THE DESIGN, CHARACTERIZATION AND UTILITY OF SELF-ASSEMBLING HYDROPHOBIC COLLAGEN PEPTIDES

By Kenneth N. McGuinness

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Computational Biology and Molecular Biophysics

Written under the direction of

Vikas Nanda PhD

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2016

ABSTRACT OF THE DISSERTATION

THE DESIGN, CHARACTERIZATION AND UTILITY OF SELF-ASSEMBLING HYDROPHOBIC COLLAGEN PEPTIDES

By Kenneth N. McGuinness

Dissertation Director:

Vikas Nanda PhD

This thesis will focus on predicting and characterizing the self-assembly of collagen mimetic peptides and their interactions with natural proteins with varying secondary structure and exposed surface hydrophobicity. Chapter 1 will provide an overview of the forces that affect peptide self-assembly, and describe why the collagen triple-helix is a good model system for studying self-assembly. In Chapter 1, a set of synthetic collagen peptides, that were designed using diffusion limited aggregation (DLA) simulations, are used to probe the role of hydrophobic forces that govern protein self-assembly. DLA simulations predicted that length and pattern of the hydrophobic domain is an important driver of morphology. In Chapter 1, length of hydrophobic sequence is shown to be potentially more important in directing peptide self-assembly. Sequences with four, five and six consecutive hydrophobic residues results in the formation of discs, discs and fibers, and all fibers, respectively. In Chapter 2, the discs formed in Chapter 1 are used to probe the nature of protein-protein interactions. The utilization of peptide nanostructures as scaffolds and substrates for creating new materials and understanding the surface properties of

existing proteins is explored. Mixtures of peptide nanostructures and natural proteins provide evidence that peptide nanostructures can be used as membrane surrogates for membrane proteins, sites of nucleation for fiber assembly, and as a probe for surface hydrophobicity. Chapter 2 highlights that hydrophobicity can potentially be used to target protein structures generally, e.g. alpha helices and beta sheets. Appendix 1 takes a closer look into the morphology of H4 discs in solution. Appendix 2 sheds light on the interactions between molecules with the same and opposite chirality. This study utilizes the 'chemically nude' $(PPG)_{10}$ triple-helix to show that self-assembly into nanostructures is geometrically favorable only between opposite-handed triple helices. The implications of favorable packing preference between opposite-handed helices will be discussed.

Acknowledgments and Dedications

My PhD journey started on an interview at a local biotechnology company when one of the interviewers asked what my plans were for the next 5 years. My response of wanting to work for a few years and then apply to graduate school elicited an email from the interviewer to Dr. Daniel Hseih, a graduate student of the program I wanted to join. I am grateful I did not get the job and for the amazing consolation prize of an email that changed my life. Upon hearing of my interests, Daniel, a graduate student in Dr. Vikas Nanda's lab, suggested that I directly contact his advisor. Luckily, Dr. Nanda agreed to meet with me. Only later did I find out that Dr. Nanda thought I was a friend of Daniel. The first time I met Dr. Nanda I was consumed by his cordial nature. Dr. Nanda agreed to take me as a researcher for my final undergraduate semester. I'm grateful for the initial lab members, Sumana and Dr. Srinivas Annavarapu, Yolanda Hom, Dr. Fei Xu, John Wildeman, Mihir Joshi, and Dr. Daniel Hseih for their welcoming hospitality and guidance. At the end of the semester, I spoke to Dr. Nanda about my aims for entering graduate school the following year. He agreed to allow me to continue researching in his lab for the upcoming year and write a recommendation letter to graduate school. He wrote not only one, but two letters to the BioMaPS graduate school entrance committee. I would like to thank Dr. Daniel Weinstock for his help in the graduate school admission process and all his help during my first year. I would like to thank Jean Baum for taking me in as a rotation student and the Baum lab members Dr. Ana Monica Nunes, Dr. Maria Janowska, and Gina Moriarty for helping me navigate NMR and get through graduate school a little easier. I would like to thank Dr. Kathyrn Uhrich for taking me in as a lab rotation student and her students for helping me. I would like to thank Dr. Gyan Bhanot for his guidance and time during my first year. I would like to thank the BioMaPS staff for their tireless

help. I would like to thank Dr. Steven Burley for his initiative, and Dr. Gail Ferstandig Arnold for her countless words of wisdom and constantly helping me and listening to my struggles and always being there for support. It wouldn't be without her that today I would feel comfortable calling Dr. Vikas Nanda, "Vik". She has been my guide throughout graduate school and chief cheerleader, thank you Gail! I would like to thank Dr. Disha Patel, Dr. Joe Bauman and Jennifer Wiseman for their help in the lab and our light hearted conversations throughout the day, Dr. Steve Tuske for our inspirational conversations and suggestions for when times were tough. I would like to thank Ian Bezar for his calm mannerisms and our lunches together, and conversations, especially over eating Indian food and commiserating about science and hockey. I would like to thank David Gould for his friendship and support. I would like to thank Dr. Mike Seiler, Dr. Liyang Diao, Gautaum Singh, and Dr. Iwen Fu for their friendships, heart-to-hearts about graduate school and life, and rock climbing. I would like to thank Stefjord Todolli, my roommate and my friend for his loyal nature, his ears and his guidance throughout graduate school. I hope we continue to stay connected. I would like to my colleague Doug Pike for his friendship, continued encouragement and feedback on life and science and his endearing humor about randomness. I would like to thank my colleague Patrick Nosker for his friendship, our trips to California and Maryland and his passion for questioning and encouraging criticism. I would like to thank Jose James for his friendship and endurance inside the lab – you inspire me to continue to seek. I would like to thank my colleague Sandeep Belure for his friendship and constant seeking for betterment. I would like to thank my friend Swapnil Devarkar – I hope we stay connected. I would like to thank my girlfriend Elizabeth Schmidt for helping me seek to better myself and always seeking spiritual growth inside and outside of our relationship as well her hands for helping me edit my dissertation. I would like to thank my committee member Dr. Sanjeeva Murthy, for his constructive criticisms,

for our car rides to Ithica and Oak Ridge National Labs, his guidance in X-ray methods and life lessons, and his very valuable ears. To my advisor Dr. Vikas Nanda, if I were able to choose a third father – it would be you. You are an ever flowing fountain of perseverance, openness, positivity, creativity and overall an amazing individual that I am blessed to call my advisor. I dedicate this thesis to my family, especially my mother Melissa Burger – I wouldn't be here without her. My father Brendan McGuinness and my 'Poppa' Peter Burger for continuing to build a better relationship with me. I would like to thank my brother Thomas Burger for his playful nature, it is ever endearing to my heart – keep exploring. I would like to acknowledge my brother Kevin McGuinness, I will always love you no matter what. Last but not least, my committee members Dr. Ann Stock, Dr. Meenakshi Dutt, and Dr. David Shreiber for helping me to find my way as a graduate student and for providing me with very helpful feedback on my project, presentations, and dissertation. The process of obtaining a PhD cannot solely be measured by only the number of words written or not written, experiments performed or not performed, articles read or not read, hours alone or with others, nor with any stick marked with any number of meaningful events. Although it is necessary to put, in some form or another, these meaningful events on a single document for the purpose of furthering ones career, a broader perspective must be sought. I am the sum of all my experiences and for that I am truly grateful.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Atomic Force Microscopy	AFM
Alpha-synuclein	AS
Circular Dichroism	CD
Collagen Mimetic Peptides	CMP
Collagen type I	COLI
Congo Red	CR
Diffusion Limited Aggregation	DLA
Dithiothreitol	DTT
Molecular Dynamics	MD
Phosphotungstic acid	PTA
Reaction Center-Light Harvesting Complex 1	RC-LHC1
Shrimp tropomyosin	ST
Small-angle neutron scattering	SANS
Small-angle X-ray scattering	SAXS
Sypro Orange	SO
Transmission Electron Microscopy	TEM
Wide-angle X-ray scattering	WAXS

Prior Publications

Portions of Chapter 1 were reprinted with permission from Morphological Diversity and Polymorphism of Self-Assembling Collagen Peptides Controlled by Length of Hydrophobic Domains Kenneth McGuinness, I. John Khan, and Vikas Nanda *ACS Nano* 2014 **8** (12), 12514-12523 DOI: 10.1021/nn505369d Copyright 2014 American Chemical Society."

Portions of Chapter 2 are currently being prepared for submission for publication

Portions of Appendix B were reprinted with permission from Self-Assembly of Left- and Right-Handed Molecular Screws Fei Xu, I. John Khan, Kenneth McGuinness, Avanish S. Parmar, Teresita Silva, N. Sanjeeva Murthy, and Vikas Nanda *Journal of the American Chemical Society* 2013 ***135*** (50), 18762-18765 DOI: 10.1021/ja4106545 Copyright 2013 American Chemical Society

Introduction

Molecular scaffolds display a surface with binding motifs to support substrate stability, function, and organization. Proteins have evolved surfaces that are capable of interacting with nucleic acids, lipids, carbohydrates and other proteins either as scaffold or substrate agents via many noncovalent interactions between charged, polar, and hydrophobic groups. These interactions make up the architecture of a cell and provide the basis for gene regulation, cell signaling, tissue support, and multiple length scales of organization. Complex scaffolding-substrate organizing schemes can be seen at the level of chromosomes all the way to the macromolecular protein organization found in bones, skin and muscle. Designing protein scaffolds that mimic nature or display novel surfaces capable of targeting a specific substrate is of considerable interest to biotechnology and biomedicine. Inspired by nature, tremendous progress has been made rationally designing sequences that fold into alpha-helices, beta-strands, and collagen peptides that then self-assemble via amphipathic, charge, and hydrophobic interactions into nanostructures capable of acting as scaffolds.[1] Self-assembling peptide nanostructures show great promise and have been designed for vaccines,[2] antimicrobial materials,[3] cellular support,[4] therapeutic delivery vehicles[5] and a variety of other biotechnologies.[6] Choosing a peptide that self-assembles is a challenge due to the 20^N (where N is the sequence length) possible peptide sequences from which to choose. To traverse this vast peptide space experimentally would be too cost-prohibitive. Therefore, computational methods able to predict self-assembling sequences are needed to help guide the experimentalists' choice. Coarse-grained molecular dynamics (MD) simulation methods are becoming popular tools for simulating protein and peptide behavior. Coarse-grained MD simulations reduce the

number of atomistic degrees of freedom considered, allowing for more efficient sampling in time and space with decreased accuracy compared to all-atom MD[7]. Coarse-grained simulations have made possible predicting the self-assembly behavior of all dipeptide[8] and tripeptide[9] sequences. However, predicting the self-assembly behavior of larger peptides remains computationally challenging. This thesis describes a coarse-grained method for predicting the self-assembly of 90 amino acid collagen peptides using diffusion limited aggregation (DLA). The DLA studies consider a group of amino acids as either hydrophobic or hydrophilic. Simulations provided a testable hypothesis such that pattern and length of sequence hydrophobicity were considered both important factors controlling morphology of self-assembly. In addition to the difficulties of predicting the morphology of self-assembly, experimentally confirming the mechanism by which self-assembly occurs is a challenge due to limited atomic level structural data. As a result, synthetic sequences were selected based on DLA predictions and resulting self-assembled nanostructures were subjected to numerous indirect experimental tests. Empirically, it is shown that the length of sequence hydrophobicity is a more pronounced indicator of assembly morphology then pattern. Nanostructure morphology and peptide stability are impacted by solvent polarity. Solvent pH shows an impact on self-assembly morphology and not peptide stability. Insights gained from probing the nature of the self-assembled nanostructures can be used to design scaffolds and substrates for natural and synthetic proteins with hydrophobic surfaces. Furthermore, the interactions between natural proteins and nanostructures adds insight into the evolution of protein surfaces and their impact on protein-protein interaction specificity.

Chapter 1

Morphological Diversity and Polymorphism of Self-Assembling Collagen Peptides Controlled by Length of Hydrophobic Domains

1.1 Abstract

Synthetic collagen mimetic peptides are used to probe the role of hydrophobic forces in mediating protein self-assembly. Higher-order association is an integral property of natural collagens, which assemble into fibers and meshes that comprise the extracellular matrix of connective tissues. The unique triple-helix fold fully exposes two-thirds of positions in the protein to solvent, providing ample opportunities for engineering interaction sites. Inclusion of just a few hydrophobic groups in a minimal peptide promotes a rich variety of self-assembly behaviors, resulting in hundred-nanometer to micron size discs and nanofibers. Morphology depends primarily on the length of hydrophobic domains. Peptide discs contain lipophilic domains capable of sequestering small hydrophobic dyes. Combining multiple peptide types result in composite structures of discs and fibers ranging from stars to plates-on-a-string. These systems provide valuable tools to shed insight into the fundamental principles underlying hydrophobicity-driven higher-order protein association that will facilitate the design of self-assembling systems in biomaterials and nano-medical applications.

1.2 Introduction

Protein complex interfaces often incorporate clusters of hydrophobic amino acids, serving as sticky surfaces for adhesion.[10-13] Surface hydrophobicity is tightly regulated in natural proteins,[14, 15] and inappropriate presentation of residues on the surface can lead to aggregation and pathogenesis. In sickle cell anemia, a single hydrophilic to hydrophobic amino acid substitution on the surface of hemoglobin results in massive protein aggregates that distort red blood cells.[16] Hydrophobic stretches in protein sequences can contribute to amyloidogenesis.[17-22]

Hydrophobicity plays an important role in protein engineering - optimizing core interactions is a critical step in *de novo* design.[23-28] Adjusting surface hydrophobicity can be used to enhance protein solubility.[28, 29] The extensive use of hydrophobicity in self-assembling block copolypeptides,[30, 31] peptide amphiphiles,[32-35] small peptides[36-39] and foldamers,[40, 41] demonstrates the tremendous utility of this force in the broader field of nano-structural molecular engineering.

The collagen triple-helix fold is ideal for studying hydrophobic self-assembly

The triple-helix fold of collagen is an attractive molecular scaffold for studying the role of hydrophobicity in protein-protein interactions and higher-order assembly. Unlike nearly all other protein folds, which assemble around a hydrophobic core, the three strands of the triple-helix associate through a network of backbone-backbone interchain hydrogen bonds. As a result, hydrophobic groups on the surface are not expected to interfere with core folding, as seen in off-pathway intermediates of globular proteins.[42-45] The collagen triple helix is formed by tandem repeats of X-Y-Gly triplets, where X and Y residues are solvent exposed, and where glycine is critical for core assembly.[46] In natural collagens,

hydrophobic residues at X and Y positions are significantly under-represented relative to globular proteins (approximately 8% of residues in type I collagen (Figure 1.1), compared to 40% for myoglobin - a typical globular fold). Those that are present occur in periodic clusters every 231 - 237 residues, promoting assembly of triple-helices into higher-order fibers with a characteristic 67 nm spacing.[47-49] As with other proteins, hydrophobicity of collagen appears to be tightly regulated, playing an important role in the process of self-assembly.

GPMGPSGPRGLOGPOGAOGPQGFQGPOGEOGEOGASGPMGPRGPOGPOGKNGDDGEAGKOGROG
ERGPOGPQGARGLOGTAGLOGMKGHRGFSGLDGAKGDAGPAGPKGEOGSOGENGAPGQMGPRGL
OGERGROGAOGPAGARGNDGATGAAGPOGPTGPAGPOGFOGAVGAKGEAGPQGPRGSEGPQGVR
GEOGPOGPAGAAGPAGNOGADGQOGAKGANGAOGIAGAOGFOGARGPSGPQGPGGPOGPKGNSG
EOGAOGSKGDTGAKGEOGPVGVQGPOGPAGEEGKRGARGEOGPTGLOGPOGERGGOGSRGFOGA
DGVAGPKGPAGERGSOGPAGPKGSOGEAGROGEAGLOGAKGLTGSOGSOGPDGKTGPOGPAGQD
GROGPOGPOGARGQAGVMGFOGPKGAAGEOGKAGERGVOGPOGAVGPAGKDGEAGAQGPOGPAG
PAGERGEQGPAGSOGFQGLOGPAGPOGEAGKPGEQGVOGDLGAOGPSGARGERGFOGERGVQGP
OGPAGPRGANGAOGNDGAKGDAGAOGAOGSQGAOGLQGMOGERGAAGLOGPKGDRGDAGPKGAD
GSOGKDGVRGLTGPIGPOGPAGAPGDKGESGPSGPAGPTGARGAOGDRGEOGPOGPAGFAGPOG
ADGQOGAKGEOGDAGAKGDAGPOGPAGPAGPOGPIGNVGAOGAKGARGSAGPOGATGFOGAAGR
VGPOGPSGNAGPOGPPGPAGKEGGKGPRGETGPAGROGEVGPOGPOGPAGEKGSOGADGPAGAO
GTOGPQGIAGQRGVVGLOGQRGERGFOGLOGPSGEOGKQGPSGASGERGPOGPMGPOGLAGPOG
ESGREGAPGAEGSOGRDGSOGAKGDRGETGPAGPOGAOGAOGAOGPVGPAGKSGDRGETGPAGP
AGPVGPVGARGPAGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPOGPOGSOGEQGPSGASGPA
GPRGPOGSAGAOGKDGLNGLOGPIGPOGPRGRTGDAGPVGPOGPOGPOGPOGPO

Location of D-periodic hydrophobic triplets.

position	Sequence	Δposition
233	**GAOGFOGAR**	233
464	**GSOGFQGLP**	231
698	**GATGFOGAA**	234
935	**GHRGFSGLQ**	237

Figure 1.1. COLIαI Sequence analysis of the triple-helical region. Hydrophobic residues in green. Repeating phenylalanine residues highlighted in yellow shows the canonical 234 amino acid spacing corresponding to 67 nm D-periodic banding.

The design of hydrophobic CMPs

We specifically probe the effect hydrophobicity plays in collagen assembly by designing CMPs incorporating contiguous blocks of hydrophobic residues at non-glycine positions (Table 1). Self-assembly of CMPs into fibers or plates is observed, where subtle changes in the hydrophobic residue pattern lead to dramatic switches in nanoscale morphology.

Table 1. Model hydrophobic collagen peptide sequences

	Sequence	Morphology **P**late / **F**iber
	(LIG)	N/A
	$(LIG)_3$	N/A
H2	$(POG)_3$POGPOG**LI**G$(POG)_4$	P
H3	$(POG)_3$POGP**I**G**LI**G$(POG)_4$	P
H4	$(POG)_3$POG**LI**G**LI**G$(POG)_4$	P
SH4	$(POG)_2$**LI**G$(POG)_4$**LI**G$(POG)_4$	P
QH4	$(POG)_3$QQG**LI**G**LI**G$(POG)_4$	P
H5	$(POG)_3$P**I**G**LI**G**LI**G$(POG)_4$	F/P
YH5	YGG$(POG)_3$P**I**G**LI**G**LI**G$(POG)_4$	N/A
H6	$(POG)_3$**LI**G**LI**G**LI**G$(POG)_4$	F

Simulations of CMP self-assembly

We expected that both the pattern and extent of hydrophobicity along the triple-helix would be important parameters in directing self-assembly. Given ten amino acid triplets (X-Y-Gly) per thirty-residue peptide, there are nominally $2^{10} = 1024$ possible combinations of hydrophobic and hydrophilic triplets. To rapidly identify those sequence capable of forming unique nanostructures, we implemented a coarse-grained model of self-assembly, simulating the process of diffusion-limited aggregation (DLA) [50, 51]. More sophisticated

simulation approaches of hydrophobic peptide self-assembly have been developed that include internal degrees of freedom within and between polypeptide chains,[52, 53] but we chose rigid body docking of rods as the CMPs studied here are within an order of magnitude shorter than the persistence length of collagen.[54]

CMP self-assembly modeled using Diffusion Limited Aggregation (DLA)

In DLA, higher-order structures emerge through the step-wise accretion of monomers onto a growing assembly (Figure 1.1 B, C); this model assumes that monomer concentrations are sufficiently low that diffusion is the limiting factor in assembly kinetics. This method was attractive as it had been shown to replicate experimentally observed morphologies for systems of self-assembling peptides[55, 56] and structural proteins elastin[57] and collagen.[58] Although we did not know *a priori* whether assembly of CMP peptides would proceed in a diffusion-limited or activated manner, such DLA-based simulations were able to explain experimentally observed features of assembled collagen fibers.[59, 60]

1.3 Results and Discussion

We adapted a DLA simulation protocol from previous collagen studies.[58] The triple-helix was treated as a rigid rod composed of ten spheres, each sphere representing a hydrophobic (*h*) or polar (*p*) X-Y-Gly triplet positioned along the triple-helix surface. The simulations were performed using a 3D hexagonal lattice (Figure 1.1 A). A seed rod was placed at the center of the simulation and additional rods were released and allowed to move randomly until either contacting the aggregate or disregarded due to moving too far away from the aggregate (Figure 1.1 A). Simulations of each of the 1024 possible combinations were repeated 1000 times to assess convergence of nanoscale morphs. A suitable interaction

state was defined as greater than two hydrophobic groups in contact in the surrounding hexagonal lattice. Representative assembly structures were chosen for presentation. The peptide H4 was modeled after the DLA rod pppphhpppp. The peptide H6 was modeled after ppphhhpppp. The peptide SH4 was modeled after pphppppphpp.

