through which the packing of peptides adjusted itself in order to accommodate the increase in peptide length (**Figure 2-4**). First, the tilt angle of the core peptide with respect to the plane normal to the filament axis became progressively more negative with increasing peptide length (**Figure 2-4**-c). Despite the differences in cross-sectional geometry for the respective filaments, the helix crossing angles were remarkably similar and lay within the range from 95° to 105° (**Figure 2-4**-d). We postulate that the arginine clasp interaction required that the crossing angle was within this range in order to effectively interact with the C-terminus of an adjacent helix. Adjustment of the tilt angle of helical protomers provided a mechanism to satisfy this geometric constraint. Second, for the longer peptide sequences, the lateral interaction between protofilaments could not be maintained between nearest-neighbor helices (**Figure 2-3**-a and **Figure 2-3**-d). For the 29-24-3 peptide, the interaction occurred between the C-terminus of peptide (i_a) and the RxxxR motif of peptide (i_b+1), in contrast to the C5-symmetric structure of 15-10-3 that involved helices i_a and i_b. Similarly, the filament structure of 36-31-3 revealed that the C-terminus of peptide (i) interacted with the axially offset RxxxR motif of peptide (i+5) within the 1-start (C1) symmetry of the assembly. The length and flexibility of the arginine side chains enabled the conservation of the RxxxR lateral interaction between protofilaments despite differences in peptide length. Finally, while the local helix crossing angle associated with the arginine clasp interaction was maintained at 94°, the distal segments of the 36-31-3 helices bend at a more acute crossing angle of 80° in order to maintain the overall helical symmetry of the filament. These observations suggested that peptide length determined the helical symmetry of the respective cross-α nanotubes through selection of number and relative orientation of protofilaments that could maintain the structurally critical cohesive interactions at the axial and lateral interfaces (**Figure 2-2**).

2.2.3- Effect of arginine position

The two arginine residues in the RxxxR clasp motif were positioned on the same side of the helix in an i,i+4 orientation. This placement enabled lateral association between protofilaments in the structures of the cross-α nanotubes described above. However, a structural arrangement in which two arginines are placed in an i,i+3 orientation, that is, as an RxxR motif, could potentially promote a similar interaction. The two arginines in the RxxR motif remain on the same face of the helical protomer, but would be arranged in a different orientation (angular displacement = -60° and axial displacement = 4.5 Å) in comparison to the RxxxR motif (angular displacement = +40° and axial displacement = 6.0 Å). Previous research demonstrated that appropriately selected residues in an i,i+3 orientation could mediate stabilizing intra- and inter-helix interactions.[71]

The Form I sequence was modified to test the effect of arginine placement on the structure of the corresponding peptide assemblies (**Figure 2-5**). Arginines were placed in an i,i+3 (RxxR) arrangement within a heptad repeat located proximal to the N-terminus. The sequence at the axial interface was conserved in order to maintain the cross-α stacking interactions within a protofilament. The resultant peptide was designated as 29-20-2 using the same (x-y-z)

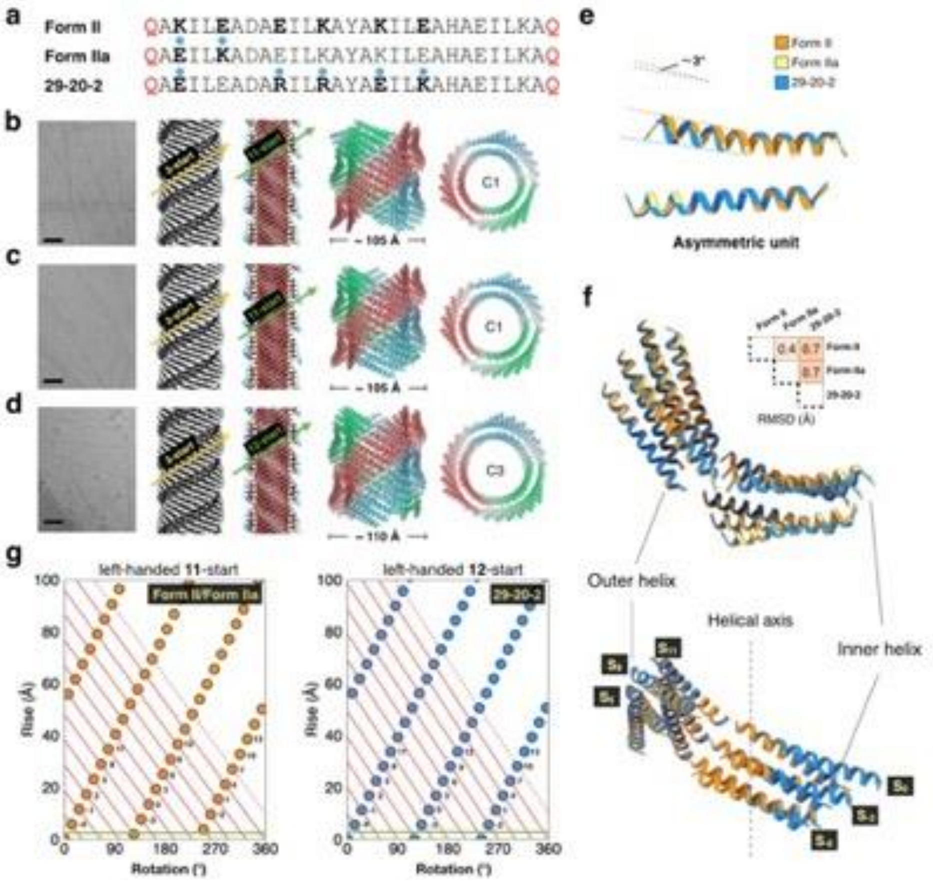


Figure 2-5: Structural Comparison of the Form II-like peptides.
a, Sequences of Form II, Form IIa and 29-20-2 peptides. Residues changed from original the Form II sequences are highlighted with blue dots. **b-d,** Cryo-EM of peptides Form II, Form IIa, and 29-20-2. Representative raw micrographs are shown on the left (scale bar 500 Å) out of a total number of 171, 233, and 347 images recorded, respectively. The cryo-EM reconstructions are shown in the middle, one view from outside highlighting right-handed 3-start helices, and another view from the inside of the nanotube lumen highlighting the left-handed 11-start or 12-start helices. The side and top views of the atomic models are shown on the right. **e,** Structural alignment of the asymmetric unit (helix-dimer) of the three peptide filaments. **f,** Structural alignment of six asymmetric units of the three peptide filaments. The alignment RMSD is shown. **g,** Helical nets for the Form II and 29-20-2 filaments. The helical nets show the unrolled surface lattice viewed from the outside of the filament. The right-handed 3-start helices can be clearly discerned for each assembly, although not explicitly highlighted. The positions of the left-handed 11- and 12-start helices are identified and are defined by the red lines crossing the green horizontal line. Subunits labeled n and n+11 are connected with red lines

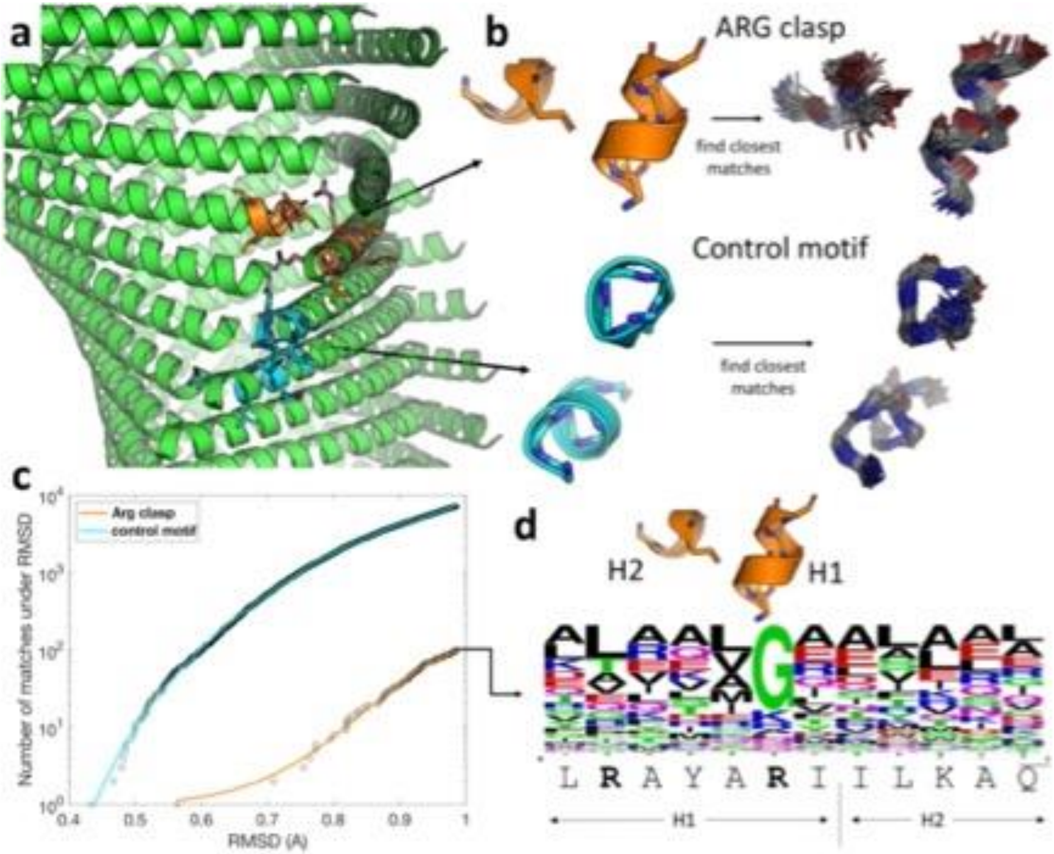
nomenclature described above, with the value of z equal to 2 for the RxxR motif (Supplementary Fig. 6 and 7). The 29-20-2 peptide adopted an α-helical conformation in solution and formed high aspect-ratio tubular filaments (Supplementary Figure2-8 and Supplemental Figure2-9). However, the observed widths of the filaments were significantly larger, circa 12 nm, than the corresponding values for peptides in the (x-y-3) series. A 3.8-Å resolution helical reconstruction from the corresponding cryo-EM images afforded an atomic model for the 29-20-2 filament (Supplementary Table 2). Unexpectedly, the filament formed a double-walled cross-α nanotube that was similar in structure to that of the Form II filament and a structurally related mutant (**Figure 2-5**). Remarkably, the arginine residues in RxxR motif were not involved in the lateral interactions between cross-α helical protofilaments. Instead, cohesive interactions were mediated through hydrogen-bonding between the N- and C-terminal glutamine residues from helices on laterally adjacent protofilaments, as in the Form II-related structures (**Figure 2-1**).

Each nanotube was derived from the coalescence of three protofilaments derived from the dimeric protomer. However, these were arranged in different symmetries; C3 for 29-20-2 and C1 for Form II. In either case, the hydrophobic stacking interactions were oriented along the direction of the three 3-start helices of the respective nanotubes and coincided with the cross-α protofilaments (**Figure 2-5-g**). Among all three structures, the helices within the dimeric asymmetric unit could be superimposed with only minimal differences (**Figure 2-5-e**). Structural overlays of laterally interacting protofilaments indicated a conservation of the interfaces between the three filament structures **(Figure 2-5-f**). The main difference between the structures was manifested in terms of the lateral interactions between the termini of the protofilaments. These interactions occurred along the direction of the left-handed 12-start and 11-start helices for 29-20-2 and Form II, respectively, as illustrated in the helical net diagrams of the corresponding assemblies (**Figure 2-5-g**). The lability of helical symmetry is common among filamentous peptide assemblies and has been observed for structures in which the interfacial interactions are largely conserved. These differences can arise solely as a result of very minor changes in packing preferences of protomers in the filament.[76-80]

The structural analysis of the 29-20-2 filament provided evidence that the formation of the arginine clasp interaction required not only the presence of the two arginines but also the correct orientation, RxxxR, of the corresponding residues within the sequences of the respective helical protomers. The relatively minor substitution of an RxxR motif for an RxxxR motif drove the lateral association down an alternative pathway that resulted in the formation of a double-walled nanotube in which lateral interactions occurred between terminal glutamine residues. The latter interaction must be relatively weak as it can only sustain lateral association for a peptide having this single length. In contrast, the RxxxR motif could mediate lateral interactions over a wider range of peptide lengths from 2 to 5 heptad repeats. However, the axial interactions along the contour length of a protofilament for the Form II-like assemblies appeared quite robust and comparable in energy and buried surface area to that of Form I protofilaments.

2.2.4- Native designability of helical interfaces

The conserved structural context of the RxxxR motif in the Form I-like assemblies raised the question of its native designability, i.e., whether it occurred as a structural feature within natural proteins in which the helices were oriented similarly to those participating in the lateral interactions of Form I-like filaments. To address this question, we defined a local lateral interface motif from each of the filament structures solved here and identified sub-structures in the PDB that displayed similar backbone geometries. The results of this analysis are reported for the 36-31-3 filament (**Figure 2-6**), although similar results were obtained for the other Form I-like assemblies. The motif comprised a seven-residue helical fragment containing the RxxxR motif and a five-residue fragment from the C-terminus of a laterally interacting helix **(Figure 2-6-a)**. Despite this compact size, the closest matching geometries from the PDB for this lateral interaction exhibited root mean squared deviations (RMSDs) over backbone atoms from the query between 0.56-0.76 Å. Furthermore, even these were rare geometries, as the RMSD of subsequent closest matches rose



<u>Figure 2-6:</u> Structural analysis of arginine clasp designability.
a, The 36-31-3 assembly showing the definition of the motif used to interrogate for the presence of the arginine clasp in native structures (orange), as well as a control motif of the same size and complexity (cyan). **b,** Structural alignments of resulting 100 closest matches for both motifs. The control motif, corresponding to the helix stacking interaction within a protofilament, clearly corresponds to a well-populated structural attractor, while the arginine clasp motif brings in much more diverse hits. **c,** RMSD divergence plots, showing the number of matches below a given RMSD cutoff for both motifs. Matches to the Arg clasp motif are rare as the RMSD required to get up to a certain number of matches grows much more rapidly than for the control. **d,** Sequence logo of the closest non-redundant sequence matches to the Arg clasp fragment, showing the query sequence underneath. The native statistics do not bear out the Arg clasp sequence motif, and the one conserved position corresponds to a Gly rather than an Arg. This preference is due to close helix-helix interactions (see Supplementary Fig. 10) and represents a different structural circumstance than the close-to-right angle docking of one helix into another, as seen in our assemblies.

quickly (**Figure 2-6-c**). To demonstrate the rarity of this geometry, we compared these statistics with those of a control motif of equal complexity, corresponding to a seven-residue helix fragment, as described above, but a different five-residue helix fragment from an axially adjacent partner within the same protofilament (**Figure 2-6-a**). In all cases, this reference motif gave much more robust statistics, with many low-RMSD matches and sequence preferences consistent with the amino-acid choices in the assembly. The latter results suggested that the stacking interactions within a protofilament were locally consistent with typical helix-helix interactions, which implies that the axial helix stacking interactions are well represented in naturally occurring proteins and, therefore, are natively designable. On the other hand, the sequence statistics of the closest hits to the lateral interaction motif did not reveal a strong RxxxR signal. The only strongly conserved position corresponded to the second Arg, but the sequence logo analysis indicated that its overwhelming preference was for Gly (**Figure 2-6-d**). Upon closer inspection, we found the nature of this preference to be entirely unrelated to the arginine clasp. Rather, some of the closest matches arose in the context of an internal helix-helix interaction between two closely approaching helices, necessitating a Gly at the position above (Supplementary Figure 10). The results suggested that the arginine clasp motif, which involves a helix terminus interacting with the side of another helix at a nearly 100° angle, was not common within the structures of naturally occurring proteins and, consequently, not natively designable.

This outcome led us to speculate whether structural motifs other than RxxxR could mediate the lateral interaction between helical stacks and promote nanotube formation. The 36-31-3 filament was employed as the starting point for a computational optimization using the design engine dTERMen.[81, 82] A solubility constraint was applied in which the occupancy of the solvent-contacting b/e/g residues within the heptad repeats was constrained to polar residues. Computational optimization returned a sequence in which the RxxxR motif was replaced with an

LxxxL motif and the C-terminal glutamine was replaced with a leucine residue (**Figure 2-7-a).** In the corresponding model, the network of hydrogen-bonding interactions at the lateral interface was replaced with hydrophobic interactions. In a control experiment, a second computational optimization was performed in which positions 6, 7, 10, and 36 were constrained to those in the original peptide sequence. The arginine residues at positions 6 and 10 defined the RxxxR motif, while residues at positions 7 and 36 (Ala and Gln, respectively) appeared to be important for organizing this motif. A comparison between the two sequences, 36-31-3_LL and 36-31-3_RR, could provide additional insight into importance of structural preservation of the arginine clasp interaction. Even in the absence of explicit sequence constraints, the resultant computational

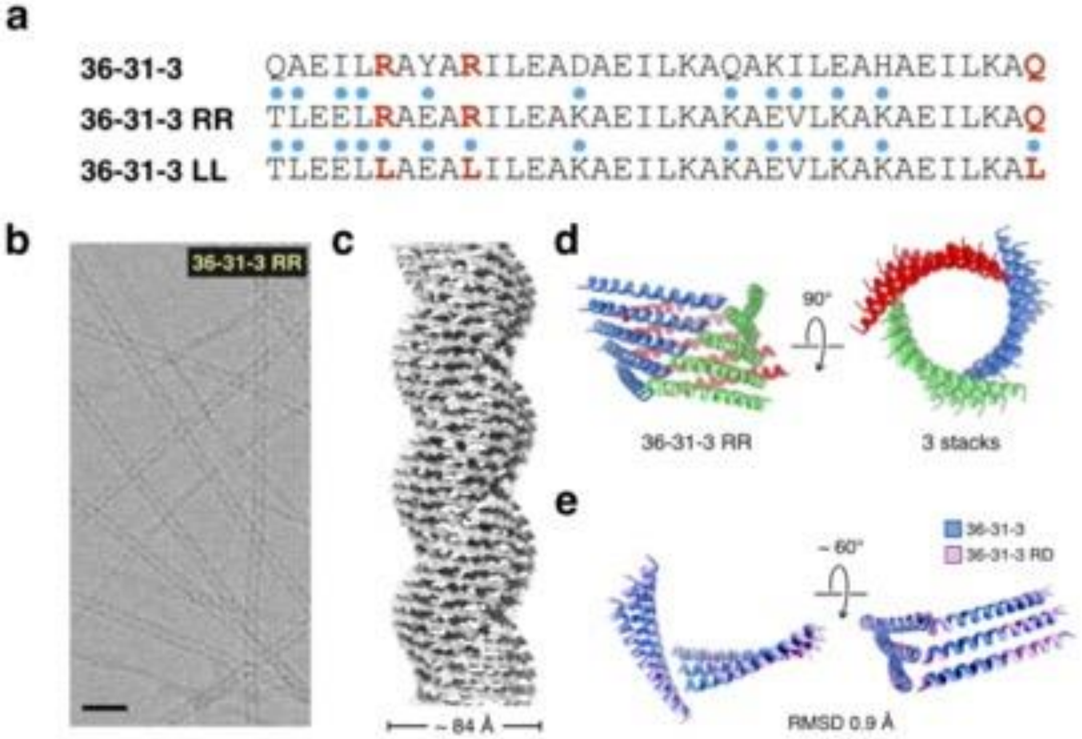


Figure 2-7: Cryo-EM of the computationally re-designed (36-31-3) peptide nanotube.
a, Sequence computationally redesigned peptides 36-31-3_RR and 36-31-3_LL in comparison to the parent sequence. Residues changed from original the **36-31-3** sequences are highlighted with blue dots on top. **b**, Representative cryo-EM image of 36-31-3_RR peptide filaments out of a total number of 166 images recorded. Scale bar is 20 nm. **c**, Cryo-EM reconstruction of the 36-31-3_RR filaments. **d**,Side view and top view of the atomic model. **e**, Alignment of two adjacent cross-α stacks between 36-31-3 (blue) and 36-31-3_RR (magenta).

models retained most of the original amino acid residues at the a/d and c/f positions that mediated the heterotypic hydrophobic interaction at the axial stacking interfaces of the protofilament (**Figure 2-7-a**). Coupled with the structural similarity of axial stacking interactions between the Form I-like and Form II-like assemblies, these results provided further support that the helix-helix interactions along the axial interface were natively designable (**Figure 2-6-c**).

The computationally designed peptides, 36-31-3_LL and 36-31-3_RR, displayed an α-helical conformation and formed high aspect-ratio filaments of similar width to those of 36-31-3 (**Figure 2-2-5**). However, filaments derived from the 36-31-3_LL peptide unraveled into individual protofilaments over time (Supplementary Figure 2-5e). In contrast, 36-31-3_RR filaments persisted over weeks in solution **(Figure 2-7-b and Figure 2-2-5f).** The differences in stability between the two computationally designed filaments suggested that the cohesive interactions were much weaker in the absence of the arginine clasp interaction. The most probable explanation resided in the replacement of the hydrogen-bonding network of the arginine clasp interaction within 36-31-3_RR with the network of hydrophobic interactions within 36-31-3_LL. In order to confirm that the arginine clasp was conserved, filaments of 36-31-3_RR were analyzed by cryo-EM. A 4.5-Å resolution reconstruction of the 36-31-3_RR filaments displayed a nearly identical structure to that of the original 36-31-3 filaments, with very slight differences in helical symmetry for the rise (2.49 versus 2.51 Å/subunit) and rotation (124.5 versus 124.0 degrees/subunit) for the respective 1-start helices of the 3D models. Accordingly, repeating the statistical analysis of this assembly's RxxxR motif produced very similar results as for the original 36-31-3 filament, indicating the native rarity of the geometry.

2.2.5- Structural comparison with other filamentous assemblies

Tayeb-Fligelman, et al., reported the initial example of a cross-α fibril structure from crystallographic analysis of PSMα3, a short cytolytic peptide secreted from virulent strains of S. aureus.[83] The helices were observed to pack in a stacked bilayer array, in which individual peptides were oriented perpendicular to the long axis of the crystallographically defined filament. In contrast to the cross-α nanotubes described here, translational symmetry restricted coiling of the cross-α filaments. Moreover, the composition of the hydrophobic interface was rich in phenylalanine residues and did not display the KIH packing characteristically associated with coiled coils. Subsequently, Zhang, et al.,[84]reported a series of peptides based on designed coiled-coil sequences that assembled into filaments. Crystallographic analysis demonstrated that the peptides formed cross-α super-helical arrays through stacking of parallel dimers in an alternating antiparallel orientation. The filaments displayed limited interfaces over which KIH packing was observed between protomers, in contrast to the cross-α protofilaments in the nanotubes that we have described. Supercoiling of filaments was only observed under conditions in which the cross-α orientation was not maintained in the extended structure. In addition, our cross-α nanotubes differed from classical coiled-coils, in that the helical packing in a protofilament displayed an open or extended mode of association that is reminiscent of solenoidal tandem repeat proteins.[85-88]Most coiled-coils form closed assemblies of defined oligomerization state, which are often, but not exclusively, based on cyclic symmetry in which the helical protomers are aligned parallel to the super-helical axis.[89, 90] Surprisingly, the closest structural analogues of these cross-α nanotubes were observed among helical supramolecular assemblies derived from oligourea foldamers having similar polar patterning based on pentad repeat sequences.[91] Crystallographic characterization of the latter assemblies revealed that the corresponding oligourea nanotubes were composed of protofilaments corresponding to stacked helical subunits that propagated along right-handed 2-start or 6-start helices.

Cross-β assemblies[92]represent a natural point of comparison to the cross-α nanotubes. The respective protofilaments differ in that cross-β architecture is primarily stabilized through main-chain hydrogen-bonding interactions along the contour length of the assembly. In contrast, the cohesive interactions between protomers within the cross-α protofilaments depended exclusively on side chain interactions. Surprisingly, few near-atomic resolution structures of synthetic cross-β nanotubes derived from designed β-strand sequences have been reported, presumably due to the propensity for structural polymorphism and concomitant difficulty in isolation of monomorphic variants.[93-96]However, the formation of cross-β nanotubes of uniform diameter has been reported for conformationally constrained cyclic peptide sequences.[97, 98] More recently, cryo-EM helical reconstruction of Orb2 filaments from D. melanogaster brain revealed a minimal diameter nanotube composed of three identical cross-β protofilaments in a C3-symmetric arrangement.[99]Individual cross-β protofilaments were based on β-hairpin protomers interacting through a combination of main-chain and side-chain hydrogen-bonding interactions. Similar to the cross-α nanotubes, side-chain hydrogen bond formation provided the cohesive interactions between protofilaments that stabilized the cross-β assembly. These results suggested that directional non-covalent interactions, e.g., the formation of complementary hydrogen-bonded networks or metal ion coordination,[100, 101] could provide a mechanism for rational design of defined higher-order structure in synthetic peptide and protein assemblies.

2.3- Conclusion

The de novo design of peptide and protein assemblies requires not only the control of interfacial interactions, but also the folding energetics of the individual subunits.[102, 103] In this study, the protomer sequences were based on the highly designable coiled-coil motif, which has been demonstrated to display robust folding energetics and structural stability in sequence space. The axial stacking interactions within a protofilament were highly conserved between assemblies differing in both the number of protofilaments and nature of the lateral interactions between protofilaments. The structure of the axial interfaces within a protofilament were consistent with helix-helix interactions in canonical coiled-coils, even though the open-ended supramolecular architecture of the helical assemblies was quite distinct from the closed cyclic oligomers that are typically associated with such helical bundles. In contrast, the lateral interactions between protofilaments have not been commonly observed within the structures of native proteins having similar helix-helix orientations. Notably, we corroborated that the lateral interface between Form I-like protofilaments was based on the unusual interaction of an RxxxR motif with the C-terminal residues of a helix in a structurally adjacent protofilament. Our investigation of the designability of this arginine clasp motif suggested that it does not appear to have been evolutionarily sampled as a mechanism to mediate helix-helix interactions in native protein structures and cannot be considered as natively designable. Nevertheless, the RxxxR motif appeared robust as a design element within a structurally related family of cross α-helical assemblies and was preferred to other more designable interaction motifs, such as the common LxxxL motif, within an otherwise identical peptide sequence context. We demonstrated that within a designed series of coiled-coil peptides that the RxxxR motif could be employed to control nanotube structure through peptide length. The absence of the RxxxR motif drove the assembly down an alternative pathway, which suggested that the arginine clasp interaction was not robust outside of this relatively limited sequence context.[104]

From this perspective, we propose that the RxxxR clasp may be representive of a more general class of local interaction motifs that can potentially mediate quaternary interactions within a specific structural context in peptide and protein assemblies.[105]The influence of these localized interaction motifs may be magnified due to the presence of multiple copies in a structurally similar context along the contour length of the helical assembly, which enables selection for specific modes of self-association of protofilaments into higher order structures. Similar local interaction motifs, e.g., Phe-Phe diads[106-10807-109107-109108-110108-110108-110108-110108-110] and glutamine/asparagine ladders,[107, 109] among others, have been identified that mediate higher-order interactions in cross-β assemblies. The relatively low information content of such short sequence motifs hinders a systematic sequence-based analysis of their role in higher order assembly processes. The challenge will be to effectively identify such motifs in order to exploit them for de novo design of helical assemblies. This process requires the application of high-resolution structural analysis in order to identify the presence of such structurally determinative interactions. Cryo-EM is the primary method for structural characterization of helical filaments. More broadly applied structural investigations of naturally occurring and synthetic peptide and protein filaments represent the most promising approach to identify local structural interactions that contribute to specificity and stability within these assemblies.

2.4- Material and methods

2.4.1- Peptide synthesis, Purification and assembly

Chemical reagents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) or Anaspec, Inc. (Fremont, CA) unless otherwise specified. Peptides 15-10-3, 22-17-3, 29-24-3, 36-

31-3, 36-31-3_RR, 36-31-3_LL and 29-20-2 were obtained from SynPeptide Co., LTD. (Shanghai, China). All purchased chemical reagents and peptides ordered were used without further purification. The Form II and Form IIa peptides were prepared via microwave-assisted solid phase peptide synthesis on a CEM Liberty Blue Automated Microwave Peptide Synthesizer as the N-acetyl, C-amide capped derivatives. A PAL-PEG-PS resin from Applied Biosystems (Foster City, CA) was used for synthesis of both peptides. Standard Fmoc protection chemistry was utilized in conjunction with coupling cycles consisting of HBTU/OXIMA-mediated activation protocols and base-induced deprotection (20% piperidine in N, N-dimethylformamide with 0.1 M hydroxybenzotriazole) of the Fmoc group. After synthesis, the DMF/resin mixture was filtered and rinsed with acetone then air-dried. The peptides were cleaved from the resin by incubation at room temperature for 3 hours in a cocktail consisting of 92.5% trifluoroacetic acid (TFA), 2.5% distilled water, 2.5% triisopropylsilane, and 2.5% 2,2’- (ethylenedioxy)-diethanethiol. Cleavage was followed by filtration and subsequent precipitation in diethyl ether. The precipitate was allowed to desiccate overnight. The crude peptides were resolubilized in 3 mL of a 50:50 mixture of acetonitrile and water (0.1% TFA additive) and purified using a Shimadzu LC-20AP reversed-phase high-pressure liquid chromatography (HPLC) equipped with a C18 column. The peptides were eluted with a linear gradient of water-acetonitrile with 0.1% TFA. Peptide mass was confirmed using ESI mass spectrometry. Purified HPLC fractions were lyophilized, sealed, and stored at -30 °C.

Stock solutions of all peptides (3 mg/mL) were prepared by solubilizing purified, lyophilized peptide (0.75 mg) in 250 μL of acetate buffer (10 mM, pH 4.0). The solution was titrated with dilute sodium hydroxide solution to adjust the final pH value to 4.0. The solutions were allowed to assemble from 24 hours to 2 weeks depending on the sequence. Peptides 36-31-

6_RR and 36-31-6_LL were thermally annealed using the following thermal cycler protocol: (1) rapid heating to 90 °C for 30 minutes and (2) cooling to 25 °C at a rate of 0.2 °C/minute.

2.4.2- Circular dichroism spectropolarimetry

CD measurements were performed on a Jasco J-1500 CD spectropolarimeter using 0.10 mm thick quartz plates (Hellma Analytics). Three CD spectra were collected and averaged for each peptide sample. Each spectrum was acquired at a scan rate of 100 nm/min from 190-260 nm with a bandwidth of 2 nm and a data pitch of 0.2 nm.

2.4.3- Negative Stain Transmission electron microscopy Analysis

TEM Grids were prepared using dilute solutions of peptide (3 mg/mL) in aqueous buffer (10 mM acetate, pH 4.0). Samples were prepared by depositing a peptide solution (4 μL) onto a 200-mesh carbon-coated copper grid from Electron Microscopy Services (Hatfield, PA). After 90 s of incubation on the grid, the excess liquid was wicked away, leaving a thin film of sample. An aliquot (4 μL) of negative stain solution (1% uranyl acetate) was deposited onto the thin film. After 1 min of staining, the remaining moisture was wicked away, and the grid was dried overnight in a desiccator. Electron micrographs were captured on a Hitachi HT-7700 transmission electron microscopy, equipped with a tungsten filament and AMT CCD camera, operating at an accelerating voltage of 80 kV.

2.4.4- Cryo-electron microscopy

The peptide sample was applied to glow-discharged Quantifoil (1.2/1.3 or 2/2) or lacey carbon grids, and then plunge frozen using a Vitrobot Mark IV (FEI). Cryo-EM data sets on peptides 15-10-3, 29-24-3, 29-20-2 and 36-31-3_RR were collected on a 200 keV Talos Arctica with a K2 camera (Emory University) at 1.04 Å/pixel and a total dose of ca. 55 e/Å^2. The cryo-

EM images for the dataset of peptide 36-31-3 were collected on a 300 keV Titan Krios with a K2 camera (National Cryo-EM Facility at NCI) at 1.06 Å/pixel and a total dose of ca. 50 e/Å^2. Cryo-EM datasets on peptides Form II and Form IIa were collected on a 300 keV Titan Krios with a K3 camera (University of Virginia) at 1.08 Å/pixel and a total dose of ca. 51 e/Å^2. Cryo-EM movies were recorded in counting mode using EPU v2.4 (Thermo Fisher). The micrographs were first motion corrected and dose weighted by MotionCorr v2,[110] and then a CTF correction was applied by multiplying the images with the theoretical CTF using CTFFIND3. Filament images corresponding to ca. 20 electrons/Å^2 were extracted using e2helixboxer (EMAN2). [111 111 111 111 111 111 111 111 111 112 112 113 113 113 113 113] For each peptide filament, a list of possible helical symmetries was calculated from the averaged power spectrum of peptide particles. To determine the correct helical symmetry from the list, first if the full dataset has more than 30,000 particles, a subset containing 30,000 particles was generated. And then the initial volume was generated from those 30,000 particles with random assigned azimuthal angles. After that, possible helical symmetries were tested in Spider v22.10 using IHRSR[112, 113] by trial and error until recognizable amino acid side chain densities could be observed in the correct symmetry (see Supplementary Fig. 11 and 12 for representative examples of this process). Specifically, the IHSRS reconstructions were divided into 3 steps, where 4x binned, 2x binned, un-binned particles were used to accelerate the progress. Helical symmetries were tightly locked in the first two steps, and then slightly relaxed in the third step. All helical parameters, including rise, rotation and point-group symmetry were imposed at the end of each IHRSR cycle. After determining the correct symmetry, final reconstructions using the all particles were run in both RELION[114] v3.0 and Spider[115] v22.10. The resulting volumes from both approaches were sharpened with the same negative B-factor, and then selected by eye based on the quality of the side chain densities. The statistics are listed in Supplementary Tables 1 and 2.

2.4.5- Model Building

As the cryo-EM maps of all peptide samples reach to 4.5 Å resolution or higher, right-handed α-helices were used to determine the helical hand of the EM volumes. It was clear from the map that each peptide molecule forms a single α-helix. Therefore, an α-helix model was generated in UCSF Chimera[116] v1.14 from the corresponding peptide sequence and then docked in the respective EM map. This single α-helix model was adjusted manually in Coot[117] v0.8 to best fit into the map. Finally, this adjusted single α-helix model was used to generate a model filament using the determined helical symmetry, which was then refined against the full cryo-EM map using real space refinement in PHENIX[3] v1.16. Final geometries of the atomic models were validated with the MolProbity[118] implementation in PHENIX. The refinement statistics are shown in Supplementary Tables 1 and 2. Cryo-EM maps and atomic coordinates have been deposited with the Electron Microscopy Data Bank and Protein Data Bank with accession codes given in Supplementary Tables 1 and 2. Model versus map FSC calculations were employed to estimate the resolution of the reconstructions and are reported in Supplementary Fig. 13.

2.4.6-Motif Analysis

A search database was created by filtering the PDB, as of 04/19/2020, for the following: X-ray structures with resolution of 2.6 Å or better, containing at least 60% protein residues, with no more than 5000 residues and no more than 26 chains in the biological unit. This resulted in a total of 146,052 biological-unit entries. All-to-all sequence clustering of chains within these biological units was then performed at 50% sequence identity using the USEARCH tool v10. [119]
119 119 119 119 119 119 119 119 120 120 120 120 120 120 120 Using this information, a minimally non-redundant

subset of biological units was selected by eliminating a biological unit A if another one in the set contained a super set of its chains (i.e., had chains belonging to the same clusters as chains of A and additional chains). Resolution was used as the tie-breaker when two biounits had equivalent sets of chains. In the end, 23,915 biounits survived with a total of 176,442 chains and 43,449,344 residues. This set of structures was used to search for the Arg clasp structural motif (or the control motif) using the MASTER program.[120, 121] All non-redundant matches (i.e., less than 50% sequence identity in the matching region—the matching fragments themselves +/- 15 residues on each side) with backbone RMSD below 1.0 Å were sought and used for analysis.

2.4.7-Computational design

dTERMen[122] was used to perform design. dTERMen is a general-purpose method for computational protein design that is based on the concept of structural degeneracy of proteins. Central to this approach is the concept of a TERM (short for tertiary motif)—a fragment of protein structure–consisting of one or more disjoint structural segments, that is reused across unrelated proteins. This reuse, which we described in several studies, [122, 123 122, 123 122, 123 122, 123 122, 123 122, 123 122, 123 122, 123 122, 123 123, 124 123, 124 123, 124 123, 124 123, 124 123, 124 123, 124] allows one to directly mine the PDB for sequence-structure relationships. Specifically, repeated TERM instances allow one to identify sequence features necessary for stabilizing the corresponding structural motif. Thus, dTERMen works by automatically breaking the desired target structure T down into its constituent TERMs, deducing sequence-structure relationships for each using rapid database searches (via the underlying method MASTER,[81]) and combines these to build a model for what sequences would most likely stabilize target structure T.

A python-based implementation of dTERMen was employed in this study, which makes extensive use of the structure search engine MASTER to identify TERM matches and extract their sequence statistics. The structural database for underlying MASTER searches used in this study was prepared on 01/22/2019 following the procedure above. As the design template, a 15-chain portion of the 36-31-3 assembly was used (shown in color in Fig. 7c), with one chain (chain A) fully surrounded by all symmetry mates with which it could make any appreciable interactions. The symmetry-based design feature in dTERMen was used (--image flag), with chain A as the central unit and all others being images. The energy table resulting from running dTERMen described the sequence landscape compatible with folding to the 36-31-3 geometry, based on structural statistics of constituent motifs. Given this table, integer linear programing (ILP) was used to identify the best-scoring sequence under a solubility/fold-specificity constraint, requiring solvent-facing g positions (i.e., 8, 15, 22, and 29) to be polar (i.e., one of Asp, Glu, Gly, His, Lys, Asn, Gln, Arg, Ser, Thr, Tyr). The resulting sequence, TLEELLAEALILKAKAEILKAKAEVLKAKAEILKAL, lacked the Arg clasp motif, which was not surprising in light of our analysis of this motif's natural designability. To test the importance of this motif for assembly formation, we generated another design by re-running the optimization with the additional constraint that preserved the Arg clasp motif (i.e., left identities of positions 6, 7, 10, and 36 at Arg, Ala, Arg, and Gln, respectively). This resulted in the sequence: TLEELRAEARILEAKAEILKAKAEVLKAKAEILKAQ, which was called 36-31-3_RR, while the original design was called 36-31-3_LL.