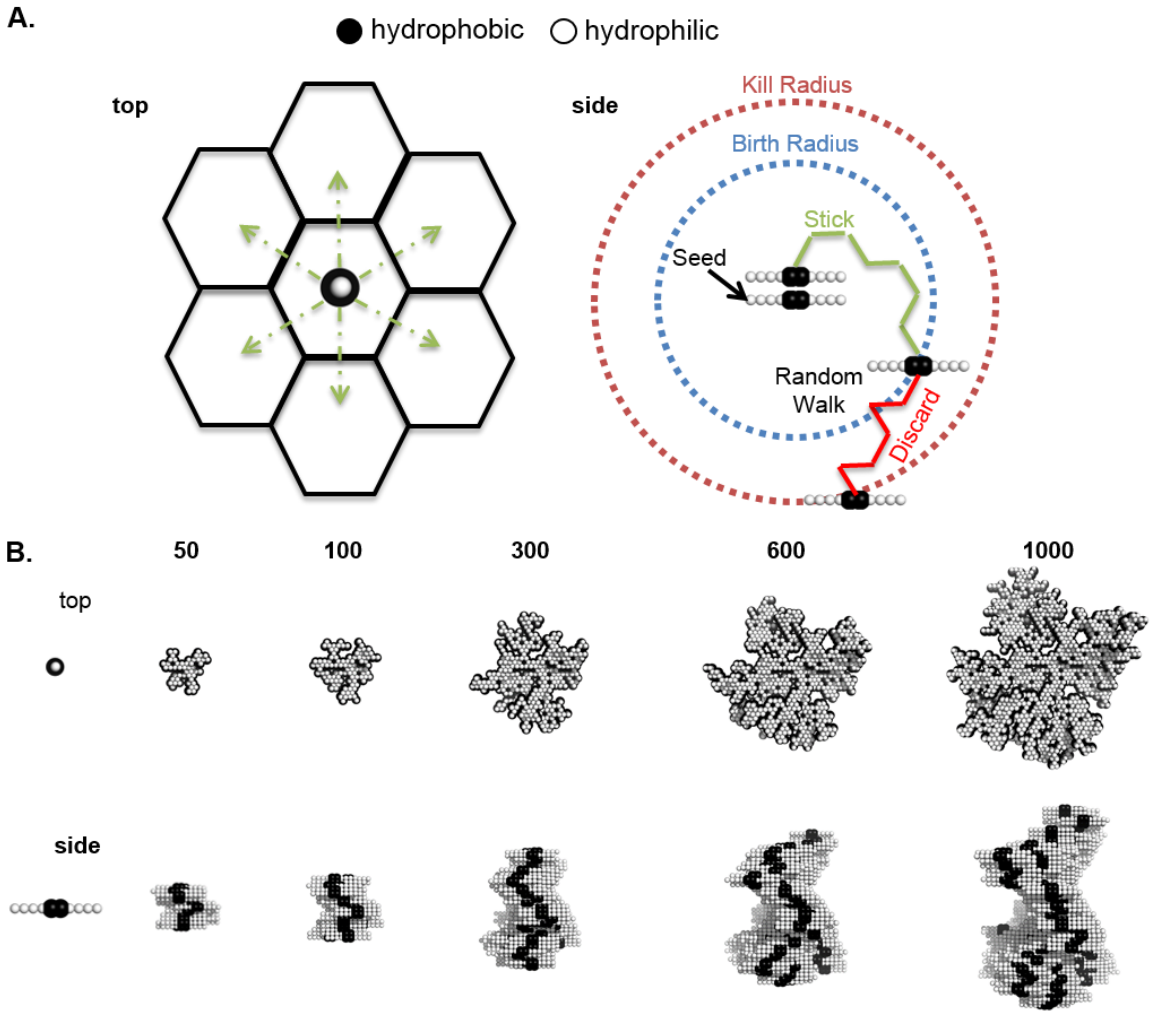

Figure 1.2. DLA simulation. (A) 3D hexagonal lattice of acceptable moves, including moves about the long axis of the monomer (top) and an example of pppphhpppp (H4) taking a random diffusion walk around the hexagonal lattice (side). Accretion of (B) pppphhpppp (H4) in steps of 50, 100, 300, 600, 1000 monomers.

From the 1024 pattern combinations, we chose to focus on *pppphhpppp*, where *p* ≡ Pro-Hyp-Gly (POG) and *h* ≡ Leu-Ile-Gly (LIG) (Figure 1.3 A). The choice of hydrophobic residues was based on the high independent frequencies of Leu and Ile at X and Y positions respectively of natural fibrillar collagens.[61]

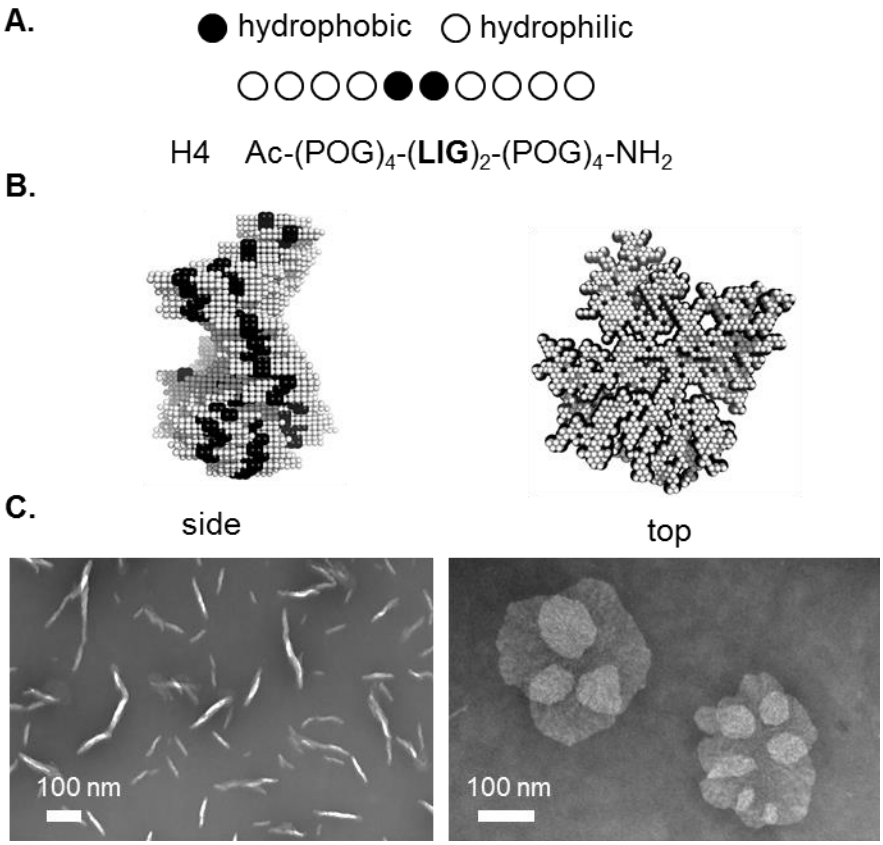

Figure 1.3. H4 peptide design and characterization (A) Coarse-grained model of the H4 sequence. Each rod consists of ten spheres, each sphere representing a triplet of amino acids. (B) Simulated DLA of 1000 pppphhpppp rods generated plate-like structures. Discs pictured would represent assemblies approximately 50 nm in diameter view from the top. (C) Electron micrographs of H4.

Leucine has also been shown to promote higher-order assembly in other CMP designs. [62, 63] The DLA simulation predicted this pattern would form plate-like structures (Figure 1.2 B, Figure 1.3 B).

CMP H4 form discs matching DLA prediction

Transmission Electron Microscopy (TEM) images of H4 were consistent with the DLA prediction of a disc-like morphology. Specifically, three features of the DLA models were observed by TEM (Figure 1.3 C): (1) H4 discs were frequently observed to extend end-on from the hydrophobic carbon coated copper EM grid, consistent with a hydrophobic disc edge adhering to the surface; (2) thickness of these discs was estimated to be ~10 nm, equivalent to the length of one peptide, which would be expected if the peptides were assembling perpendicular to the plane of the disc; (3) rosette structures were observed in multiple preparations of H4, consistent with fractal dendritic DLA-type aggregation in two dimensions.[55] The lack of sharper, well-defined dendritic structures may have been due to fluidity of the disc phase, allowing internal rearrangements to minimize the surface energy of the aggregate. An alternate process for rosette formation would be the association of fully assembled discs, rather than accretion of monomers through DLA. We were not able to discriminate between these two mechanisms based on the existing observations.

H_4 discs are flexible

Although rosettes were consistently found, the dominant H4 morphology was a rounded disc. H4 disc diameters ranged from 50 nm to 1 micron. In many cases, the plates appeared to be flexible, depositing on the grid in shapes resembling curved 'potato chips' (Figure 1.4,

1.5), indicating these structures lacked the crystallinity and rigidity found in other designed CMP plates.[64-66]

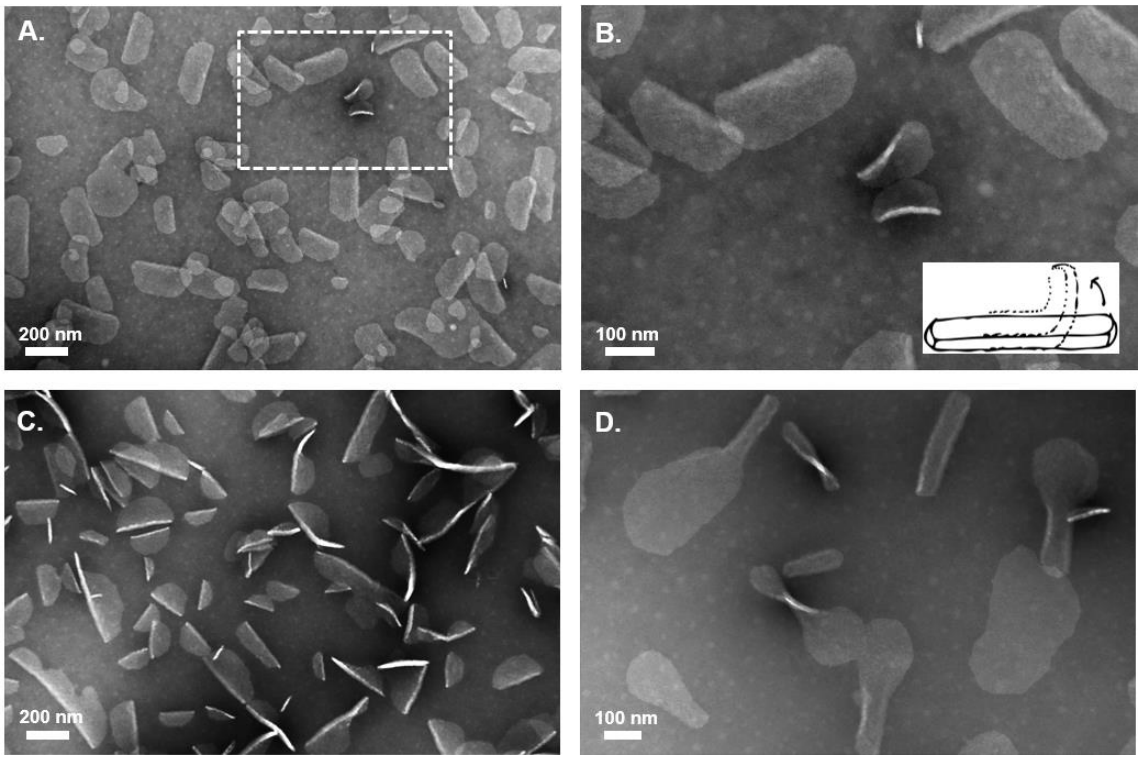

Figure 1.4. TEM of H4 highlighting flexibility. (A) This field shows numerous discs alone or overlapping adjacent discs. (B) Close-up of two H4 discs in field A curving as shown in the drawing. (C) Multiple curved discs and discs perpendicularly aligned to the electron beam. (D) Another field highlighting disc flexibility.

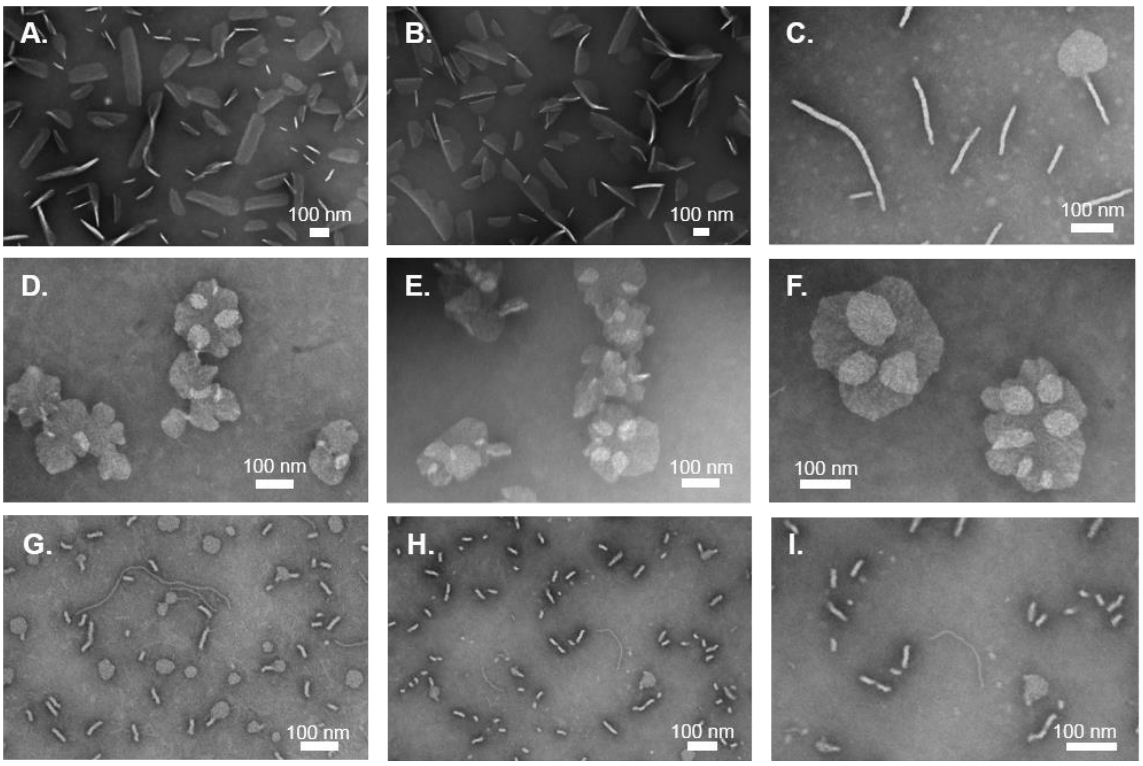

Figure 1.5. TEM of H4 highlighting populations of discs and fibers (A-C) discs. Bright lines are discs on their edge, perpendicular to the grid (D-F) rosettes (G-I) contain minor fiber populations among a field of discs. Panel I is a close-up of panel H.

H4 discs are composed of packed triple-helices

Using circular dichroism (CD) spectroscopy, we confirmed that the H4 peptide assembled into a triple-helix with a cooperative unfolding transition at 33 °C (Figure 1.6 A, B). Observed d-spacings from wide-angle X-ray scattering (WAXS) of supersaturated H4 solutions further supported the presence of triple-helices in the aggregate phase. Bragg diffraction peaks at 11.7 and 12.7 Å (Figure 1.6 C) were consistent with previous reports for the lateral packing of collagen triple-helices.[64, 67]

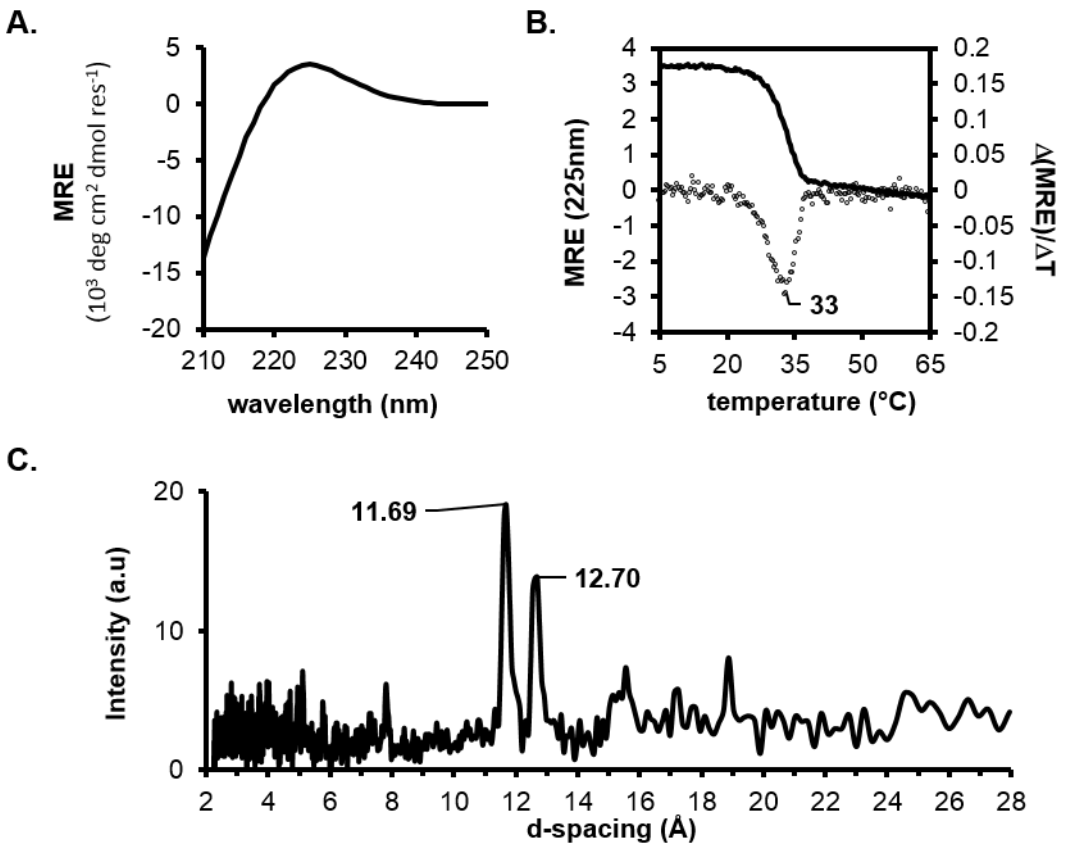

Figure 1.6. H4 forms triple-helices (A) CD spectrum shows positive MRE (mean residue ellipticity) band at 225 nm, consistent with triple-helical structure. (B) Thermal denaturation profile shows cooperative unfolding of H4 with a Tm of 33°C (C) WAXS of H4 shows interhelical d-spacing consistent with fiber packing observed in natural collagens.

The thickness of H4 discs that deposited end-on ascertained by TEM matched the length of triple-helix. This was confirmed by atomic force microscopy (AFM) (Figure 1.7 A). Analysis of many fields revealed height dimensions that occurred in multiples of 7-8 nm (Figure 1.7 B, C), close to the expected length of a CMP of 8.6 nm (30 residues × 0.286 nm rise/residue). Over multiple experiments, a wide range of heights from 6 - 500 nm were observed for H4 (Figure 1.8), all maintaining an overall disc-like morphology, suggesting that significant stacking of discs could occur.

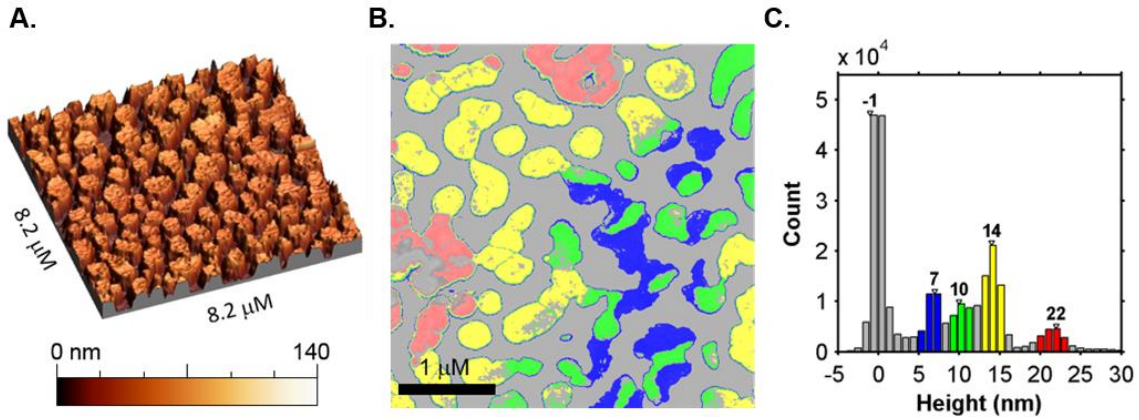

Figure 1.7. AFM of H4 discs (A) 3D representation of discs (B) Coloring of AFM field based on height. (C) Histogram of H4 heights observed in (B).

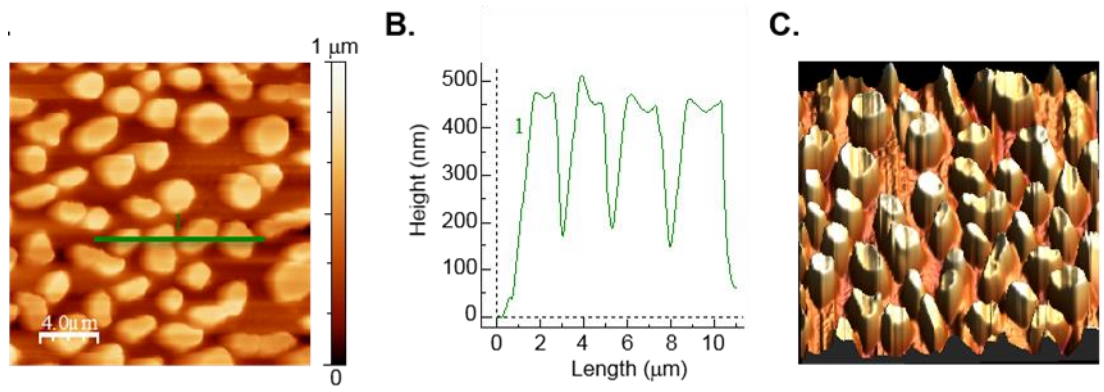

Figure 1.8. (A) AFM of H4 with multiple discs (15-20) stacked on top of each other. (B) Height profile of (A) and (C) 3D view of (A).

Pattern of hydrophobicity effects triple-helix stability not disc morphology

Changing the pattern and keeping hydrophobicity constant for a collagen sequence by separating hydrophobic triplets (*pphpppphpp*) (Figure 1.9 A) was predicted in DLA simulations to produce banded fiber-like structures that have tapered ends (Figure 1.9 B) similar to natural collagen.[68] However, electron microscopy images revealed the synthetic CMP version - SH4 (Table 1) - formed discs (Figure 1.9 C). SH4, for split H4, discs were

10-500 nm in diameter and were observed to bind edge-wise to the TEM grid similarly to H4 discs (Figure 1.9 C). SH4 folded into a slightly more stable triple-helix than H4 (Tm of 33 versus 36 °C) (Figure 1.10) suggesting that the location of destabilizing groups within the triple-helix is an important sequence selection factor.