2.5 -Supplementary Figures

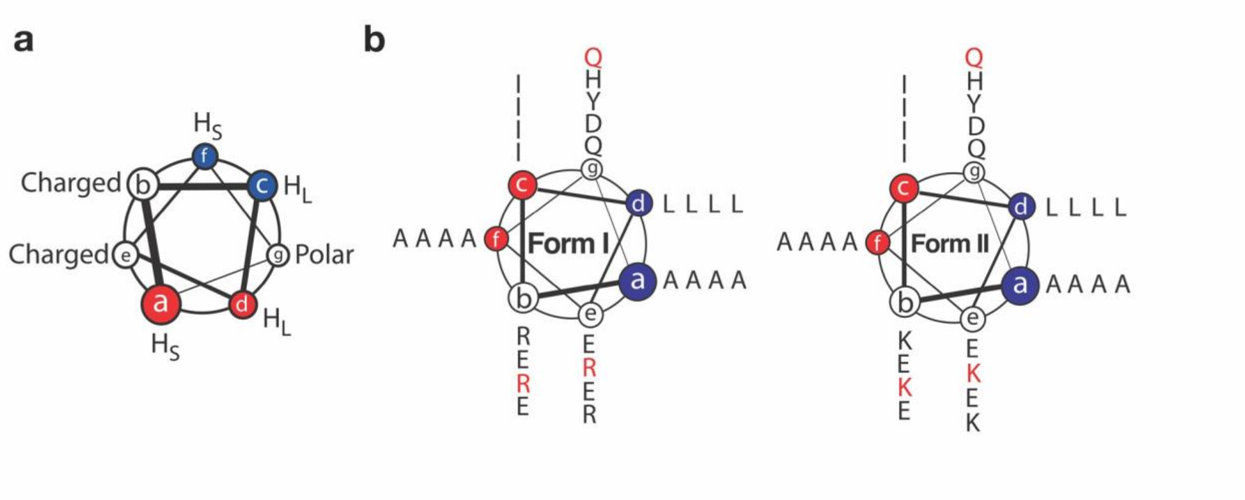


Figure 2-2-1: Helical wheel diagrams representing peptide sequence designs in coiled-coil space. **a**, General design of Form peptide sequences based on the heptad repeat pattern of Type III coiled-coils. The two offset hydrophobic faces corresponding to the *a*/*d* and *c*/*f* residues are highlighted in red and blue, respectively. Selective heterotypic association between the *a*/*d* and *c*/*f* faces results in a two residues displacement between interacting peptides along the helix-helix interface. (H_L, large hydrophobic residues; H_S, small hydrophobic residues) **b**, Heptad sequences of Form I and Form II parent peptides threaded onto a coiled-coil helical wheel diagram. The residues involved in the arginine clasp interaction (R13, R17, Q29) are highlighted in red for the Form I sequence, as well as the corresponding positions in the Form II sequence.

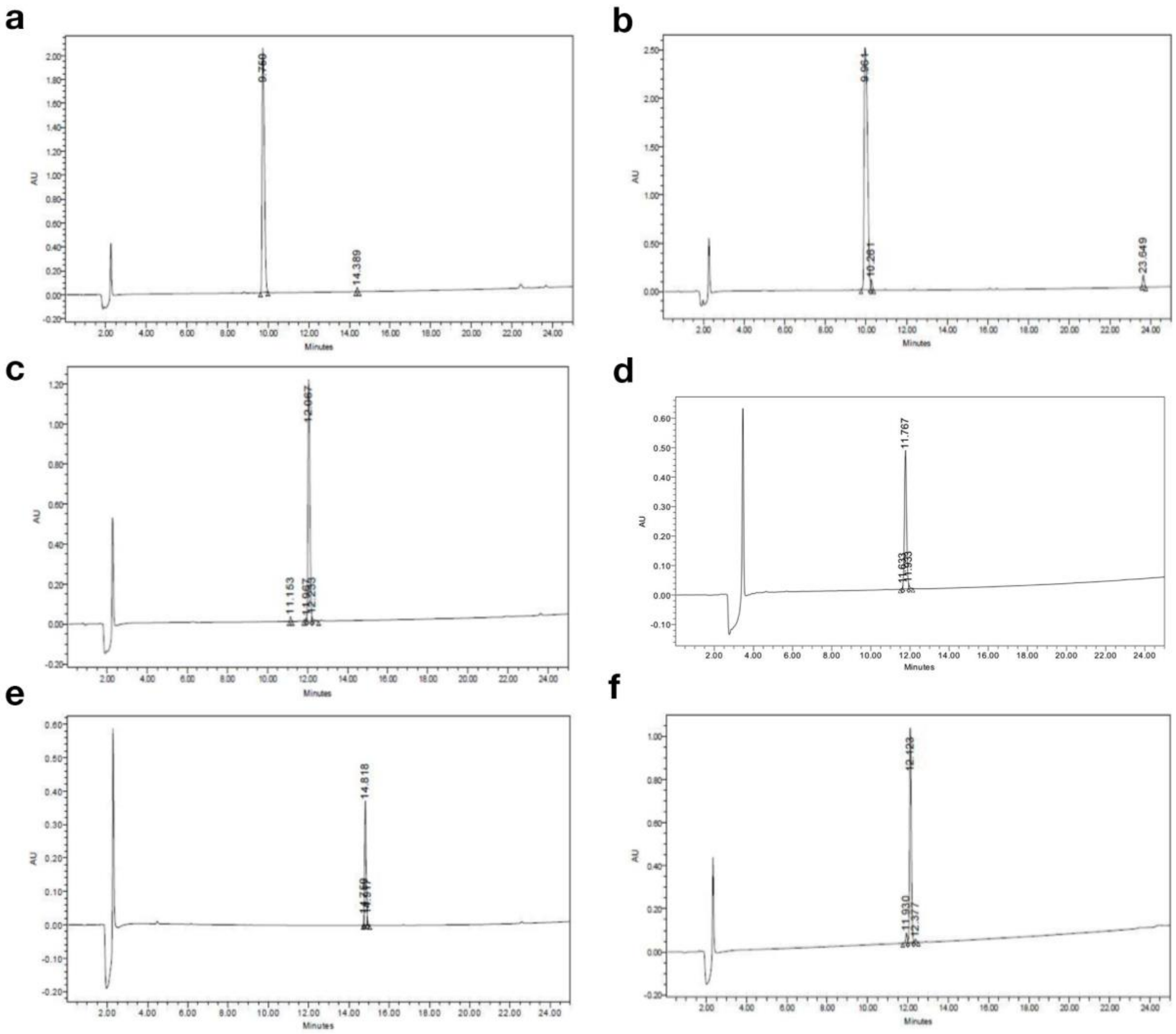


Figure 2-2-2: Analytical HPLC traces of Form I-like peptides (a, 15-10-3; b, 22-17-3; c, 29-24-3; d, 36-31-3; e, 36-31-3_LL; f, 36-31-3_RR).

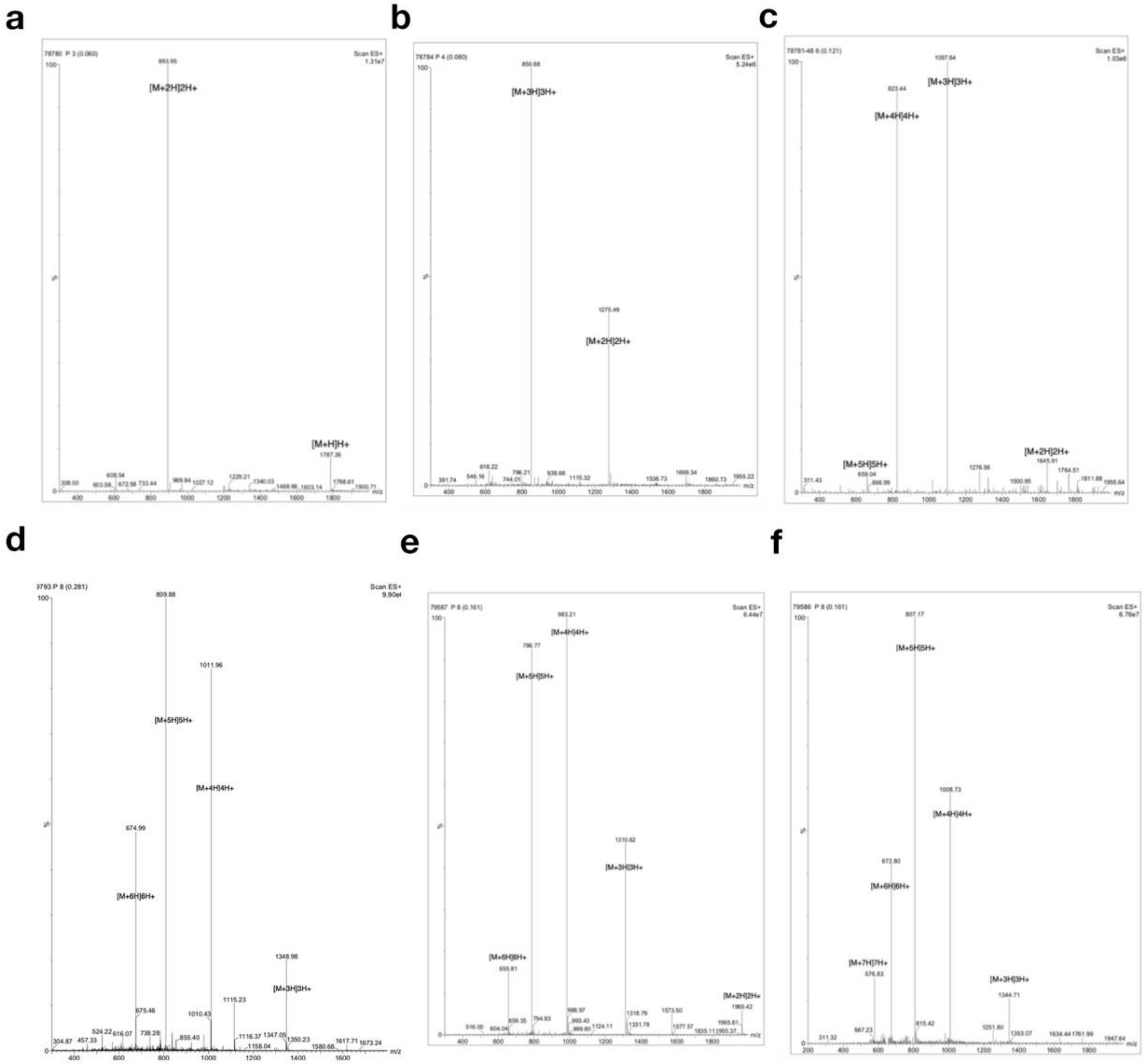


Figure 2-2-3: Electrospray mass spectra of Form I-like peptides (**a**, 15-10-3; **b**, 22-17-3; **c**, 29-24-3; **d**, 36-31-3; **e**, 36-31-3_LL; **f**, 36-31-3_RR).

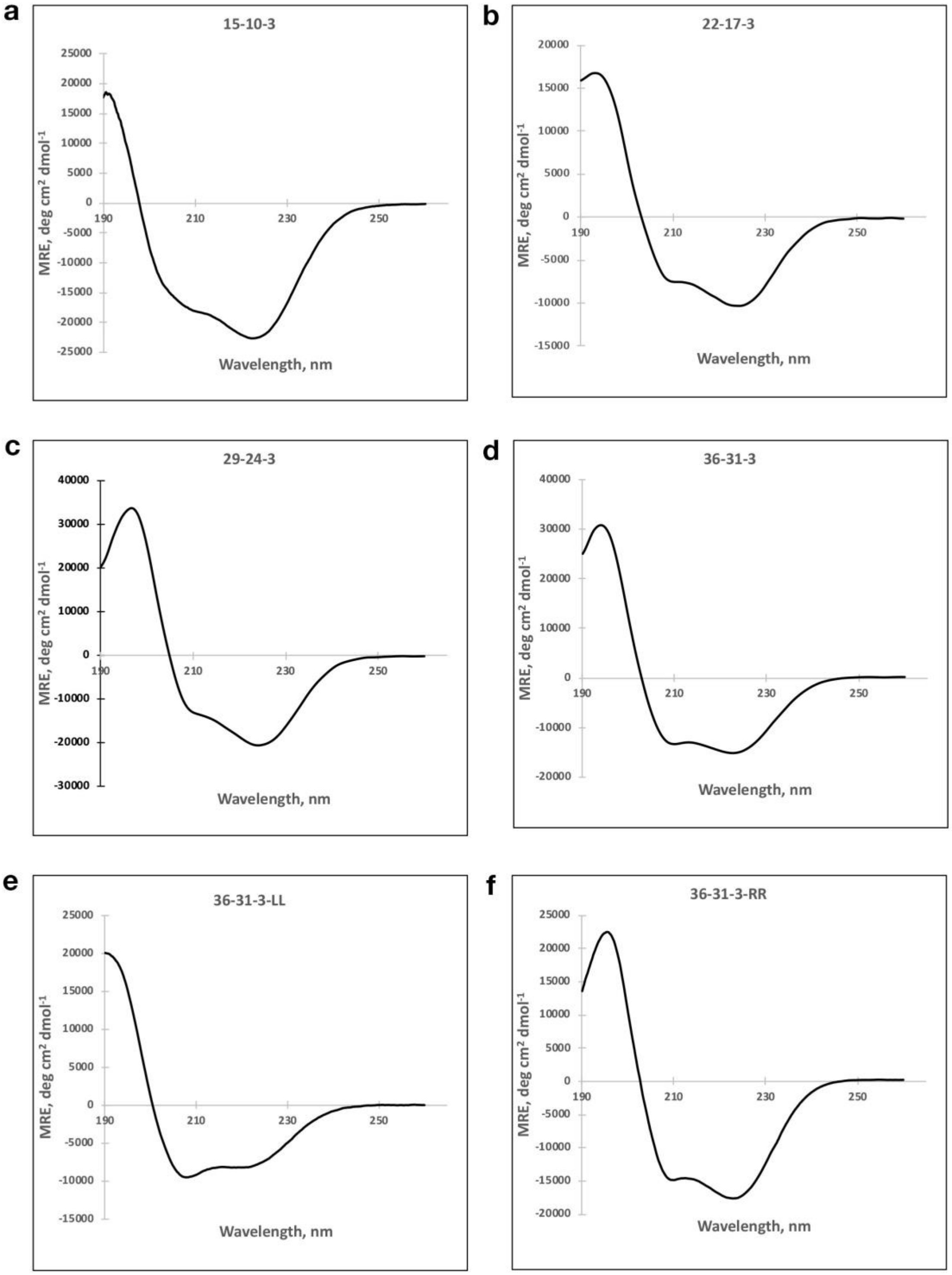


Figure 2-2-1: Circular dichroism spectra of Form I-like peptides (**a**, 15-10-3; **b**, 2217-3; **c**, 29-24-3; **d**, 36-31-3; **e**, 36-31-3_LL; **f**, 36-31-3_RR).

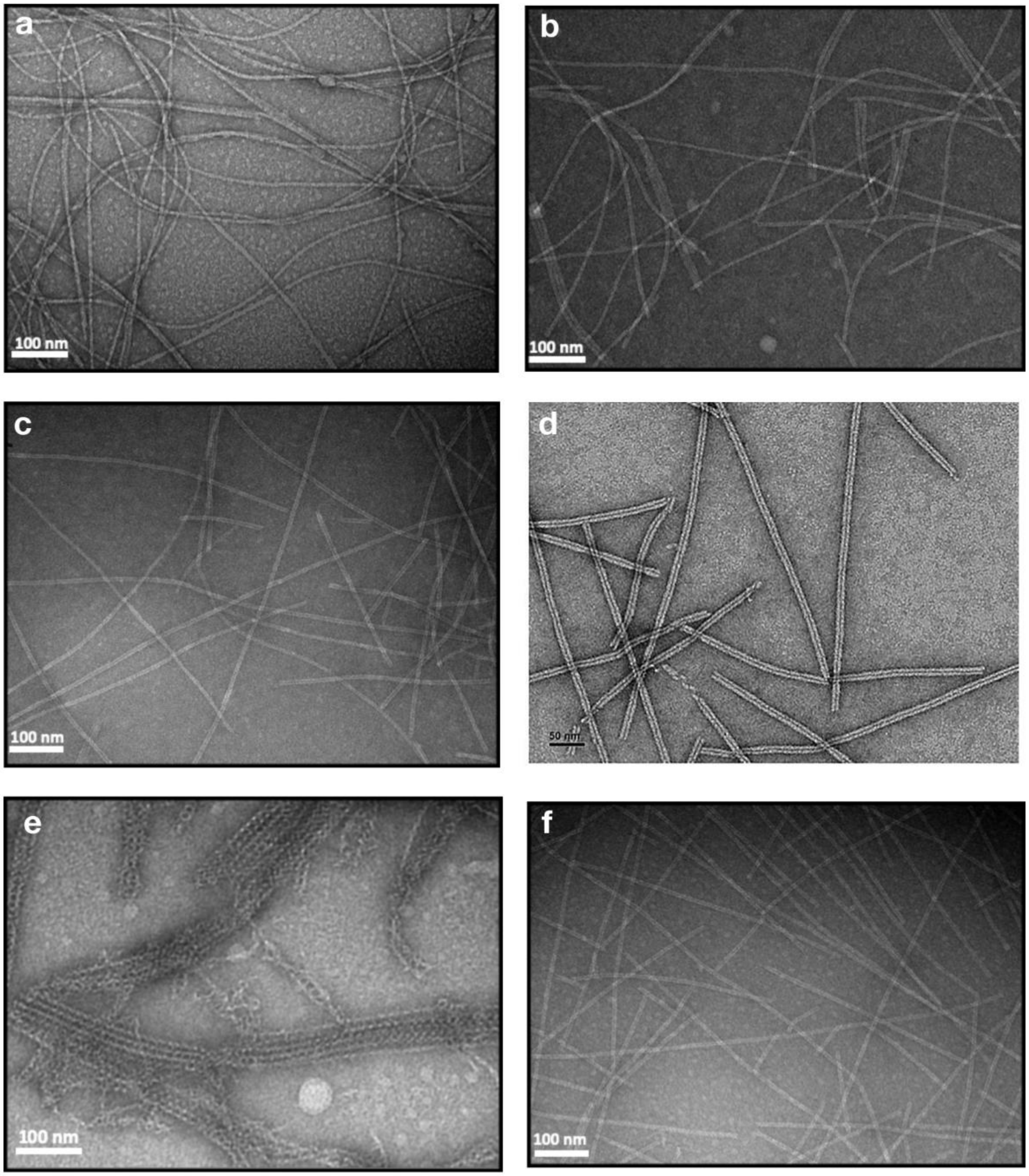


Figure 2-2-5: Representative negative-stain TEM images of Form I-like peptides (**a**, 15-10-3; **b**, 22-17-3; **c**, 29-24-3; **d**, 36-31-3; **e**, 36-31-3 LL; **f**, 36-31-3 RR). Total number of collected TEM images for each sample: 30 (15-10-3); 30 (22-17-3); 30 (29-24-3); 14 (36-31-3); 10 (36-31-3_LL); 15 (36-31-3_RR).

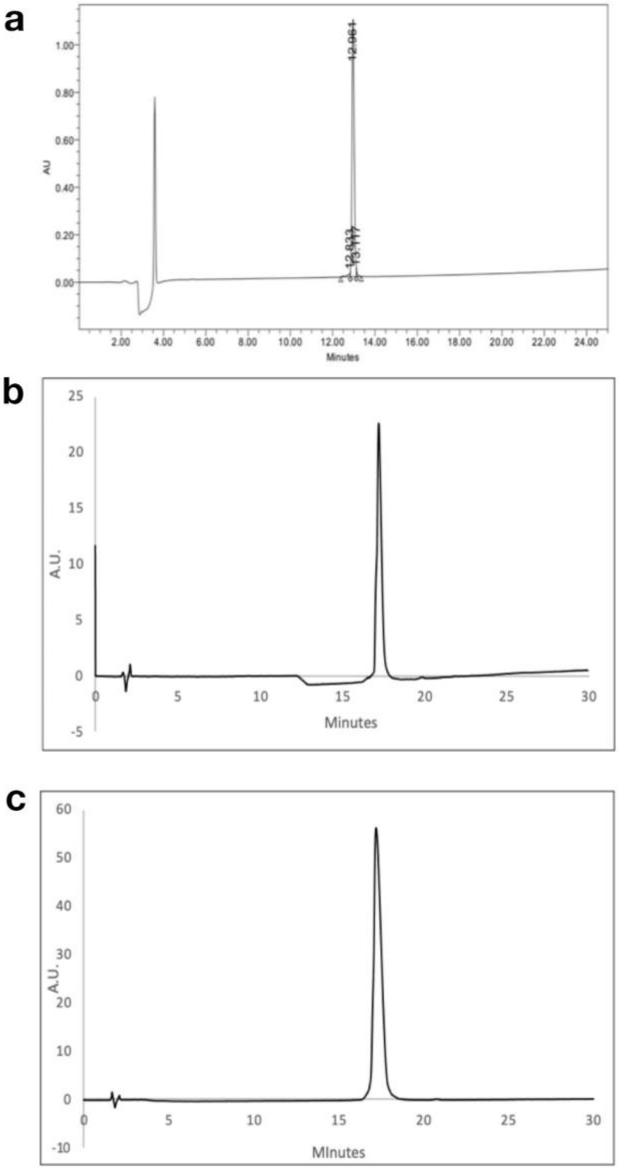


Figure 2-2-2: Analytical HPLC traces of Form II-like peptides (**a**, 29-20-2; **b**, Form II; **c**, Form IIa).

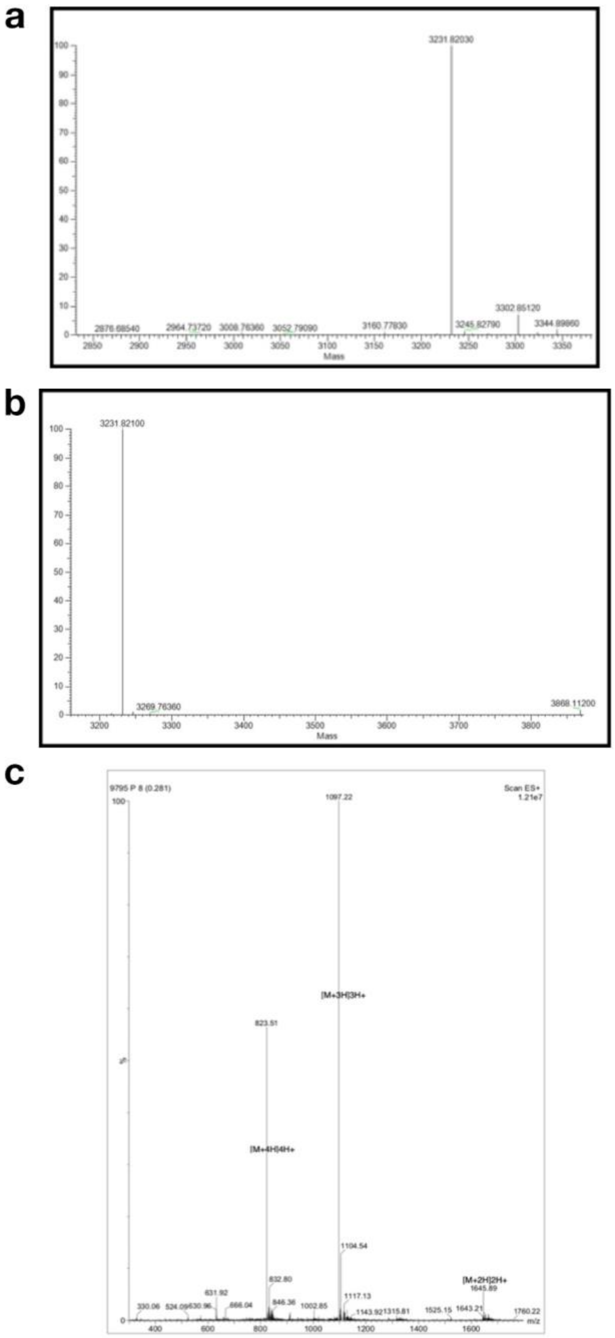


Figure 2-2-3: Electrospray mass spectra of Form II-like peptides (**a**, Form II; **b**, Form IIa; **c**, 29-20-2).

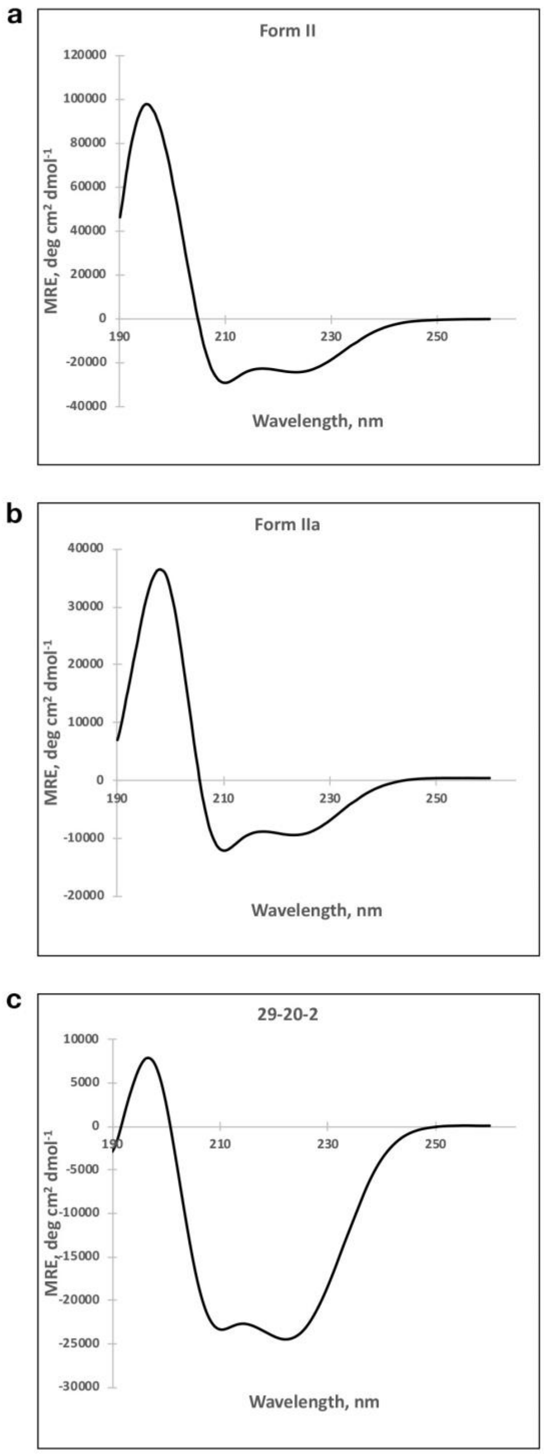


Figure 2-2-4: Circular dichroism spectra of Form II-like peptides (**a**, Form II; **b**, Form IIa; **c**, 29-20-2).

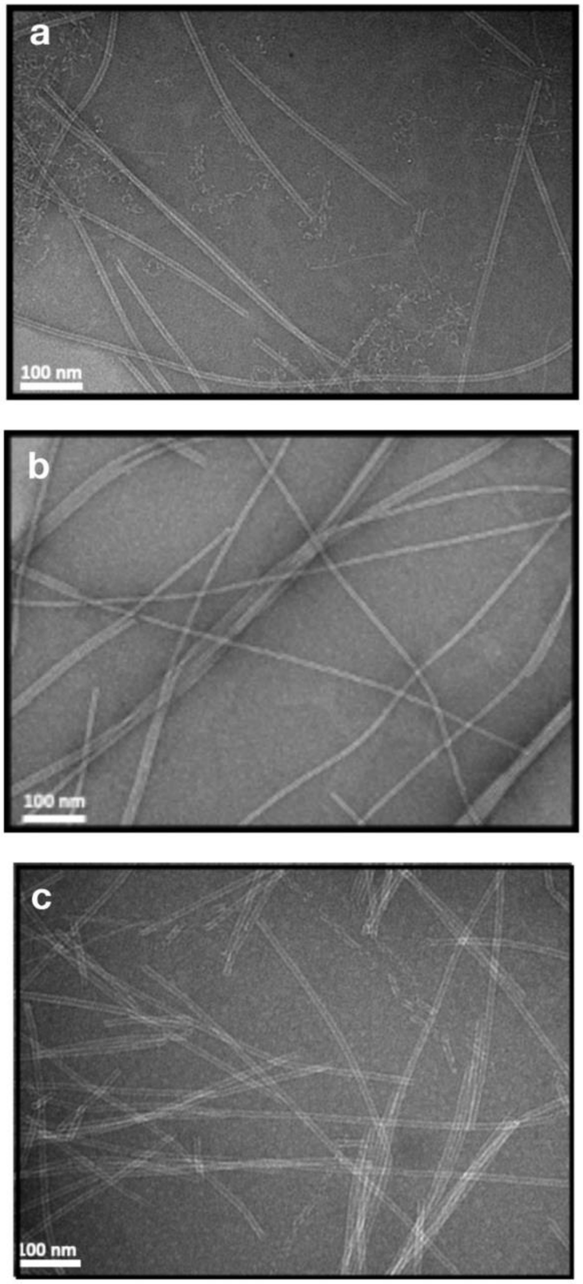


Figure 2-2-5: Representative negative stain TEM images of Form II-like peptides (**a**, Form II; **b**, Form IIa; **c**, 29-20-2). Total number of collected TEM images for each sample: 12 (Form II); 12 (Form IIa); 20 (29-20-2).

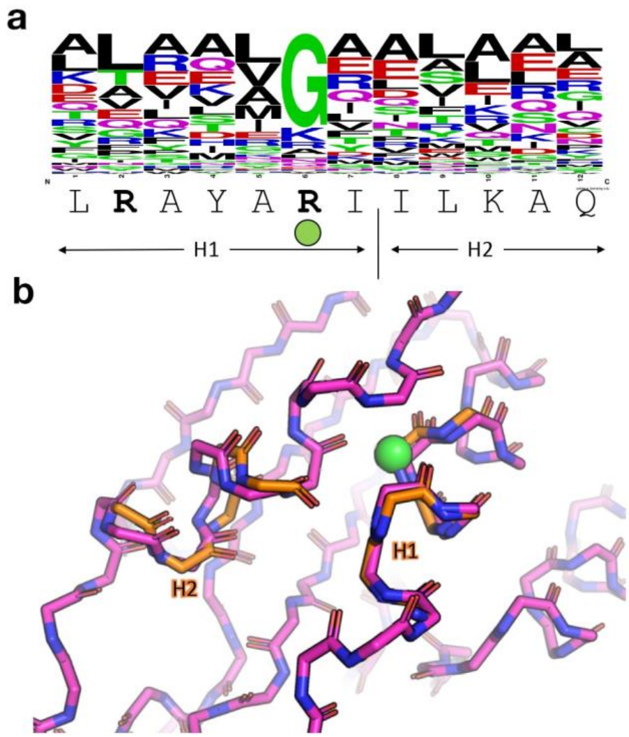
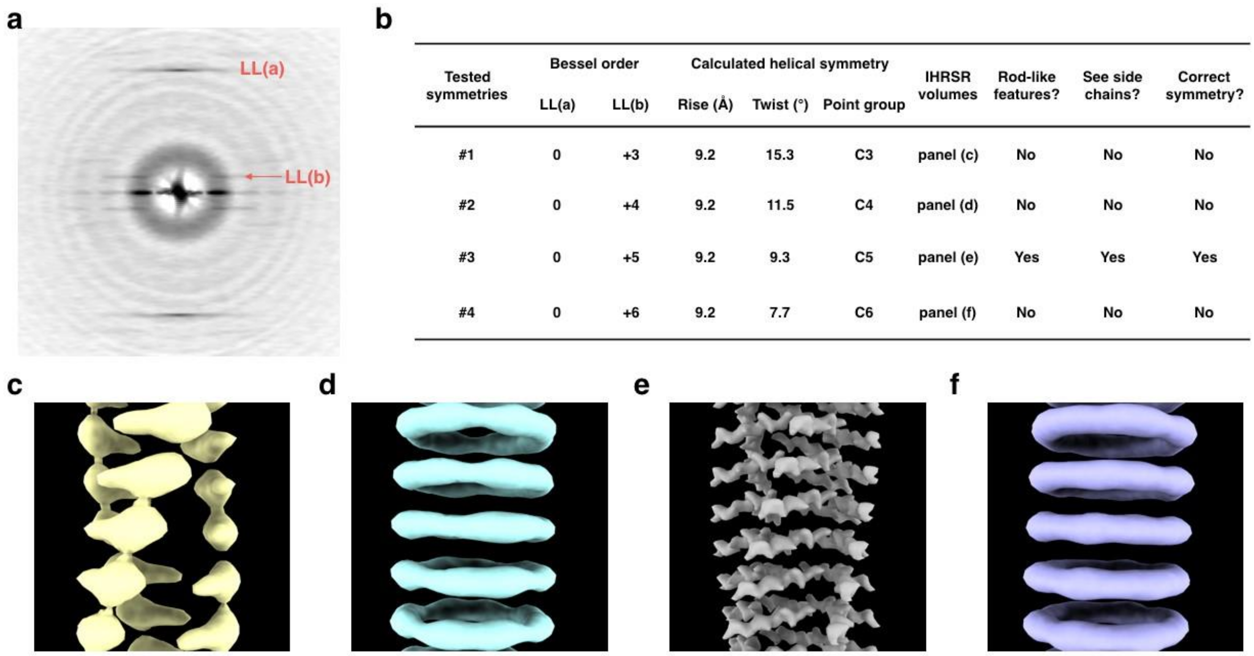
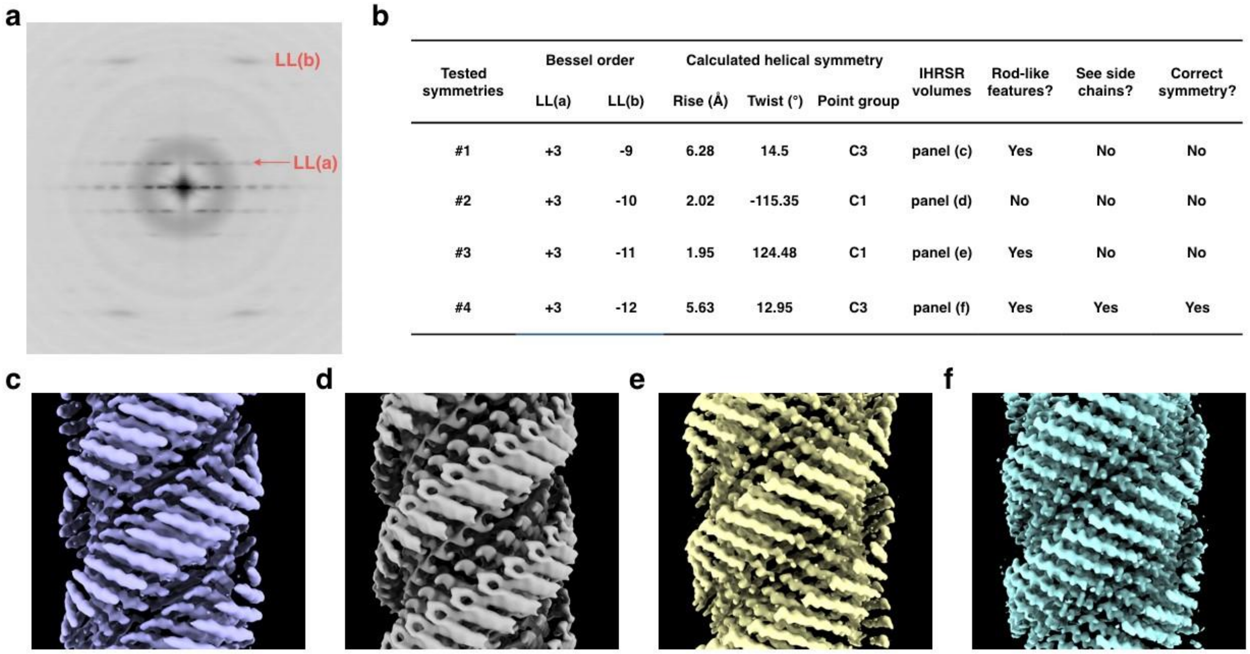


Figure 2-2-6: Comparison between the arginine clasp motif and closest nonredundant matches. **a**, Sequence logo of the closest non-redundant sequence matches to the Arg clasp fragments. The green dot indicates the position of the conserved glycine residue in helix 1 (H1), which occurs in place of the second arginine in the RxxxR motif. Mutagenesis of the arginine at this position in Form I-like structures abrogates the arginine clasp interaction and results in conversion to Form II-like structures. **b**, Structural overlay of the interacting segments of the arginine clasp motif (H1 and H2, orange) in the 36-31-3 structure with the closest match (pink). The latter occurs in the context of an internal helix-helix interaction of two closelyapproaching helices, which necessitates the presence of a glycine residue (green dot).



Tested symmetries	Bessel order		Calculated helical symmetry			IHRSR volumes	Rod-like features?	See side chains?	Correct symmetry?
	LL(a)	LL(b)	Rise (Å)	Twist (°)	Point group				
#1	0	+3	9.2	15.3	C3	panel (c)	No	No	No
#2	0	+4	9.2	11.5	C4	panel (d)	No	No	No
#3	0	+5	9.2	9.3	C5	panel (e)	Yes	Yes	Yes
#4	0	+6	9.2	7.7	C6	panel (f)	No	No	No

<u>**Figure 2-2-7:**</u> Helical symmetry determination of Form I-like peptide, 15-10-3. **a,** Averaged power spectrum of the segments used in the initial reconstruction. The layer lines (LL) that were used to calculate the helical symmetry are labelled as LL(a) and LL(b). **b,** List of tested symmetries. The Bessel order assigned for each symmetry, detailed helical parameters, and additional criteria to judge the results. **c-f,** The resulting IHRSR volumes of tested symmetries listed in panel (**b**)



Tested symmetries	Bessel order		Calculated helical symmetry			IHRSR volumes	Rod-like features?	See side chains?	Correct symmetry?
	LL(a)	LL(b)	Rise (Å)	Twist (°)	Point group				
#1	+3	-9	6.28	14.5	C3	panel (c)	Yes	No	No
#2	+3	-10	2.02	-115.35	C1	panel (d)	No	No	No
#3	+3	-11	1.95	124.48	C1	panel (e)	Yes	No	No
#4	+3	-12	5.63	12.95	C3	panel (f)	Yes	Yes	Yes

Figure 2-2-8: Helical symmetry determination of the Form II-like peptide, 29-20-2. **a,** Averaged power spectrum of the segments used in the initial reconstruction. The layer lines (LL) that were used to calculate the helical symmetry are labelled as LL(a) and LL(b). **b,** List of tested symmetries. The Bessel order assigned for each symmetry, detailed helical parameters, and criteria to judge the results. **c-f,** The resulting IHRSR volumes of tested symmetries listed in panel (**b**)

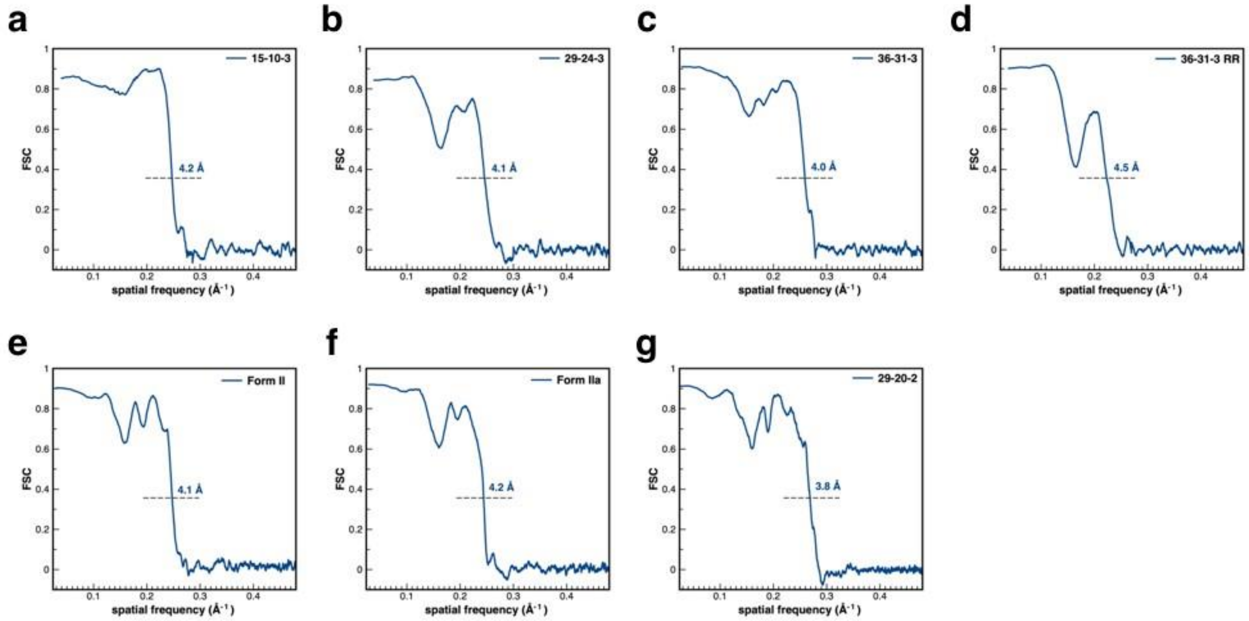


Figure 2-2-9: Fourier Shell Correlation (FSC) calculations. The model:map FSC calculation using a 0.38 criterion, which is sqrt (0.143) to estimate the resolution. (**a**, 15-10-3; **b**, 29-24-3; **c**, 36-31-3; **d**, 36-31-3 RR; **e**, Form II; **f**, Form IIa; **g**, 29-20-2).