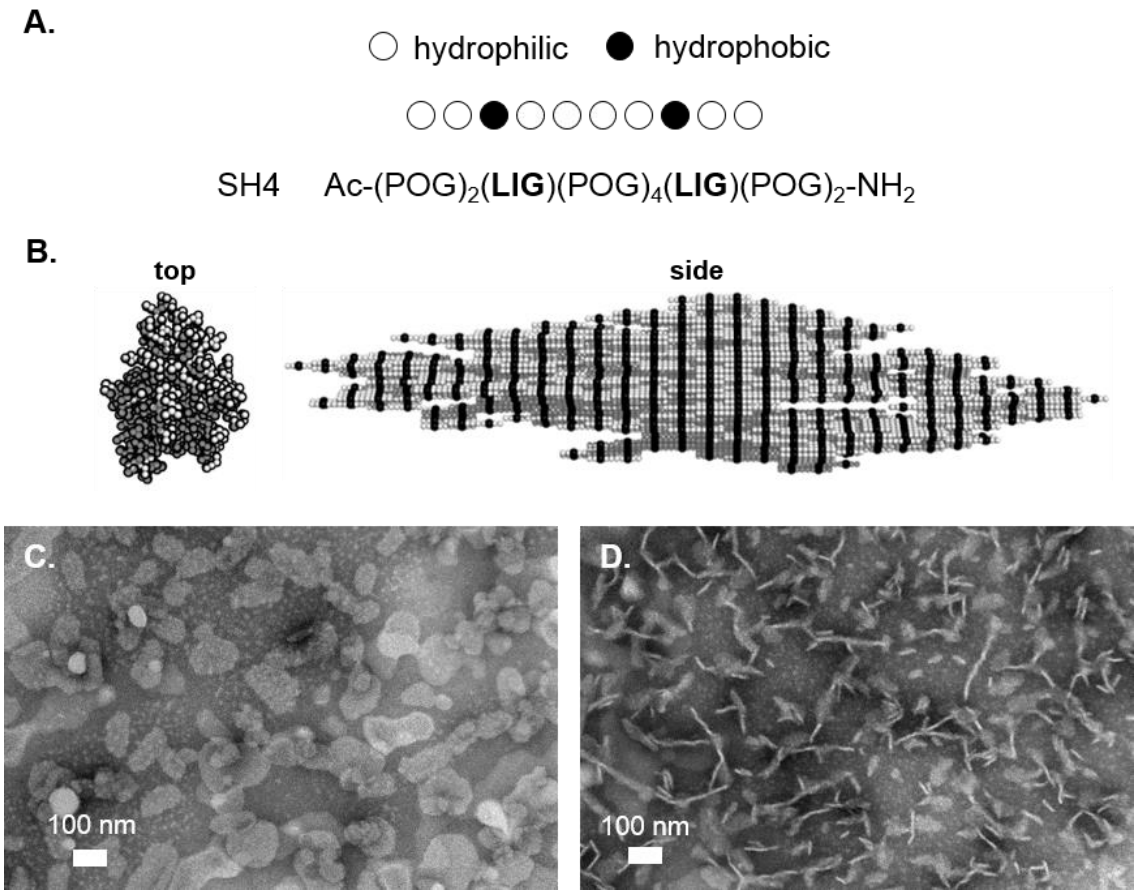

Figure 1.9. SH4 design and characterization. (A) Simulated and synthetic sequence design of collagen peptide SH4. (B) DLA predicted fiber morphology. (C) Electron micrograph of SH4 discs. (D) Electron micrograph of SH4 discs on edge.

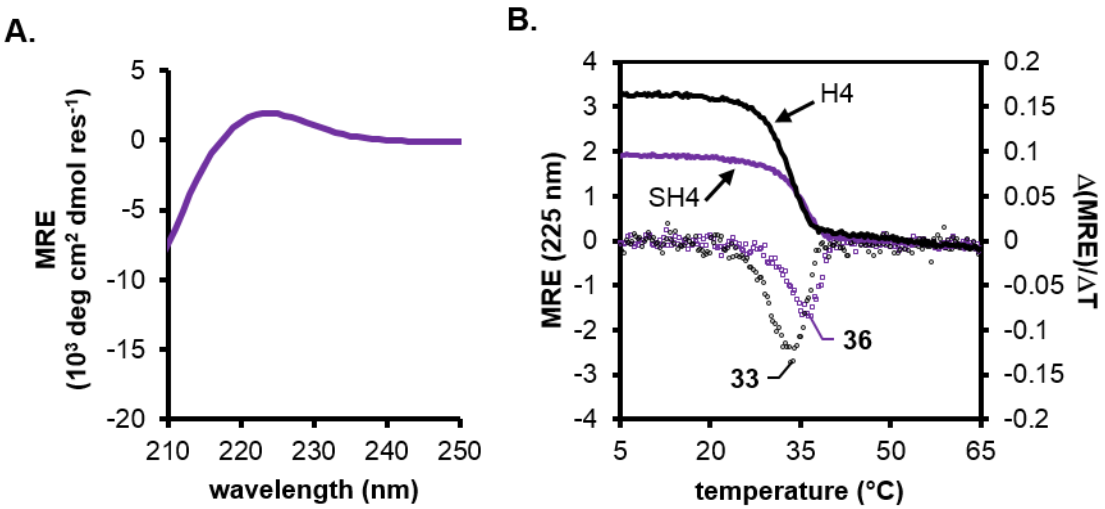

Figure 1.10. Circular Dichroism measurements of SH4 (A) Wavelength scan of SH4. (B) Thermal denaturation plots of SH4 compared to H4.

Reduced hydrophobicity results in disc formation (H3-2)

A reduction in sequence hydrophobicity was predicted to maintain a disc-like morph due to limited registry interactions between triple-helices that would minimize surface exposed hydrophobic residues. Replacing one leucine in H4 with proline produced H3 (Table 1), which folded into a triple-helix, Tm of 41°C (Figure 1.11, 1.12), and self-assembled into plates (Figure 1.13 A). Similarly, H2 formed a triple-helix with a Tm of 50°C (Figure 1.11, 1.12) and was shown by TEM to form plates similar to H4 (Figure 1.13 B).

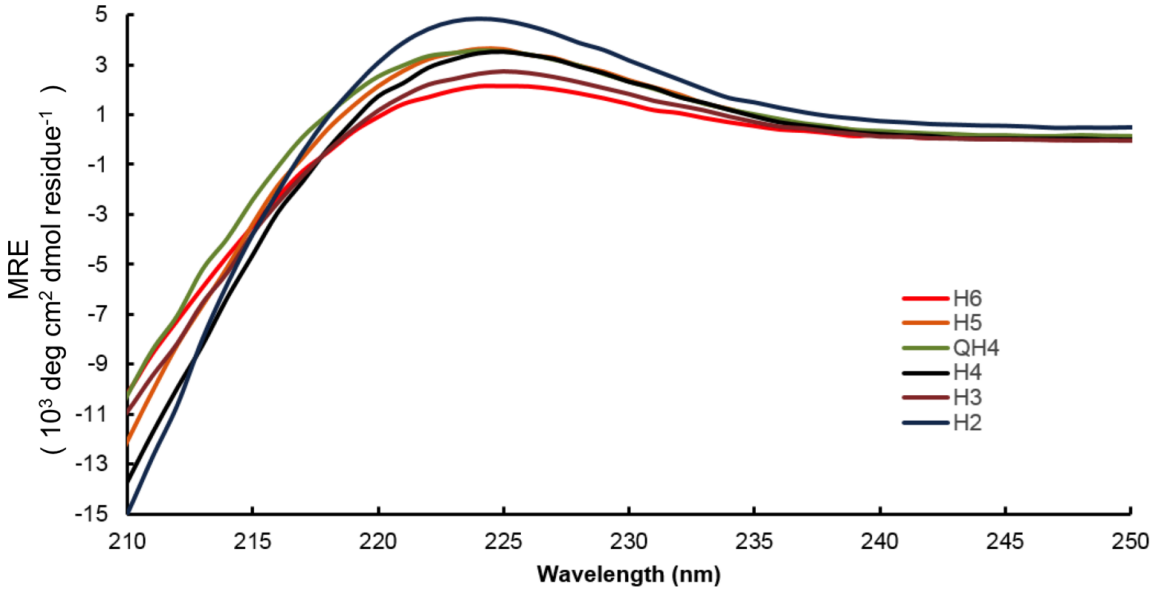

Figure 1.11. Circular dichroism wavelength spectra of peptides H6-2, and QH4 studied.

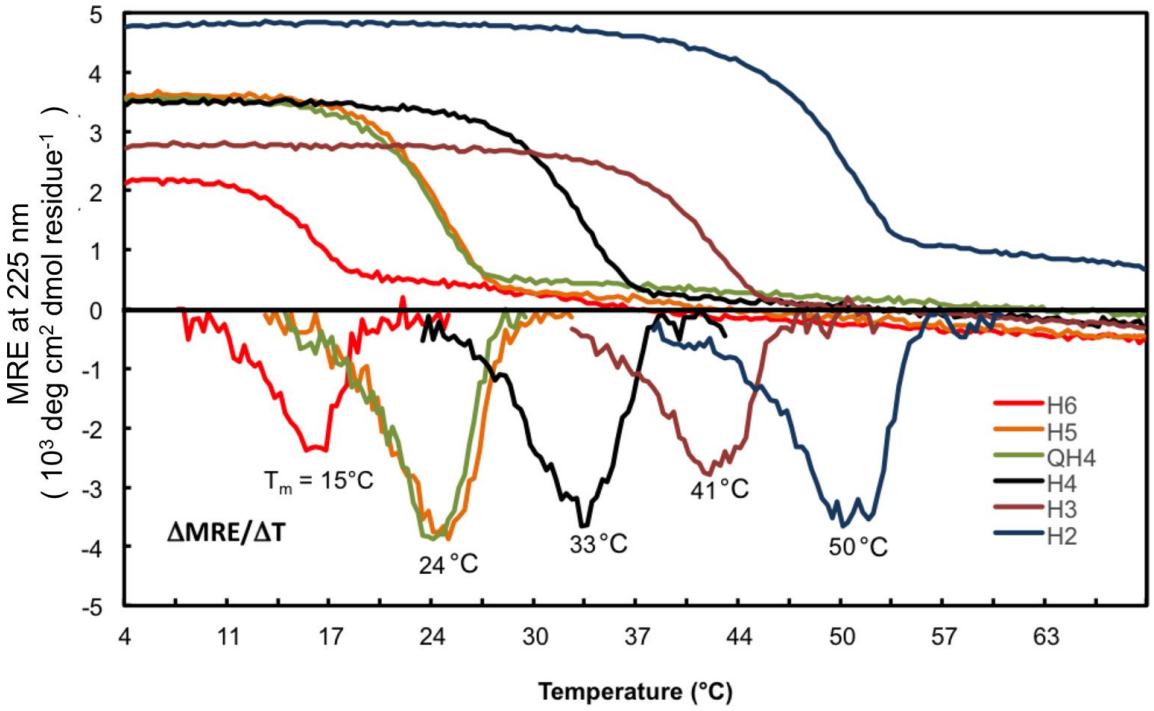

Figure 1.12. Thermal denaturation plots of peptides H6-2, and QH4 studied. Melting temperatures were determined from first derivative analysis (below).

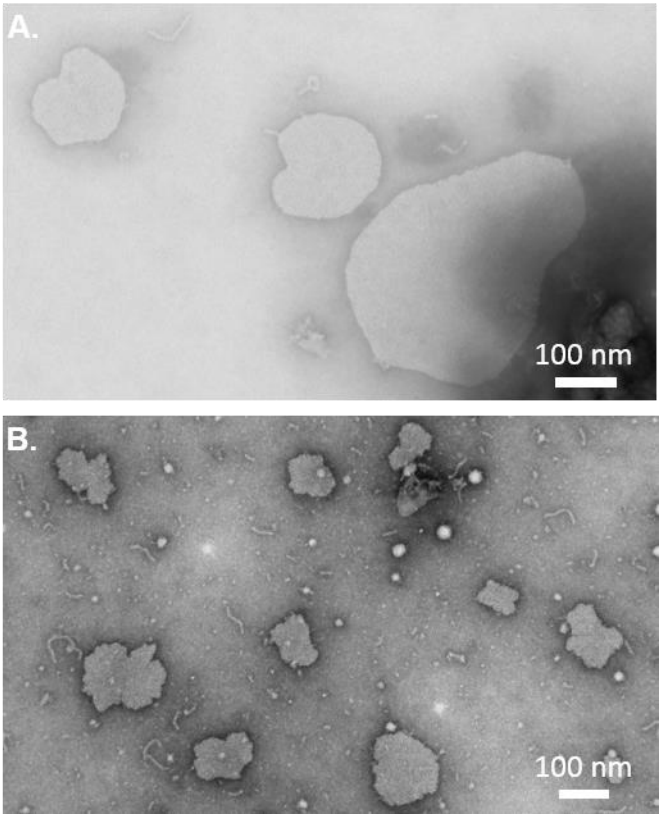

Figure 1.13. TEM micrographs of A) H3 and B) H2.

CMP H6 forms nanofibers

Increasing the hydrophobicity of a collagen sequence by adding an additional hydrophobic triplet (*ppphhhpppp*) was also predicted in DLA simulations to produce disc-like structures with variable thickness due to staggered interactions between adjacent monomer rods (Figure 1.14 A, B). Surprisingly, the CMP equivalent of this pattern - H6 - formed uniform fibers, strikingly different from the model. Fibers observed by TEM varied from ~ 50 nm to several microns in length, and 6-11 nm in width (Figure 1.14 C1, C3). Upon heating and re-annealing, H6 fibers formed straight, aligned clusters several microns in length (Figure 1.14 C2). Solutions of H6 after three years of incubation at 4 °C also formed long fibers

(Figure 1.15 A, B). Occasional evidence of branching and punctate structures was observed, possibly indicating pre-fibril intermediates or small discs (Figure 1.15 C, D).

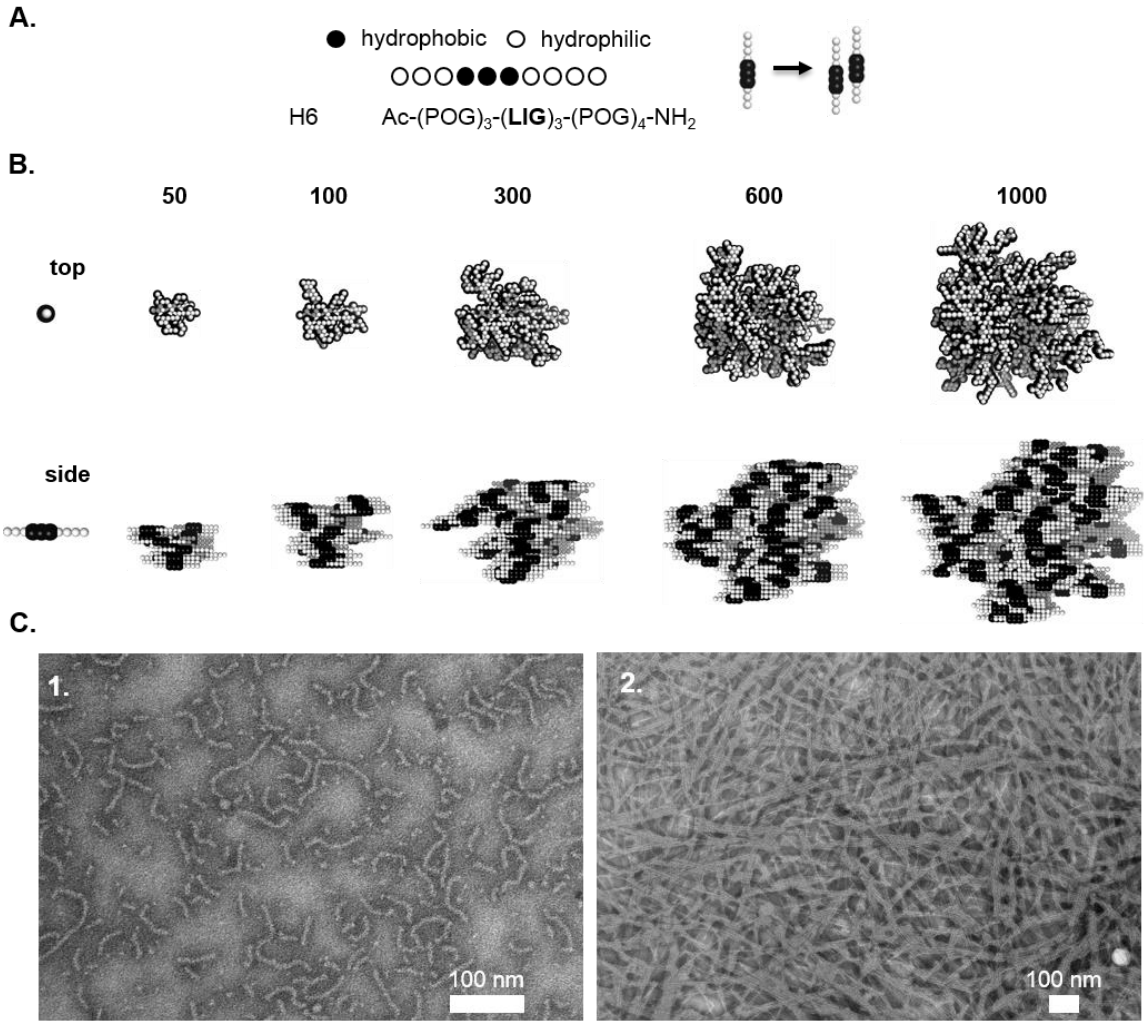

Figure 1.14. H6 peptide design and characterization (A) Simulated and synthetic sequence design of collagen peptide H6. Each sphere represents a triplet of amino acids per chain. (B) DLA simulation (C) Electron micrographs of H6 incubated at 4 °C (C1) and after heating and annealing (C2)

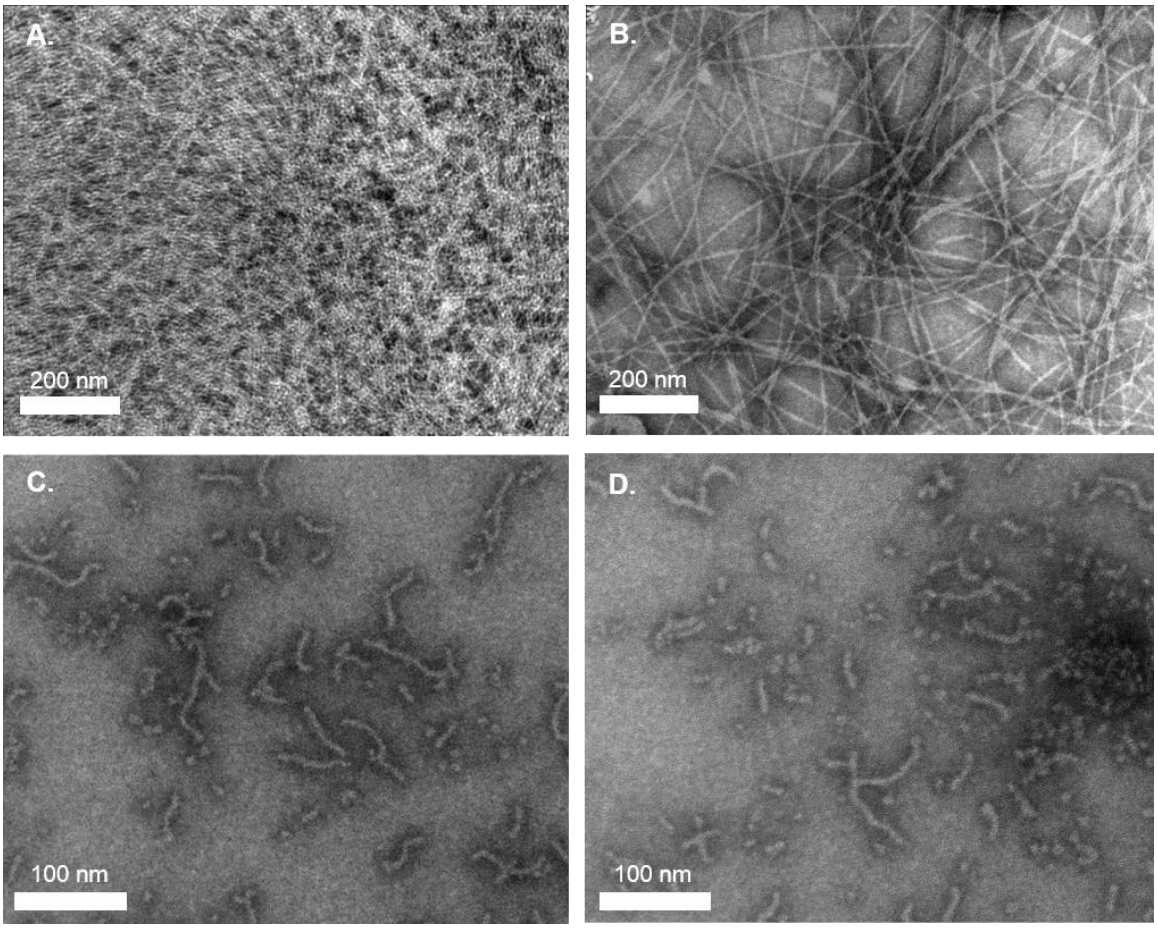

Figure 1.15. H6 equilibrium and early stage morphology (A-B) Representative electron micrographs of H6 incubated for longer than three years at 4 °C (C-D) Representative electron micrographs of H6 branching and punctate structures.

H6 does not bind Congo Red

One potential explanation for the unexpected fiber morphology was the misfolding or unfolding of H6 leading to different modes of aggregation. Although H6 clearly formed triple-helices as measured by CD, the peptides were marginally stable with a T_m ~15 C ° (Figure 1.11, 1.12). Based on its physiochemical properties, the $(LIG)_3$ hydrophobic domain of H6 has a high computed propensity to form amyloids,[69] and previous examples of cross-beta structures formed by amyloidogenic regions within CMPs have been described.[19, 70] However, H6 fibers did not bind Congo Red, a common amyloid probe[71]

(Figure 1.16), and no evidence of the twisted cross-beta structure was noticed in TEM images. H6 fibers were morphologically similar to those formed by other CMPs where triple-helical structure was required for higher-order assembly.[72]

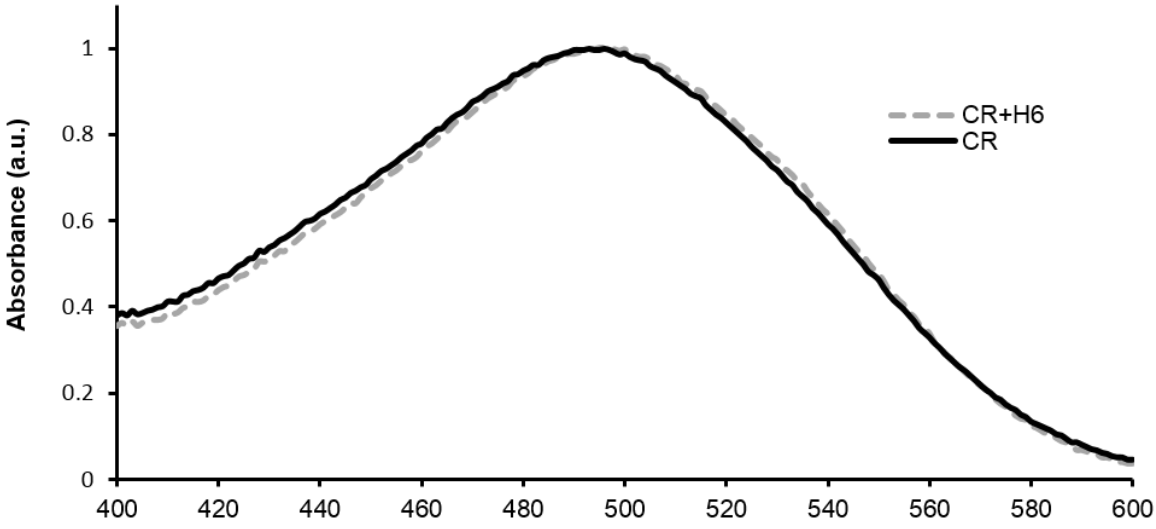

Figure 1.16. Congo Red binding assay on H6.