2.6- Supplementary Tables

Supplementary Table 2-1: Cryo-EM and Refinement Statistics of Form I Related Nanotubes

Parameter	**15-10-3**	**29-17-3 (Form I)***	**29-24-3**	**36-31-3**	**36-31-3 RR**
Data collection and processing					
Voltage (kV)	200	300	200	300	200
Electron exposure (e^{-} $Å^{-2}$)	55	50	55	50	55
Pixel size (Å)	1.04	1.02	1.04	1.06	1.04
Final particle images (n)	11,334	62,122	12,869	66,079	39,805
Helical symmetry					
Point group	C5	C1	C4	C1	C1
Helical rise (Å)	9.21	2.20	7.96	2.50	2.51
Helical twist (°)	9.3	-87.1	10.9	124.0	124.5
Map resolution (Å)					
Model:map FSC (0.38)	4.2	3.6	4.1	4.0	4.5
d_{99}	4.3	3.7	4.2	4.0	4.6
Refinement and Model validation					
Map-sharpening B-factor ($Å^2$)	-100	-100	-120	-200	-120
Bond lengths rmsd (Å)	0.006	0.093	0.006	0.007	0.006
Bond angles rmsd (°)	0.557	1.875	0.726	0.797	0.810
Clashscore	2.5	9.2	7.2	2.0	6.3
Poor rotamers (%)	0	0	0	0	0
Ramachandran Favored (%)	100	96.5	99.4	100	100
Ramachandran Outlier (%)	0	0	0.6	0	0
MolProbity score	1.04	1.72	1.40	0.97	1.35
Deposition ID					
PDB (model)	6WKX	3J89	6WKY	6WL1	6WL0

EMDB (map)	EMD-21812	EMD-6123	EMD-21813	EMD-21815	EMD-21814

*Filament structure reported on 2015.

Supplementary Table 2-2: Cryo-EM and Refinement Statistics of Form II Related Nanotubes

Parameter	**Form II**	**Form IIa**	**29-20-2**
Data collection and processing			
Voltage (kV)	300	300	200
Electron exposure (e^{-} Å^{-2})	51	51	55
Pixel size (Å)	1.08	1.08	1.04
Final particle images (n)	408,751	288,990	67,941
Helical symmetry			
Point group	C1	C1	C3
Helical rise (Å)	1.93	1.92	5.63
Helical twist (°)	124.36	124.37	12.90
Map resolution (Å)			
Model:map FSC (0.38)	4.1	4.2	3.8
d_{99}	4.4	4.6	4.1
Refinement and Model validation			
Map-sharpening B-factor (Å^2)	-189	-176	-100
Bond lengths rmsd (Å)	0.006	0.006	0.006
Bond angles rmsd (°)	0.805	0.530	0.559
Clashscore	2.1	14.7	8.5
Poor rotamers (%)	0	0	0
Ramachandran Favored (%)	100	98.2	98.8
Ramachandran Outlier (%)	0	0	0
MolProbity score	0.98	1.67	1.46
Deposition ID			

PDB (model)	6WL8	6WL9	6WL7
EMDB (map)	EMD-21817	EMD-21818	EMD-21816

2.7- References

(1)Papapostolou, D.; Smith, A. M.; Atkins, E. D.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N. Engineering nanoscale order into a designed protein fiber. *Proc Natl Acad Sci U S A* **2007**, *104* (26), 10853-10858. DOI: 10.1073/pnas.0700801104.

(2)Sharp, T. H.; Bruning, M.; Mantell, J.; Sessions, R. B.; Thomson, A. R.; Zaccai, N. R.; Brady, R. L.; Verkade, P.; Woolfson, D. N. Cryo-transmission electron microscopy structure of a gigadalton peptide fiber of de novo design. *Proc Natl Acad Sci U S A* **2012**, *109* (33), 13266-13271. DOI: 10.1073/pnas.1118622109.

(3)Burgess, N. C.; Sharp, T. H.; Thomas, F.; Wood, C. W.; Thomson, A. R.; Zaccai, N. R.; Brady, R. L.; Serpell, L. C.; Woolfson, D. N. Modular Design of Self-Assembling Peptide-Based Nanotubes. *J Am Chem Soc* **2015**, *137* (33), 10554-10562. DOI: 10.1021/jacs.5b03973 From NLM.

(4)Potekhin, S. A.; Melnik, T. N.; Popov, V.; Lanina, N. F.; Vazina, A. A.; Rigler, P.; Verdini, A. S.; Corradin, G.; Kajava, A. V. De novo design of fibrils made of short alpha-helical coiled coil peptides. *Chem Biol* **2001**, *8* (11), 1025-1032. DOI: 10.1016/s1074-5521(01)00073-4 From NLM.

(5)Melnik, T. N.; Villard, V.; Vasiliev, V.; Corradin, G.; Kajava, A. V.; Potekhin, S. A. Shift of fibril-forming ability of the designed alpha-helical coiled-coil peptides into the physiological pH region. *Protein Eng* **2003**, *16* (12), 1125-1130. DOI: 10.1093/protein/gzg138.

(6)Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P. Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif. *Tetrahedron* **2004**, *60* (34), 7237-7246. DOI: https://doi.org/10.1016/j.tet.2004.06.068.

(7)Kajava, A. V.; Potekhin, S. A.; Corradin, G.; Leapman, R. D. Organization of designed nanofibrils assembled from alpha-helical peptides as determined by electron microscopy. *J Pept Sci* **2004**, *10* (5), 291-297. DOI: 10.1002/psc.520.

(8)Walshaw, J.; Woolfson, D. N. Open-and-shut cases in coiled-coil assembly: alpha-sheets and alpha-cylinders. *Protein Sci* **2001**, *10* (3), 668-673. DOI: 10.1110/ps.36901 From NLM Medline.

(9)Egelman, E. H.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.; Baker, D.; Conticello, V. P. Structural plasticity of helical nanotubes based on coiled-coil assemblies. *Structure* **2015**, *23* (2), 280-289. DOI: 10.1016/j.str.2014.12.008.

(10)Ghosh, T. S.; Chaitanya, S. K.; Sankararamakrishnan, R. End-to-end and end-to-middle interhelical interactions: new classes of interacting helix pairs in protein structures. *Acta Crystallogr D Biol Crystallogr* **2009**, *65* (Pt 10), 1032-1041. DOI: 10.1107/S0907444909027012.

(11)Marqusee, S.; Baldwin, R. L. Helix stabilization by Glu-...Lys+ salt bridges in short peptides of de novo design. *Proc Natl Acad Sci U S A* **1987**, *84* (24), 8898-8902. DOI: 10.1073/pnas.84.24.8898 From NLM.

(12)Lyu, P. C.; Gans, P. J.; Kallenbach, N. R. Energetic contribution of solvent-exposed ion pairs to alpha-helix structure. *J Mol Biol* **1992**, *223* (1), 343-350. DOI: 10.1016/0022-2836(92)90735-3.

(13)Huyghues-Despointes, B. M.; Scholtz, J. M.; Baldwin, R. L. Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci* **1993**, *2* (1), 80-85. DOI: 10.1002/pro.5560020108.

(14)Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry* **1993**, *32* (37), 9668-9676. DOI: 10.1021/bi00088a019.

(15)Cid, H.; Gazitua, F.; Bunster, M. Interactions that regulate the helical fold in proteins. *Biol Res* **1996**, *29* (2), 213-225.

(16)Andrew, C. D.; Penel, S.; Jones, G. R.; Doig, A. J. Stabilizing nonpolar/polar side-chain interactions in the alpha-helix. *Proteins* **2001**, *45* (4), 449-455. DOI: 10.1002/prot.1161.

(17)Andrew, C. D.; Bhattacharjee, S.; Kokkoni, N.; Hirst, J. D.; Jones, G. R.; Doig, A. J. Stabilizing interactions between aromatic and basic side chains in alpha-helical peptides and proteins. Tyrosine effects on helix circular dichroism. *J Am Chem Soc* **2002**, *124* (43), 12706-12714. DOI: 10.1021/ja027629h.

(18)Shi, Z.; Olson, C. A.; Kallenbach, N. R. Cation-pi interaction in model alpha-helical peptides. *J Am Chem Soc* **2002**, *124* (13), 3284-3291. DOI: 10.1021/ja0174938.

(19)Butterfield, S. M.; Patel, P. R.; Waters, M. L. Contribution of aromatic interactions to alpha-helix stability. *J Am Chem Soc* **2002**, *124* (33), 9751-9755. DOI: 10.1021/ja026668q.

(20)Meuzelaar, H.; Vreede, J.; Woutersen, S. Influence of Glu/Arg, Asp/Arg, and Glu/Lys Salt Bridges on α-Helical Stability and Folding Kinetics. *Biophysical Journal* **2016**, *110* (11), 2328-2341. DOI: 10.1016/j.bpj.2016.04.015 (acccessed 2020/05/15).

(21)Crick, F. The packing of [alpha]-helices: simple coiled-coils. *Acta Crystallographica* **1953**, *6* (8-9), 689-697. DOI: doi:10.1107/S0365110X53001964.

(22)O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254* (5031), 539-544. DOI: 10.1126/science.1948029.

(23)Alber, T. Structure of the leucine zipper. *Curr Opin Genet Dev* **1992**, *2* (2), 205-210. DOI: 10.1016/s0959-437x(05)80275-8 PubMed.

(24)Russ, W. P.; Engelman, D. M. The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol* **2000**, *296* (3), 911-919. DOI: 10.1006/jmbi.1999.3489.

(25)Dieckmann, G. R.; DeGrado, W. F. Modeling transmembrane helical oligomers. *Curr Opin Struct Biol* **1997**, *7* (4), 486-494. DOI: 10.1016/s0959-440x(97)80111-x.

(26)Walters, R. F.; DeGrado, W. F. Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci U S A* **2006**, *103* (37), 13658-13663. DOI: 10.1073/pnas.0605878103 From NLM.

(27)Zhang, S. Q.; Kulp, D. W.; Schramm, C. A.; Mravic, M.; Samish, I.; DeGrado, W. F. The membrane- and soluble-protein helix-helix interactome: similar geometry via different interactions. *Structure* **2015**, *23* (3), 527-541. DOI: 10.1016/j.str.2015.01.009 From NLM.

(28)Walther, D.; Eisenhaber, F.; Argos, P. Principles of helix-helix packing in proteins: the helical lattice superposition model. *J Mol Biol* **1996**, *255* (3), 536-553. DOI: 10.1006/jmbi.1996.0044.

(29)Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **2007**, *372* (3), 774-797. DOI: 10.1016/j.jmb.2007.05.022 From NLM.

(30)Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **2000**, *85* (4), 225-234. DOI: 10.1016/s0304-3991(00)00062-0 From NLM.

(31)Egelman, E. H. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J Struct Biol* **2007**, *157* (1), 83-94. DOI: 10.1016/j.jsb.2006.05.015 From NLM Medline.

(32)Egelman, E. H. Reconstruction of helical filaments and tubes. *Methods Enzymol* **2010**, *482*, 167-183. DOI: 10.1016/s0076-6879(10)82006-3 From NLM.

(33)He, S.; Scheres, S. H. W. Helical reconstruction in RELION. *Journal of Structural Biology* **2017**, *198* (3), 163-176. DOI: https://doi.org/10.1016/j.jsb.2017.02.003.

(34)Lu, A.; Li, Y.; Yin, Q.; Ruan, J.; Yu, X.; Egelman, E.; Wu, H. Plasticity in PYD assembly revealed by cryo-EM structure of the PYD filament of AIM2. *Cell Discov* **2015**, *1*. DOI: 10.1038/celldisc.2015.13 From NLM.

(35)Shen, H.; Fallas, J. A.; Lynch, E.; Sheffler, W.; Parry, B.; Jannetty, N.; Decarreau, J.; Wagenbach, M.; Vicente, J. J.; Chen, J.; et al. De novo design of self-assembling helical protein filaments. *Science* **2018**, *362* (6415), 705-709. DOI: 10.1126/science.aau3775.

(36)Zheng, W.; Pena, A.; Low, W. W.; Wong, J. L. C.; Frankel, G.; Egelman, E. H. Cryoelectron-Microscopic Structure of the pKpQIL Conjugative Pili from Carbapenem-Resistant Klebsiella pneumoniae. *Structure*. DOI: 10.1016/j.str.2020.08.010 (acccessed 2020/09/21).

(37)Xu, J.; Dayan, N.; Goldbourt, A.; Xiang, Y. Cryo-electron microscopy structure of the filamentous bacteriophage IKe. *Proceedings of the National Academy of Sciences* **2019**, *116* (12), 5493-5498. DOI: 10.1073/pnas.1811929116.

(38)Tarafder, A. K.; von Kügelgen, A.; Mellul, A. J.; Schulze, U.; Aarts, D. G. A. L.; Bharat, T. A. M. Phage liquid crystalline droplets form occlusive sheaths that encapsulate and protect infectious rod-shaped bacteria. *Proceedings of the National Academy of Sciences* **2020**, *117* (9), 4724-4731. DOI: 10.1073/pnas.1917726117.

(39)Costa, T. R. D.; Ilangovan, A.; Ukleja, M.; Redzej, A.; Santini, J. M.; Smith, T. K.; Egelman, E. H.; Waksman, G. Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* **2016**, *166* (6), 1436-1444.e1410. DOI: 10.1016/j.cell.2016.08.025 (acccessed 2020/07/14).

(40)Zhou, J.; Panaitiu, A. E.; Grigoryan, G. A general-purpose protein design framework based on mining sequence-structure relationships in known protein structures. *Proc Natl Acad Sci U S A* **2020**, *117* (2), 1059-1068. DOI: 10.1073/pnas.1908723117 From NLM.

(41)Frappier, V.; Jenson, J. M.; Zhou, J.; Grigoryan, G.; Keating, A. E. Tertiary Structural Motif Sequence Statistics Enable Facile Prediction and Design of Peptides that Bind Anti-apoptotic Bfl-1 and Mcl-1. *Structure* **2019**, *27* (4), 606-617.e605. DOI: 10.1016/j.str.2019.01.008 From NLM.

(42)Tayeb-Fligelman, E.; Tabachnikov, O.; Moshe, A.; Goldshmidt-Tran, O.; Sawaya, M. R.; Coquelle, N.; Colletier, J. P.; Landau, M. The cytotoxic Staphylococcus aureus PSMalpha3 reveals a cross-alpha amyloid-like fibril. *Science* **2017**, *355* (6327), 831-833. DOI: 10.1126/science.aaf4901.

(43)Zhang, S. Q.; Huang, H.; Yang, J.; Kratochvil, H. T.; Lolicato, M.; Liu, Y.; Shu, X.; Liu, L.; DeGrado, W. F. Designed peptides that assemble into cross-alpha amyloid-like structures. *Nat Chem Biol* **2018**, *14* (9), 870-875. DOI: 10.1038/s41589-018-0105-5.

(44)Hughes, S. A.; Wang, F.; Wang, S.; Kreutzberger, M. A. B.; Osinski, T.; Orlova, A.; Wall, J. S.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Ambidextrous helical nanotubes from self-assembly of designed helical hairpin motifs. *Proc Natl Acad Sci U S A* **2019**, *116* (29), 14456-14464. DOI: 10.1073/pnas.1903910116 From NLM Medline.

(45)Kobe, B.; Kajava, A. V. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends in Biochemical Sciences* **2000**, *25* (10), 509-515. DOI: 10.1016/S0968-0004(00)01667-4 (acccessed 2020/05/16).

(46)Kajava, A. V. Tandem repeats in proteins: From sequence to structure. *Journal of Structural Biology* **2012**, *179* (3), 279-288. DOI: https://doi.org/10.1016/j.jsb.2011.08.009.

(47)Lupas, A. N.; Bassler, J. Coiled Coils - A Model System for the 21st Century. *Trends Biochem Sci* **2017**, *42* (2), 130-140. DOI: 10.1016/j.tibs.2016.10.007.

(48)Woolfson, D. N. Coiled-Coil Design: Updated and Upgraded. *Subcell Biochem* **2017**, *82*, 35-61. DOI: 10.1007/978-3-319-49674-0_2 From NLM Medline.

(49)Collie, G. W.; Pulka-Ziach, K.; Lombardo, C. M.; Fremaux, J.; Rosu, F.; Decossas, M.; Mauran, L.; Lambert, O.; Gabelica, V.; Mackereth, C. D.; et al. Shaping quaternary assemblies of water-soluble non-peptide helical foldamers by sequence manipulation. *Nat Chem* **2015**, *7* (11), 871-878. DOI: 10.1038/nchem.2353.

(50)Gallardo, R.; Ranson, N. A.; Radford, S. E. Amyloid structures: much more than just a cross-β fold. *Current Opinion in Structural Biology* **2020**, *60*, 7-16. DOI: https://doi.org/10.1016/j.sbi.2019.09.001.

(51)Zhang, W.; Falcon, B.; Murzin, A. G.; Fan, J.; Crowther, R. A.; Goedert, M.; Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **2019**, *8*. DOI: 10.7554/eLife.43584 From NLM.

(52)Guenther, E. L.; Ge, P.; Trinh, H.; Sawaya, M. R.; Cascio, D.; Boyer, D. R.; Gonen, T.; Zhou, Z. H.; Eisenberg, D. S. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nat Struct Mol Biol* **2018**, *25* (4), 311-319. DOI: 10.1038/s41594-018-0045-5.

(53)Close, W.; Neumann, M.; Schmidt, A.; Hora, M.; Annamalai, K.; Schmidt, M.; Reif, B.; Schmidt, V.; Grigorieff, N.; Fandrich, M. Physical basis of amyloid fibril polymorphism. *Nat Commun* **2018**, *9* (1), 699. DOI: 10.1038/s41467-018-03164-5.

(54)Chen, K. H.; Corro, K. A.; Le, S. P.; Nowick, J. S. X-ray Crystallographic Structure of a Giant Double-Walled Peptide Nanotube Formed by a Macrocyclic beta-Sheet Containing Abeta16-22. *J Am Chem Soc* **2017**, *139* (24), 8102-8105. DOI: 10.1021/jacs.7b03890.

(55)Hervas, R.; Rau, M. J.; Park, Y.; Zhang, W.; Murzin, A. G.; Fitzpatrick, J. A. J.; Scheres, S. H. W.; Si, K. Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in Drosophila. *Science* **2020**, *367* (6483), 1230-1234. DOI: 10.1126/science.aba3526.

(56)Brodin, J. D.; Smith, S. J.; Carr, J. R.; Tezcan, F. A. Designed, Helical Protein Nanotubes with Variable Diameters from a Single Building Block. *J Am Chem Soc* **2015**, *137* (33), 10468-10471. DOI: 10.1021/jacs.5b05755.

(57)Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nat Chem* **2012**, *4* (5), 375-382. DOI: 10.1038/nchem.1290.

(58)Zhang, J.; Zheng, F.; Grigoryan, G. Design and designability of protein-based assemblies. *Current Opinion in Structural Biology* **2014**, *27*, 79-86. DOI: https://doi.org/10.1016/j.sbi.2014.05.009.

(59)Görbitz, C. H. Nanotube Formation by Hydrophobic Dipeptides. *Chemistry – A European Journal* **2001**, *7* (23), 5153-5159. DOI: 10.1002/1521-3765(20011203)7:23<5153::AID-CHEM5153>3.0.CO;2-N (acccessed 2020/05/14).

(60)Reches, M.; Gazit, E. Casting Metal Nanowires Within Discrete Self-Assembled Peptide Nanotubes. *Science* **2003**, *300* (5619), 625-627. DOI: 10.1126/science.1082387.

(61)Lu, K.; Jacob, J.; Thiyagarajan, P.; Conticello, V. P.; Lynn, D. G. Exploiting Amyloid Fibril Lamination for Nanotube Self-Assembly. *Journal of the American Chemical Society* **2003**, *125* (21), 6391-6393. DOI: 10.1021/ja0341642.

(62)Kurt, T. D.; Aguilar-Calvo, P.; Jiang, L.; Rodriguez, J. A.; Alderson, N.; Eisenberg, D. S.; Sigurdson, C. J. Asparagine and glutamine ladders promote cross-species prion conversion. *Journal of Biological Chemistry* **2017**, *292* (46), 19076-19086. DOI: 10.1074/jbc.M117.794107.

(63)Mindell, J. A.; Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology* **2003**, *142* (3), 334-347. DOI: https://doi.org/10.1016/S1047-8477(03)00069-8.

(64)Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **2007**, *157* (1), 38-46. DOI: 10.1016/j.jsb.2006.05.009.

(65)Frank, J.; Radermacher, M.; Penczek, P.; Zhu, J.; Li, Y.; Ladjadj, M.; Leith, A. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **1996**, *116* (1), 190-199. DOI: 10.1006/jsbi.1996.0030.

(66)Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **2012**, *180* (3), 519-530. DOI: https://doi.org/10.1016/j.jsb.2012.09.006.

(67)Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-1612. DOI: 10.1002/jcc.20084.

(68)Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **2010**, *66* (4), 486-501. DOI: doi:10.1107/S0907444910007493.

(69)Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **2007**, *35* (Web Server issue), W375-383. DOI: 10.1093/nar/gkm216.

(70)Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26* (19), 2460-2461. DOI: 10.1093/bioinformatics/btq461 (acccessed 5/15/2020).

(71)Zhou, J.; Grigoryan, G. Rapid search for tertiary fragments reveals protein sequence–structure relationships. *Protein Science* **2015**, *24* (4), 508-524. DOI: 10.1002/pro.2610 (acccessed 2020/05/15).

(72)Zhou, J.; Grigoryan, G. A C++ library for protein sub-structure search. *bioRxiv* **2020**, 2020.2004.2026.062612. DOI: 10.1101/2020.04.26.062612.

(73)Mackenzie, C. O.; Zhou, J.; Grigoryan, G. Tertiary alphabet for the observable protein structural universe. *Proceedings of the National Academy of Sciences* **2016**, *113* (47), E7438-E7447. DOI: 10.1073/pnas.1607178113.

(74)Mackenzie, C. O.; Grigoryan, G. Protein structural motifs in prediction and design. *Current Opinion in Structural Biology* **2017**, *44*, 161-167. DOI: https://doi.org/10.1016/j.sbi.2017.03.012.

(75)Zheng, F.; Grigoryan, G. Sequence statistics of tertiary structural motifs reflect protein stability. *PLoS One* **2017**, *12* (5), e0178272. DOI: 10.1371/journal.pone.0178272.

(76)Tsutsumi, M.; Otaki, J. M. Parallel and Antiparallel β-Strands Differ in Amino Acid Composition and Availability of Short Constituent Sequences. *Journal of Chemical Information and Modeling* **2011**, *51* (6), 1457-1464. DOI: 10.1021/ci200027d.

(77)Pashuck, E. T.; Seeman, N.; Macfarlane, R. Self-assembly of bioinspired and biologically functional materials. *MRS Bulletin* **2020**, *45* (10), 832-840. DOI: 10.1557/mrs.2020.249.

(78)Lassak, K.; Ghosh, A.; Albers, S. V. Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res Microbiol* **2012**, *163* (9-10), 630-644. DOI: 10.1016/j.resmic.2012.10.024 From NLM.

(79)Whitesides, G. M.; Grzybowski, B. Self-assembly at all scales. *Science* **2002**, *295* (5564), 2418-2421. DOI: 10.1126/science.1070821 From NLM.

(80)Krause, M.; Gautreau, A. Steering cell migration: lamellipodium dynamics and the regulation of directional persistence. *Nat Rev Mol Cell Biol* **2014**, *15* (9), 577-590. DOI: 10.1038/nrm3861 From NLM.

(81)Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A. Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc Natl Acad Sci U S A* **1993**, *90* (8), 3334-3338. DOI: 10.1073/pnas.90.8.3334 From NLM.

(82)Ananthakrishnan, R.; Ehrlicher, A. The forces behind cell movement. *Int J Biol Sci* **2007**, *3* (5), 303-317. DOI: 10.7150/ijbs.3.303 From NLM.

(83)Bromley, E. H.; Channon, K.; Moutevelis, E.; Woolfson, D. N. Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. *ACS Chem Biol* **2008**, *3* (1), 38-50. DOI: 10.1021/cb700249v From NLM.

(84)Mendes, A. C.; Baran, E. T.; Reis, R. L.; Azevedo, H. S. Self-assembly in nature: using the principles of nature to create complex nanobiomaterials. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **2013**, *5* (6), 582-612. DOI: 10.1002/wnan.1238 From NLM.

(85)Lazzaro, B. P.; Zasloff, M.; Rolff, J. Antimicrobial peptides: Application informed by evolution. *Science* **2020**, *368* (6490). DOI: 10.1126/science.aau5480 From NLM.

(86)Wei, G.; Su, Z.; Reynolds, N. P.; Arosio, P.; Hamley, I. W.; Gazit, E.; Mezzenga, R. Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. *Chem Soc Rev* **2017**, *46* (15), 4661-4708. DOI: 10.1039/c6cs00542j From NLM.

(87)Tang, A.; Wang, C.; Stewart, R.; Kopecek, J. Self-assembled peptides exposing epitopes recognizable by human lymphoma cells. *Bioconjug Chem* **2000**, *11* (3), 363-371. DOI: 10.1021/bc990133d From NLM.

(88)Tao, K.; Makam, P.; Aizen, R.; Gazit, E. Self-assembling peptide semiconductors. *Science* **2017**, *358* (6365). DOI: 10.1126/science.aam9756 From NLM.

(89)Ryan, K.; Beirne, J.; Redmond, G.; Kilpatrick, J. I.; Guyonnet, J.; Buchete, N. V.; Kholkin, A. L.; Rodriguez, B. J. Nanoscale Piezoelectric Properties of Self-Assembled Fmoc-FF Peptide Fibrous Networks. *ACS Appl Mater Interfaces* **2015**, *7* (23), 12702-12707. DOI: 10.1021/acsami.5b01251 From NLM.

(90)Nilsson, K. P.; Rydberg, J.; Baltzer, L.; Inganäs, O. Self-assembly of synthetic peptides control conformation and optical properties of a zwitterionic polythiophene derivative. *Proc Natl Acad Sci U S A* **2003**, *100* (18), 10170-10174. DOI: 10.1073/pnas.1834422100 From NLM.

(91)Gao, X.; Matsui, H. Peptide-Based Nanotubes and Their Applications in Bionanotechnology. *Adv Mater* **2005**, *17* (17), 2037-2050. DOI: 10.1002/adma.200401849 From NLM.

(92)Lampel, A.; McPhee, S. A.; Kassem, S.; Sementa, D.; Massarano, T.; Aramini, J. M.; He, Y.; Ulijn, R. V. Melanin-Inspired Chromophoric Microparticles Composed of Polymeric Peptide Pigments. *Angew Chem Int Ed Engl* **2021**, *60* (14), 7564-7569. DOI: 10.1002/anie.202015170 From NLM.

(93)Guilbaud, J. B.; Saiani, A. Using small angle scattering (SAS) to structurally characterise peptide and protein self-assembled materials. *Chem Soc Rev* **2011**, *40* (3), 1200-1210. DOI: 10.1039/c0cs00105h From NLM.

(94)Zhang, S.; Holmes, T. C.; DiPersio, C. M.; Hynes, R. O.; Su, X.; Rich, A. Self-complementary oligopeptide matrices support mammalian cell attachment. *Biomaterials* **1995**, *16* (18), 1385-1393. DOI: 10.1016/0142-9612(95)96874-y From NLM.

(95)Wang, F.; Gnewou, O.; Solemanifar, A.; Conticello, V. P.; Egelman, E. H. Cryo-EM of Helical Polymers. *Chem Rev* **2022**. DOI: 10.1021/acs.chemrev.1c00753 From NLM.

(96)Kocabey, S.; Ceylan, H.; Tekinay, A. B.; Guler, M. O. Glycosaminoglycan mimetic peptide nanofibers promote mineralization by osteogenic cells. *Acta Biomater* **2013**, *9* (11), 9075-9085. DOI: 10.1016/j.actbio.2013.07.007 From NLM.

(97)Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; et al. Structure of the toxic core of α-synuclein from invisible crystals. *Nature* **2015**, *525* (7570), 486-490. DOI: 10.1038/nature15368 From NLM.

(98)Wlodawer, A.; Dauter, Z. `Atomic resolution': a badly abused term in structural biology. *Acta Crystallogr D Struct Biol* **2017**, *73* (Pt 4), 379-380. DOI: 10.1107/s205979831700225x From NLM.

(99)Colvin, M. T.; Silvers, R.; Ni, Q. Z.; Can, T. V.; Sergeyev, I.; Rosay, M.; Donovan, K. J.; Michael, B.; Wall, J.; Linse, S.; et al. Atomic Resolution Structure of Monomorphic Aβ42 Amyloid Fibrils. *J Am Chem Soc* **2016**, *138* (30), 9663-9674. DOI: 10.1021/jacs.6b05129 From NLM.

(100)Dubochet, J. Cryo-EM--the first thirty years. *J Microsc* **2012**, *245* (3), 221-224. DOI: 10.1111/j.1365-2818.2011.03569.x From NLM.

(101)Spaulding, C. N.; Schreiber, H. L. t.; Zheng, W.; Dodson, K. W.; Hazen, J. E.; Conover, M. S.; Wang, F.; Svenmarker, P.; Luna-Rico, A.; Francetic, O.; et al. Functional role of the type 1 pilus rod structure in mediating host-pathogen interactions. *Elife* **2018**, *7*. DOI: 10.7554/eLife.31662 From NLM.

(102)Egelman, E. H. Cryo-EM of bacterial pili and archaeal flagellar filaments. *Curr Opin Struct Biol* **2017**, *46*, 31-37. DOI: 10.1016/j.sbi.2017.05.012 From NLM.

(103)Wyckoff, R. W.; Corey, R. B. THE ULTRACENTRIFUGAL CRYSTALLIZATION OF TOBACCO MOSAIC VIRUS PROTEIN. *Science* **1936**, *84* (2188), 513. DOI: 10.1126/science.84.2188.513 From NLM.

(104)Harris, J. R. Transmission electron microscopy in molecular structural biology: A historical survey. *Arch Biochem Biophys* **2015**, *581*, 3-18. DOI: 10.1016/j.abb.2014.11.011 From NLM.

(105)Butler, P. J. Self-assembly of tobacco mosaic virus: the role of an intermediate aggregate in generating both specificity and speed. *Philos Trans R Soc Lond B Biol Sci* **1999**, *354* (1383), 537-550. DOI: 10.1098/rstb.1999.0405 From NLM.

(106)Freeman, R.; Leonard, K. R. Comparative mass measurement of biological macromolecules by scanning transmission electron microscopy. *J Microsc* **1981**, *122* (Pt 3), 275-286. DOI: 10.1111/j.1365-2818.1981.tb01267.x From NLM.

(107)Werneburg, G. T.; Thanassi, D. G. Pili Assembled by the Chaperone/Usher Pathway in Escherichia coli and Salmonella. *EcoSal Plus* **2018**, *8* (1). DOI: 10.1128/ecosalplus.ESP-0007-2017 From NLM.

(108)Pushparaj, V. L.; Shaijumon, M. M.; Kumar, A.; Murugesan, S.; Ci, L.; Vajtai, R.; Linhardt, R. J.; Nalamasu, O.; Ajayan, P. M. Flexible energy storage devices based on nanocomposite paper. *Proc Natl Acad Sci U S A* **2007**, *104* (34), 13574-13577. DOI: 10.1073/pnas.0706508104 From NLM.

(109)Hospenthal, M. K.; Redzej, A.; Dodson, K.; Ukleja, M.; Frenz, B.; Rodrigues, C.; Hultgren, S. J.; DiMaio, F.; Egelman, E. H.; Waksman, G. Structure of a Chaperone-Usher Pilus Reveals the Molecular Basis of Rod Uncoiling. *Cell* **2016**, *164* (1-2), 269-278. DOI: 10.1016/j.cell.2015.11.049 From NLM.

(110)Doig, A. J. Stability and design of alpha-helical peptides. *Prog Mol Biol Transl Sci* **2008**, *83*, 1-52. DOI: 10.1016/s0079-6603(08)00601-6 From NLM.

(111)Zheng, W.; Wang, F.; Taylor, N. M. I.; Guerrero-Ferreira, R. C.; Leiman, P. G.; Egelman, E. H. Refined Cryo-EM Structure of the T4 Tail Tube: Exploring the Lowest Dose Limit. *Structure* **2017**, *25* (9), 1436-1441.e1432. DOI: 10.1016/j.str.2017.06.017 From NLM.

(112)Lupas, A. N.; Bassler, J.; Dunin-Horkawicz, S. The Structure and Topology of α-Helical Coiled Coils. *Subcell Biochem* **2017**, *82*, 95-129. DOI: 10.1007/978-3-319-49674-0_4 From NLM.

(113)Busiek, K. K.; Margolin, W. Bacterial actin and tubulin homologs in cell growth and division. *Curr Biol* **2015**, *25* (6), R243-r254. DOI: 10.1016/j.cub.2015.01.030 From NLM.

(114)Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry* **2000**, *39* (30), 8728-8734. DOI: 10.1021/bi000246g From NLM.

(115)Crisma, M.; Formaggio, F.; Moretto, A.; Toniolo, C. Peptide helices based on alpha-amino acids. *Biopolymers* **2006**, *84* (1), 3-12. DOI: 10.1002/bip.20357 From NLM.

(116)Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G. Computational design of a protein crystal. *Proc Natl Acad Sci U S A* **2012**, *109* (19), 7304-7309. DOI: 10.1073/pnas.1112595109 From NLM.

(117)Walshaw, J.; Woolfson, D. N. Extended knobs-into-holes packing in classical and complex coiled-coil assemblies. *J Struct Biol* **2003**, *144* (3), 349-361. DOI: 10.1016/j.jsb.2003.10.014 From NLM.

(118)Xu, C.; Liu, R.; Mehta, A. K.; Guerrero-Ferreira, R. C.; Wright, E. R.; Dunin-Horkawicz, S.; Morris, K.; Serpell, L. C.; Zuo, X.; Wall, J. S.; et al. Rational design of helical nanotubes from self-assembly of coiled-coil lock washers. *J Am Chem Soc* **2013**, *135* (41), 15565-15578. DOI: 10.1021/ja4074529 From NLM.

(119)Cheung, G. Y.; Kretschmer, D.; Queck, S. Y.; Joo, H. S.; Wang, R.; Duong, A. C.; Nguyen, T. H.; Bach, T. H.; Porter, A. R.; DeLeo, F. R.; et al. Insight into structure-function relationship in phenol-soluble modulins using an alanine screen of the phenol-soluble modulin (PSM) α3 peptide. *Faseb j* **2014**, *28* (1), 153-161. DOI: 10.1096/fj.13-232041 From NLM.

(120)Zhang, S. Q.; Huang, H.; Yang, J.; Kratochvil, H. T.; Lolicato, M.; Liu, Y.; Shu, X.; Liu, L.; DeGrado, W. F. Designed peptides that assemble into cross-α amyloid-like structures. *Nat Chem Biol* **2018**, *14* (9), 870-875. DOI: 10.1038/s41589-018-0105-5 From NLM.

(121)Laabei, M.; Jamieson, W. D.; Yang, Y.; van den Elsen, J.; Jenkins, A. T. Investigating the lytic activity and structural properties of Staphylococcus aureus phenol soluble modulin (PSM) peptide toxins. *Biochim Biophys Acta* **2014**, *1838* (12), 3153-3161. DOI: 10.1016/j.bbamem.2014.08.026 From NLM.

(122)Kreutzberger, M. A. B.; Wang, S.; Beltran, L. C.; Tuachi, A.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Phenol-soluble modulins PSMα3 and PSMβ2 form nanotubes that are cross-α amyloids. *Proc Natl Acad Sci U S A* **2022**, *119* (20), e2121586119. DOI: 10.1073/pnas.2121586119 From NLM.

(123)Kreutzberger, M. A. B.; Wang, S.; Beltran, L. C.; Tuachi, A.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Phenol-soluble modulins PSMalpha3 and PSMbeta2 form nanotubes that are cross-alpha amyloids. *Proc Natl Acad Sci U S A* **2022**, *119* (20), e2121586119. DOI: 10.1073/pnas.2121586119 From NLM Medline.

Chapter 3: Self-assembly of a β-sheet nanotube from an amphipathic oligopeptide

This chapter has been adapted with permission from Wang F, Gnewou O, Wang S, Osinski T, Zuo X, Egelman EH, Conticello VP. Deterministic chaos in the self-assembly of β sheet nanotubes from an amphipathic oligopeptide. Matter. 2021 Oct 6;4(10):3217-3231. doi: 10.1016/j.matt.2021.06.037. PMID: 34632372; PMCID: PMC8494133. **Copyright © 2022 Matter, Elsevier** (https://www.sciencedirect.com/science/article/pii/S2590238521003076)

3.1- Introduction

Synthetic peptide filaments have been studied for nearly three decades.[1-3] Despite extensive research effort, the design of self-assembling peptides remains largely empirical and is primarily drawn from sequence-structure correlations established from analysis of soluble proteins and discrete oligomers. β-sheet filaments were among the first synthetic peptide assemblies to be reported.[4-6] The peptide designs were based on an alternating non-polar/polar residue pattern that biased the sequences towards the adoption of an amphipathic β-strand conformation.[1] Self-association of β-strands through hydrogen bonding interactions resulted in the formation of cross-β filaments.[7-9] In the case of amphipathic β-strand peptides, this process was accompanied by pairwise interaction of β-sheets in which the hydrophobic residues were buried at the interface.[10] Thus far, high-resolution structural analysis of designed β-sheet assemblies has been limited to solid-state NMR methods,[11-13,] which cannot easily determine long-range structure. These supramolecular structural parameters exert a significant influence on materials properties, such as liquid crystallinity, gelation, and mechanical response.[14, 15] In addition, long-range structure influences the density, symmetry, and spatial availability of bioactive functional groups along the contour length of a filament, and, consequently, the ability of the material to predictably and productively engage in cellular interactions at bio interfaces. In the case of cross-β filaments, these considerations become particularly acute as structural analyses of amyloidogenic peptide assemblies[16] have demonstrated that structural polymorphism is often observed and depends critically on the assembly conditions.[17-21]

In order to understand the structural principles that underlie the design of synthetic β-sheet assemblies, we focused on an analysis of the self-complementary, amphipathic octapeptide sequence, Ac-FKFEFKFE-NH_2 (KFE8).[22, 23] The peptide design was based on an electrostatically self-complementary amphipathic sequence that was proposed to favor the formation of antiparallel β-sheet. Facile self-assembly occurred in aqueous solution, which resulted in formation of filamentous structures. A structural model was proposed for the KFE8 assemblies in which two cross-β filaments self-associate into left-handed helical ribbons, in which the helical hand was confirmed from AFM measurements. However, this model was insufficient to rationalize the observed curvature of the ribbon since it was based on symmetric face-to-face packing of antiparallel β-sheet. Assemblies of KFE8 and related sequences have been employed as synthetic biomaterials in tissue-engineering and immunotherapy.[24, 25] For these applications, detailed knowledge of the molecular structure of the assemblies is a critical consideration for the development of effective strategies to control interactions at the cell-biomaterial interface. We report here the cryo-EM structural analysis of KFE8 assembled under two different sets of conditions and identify the presence not only of helical ribbons but also of different populations of helical nanotubes. These studies demonstrate the chaotic nature of peptide-self-assembly and its potential impact on biomaterials applications. Peptide-based assemblies have been proposed as designer biomaterials, in which control of sequence affords more reliable control of structure and function. However, recent high-resolution structural analyses of peptide and protein filaments have demonstrated that reliable prediction of supramolecular structure from sequence information remains a significant challenge to de novo design efforts.[26-31] In addition, our results indicate that preparative conditions exert a significant influence on supramolecular structure and can direct the self-assembly process down different folding landscapes.

3.2- Result

3.2.1- Different morphology of KFE8 nanotubes under cryo-EM

To determine the atomic structure of KFE8 assemblies, we imaged the peptide sample under two different conditions: the "ambient" sample was incubated at room temperature, while the "annealed" sample was heated to 90°C and slowly cooled to 25°C. In the ambient sample, consistent with other reports,[22, 32] abundant ribbon-like filaments were seen after 18 hours of

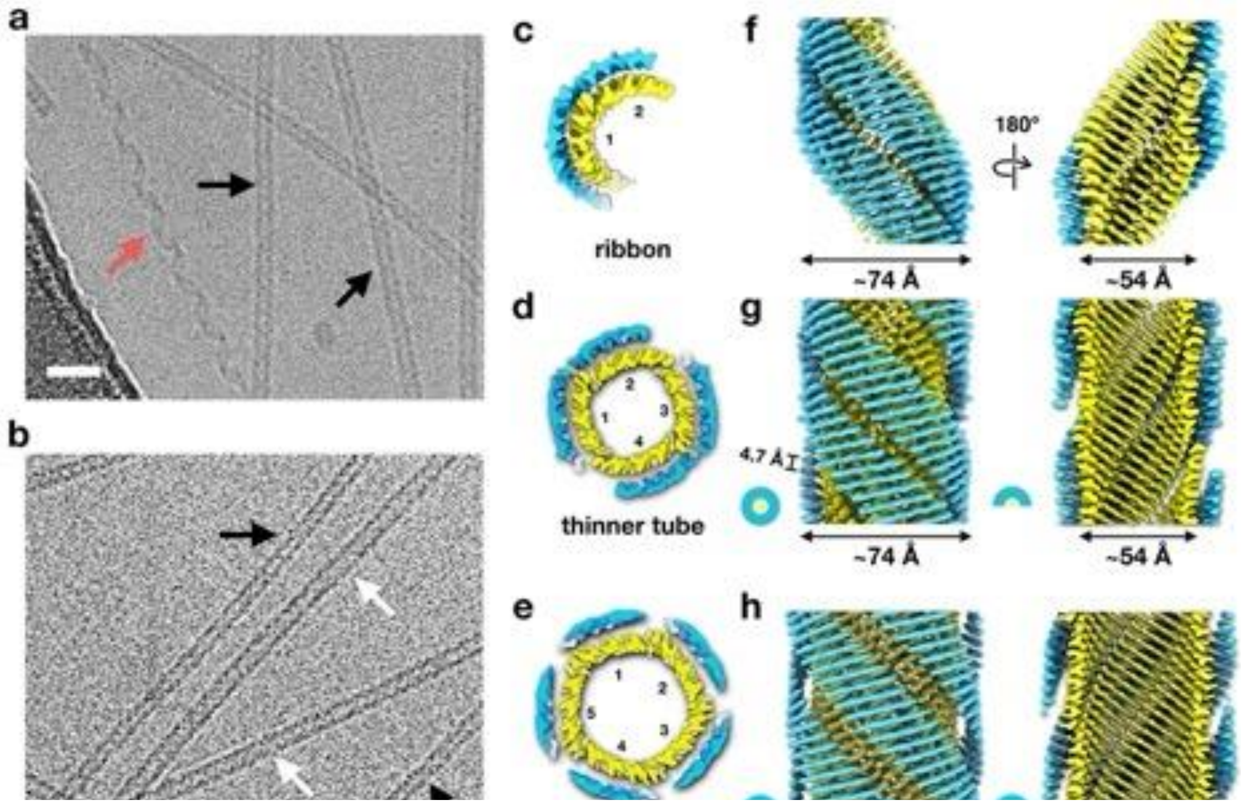


Figure 3-1: Cryo-EM of KFE8

a) Cryo-electron micrograph of the KFE8 tubes assembled at ambient temperature. Scale bar, 20nm. Black arrow points to the thinner tube and red arrow points to the ribbon. **b)** Cryo-electron micrograph of the annealed KFE8 tubes. The peptide was heated immediately after solubilization to 95 °C and slowly cooled to 25 °C. Black arrows point to the thinner tubes and white arrows point to the thicker tubes. **c-e**) Top view of the ribbon, thinner tube and thicker tube. **f-h**) Side views of the 3D reconstructions of the ribbon (f), thinner tube (g) and the thicker tube (h). The outer surface is shown on the left for each of these, while a cut-away view revealing the surface of the lumen is shown on the right incubation (**Figure 3-3-1** and **Figure 3-3-3**). With longer incubations, the ribbons gradually disappeared, and the presence of nanotubes was observed. After two weeks, cryo-EM imaging indicated that ~98% of the filaments were 7.5 nm diameter nanotubes and ~2% were ribbons (**Figure 3-1-a**). In the annealed sample, initial cryo-EM imaging indicated a stable, mixed population of nanotubes in which ~30% were 7.5 nm in diameter, as seen in the ambient sample, while ~70% were 8.5 nm in diamete**r (Figure 3-1-b**). The corresponding morphological changes correlated with conformational transitions observed in the circular dichroism (CD) spectra (**Figure 3-3-2**).

Three-dimensional reconstruction from two-dimensional cryo-EM images has now reached the stage in structural biology where near-atomic resolution is almost routinely achieved.[33, 34] CryoEM has emerged more recently as a powerful tool in understanding peptide nanotube architectures.[26, 28, 35-37] We have taken advantage of the helical symmetry in these tubes to greatly improve the signal-to-noise ratio by helical averaging.[38] The cryo-EM reconstructions clearly show that both the thinner tube (7.5 nm) and the thicker tube (8.5 nm) are double-walled structures (**Figure 3-1-c-e**), in which thinner and thicker tubes consist of four and five cross-β sheet sandwiches, respectively. The ribbon, having two cross-β sheet sandwiches, is identical to half of the thinner tube **(Figure 3-1-h**, **Figure 3-3-4**). In all structures the KFE8 peptides pack tightly in the inner wall, with 4-fold and 5-fold rotational point group symmetry in the thinner and thicker tubes, respectively. Interestingly, the four outer sheets in the thinner tube dimerize into two protofilaments, such that the 4-fold symmetry is broken in the outer wall (**Figure 3-1-d**). As discussed below, the 2-fold symmetry can be broken further through a registry shift between protofilaments. In contrast to the inner wall, the outer sheets in the thicker tube do not make contacts with each other **(Figure 3-1-h**).

3.2.2-The doubled-wall nanotubes are made of both parallel and anti-parallel sheets.

Initial reconstructions of the thinner tube volume showed well defined amino acid side chains and the typical zig-zag pattern for the Cα trace expected in a β-sheet. In contrast, such patterns were not seen in the outer wall (**Figure 3-3-5**). Based on the ~3.6 Å resolution map of the inner wall, and prior knowledge that this nanotube has a left-handed protofilament seen by AFM imaging,[22, 23] we were able to build a reliable atomic model of parallel sheets in the inner wall, in contrast to the antiparallel arrangement that was previously predicted by molecular dynamics

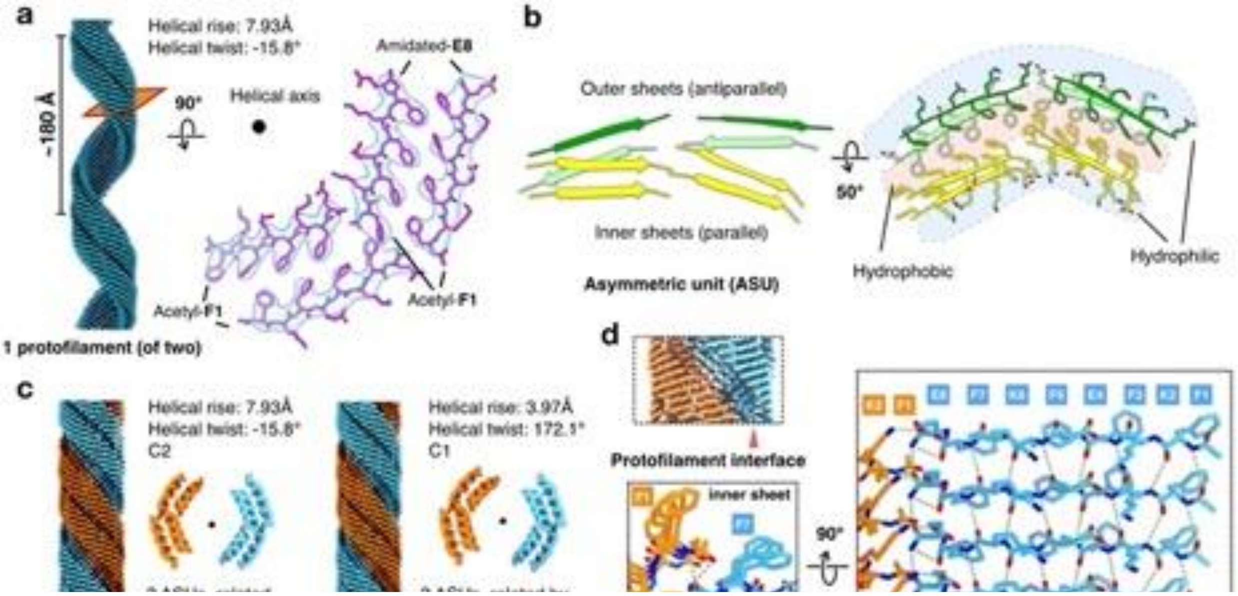


Figure 3-2: **The thinner tube is made of two ribbons**

a) The cryo-EM structure of one of two protofilaments that form the thinner tube. The surface of this protofilament from the helical reconstruction is on the left. The high-resolution crosssectional density map is on the right, with the atomic model built into the map shown.

b) The asymmetric unit (ASU) of the thinner tube contains eight strands: the parallel inner strands are in yellow, while the anti-parallel outer strands are in green and light green. The hydrophobic interactions within the ASU involve the interdigitated phenylalanine residues.

c) The thinner tubes can be sorted into two classes, where the two protofilaments are related by a C_2 symmetry (class 1) or a screw symmetry (class 2).

d) The protofilament interface exists in the inner wall only. The hydrogen bonds are highlighted between protofilaments, and within parallel sheets.

simulations.[22, 23] Strikingly, when we tried to build parallel sheets into the outer wall, strong clashes existed between the O-atoms such that the typical β-sheet hydrogen bonds could not form (**Figure 3-3-6**). Therefore, we built antiparallel sheets into the outer wall, which resulted in the absence of clashes and formation of reasonable inter-strand hydrogen bonds.

To validate this hypothesis of inner parallel and outer antiparallel sheets, a single protofilament refinement was performed for the thinner tube (**Figure 3-2-a**). The reconstruction reached high resolution for both the inner and outer wall, and clearly revealed both an inner parallel sheet and an outer antiparallel sheet (**Figure 3-2-a-b**). We were able to visualize a nearly identical arrangement for an independent reconstruction of the ribbon filaments (**Figure 3-3-4)**The asymmetric unit (ASU) of the thinner tube and the ribbon consists of eight β-strands: four inner strands in two parallel sheets and four outer strands in two antiparallel sheets. The phenylalanine residues create a hydrophobic interface that stabilizes the interaction between the inner and outer walls. In contrast, hydrophilic residues line the inner lumen and cover the outer surface (Error! Reference source not found.**-b**). We used prior knowledge about the structure of the ribbon to determine that the thinner tube exists in either of two states in which the two ribbons are related by either a C_2 symmetry or a 2_1-like screw symmetry (**Figure 3-2-c**). In the C_2 symmetry the two protofilaments are in register. In the screw symmetry they are shifted by a strand, such that the polarity of the strands is reversed through rotation of one protofilament into the other. While no contacts are observed between the two protofilaments in the outer wall, the contacts between protofilaments in the inner wall are only made via a few hydrogen bonds (**Figure 3-2-d**). Despite the unpredicted antiparallel on parallel packing of β-sheets at the outer-inner interface, structural homology searches of the PDB indicated that this arrangement occurred at a similar frequency within native protein structures when accounting for the statistical preference for antiparallel versus parallel sheet (**Figure 3-3-7**).[39]

Similar results were observed in the structural analysis of the thicker tube. The structure of the inner wall was determined at high resolution (~3.6 Å) and found to be based on parallel β-sheet. Due to difficulties in signal subtraction and the absence of the corresponding ribbon filaments, we could not obtain a high-resolution structure for the outer wall (**Figure 3-3-a**). However, using a similar analysis of steric clash as employed for the thinner tube, we could demonstrate that parallel sheet is impossible for the outer wall (**Figure 3-3-6**). The ASU in the thicker tube consists of four strands: two in the inner parallel sheet and two in the outer antiparallel sheet, with similar interfaces seen in the thinner tube (**Figure 3-3-b**). As with the thinner tubes, the

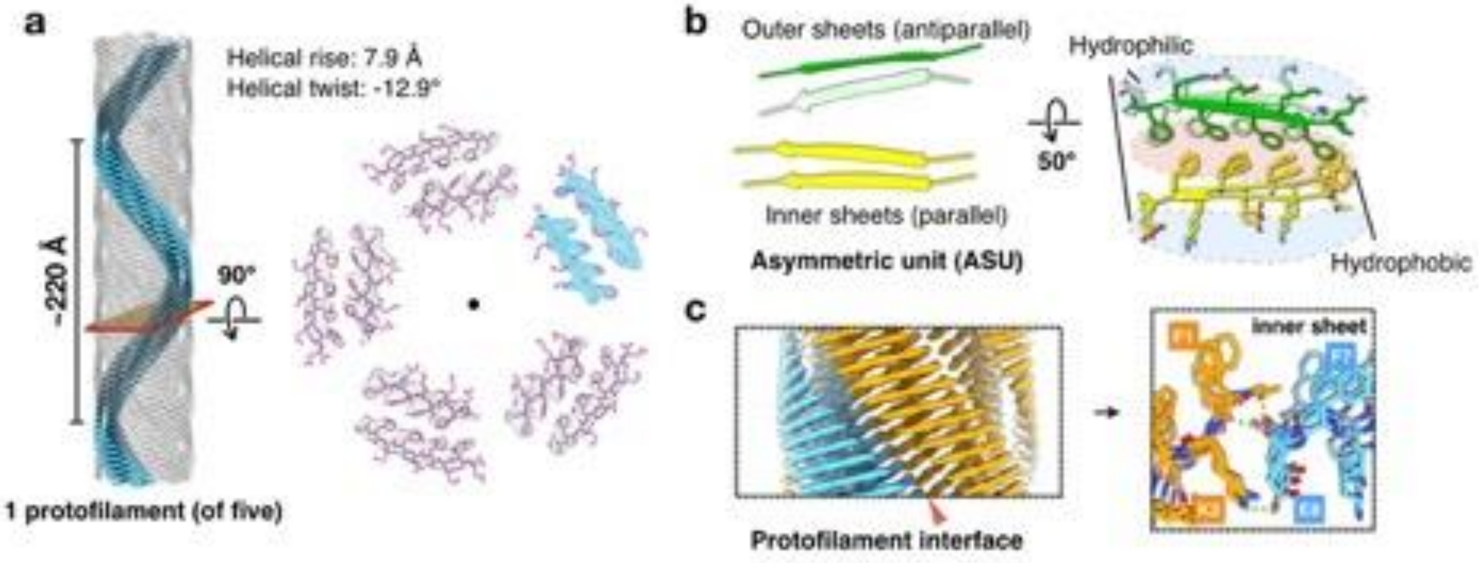


Figure 3-3: The thicker tube forms after annealing

a) The thicker tube cryo-EM structure. The 5-fold averaged helical reconstruction is on the left, with a single protofilament shown in cyan. The cross-sectional density map is on the right, with an atomic model built into the map.

b) The asymmetric unit (ASU) of the thicker tube. The hydrophobic interactions within the ASU arise from the interdigitated phenylalanine residues.

c) The protofilament interface exists only in the inner wall. The hydrogen bonds between protofilaments are shown as dotted lines.

contacts between protofilaments are mediated exclusively through hydrogen-bonding interactions between strands in the inner wall (**Figure 3-3-c**).

3.2.3- Different helical packing from the same 8-residue peptide

A structural comparison of the thinner and thicker tubes revealed subtle differences in protomer packing arrangements. The tilt angle of the peptides in the thinner tube, defined with respect to the plane normal to the filament axis, was observed to be 31° and 14° for the inner and outer sheets, respectively. The corresponding tilt angles in the thicker tube are 30° and 13°. The difference in tilt angle between the inner and outer sheet was conserved at 17° for both tubes (**Figure 3-4-a**). A structural comparison of two β-strand peptides in the inner wall indicated that packing arrangement was nearly identical between the thinner and the thicker tubes (RMSD of only 0.2 Å). However, when comparing ten peptide strands, the RMSD increases 4.5 times to 0.9 Å, which suggests that a slight difference in twist angle exists between the adjacent strands in the parallel β-sheet between the thinner (-7.9° per strand) and thicker (-6.4° per strand) tubes (**Figure 3-4-b**).

Next, we investigated the protofilament interfaces between the two inner sheets. The hydrogen bonds are quite similar between the thinner and thicker tubes (**Figure 3-2-d, Figure 3-3-c**). When we aligned one sheet of the two, we observed a ~7° rotation in the other sheet. As with the difference in twist within the inner sheets (-7.9° vs -6.4° per strand), this shows that the protofilament interface has a plasticity for such contacts. A similar phenomenon has been reported in our recent study of a-helical peptide assemblies.[37] The difference in β-sheet orientation for the inner (parallel) and outer (antiparallel) walls of the nanotubes leads to different interfacial packing arrangements between the inner and outer β-sheets. The thinner tube displays four distinct interfacial arrangements. In contrast, the thicker tube consists of two interfaces; each of which are

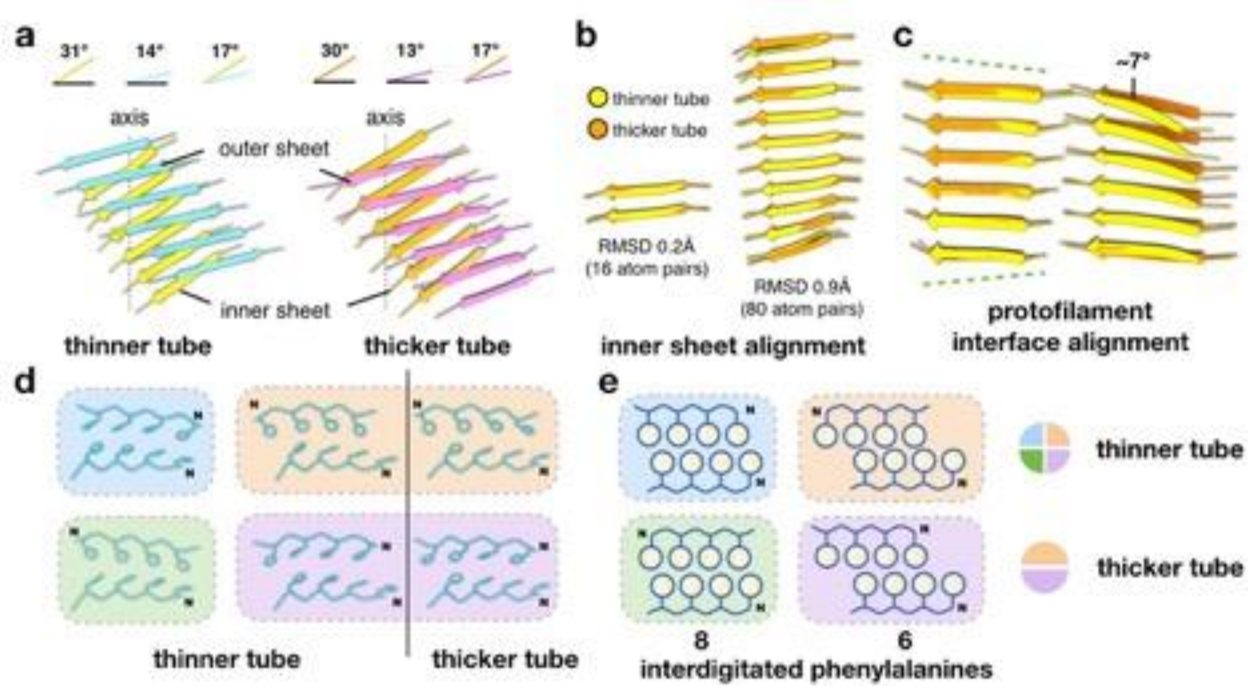


Figure 3-4: The different helical packings of the same 8-residue peptide

a) One stack containing one sheet of the inner wall and one sheet of the outer wall (thinner tube on the left and thicker tube on the right). The tilt angles of the sheet from the plane normal to the helical axis are indicated on top.

b) Structural alignment of a single inner β-sheet containing 2 strands (left) and 10 strands (right).

c) Structural comparison of two adjacent inner β-sheets containing 6 strands. The alignment was made on the left sheets, resulting in the twist differences on the right. **d-e**) The inner-outer strand pairs in the thinner and thicker tube. A cartoon illustration is shown in (e). The N-terminus of peptide is labeled.

similar to two of the interfaces in the thinner tube (**Figure 3-4-d-e**). These interfaces comprise interdigitated phenylalanine residues, which are arranged similarly to the pairwise β-zipper packing of strands in amyloid fibers.[40, 41]

3.3- Discussion

We have shown that plasticity in the structure of the β-sheets formed by the KFE8 peptide, and plasticity in the interfaces between these sheets, leads to a polymorphism in the assembly of ribbons and nanotubes formed by this peptide. A plasticity in the packing of proteins was invoked many years ago in the “theory of quasi-equivalence” to explain how identical copies of proteins might be situated in non-identical environments to form icosahedral viral capsids.[42] This plasticity can be seen in the polymerization of tubulin in vitro, where microtubules are formed with between eight to 17 protofilaments, while in vivo this number is almost always 13.[43] Similarly, amyloidogenic peptides have been demonstrated to polymerize into structurally distinct filaments when assembled in vitro versus in vivo.[19-21] In addition, the amyloidogenic peptide sequence Ab(16-22)[44] has been observed to assemble in vitro either nanotubes or ribbons in a manner that depends solely on concentration.[45] Therefore, even subtle variations in assembly conditions may result in critical differences in supramolecular structure that have significant consequences for functional applications. Analytical models have been developed that can account in approximate terms for differences in observed self-assembly behavior for synthetic b-sheet peptide peptides in terms of the free energies of the respective assemblies.[10, 46, 47] However, our results suggest that we do not fully understand at present the structural complexity of the folding landscape, even for assemblies based on relatively simple peptide sequences.

The KFE8 peptide was designed to form a synthetic cross-β fibril assembly in which the alternating hydrophilic-hydrophobic sequence reinforced the dyad fiber repeat of the β-sheet.[1] Polar sequence patterning has been employed frequently as a design principle to bias peptide conformation and direct protein self-association.[48-51] In FKE8, the pattern of polar, charged residues was chosen to reinforce an antiparallel orientation through introduction of a network of complementary electrostatic interactions between adjacent β-strands within the presumed cross-β filament that would arise from peptide self-assembly.[22, 23] Structural analyses of β-sheets in native proteins indicated a preference for the antiparallel orientation,[52] particularly for sequences that comprise alternating patterns of polar and non-polar residues.[53] Interestingly, it was observed some time ago that antiparallel sheets tend to be found on the outside of proteins and parallel ones in the interior, and this was attributed to the greater conformational flexibility of the antiparallel sheet.[54] In both the thinner and thicker tubes, the antiparallel sheets are located on the outside, while the inner layer of the tube consists solely of parallel β-sheet. In the latter case, close contact occurs between parallel strands in the inner sheet despite the fact that the electrostatic interactions should be repulsive at the strand-strand interfaces. However, the respective polar surfaces of the bilayer sheets should be highly hydrated and consequently, charge-charge interactions may not play a determinative role in selection of strand orientation. A similar situation has been observed for nanotube formation from surfactant-like synthetic peptide sequences, for which theoretically high surface charge densities do not abrogate selfassembly.[55-62] While it has been known that the sequence of a peptide or protein will influence whether a parallel or antiparallel β-sheet is formed,[52, 53, 63] the same KFE8 peptide sequence readily forms both parallel sheets and antiparallel sheets within the same protofilaments (Figures 2 and 3). Statistical analyses of the PDB (Figure S7) indicate that the antiparallel on parallel sheet packing within the protofilament has clear precedent among native protein structures, however the preference for this mode of sheet lamination, vis-à-vis either the parallel-parallel or antiparallel packing arrangements, remains unclear at this point. However, this packing arrangement does account for the observed curvature of the bilayer ribbons that has been reproducibly observed in the literature for KFE8 assemblies.[22, 23, 32]

3.4- Conclusion

` We hypothesize that the self-assembly of the KFE8 ribbons and nanotubes is an instance of chaotic behavior, where small differences in nucleation events lead to quite different assemblies being formed. While the process of self-assembly is deterministic (as opposed to something that is truly stochastic such as quantum uncertainty), this does not make them predictable. What are the implications of this for the design of peptide assemblies? Clearly, computational and theoretical approaches to understand peptide assemblies[23] are unable at this point to yield correct structures due to the complexity of the problem, inadequate force fields, and limited computational power. Since these assemblies obey deterministic laws of chemistry and physics, we do not suggest that this will always be the case. But experimental techniques, especially cryo EM, will have an extremely important role to play until we can be in a position to accurately simulate such assemblies in silico. Peptide assemblies, such as those derived from KFE8, are being actively investigated as medical biomaterials. Our results suggest that the structures of these assemblies depend acutely on preparative conditions. Structural variations in the assemblies may impact interactions at cell-biomaterial interfaces in a manner that is difficult to predict but critical for success in biomedical applications. In addition, as has been previously demonstrated for other peptide assemblies,[26, 37] sequence-structure correlations derived from the analysis of soluble, globular proteins may provide limited predictive insight into the reliable design of filamentous peptide nanomaterials. Therefore, high-resolution structural analyses of native and synthetic assemblies are essential to building a structural database that can be employed as a more reliable resource of information to guide the sequence design of self-assembling peptides.

3.5- Material and methods

3.5.1- Peptide synthesis and assembly

KFE8 peptide, FKFEFKFE, was obtained from GenScript USA Inc. (Piscataway, NJ) as the N-acetyl, C-amide derivative at ≥ 95 % purity by analytical HPLC. Stock solutions of the KFE8 peptide (3 mg·mL^{-1}) were prepared by solubilizing 0.75 mg of purified, lyophilized peptide in 250 μL of sodium acetate buffer (10 mM, pH 4.0) via sonication for 4-5 minutes.[1] The solutions were allowed to either incubate at ambient temperature or were subjected to the following annealing protocol in a thermal cycler: (1) rapid heating to 90° C for 30 minutes and (2) cooling to 25°C at a rate of 0.2°C/minute. The ambient sample consisted primarily of ribbons at short incubation times ($\leq$ 18 hours). The population of nanotubes slowly increased upon incubation at ambient temperature, eventually becoming the major population over a period of 2 weeks. Few ribbons were observed in the annealed sample and the population of nanotubes did not appear to change upon extended incubation at ambient temperature.

3.5.2- Circular dichroism spectropolarimetry

CD measurements were performed on a Jasco J-1500 CD spectropolarimeter using 0.20 mm thick quartz plates from Hellma USA Inc. (Plainview, NY). Three spectra were collected and averaged in a wavelength range from 190-260 nm at a scanning range of 100 nm/min with a bandwidth of 2 nm and a data pitch of 0.2 nm.

3.5.3- Negative Stain TEM Analysis

Grids were prepared using diluted solutions of peptide (3 mg·mL^{-1}) in aqueous buffer (10 mM acetate, pH 4.0). Samples were prepared by depositing 4 μL of peptide solution onto a 200-mesh carbon-coated copper grid from Electron Microscopy Services (Hatfield, PA). After 90 s of incubation on the grid, moisture was wicked away, leaving a thin film of sample. An aliquot (4 μL) of negative stain solution (1% uranyl acetate) was deposited onto the thin film. After 60 s of staining, the remaining moisture was wicked away, and the grid dried in a tabletop desiccator in vacuo for at least 5 min. Electron micrographs were recorded on a Hitachi HT-7700 transmission electron microscopy with a tungsten filament and AMT CCD camera, at an accelerating voltage of 80 kV.

3.5.4- Motif Analysis

A search database was created by filtering the PDB, as of 9 Dec., 2020, using the following criteria: X-ray structures with resolution of 2.6 Å or better with no more than 5000 residues and no more than 26 chains in the biological unit. This search resulted in a total of 185,302 biological-unit entries. This set of structures was employed as the database in structural homology searchers using computational algorithm MASTER.[2, 3] A minimal assembly motif was constructed from four b-strands (2 inner + 2 outer) derived from the cross-b interface from the single protofilament reconstruction of the thinner tube . In order to eliminate artifacts due to the peptide termini, the N-terminal Phe and C-terminal Glu residues of KFE8 were trimmed in PyMol to generate a query structure (KFEFKF) that consisted of 4 b-strands, each of six residues. Two structural variants of KFEFKF were generated in which the orientations of inner and outer strands were changed to antiparallel and parallel respectively. The respective structures comprise antiparallel/antiparallel and parallel/parallel interfaces, which were employed as control motifs in structural homology searches. These controls generated by aligning the strands with the desired orientation to the strands with the undesired orientation in PyMol. The three motifs (KFEFKF and two controls) were employed as query structures in MASTER searches of the compiled database. All unique and non-redundant matches (i.e., less than 50% sequence identity than the queued sequence) with full-backbone RMSD below 2.0 Å were sought and used in the subsequent analysis.

3.5.5- Cryo-electron microscopy imaging and analysis

The peptide sample (ca. 2-2.5 □l) was applied to glow-discharged lacey carbon grids, and then plunge frozen using an EM GP Plunge Freezer (Leica). The cryo-EMs were collected on a 300 keV Titan Krios with a K3 camera (University of Virginia) at 1.08 Å/pixel and a total dose of ca. 50 e/Å^2. The micrographs were first motion corrected and dose weighted by MotionCorr v2,[4] and then CTF correction was applied by multiplying the images with the theoretical CTF. Filament images corresponding to ca. 20 electrons/Å^2 were extracted using EMAN2.[5] For each peptide filament, a list of possible helical symmetries was calculated from the averaged power spectrum of peptide particles (Supp Fig. 5,8). To determine the correct helical symmetry from the list, first if the full dataset has more than 30,000 particles, a subset containing 30,000 particles was generated. And then the initial volume was generated from those 30,000 particles with random assigned azimuthal angles. After that, possible helical symmetries were tested in Spider using IHRSR[6, 7] by trial and error until recognizable amino acid side chain densities could be observed in the correct symmetry. Specifically, the IHSRS reconstructions were divided into 3 steps, where 4x binned, 2x binned, un-binned particles were used to accelerate the progress. Helical symmetries were tightly locked in the first two steps, and then slightly relaxed in the third step. All helical parameters, including rise, rotation and point-group symmetry were imposed at the end of each IHRSR cycle. Since the KFE8 nanotubes are made of only β-sheet, the helical hand cannot be easily determined from the cryo-EM volume directly, unlike volumes that contain an α-helix, in which the hand is obvious when the resolution is 4.5 Å or better. Therefore, the hand is determined using prior knowledge of previously reported AFM results, where left-handed ribbons were reported.[8, 9]

3.5.6- Additional refinement of 1-protofilament

A detailed workflow of the thinner tube is shown in Supp Fig. 5. After determining the correct symmetry for the inner wall, a reconstruction using all particles was run in Spider.[6] The resulting volumes were then imported into RELION[10] for a 3D classification. Interactive 3D classifications were run in RELION for a potential separation of the two classes (T = 4 to T = 200), but the outcome was not ideal. Therefore, we used RELION “particle subtraction” to generate an image stack containing the signal of 1-protofilament, and exported the stack back to Spider for IHRSR. This ended up with a ~3.4 Å resolution reconstruction for the 1-protofilament of the thinner tube (Supp Fig. 9). A C2 and 2_1-like reference were generated from the 1-protofilament volume and these models were used for reference-based sorting. All maps were filtered to 3.0 Å without B-factor in Spider and then used as input in a new developed map sharpening tool, DeepEMhacer.[11] This sharpening result is slightly better than maps sharpened with a negative B-factor. The statistics are listed in Supplemental Tables 1.

3.5.7- Model building

The eight residue KFE8 model was generated in Chimera[12] from the corresponding peptide sequence and then docked in the EM maps. The ASU model was adjusted manually in Coot[13] to best fit into the map. Then, this adjusted ASU model was used to generate a model filament using the determined helical symmetry, which was then refined against the full cryo-EM map using real space refinement in PHENIX[13]. The refinement statistics are shown in Supplemental Tables 1. Cryo-EM maps and atomic coordinates have been deposited with the Electron Microscopy Data Bank and Protein Data Bank with accession codes given in Supplemental Table 1. Model versus map FSC calculations were employed to estimate the resolution of the reconstructions and are reported in Supplemental Fig. 9.

3.6 – Supplemental figures and table

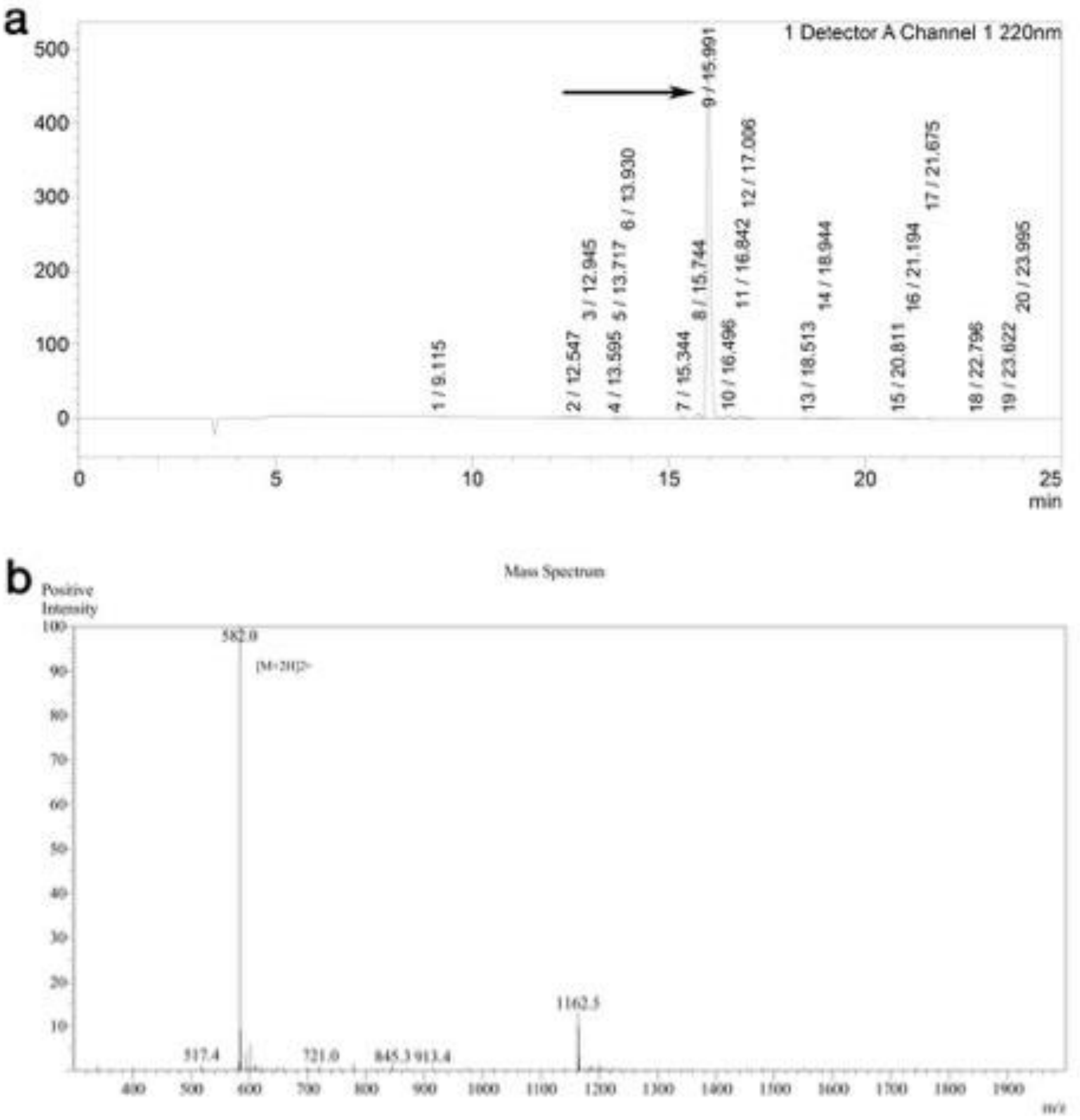


Figure 3-3-1: a, Analytical HPLC trace of KFE8, Ac-FKFEFKFE-NH_2. Arrow indicates the position of the major peak that corresponds to the KFE8 peptide (ca. 95.5% purity). **b**, Electrospray-ionization mass spectrum of KFE8 (calculated mass, 1162 Da; experimental mass, 1163 Da (m+1)).