H6 stability not morphology affected by solvent polarity

This section focuses on the effects of solvent polarity on H6 stability and assembly. Nonpolar solvents have been observed to stabilize the triple-helix fold[73]. In addition to stabilizing the triple-helical fold, we hypothesize that changing the solvent polarity will also effect H6 assembly. Circular dichroism of H6-methanol solutions showed that methanol has a stabilizing effect on the triple-helix at and above 20% (v/v) (Figure 1.17 A, B). Of note, a solution of H6 in 100% methanol resulted in a T_m of ~48 C °, which is close to the boiling point of methanol 64.7 C °. H6 solutions of 1-30% (v/v) of methanol were observed by TEM to form fibers (Figure 1.17 C). Additional experiments using stronger nonpolar solvents are predicted to dramatically change H6 assembly morphology by either

preventing assembly or changing the morphology to a disc. Furthermore, the method of solvent addition is also predicted to change the stability and assembly morphology.

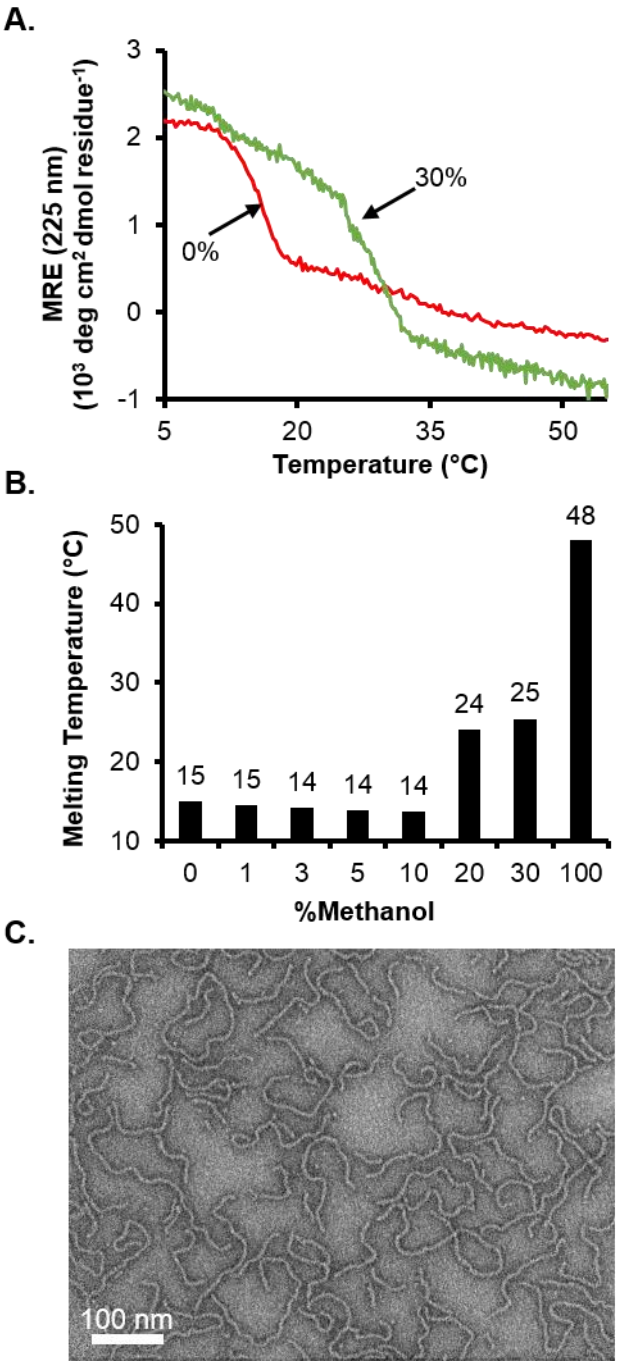

Figure 1.17. Effects of solvent polarity on H6. (A) Circular dichroism thermal denature plot of H6 with 0 (red) and 30% (green) methanol, (B) Melting temperatures of H6 with 0, 1, 3, 5, 10, 20, 30 and 100% methanol, and (C) Electron micrograph of H6 in 30% methanol.

H6 helical packing model

If adjacent rods packed at an angle, then the resulting fiber would be a super helix, or a helical tape, of CMPs where fiber thickness matches the length of a peptide (Figure 1.18).[74] This type angular packing was not incorporated in the DLA coarse-grained simulation, so this fiber morph would not be predicted. AFM of H6 reveals an undulating height profile varying from 1.5 to 6 nm, consistent with a helical tape (Figure 1.19).

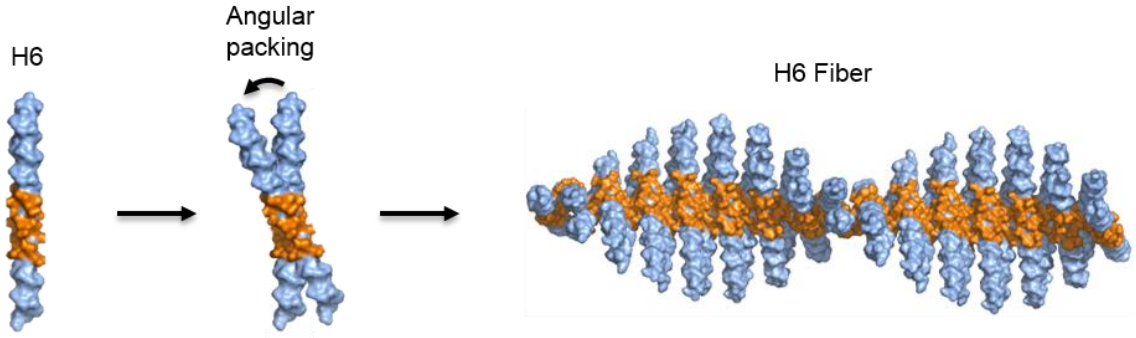

Figure 1.18. Helical tape assembly model for H6 fiber.

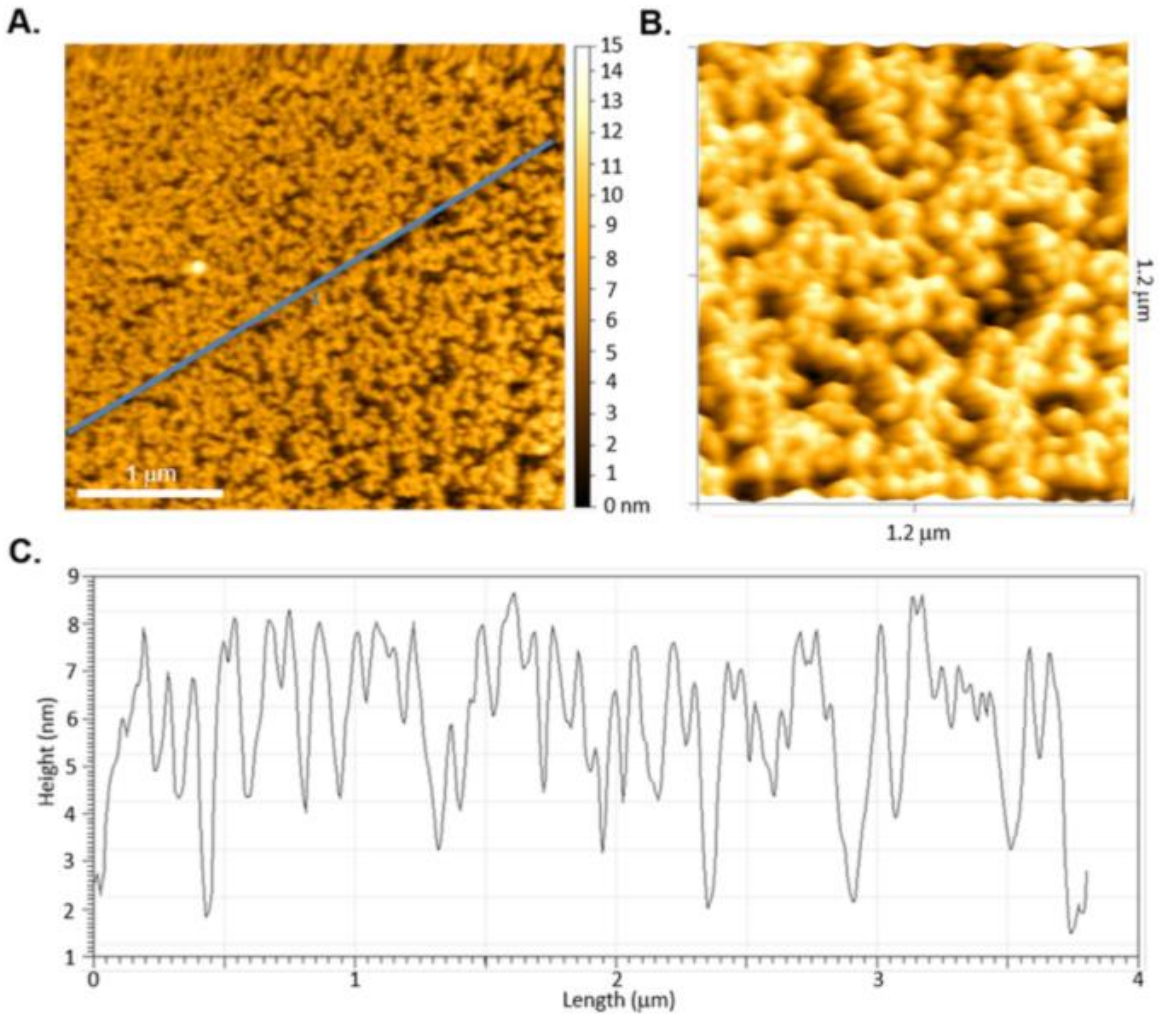

Figure 1.19. H6 height characterization. (A) H6 AFM (B) 3D AFM highlighted view of H6 (C) Height profile of blue line in (A).

Morphology of self-assembled structure is determined by hydrophobicity

To determine whether the fiber morphology arises from the low structural stability of H6 or its increased hydrophobicity, we synthesized the peptide QH4 (Table 1) consisting of two hydrophobic triplets and one Gln-Gln-Gly. The polar amino acid glutamine is well tolerated at both the X and Y positions of the triple-helix, but does not confer the same stability as proline or hydroxyproline,[61] making it a suitable proxy for the low stability of H6 while maintaining the hydrophobicity of H4. QH4 folded into a triple-helix (Figure

1.10) and despite a 9°C lower T_m than H4 (Figure1.11), the peptides formed plates (Figure 1.20), thus supporting a model where length of hydrophobic region is the primary determinant of assembly morphology. QH4 plates resembled those of the florets seen in H4.

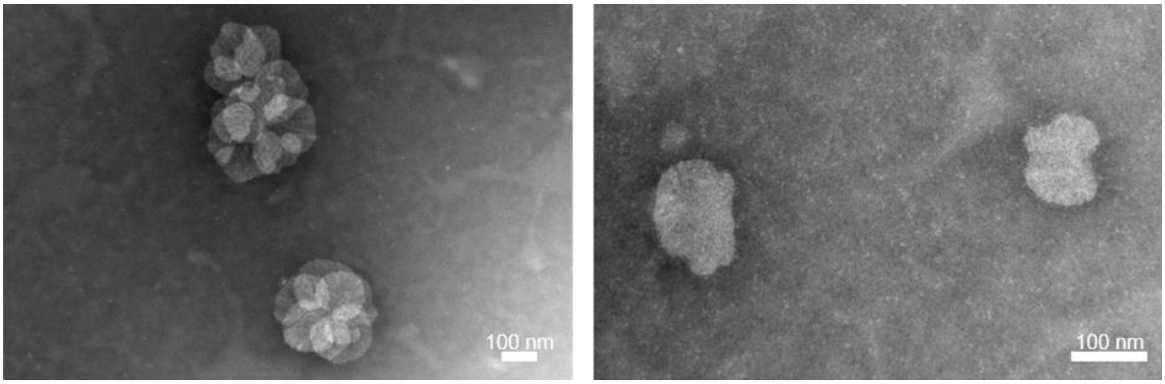

Figure 1.20. TEM of QH4.

Polymorphic behavior of H5

The collective properties of existing CMPs pointed to a morphological bifurcation point related to hydrophobicity between H4 and H6. Below this point, the disc morph dominated, with only minor populations of fibers seen in H2, H3 and H4 TEM fields. Above this point, fibers were predominantly formed. To further pinpoint where this transition occurred, we examined H5 (Table 1). H5 folded into a triple-helix with the stability of 24°C (Figure 1.10, 11). Consistent with it sitting on a saddle point between two morphologies, H5 was observed by TEM to assemble into major populations of *both* fibers and plates (Figure 1.21, 1.22). Fiber widths and lengths were calculated to be 7.5 - 9.5 nm and several microns, respectively. H5 disc diameters were 50 – 200 nm. AFM confirmed the presence of fibers and plate morphs with fiber heights between 1-3 nm and plates with heights 6 – 11 nm (Figure 1.21 B). Similarly to H6, H5 did not bind Congo Red (Figure 1.23).

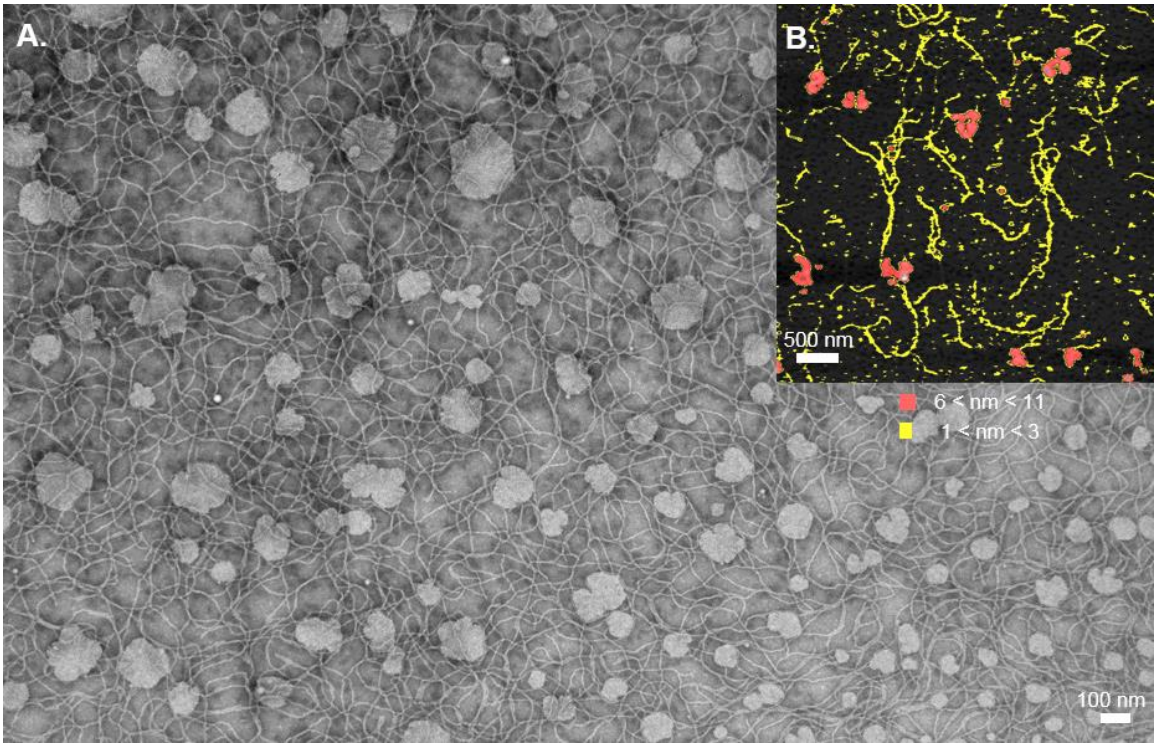

Figure 1.21. Characterization of H5. (A) TEM (B) AFM where yellow signifies 1-3 nm height features and red signifies 6-11 nm height features.

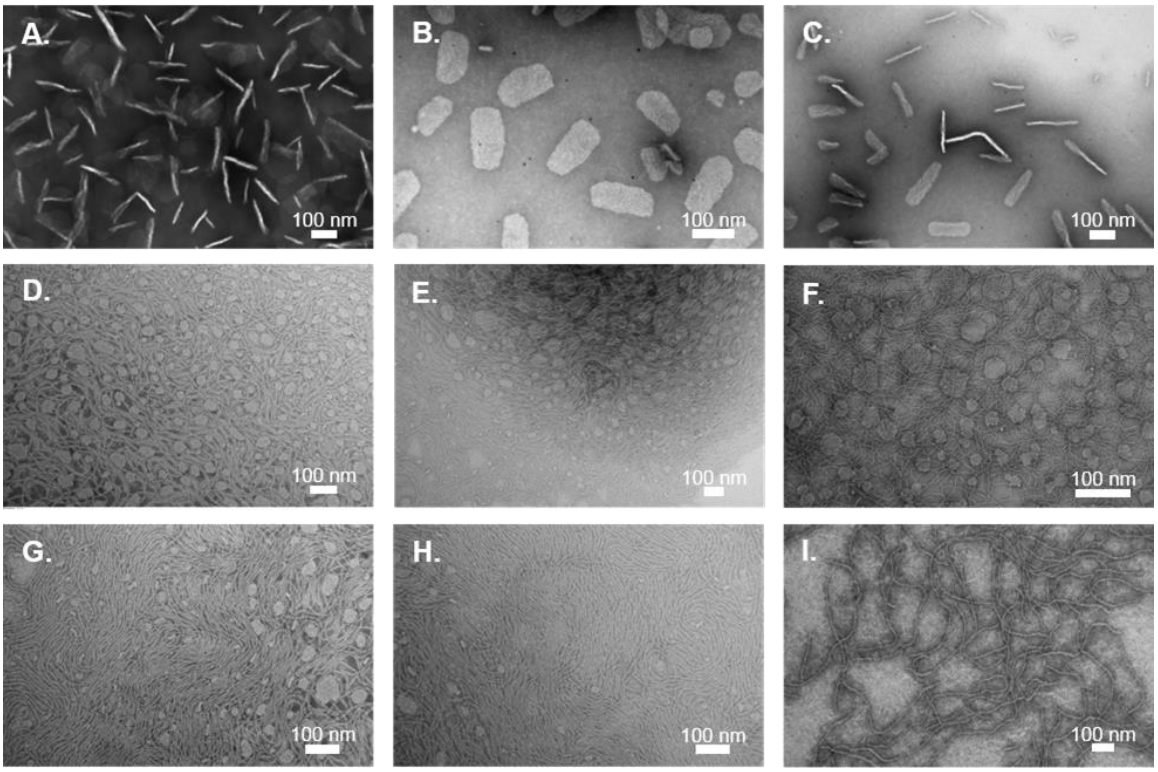

Figure 1.22. TEM of H5 morphologies. (A-C) TEM fields of primarily discs, (D-F) TEM fields of discs and fibers and (G-I) TEM fields that were primarily fibers.

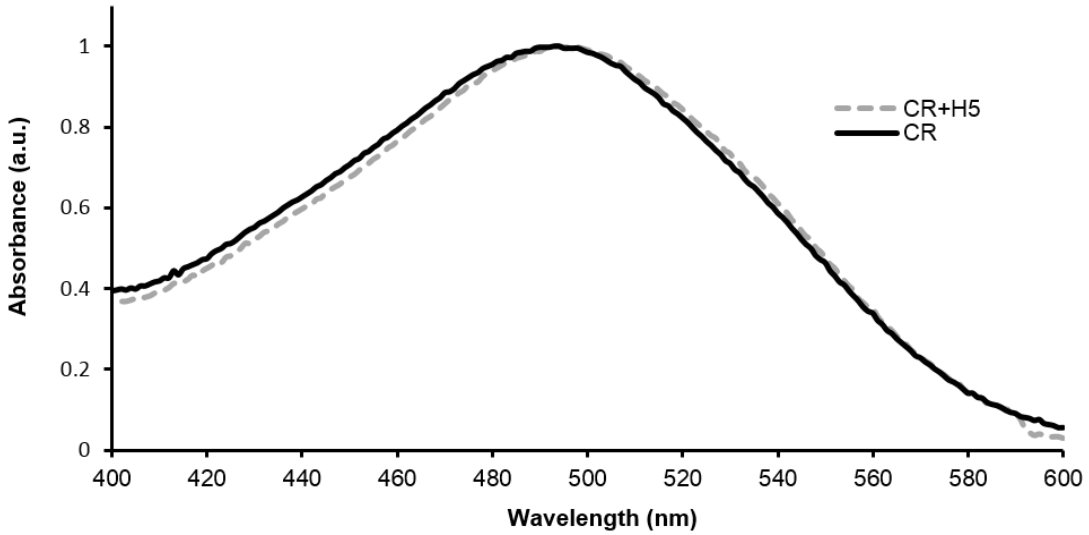

Figure 1.23. Congo Red binding assay for H5.

Solvent polarity impacts H5 self-assembly morphology and triple-helix stability

Molecular switches driven by environmental factors such as pH, temperature, and solvent polarity are attractive for building stimuli responsive materials[75]. We predict that solvent polarity can be used as a stimuli to convert H5 into either discs or fibers only. Electron microscopy of H5 solutions of 20% (v/v) acetonitrile reveals that the primary morphology are sheets and cylinders (Figure 1.24), which we suspect are hollow and are formed due to sheets folding on itself.