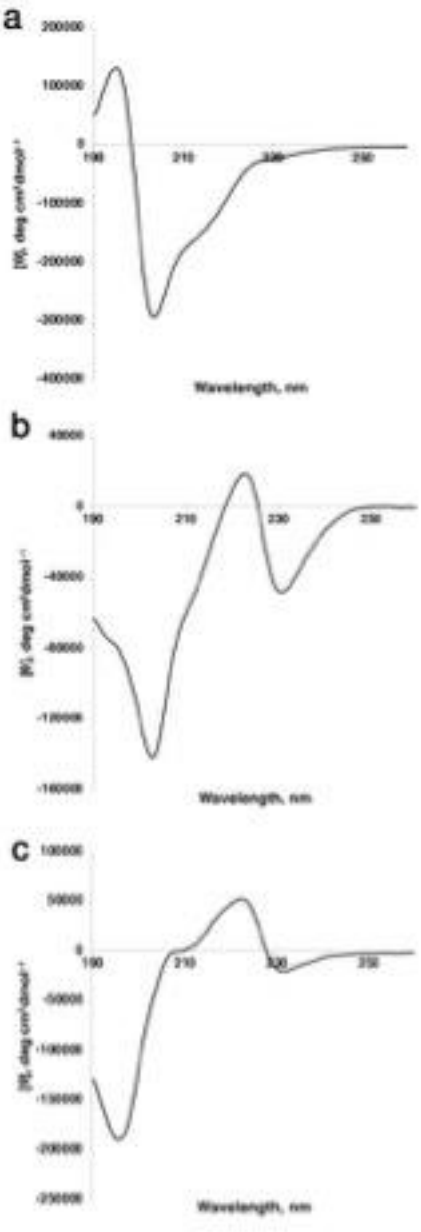


Figure 3-3-2: Circular dichroism spectra of KFE8 peptides assemblies. **a**, ambient sample incubated for 18 h (mostly ribbons). **b**, ambient sample incubated for two weeks (primarily thin nanotubes). **c**, annealed sample (primarily thicker tubes).

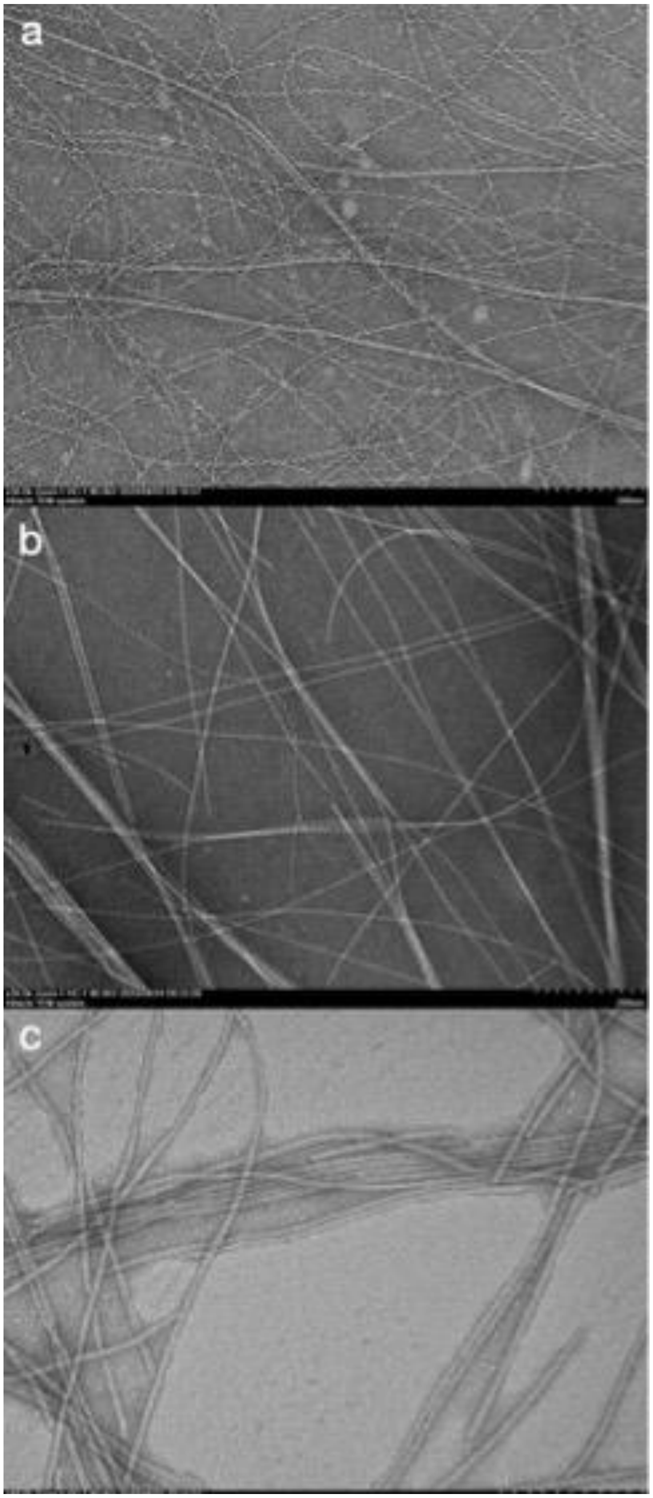


Figure 3-3-1: Representative negative-stain TEM images of KFE8 peptide assemblies. **a**, ambient sample incubated for 18 h. **b**, ambient sample incubated for two weeks. **c**, annealed sample (primarily thicker tubes). Total number of collected TEM images for each sample: 7 (**a**); 15 (**b**); 8 (**c**).

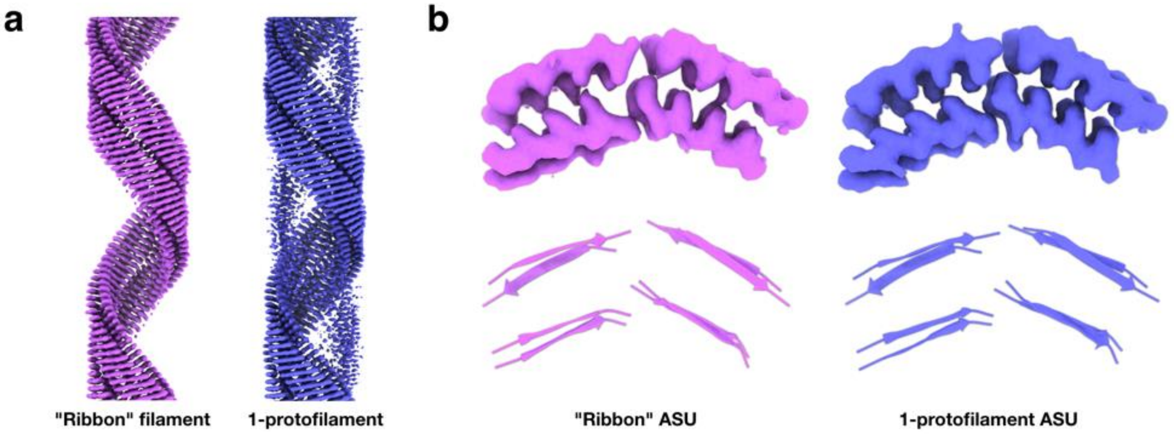


Figure 3-2-2: Cryo-EM of the ribbon filament and one protofilament of the thinner tube. **a,** The cryo-EM reconstructions of ribbon filament (~3.8 Å resolution, left) and protofilament (~3.4 Å resolution, right). **b,** The densities corresponding to a single ASU in both the ribbon filament and thinner tube protofilament are shown at the top. The ASU models are shown at the bottom.

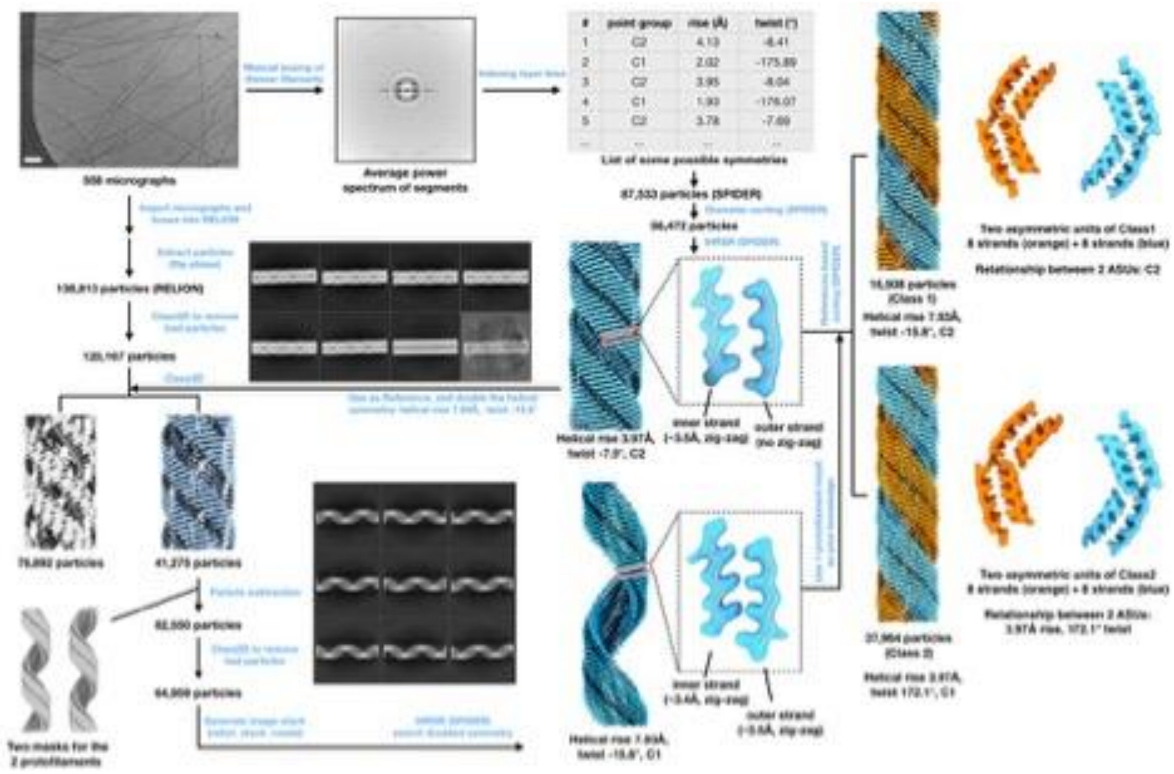


Figure 3-3-3: The cryo-EM helical processing workflow for the thinner tube.

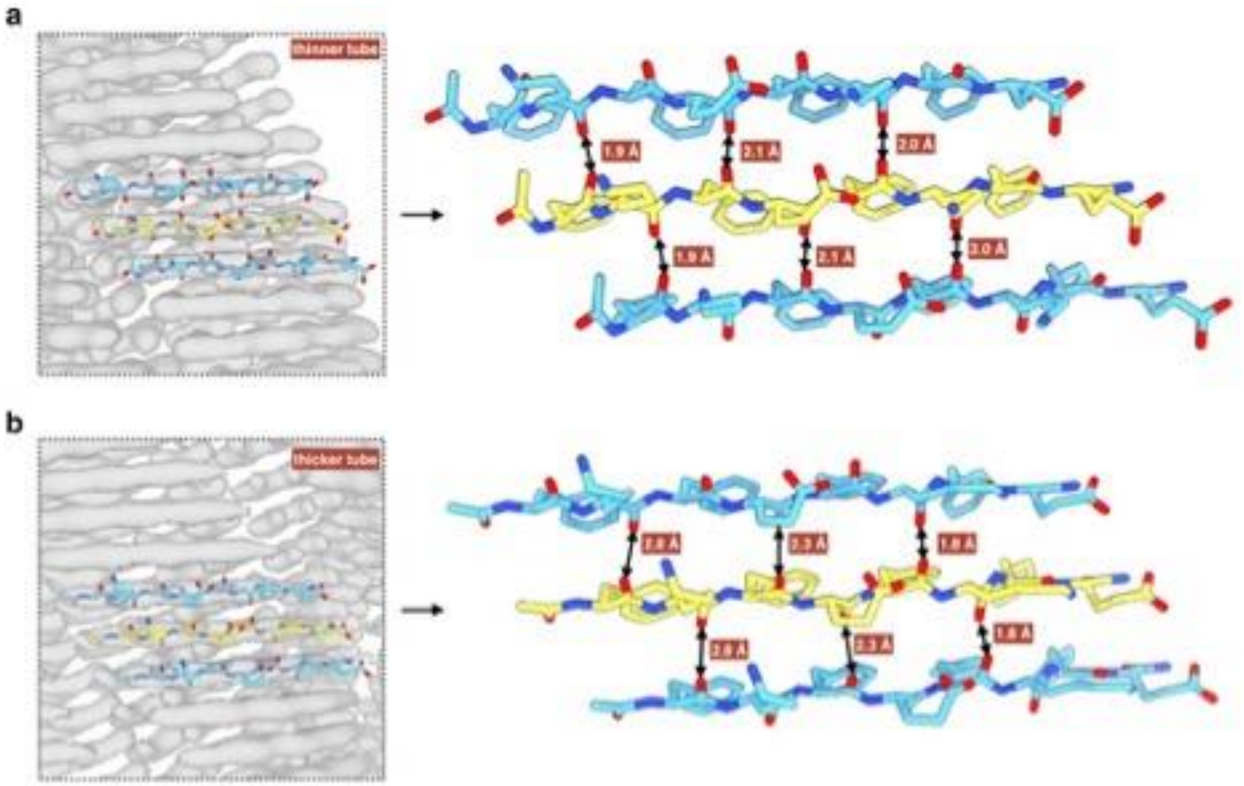


Figure 3-3-4: Steric clashes when fitting parallel sheet into the outer wall. In both thinner tube (**a**) and thicker tube (**b**), one can easily dock the first peptide (yellow) into the outer wall density. Then, assuming the outer wall is a parallel sheet, the upper and lower peptide will be related by helical symmetry (blue). Between the yellow and blue peptides, the small distances between oxygen atoms (labeled) would lead to steric clashes rather than the typical hydrogen bonds formed in β-sheets.

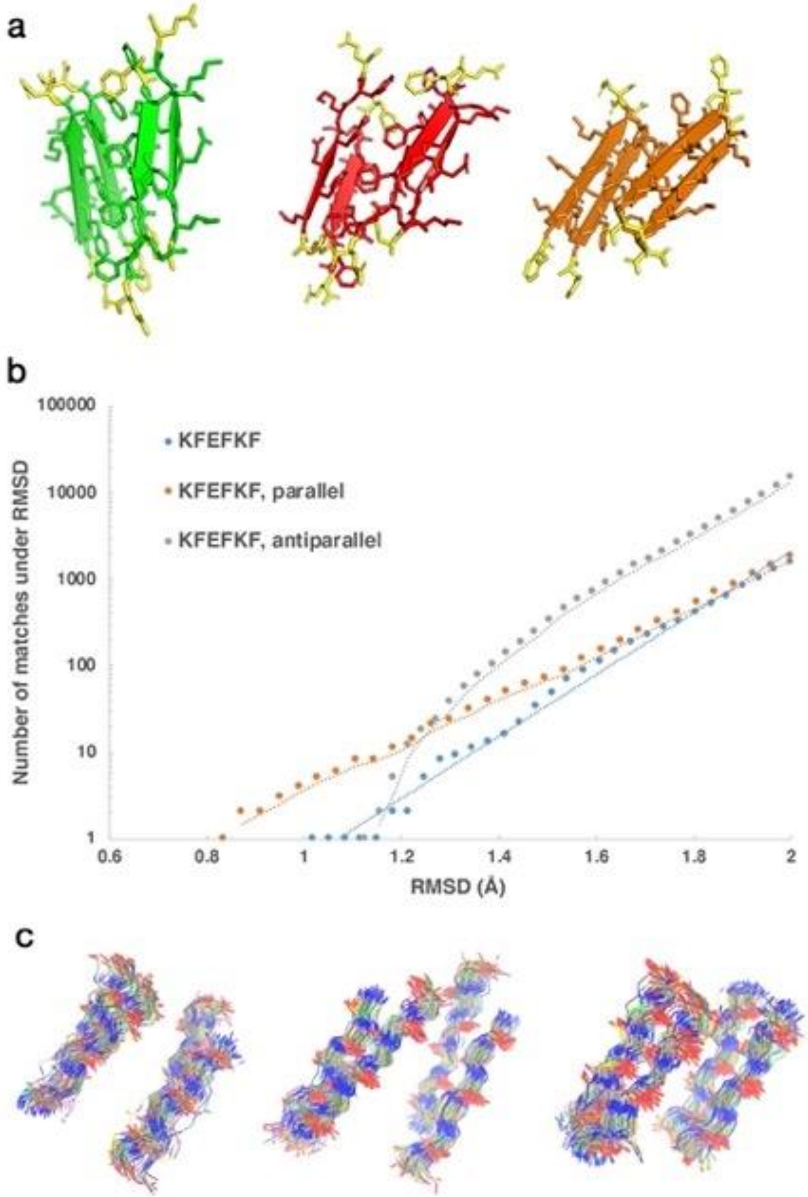


Figure 3-3-5: **a**, Minimal four-strand (2 inner + 2 outer) motifs employed as structural queries employed in MASTER(Kreutzberger, 2022 #2) searches of the Protein Data Bank (PDB): thin nanotube interface (green), antiparallel control (red), and parallel control (orange). The eight residue β-strands were trimmed at the N- and C-terminus to generate six core residue strands consisting of the sequence KFEFKF. **a**, RMSD divergence plots, showing the number of matches below a given RMSD cutoff for the thin tube interfacial motif (KFEFKF) motif and control motifs having parallel or antiparallel orientation. Threshold RMSDs lie within the expected ranges for motifs having a similar number of residues within the query structure. Matches to the thin tube motif and the parallel control motif are more rare than those to the antiparallel control motif, as the RMSD required to get up to a certain number of matches grows much more rapidly in the latter case. However, the greater number of matches for the antiparallel control may reflect the greater frequency of occurrence of antiparallel sheets in native protein structures.(Zhang, 1993 #39) Within the limitations of this analysis, the antiparallel-parallel sheet interfacial packing observed for the KFE8 assemblies does not seem any less common than the parallel control. Similar interfaces are observed in native protein sequences. **c**, Structural alignments of resulting 100 closest matches from the MASTER searches for the thin tube interface (left), antiparallel control motif (center), and parallel control motif (right).

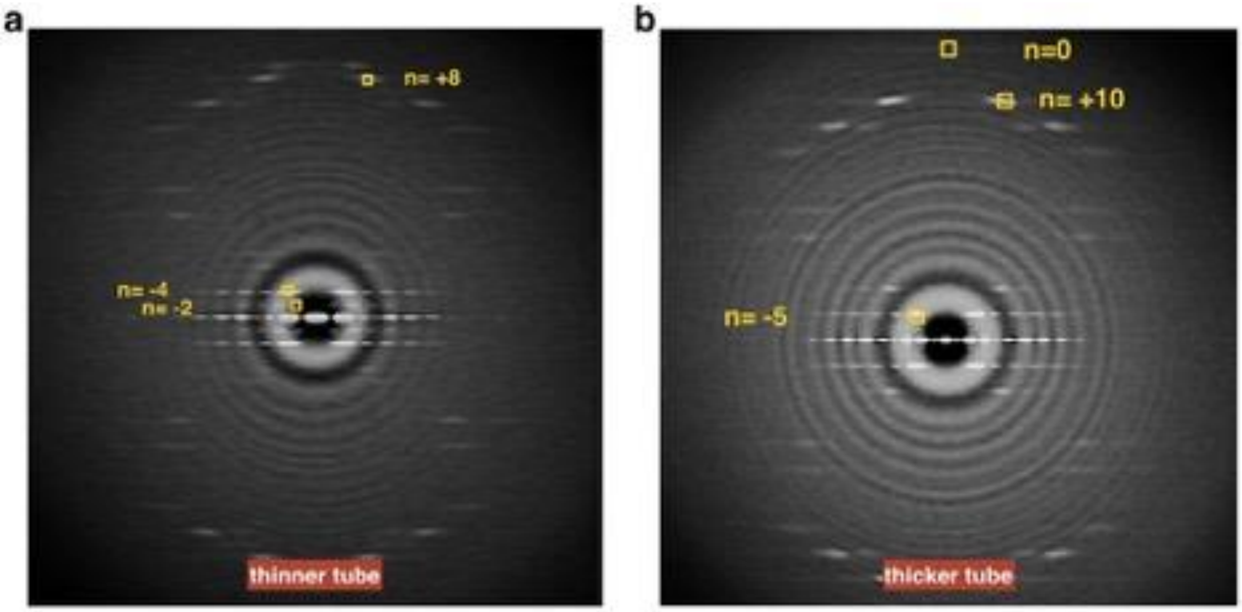


Figure 3-3-7: Power spectra of the thinner and thicker tube. The averaged power spectrum calculated from the particle stacks of the thinner tube (**a**) and the thicker tube (**b**). The layer-line indexing is also labeled: the Bessel order reflects th

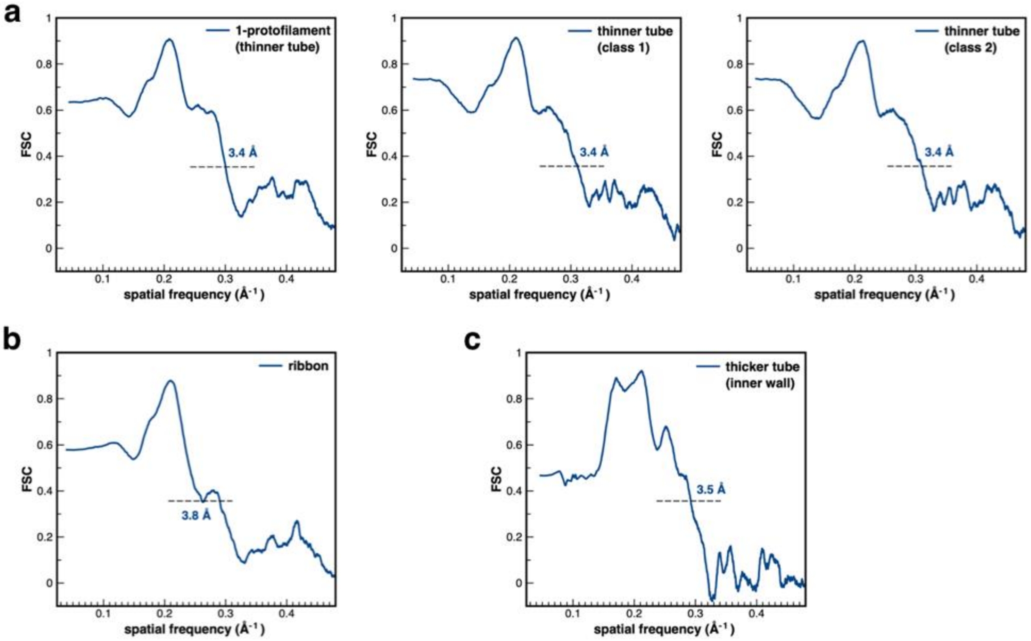


Figure 3-3-6: Fourier Shell Correlation (FSC) calculations. **a,** From left to right: protofilament of the thinner tube, thinner tube (class 1) and thinner tube (class 2). Model:map FSC calculations using a 0.38 criterion, which is sqrt (0.143). **b,** Ribbon filament. **c,** Thicker tube.

Supplemental Table 3-1:. Cryo-EM and Refinement Statistics of KFE8 Nanotubes

Parameter	**Ribbon filament**	**1-protofilament (thinner tube)**	**Thinner tube (class 1)**	**Thinner tube (class 2)**	**Thicker tube**
Data collection and processing					
Voltage (kV)	300	300	300	300	300
Electron exposure (e^{-} Å^{-2})	50	50	50	50	50
Pixel size (Å)	1.08	1.08	1.08	1.08	1.08
Particle images (n)	26,198	64,959	15,486	31,760	50,655
Shift between particles (pixel)	11	14	8	8	10
Helical symmetry					
Point group	C1	C1	C2	C1	C5*****
Helical rise (Å)	7.93	7.93	7.93	3.97	3.95
Helical twist (°)	-15.8	-15.8	-15.8	172.1	-6.4
Map resolution (Å)					
Software (final reconstruction)	SPIDER	RELION/SPIDER	SPIDER	SPIDER	SPIDER
Model:map FSC (0.38)	3.8	3.4	3.4	3.4	3.5*****
Refinement and Model validation					
Map-sharpening method	DeepEMhacer (ref: https://doi.org/10.1101/2020.06.12.148296)				
Bond lengths rmsd (Å)	0.007	0.008	0.008	0.008	0.008
Bond angles rmsd (°)	0.925	0.856	0.864	0.865	0.803
Clashscore	22.8	23.2	31.0	31.2	22.7
Ramachandran Favored (%)	100	100	100	100	100
Ramachandran Outlier (%)	0	0	0	0	0
MolProbity score	2.32	2.74	2.86	2.84	2.38
Deposition ID					
PDB (model)	Xxxx	Xxxx	Xxxx	Xxxx	Xxxx
EMDB (map)	EMD-00000	EMD-00000	EMD-00000	EMD-00000	EMD-00000

3.7- References

1. Bowerman, C.J., and Nilsson, B.L. (2012). Review self-assembly of amphipathic β-sheet peptides: Insights and applications. Peptide Science 98, 169-184. doi:https://doi.org/10.1002/bip.22058.

2. Zhang, S. (2017). Discovery and design of self-assembling peptides. Interface Focus 7, 20170028. doi:doi:10.1098/rsfs.2017.0028.

3. Beesley, J.L., and Woolfson, D.N. (2019). The de novo design of α-helical peptides for supramolecular self-assembly. Current Opinion in Biotechnology 58, 175-182. doi:https://doi.org/10.1016/j.copbio.2019.03.017.

4. Zhang, S., Holmes, T., Lockshin, C., and Rich, A. (1993). Spontaneous assembly of a selfcomplementary oligopeptide to form a stable macroscopic membrane. Proceedings of the National Academy of Sciences 90, 3334-3338. doi:10.1073/pnas.90.8.3334.

5. Aggeli, A., Bell, M., Boden, N., Keen, J.N., Knowles, P.F., McLeish, T.C.B., Pitkeathly, M., and Radford, S.E. (1997). Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheet tapes. Nature 386, 259-262. doi:10.1038/386259a0.

6. Schneider, J.P., Pochan, D.J., Ozbas, B., Rajagopal, K., Pakstis, L., and Kretsinger, J. (2002). Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide. Journal of the American Chemical Society 124, 15030-15037. doi:10.1021/ja027993g.

7. EANES, E.D., and GLENNER, G.G. (1968). X-RAY DIFFRACTION STUDIES ON AMYLOID FILAMENTS. Journal of Histochemistry & Cytochemistry 16, 673-677. doi:10.1177/16.11.673.

8. Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.C.F. (1997).

Common core structure of amyloid fibrils by synchrotron X-ray diffraction11Edited by F. E. Cohen. Journal of Molecular Biology 273, 729-739. doi:https://doi.org/10.1006/jmbi.1997.1348.

9. Diaz-Avalos, R., Long, C., Fontano, E., Balbirnie, M., Grothe, R., Eisenberg, D., and Caspar, D.L.D. (2003). Cross-beta Order and Diversity in Nanocrystals of an Amyloidforming Peptide. Journal of Molecular Biology 330, 1165-1175. doi:https://doi.org/10.1016/S0022-2836(03)00659-4.

10. Aggeli, A., Nyrkova, I.A., Bell, M., Harding, R., Carrick, L., McLeish, T.C., Semenov, A.N., and Boden, N. (2001). 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 98, 11857-11862. doi:10.1073/pnas.191250198.

11. Lee, M., Wang, T., Makhlynets, O.V., Wu, Y., Polizzi, N.F., Wu, H., Gosavi, P.M., Stohr, J., Korendovych, I.V., DeGrado, W.F., et al. (2017). Zinc-binding structure of a catalytic amyloid from solid-state NMR. Proc Natl Acad Sci U S A 114, 6191-6196. doi:10.1073/pnas.1706179114.

12. Nagy-Smith, K., Moore, E., Schneider, J., and Tycko, R. (2015). Molecular structure of monomorphic peptide fibrils within a kinetically trapped hydrogel network. Proc Natl Acad Sci U S A 112, 9816-9821. doi:10.1073/pnas.1509313112.

13. Cormier, A.R., Pang, X., Zimmerman, M.I., Zhou, H.X., and Paravastu, A.K. (2013). Molecular structure of RADA16-I designer self-assembling peptide nanofibers. ACS Nano 7, 7562-7572. doi:10.1021/nn401562f.

14. Knowles, T.P., Fitzpatrick, A.W., Meehan, S., Mott, H.R., Vendruscolo, M., Dobson, C.M., and Welland, M.E. (2007). Role of Intermolecular Forces in Defining Material Properties of Protein Nanofibrils. Science 318, 1900-1903. doi:10.1126/science.1150057.

15. Schleeger, M., vandenAkker, C.C., Deckert-Gaudig, T., Deckert, V., Velikov, K.P., Koenderink, G., and Bonn, M. (2013). Amyloids: From molecular structure to mechanical properties. Polymer 54, 2473-2488. doi:https://doi.org/10.1016/j.polymer.2013.02.029. 16. Ke, P.C., Zhou, R., Serpell, L.C., Riek, R., Knowles, T.P.J., Lashuel, H.A., Gazit, E., Hamley, I.W., Davis, T.P., Fändrich, M., et al. (2020). Half a century of amyloids: past, present and future. Chemical Society Reviews 49, 5473-5509. doi:10.1039/C9CS00199A.

17. Close, W., Neumann, M., Schmidt, A., Hora, M., Annamalai, K., Schmidt, M., Reif, B., Schmidt, V., Grigorieff, N., and Fandrich, M. (2018). Physical basis of amyloid fibril polymorphism. Nat Commun 9, 699. doi:10.1038/s41467-018-03164-5.

18. Guenther, E.L., Ge, P., Trinh, H., Sawaya, M.R., Cascio, D., Boyer, D.R., Gonen, T., Zhou, Z.H., and Eisenberg, D.S. (2018). Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. Nat Struct Mol Biol 25, 311-319. doi:10.1038/s41594-018-0045-5.

19. Zhang, W., Falcon, B., Murzin, A.G., Fan, J., Crowther, R.A., Goedert, M., and Scheres, S.H. (2019). Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. Elife 8. doi:10.7554/eLife.43584.

20. Kollmer, M., Close, W., Funk, L., Rasmussen, J., Bsoul, A., Schierhorn, A., Schmidt, M., Sigurdson, C.J., Jucker, M., and Fändrich, M. (2019). Cryo-EM structure and polymorphism of Aβ amyloid fibrils purified from Alzheimer’s brain tissue. Nature Communications 10, 4760. doi:10.1038/s41467-019-12683-8.

21. Bansal, A., Schmidt, M., Rennegarbe, M., Haupt, C., Liberta, F., Stecher, S., PuscalauGirtu, I., Biedermann, A., and Fändrich, M. (2021). AA amyloid fibrils from diseased tissue are structurally different from in vitro formed SAA fibrils. Nature Communications 12, 1013. doi:10.1038/s41467-021-21129-z.

22. Marini, D.M., Hwang, W., Lauffenburger, D.A., Zhang, S., and Kamm, R.D. (2002). LeftHanded Helical Ribbon Intermediates in the Self-Assembly of a β-Sheet Peptide. Nano Letters 2, 295-299. doi:10.1021/nl015697g.

23. Hwang, W., Marini, D.M., Kamm, R.D., and Zhang, S. (2003). Supramolecular structure of helical ribbons self-assembled from a β-sheet peptide. The Journal of Chemical Physics 118, 389-397. doi:10.1063/1.1524618.

24. Rudra, J.S., Sun, T., Bird, K.C., Daniels, M.D., Gasiorowski, J.Z., Chong, A.S., and Collier, J.H. (2012). Modulating Adaptive Immune Responses to Peptide Self-Assemblies. ACS Nano 6, 1557-1564. doi:10.1021/nn204530r.

25. Wan, S., Borland, S., Richardson, S.M., Merry, C.L.R., Saiani, A., and Gough, J.E. (2016). Self-assembling peptide hydrogel for intervertebral disc tissue engineering. Acta Biomaterialia 46, 29-40. doi:https://doi.org/10.1016/j.actbio.2016.09.033.

26. Egelman, E.H., Xu, C., DiMaio, F., Magnotti, E., Modlin, C., Yu, X., Wright, E., Baker, D., and Conticello, V.P. (2015). Structural plasticity of helical nanotubes based on coiled-coil assemblies. Structure 23, 280-289. doi:10.1016/j.str.2014.12.008.

27. DiMaio, F., Song, Y., Li, X., Brunner, M.J., Xu, C., Conticello, V., Egelman, E., Marlovits, T., Cheng, Y., and Baker, D. (2015). Atomic-accuracy models from 4.5-A cryo-electron microscopy data with density-guided iterative local refinement. Nat Methods 12, 361365. doi:10.1038/nmeth.3286.

28. Hughes, S.A., Wang, F., Wang, S., Kreutzberger, M.A.B., Osinski, T., Orlova, A., Wall, J.S., Zuo, X., Egelman, E.H., and Conticello, V.P. (2019). Ambidextrous helical nanotubes from self-assembly of designed helical hairpin motifs. Proc Natl Acad Sci U S A 116, 1445614464. doi:10.1073/pnas.1903910116.

29. Chen, Y.X., Ing, N.L., Wang, F., Xu, D., Sloan, N.B., Lam, N.T., Winter, D.L., Egelman, E.H.,

Hochbaum, A.I., Clark, D.S., et al. (2020). Structural Determination of a Filamentous Chaperone to Fabricate Electronically Conductive Metalloprotein Nanowires. ACS Nano 14, 6559-6569. doi:10.1021/acsnano.9b09405.

30. Feng, Z., Wang, H., Wang, F., Oh, Y., Berciu, C., Cui, Q., Egelman, E.H., and Xu, B. (2020).

Artificial Intracellular Filaments. Cell Reports Physical Science 1, 100085. doi:https://doi.org/10.1016/j.xcrp.2020.100085.

31. Wang, F., Gnewou, O., Modlin, C., Beltran, L.C., Xu, C., Su, Z., Juneja, P., Grigoryan, G., Egelman, E.H., and Conticello, V.P. (2021). Structural analysis of cross α-helical nanotubes provides insight into the designability of filamentous peptide nanomaterials. Nature Communications 12, 407. doi:10.1038/s41467-020-20689-w.

32. Bowerman, C.J., Ryan, D.M., Nissan, D.A., and Nilsson, B.L. (2009). The effect of increasing hydrophobicity on the self-assembly of amphipathic beta-sheet peptides. Mol Biosyst 5, 1058-1069. doi:10.1039/b904439f.

33. Subramaniam, S., Earl, L.A., Falconieri, V., Milne, J.L., and Egelman, E.H. (2016). Resolution advances in cryo-EM enable application to drug discovery. Curr Opin Struct Biol 41, 194-202. doi:10.1016/j.sbi.2016.07.009.

34. Kuhlbrandt, W. (2014). Cryo-EM enters a new era. eLife 3, e03678. doi:10.7554/eLife.03678.

35. Feng, Z., Wang, H., Wang, F., Oh, Y., Berciu, C., Cui, Q., Egelman, E.H., and Xu, B. (2020). Artificial Intracellular Filaments. Cell Rep Phys Sci 1. doi:10.1016/j.xcrp.2020.100085.

36. Shen, H., Fallas, J.A., Lynch, E., Sheffler, W., Parry, B., Jannetty, N., Decarreau, J., Wagenbach, M., Vicente, J.J., Chen, J., et al. (2018). De novo design of self-assembling helical protein filaments. Science 362, 705-709. doi:10.1126/science.aau3775.

37. Wang, F., Gnewou, O. , Modlin, C. , Beltran, L. C., Xu, C., Su, Z., Juneja, P., Grigoryan, G., Egelman, E.H. , Conticello, V.P. (2021). Cross α-Helical Nanotubes: Mining the Designability Landscape for Filamentous Peptide. Nature Communications.

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

39. Tsutsumi, M., and Otaki, J.M. (2011). Parallel and Antiparallel β-Strands Differ in Amino Acid Composition and Availability of Short Constituent Sequences. Journal of Chemical Information and Modeling 51, 1457-1464. doi:10.1021/ci200027d.

40. Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.O., Riekel, C., Grothe, R., and Eisenberg, D. (2005). Structure of the cross-beta spine of amyloid-like fibrils. Nature 435, 773-778. doi:10.1038/nature03680.

41. Eisenberg, D.S., and Sawaya, M.R. (2017). Structural Studies of Amyloid Proteins at the Molecular Level. Annual Review of Biochemistry 86, 69-95. doi:10.1146/annurevbiochem-061516-045104.

42. Caspar, D.L., and Klug, A. (1962). Physical principles in the construction of regular viruses. Cold Spring Harb.Symp.Quant.Biol. 27, 1-24.

43. Chretien, D., Metoz, F., Verde, F., Karsenti, E., and Wade, R.H. (1992). Lattice defects in microtubules: protofilament numbers vary within individual microtubules. Journal of Cell Biology 117, 1031-1040.

44. Lu, K., Jacob, J., Thiyagarajan, P., Conticello, V.P., and Lynn, D.G. (2003). Exploiting Amyloid Fibril Lamination for Nanotube Self-Assembly. Journal of the American Chemical Society 125, 6391-6393. doi:10.1021/ja0341642.

45. Narayanan, T., Rüter, A., and Olsson, U. (2021). SAXS/WAXS Investigation of Amyloidβ(16-22) Peptide Nanotubes. Frontiers in Bioengineering and Biotechnology 9. doi:10.3389/fbioe.2021.654349.

46. Nyrkova, I.A., Semenov, A.N., Aggeli, A., and Boden, N. (2000). Fibril stability in solutions of twisted -sheet peptides: a new kind of micellization in chiral systems. The European Physical Journal B - Condensed Matter and Complex Systems 17, 481-497. doi:10.1007/s100510070127.

47. Rüter, A., Kuczera, S., Stenhammar, J., Zinn, T., Narayanan, T., and Olsson, U. (2020). Tube to ribbon transition in a self-assembling model peptide system. Physical Chemistry Chemical Physics 22, 18320-18327. doi:10.1039/D0CP03204B.

48. Kamtekar, S., Schiffer, J., Xiong, H., Babik, J., and Hecht, M. (1993). Protein design by binary patterning of polar and nonpolar amino acids. Science 262, 1680-1685. doi:10.1126/science.8259512.

49. West, M.W., and Hecht, M.H. (1995). Binary patterning of polar and nonpolar amino acids in the sequences and structures of native proteins. Protein Science 4, 2032-2039. doi:https://doi.org/10.1002/pro.5560041008.

50. Xiong, H., Buckwalter, B.L., Shieh, H.M., and Hecht, M.H. (1995). Periodicity of polar and nonpolar amino acids is the major determinant of secondary structure in selfassembling oligomeric peptides. Proceedings of the National Academy of Sciences 92, 6349-6353. doi:10.1073/pnas.92.14.6349.

51. West, M.W., Wang, W., Patterson, J., Mancias, J.D., Beasley, J.R., and Hecht, M.H. (1999). De novo amyloid proteins from designed combinatorial libraries. Proceedings of the National Academy of Sciences 96, 11211-11216. doi:10.1073/pnas.96.20.11211.

52. Dou, Y., Baisnée, P.-F., Pollastri, G., Pécout, Y., Nowick, J., and Baldi, P. (2004). ICBS: a database of interactions between protein chains mediated by β-sheet formation. Bioinformatics 20, 2767-2777.

53. Mandel-Gutfreund, Y., and Gregoret, L.M. (2002). On the Significance of Alternating Patterns of Polar and Non-polar Residues in Beta-strands. Journal of Molecular Biology 323, 453-461. doi:https://doi.org/10.1016/S0022-2836(02)00973-7.

54. Salemme, F.R. (1983). Structural properties of protein beta-sheets. Prog Biophys Mol Biol 42, 95-133. doi:10.1016/0079-6107(83)90005-6.

55. Vauthey, S., Santoso, S., Gong, H., Watson, N., and Zhang, S. (2002). Molecular selfassembly of surfactant-like peptides to form nanotubes and nanovesicles. Proc Natl Acad Sci U S A 99, 5355-5360. doi:10.1073/pnas.072089599.

56. von Maltzahn, G., Vauthey, S., Santoso, S., and Zhang, S. (2003). Positively Charged Surfactant-like Peptides Self-assemble into Nanostructures. Langmuir 19, 4332-4337. doi:10.1021/la026526+.

57. Bucak, S., Cenker, C., Nasir, I., Olsson, U., and Zackrisson, M. (2009). Peptide nanotube nematic phase. Langmuir 25, 4262-4265. doi:10.1021/la804175h.

58. Castelletto, V., Nutt, D.R., Hamley, I.W., Bucak, S., Cenker, C., and Olsson, U. (2010). Structure of single-wall peptide nanotubes: in situ flow aligning X-ray diffraction. Chem Commun (Camb) 46, 6270-6272. doi:10.1039/c0cc00212g.

59. Middleton, D.A., Madine, J., Castelletto, V., and Hamley, I.W. (2013). Insights into the molecular architecture of a peptide nanotube using FTIR and solid-state NMR spectroscopic measurements on an aligned sample. Angew Chem Int Ed Engl 52, 1053710540. doi:10.1002/anie.201301960.

60. Zhao, Y., Wang, J., Deng, L., Zhou, P., Wang, S., Wang, Y., Xu, H., and Lu, J.R. (2013). Tuning the self-assembly of short peptides via sequence variations. Langmuir 29, 1345713464. doi:10.1021/la402441w.

61. Zhao, Y., Deng, L., Wang, J., Xu, H., and Lu, J.R. (2015). Solvent Controlled Structural Transition of KI4K Self-Assemblies: from Nanotubes to Nanofibrils. Langmuir 31, 1297512983. doi:10.1021/acs.langmuir.5b02303.

62. Zhao, Y., Yang, W., Wang, D., Wang, J., Li, Z., Hu, X., King, S., Rogers, S., Lu, J.R., and Xu, H. (2018). Controlling the Diameters of Nanotubes Self-Assembled from Designed Peptide Bolaphiles. Small 14, 1703216. doi:10.1002/smll.201703216.

63. Cheng, P.N., Pham, J.D., and Nowick, J.S. (2013). The supramolecular chemistry of betasheets. Journal of the American Chemical Society 135, 5477-5492. doi:10.1021/ja3088407.

3.8 Supplemental References

(1)Bowerman, C. J.; Ryan, D. M.; Nissan, D. A.; Nilsson, B. L. The effect of increasing hydrophobicity on the self-assembly of amphipathic β-sheet peptides. *Molecular BioSystems* **2009**, *5* (9), 1058-1069, 10.1039/B904439F. DOI: 10.1039/B904439F.

(2)Zhou, J.; Grigoryan, G. Rapid search for tertiary fragments reveals protein sequence–structure relationships. *Protein Science* **2015**, *24* (4), 508-524. DOI: 10.1002/pro.2610 (acccessed 2020/05/15).

(3)Zhou, J.; Grigoryan, G. A C++ library for protein sub-structure search. *bioRxiv* **2020**, 2020.2004.2026.062612. DOI: 10.1101/2020.04.26.062612.

(4)Mindell, J. A.; Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology* **2003**, *142* (3), 334-347. DOI: https://doi.org/10.1016/S1047-8477(03)00069-8.

(5)Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **2007**, *157* (1), 38-46. DOI: 10.1016/j.jsb.2006.05.009.

(6)Frank, J.; Radermacher, M.; Penczek, P.; Zhu, J.; Li, Y.; Ladjadj, M.; Leith, A. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **1996**, *116* (1), 190-199. DOI: 10.1006/jsbi.1996.0030.

(7)Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **2000**, *85* (4), 225-234. DOI: 10.1016/s0304-3991(00)00062-0 From NLM.

(8)Marini, D. M.; Hwang, W.; Lauffenburger, D. A.; Zhang, S.; Kamm, R. D. Left-Handed Helical Ribbon Intermediates in the Self-Assembly of a β-Sheet Peptide. *Nano Letters* **2002**, *2* (4), 295-299. DOI: 10.1021/nl015697g.

(9)Hwang, W.; Marini, D. M.; Kamm, R. D.; Zhang, S. Supramolecular structure of helical ribbons self-assembled from a β-sheet peptide. *The Journal of Chemical Physics* **2002**, *118* (1), 389-397. DOI: 10.1063/1.1524618 (acccessed 2021/01/12).

(10)Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **2012**, *180* (3), 519-530. DOI: https://doi.org/10.1016/j.jsb.2012.09.006.

(11)Sanchez-Garcia, R.; Gomez-Blanco, J.; Cuervo, A.; Carazo, J.; Sorzano, C.; Vargas, J. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *bioRxiv* **2020**, 2020.2006.2012.148296. DOI: 10.1101/2020.06.12.148296.

(12)Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-1612. DOI: 10.1002/jcc.20084.

(13)Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **2010**, *66* (4), 486-501. DOI: doi:10.1107/S0907444910007493.

Chapter 4: Cryo-EM of helical Polymers

This chapter has been adapted with permission from Wang F, Gnewou O, Solemanifar A, Conticello VP, Egelman EH. Cryo-EM of Helical Polymers. Chem Rev. 2022 Feb 8. doi: 10.1021/acs.chemrev.1c00753. Epub ahead of print. PMID: 35133794.

4.1- Introduction

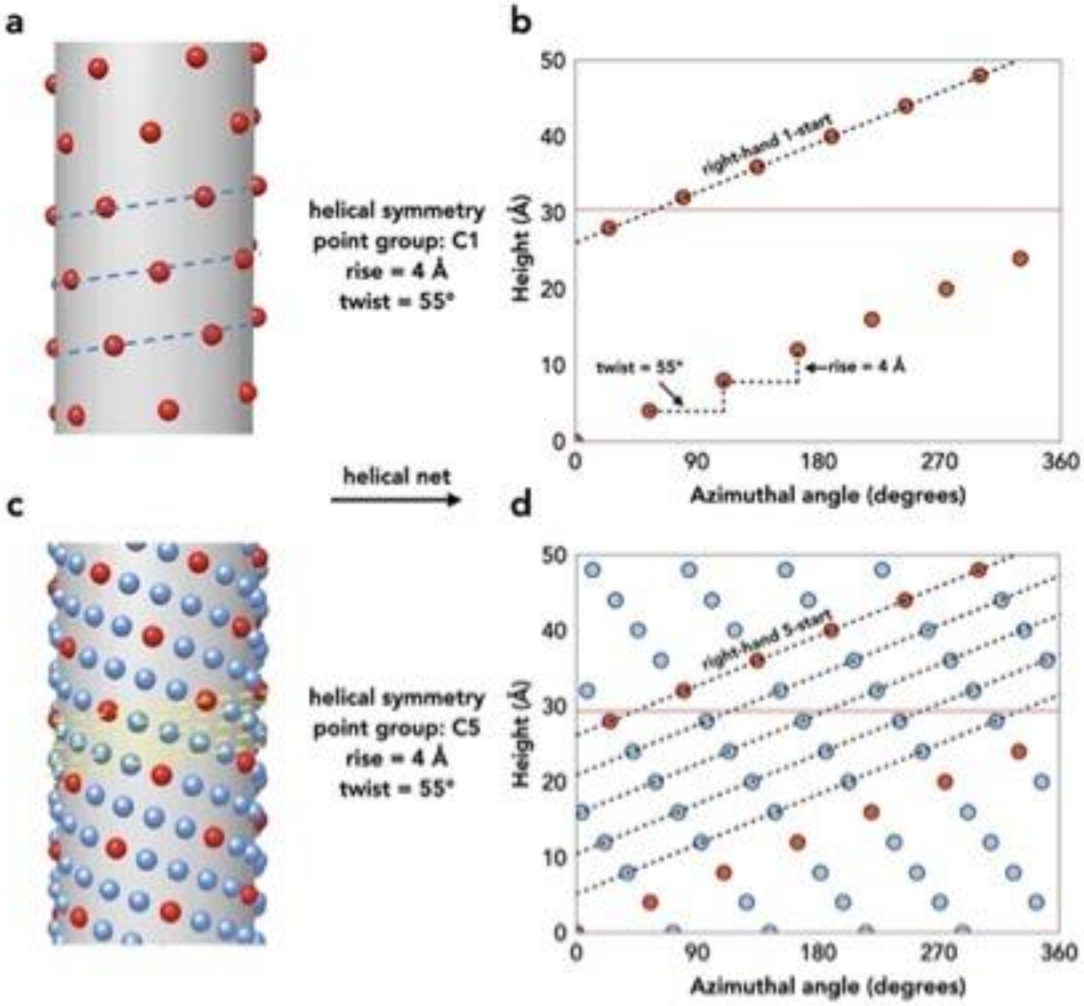
It is probably fair to say that most protein (by mass) in both prokaryotic and eukaryotic cells is found in the form of helical polymers. In prokaryotes some well-known examples are the actin and tubulin homologs[1] (e.g., MreB, ParM, FtsZ, etc.), flagellin and pili[2]. In flagellated bacteria, flagellin can be the single most abundant protein. In response to large amounts of DNA damage, the RecA protein, which forms a helical polymer on DNA, can be the single most abundant protein in bacterial cells[3]. In eukaryotes, in addition to actin (the single most abundant protein) and tubulin there are the intermediate filaments, collagen, etc., and we are still learning about other polymers within cells. For example, we now understand that nucleated polymerization of the helical inflammasome initiates a pathway leading to pyroptosis[4]. The first virus ever isolated, tobacco mosaic virus (TMV), is a helical polymer of a protein bound to single-stranded RNA[5,6]. The early structural characterization of viruses led to the realization that the two simplest ways that multiple copies of a single protein could be assembled into a capsid involved either a helix or an icosahedron[7,8]. The frequently overlooked reason for the ubiquity of helical polymers in biology is that a helical assembly is the most likely consequence of any arbitrary favorable interaction between two copies of the same molecule when this interaction is extended to multiple subunits[9].

It is no coincidence, therefore, that the first application of three-dimensional reconstruction from electron microscopic images to a biological system was for the helical tail of an icosahedral bacteriophage[10]. While the resolution obtained from negatively stained samples in this pioneering work in 1968 was about 35 Å, it nevertheless opened up the possibility that two-dimensional EM images might be routinely used to generate three dimensional reconstructions of biological assemblies. A simple measure of progress is that the same bacteriophage T4 tail tube has now been reconstructed at 3.4 Å resolution using cryoEM[11], which is an increase in information content by about 1,000 from the work done in 1968. In fact, this comparison illustrates the general progress in cryo-EM, where a near-atomic level of resolution has now become the standard, and not the exception. To put this in perspective, a search of the Electron Microscopy Data Bank shows 440 helical structures deposited with better than 5.0 Å resolution, and the number of such deposits per year[12] shows the exponential growth reflecting the overall "Resolution Revolution" in the field[13-15].

In addition to their ubiquity, helical polymers have the attractive feature that, in principle, a single helical polymer provides all the information in a single projection image to reconstruct in three-dimensions. This is due to the fact that the helical symmetry is generating what is effectively a single-axis tilt series showing all of the different projections of a single subunit as it is rotated about the helical axis. One thus has no need for other views, such as down the helical axis, in the same way that medical tomography of an asymmetric volume (e.g., a skull or a chest) can obtain all the needed three-dimensional information from collecting the projections as a source and detector are rotated around a single axis. In practice, the high degree of noise present in cryo-EM images (with a signal-to-noise ratio <<1) due to electron shot statistics, resulting from the low dose needed to minimize radiation damage, requires that extensive averaging must be used, so one may need to collect images from hundreds or thousands of polymers to reach a near-atomic level of resolution. Further, the idealization that a single projection contains all of the information needed may only be true in the absence of what is called in Fourier space "Bessel-overlap" which may arise at relatively low resolution when there are a limited number of subunits in a true repeat. Thus, if one has a helix with exactly seven subunits per turn, then a projection image of a single filament will only show seven different projection angles for the component subunit. This would be equivalent to a single-axis tilt series with tilt increments of 51.4°, yielding a very low-resolution reconstruction. For example, if the filament had a diameter of 100 Å, then the resolution obtainable would only be ~ 90 Å, using the relationship that $d = D\Delta\alpha$, where d is the resolution, D is the diameter, and $\Delta\alpha$ is the tilt increment in radians. Since there is no reason why any polymer, outside of a crystal, should have an integer number of subunits per turn[16], consider a filament with 7.1 subunits per turn. A single filament would generate 71 different projections of the component subunit, equivalent to a single-axis tilt series with 5.07° increments. For the 100 Å diameter filament, the resolution obtainable would be 8.8 Å, still quite far from what is needed for building atomic models. We expect that filaments will have a random azimuthal orientation in cryo-EM images, so multiple filaments can thus be used to finely sample the azimuthal orientation space, yielding the information needed for high resolution.

It is helpful whenever discussing helical polymers to introduce the "helical net" (**Figure 4-1**) and some terminology essential for describing the symmetry. Helical symmetry, or screw symmetry, involves a coupled rotation ($\Delta\theta$) about a helical axis with a translation (Δz) along the helical axis. In addition, there can be a C_n rotational point group symmetry about the helical axis. In the absence of a rotational symmetry, any helical lattice can therefore be completely described by the parameters of the 1-start helix that passes through every subunit, where the pitch of the 1-

start helix is $\Delta z * 360^\circ/\Delta\theta$. The "start" number of the helix is simply how many times a horizontal line in the helical net intersects the helical lines (**Figure 4-1b,d**).



<u>Figure 4-1:</u> Helical symmetry is best understood in terms of the helical net.

(a) An illustration of a helix without point group symmetry (rise: 4 Å, twist: 55°, C_1). The asymmetric units (ASU) are represented by spheres. The right-handed 1-start helix that passes through every subunit is shown as a dashed line.

(b) The helical net of **a**, using the convention that the surface is unrolled, and we are viewing it from the outside. The helix crosses the horizontal red line once, so the helix is called a 1-start. The helical rise and helical twist along the right-handed 1-start helix are labeled.

(c) An illustration of a helix with point group symmetry (rise: 4 Å, twist: 55°, C_5). Because of the rotational point group symmetry, any rotation of this structure by multiples of 72° is an identity operation. The asymmetric units are represented by spheres. The subunits along a single 5-start strand are shown in red, and the dashed lines show the 5-start helices.

(d) The helical net of **c**, using the convention above. The right-handed 5-start helices cross the horizontal red line five times, hence the name 5-start.

In a three-dimensional crystal the only allowed helical symmetries have 2, 3, 4 or 6 subunits per turn. But outside of a crystal there is absolutely no reason why the number of subunits per turn ($360°/\Delta\theta$) should be integer. Formalisms that are based upon describing helical symmetry as integer ratios[17] have a helical repeat, where one subunit can be translated along the helical axis by this repeat and be exactly superimposed upon another subunit. But such formalisms are ill-conditioned due to the fact that infinitesimally small changes in the helical twist result in huge changes in the helical repeat[16], and it is much better to simply describe the symmetry in terms of the two continuously variable real numbers, $\Delta\theta$ and Δz. When there is a rotational point group symmetry present, there cannot be a 1-start helix that passes through every subunit. For a C_n point group symmetry, there will only be n-start helices and multiples of n. For example, if there is a C_5 rotational symmetry (Fig. 1), then the structure will only have 5-start, 10-start, 15-start, etc. helices. The convention that exists is that righthanded helices are generated by a positive rotation, while left-handed helices result from a negative rotation. The original convention[17] for the helical net involved unrolling the surface of a cylinder and looking at this two-dimensional lattice from the inside of the cylinder, and this convention was used in much of the earlier literature. Since we actually look at most helical structures, whether using three-dimensional reconstruction, AFM, metal-shadowing, or SEM, from the outside, it makes more sense to reverse the original convention (**Figure 4-1-b,d**).

In addition to a point group rotational symmetry, there can be a dihedral symmetry such that the polymer has a 2-fold axis perpendicular to the helical axis. For polymers built from intrinsically asymmetric subunits such as peptides or proteins, this requires that the asymmetric unit in the polymer must be at least a dimer. When such a 2-fold symmetry exists along with an n-fold rotational symmetry, this can be represented as an overall D_n dihedral symmetry.

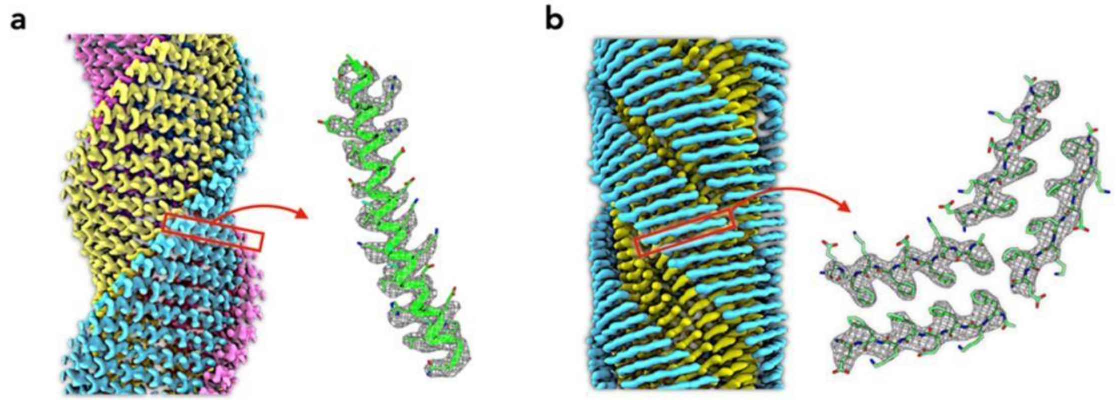
For over 30 years since the introduction of three-dimensional electron microscopy[10] in 1968 the main approach used for the reconstruction of helical polymers was the Fourier-Bessel method[17]. An alternative approach, called the Iterative Helical Real Space Reconstruction (IHRSR) method[18] has now dominated the field. A number of reviews have discussed the advantages of the IHRSR method and the limitations of the Fourier-Bessel approach[16,19-21]. Since no one is currently promoting Fourier-Bessel methods as superior to real-space approaches, discussing this further is beating a dead horse. However, an early misconception was that one of the advantages of the real-space approach was that one did not need to understand FourierBessel analysis. This, as discussed below, turns out to be wrong, and the greatest challenge in applications of IHRSR is usually determining the helical symmetry[12,22] which is best done through an understanding of the Fourier-Bessel formalism.

While all of the early applications of IHRSR were made with what is now legacy software, the SPIDER package[23,24], the IHRSR algorithm has been implemented in comprehensive modern packages such as Relion[25] and cryoSPARC[26,27].

4.2- Soft Matter

While there is a vast literature and associated prodigious research effort on supramolecular assemblies of nanotubes from peptides and small molecules[28], there have been only a handful of high-resolution cryo-EM structures from these complexes which we will discuss. One of the reasons that three-dimensional EM arose in the biological sciences as opposed to such areas as materials science or metallurgy (both areas where EM has also been extensively used) is that most biological specimens can be viewed as weakly scattering in terms of interactions with electrons having an energy of 100-300 kV. Thus, an electron passing through the sample will either not interact at all, or have a single interaction. In contrast, for a strongly scattering material such as a metal, an electron will typically have multiple interactions with the sample before it reaches the detector. This multiple scattering (or dynamic scattering) means that the image that is obtained can be quite complex and not simply related to the sample being examined. However, in cryo-EM of most macromolecular complexes it is reasonable to assume that the image, ignoring for the moment the Contrast Transfer Function (CTF), corresponds to the projected density of the sample. This relatively simple relationship leads to the ability to go from two-dimensional images to true atomic resolution volumes[29,30] using techniques such as weighted back-projection, algebraic reconstruction or Fourier methods.

Perhaps as a consequence of this history, the current capabilities of cryo-EM to routinely reach a near-atomic level of resolution for protein polymers are still largely unknown in materials science and chemistry. For example, there has been a mistaken belief that the "microbial nanowires" conducting electrons away from bacteria such as *Geobacter sulfurreducens* are Type IV pili, when it has now been shown by cryo-EM at near-atomic resolution that these filaments are actually a polymer of cytochromes[31]. Prior to the cytochrome publication, a 2017 paper in a chemistry journal stated that determining the actual structure of these filaments remains challenging due to the "low-resolution of electron microscopy techniques"[32]. There have now been a number of published high resolution cryoEM structures of synthetic polymers formed from peptides[33-37] that should begin to change this perception about resolution in chemistry and materials science, and we will focus on these papers in this review.



Cross-α nanotube (36-31-3 peptide, ~3.8 Å)

Cross-β nanotube (KFE8, ~3.5 Å)

Figure 4-2: Cryo-EM has become the method of choice for helical polymers.
Cryo-EM maps better than 4 Å resolution, now routinely attainable, clearly resolve some side chains of the peptide or peptide-like compounds, and one has the ability to build atomic models *de novo*.

(a) Cryo-EM reconstruction of a cross-α nanotube at 3.8 Å resolution[33]. The α-helices form stacks, and the three stacks forming the tube are each shown in a different color. The ASU is a single α-helical peptide containing 36 residues. The density from a single subunit is shown on the right as a grey mesh, and a ribbon diagram of the atomic model has been fit into this density.

(b) Cryo-EM reconstruction of a cross-β nanotube at 3.5 Å resolution[38]. The β-sheets on the inside (yellow) are parallel, and those on the outside (cyan) are antiparallel. The ASU contains four peptides, shown on the right, forming four β-strands. Although each 8-residue peptide is chemically identical, each of the four peptides in the ASU is in a different environment.

As shown in **Figure 4-2**, when one reaches a resolution of ~ 4 Å, there is no ambiguity in building atomic models into these peptide maps. The determination of the absolute hand can still be challenging, however. Since cryo-EM images are projections, information about the hand is lost, as a structure and the mirror-image of the structure will both give rise to indistinguishable projections. This can be understood at the anatomical level, where an x-ray of a left hand with the palm up is the same as the x-ray of a right hand with the palm down (**Figure 4-3)**. If one knew whether the palm was up or down, one could distinguish which was which. But in the absence of that knowledge there would be an enantiomorphic ambiguity.

<u>Figure 4-3:</u> X-ray of a human hand

The loss of information about the absolute hand in projection images is best illustrated with a clinical x-ray of the human hand, since such x-ray images, like cryo-EM images, are projections. From a radiograph on the left, one cannot tell whether this is a right hand, palm down, or a left hand, palm up. The image on the right is simply the mirror image of the one on the left, and this could be either a left hand, palm down, or a right hand, palm up. Radiograph courtesy of Department of Radiology, University of California San Diego

In cryo-EM, reconstructing a left-handed helix assuming that it is right-handed, or the opposite, should just generate a reconstructed volume that is the perfect mirror image of the true structure. One only needs to mirror this volume to correct this. A cryo-EM paper[39] looking at the assembly of amyloid filaments formed by a 101-residue fragment of a protein argued that the correct hand could be determined directly since the reconstructed volume using a twist of -7.68°

had a better resolution than one generated with a twist of 7.68°. This result cannot possibly be true, and the most likely explanation for their different resolutions involved the starting models for the left- and right-handed reconstructions. For example, one might have started both reconstructions with a volume having a left-handed twist, and it is therefore not surprising that the reconstruction with the left-handed twist had a better resolution. Righthanded versus left-handed versions of chiral monomers must necessarily be non-identical, just as a right-handed α-helix composed of L-amino acids must be different from a left-handed αhelix composed of D-amino acids (never found in nature). But the projections of a right-handed α-helix composed of L-amino acids would be indistinguishable from the projections of a lefthanded α-helix composed of D-amino acids. A clever approach to see what such true enantiomorphs would look like was done by mirroring the entire PDB to create a library of 2.8M D-peptides[40]. These could then be used to create D-analogs of a given peptide, which would resist proteases and could be used pharmacologically.

For natural proteins or synthetic peptides built from L-amino acids, one can thus establish the absolute hand of a reconstruction if the structure of the protomer contains a single α-helix, knowing that this must be right-handed. But β-sheets are more problematic, starting with the problem that one needs to consider whether the twist of a strand is being specified (typically right-handed) or the twist of the sheet (typically left-handed) is being described[41,42]. In any case, while most proteins contain left-handed sheets composed of strands with a right-handed twist, an analysis of the PDB shows the presence of some left-handed twist strands[43], so the overall twist of the sheet may not be 100% reliable in assigning the correct hand to a reconstruction. For one amyloid-forming protein, AFM imaging revealed the presence of both left- and right-handed fibrils[39]. When protein secondary structure is not present determining the absolute hand may be quite problematic, and one may need to use other techniques such as atomic force microscopy (AFM) or cryo-electron tomography to establish the hand. However, the simplest applications of both of these techniques are fairly low resolution, and for relatively thin and smooth nanotubes the helical hand may not be easily seen. AFM image analysis[44] and sub-tomogram averaging[45] may need to be employed in these cases. A study of chirality in nanofibers self-assembled from short amphiphilic peptides used TEM, scanning electron microscopy (SEM), AFM, molecular dynamics simulations (MD) and circular dichroism spectroscopy (CD)[46]. The hand was quite clear from AFM for some of the fibril forms (but not for the thinnest ones), but it did not appear clear from SEM for any of the fibers. Since the TEM images were from negative stain, they (like cryo-EM images) will not show the hand[47]. The CD signals appeared to be determined by the chirality of isoleucine residues, and thus have limited generalizability to other peptide assemblies either in terms of molecular chirality or supramolecular handedness. While we discuss below how symmetry determination may be a very challenging problem when reconstructing polymers, the determination of the absolute hand may actually be more challenging in some cases in the absence of recognizable secondary structure, as in certainly the case for most synthetic soft matter nanotubes. Thus, in the absence of true atomic resolution (~ 1.3 Å or better), directly visualizing the chirality of molecules may be impossible.

There have been other applications of cryo-EM to look at designed protein polymers[48-52] (as opposed to naturally occurring ones), but almost all were at very limited resolutions with likely questions about whether the correct symmetry had been found[12]. The more recent one from Baker and colleagues[50] was done at a near-atomic level of resolution, generating confidence that the symmetry was correct.

4.3- Polymorphism

Most naturally occurring biological polymers, such as actin filaments, microtubules, TMV, etc. exist with a defined symmetry, although filaments such as actin[53,54] and amyloid[55] can have a large

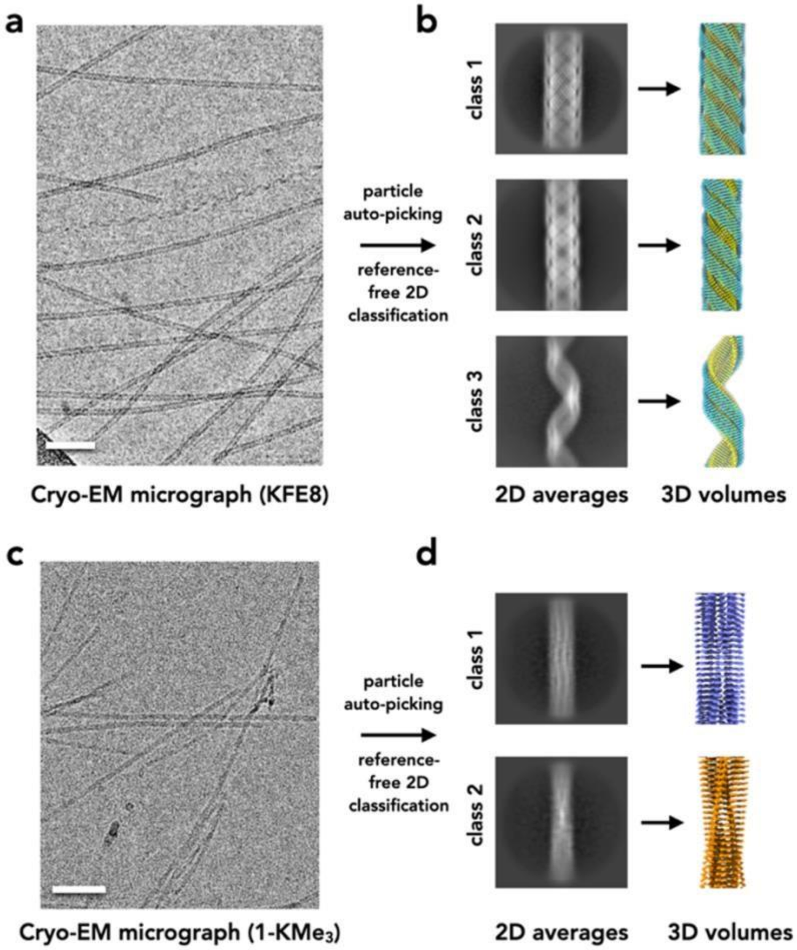


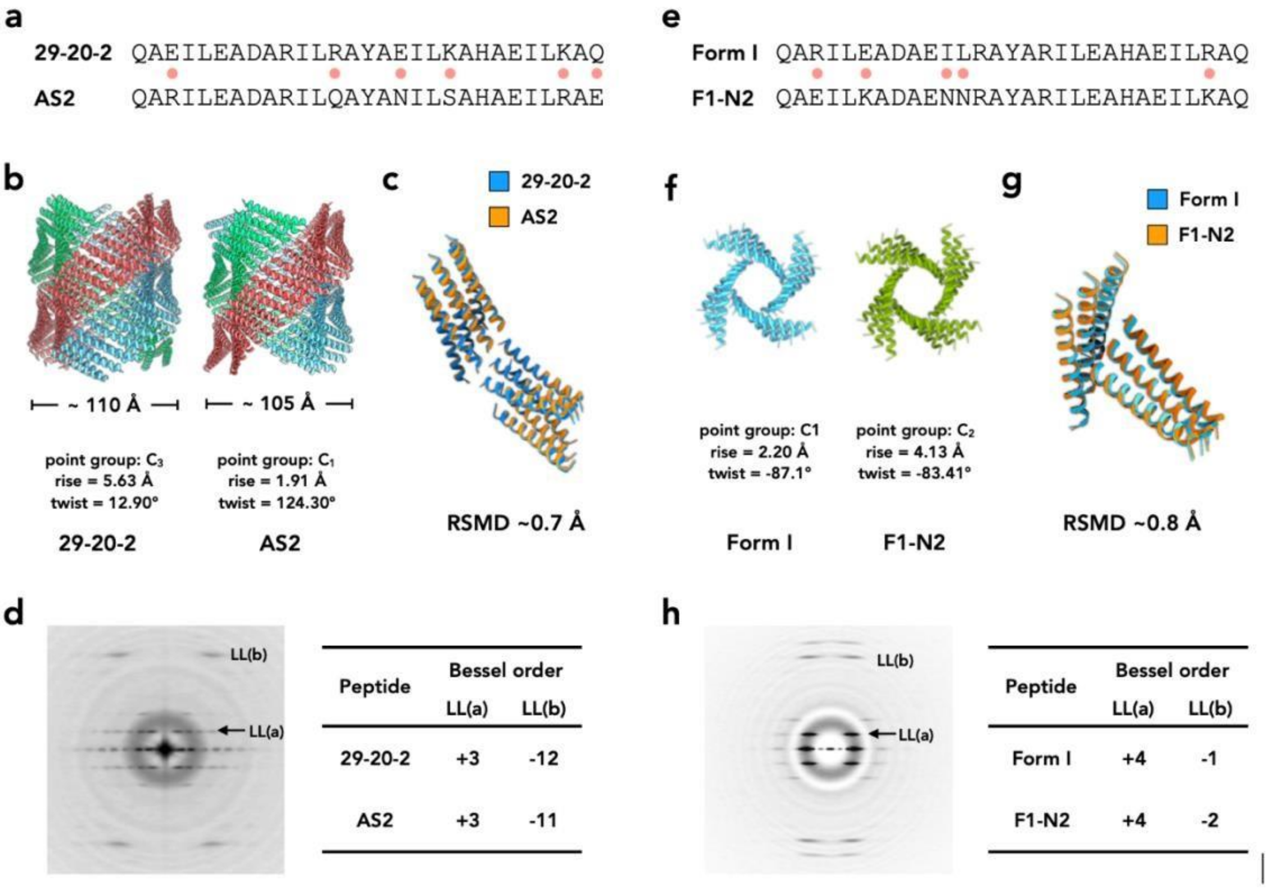
Figure 4-4: **Polymorphism is the norm rather than the exception with many peptide assemblies, and cryo-EM can sort out heterogeneity of polymers.**

(a) & **(c),** representative cryo-EMs of KFE8 nanotube[38] (**a**) and 1-KMe$_3$ nanotube[34] (**c**). Scale bar is 50 nm.

(b) & (**d**), After automatic particle picking and reference-free 2D classfications, helical polymers are grouped into different classes. The differences are obvious by looking at the averages, and helical symmetry determination and high-resolution reconstructions can subsequently be done for each class. In (**b**) class 1 corresponds to tubes with five sheets, class 2 corresponds to tubes formed from four sheets, and class 3 is a ribbon containing a single sheet. In (**d**), class 1 corresponds to a filament with seven peptides per plane, and class 2 corresponds to a filament with six peptides per plane.

variability in twist about that defined symmetry. When microtubules are polymerized *in vitro*, however, they can have a large variability in the number of protofilaments[56], with the number of protofilaments varying within the same microtubule. Similarly, amyloid filaments polymerized *in vitro* have been shown to be highly polymorphic[57] and have raised questions about the validity of these *in vitro* polymers for understanding specific neurodegenerative diseases. When one works with the *in vitro* self-assembly of synthetic peptides, the potential for polymorphism is huge. A recent paper on the assembly of an octapeptide[37] showed that tubes could form with either four or five protofilaments, in addition to ribbons corresponding to a single protofilament **(Figure 4-4-a,b).** Within the context of designed peptide assemblies, this relatively limited degree of polymorphism can actually be seen as relatively well-behaved.

The potential polymorphism often manifests itself as different filamentous morphologies present within a sample that correspond to different helical symmetries (Fig. 4). However, it is also observed that closely related peptides can form nanotubes with similar morphologies, but different symmetries. Despite the fact that the symmetry is different, the interfaces between protomers has been found to be well conserved (**Figure 4-5**). Of course, to show this one must first determine the helical symmetry for each class of filaments. A similar phenomenon was also observed with proteins: for example, when green fluorescent protein (GFP) tags were added to pyrin domains (PYDs), the PYDs assembled into helical polymers with a different helical symmetry from those formed without the tags, but the local environment for the PYDs was quite conserved between both filament forms[58].



Peptide	Bessel order	
	LL(a)	LL(b)
29-20-2	+3	-12
AS2	+3	-11

Peptide	Bessel order	
	LL(a)	LL(b)
Form I	+4	-1
F1-N2	+4	-2

Figure 4-5: Helical polymers made from small subunits can easily adopt different helical symmetries while maintaining relatively conserved interactions.

Two peptide examples are shown here: (a)-(d) nanotube 29-20-2 compared with AS2; (e)-(h) nanotube Form I compared with F1-N2.

(a) & **(e),** The amino acid sequences of the peptides. The dots between the sequences indicate the residues that are non-identical in each of the pairs.

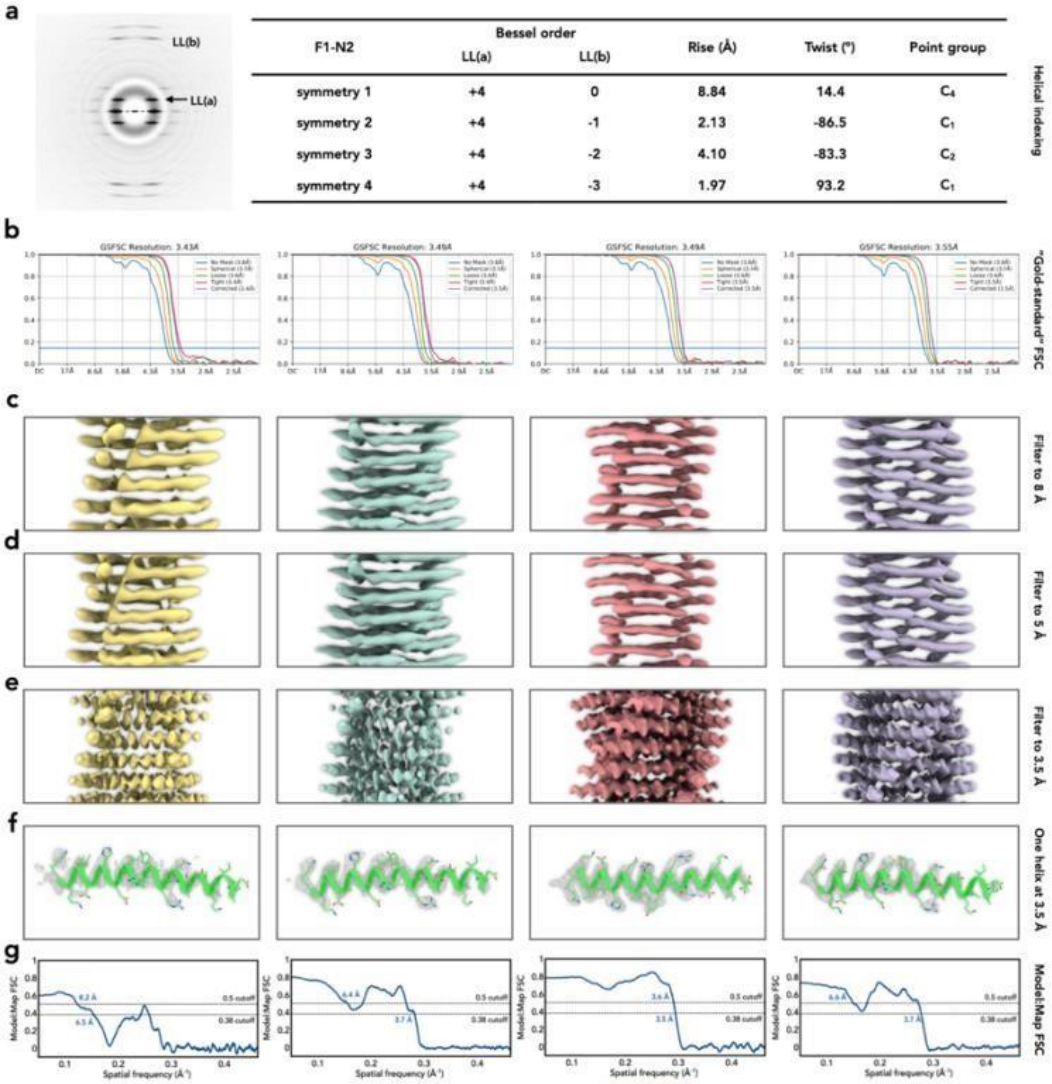
(b) & **(f),** The determined helical symmetry and atomic models of the nanotubes (PDB 7RX4 for AS2 and 7RX5 for F1-N2).

(c) & **(g),** The interface comparison between two nanotubes shows that relatively conserved contacts are maintained, even though the helical symmetry has changed.

(d) & **(h),** The averaged power spectrum (only showing one of two, as they look similar) of the nanotubes, and the helical indexing.