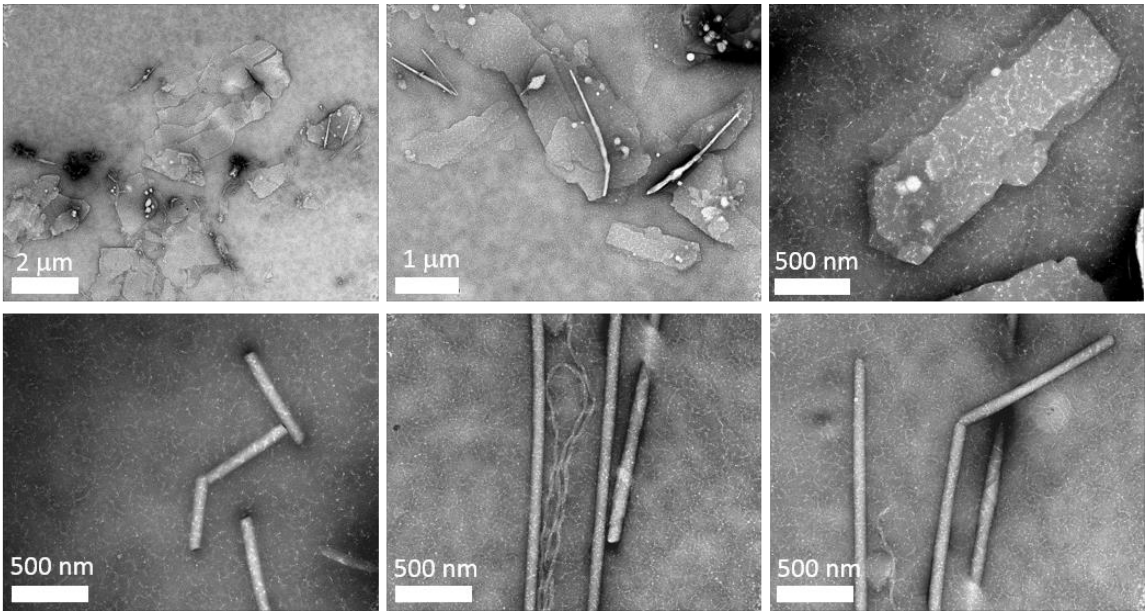

Figure 1.24. H5 morphology in acetonitrile. (top) Representative electron micrographs of sheets formed by H5 in 20% (v/v) acetonitrile. (bottom) Representative electron micrographs of cylinders and fibers formed by H5 in 20% acetonitrile (v/v).

H3-4 discs sequester hydrophobic dyes

Self-assembling amphiphilic peptides and peptidomimetics can produce lipophilic domains capable of sequestering small molecules and acting as vehicles for drug delivery.[5, 76, 77] The H4 discs are similarly able to bind solvatochromatic dyes such as Sypro Orange (SO) and

Nile Red (NR), which has enhanced quantum yield when bound to hydrophobic surfaces. Confocal microscope images of H4 + SO (Figure. 1.25 A) and H4 + NR (Figure 1.26) show effective disc sequestration of the dye. H3 was also observed to bind NR (Figure 1.27). SO intensity increases when discs are heated above the T_m of H4 (Figure 1.25 B), which is typically observed for globular proteins upon unfolding.[78] This suggests that denaturation of the H4 discs exposes greater hydrophobic surface area for SO binding.

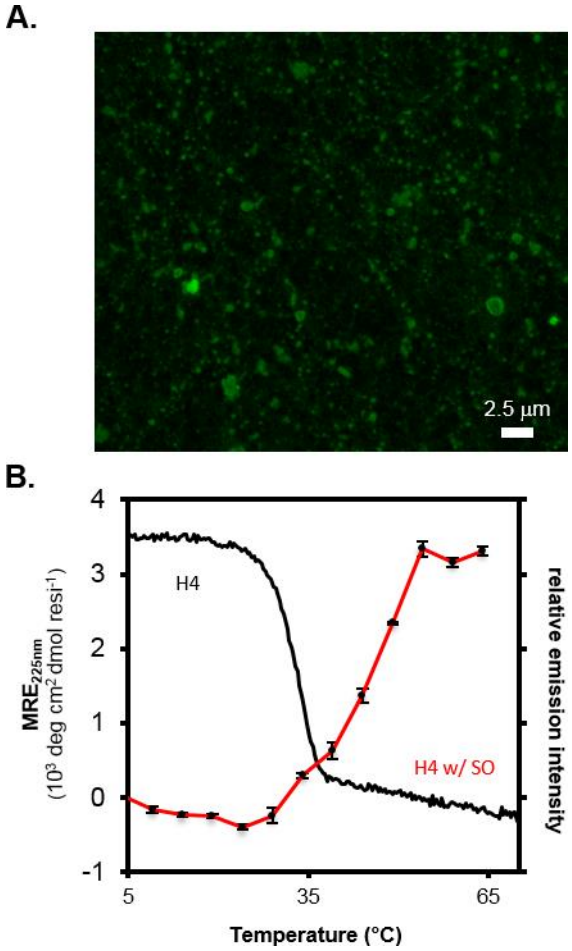

Figure 1.25. Sypro Orange binding assay on H4. (A) Confocal fluorescence microscopy of SO-H4 solution. (B) Differential scanning fluorimetry of H4 with Sypro Orange (red) and circular dichroism thermal denaturation melt of H4 (black).

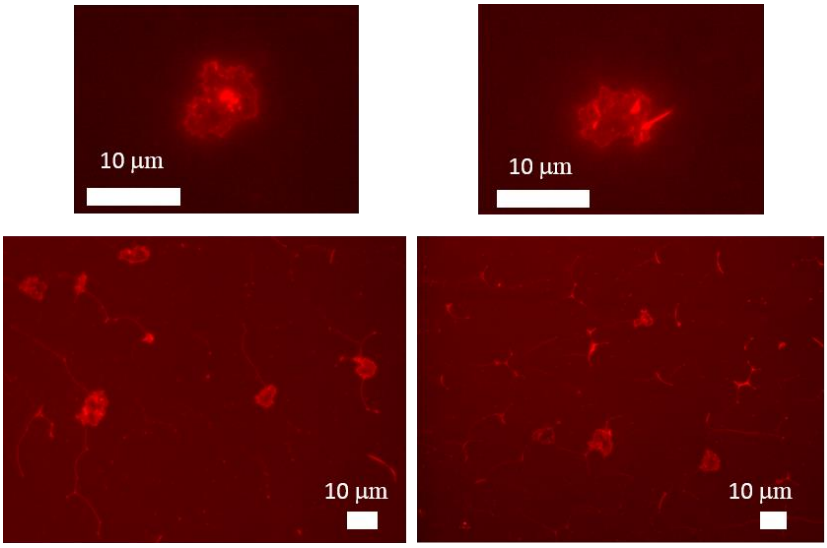

Figure 1.26. Nile Red binding assay on H4.

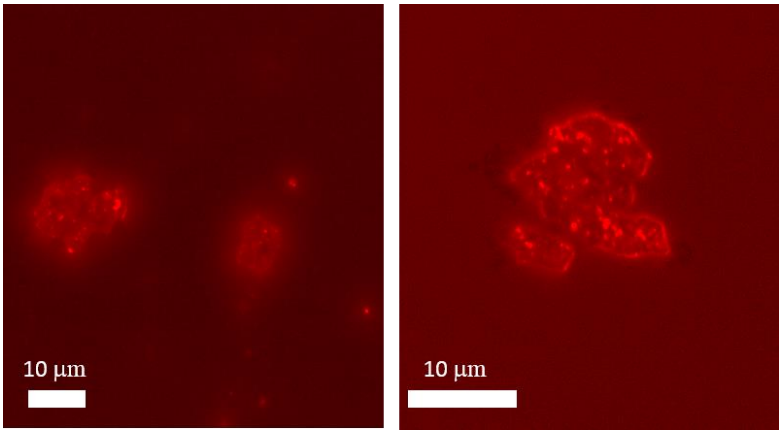

Figure 1.27. Nile Red binding assay on H3.

Solvent pH effects H4 assembly morphology not triple-helix stability

This section investigates the affect solvent pH has on H4 assembly. The effects of pH on H4 assembly is of interest because we want to investigate ways of releasing the hydrophobic molecules, such as those in the last section. Electron microscopy of H4 prepared in acidic conditions showed the formation of sheets (Figure 1.28 A, B) and cylinders (Figure 1.28 B). Cylinders are suspected to be formed by the folding of sheets

and are hypothesized to be hollow. As expected, H4 folded into a triple-helix (Figure 1.29 A) with the same stability in both neutral and acidic conditions 34 °C (Figure 1.29 B) because of not having ionizable amino acids.[79, 80] The observed pH induced morphology shift of H4 assembly from discs to sheets maybe due to increased water release from disc surfaces, leading to a more energetically favorable binding surface for peptides.[81]

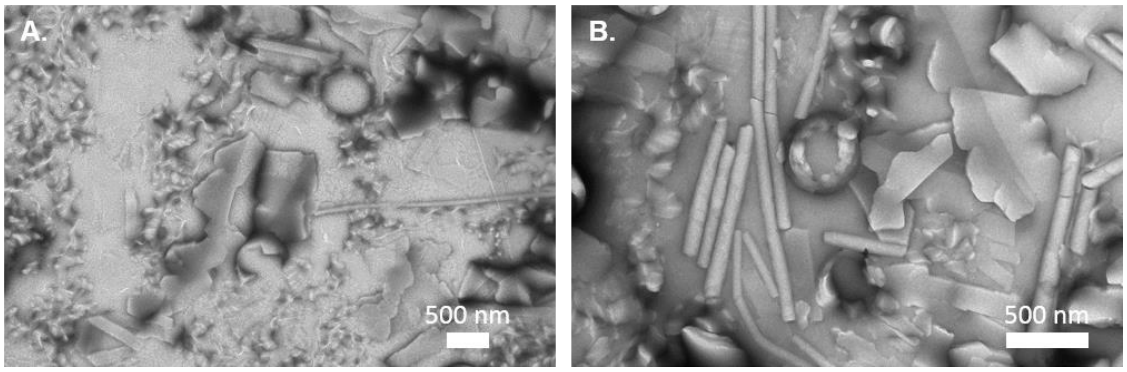

Figure 1.28. Morphology of H4 in acidic conditions. (A) Representative electron micrograph of 0.2 mM H4, and (B) Representative electron micrograph of 6.0 mM H4. Both solutions contain 0.2 N acetic acid at pH 3.3.

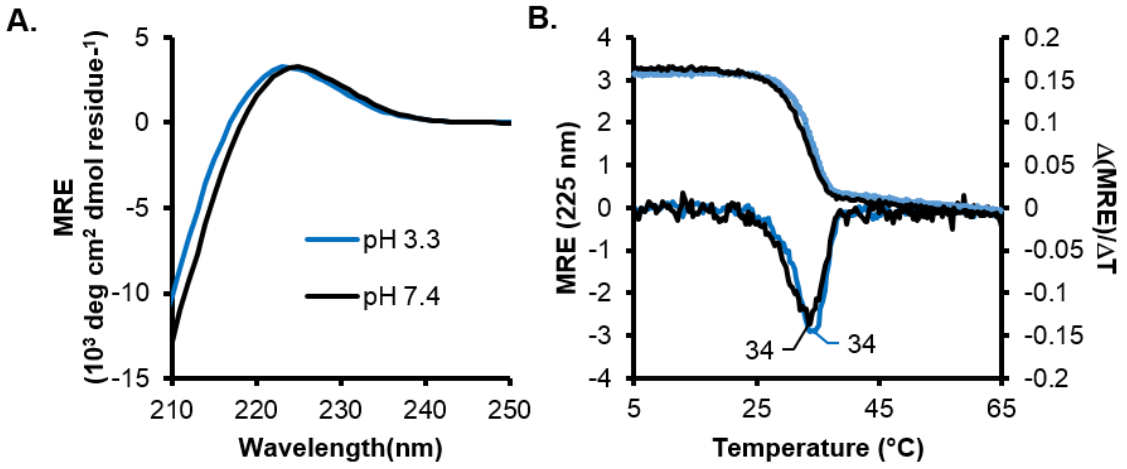

Figure 1.29. Circular dichroism of H4 in acidic conditions. (A) Wavelength scan of H4 in neutral and acidic conditions, and (B) Thermal denaturation of H4 in neutral and acidic conditions.

Multi-component peptide nanostars and plates-on-strings

Natural collagen proteins make use of hydrophobic interactions to bind other proteins such as integrin receptors.[82] Likewise, we would expect the hydrophobic CMPs to interact with other molecules that also present exposed hydrophobic surfaces. We explore two cases of multi-component interactions – H4/H6 nanostars and H4/amyloid peptide tandem plate assemblies.

As postulated earlier, one possible structural model for the fiber morph is that of a helical tape[74] formed by CMP rods that pack at an angle. Assuming H4 discs have hydrophobic edges and H6 fibers have hydrophobic ends, combining the two would be expected to generate conjoined nanostar-like structures with fibers alighting from the edges of discs (Figure 1.30 A). Consistent with this, EM images of H6:H4 mixtures revealed fibers branching from discs (Figure 1.30 B, 1.31). We note similar behavior in some H5 TEM fields where fibers are seen originating from the edges of discs.

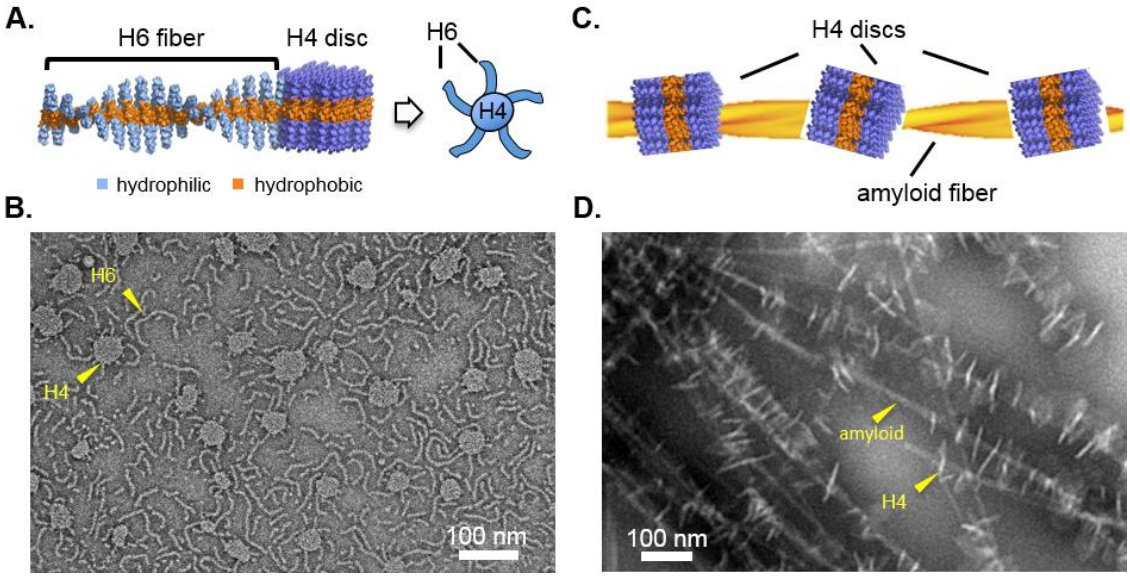

Figure 1.30. Multi-component assembly H4-H6 and H4-NdQ. (A) Hydrophobic interactions between H6 fiber and H4 disc models suggest formation of nanostar structures

when mixed. (B) TEM of a 2:1 ratio of H6:H4 peptides confirms fibers matching H6 morphology attached to the disc edges. (C) The NdQ peptide forms an extended hydrophobic amyloid fiber that should interact with the hydrophobic edges of H4 discs. (D) TEM of a 1:1 ratio of H4 discs mixed with amyloids fibers shows tandem discs aligned perpendicular to both the grid surface and the long axis of the amyloid fibril, consistent with the proposed interaction model.

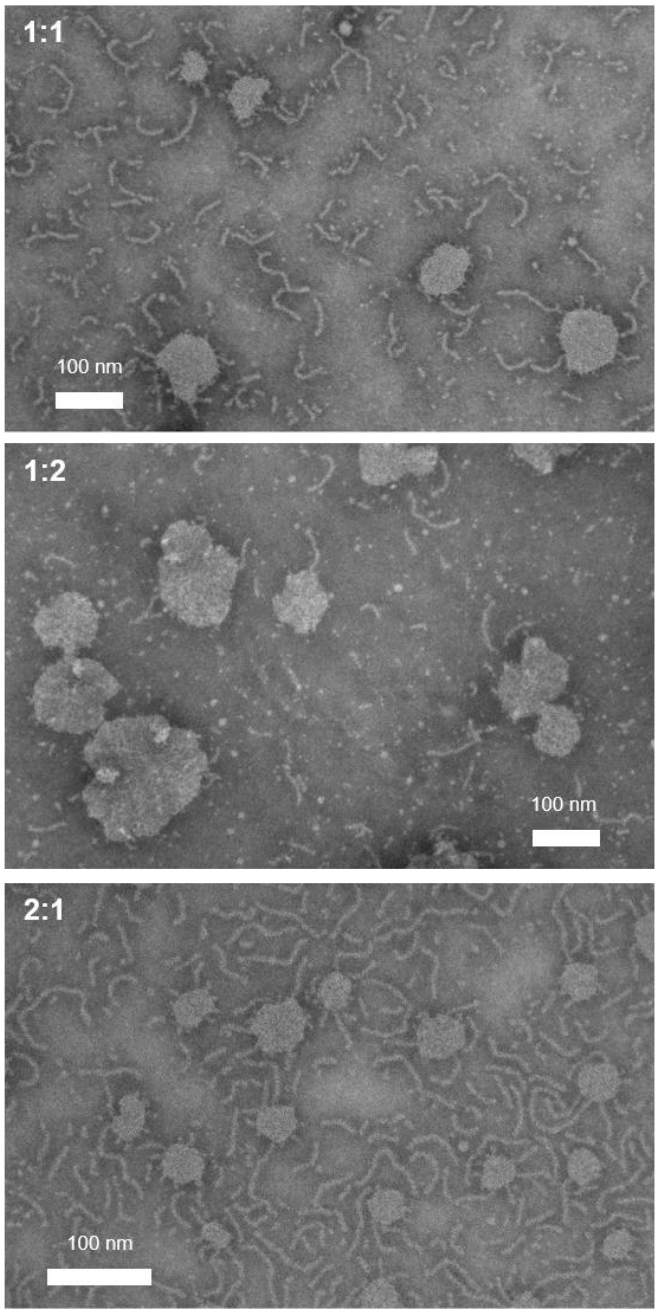

Figure 1.31. TEM of ratios of H6-H4 mixtures.

The observation that H6 fibers only interact at their ends suggests the packing geometry of the fiber shields the central hydrophobic residues. If an alternative type of fiber were to present a hydrophobic surface, then discs might be expected to attach along the fiber length rather than only at the ends (Figure 1.30 C). Using a short hydrophobic amyloidogenic peptide, NdQ, we were able to generate extended beta-helical structures that would presumably present hydrophobic side chains along the full length of the fiber (Figure 1.32). When these were combined with H4, multiple tandem discs attached perpendicularly to the fibers (Figure 1.30 D). In several instances, discs would bundle individual fibers together, aligning them over micron distances.

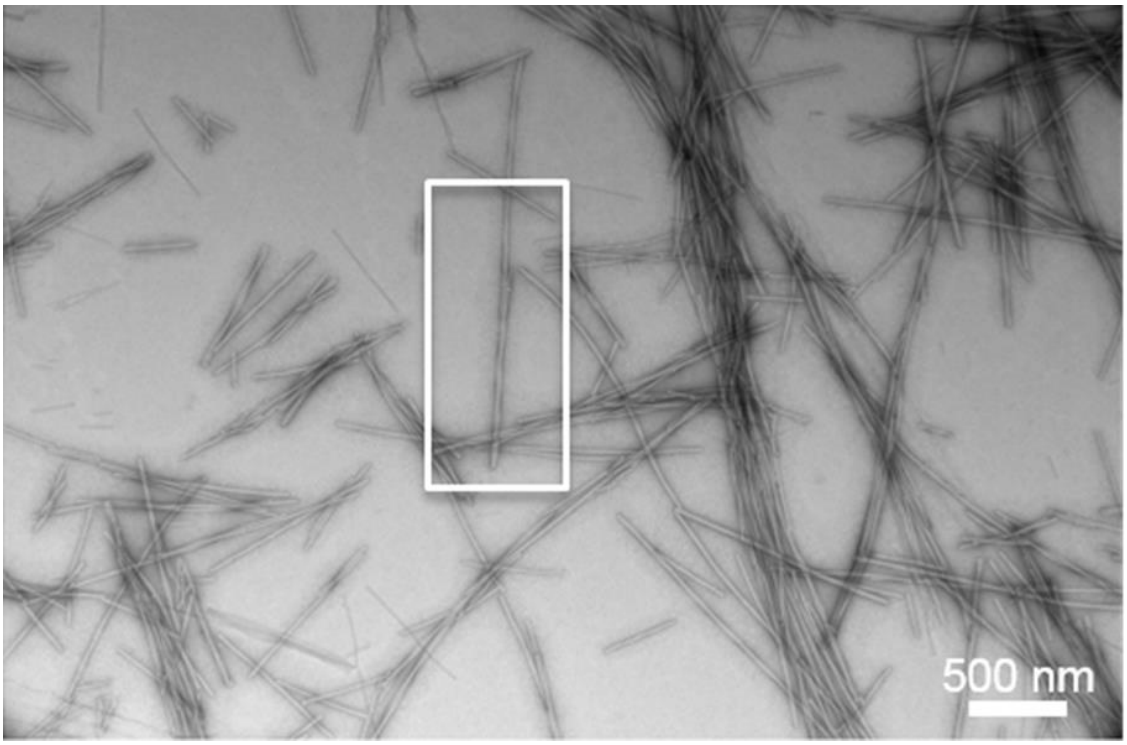

Figure 1.32. TEM of the amyloid forming peptide NdQ. (Box) shows characteristic amyloid fiber betal-helical twist.

1.4 Conclusions

Collagen is on the cusp of solubility, which is functionally advantageous given its role in higher-order assembly processes in biology. Small changes in sequence hydrophobicity induce rapid assembly and large changes in morphology. This sequence-dependent sensitivity of CMP assembly has been observed many times before: increasing core POG repeat length of metal-ligated CMP assemblies results in shifts from meshes to florettes to saddle-like structures[83] and increasing the number of core bipyridines lead to the formation of fibers,[84] disks[85, 86] and then to hollow spheres.[86] Electrostatically-driven assembly of CMPs containing acidic and basic regions show similar sequence sensitive polymorphism, where arginine and glutamic acid formed D-periodic fibers[87] whereas lysine and aspartic acid formed hydrogels.[88] This is a general property expected of self-assembling system where the sum of small forces in local molecular interactions exerts significant influence on the properties at the aggregate level.