4.4- Symmetry Determination

Perhaps the greatest challenge in reconstructing helical polymers, whether they are native biological assemblies or synthetic *in vitro* constructs, is determining the correct helical symmetry[12,19-22,59]. We now understand that there is no simple metric that allows one to decide whether they have chosen the correct symmetry. For example, it would be nice to be able to say that if the resolution as determined by the FSC (Fourier Shell Correlation) reaches 4.0 Å or better, then one knows that they have the correct symmetry. But this is not the case, and we have seen many examples where the FSC is 4.0 Å or better but the map is uninterpretable, a consequence of using the wrong symmetry. We show **(Figure 4-6)** how four different helical symmetries (only one of which is correct) generate "gold standard" FSC values that are all virtually identical (~ 3.5 Å). This gets to one of the problems with the use of the FSC, as it is not really a measure of true resolution but rather a measure of reproducibility. This is discussed in some detail in an earlier review[60] so we will not dwell on this here. But what becomes obvious is that if the wrong symmetry is imposed, reproducible "detail" can be present at seemingly high resolution that is actually a complete artifact. One can also show that the wrong symmetry will generate a reasonable map in some cases, but the resolution is limited. For example, a reconstruction of a bacterial conjugation pilus reached 5.0 Å using the wrong symmetry[61]. However, with the correct symmetry the resolution obtained was 3.9 Å. A comparison of the two maps at 5.0 Å resolution showed that they were extremely similar, so if higher resolution had not been reached one would have been unable to decide which was the correct symmetry. How is this ambiguity possible? A comparison of the two helical nets[61] shows that the local packing for these two different



F1-N2	Bessel order		Rise (Å)	Twist (°)	Point group
	LL(a)	LL(b)			
symmetry 1	+4	0	8.84	14.4	C_4
symmetry 2	+4	-1	2.13	-86.5	C_1
symmetry 3	+4	-2	4.10	-83.3	C_2
symmetry 4	+4	-3	1.97	93.2	C_1

Figure 4-6: Ambiguities frequently exist in helical symmetry determination.

(a), The averaged power spectrum of nanotube F1-N2, and four possible helical symmetries generated from indexing the power spectrum.

(b), The "gold-standard" FSCs after applying those four different symmetries in helical reconstructions are nearly indistinguishable, despite the fact that three of these symmetries are wrong.

(c)-(e), The reconstruction volumes for those four different symmetries filtered to resolutions of 8 Å (c), 5 Å (d), and 3.5 Å (e). At 5 Å resolution or worse all four symmetries generate maps that might be considered reasonable for α-helical subunits.

(f), The density of a single helix is shown from the 3.5 Å map. The comparison between the density and the atomic model shows that symmetry 3 is the correct one.

(g), The model:map FSC for those four maps with different symmetries and corresponding atomic models. The commonly used arbitrary cutoffs (0.5 and 0.38) are shown. Using the FSC=0.38 criterion (where 0.38 =√ 0.143), only one of these (symmetry 1) might be excluded as obviously incorrect.

symmetries are not that different. Thus, one can align images locally using the incorrect symmetry without doing too much damage to the resolution. However, the further one is from the center of the image the greater the mismatch will be between the actual symmetry and the incorrectly assumed symmetry, which is ultimately what is limiting the resolution. There are some helical structures where one might never reach high resolution due to disorder, heterogeneity, etc., and one thus must be concerned in these cases that the correct symmetry cannot be established with high confidence. What the mating pilus example illustrates is that one must explore all possible choices of symmetry before concluding that the correct one has been found.

So how does one go about determining possible symmetries? The best approach involves starting with an averaged power spectrum from the images (**Figure 4-5-d,h**; **Figure 4-6-a**), which is different from the power spectrum of an averaged image. The averaged power spectrum is an incoherent average that is invariant under the shifts of images that are needed to generate a coherent real-space average. But the process of generating a coherent real-space average can introduce artifacts. For example, if one averages together images using a picture of Albert Einstein as a template, then the average will look like Einstein[62], and the power spectrum of this average will look like the power spectrum of Einstein. However, the averaged power spectrum will show no such artifact, and will actually represent the information in the raw images. Similarly, an average may be based upon the alignment of one very strong feature (such as a long-pitch helix), but other helices in the image are blurred out. The power spectrum of the averaged image will then only show this long-pitch helix, while an averaged power spectrum from the raw images will show features arising from all of the helices.

In rare cases, the averaged power spectrum, combined with a knowledge about the diameter of the filaments[20,21,63], may lead to an unambiguous determination of the helical symmetry. In the more likely case, there will be ambiguities. For example, a helical virus had at least 10 possible symmetries[64], while 1-KMe$_3$ nanotube had > 30 possible symmetries[34]. What can be seen in Fig. 6b is that metrics, such as the putative resolution determined by the FSC, may not correctly distinguish the correct symmetry from the incorrect one. Using atomic models built into the maps, one can only exclude one of the symmetries based upon a poor map:model FSC **(Figure 4-6-g).** But in this test, one first must build an atomic model into the map. These problems make fully automated searches impossible at the moment, and one still needs the active involvement of an observer who knows what a protein or peptide map should look like at different resolutions.

4.5- Chaos

The mathematician and meteorologist Edward Lorenz is widely quoted as having described chaos in the temporal evolution of a system thusly: “When the present determines the future, but the approximate present does not approximately determine the future.” Chaotic systems are therefore those that are incredibly sensitive to small changes. We have suggested that many peptide assemblies behave in such a chaotic manner, as small changes in sequence or assembly conditions can generate huge changes in the structures that are formed[37]. A very striking example of this involves a 29-residue α-helical peptide that self-assembles into helical nanotubes[36]. The semi-conservative mutation of a single arginine to a lysine resulted in: 1) tubes having approximately twice the diameter; 2) the asymmetric unit changing from a single peptide to two peptides, where the two peptides were now in different environments from each other; 3) the structure changing from four stacks of helices to three stacks. While these changes are not mystical or due to supernatural forces, they reflect quaternary interactions that are far beyond our current abilities to predict.

Similarly, while it is known that pH can determine whether a peptide forms parallel or antiparallel β-sheets[65], and that certain peptide sequences will lead to parallel β-sheets while other sequences will lead to antiparallel sheets[66], we have shown that the same peptide will assemble into parallel sheets on the inside of a nanotube and antiparallel sheets on the outside of the same tube[37]. We have called this an instance of "deterministic chaos" and it is yet another example of how the quaternary interactions defy predictability with current knowledge and tools.

4.6- Conclusion

The prior discussion highlights the challenges associated with cryo-EM analysis of helical polymers at near-atomic resolution. However, in order to gain perspective, it is instructive to look back ten years to gauge progress. At that point in time, direct electron detection cameras for electron microscopes were still being developed and were not widely available. These cameras enabled image capture at high detective quantum efficiency, with a vastly improved signal-to-noise ratio, and rapid acquisition times. This development was a necessary prerequisite for the "Resolution Revolution" in cryo-EM. Previously, few structures of biologically derived helical filaments and no structures of designed filaments had been solved to near-atomic resolution using cryo-EM helical reconstruction. The ensuing decade has led to an exponential growth in the number of reported high-resolution structures, but it is clear that the growth curve is still at a very early stage. Currently, high quality cryo-EM data sets can be routinely obtained for synthetic soft materials. The availability of these data sets now allows researchers to address the more challenging technical aspects of cryo-EM analysis, including determination of more effective methods to assign helical hand and determine helical symmetry. These efforts should result in greater numbers of near-atomic resolution structures of synthetic helical filaments. This structural information should facilitate research efforts aimed at the development of methods for reproducible and predictable *de novo* design of selfassembling soft materials. In the short term, research will continue to focus on assemblies derived from sequence-specific chiral oligomers, especially peptide and peptido-mimetic protomers. These substrates have the advantage that sequence-structure correlations have been established that can serve as the input for molecular design efforts. In addition, a huge number of non-peptidic soft matter nanotubes have been designed for many applications, such as optoelectronic materials[67-69]. Very little structural information is available at high resolution for these types of assemblies, but it is clear that this information is critical for understanding the structural features that underlie their technological application and will be critical for materials informatics approaches to optimize function.

4.7- References

(1) Busiek, K. K.; Margolin, W. Bacterial actin and tubulin homologs in cell growth and division. *Current biology : CB* **2015,** *25* (6), R243.

(2) Egelman, E. H. Cryo-EM of bacterial pili and archaeal flagellar filaments. *Curr Opin Struct Biol* **2017,** *46*, 31.

(3) Roca, A. I.; Cox, M. M. The RecA protein: structure and function. *Critical Reviews in Biochemistry and Molecular Biology* **1990,** *25,* 415.

(4) Lu, A.; Magupalli, V. G.; Ruan, J.; Yin, Q.; Atianand, M. K.; Vos, M. R.; Schroder, G. F.; Fitzgerald, K. A.; Wu, H.; Egelman, E. H. Unified Polymerization Mechanism for the Assembly of ASC-Dependent Inflammasomes. *Cell* **2014,** *156* (6), 1193.

(5) Franklin, R. E.; Holmes, K. C. The helical arrangement of the protein subunits in tobacco mosaic virus. *Biochim Biophys Acta* **1956,** *21* (2), 405.

(6) Bernal, J. D.; Fankuchen, I. X-Ray and Crystallographic Studies of Plant Virus Preparations. Iii. *The Journal of general physiology* **1941,** *25* (1), 147.

(7) Crick, F. H.; Watson, J. D. Structure of small viruses. *Nature* **1956,** *177* (4506), 473.

(8) Caspar, D. L.; Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb.Symp.Quant.Biol.* **1962,** *27*, 1.

(9) Egelman, E. H. A tale of two polymers: new insights into helical filaments. *Nat.Rev.Mol.Cell Biol.* **2003,** *4* (8), 621.

(10) DeRosier, D. J.; Klug, A. Reconstruction of three-dimensional structures from electron micrographs. *Nature* **1968,** *217*, 130.

(11) Zheng, W.; Wang, F.; Taylor, N. M. I.; Guerrero-Ferreira, R. C.; Leiman, P. G.; Egelman, E. H. Refined Cryo-EM Structure of the T4 Tail Tube: Exploring the Lowest Dose Limit. *Structure* **2017,** *25* (9), 1436.

(12) Egelman, E. H.; Wang, F. Cryo-EM is a powerful tool, but helical applications can have pitfalls. *Soft Matter* **2021,** *17*, 3291.

(13) Egelman, E. H. The Current Revolution in Cryo-EM. *Biophys J* **2016,** *110* (5), 1008.

(14) Kuhlbrandt, W. Cryo-EM enters a new era. *eLife* **2014,** *3*, e03678.

(15) Kuhlbrandt, W. Biochemistry. The resolution revolution. *Science* **2014,** *343* (6178), 1443.

(16) Egelman, E. H. The iterative helical real space reconstruction method: Surmounting the problems posed by real polymers. *Journal of Structural Biology* **2007,** *157* (1), 83.

(17) Klug, A.; Crick, F. H.; Wyckoff, H. W. Diffraction by helical structures. *Acta Crystallographica* **1958,** *11*, 199.

(18) Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using singleparticle methods. *Ultramicroscopy* **2000,** *85*, 225.

(19) Egelman, E. H. Three-dimensional reconstruction of helical polymers. *Arch Biochem Biophys* **2015,** *581*, 54.

(20) Egelman, E. H. Reconstruction of helical filaments and tubes. *Methods in Enzymology* **2010,** *482*, 167.

(21) Fromm, S. A.; Sachse, C. Cryo-EM Structure Determination Using Segmented Helical Image Reconstruction. *Methods Enzymol* **2016,** *579*, 307.

(22) Egelman, E. H. Ambiguities in Helical Reconstruction. *eLife* **2014,** *3:e04969 doi:10.7554/eLife.04969*.

(23) Frank, J.; Radermacher, M.; Penczek, P.; Zhu, J.; Li, Y.; Ladjadj, M.; Leith, A. SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology* **1996,** *116* (1), 190.

(24) Frank, J.; Shimkin, B.; Dowse, H. SPIDER - A modular software system for electron image processing. *Ultramicroscopy* **1981,** *6*, 343.

(25) He, S.; Scheres, S. H. W. Helical reconstruction in RELION. *J Struct Biol* **2017,** *198* (3), 163.

(26) Punjani, A.; Zhang, H.; Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nature Methods* **2020,** *17* (12), 1214.

(27) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **2017,** *14* (3), 290.

(28) Shimizu, T. In *Smart Soft-Matter Nanotubes: Preparation, Functions, and Applications*; Springer Singapore: Singapore, 2021, DOI:10.1007/978-981-16-2685-2_1 10.1007/978981-16-2685-2_1.

(29) Nakane, T.; Kotecha, A.; Sente, A.; McMullan, G.; Masiulis, S.; Brown, P.; Grigoras, I. T.; Malinauskaite, L.; Malinauskas, T.; Miehling, J.et al. Single-particle cryo-EM at atomic resolution. *Nature* **2020,** *587* (7832), 152.

(30) Yip, K. M.; Fischer, N.; Paknia, E.; Chari, A.; Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **2020,** *587* (7832), 157.

(31) Wang, F.; Gu, Y.; O'Brien, J. P.; Yi, S. M.; Yalcin, S. E.; Srikanth, V.; Shen, C.; Vu, D.; Ing, N. L.; Hochbaum, A. I.et al. Structure of Microbial Nanowires Reveals Stacked Hemes that Transport Electrons over Micrometers. *Cell* **2019,** *177* (2), 361.

(32) Shu, C.; Xiao, K.; Cao, C.; Ding, D.; Sun, X. Predicting and Interpreting the Structure of Type IV Pilus of Electricigens by Molecular Dynamics Simulations. *Molecules* **2017,** *22* (8).

(33) Wang, F.; Gnewou, O.; Modlin, C.; Beltran, L. C.; Xu, C.; Su, Z.; Juneja, P.; Grigoryan, G.; Egelman, E. H.; Conticello, V. P. Structural analysis of cross alpha-helical nanotubes provides insight into the designability of filamentous peptide nanomaterials. *Nature Communications* **2021,** *12* (1), 407.

(34) Feng, Z.; Wang, H.; Wang, F.; Oh, Y.; Berciu, C.; Cui, Q.; Egelman, E. H.; Xu, B. Artificial Intracellular Filaments. *Cell Rep Phys Sci* **2020,** *1* (7).

(35) Hughes, S. A.; Wang, F.; Wang, S.; Kreutzberger, M. A. B.; Osinski, T.; Orlova, A.; Wall, J. S.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Ambidextrous helical nanotubes from selfassembly of designed helical hairpin motifs. *Proc Natl Acad Sci U S A* **2019,** *116* (29), 14456.

(36) Egelman, E. H.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.; Baker, D.; Conticello, V. P. Structural plasticity of helical nanotubes based on coiled-coil assemblies. *Structure* **2015,** *23* (2), 280.

(37) Wang, F.; Gnewou, O.; Wang, S.; Osinski, T.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Deterministic chaos in the self-assembly of β-sheet nanotubes from an amphipathic oligopeptide. *Matter* **2021**, DOI:10.1016/j.matt.2021.06.037 10.1016/j.matt.2021.06.037.

(38) Wang, F.; Gnewou, O.; Wang, S.; Osinski, T.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Deterministic chaos in the self-assembly of β-sheet nanotubes from an amphipathic oligopeptide. *Matter*, DOI:10.1016/j.matt.2021.06.037 10.1016/j.matt.2021.06.037.

(39) Wu, X.; Ma, Y.; Zhao, K.; Zhang, J.; Sun, Y.; Li, Y.; Dong, X.; Hu, H.; Liu, J.; Wang, J.et al. The structure of a minimum amyloid fibril core formed by necroptosis-mediating RHIM of human RIPK3. *Proc Natl Acad Sci U S A* **2021,** *118* (14).

(40) Garton, M.; Nim, S.; Stone, T. A.; Wang, K. E.; Deber, C. M.; Kim, P. M. Method to generate highly stable D-amino acid analogs of bioactive helical peptides using a mirror image of the entire PDB. *Proc Natl Acad Sci U S A* **2018,** *115* (7), 1505.

(41) Chothia, C. Conformation of twisted beta-pleated sheets in proteins. *J Mol Biol* **1973,** *75* (2), 295.

(42) Salemme, F. R. Structural properties of protein beta-sheets. *Prog Biophys Mol Biol* **1983,** *42* (2-3), 95.

(43) Caudron, B.; Jestin, J.-L. Geometrical criteria for left-handed twists within protein betastrands. *Journal of Biophysical Chemistry* **2014,** *Vol.05No.01,* 8.

(44) Heath, G. R.; Kots, E.; Robertson, J. L.; Lansky, S.; Khelashvili, G.; Weinstein, H.; Scheuring, S. Localization atomic force microscopy. *Nature* **2021,** *594* (7863), 385.

(45) Pyle, E.; Zanetti, G. Current data processing strategies for cryo-electron tomography and subtomogram averaging. *The Biochemical journal* **2021,** *478* (10), 1827.

(46) Wang, M.; Zhou, P.; Wang, J.; Zhao, Y.; Ma, H.; Lu, J. R.; Xu, H. Left or Right: How Does Amino Acid Chirality Affect the Handedness of Nanostructures Self-Assembled from Short Amphiphilic Peptides? *Journal of the American Chemical Society* **2017,** *139* (11), 4185.

(47) Egelman, E. H.; Amos, L. A. Electron microscopy of helical filaments: rediscovering buried treasures in negative stain. *BioEssays* **2009,** *31* (9), 909.

(48) Li, Z.; Chen, S.; Gao, C.; Yang, Z.; Shih, K.-C.; Kochovski, Z.; Yang, G.; Gou, L.; Nieh, M.-P.; Jiang, M.et al. Chemically Controlled Helical Polymorphism in Protein Tubes by Selective Modulation of Supramolecular Interactions. *Journal of the American Chemical Society* **2019,** *141* (49), 19448.

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

(50) Shen, H.; Fallas, J. A.; Lynch, E.; Sheffler, W.; Parry, B.; Jannetty, N.; Decarreau, J.; Wagenbach, M.; Vicente, J. J.; Chen, J.et al. De novo design of self-assembling helical protein filaments. *Science* **2018,** *362* (6415), 705.

(51) Rao, G.; Fu, Y.; Li, N.; Yin, J.; Zhang, J.; Wang, M.; Hu, Z.; Cao, S. Controllable Assembly of Flexible Protein Nanotubes for Loading Multifunctional Modules. *ACS Appl Mater Interfaces* **2018,** *10* (30), 25135.

(52) Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X.et al. Precise and Reversible Protein-Microtubule-Like Structure with Helicity Driven by Dual Supramolecular Interactions. *Journal of the American Chemical Society* **2016,** *138* (6), 1932.

(53) Egelman, E. H.; Francis, N.; DeRosier, D. J. F-actin is a helix with a random variable twist. *Nature* **1982,** *298*, 131.

(54) Galkin, V. E.; Orlova, A.; Vos, M. R.; Schroder, G. F.; Egelman, E. H. Near-atomic resolution for one state of f-actin. *Structure* **2015,** *23* (1), 173.

(55) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017,** *547* (7662), 185.

(56) Chretien, D.; Metoz, F.; Verde, F.; Karsenti, E.; Wade, R. H. Lattice defects in microtubules: protofilament numbers vary within individual microtubules. *Journal of Cell Biology* **1992,** *117* (5), 1031.

(57) Zhang, W.; Falcon, B.; Murzin, A. G.; Fan, J.; Crowther, R. A.; Goedert, M.; Scheres, S. H. Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *eLife* **2019,** *8*.

(58) Lu, A.; Li, Y.; Yin, Q.; Ruan, J.; Yu, X.; Egelman, E. H.; Wu, H. Plasticity in PYD assembly revealed by cryo-EM structure of the PYD filament of AIM2. *Cell Discovery* **2015,** *1*, 15013.

(59) Yu, X.; Egelman, E. H. Helical filaments of human Dmc1 protein on single-stranded DNA:

a cautionary tale. *Journal of Molecular Biology* **2010,** *401* (3), 544.

(60) Subramaniam, S.; Earl, L. A.; Falconieri, V.; Milne, J. L.; Egelman, E. H. Resolution advances in cryo-EM enable application to drug discovery. *Curr Opin Struct Biol* **2016,** *41*, 194.

(61) Zheng, W.; Pena, A.; Low, W. W.; Wong, J. L. C.; Frankel, G.; Egelman, E. H. CryoelectronMicroscopic Structure of the pKpQIL Conjugative Pili from Carbapenem-Resistant Klebsiella pneumoniae. *Structure* **2020,** *28* (12), 1321.

(62) Henderson, R. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proc Natl Acad Sci U S A* **2013,** *110* (45), 18037.

(63) Desfosses, A.; Ciuffa, R.; Gutsche, I.; Sachse, C. SPRING - an image processing package for single-particle based helical reconstruction from electron cryomicrographs. *J Struct Biol* **2014,** *185* (1), 15.

(64) DiMaio, F.; Yu, X.; Rensen, E.; Krupovic, M.; Prangishvili, D.; Egelman, E. H. A Virus that Infects a Hyperthermophile Encapsidates A-Form DNA. *Science* **2015,** *348*, 914.

(65) Tseng, W.-H.; Chen, S.-H.; Hiramatsu, H. pH-controlled stacking direction of the βstrands in peptide fibrils. *Scientific Reports* **2020,** *10* (1), 22199.

(66) Lifson, S.; Sander, C. Antiparallel and parallel β-strands differ in amino acid residue preferences. *Nature* **1979,** *282* (5734), 109.

(67) Achyuthan, K. E.; Lu, L.; Lopez, G. P.; Whitten, D. G. Supramolecular photochemical selfassemblies for fluorescence "turn on" and "turn off" assays for chem-bio-helices. *Photochem Photobiol Sci* **2006,** *5* (10), 931.

(68) Eisele, D. M.; Cone, C. W.; Bloemsma, E. A.; Vlaming, S. M.; van der Kwaak, C. G. F.; Silbey, R. J.; Bawendi, M. G.; Knoester, J.; Rabe, J. P.; Vanden Bout, D. A. Utilizing redoxchemistry to elucidate the nature of exciton transitions in supramolecular dye nanotubes. *Nature chemistry* **2012,** *4* (8), 655.

(69) Caram, J. R.; Doria, S.; Eisele, D. M.; Freyria, F. S.; Sinclair, T. S.; Rebentrost, P.; Lloyd, S.; Bawendi, M. G. Room-Temperature Micron-Scale Exciton Migration in a Stabilized Emissive Molecular Aggregate. *Nano Lett* **2016,** *16* (11), 6808.

Chapter 5: Insight into the mechanism of formation of a β-sheet nanotube from an amphipathic oligopeptide

5.1- Introduction

Peptides assemblies can emulate a diversity of functions that include controlled release and delivery, cargo transport, locomotion, energy transduction, and actuation[1-7]. This degree of functional activities is highly desirable and probably necessary for the understanding and improving upon the complex functions observed in native biological assemblies[8-12], e.g., extracellular protein filaments and filamentous viruses. Among the many different types of assemblies, cross-β amyloid-like assemblies have been extensively studied over the years. The high level of scrutiny can be attributed to the fact that the association of small peptides and proteins molecules into cross-β amyloid has been correlated with pathologies that include Alzheimer's disease, Creutzfeldt−Jakob disease, and type II diabetes field[13, 14] leading to them having conserved, functional motif in many organisms[15, 16]. Studies focusing on β-sheet fibril structures have uncovered that many amyloidogenic peptide sequences display high levels of structural polymorphism[17]. This situation presumably arises as a consequence of the cross-β filament structure[18-21], which primarily relies on non-sequence-specific hydrogen bonding interactions between main-chain donors and acceptors to stabilize the structural interfaces between protomers[22]. Others have looked at the pathway leading to the formation and maturation of the resulting filaments.

Single (1SN) and double (2SN) step nucleation has mostly been used as models to rationalize assembly formation. In the single nucleation model (1SN)[23-25], it is assumed that individual peptide monomers directly nucleate in solution to form the resulting structure. The double step model (2SN)[26-30], however, includes the presence of an intermediate step in which a metastable particle phase is formed prior to fibrillation[31]. Supported by Ostwald's rule of stages[32], the aggregation of the peptide oligomers inside a particle-like environment has been adopted as precursors for the first nucleation step, which can then proceed to form the final fibrillar structure through 3 potential pathways: They can either (a) be incorporated as the assembly phases developed[33-35], (b) coexist with the assembled phases[36, 37], or (c) persist if assembly nucleation is slow or thermodynamically disfavored[38, 39] until the formation of the final fibers[30]. It is increasingly important to understand both the structural interactions that govern their assemblies as well as the kinetic and thermodynamic parameters that direct the pathway of assembly in order to aid in the construction of functional biomaterials[40].

We recently reported the structural analysis of an amphipathic oligopeptide Ac-FKFEFKFE-NH2 (KFE8)[41]. It was shown to form left-handed bilayer helical ribbons in which face-to-face packing of antiparallel β-sheets were proposed for the underlying structure. We found that these ribbons coalesced into helical nanotubes, slowly at ambient temperature and rapidly upon heating. Two distinct populations of nanotubes were observed based on the presence of four or five protofilaments, respectively. Both the four-protofilament and five-protofilament tubes were shown to be based on a bilayer structure as a consequence of the facially amphipathic sequence[42] of KFE8 in a β-strand conformation. The structure of the five-protofilament tubes is based on the coalescence of five individual left-handed ribbons. In contrast, the structure of the four-protofilament tube is based on the interaction between two paired protofilaments that associate to break higher-order symmetry. Furthermore, we uncovered that, in either case, the inner layer is based on parallel β-sheets while the outer layer is based on antiparallel β-sheets. We independently confirmed the presence of this unusual packing arrangement in the helical ribbons that we assumed served as precursors to the nanotubes.

Here, we sought to understand the molecular self-assembly mechanism of the previously characterized cross-β nanotubes using complementary experimental techniques. We first utilized different time-resolved spectroscopic approaches such as static light scattering to interrogate the kinetics associated with the quaternary structure variation of assemblies under different conditions. We then took advantage of structural characterization techniques such as atomic force microscopy to gain more insight into the different steps that occur during nanotube formation until the equilibrium is reached. We ultimately demonstrate the presence of an additional population of filaments that appear to be precursors to the formation of ribbons that later coalesce in the different diameter nanotubes described above.

5.2- Results

5.2.1- There are multiple different populations of assemblies over time

AFM imaging of the two samples of peptide KFE8, prepared using either the thermal annealing and room temperature assembly processes, provided another illustration of the distinct assemblies population **(Figure 5-1).** As previously observed by electron microscopy, a thinner and thicker tube with relative abundance depending on the assembly conditions (i.e., annealed or at room temperature) can be observed. In addition to these fully formed tubular filaments, ribbon-like assemblies can be detected after overnight incubation at ambient temperature, which was compatible with the previously observed assembly behavior of KFE8 using electron microscopy **(Figure 5-1).** Surprisingly, we also detected the presence of additional protofilaments of various widths in the background of the AFM imaged under annealing conditions **(Figure 5-1-c)** and at about 5h after assembly at room temperature **(Figure 5-1-d)**. These protofilaments had not been previously observed using electron microscopy, which is most likely due to both their small diameter and lower contrast in comparison to the wider diameter tubes and ribbons. AFM imaging indicated the presence of heterogeneity as revealed through the detection of multiple species that

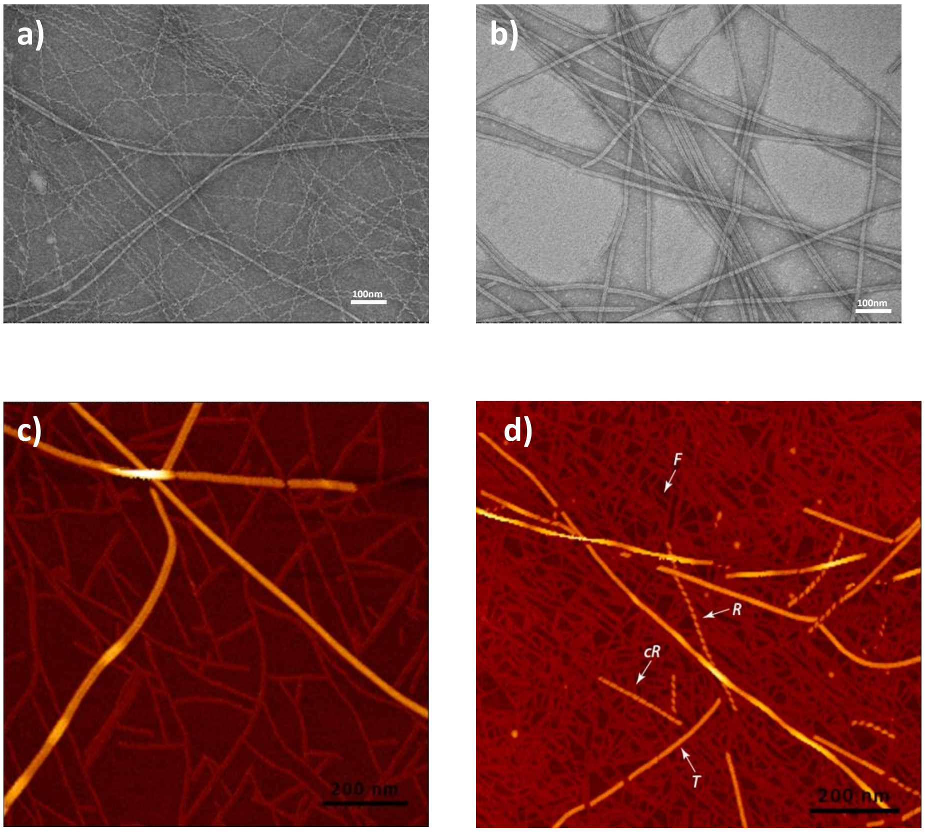


Figure 5-1: TEM and AFM of the assemblies

TEM micrographs of the FKE8 peptide assembled and imaged after 24h **(a)** room temperature, **(b)** annealed. AFM images of the FKE8 peptide assembled and imaged after 24h **(c)** annealed and **(d)** room temperature with arrow showing the different population of tubes: F for filaments, R for Ribbons, T for tubes and Rc for condensed ribbons.

might be struccturally related through conversion processes. As illustrated in **Figure 5-1-d,** AFM analysis at early stages of the self-assembly show different populations with extensions at both extremities of the tube: One bearing protofilaments on one end while the other has ribbons. All these observations indicate the existence of a multiple structuration process of this peptide.

The above observations led us to hypothesize that the lowest energy conformation (nanotubes) deriving from the coalescence of ribbons occurred as a consequence of the self-association of protofilaments present in the background of the AFM images in **Figure 5-1**. This suggests the presence of a nucleation process in their pathway of formation, probably originating fron the dots observable on the micrograph **(Figure 5-1-d).**

5.2.2- Conformational rearrangement of the peptides

In order to explore the time scale of the formation of these different sets of assemblies, we followed the static light scattering (SLS) signal at 450 nm as a function of time for two distinct initial peptide concentrations (Figure 5-2-a,100 and 430 μM). This experiment was conducted under slight heating conditions (37 °C) to speed up the self-assembly process. From these results, we can note two interesting behaviors. First, for both concentrations, the kinetic curved did not reveal a lag time. We can see an immediate rise of the signal from the initial time of assemblies up to 6 h after, followed by a stabilization of the curves indefinitely. This could either indicate the presence of a limiting nucleation step that is too fast to be captured by the signal, or the absence of such step altogether. The second behavior relates to the rate of increase of the SLS signal. The rate of formation of the final nanotube appears to be quasi-independent of peptide concentration, which suggests that the formation of assemblies can be described using a pseudo-first-order kinetic

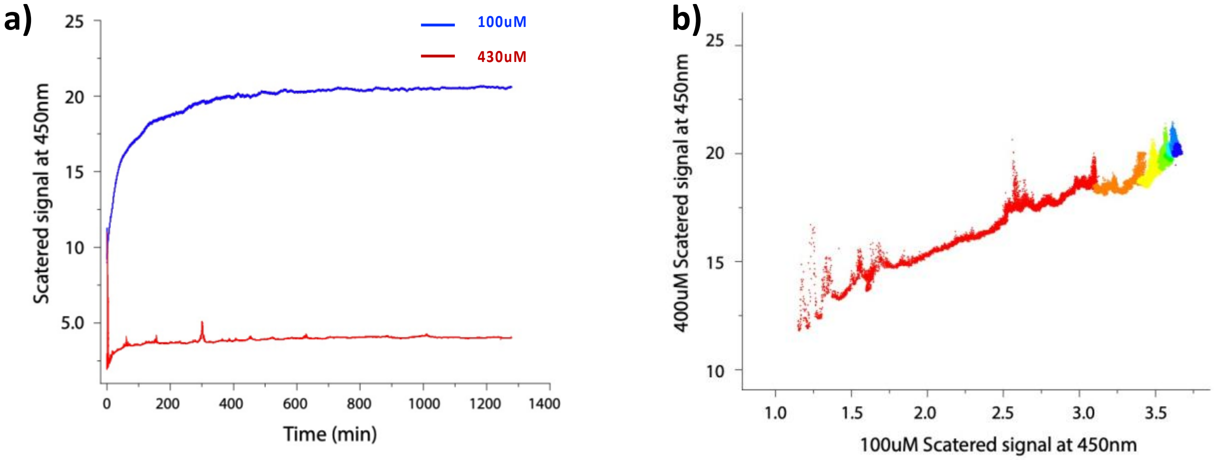


Figure 5-2: Static Light Scattering experiment

(a) Static light scattering (SLS) Kinetics of the FKE8 pepides at 2 different concentration (100uM and 430uM). The scattering intensity is measured at 450nm as a function of time. **(b)** Scattering signal time correlation between 100uM and 430uM approximation. Moreover, once the kinetic curves are normalized and the scattering signals are correlated, the fitting is almost perfect. **(Figure 5-2-b)**

These findings raised the question of whether the structural rearrangement of the β-sheets could be detected using circular dichroism spectroscopy. To do so, we decided to follow the same conditions as in the SLS experiments by incubating the sample at 37°C in other to maintain consistency. We then took out a sample at different time intervals (T = 0 h, 15 min, 1 h, 3.5 h, and 24h, 2 days) and combined the graphs to visualize the evolution of the assemblies. Circular dichroism results revealed that a spectroscopic change, presumably corresponding to a conformational transition during self-assembly, could be correlated to a morphological change within the peptide assemblies. We could distinguish a near-isodichroic point at 195 nm indicative

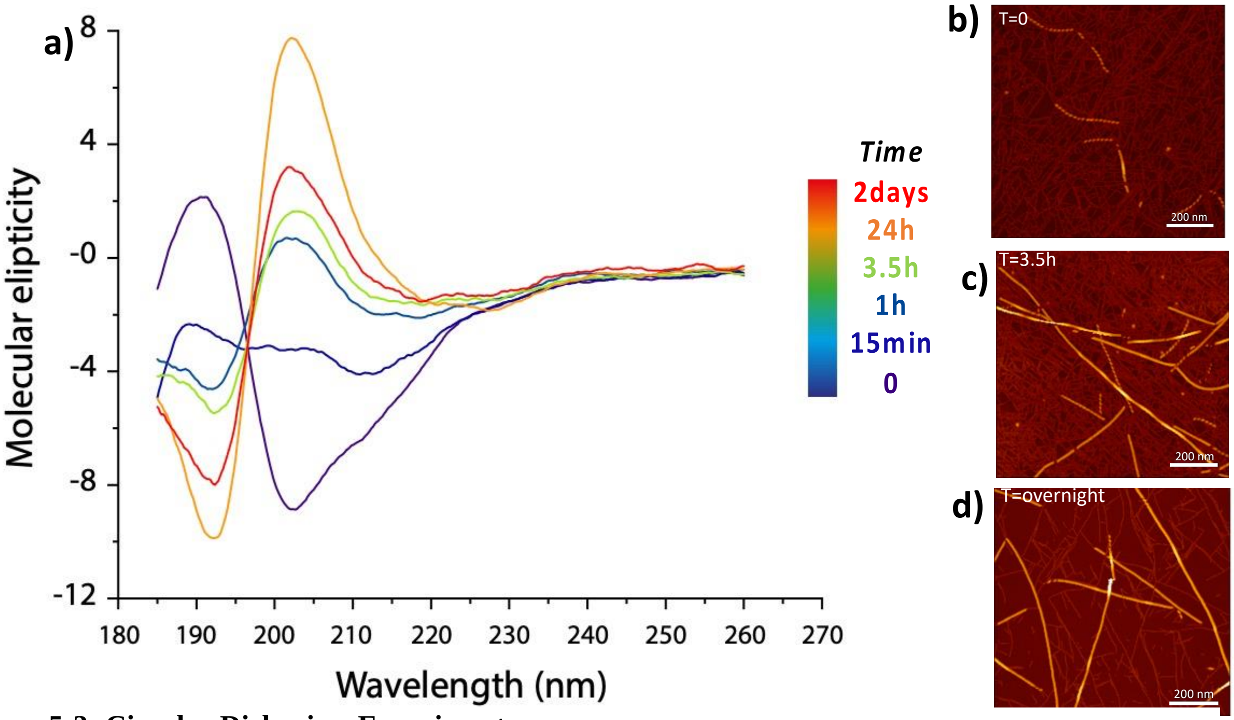


Figure 5-3: Circular Dichroism Experiment

Right side: (a) Circular dichroism (CD) spectroscopy of the FKE8 peptide showing the change in quaternary structures of the assemblies over the course of 2 days. Left side: AFM images of sample taken at specific time point to verify the state of the peptides (b) at t = 0, (c) at t = 3.5 h, and (d) 18h after assembly of the presence of only two species at equilibrium. In addition, we observed a shift from negative to positive signal at 205 nm, with the reverse spotted at 225 nm. These tracers display an atypical representation of β-sheet that indicates a change in quaternary structure within the representative assemblies. The variation of molecular ellipticity as a function of time suggests a β structuration of the peptide starting at t = 0, which undergoes a secondary structural rearrangement as a function of time compatible with the formation of a highly twisted antiparallel β-sheet **(Figure 5-2-c).** This structural rearrangement could also be visualized by AFM. We imaged the sample 3 times (T= 0 h, 3.5 h, and 18h as they were assumed to be intervals at which we could observe the most difference between the assemblies. The t = 0 h images show a vast amount of protofilaments on top of which a small number of ribbons can be seen. At t = 3.5 h, there is the coexistence of protofilaments in large amounts in the background with ribbons and tubes simultaneously. The final overnight sample no longer has ribbons but still contains a small number of protofilaments with fully formed tubes of different diameters **(Figure 5-3).** The density of the protofilaments in the background seems to be reducing over time, as seen in the background of **Figure 5-3** when comparing images b and d. This indicates that the protofilaments are being consumed and succinctly rearranged over the first 6 hours into ribbons and then tubes. The presence of a signal at t = 0 h, however, raised the question of whether or not a rapid initial nucleation step led to the protofilaments at an early stage of the process.

5.2.3- Reversibility of the assembly process until equilibrium

Before determining if the sample's size evolution is associated with an increase in β-sheet content, we needed to clarify the time at which the protofilaments visible in the background started appearing and rearranging. We monitored the secondary structure of the lyophilized peptide in the solid powder form, that is, prior to solution self-assembly, using solid-state Fourier Transform

Infrared (FTIR) spectroscopy. As shown in Figure 5-4-a, the FTIR spectra reveal the presence of unique peaks in the amide I region between 1610 and 1649cm^{-1} with 1633 cm^{-1} more explicitly corresponding to the presence of β-sheet with some turns between 1660 and 1685 cm^{-1} (1680 cm^{-1}). Interestingly the result of a time point FT_IR study performed on the pellet from ultracentrifugation of the sample at different timepoint suggests that the β-sheet content of the sample does not significantly evolve as a function of time, indicating no variation of β-sheet content during the polymerization (**Figure 5-5-1**). This last observation was confirmed through fluorescence ThT measurements. No change in ThT fluorescence was observed during polymerization (**Figure 5-4-b),** indicating the absence of *de novo* formation of β-sheet structures.