Certain features of the H-series CMPs could serve as starting points for developing functional materials. Given the observation that single amino acid changes can drive dramatic variations in the self-assembly phenotype, it should be possible to design switchable peptides that interconvert between two nanoscale morphologies based on environmental perturbations, such as solvent pH and polarity, leading to potential biosensor or controlled-release applications. Additionally, the ability to control structure at the scale of microns and above is important for engineering materials that interact with cells.[89, 90] H4 and H6 CMPs form structures on this length scale and could be further developed to provide cell-scale interactions. The ability of these peptides to interact with other proteins suggests a modular design strategy where structure and function are separated into different

molecules. A structural matrix scaffold that forms micron scale fibers such as the amyloids or natural collagens could be decorated with peptides that provide chemical or biological functionality.

In the early evolution of natural collagen, such a propensity for polymorphism might have been advantageous, allowing smaller molecules to rapidly explore and optimize the assembly of functionally useful nanoscale forms. In extant collagens which are generally much larger, hydrophobicity is tightly constrained, with only two triplets in Type I collagen out of three-hundred having a tandem pair of hydrophobic amino acids at X and Y positions akin to the H2 peptide. Longer hydrophobic stretches such as those incorporated into H3-H6 are extremely rare, with H4-like LIGLIG sequences in Types V and XI collagens. How such hydrophobic sites function in the context of natural collagen *versus* model peptides will be an interesting area of future study.

Higher resolution structural information will be critical in order to fully understand how hydrophobicity drives nanoscale morphology. A change in the packing angle between adjacent triple helices would explain the difference between the fiber and disc morphs. However, the relative amounts of these two species going from H2 to H6 could depend both on energetic and kinetic factors in assembly.

Even within the scope of simple *h*/*p* CMP designs, this work has only just scratched the surface of the sequence diversity available. In addition to length of the hydrophobic core, further designs incorporating different patterns of hydrophobic/hydrophilic residues can be designed to potentially produce a rich library of higher order structures. Further cycles of simulation, design and experimental characterization will likely improve the accuracy of

predictive tools and enhance our understanding of hydrophobic forces in protein self-assembly.

1.5 Materials and Methods

Diffusion Limited Aggregation (DLA)

The DLA simulations were conducted in a 3D hexagonal lattice. The seed for each DLA simulation was a discretized rigid rod composed of ten spheres. Each sphere represented an XYG triplet of amino acids and was either hydrophobic (*h*) or polar (*p*). A seed rod was placed at the center of the simulation and additional rods were released and allowed to move randomly until either contacting the aggregate or until they moved too far away from the aggregate and were discarded (Figure 1.2 A). Simulations of each of the 1024 possible sequence combinations were repeated 1000 times to assess convergence of nanoscale morphs. A suitable interaction state was defined as greater than two hydrophobic groups in contact. Representative assembly structures were chosen for presentation. The peptide H4 was modeled after the DLA rod *pppphhpppp*. The peptide H6 was modeled after *ppphhhpppp*. JAVA source code for the simulations are on the ACS Nano website (http://pubs.acs.org/journal/ancac3).

Peptide synthesis and purification

Peptides were synthesized using solid phase FMOC chemistry, purified by reverse phase HPLC and verified using mass spectrometry at the Tufts University Core Facility (http://tucf.org). H6, H5, H4, H3, QH4, and SH4 were purified to >95% purity by reverse phase HPLC. Due to issues with synthesis, H2 purity was low, ~ 80%. N- and C- termini were acetylated and amidated, respectively. Peptides were dialyzed in filtered deionized water, lyophilized and kept in -20 °C. Peptides were weighed and suspended in 10 mM

phosphate buffer (sodium phosphate monobasic monohydrate and sodium phosphate dibasic heptahydrate) pH 7.4. H6 methanol solutions were prepared by solvating 10 mM phosphate buffer pH 7.4. Collagen peptide concentrations were confirmed by measuring the absorbance at 214 nm, using an extinction coefficient of 2200 M^{-1} cm^{-1} per peptide bond, on an AVIV Model 14DS UV-VIS Spectrophotometer. Peptide NdQ, sequence NYFYSLFdQG (dQ = D-glutamine), was purified to > 95% and solvated in 50 mM Tris pH 8.0 buffer with 30 mM NaCl, 2.5% glycerol, and 10 mM DTT.

Circular Dichroism (CD)

CD experiments were performed on an Aviv Model 400 Spectrophotometer. Optically matched 1 mm path length quartz cuvettes (Model 110-OS, Hellman USA) were used. Peptide solutions were prepared and kept at 4 °C for a minimum of 48 hours prior to measurement. Time course CD measurements monitored at 225 nm show that after 48hrs the triple-helical conformation has reached a plateau (Figure 1.30). It is interesting to note the longer lag time of folding for H6. Before measurement, peptide concentrations were diluted to 0.20 mM to reduce excess absorption at lower wavelengths. Wavelength scans were taken from 260-200 nm at 4 °C. Thermal denaturation melts were monitored at 225 nm with a temperature step of 0.3 °C and a two minute equilibration period from 4 – 80 °C. Savitzky– Golay smoothing of the denaturation profiles was carried out over a span of 11 points using a third-order polynomial.[91] The melting temperature (T_m) of each melt was assigned to the extrema of the first derivative of the denaturation profiles.

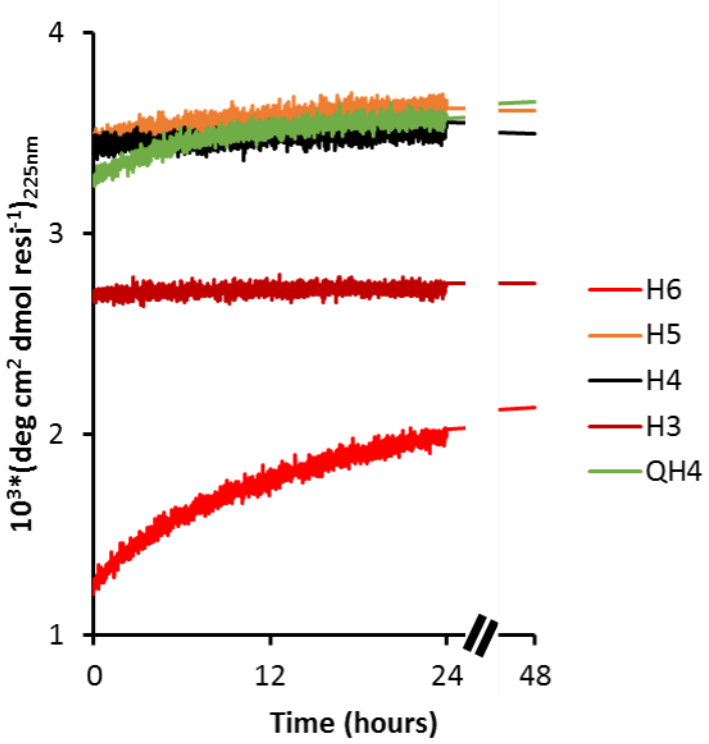

Figure 1.33. CD folding kinetics of the H-series peptides.

Congo Red binding assay

A stock 5 μM Congo Red (CR) (C6767, Sigma Aldrich) solution was prepared using deionized water. Concentration of CR was measured using the absorption value at 505 nm and the extinction coefficient of 59,300 M^{-1} cm^{-1}.[71] A volume of 12 μL of 3 mM peptide solution was mixed with 280 μL of stock CR solution. Wavelength measurements of peptide solutions with and without CR, CR alone and buffer were taken from 400 – 600 nm.

Transmission Electron Microscope (TEM)

Collagen peptide samples were prepared at concentrations ranging from 2 mM to 10 mM and incubated at 4 °C for two weeks prior to imaging. NdQ was prepared at concentrations

of 0.5 mM and fibers were observed within three hours. H6 and H4 samples were combined 1:1, 1:2 and 2:1 wt./wt. ratios with a total weight of 3.6 mg and dissolved in pH 7.0 phosphate buffer. NdQ and H4 samples were combined at 1:1 molar ratio. A volume of 5 μL of peptide solution was deposited on a copper grid coated with carbon (Electron Microscopy Sciences CF-400-Cu), then a time of two minutes was passed to ensure deposition, and excess peptide was wicked away using filter paper. The grid was then negatively stained with 1% phosphotungstic acid (PTA) pH 7.0 for two minutes and excess stain was wicked away. Afterward, the grids were left to dry for at least two minutes in ambient conditions before imaging. Images were captured using a Philip 420 Electron Microscope at 80 kV.

Atomic Force Microscopy (AFM)

For H4 and H6 solutions the method of Xu *et al*. was followed.[64] Briefly, silicon wafers were cleaned with a series of alcohols and subjected to sonication and incubation. The silicon wafers were then dried with nitrogen, placed on a glass slide with adhesive and cooled before sample deposition. Excess sample was then removed using deionized water and dried for 20 minutes at room temperature. Three scans were taken for each measurement using an Agilent 5500 in tapping mode, with a Multi75a probe (Bruker MPP-21120-10).

H5 solutions were visualized using a MFP-3D-Bio AFM with a Nikon microscope (Asylum Research, Santa Barbara, CA 93117). Measurements were made using contact mode in air with a SNL-10 probe (Bruker) on samples that were prepared on freshly cleaved mica. Samples were rinsed with deionized water and dried with argon prior to

imaging. Data were analyzed using a combination of Gwyddion,[92] WSXM,[93] and MATLAB[94] software packages.

Wide-angle X-Ray Scattering (WAXS)

Supersaturated solutions of H4 were prepared by placing peptide powder in a glass capillary (Special glass 1 mm O.D.; Charles Supper Company, Natick, MA) and then solvent was added on top of the powder. The solvent was allowed to soak the powder for at least two weeks at 4 °C to ensure sufficient hydration.

The capillary was mounted in a Rigaku Geigerflex X-ray diffractometer with a sealed Cu target X-ray tube and Rigaku Osmic mirror focusing monochromator (CuKα; $\lambda = 1.5418$ Å). The nominal sample to detector distance was set at 2.2 cm. Data were collected by scanning along 2θ.

1.6 Acknowledgements

We thank the National Institutes of Health grant DP2 OD-006478 and the GAANN fellowship for supporting this research. We also thank John Kahn for performing and analyzing WAXS data, Daniel Grisham with assistance with image processing and analysis, Patrick Nosker for the amyloid peptides, Michele Vittadello, Hiroshi Matsui, and Michael Kopka for access to AFM, data collection and assistance with analysis, Smita Patel for access to fluorimetry, Jonathan Boyd from Leica-Microsystems for access to confocal microscopy, and Srinivas Annavarapu for helpful discussions.

Chapter 2

Hydrophobic collagen peptide discs act as substrates and scaffolds for the noncovalent organization and higher order assembly of natural proteins

2.1 Abstract

Hydrophobic collagen peptide discs are used to drive supramolecular assembly of natural proteins. Hydrophobic driven multi-component assembly is prevalent in the cellular environment and difficult to design. Using a scaffold-substrate scheme involving synthetic hydrophobic collagen peptide discs, we show that higher-order assembly and organization of natural proteins is possible. A series of natural fibrous proteins with varying secondary structure and exposed surface hydrophobicity are shown to scaffold organized binding of discs and subsequent binding induces bundling of the fibrous proteins. The discs served as a scaffold for organizing and encapsulating a light harvesting membrane protein. These interactions provide insight into the surface hydrophobicity of natural proteins and a means to design nascent interaction sites between natural and synthetic proteins to build advanced biomaterials for nano-medical and bio-solar applications.

2.2 Objective

We have designed a set of collagen mimetic peptides (CMP) (H-Series) that self-assembled into discs (QH4, SH4, H2-4) and fibers (H6), (where H# denotes length of hydrophobic amino acids).[95] H4 discs were shown to be polydisperse in size and shape, highly flexible and to ostensibly present hydrophobic groups that allowed for edge-wise interactions with other hydrophobic surfaces. These interactions included perpendicular alignment with the electron microscopy grid and the surface of a synthetic amyloid fiber (NdQ). The resulting ‘plates-on-a-string’ H4-NdQ and ‘seastar’ H4-H6 composite nanostructure inspired the co-incubation of discs with natural proteins displaying hydrophobic surfaces to test whether

these interactions were specific to only synthetic amyloid-like and collagen peptide nanostructures or a general phenomenon.

Interactions between the H-Series discs and the natural fibrous proteins collagen type I (COLI), shrimp tropomyosin (ST), and alpha-synuclein (AS) are shown here to promote long-range organization of discs and in the case of ST and COLI, binding of discs mediated the assembly of the fibrous protein itself.

Additionally, parallel edge-wise interactions between the H4 discs and H6 fiber edges resulted in 'nanostar' composite nanostructures, suggesting H4 can also serve as a scaffolding agent for proteins with hydrophobic edges. H4 discs are shown to organize the membrane protein Reaction Center Light Harvesting Complex 1 (RC-LHC1).

We show, using transmission electron microscopy (TEM), that H-Series discs can act as a substrate and scaffold when interacting with natural proteins. Similar perpendicular and parallel edge-wise interactions were seen between discs and natural proteins as those between H4 and synthetic peptide nanostructures.

2.3 Results

COLI type I as a synthetic CMP organization scaffold

Collagen type I (COLI) is a triple-helical fibrous structural scaffold within the extra cellular matrix. The surface of the COLI triple-helix exposes hydrophobic groups that mediate cellular function [96] as well as *in-vitro* fiber assembly.[97] Neutral pH and heat is required for *in-vitro* assembly. Targeting COLI with structurally mimetic peptides has shown useful for binding and monitoring the activity of collagen *in-vivo.* [98] The structural similarity, length (~300 nm), and scaffolding properties of the COLI triple-helix makes it an attractive

candidate for orientating H-Series discs. Additionally, hydrophobic groups are thought to be very important in collagen fibril self-assembly.[95] We expected edge-wise deposition of the discs similarly to the AS and NdQ fibers previously studied. Additionally, association behavior of the discs was thought to be dependent on the length of the hydrophobic region of the discs.

CMP induced COLI assembly

Co-incubation of non-fibrillar COLI and the H-series discs (QH4, SH4, and H2-H4) resulted in several binding modes where the COLI triple-helix acted as a scaffold for disc organization and the discs acted as a scaffold for COLI assembly (Figure 2.1 A). The following interactions between the COLI triple-helix (TH) and the H-series discs were observed: (1) small discs lined along the TH (Figure 2.1 B), suggesting that the dispersed hydrophobic regions on the TH are sufficient to accommodate a certain size discs, and the available hydrophobic regions on the TH act as a sink for a particular sized disc, (2) large discs were seen to order parallel and perpendicurally along the THs (Figure 2.1 C), (3) long range ordering of discs aligned perpendicularly along the TH (Figure 2.1 D), and (4) multiple discs aligned perpendicularly to mediate TH bundle formation (Figure 2.1E). We note H4 bundling of NdQ fibers was seen previously. We note that discs were observed to be the most ordered in regions where collagen triple-helices were seen to be uniformily aligned along the TEM grid, the discs followed an angular interaction dependence similar to those seen for the flagella samples. Interestingly H2 was seen to predominantly interact laterally with the THs suggesting the size of the exposed hydrophobic surface patch is an important factor for binding preference.

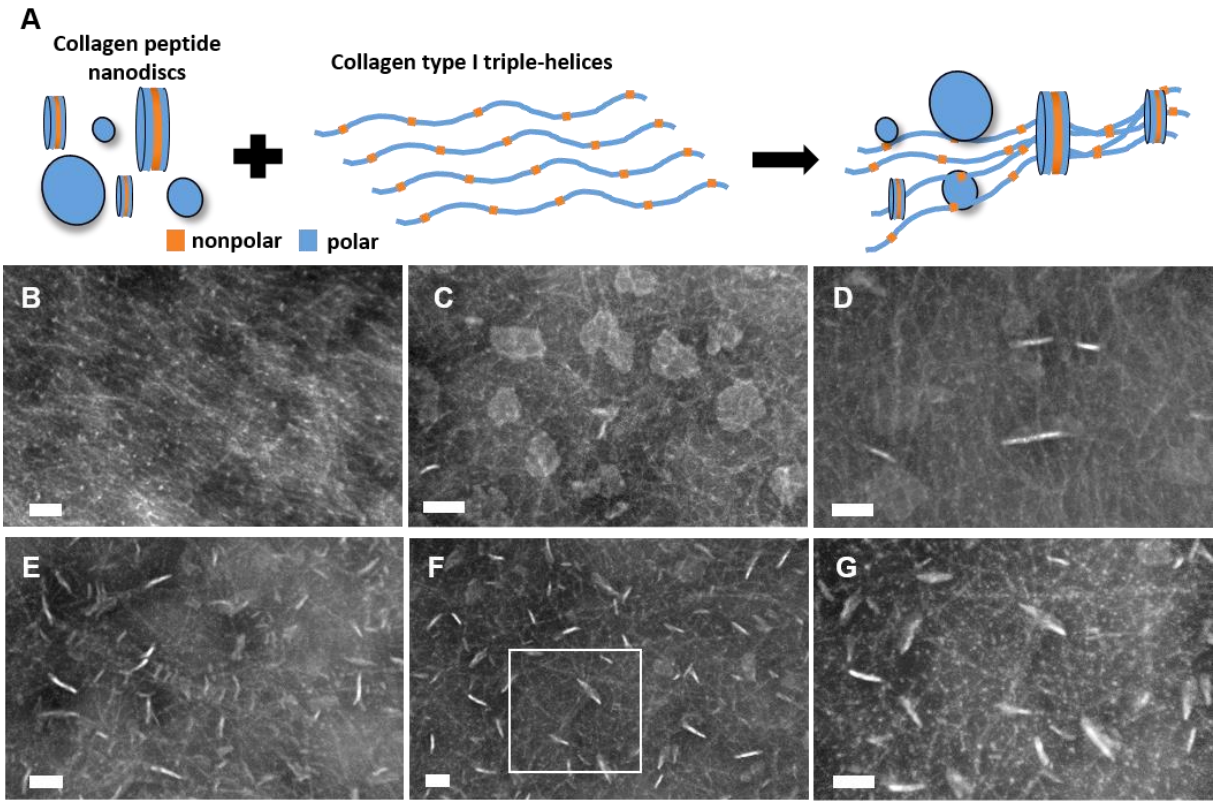

Figure 2.1. H-series discs interacting with COLI triple-helix. (A) interaction scheme between collagen peptide discs and COLI, highlighting the different binding modes observed (B) small H4 discs binding along COLI, (C) large H2 discs laterally along COLI, (D) SH4 edge-wise along COLI, (E) multiple QH4 discs edge-wise binding to multiple COLI, F) H3 bundling of COLI and (G) zoom in of white box in E. Scale bars are 100 nm.

H4 discs bind natural amyloid fibers

Inspired by H4 interactions previous interactions with NdQ, we asked whether H4 binding to amyloids was a general phenomenon. Alpha-synuclein (AS), an amyloid forming protein indicated in Parkinson's disease was expected to provide a sufficient hydrophobic surface along its fibers for H4 discs to bind. The formation of AS fibers is dependent on 12 hydrophobic residues forming the core of the AS fiber. [99] H4 discs were observed by TEM to align perpendicularly and parallel along AS fibers (Figure 2.2) Discs were also seen binding multiple AS fibers together. However compared to H4-NdQ, NdQ fiber surface

was more hydrophobic then AS surface, or the folding and assembly of mode of fibers (cross or straight) for NdQ versus AS lead to differing presentations of hydrophobic groups to interact with H4.

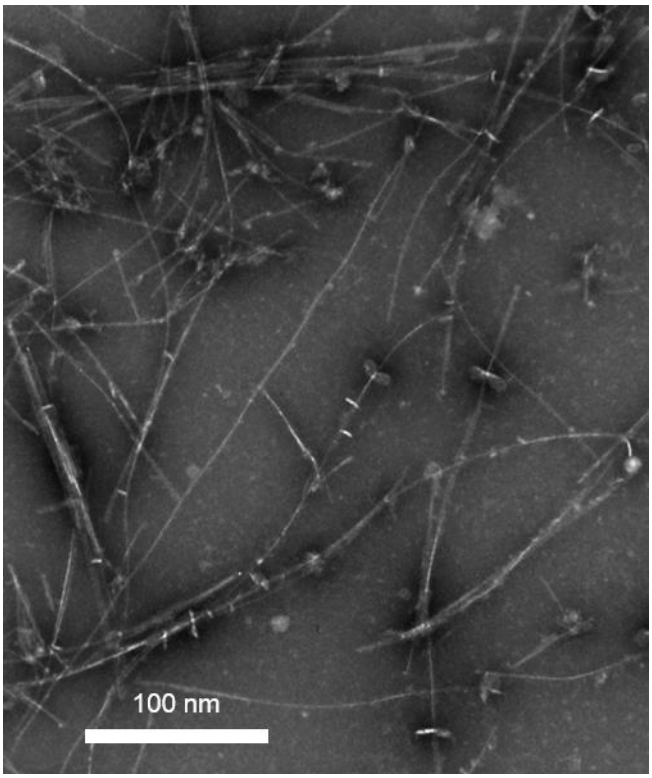

Figure 2.2. TEM of H4 discs interactions with alpha-synuclein.

H4 as scaffold for Reaction Center Light Harvesting Complex 1

Membrane proteins are important for mammalian cellular activity including ion, water and protein transport. They make up ~25% of the human proteome and account for over 50% of drug targets.[100] Methods using liposomes and protein discs have allowed for structural information to be obtained about membrane proteins. Membrane proteins called Light Harvesting Complexes (LHC) allow plants and certain bacteria to perform photosynthesis. Harnessing the power of the LHC would be beneficial to biotechnology (e.g. biocells). We

aim to increase the tool belt for studying LHC membrane proteins. Membrane proteins have hydrophobic edges that are expected to bind to the hydrophobic sided CMP discs. We provide a proof of concept that collagen mimetic nanostructures can be used to organize and encapsulate membrane proteins. We choose to study the Reaction Center –Light Harvesting Complex 1 (RC-LHC1) isolated from the purple bacteria *Rhodobacter sphaeroides* because, when negatively stained, presented an easily distinguishable hexagonal pattern that provided a stringent test for interactions. RC-LHC1 was seen to predominantly bind to the edge of H4 discs when incubated with H4 discs preformed (Figure 2.3 A). The internal hexagonal core that is the RC was confirmed in the presence of H4 discs (Figure 2.3 A). The diameter of the RC-LHC1 was measured to be 11.07 +/- 0.7 nm, which reasonably agrees with previous results.[101] Two forms of the membrane proteins were seen, with and without the hexagonal core visible. Those without the hexagonal core visible suggest the core was either not present or the stain did not bind well leading to lack of constrast. (Figure 2.3 A 1 and 2). Additional interactions observed suggest H4 discs are capable of enveloping the RC-LHC1 (Figure 2.3 B, C). The hydrophobic region of the RC-LHC1 being thicker than that of the H4 discs is hypothesized to be the cause for the predominant surface orientation of the RC-LHC1 along the H4 discs (Figure 2.3 D). Hydrophobic mismatch between membrane spanning proteins and lipid membranes is known to drive proteins to orient along the surface of the membrane.[102]

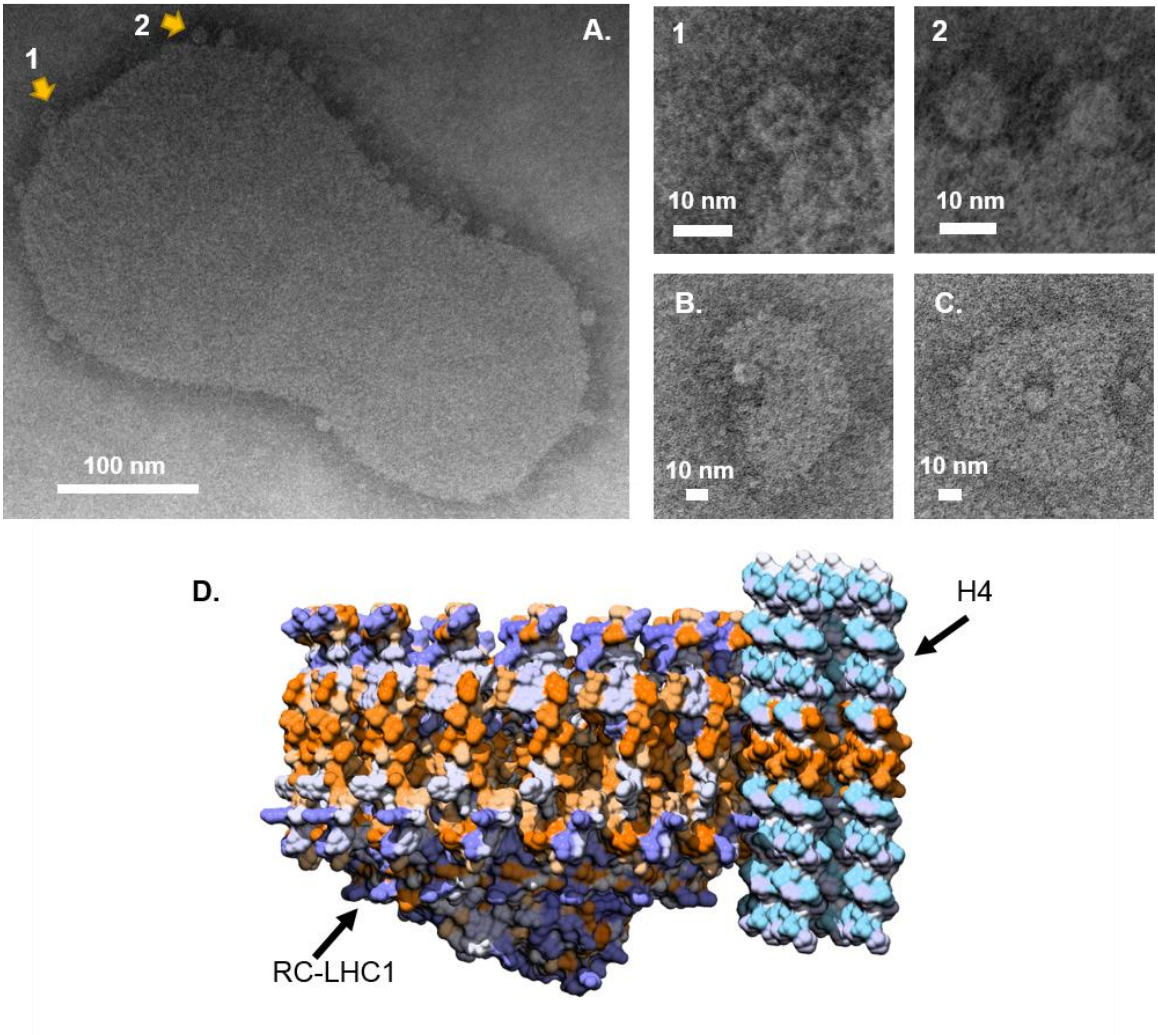

Figure 2.3. H4 and RC-LHC-I interactions. A) Electron micrograph of RC-LHC1 binding predominantly on the surface of an H4 discs. 1) Hexagonal shape of reaction-center 2) RC-LHC1 presumably without RC. B) H4 discs enveloping RC-LHC1 and C) RC-LHC1 presumable at the center of a disc. D) Model of H4 discs binding to the RC-LHC1 membrane protein. Blue to orange represents increasing hydrophobicity.We thank Dr. Klaus Schulten for providing the RC-LHC1 structure in panel D.

H4 discs organized by and staple bacilli flagella together

Serendipitously, several H4 samples were contaminated with rod-shaped uniflagella bacilli. Bacteria flagella filament are composed of the protein flagellin that assemble vertically exposing the helical regions to solvent. The flagellin protein is sought to be used as an adjuvant with vaccines. [103] H4 discs were observed to align perpendicularly along the

entire flagella (> 8 microns), following the curvature (Figure 2.4 A-B). Discs bound perpendicularly to bundle multiple flagella together as ‘staples’ microns in distance (Figure 2.4A, C-D). Additionally, discs are seen binding side-ways to bundled flagella (Figure 2.4 E).

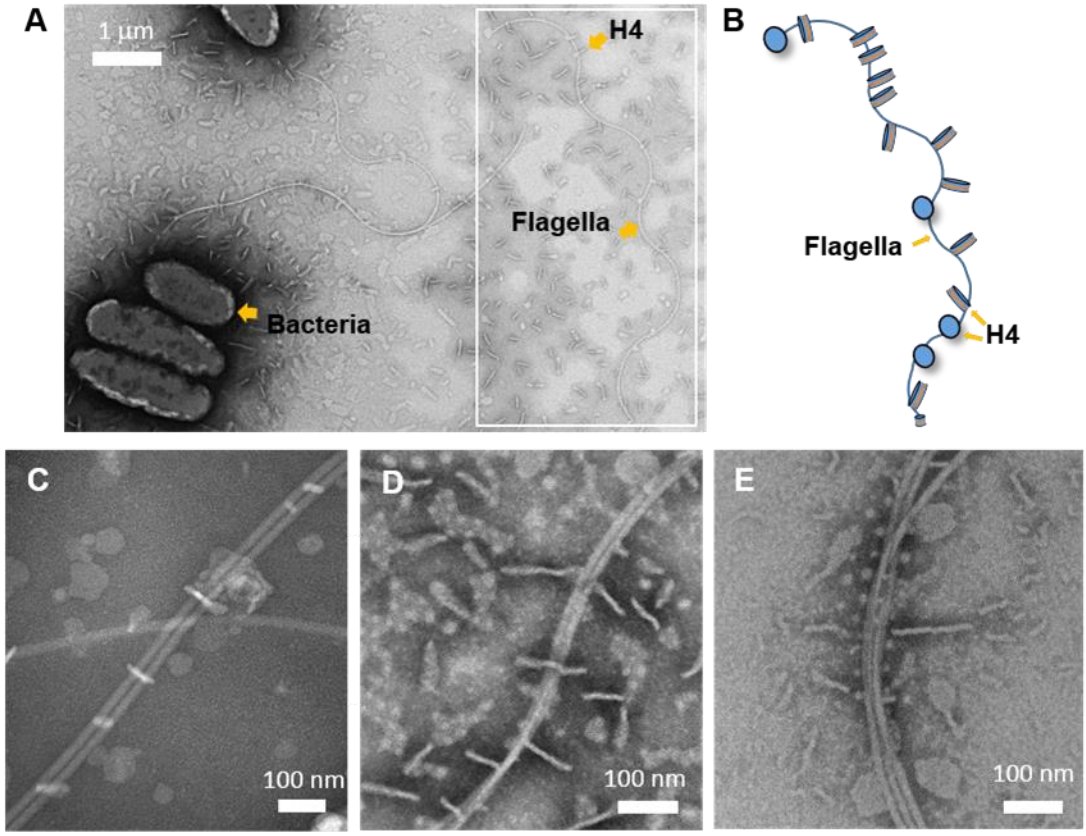

Figure 2.4. H4 discs and bacteria flagella interactions. A) electron micrograph highlighting long range ordering of H4 discs along flagella, B) trace the flagella H4 interactions and white box in A, C-D) H4 discs bundling flagella, E) H4 discs aligned along bundled flagella

H4 promotes tropomyosin fibrillogenesis

Considering we observed that H4 discs bound to the hydrophobic edges of an alpha-helical membrane protein and by chance the alpha-helix rich proteins of bacteria flagella, flagellin, we decided to test the classic alpha-helix coiled-coil Tropomyosin (TM) for binding capabilities. TM is a two-chained alpha-helical coiled-coil protein that is known to

polimerize, via head-to-tail interactions, to form a chain that stabilizes actin filaments within muscle.[104] Tropomyosin presents hydrophobic groups to bind partners (e.g troponin). The conformational flexibility of TM was expected to provide a surface able to scaffold H4 disc interactions.

The addition of H4 to TM resulted in banded fibers. The following interactions suggests that the H4-TM composite fiber formation occurs step-wise (Figure 2.5 A) - 1) varying thickness of fibers with small discs aligned latterally (Figure 2.5 B), 2) chains were observed to append to an existing fiber, (Figure 2.5 C,E,F) 3) fibers were seen to be 'zipping' up (Figure 2.5 C,E), 4) frayed edges of the fiber, (Figure 2.5 C,E) 5) mature fibers have two type of banding patterns of 15 and 50 nm spacing (Figure 2.5 D). Fiber formation is suspected to be a result of polimerized tromopomyosin chains interacting with a combination of free H4 triple-helices and discs. As the fiber maturation occurs, discs are suspected to combine to form larger discs and/or rearrange to allow for banding to occur. Discs also bound perpendicullarly to a bundle, laterally to an existing fiber and fiber was observed to twist (Figure 2.5 F). The observed 50 nm spacing is longer then an actual tropomyosin (~40 nm), suggesting the fiber is paracrystalline, with ordered arrangments of either tropomyosin monomers or chains. The 15 nm spacing between discs is suspected to be the minimum distance discs can stack given water is in between the surfaces. The fibers resemble Segment-long spacing crystallites[105] hypothesized to be the precursor morphology for collagen fibrillogenesis. The banding of the fibers are reminiscent of the stacking of other CMP nanostructures.[106] The stability of tropomyosin makes it an attractive protein for incorporation into biomaterials, similar to metal rods strengthen cement.

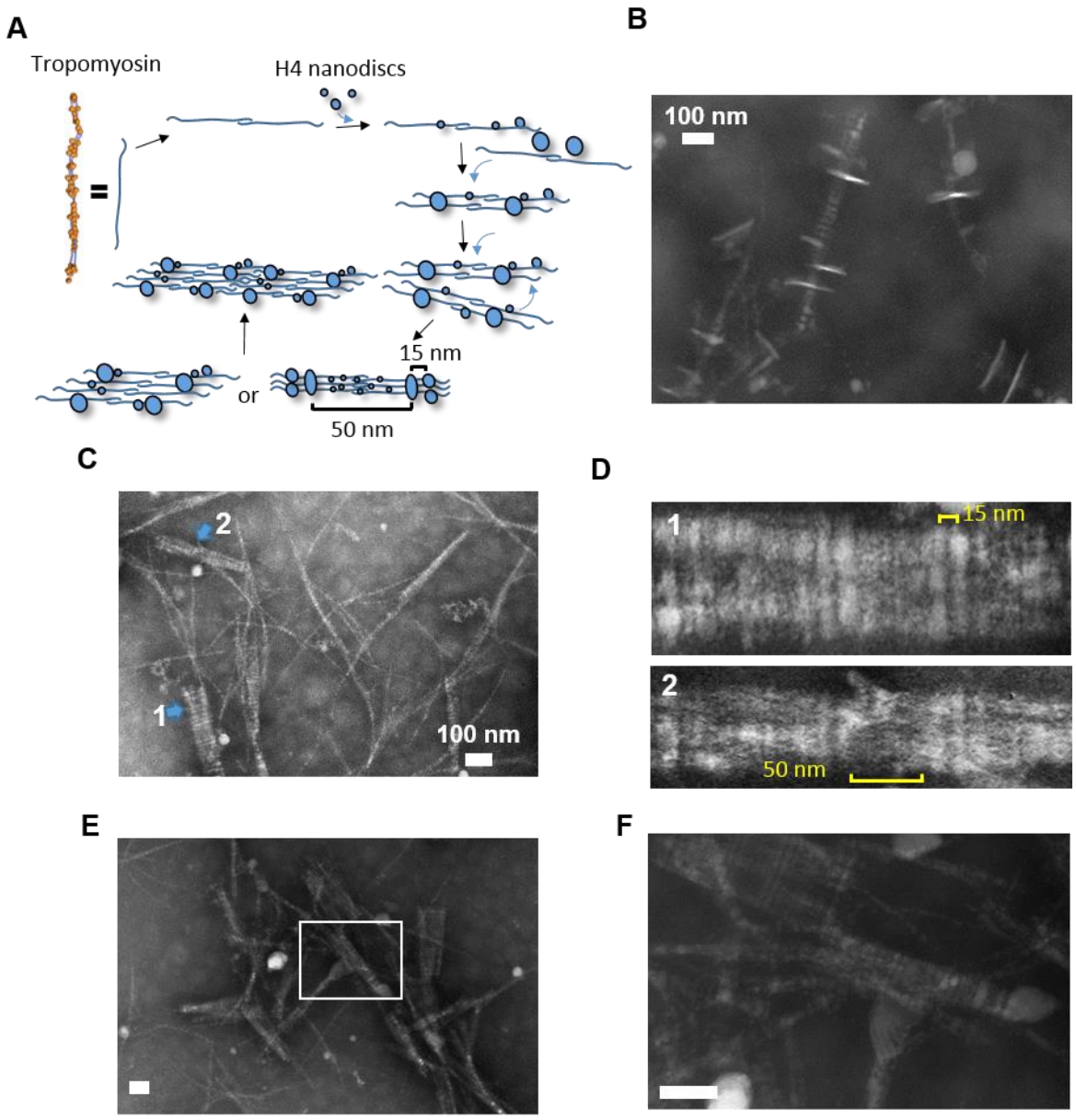

Figure 2.5. H4 discs and tropomyosin interactions. A) Schema of interactions, B) Discs aligning edge-wise to tropomyosin fibers, C) TEM field of H4 discs interacting tropomyosin, D) zoom in of C, highlighting the 15 and 50 nm banding patterns, E) Additional field of H4 discs interacting with TM, F) zoom in of E, highlighting zippering, twisting and addition of tropomyosin/H4 coated strands to an existing fiber. Tropomyosin is ~ 40 nm in length.

2.4 Discussion and Conclusion

Protein interaction promiscuity is an evolutionarily favorable trait that allows a single protein to perform multiple biologically relevant tasks.[107] The H-series discs were shown to bind multiple partners that were structurally different and displayed hydrophobic surfaces to differing extents. Our findings agree with the idea that once a protein has evolved a good binding site it will be utilized in many ways,[108, 109] decreasing its specificity. The chosen natural proteins studied here interact through hydrophobic interactions with cognate binding partners. By keeping the H-Series discs as the common denominator we have ostensibly shown that hydrophobic interactions serve as the major driving force for protein-disc interactions. It has been shown that nearby charged groups strengthen hydrophobic interactions.[110] Each of the chosen proteins have an abundance of charged groups that are surface accessible and therefore may play an important role in the extent of protein-disc interactions. For example, tropomyosin is known to have a highly charged surface and is seen to have the highest order of interactions with H-series discs.

The H-series discs have been shown to interact with the collagen type I triple-helix, the tropomyosin alpha-helix, the alpha-synuclein beta-sheet, protein complex of bacterium flagella, and the alpha-helical membrane protein Reaction Center Light Harvesting Complex I. Fibrous proteins served as scaffolds for disc organization on micron length scales and discs served as scaffolds for the higher order assembly of the fibrous proteins. The discs acted as scaffold for the RC-LHC1 organization.

The H-series is one of many sets of self-assembling nanostructures that can be used to assist multicomponent supramolecular assembly. There exists a rich library of other self-assembling collagen triple-helices that form a variety of nanostructures, e.g. fibers, sheets,

hydrogels, and discs, based on metal coordination,[84-86, 111] chirality,[64] electrostatics,[65, 66, 72, 87, 88] aromatic interactions,[49] hydrophobicity[62] that possibly could be used to drive supramolecular assembly of proteins in a context dependent manner. Combining collagen triple-helices that orthogonally self-assemble into nanostructures has provided new nanostructures, multi-layer sheets[112] with the combined properties of the individual nanostructure components that could also be used to enhance functionality and provide additional modes for guiding supramolecular assemblies.

The H-series discs are promiscuous binders. Improved designs targeting a more specific surface would potentially benefit purification techniques. Further design of either membrane proteins or the H-series discs is predicted to increase specificity and decrease the potential occurrence of hydrophobic mismatch. Recently, it has been shown that the beta barrel protein OMP A can withstand robust hydrophobic mutations and remain functional.[113] This method may potentially be used to redesign membrane proteins to specifically interact with the H-Series discs. Targeting proteins with hydrophobic nanostructures has major potential implications in enhancing hydrophobicity of biomaterials by allowing the storage of hydrophobic drugs compounds in hydrogels. Also, nanostructures can be used to probe the surface features of proteins. The improper presentation of hydrophobic groups is known to be pathogenic and this work further highlights the need for natural protein systems to control the presentation of hydrophobic groups.

2.5 Materials and Methods

Peptide synthesis and purification

Peptides were synthesized using solid phase FMOC chemistry, purified by reverse phase HPLC and verified using mass spectrometry by either LifeTein, LLC (www.lifetein.com) or Tufts University Core Facility. H4, H3 were purified to >95% purity by reverse phase HPLC. Due to issues with synthesis, H2 purity was low, ~ 80%. N- and C- termini were acetylated and amidated, respectively. Peptides were dialyzed in filtered deionized water, lyophilized and kept at -20 °C. Peptides were weighed and resuspended in 10 mM phosphate buffer, sodium phosphate monobasic monohydrate and sodium phosphate dibasic heptahydrate, pH 7.4.

Shrimp Tropomyosin (ST) extraction and purification

Tropomyosin was extracted and purified in multiple steps by Doug Pike under the supervision of Sarah Hitchcock-Degregori from shrimp purchased from Stop and Shop®. All steps were performed at 4 °C except where noted. 400 g of fresh, shelled, and deveined shrimp were minced in a food processor. Care was taken to ensure not to over mince shrimp so that shrimp did not fall through a cheesecloth. Minced shrimp were then washed by stirred for five minutes in a 800 ml dilute salt solution of 20 mM KCL, 1 mM $KHCO_3$, 0.1 mM $CaCL_{12}$, 0.1 M DTT to remove sarcoplasmic proteins. The suspension was then filtered either through a double layer of cheese cloth lightly over a funnel or centrifuged a few minutes at 5,000 rpm. This wash was repeated four times. 800 ml of 95% ethanol was then added to the residue and filtered after 10 minutes. Ethanol extraction was repeated two more times. Residue was then transferred to ice and washed three times in the chemical hood with 800 ml of chilled diethyl ether. Residue was then dried at room temperature for

2 to 3 hours or overnight. Protein powder was then stored at -20 °C. Powder was then dissolved in an 800 ml solution of 1 M KCL, 25 mM Tris (pH 8.0), 0.1 mM $CaCl_2$ and 1 mM DTT, filtered through a cheese cloth and placed in 400 ml of 1 M KCL. The solution was then dialyzed against six liters of 5 mM Tris (pH 7.5) containing 0.1 M $CaCl_2$, and 0.1 DTT for six hours and overnight with fresh dialysis buffer. The solution was retrieved from dialysis, brought up to 133 ml by the addition of 1 M of KCL, and pH 4.6 by dropwise addition of 7.5% acetic acid. The solution was then divided into three equal aliquots to which enzyme grade powdered ammonium sulfate was gradually to two aliquots (leaving the third alone), where one aliquot was taken to 40% and the second to 53% saturation and let to sit for 30 minutes. Solutions were then centrifuged at 30,000 rpm. Ammonium sulfate was then gradually added to all three aliquots - 0 to 40%, 40 to 53%, and 53 to 70%. Each of the three precipitates were then dissolved in 66 ml of 5 mM Tris (pH 7.5) containing 0.1 M $CaCl_2$ and 0.1 M DTT. The tropomyosin precipitate was then removed by centrifugation at 30,000 rpm and a stock solution of 6.8 mg/ml of ST in 10 mM phosphate buffer 50 mM NaCl was stored at -20 °C.

Reaction-Center Light Harvesting Complex 1 extraction and purification

Colonies of *Rhodobacter sphaeroides* M2192 and M21 were grown semiaerobically for 4 days at light intensity of 100 Wm^{-2} into mid-logarithmic phase in CB medium (Potassium Phosphate, Ammonium DL-malate, Concentrated Base[1]) in separate Erlenmeyer flasks filled to 80% capacity on a gyratory shaker at 200 rpm at 30 °C to promote semi-aerobic growth. After reaching an $OD_{680\ nm}$ of ~0.8, cells were pelleted at 11000 g and 4 °C, washed, and resuspended in 1 mM Tris (pH 7.5); a few crystals of DNaseI and protease

[1] Concentrated Base: Nitrilotriacetic acid, $MgSO_4$, $CaCl_2$-$2H_20$, $(NH_4)_6Mo_7O_{24}$-$4H_2O$, $FeSO_4-7H_2O$, Nicotinic acid, Thiamine-HCl, Biotin, Metals "44," Distilled Water to 1000mL

inhibitor cocktail (Roche) were added, and the cells were passed twice through a French press. Debris and unbroken cells were removed by centrifugation at 12000 g and the supernatant was layered onto a 5 to 35% (w/w) sucrose gradient prepared in 1 mM Tris (pH 7.5) over a 60% sucrose cushion and subjected to rate-zone ultracentrifugation for 3 h at 28000 rpm in Beckman SW-28 rotor. The broad red pigmented band was collected and stored at 4 °C.

Removal of sucrose was modified from 4,000 rpm in a swinging bucket rotor to a 2-hour 55,000 rpm centrifugation on a Beckman 70 Ti rotor. The concentrated pellet was kept and stored while the supernatant was discarded (containing sucrose). The pellet contained the broad red-pigmented band, which is composed of membrane fragments from the broken chromatophores that were collected in the previous (28k) centrifugation.