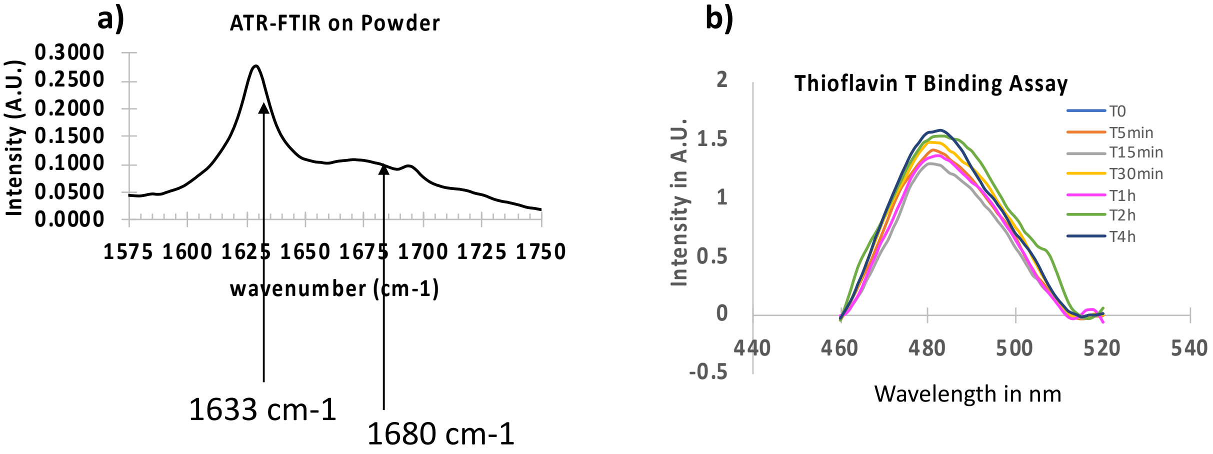


Figure 5-4: FT-IR and ThT binding experiments

(a) Solid-state FT-IR of the FKE8 peptide (in powder) showing the presence of β-sheet structure. (b) fluorescence spectroscopy experiment monitoring the binding of thioflavin T to the FKE8 peptide over a period of 4hours

Moreover, AFM imaging of the sample at different time points in the assembly process supported that the protofilaments were indeed precursors to the ribbons and tubes. **Figure 5-5** reveals snapshots of assemblies incubated at room temperature at 37°C at two different times: a) 2h and b) 5h, and after assemblies. They captured very specific events that are most likely leading to the formation of the thinner diameter nanotube. The first one (**Figure 5-5-a)** illustrates a heterogeneous mixture comprised of a fully formed tube and what appears to be a pair of doubled wall protofilaments twisting into a tube. The second one (b) shows an intra-heterogenous mix of three protofilaments of different thicknesses coming together to close the ribbon into tubes. Since the cryo-EM model previously identified the presence of predominantly thinner diameter nanotubes shown to derive from the coupling of 2 ribbons, **Figure 5-5-b** raises the question of the make-up of protofilaments precursors. They could be made of a single doubled wall stack or a pair doubled wall stack of β-sheets. In the latter case, we might be observing the folding of one pair of double wall stackstacks with two single double wall stack to form what we observed by cryo-EM to be a thinner tube

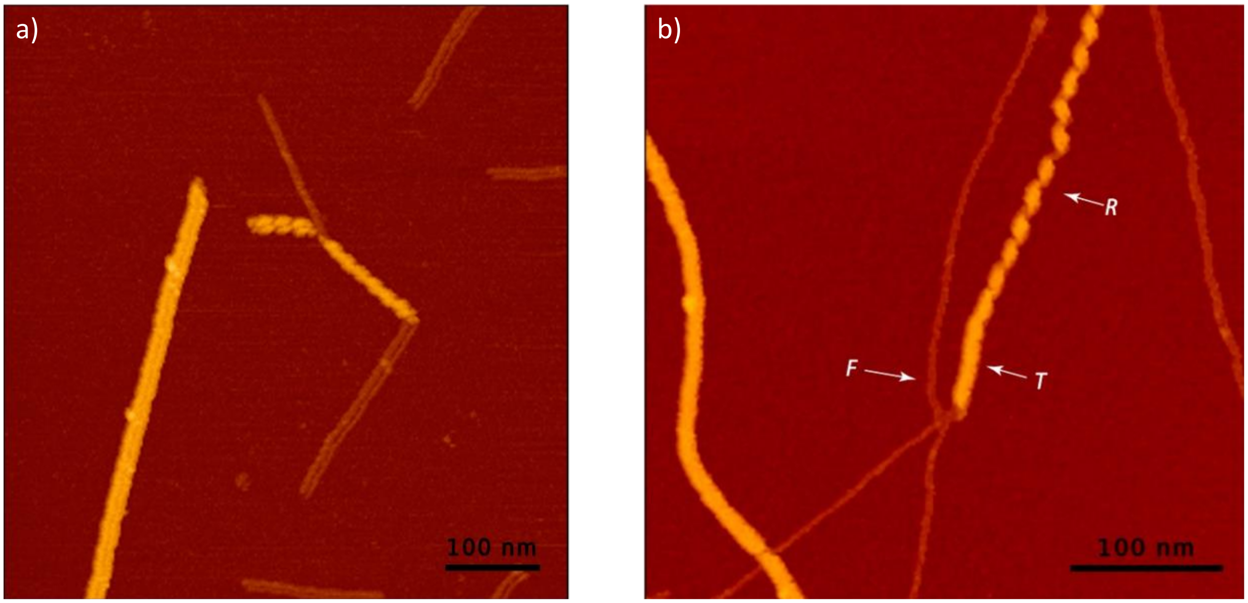


Figure 5-5: AFM of the sample at different times illustrating the sequence of events leading to the tube formation

AFM images of FKE8 peptides left to incubate at room temperature and imaged at different time points: 2h and 5h. Respectively, (a) Heterogenous mix of protofilaments int ribbons and separated fully formed tubes, (b) Intra heterogenous assembly with protofilaments closing the ribbons into the tubes.

All these observations strongly reinforced that i) the size variation corresponds to a very large assemblies' condensation with a pseudo-first-order kinetic **(Figure 5-1).** ii) a probable presence of preformed assemblies that eliminates the existence of a nucleation event leading to the winding of the protofilaments into ribbons.

To decipher the mechanism of the formation of the different populations of self-assembled species and identify the structure of the most stable species, we followed the evolution of quaternary structure by SLS through pulsed temperature variation. The temperature of the sample preparation was raised from 0 to 85 °C at a rate of one degree per minute, held at 85 °C for 5 min, and cooled back down at the same rate followed by another hold of 5 min **(Figure 5-5).** Exposing the sample to this cycle over the course of 24 h showed that the size evolution of the assemblies can be divided into two phases. The first phase, in which the size of the assemblies increases as a function of the temperature, indicates an endothermic structural rearrangement contributing to an increase in the molecular weight of the quaternary structure. Then, a second phase in which the size of the assemblies successively increases and decreases as the temperature continue to vary, pointing to a decrease in size or subsequent unraveling of the assemblies. Moreover, as shown in **Figure 5-5-a-b,** the rate of the temperature increase and decrease temperature increase and decrease rate does not significantly affect the two phases, indicating that the structural rearrangement concerted with size variation is relatively fast.

To determine which types of quaternary structure are responsible for the observed size variation by SLS, we analyzed the morphology of the assemblies immediately after peptide solubilization and at the end of the annealing protocol described in the experimental section. As shown infigure 5-6 while after heating until 85 °C and slowly cooling down to room temperature, the peptide forms tubular quaternary structures **(Figure 5-6-d)** with a width of either 7.5 nm or 8.4 nm in diameter; two types of assemblies are observed as expected before the cycle of heating and cooling: the protofilaments and the ribbons **(Figure 5-6-c)**.

Altogether these observations indicate that the protofilaments formed immediately after peptide solubilization and undergo an endothermic structural rearrangement through the ribbons to form tubes. Furthermore, the cyclic size variation during the second phase indicates that temperature induces an unfolding process. This unfolding process could correspond to a protomer/peptide release hat is then recaptured during the refolding (temperature decrease).

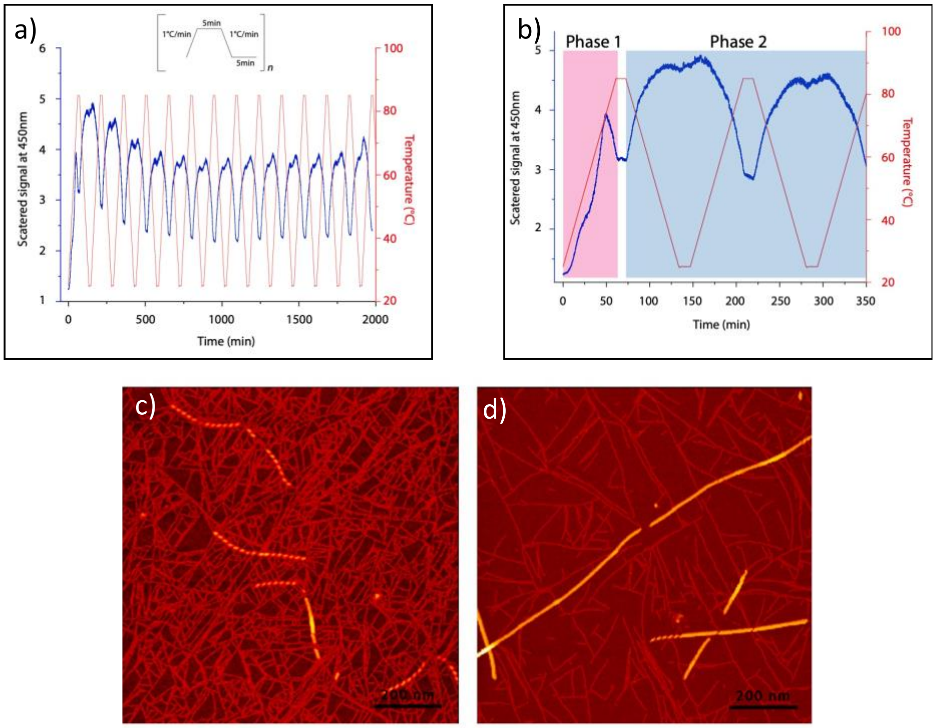


Figure 5-6: **Temperature ramping experiment**

a) Scattered intensity signal as function of periodical temperature variation. The variation of the mean average molecular weight of the peptides FKFEFKFE submitted to periodical temperature perturbation as indicated in the figure insert. **(b)** A zoom of the early stage of periodic temperature ramping. As shown the early stage of the temperature ramping can be split in two phases. The first phase (in pink) which corresponds to an endothermic irreversible polymerization process gives raise to the formation of assemblies which reversibly polymerize and depolymerize respectively at low and high temperature. **(c)** and **(d)** are respectively images of the sample (430uM in acetate ph 4) taken right before the annealing process at T0 and at the end~ T=33h

5.3- Discussion

Previous studies have established that the formation of fibers for amyloid-like assemblies occurs via a pre-fibrillation stage[43-45]. Our findings suggest the presence of partially folded assemblies containing β-sheet structures as intermediates to the final observable ribbons and nanotubes. Such behavior has been observed before in the context of prion protein PrP[46, 47], transtyretin[48, 49], and many others. There is, however, a rich history of studies supporting the path according to which early aggregates will directly nucleate into the formation of fibrils. We encountered a couple of limitations in adopting that theory to our system. First is the absence of a lag time, which indicates the nonexistence of a nucleation-elongation stage at which a low aggregate concentration is usually detected. Instead, we see new assemblies' growth rapidly until equilibrium. The central challenge will be to determine whether or not primary nucleation is taking place and its importance in the overall kinetics of the system.

Second is the fact that the proliferation of the product seems to be happening immediately after hydration of the sample. Because of its relatively low solubility, the sample needs to be mixed using a vortex mixer for a few seconds to dissolve in the buffer solution fully. The addition of mechanical stress to the system can dramatically decrease the lag phase[50], hence reducing the mechanical stability of fibrils. Another remaining question is the characterization of the already present oligomers. This is of particular relevance as understanding the mechanism of formation of these oligomers could provide insight into the lack of a lag time for this specific case but also shine light on the different sized of protofilaments observed at the early stages of the process. Finally, it will be essential to use the hereby knowledge to understand which molecular events take place at which phase still needs to be examined.

5.4- Conclusion

Several amyloidogenic sequence fragments have been isolated, assembled, and studied over the years. The studies have involved structural and mechanistic elements in aiding in our understanding of how self-assembly occurs. Over the past decade, advances have been made toward a complete description of the fibril formation process, from the monomers to oligomers to polymers. In this study, we show that the fibril formation in the FKE8 case can be reversed (disassembly process) upon variation of the temperature but eventually reach equilibrium. Additionally, we demonstrate that the self-assembly process does not always start from a fully denatured monomer.In contrast, a strong β-sheet content can be detected as early as when the peptide is in the powder form. In our case, the already present β structure assembles directly into nanotubes or will form ribbons that will coalesce into a tube of bigger diameter over time. Because we still have very little understanding of how these initial protofilaments even form, it will be essential to gain more structural information on the initial protofilament to understand the overall process better.

5.5- Materials and methods

All chemical reagents and peptides ordered were used without further purification. The peptide FKFEFKFE was obtained from GenScript (Shanghai, China).

5.5.1- Peptide synthesis

The peptide FKFEFKFE was prepared via microwave-assisted solid phase peptide synthesis on a CEM Liberty Blue Automated Microwave Peptide Synthesizer as the *N*-acetyl, *C*-amide capped derivatives. A PAL-PEG-PS resin from Applied Biosystems (Foster City, CA) was used. Standard Fmoc protection chemistry was utilized in conjunction with coupling cycles consisting of OXIMA-mediated activation protocols and base-induced deprotection (20%

piperidine in *N, N*-dimethylformamide with 0.1 M hydroxy benzotriazole of the Fmoc group. After synthesis, the DMF/resin mixture was filtered, rinsed with acetone, and then air-dried. The peptides were cleaved from the resin by incubation at room temperature for 3 hours in a cocktail consisting of 92.5% trifluoroacetic acid (TFA), 2.5% distilled water, and 2.5% triisopropylsilane, and 2.5% 2,2'- (ethylenedioxy)-ethanethiol. Cleavage was followed by filtration and subsequent precipitation in diethyl ether. The peptide/diethyl ether mixture was centrifuged at four °C at 4,000 rpm for 10 min. The supernatant was discarded. The precipitate was allowed to desiccate overnight. Following desiccation, the crude peptide gels were resolubilized in 3 mL of a 50:50 mixture of acetonitrile and water (0.1% TFA additive) and purified using Shimadzu LC-20AP reversed-phase high-pressure liquid chromatography (HPLC) equipped with a C18 column. The peptides were eluted with a linear gradient of water-acetonitrile with 0.1% TFA. Peptide mass was confirmed using ESI mass spectrometry. Purified HPLC fractions were then lyophilized, sealed, and stored at -30 °C.

5.5.2-Peptide assemblies

To conduct the experiments described, the stock solutions of the peptide FKFEFKFE (3 mg·mL^{-1}) were prepared at either 3mg/ml or 0.5mg/ml. To do so, 0.75mg and 0.5mg of purified, lyophilized peptides were respectively solubilized in 250 µL and 1000uL of 10 mM acetate buffer, pH 4.0. The solutions were then either allowed to assemble at room temperature or under thermally annealing conditions following thermal cycler protocol: (1) rapid heating to 90° C for 30 minutes and (2) cooling to 25°C at a rate of 0.2°C/minute.

5.5.3- Circular dichroism

CD measurements were performed on a Jasco J-810 CD spectropolarimeter using a 100nm circular cuvette. All spectra were collected at a scanning range of 100 nm/m,in then average from

190-260 nm with a bandwidth of 2nm and a data pitch of 0.2 nm. A 500 ul FKFEFKFE peptide was prepared and used for a complete kinetics measurement at 37ºC. Every measure was done by using 50 ul of the sample at each time point.

5.5.4- Transmission electron microscopy

TEM grids were prepared using diluted solutions of peptide (3 mg·mL) in an aqueous buffer (10 mM acetate, pH 4.0). Samples were prepared by depositing four μL of peptide solution onto a 200-mesh carbon-coated copper grid from Electron Microscopy Services (Hatfield, PA). After 1.5 minutes of incubation on the grid, moisture was wicked away, leaving only a thin film of sthe ample. Four μL of negative stain solution (1% uranyl acetate) was deposited onto the thin film to allow proper mixture. After 1 minute of staining, the remaining moisture was wicked away, and the grid dried overnight in a tabletop desiccator. Electron micrographs were captured on a Hitachi HT-7700 microscope with a tungsten filament and AMT CCD camera at an accelerating voltage of 80 kV

5.5.5- Atomic Force Microscopy

AFM experiments were performed on a NanoWizard 3 (JPK) in liquid using a nanoprobe of 10Å diameter (Peakforce-HIRS-F-B2) in a liquid bio-cell chamber in a QI mode. The peptide assemblies at 430uM or 100uM (concentration expressed in monomer equivalent) have been deposited on a freshly cleaved mica surface for 3 minutes. After two iterative washes of the mica surface with sodium acetate buffer (10mM, pH5.0) the bio-cell compartment was filled with 1ml buffer.

5.5.6- Static Light Scattering

Static light scattering for kinetics and periodic temperature ramping experiments were performed on a homemade device using a 450 nm laser beam, in 2mm quartz cuvettes. Kinetic experiments were performed at 37°C. Cyclic temperature ramping experiments was performed by cycling between 25°C and 85°C at the rate of 1°C/minute with a pause of 5 min between each phase.

5.5.7- Fourier Transform Infrared Spectroscopy

ATR-FTIR spectra were recorded on a diamond crystal of the dried powder at a 4 cm^{-1} resolution with a Bruker IFS 66 spectrophotometer equipped with a 45° N ZnSe ATR attachment. The spectra obtained resulted from the average of 30 scans and were corrected for the linear dependence on the absorption wavelength measured by ATR. The water signal was removed by subtraction of the pure water spectrum.

5.6- Supplemental figures

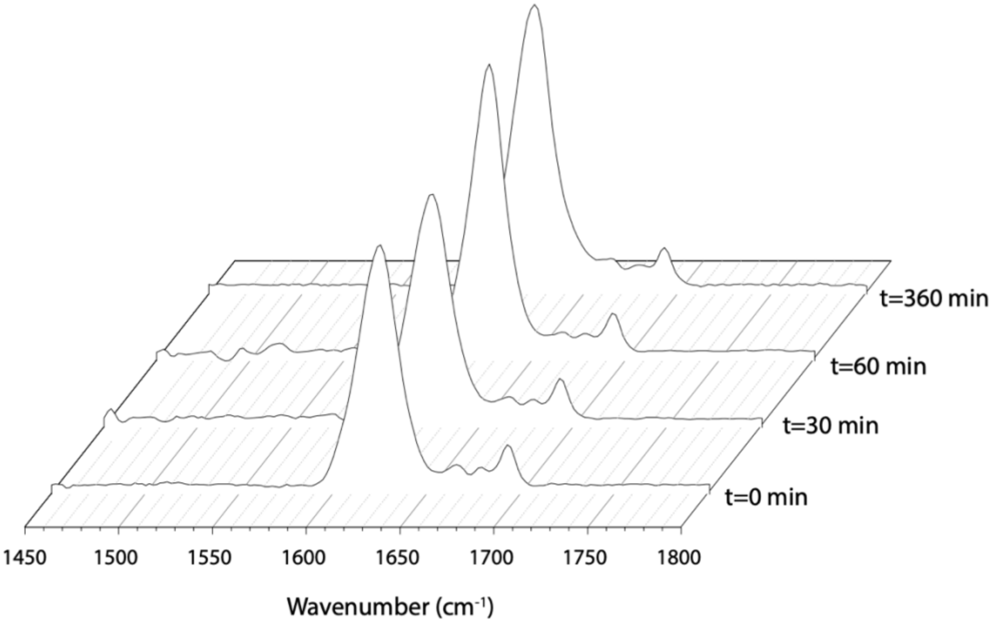


Figure 5-5-1: Time point solid sate FT-IR of the assemblies

The sample was prepared at 430uM in 10mM acetate ph4. The measurement was performed on the pellet resulting from ultracentrifugation of an alicot remove at every time point indicated in the graph: t= 0h, 30min, 60min, and 360min

5.7- References

(1)Yan, X.; He, Q.; Wang, K.; Duan, L.; Cui, Y.; Li, J. Transition of cationic dipeptide nanotubes into vesicles and oligonucleotide delivery. *Angew Chem Int Ed Engl* **2007**, *46* (14), 2431-2434. DOI: 10.1002/anie.200603387 From NLM.

(2)Xu, J.; Dayan, N.; Goldbourt, A.; Xiang, Y. Cryo-electron microscopy structure of the filamentous bacteriophage IKe. *Proc Natl Acad Sci U S A* **2019**, *116* (12), 5493-5498. DOI: 10.1073/pnas.1811929116 From NLM.

(3)de la Rica, R.; Matsui, H. Applications of peptide and protein-based materials in bionanotechnology. *Chem Soc Rev* **2010**, *39* (9), 3499-3509. DOI: 10.1039/b917574c From NLM.

(4)de la Rica, R.; Mendoza, E.; Lechuga, L. M.; Matsui, H. Label-free pathogen detection with sensor chips assembled from Peptide nanotubes. *Angew Chem Int Ed Engl* **2008**, *47* (50), 9752-9755. DOI: 10.1002/anie.200804299 From NLM.

(5)Schneider, A.; Garlick, J. A.; Egles, C. Self-assembling peptide nanofiber scaffolds accelerate wound healing. *PLoS One* **2008**, *3* (1), e1410. DOI: 10.1371/journal.pone.0001410 From NLM.

(6)Reches, M.; Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **2003**, *300* (5619), 625-627. DOI: 10.1126/science.1082387 From NLM.

(7)Omosun, T. O.; Hsieh, M. C.; Childers, W. S.; Das, D.; Mehta, A. K.; Anthony, N. R.; Pan, T.; Grover, M. A.; Berland, K. M.; Lynn, D. G. Catalytic diversity in self-propagating peptide assemblies. *Nat Chem* **2017**, *9* (8), 805-809. DOI: 10.1038/nchem.2738 From NLM.

(8)Wang, R.; Wang, Z.; Guo, Y.; Li, H.; Chen, Z. Design of a RADA16-based self-assembling peptide nanofiber scaffold for biomedical applications. *J Biomater Sci Polym Ed* **2019**, *30* (9), 713-736. DOI: 10.1080/09205063.2019.1605868 From NLM.

(9)Appavu, R.; Chesson, C. B.; Koyfman, A. Y.; Snook, J. D.; Kohlhapp, F. J.; Zloza, A.; Rudra, J. S. Enhancing the Magnitude of Antibody Responses through Biomaterial Stereochemistry. *ACS Biomater Sci Eng* **2015**, *1* (7), 601-609. DOI: 10.1021/acsbiomaterials.5b00139 From NLM.

(10)Altunbas, A.; Lee, S. J.; Rajasekaran, S. A.; Schneider, J. P.; Pochan, D. J. Encapsulation of curcumin in self-assembling peptide hydrogels as injectable drug delivery vehicles. *Biomaterials* **2011**, *32* (25), 5906-5914. DOI: 10.1016/j.biomaterials.2011.04.069 From NLM.

(11)Easterhoff, D.; DiMaio, J. T.; Doran, T. M.; Dewhurst, S.; Nilsson, B. L. Enhancement of HIV-1 infectivity by simple, self-assembling modular peptides. *Biophys J* **2011**, *100* (5), 1325-1334. DOI: 10.1016/j.bpj.2011.01.037 From NLM.

(12)Hauser, C. A.; Zhang, S. Designer self-assembling peptide nanofiber biological materials. *Chem Soc Rev* **2010**, *39* (8), 2780-2790. DOI: 10.1039/b921448h From NLM.

(13)Hamley, I. W. Peptide fibrillization. *Angew Chem Int Ed Engl* **2007**, *46* (43), 8128-8147. DOI: 10.1002/anie.200700861 From NLM.

(14)Hershko, A.; Ciechanover, A. The ubiquitin system for protein degradation. *Annu Rev Biochem* **1992**, *61*, 761-807. DOI: 10.1146/annurev.bi.61.070192.003553 From NLM.

(15)Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional amyloid formation within mammalian tissue. *PLoS Biol* **2006**, *4* (1), e6. DOI: 10.1371/journal.pbio.0040006 From NLM.

(16)Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P.; Simon, R.; Schubert, D.; et al. Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **2009**, *325* (5938), 328-332. DOI: 10.1126/science.1173155 From NLM.

(17)Close, W.; Neumann, M.; Schmidt, A.; Hora, M.; Annamalai, K.; Schmidt, M.; Reif, B.; Schmidt, V.; Grigorieff, N.; Fändrich, M. Physical basis of amyloid fibril polymorphism. *Nat Commun* **2018**, *9* (1), 699. DOI: 10.1038/s41467-018-03164-5 From NLM.

(18)Eanes, E. D.; Glenner, G. G. X-ray diffraction studies on amyloid filaments. *J Histochem Cytochem* **1968**, *16* (11), 673-677. DOI: 10.1177/16.11.673 From NLM.

(19)Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J Mol Biol* **1997**, *273* (3), 729-739. DOI: 10.1006/jmbi.1997.1348 From NLM.

(20)Diaz-Avalos, R.; Long, C.; Fontano, E.; Balbirnie, M.; Grothe, R.; Eisenberg, D.; Caspar, D. L. Cross-beta order and diversity in nanocrystals of an amyloid-forming peptide. *J Mol Biol* **2003**, *330* (5), 1165-1175. DOI: 10.1016/s0022-2836(03)00659-4 From NLM.

(21)Ke, P. C.; Zhou, R.; Serpell, L. C.; Riek, R.; Knowles, T. P. J.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fändrich, M.; et al. Half a century of amyloids: past, present and future. *Chem Soc Rev* **2020**, *49* (15), 5473-5509. DOI: 10.1039/c9cs00199a From NLM.

(22)Caplan, M. R.; Schwartzfarb, E. M.; Zhang, S.; Kamm, R. D.; Lauffenburger, D. A. Control of self-assembling oligopeptide matrix formation through systematic variation of amino acid sequence. *Biomaterials* **2002**, *23* (1), 219-227. DOI: 10.1016/s0142-9612(01)00099-0 From NLM.

(23)Knowles, T. P.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An analytical solution to the kinetics of breakable filament assembly. *Science* **2009**, *326* (5959), 1533-1537. DOI: 10.1126/science.1178250 From NLM.

(24)Morris, A. M.; Watzky, M. A.; Agar, J. N.; Finke, R. G. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"Ockham's razor" model: the Finke-Watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry* **2008**, *47* (8), 2413-2427. DOI: 10.1021/bi701899y From NLM.

(25)Cohen, S. I.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. Nucleated polymerization with secondary pathways. III. Equilibrium behavior and oligomer populations. *J Chem Phys* **2011**, *135* (6), 065107. DOI: 10.1063/1.3608918 From NLM.

(26)Lomakin, A.; Chung, D. S.; Benedek, G. B.; Kirschner, D. A.; Teplow, D. B. On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. *Proc Natl Acad Sci U S A* **1996**, *93* (3), 1125-1129. DOI: 10.1073/pnas.93.3.1125 From NLM.

(27)Krishnan, R.; Lindquist, S. L. Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* **2005**, *435* (7043), 765-772. DOI: 10.1038/nature03679 From NLM.

(28)Luiken, J. A.; Bolhuis, P. G. Primary Nucleation Kinetics of Short Fibril-Forming Amyloidogenic Peptides. *J Phys Chem B* **2015**, *119* (39), 12568-12579. DOI: 10.1021/acs.jpcb.5b05799 From NLM.

(29)Lin, D.; Luo, Y.; Wu, S.; Ma, Q.; Wei, G.; Yang, X. Investigation of the aggregation process of amyloid-β-(16-22) peptides and the dissolution of intermediate aggregates. *Langmuir* **2014**, *30* (11), 3170-3175. DOI: 10.1021/la4048165 From NLM.

(30)Hsieh, M. C.; Lynn, D. G.; Grover, M. A. Kinetic Model for Two-Step Nucleation of Peptide Assembly. *J Phys Chem B* **2017**, *121* (31), 7401-7411. DOI: 10.1021/acs.jpcb.7b03085 From NLM.

(31)Auer, S.; Ricchiuto, P.; Kashchiev, D. Two-step nucleation of amyloid fibrils: omnipresent or not? *J Mol Biol* **2012**, *422* (5), 723-730. DOI: 10.1016/j.jmb.2012.06.022 From NLM.

(32)Levin, A.; Mason, T. O.; Adler-Abramovich, L.; Buell, A. K.; Meisl, G.; Galvagnion, C.; Bram, Y.; Stratford, S. A.; Dobson, C. M.; Knowles, T. P.; et al. Ostwald's rule of stages governs structural transitions and morphology of dipeptide supramolecular polymers. *Nat Commun* **2014**, *5*, 5219. DOI: 10.1038/ncomms6219 From NLM.

(33)Vekilov, P. G. The two-step mechanism of nucleation of crystals in solution. *Nanoscale* **2010**, *2* (11), 2346-2357. DOI: 10.1039/c0nr00628a From NLM.

(34)Vivarès, D.; Kaler, E. W.; Lenhoff, A. M. Quantitative imaging by confocal scanning fluorescence microscopy of protein crystallization via liquid-liquid phase separation. *Acta Crystallogr D Biol Crystallogr* **2005**, *61* (Pt 6), 819-825. DOI: 10.1107/s090744490402949x From NLM.

(35)Childers, W. S.; Anthony, N. R.; Mehta, A. K.; Berland, K. M.; Lynn, D. G. Phase networks of cross-β peptide assemblies. *Langmuir* **2012**, *28* (15), 6386-6395. DOI: 10.1021/la300143j From NLM.

(36)Lomakin, A.; Asherie, N.; Benedek, G. B. Liquid-solid transition in nuclei of protein crystals. *Proc Natl Acad Sci U S A* **2003**, *100* (18), 10254-10257. DOI: 10.1073/pnas.1334069100 From NLM.

(37)Debeljuh, N.; Barrow, C. J.; Byrne, N. The impact of ionic liquids on amyloid fibrilization of Aβ16-22: tuning the rate of fibrilization using a reverse Hofmeister strategy. *Phys Chem Chem Phys* **2011**, *13* (37), 16534-16536. DOI: 10.1039/c1cp22256b From NLM.

(38)Mehta, A. K.; Rosen, R. F.; Childers, W. S.; Gehman, J. D.; Walker, L. C.; Lynn, D. G. Context dependence of protein misfolding and structural strains in neurodegenerative diseases. *Biopolymers* **2013**, *100* (6), 722-730. DOI: 10.1002/bip.22283 From NLM.

(39)Tomiyama, T.; Matsuyama, S.; Iso, H.; Umeda, T.; Takuma, H.; Ohnishi, K.; Ishibashi, K.; Teraoka, R.; Sakama, N.; Yamashita, T.; et al. A mouse model of amyloid beta oligomers: their contribution to synaptic alteration, abnormal tau phosphorylation, glial activation, and neuronal loss in vivo. *J Neurosci* **2010**, *30* (14), 4845-4856. DOI: 10.1523/jneurosci.5825-09.2010 From NLM.

(40)Yang, C.; Li, D.; Liu, Z.; Hong, G.; Zhang, J.; Kong, D.; Yang, Z. Responsive small molecular hydrogels based on adamantane-peptides for cell culture. *J Phys Chem B* **2012**, *116* (1), 633-638. DOI: 10.1021/jp209441r From NLM.

(41)Wang, F.; Gnewou, O.; Wang, S.; Osinski, T.; Zuo, X.; Egelman, E. H.; Conticello, V. P. Deterministic chaos in the self-assembly of β sheet nanotubes from an amphipathic oligopeptide. *Matter* **2021**, *4* (10), 3217-3231. DOI: 10.1016/j.matt.2021.06.037 From NLM.

(42)Bowerman, C. J.; Nilsson, B. L. Self-assembly of amphipathic β-sheet peptides: insights and applications. *Biopolymers* **2012**, *98* (3), 169-184. DOI: 10.1002/bip.22058 From NLM.

(43)Caughey, B.; Lansbury, P. T. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu Rev Neurosci* **2003**, *26*, 267-298. DOI: 10.1146/annurev.neuro.26.010302.081142 From NLM.

(44)Lansbury, P. T.; Lashuel, H. A. A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature* **2006**, *443* (7113), 774-779. DOI: 10.1038/nature05290 From NLM.

(45)Walsh, D. M.; Lomakin, A.; Benedek, G. B.; Condron, M. M.; Teplow, D. B. Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate. *J Biol Chem* **1997**, *272* (35), 22364-22372. DOI: 10.1074/jbc.272.35.22364 From NLM.

(46)Jackson, G. S.; Hosszu, L. L.; Power, A.; Hill, A. F.; Kenney, J.; Saibil, H.; Craven, C. J.; Waltho, J. P.; Clarke, A. R.; Collinge, J. Reversible conversion of monomeric human prion protein between native and fibrilogenic conformations. *Science* **1999**, *283* (5409), 1935-1937. DOI: 10.1126/science.283.5409.1935 From NLM.

(47)Hosszu, L. L.; Baxter, N. J.; Jackson, G. S.; Power, A.; Clarke, A. R.; Waltho, J. P.; Craven, C. J.; Collinge, J. Structural mobility of the human prion protein probed by backbone hydrogen exchange. *Nat Struct Biol* **1999**, *6* (8), 740-743. DOI: 10.1038/11507 From NLM.

(48)McCutchen, S. L.; Lai, Z.; Miroy, G. J.; Kelly, J. W.; Colón, W. Comparison of lethal and nonlethal transthyretin variants and their relationship to amyloid disease. *Biochemistry* **1995**, *34* (41), 13527-13536. DOI: 10.1021/bi00041a032 From NLM.

(49)Lai, Z.; Colón, W.; Kelly, J. W. The acid-mediated denaturation pathway of transthyretin yields a conformational intermediate that can self-assemble into amyloid. *Biochemistry* **1996**, *35* (20), 6470-6482. DOI: 10.1021/bi952501g From NLM.

(50)Cohen, S. I.; Linse, S.; Luheshi, L. M.; Hellstrand, E.; White, D. A.; Rajah, L.; Otzen, D. E.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc Natl Acad Sci U S A* **2013**, *110* (24), 9758-9763. DOI: 10.1073/pnas.1218402110 From NLM.

Chapter 6: Conclusion and outlook

6.1- Conclusion

The self-assembly of designed peptides into filaments and other higher-order structures has been the focus of intense interest because of the potential for creating new biomaterials and biomedical devices. The exquisite structure-function correlations observed in filamentous protein assemblies provide a paradigm for the design of synthetic peptide-based nanomaterials. However, the plasticity of quaternary structure in sequence space and the lability of helical symmetry present significant challenges to the *de novo* design and structural analysis of such filaments. The work shown above described a rational approach to designing self-assembling peptide nanotubes based on controlling lateral interactions between protofilaments with an unusual cross-α supramolecular architecture. Near-atomic resolution cryo-EM structural analysis of seven designed nanotubes provides insight into the designability of interfaces within these synthetic peptide assemblies. It identifies a non-native structural interaction based on a pair of arginine residues. We showed that this arginine clasp motif could robustly mediate cohesive interactions between protofilaments within these cross-α nanotubes. The structure of these resultant assemblies could be controlled through the sequence and length of the peptide subunits, which generates synthetic peptide filaments of similar dimensions to flagella and pili.

Following the analysis of cross-α helical nanotubes, we set to explore the cross-β space. These cross-β peptide assemblies have also been used as models for understanding biological processes, such as the pathological formation of amyloid. We investigate the assembly of an octapeptide sequence, Ac-FKFEFKFE-NH_2, motivated by prior studies that demonstrated this amphipathic b strand peptide self-assembled into fibrils and biocompatible hydrogels. Using high-resolution cryo-electron microscopy (cryo-EM), we determined the atomic structure for two different coexisting forms of the fibrils, containing four and five β sandwich protofilaments, respectively. Surprisingly, the inner walls in both forms are parallel β-sheets, while the outer walls are antiparallel β-sheets. Our results demonstrated the chaotic nature of peptide self-assembly. They illustrated the importance of cryo-EM structural analysis to understand the complex phase behavior of these materials at near-atomic resolution.

We followed the work on self-assembly of peptides nanotubes by studying the mechanism of formations of the cross-β peptide of the octapeptide sequence, Ac-FKFEFKFE-NH_2 that had previously been characterized. We used a combination of techniques to explore the pathways of assemblies and the nature of the different populations of assemblies present in the solution. We demonstrated that the formation of varying diameter nanotubes came from a structural rearrangement of already present protofilaments that were shown to coalesce in ribbons and tubes later in the process. These findings raised the possibility of having an elongation process without the necessary presence of a nucleation step, as previously assumed in the case of amyloid-derived filaments assemblies.

Finally, while the application of cryo-EM to helical polymers in biology has a long history, due to the vast number of helical macromolecular assemblies in viruses, bacteria, archaea, and eukaryotes, the use of cryo-EM to study synthetic soft matter non-covalent polymers has been much more limited. This has mainly been due to the lack of familiarity with cryo-EM in the materials science and chemistry communities, in contrast to the fact that cryo-EM was developed as a biological technique. Nevertheless, the relatively few structures of self-assembled peptide nanotubes and ribbons solved at near-atomic resolution by cryo-EM have demonstrated that cryo-EM should be the method of choice for structural analysis of synthetic helical filaments. In addition, cryo-EM we have aided in establishing that the self-assembly of soft matter polymers has enormous potential for polymorphism, which may be obscured by techniques such as scattering and spectroscopy. These above-described cryo-EM structures have revealed how far we currently are from being able to predict the structure of these polymers due to their chaotic self-assembly behavior.

6.2- Outlook

As one of the most common molecular self-organization phenomena in nature, the study of self-assembly behavior is of great significance for understanding the molecular interaction among bio- logical-logical molecules and creating novel functional materials. Research into self-assembling peptide systems promises significant advancement in the near future. The studies described above serve as an addition to an increasingly important library of self-assembling systems. The cryo-EM structures solved as a result of our effort will help grow the universal understanding of highly complex and still unpredictable principles that govern how proteins and peptides interact at the molecular interfaces. It will also increase the number of tools available for the engineering and the production of a novel, de novo, self-assembling peptide materials. Issues such as the presence of polymorphism in the peptide's samples and the lack of robustness in the sequence-structure space (both tertiary and quaternary) are still to consider. As a result, the computational modeling of well-designed interfaces will be crucial in our ability to control supramolecular structures through intermolecular interactions and program them to form biomaterials that incorporate bifunctionality and biocompatibility. Finally, the use of cryo-EM as the method of choice to perform structural analysis of these synthetic assemblies is more than necessary. It had become standard. The development of high-end cameras of the art software for data collection and processing will make the entire process routine for researchers across fields. The potential in this field is endless, making the thought of what is to come very excitingly.