RC-LH1 complexes were extracted from intracytoplasmic membranes with 15 mM *n*-octyl-β-D-glucopyranoside (β-OG) and 15 mM deoxycholate in 10 mM Tris-HC1, pH 7.4, plus 5 mM EDTA for 30 min at 0 °C. The mixture was vortexed every 5 min during solubilization. Non-solubilized material was removed by centrifugation (2 min in an Eppendorf centrifuge); 0.3 mL of the supernatant was layered on a 9-mL sucrose gradient with 10-40% sucrose (w/v) in 20 mM Tris-HC1, pH 7.4, and 15 mM (β-OG plus 15 mM deoxycholate. Gradients were centrifuged in a Beckman SW 41 Ti rotor (4 h, 35 000 rpm, 4°C). Pigmented band were recovered from the gradient, analyzed spectrophotometrically, and stored at -80 °C.

Extracted complexes from the pigmented band exist as detergent-protein micelles of the isolated RC-LH1 and LH1 cores surrounded by β-OG detergent (*n*-octyl-β-D-glucopyranoside). Prior to incubation with H4 discs, 80 uL of RC-LH1 and LH1 solutions

were incubated with 10 uL RNase A (10 ug/uL) for 1hr at 60 °C and at 4 °C for two months to remove residual ribosomal content.

Transmission Electron Microscopy

H-series collagen peptide samples were prepared in 10mM phosphate buffer with s pH 7.3 and incubated at 4°C for 2 weeks prior to incubation with proteins. Collagen Type I (COLI) (Lot # 879114) solutions were prepared in 0.02 N acetic acid at 3mg/ml pH 3.3. H4 acetic acid studies were performed by directly solvating H4 powder into 0.02 N acetic acid pH 3.3. Alpha-synuclein solutions in PBS pH 7.0 were prepared at 1 mg/ml and amyloid fibrils were formed via shaking. H-series and COLI were solutions were combined at 1:1 volume ratios. H4-RC-LHC1 were prepared by either solvating the H4 powder directly into a stock solution of RC-LHC1 or H4 and RC-LHC1 solutions were combined at 1:1 volume ratio. H4 and ST solutions were combined 2:1 volume ratios and incubated at 4 °C for four days prior to imaging. For imaging, a volume of 5 μL of peptide solution was deposited on a copper grid coated with carbon (Electron Microscopy Sciences CF-400-Cu), then a time of two minutes was passed to ensure deposition, and excess peptide was wicked away using filter paper. The grid was then negatively stained with 1% phosphotungstic acid (PTA) pH 7.0 for two minutes and excess stain was wicked away. Afterward, the grids were left to dry for at least two minutes in ambient conditions before imaging. Images were captured using a Philip 420 Electron Microscope at 80 kV.

2.6 Acknowledgments

This chapter is in preparation for manuscript submission. We thank the National Institutes of Health grant DP2 OD-006478 and a GAANN fellowship for supporting this research. This work was done in close collaboration with Doug Pike, Kathyrn Drzewiecki, Nora

Tofigh, Mike Kopka, Arpit Patel, Maria Janowska, Dr. Jean Baum, Dr. Robert Niederman, and Dr. Vikas Nanda.

Appendix 1

Characterizing the packing of H4 discs

A.1 Introduction

This appendix takes a closer look at the nanostructures formed by the peptide H4 in Chapter 2. H4 is an ABA triblock copolymer where domain A is polar and B is nonpolar. H4 folds into a triple-helix, which further self-assembles into discs plus a small population of fibers. TEM, AFM, and fluorescence imaging suggest that the discs are highly flexible, variable in height and are capable of binding hydrophobic molecules, respectively. Disk flexibility and height variation are hypothesized to be due to registry shifts between triple-helices, which were observed in DLA simulations of H4 disk formation. Multiple packing registries have been seen in crystal structures of collagen peptides with hydrophobic domains.[114] We aim to probe the registry shift hypothesis by investigating the packing of the nonpolar domains within the H4 discs using small-angle neutron scattering (SANS) with contrast variation using D_2O and selective deuteration of the nonpolar domain. The large diameter of the discs (> 100 nm) allows for the neutron scattering profile to be directly related to the thickness and density of the deuterated components. From the scattering profile we can infer how well packed (ordered) the hydrophobic domains are and also the extent of solvent penetration. We expect our scattering densities of the deuterated components to be broad due to the disperse packing of the nonpolar domains, leading to greater intercalation of solvent. Conversely, we would expect a narrow distribution if the nonpolar domains are orderly packed (Figure A.1).

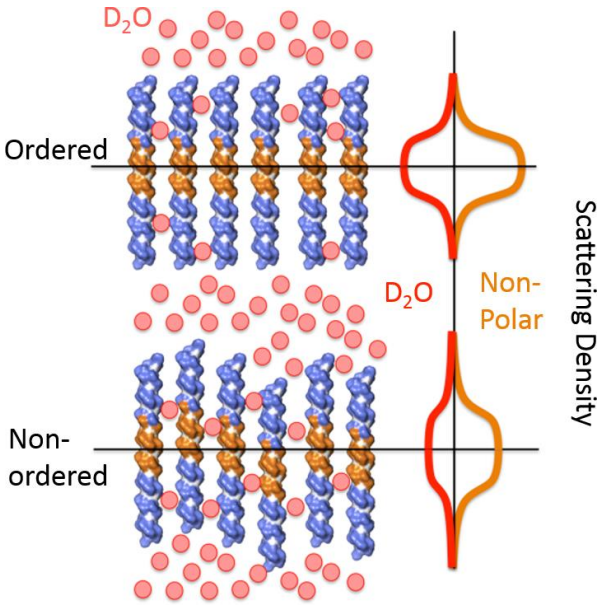

Figure A.1 Schematic of the expected results of scattering experiments. Top and bottom show the scattering profiles (right) for ordered versus non-ordered packing of nonpolar domains (left), respectively. Scatter profiles show D_2O penetration (Red) and Packing of deuterated nonpolar region (Orange).

A.2 Materials and Methods

Small Angle Neutron Scattering

Initial SANS experiments were performed on the EQ-SANS instrument, at Oak Ridge National Laboratory in Oak Ridge, TN. A 1% solution (w/v) of H4 (~3.65 mM) was prepared in 10 mM phosphate buffer two weeks prior to lyophilization, resuspension in 100% D_2O, making a 80% D_2O solution, and placement in a 2 mM banjo cuvette for measurement. It was assumed the lyophilization and resuspension process did not disrupt the underlying structures. A solution of 10 mM phosphate buffer was also lyophilized and resuspended in 100% D_2O and used as a buffer subtraction. The experiments were performed at 10°C, measuring a *Q*-range from 0.004 to 0.4 $Å^{-1}$, using a sample-to-detector distance of 4 m and a starting wavelength bands of 2.5 and 10 Å, for 1 hr per measurement.

Spectra were set to absolute intensity using a porasil standard with a 0.0245 scale factor for the 2.5 Å wavelength band, and corrected using the transmission of an empty beam, and the transmission of an empty banjo cell. A Porod fit was performed to remove the solvent background by subtracting the slope of the $I(q)*q^4$ vs q^4 plot from the intensity data. The radius of gyration (R_g), the overall size, was calculated by fitting the linear region of the Guinier plot (ln(I)vs q^2) using the Guinier approximation [115]

$$I(q) = I(0)e^{-\frac{q^2R_g^2}{3}}$$

$$R_g = \sqrt{-3A}, \quad where\ q_{max}Rg \lesssim 1.3$$

where I(q) is the measured scattering intensity, q is related to the wavelength λ and scattering angle 2θ by $q = 4\pi \sin\theta/\lambda$, and A is the slope of the linear fit. It is suggested that for a good fit, the range of q should be approximately a decade in length. For particles that are sheet-like, i.e. with one dimension longer than other two, the average thickness T can be measured by fitting the linear region of a modified Guinier (plot $\ln(q^2*I)$ vs q^2) approximation [116]

$$I(q) = I_t e^{-\frac{q^2T^2}{12}}$$

$$T = \sqrt{-12A}, \quad where\ q_{max}T \lesssim \sqrt{12}$$

where A is the slope of the linear fit, and $I_t(0)$ is the extrapolated intensity at q = 0. For rod-like particles the R_c of gyration was determined by fitting the linear region of a modified guinier for rod-like forms (plot ln(q*I) vs q^2) [116]

$$I(q) = I_c e^{-\frac{q^2R_C^2}{2}}$$

$$R_C = \sqrt{-2A}, \quad where\ q_{max}R_C \approx 1$$

$$R = \sqrt{2 * R_C}$$

where R is the radius of the rod, A is the slope of the linear region, and $I_C(0)$ is the extrapolated intensity at q = 0. A series of programs were used to analyze the results including Primus[117] and Microsoft Excel.[118]

A.3 Results

SANS experiments on H4 discs, in 80% D2O indicated a Porod exponent of ~2.7, which may indicate the presence of flexible lamellar objects in solution (Figure A.2 A). Although the Guinier plot contained a linear region suggesting a radius of gyration (Rg) of 22 nm, the Rg*qmax was 2.9, which is above the acceptable limit of 1.3. Consistent with AFM and EM studies, modified Guinier fits revealed linear regions for rods (Figure A.2 B) and discs (Figure A.2 C), supporting the existence of a mixture of disc and fiber species in solution. The R^2 values were 0.84 and 0.98 for the disc and rod fit, respectively. The radius for the rod species was calculated to be 21.8 nm and the average thickness for the discs was calculated to be 25 nm, which is longer than the expected 8.6 nm (30 residues x 0.286 nm rise/ residue) of a CMP however, within the range of heights observed by AFM.

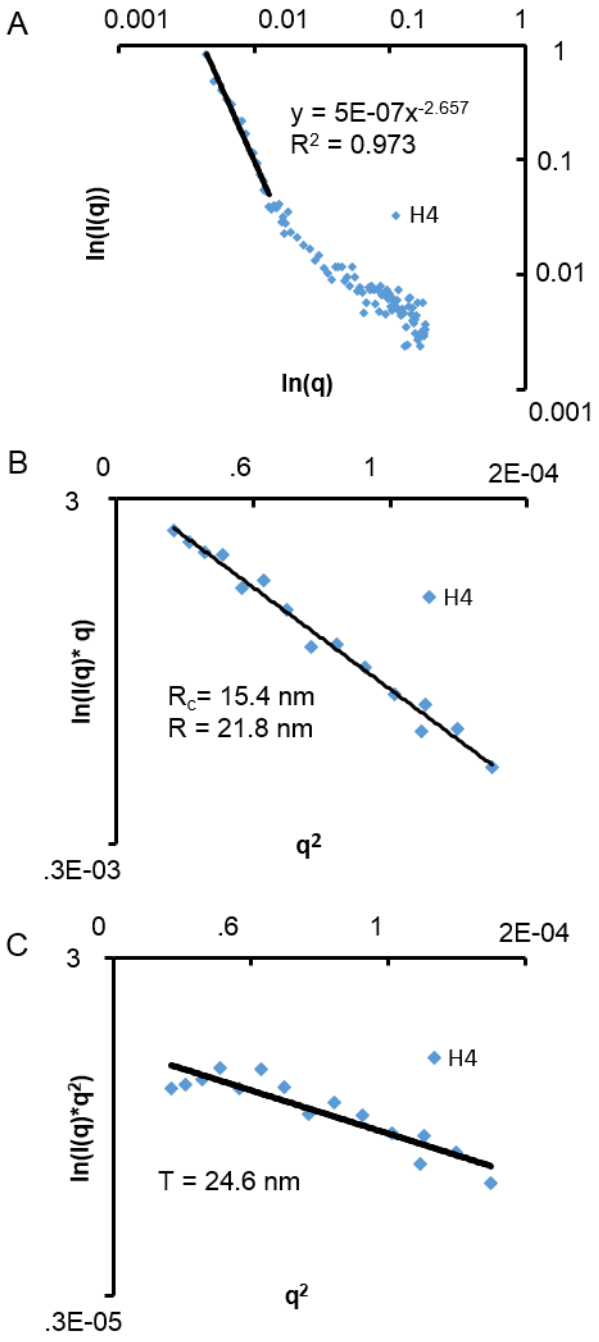

Figure A.2 Small Angle Neutron Scattering of H4 in 80% D_2O. A. Porod Plot, B. modified guinier plot – rod mode, C. modified guinier plot – disc model. q is in $Å^{-1}$

A.4 Discussion and Future Work

Preliminary SANS measurements on H4 solutions in 80% D_2O are promising for understanding the fluid properties of the H4 discs. Both types of fit give fairly straight lines, however a rod model was favored with the fitting. Due to the sample preparation at the SANS facility, we believe that the samples were not examined under equilibrium conditions, i.e. discs were not fully formed. Therefore, additional SANS experiments are required in order to be cost effective due to the high price of deuterated amino acids. Prior to performing additional SANS experiments, a series of Small-Angle X-Ray experiments are expected to help understand the conditions under which mature discs are formed and enhance the success of future SANS experiments. Due to the heterogeneity of the H4 solutions, not only in size, but potentially in morphology more experiments such as SAXS are needed to be performed so that we correctly can assign thickness and length values. Preliminary SAXS, experiments look promising (not shown).

A.5 Acknowledgments

We thank Oak Ridge National Laboratory for the use of instrument time. We thank Dr. Christopher Stanley and Dr. Ken Littrell for guidance and invaluable feedback. This work was on close in collaboration with Dr. N. Sanjeeva Murthy, and Dr. Vikas Nanda.

B. Appendix: Self-Assembly of Left and Right-Handed Molecular Screws

B.1 Introduction

This appendix investigates the rules of stereoselectivity relating to the self-assembly of homochiral and heterochiral peptides. Although natural proteins tend to prefer self or homochiral molecular recognition, this is not an absolute rule. Synthetic polypeptides exhibit no consistency in stereoselectively.[64] Despite this, general rules that relate shape complementarity to association would be useful in guiding molecular design. One such rule describes interactions between like-versus opposite-handed helical objects. A geometric analysis of the packing of coiled coils predicted that columnar associations between opposite-handed supercoils would allow for an overall tighter packing density and a greater number of intermolecular contacts than like-handed associations.[119] Optimal packing of like-handed threaded rods requires rotation of principle axes of adjacent rods, preventing tight columnar packing;[120] this same phenomenon determines helix packing in proteins.[121] Molecular simulations of opposite- and like-handed poly alanine α-helices demonstrate a preference for left-right helical dimers.[122] All of these studies support a general rule that supramolecular interactions of opposite-handed helices will be favored over like-handed assemblies. However, it is challenging to develop an appropriate experimental system that evaluates shape complementarity without being strongly influenced by the details of intermolecular interactions. Ridges-in-grooves interactions have been demonstrated on a macroscopic scale between left- and right-handed bolts.[123] At the molecular scale, a ‘chemically nude’ system is needed where the shape is a primary factor promoting close pack

The collagen mimetic peptide $(PPG)_{10}$ (P = proline, G = glycine) is a promising minimal system for evaluating the role of helix handedness on intermolecular association. Collagen is composed of three chains that are supercoiled to form a triple-helix. Except for glycine, all naturally occurring amino acids are levorotatory (L-Proline) stereoisomers (Figure B.1 A). Using dextrorotatory (D) stereoisomers (D-Proline), (Figure B.1 A) results in a mirror-image, opposite-handed triple-helix (Figure B.1 A). In the collagen mimetic peptide, $(PPG)_{10}$, proline side chains form the ridges and grooves of the triple-helix. Due to the cyclic aliphatic side chain, one might expect reduced contributions from side chain flexibility, charge-pair interactions, or hydrogen bonding that could influence molecular packing specificity. Unlike the H-Series peptides in the previous chapters, PPG_{10} is a "chemically nude" system where the complementary shape of the binding interface is the primary determinant for effective packing. Molecular simulations of collagen peptides predicted that optimal packing would occur between enantiomers similar to the packing of opposite handed screws (Figure B.1 B).[64] This prediction is shown to be true. Mixing left- and right-handed helices drastically lowered solubility, resulting in micrometer-scale sheet-like assemblies that were one peptide-length thick as characterized with atomic force microscopy.

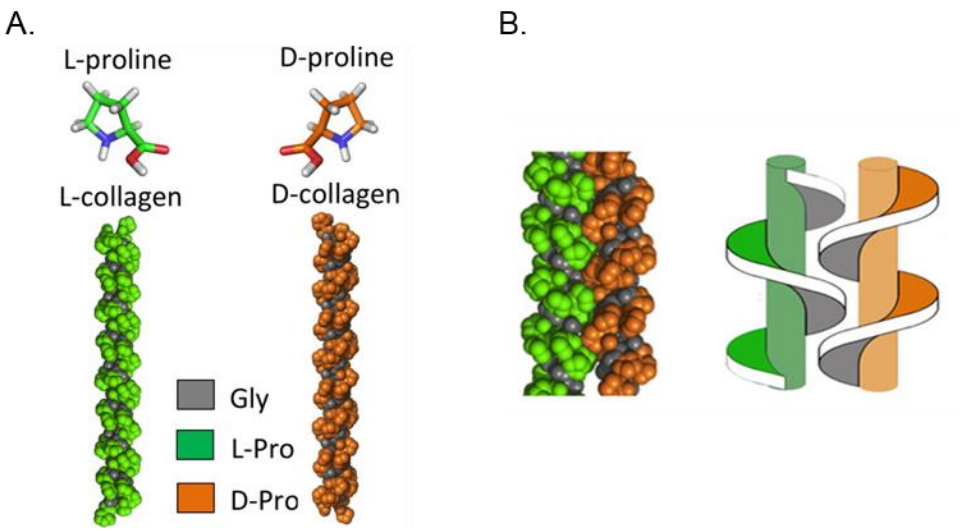

Figure B.1 Mirror-image triple helices. Structural models of $[(_{L}P_{L}PG)_{10}]_3$ and $[(_{D}P_{D}PG)_{10}]_3$ triple helices. The $[(_{L}P_{L}PG)_{10}]_3$ model was obtained from a high-resolution (1.30 Å) X-ray crystal structure (PDB: 1K6F) [124]. B. Predicted packing of L-and D- collagen peptides along with the packing of opposite handed screws.

B.2 Materials and methods

Electron microscopy (EM)

Aggregated samples were gently re-suspended and placed onto copper grids, then negatively stained with 1% phosphotungstic acid and imaged with a Philip 420 Electron Microscope.

Atomic Force Microscopy (AFM)

Silicon chips were cleaned with several solutions prior to peptide deposition: The chips were first sonicated in 100% acetone for 5 min, then soaked in 100% isopropanol for 5 minutes, followed by a final wash in 100% ethanol. The silicon chips were dried with nitrogen, then placed onto a glass slide with adhesive and chilled on ice for ten minutes prior to the deposition of 3 µl of the sample solution. Fifteen minutes after depositing the peptide, the wafer was dipped in deionized water then placed back onto the glass slide for

at least 20 minutes of drying at room temperature. Multiple scans for each measurement were taken using an Agilent Atomic Force Microscope model 5500 in tapping mode, with a Multi75a probe (Bruker MPP-21120-10) and analyzed using the Gwyddion software package version 2.32.[92]

B.3 Results

The objective of this appendix was to confirm the presence of well-ordered micrometer-scale sheets formed by mixtures of L-and D- collagen peptides that were observed in TEM (Figure B.2 A). AFM measurements confirmed the presence of sheets as well showed that these sheets had an average single layer of ~10 nm (Figure B.2 B) which is consistent with the length of a 30-residue triple helix. These heterochiral triple helices may align to form the sheet like-structure (Figure B.2 C), although higher-resolution data would be needed to confirm this model.

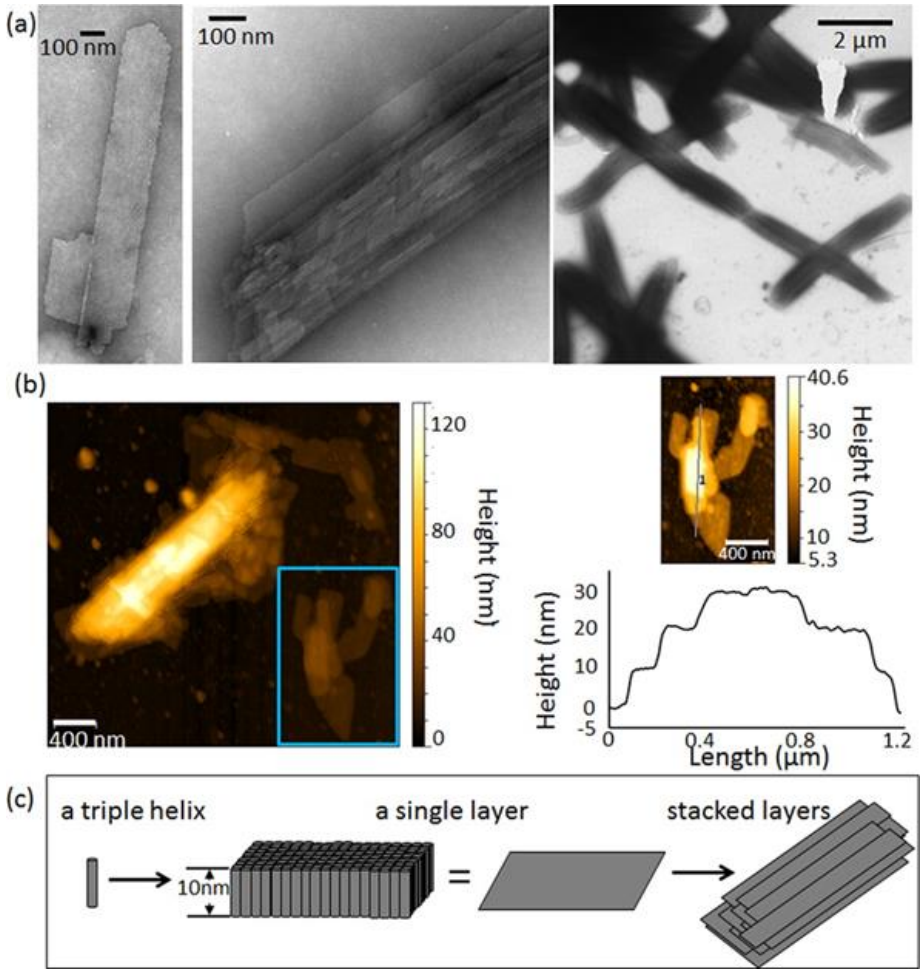

Figure B.2 Sheet morphology. (a) TEM and (b) AFM of a 1:1 l to d mixture. Contour heights varied by integer multiples of 10 nm. (c) A proposed model of assembly where triple helices align into a 10 nm thick layer, which further stack. Samples were incubated at 4 °C for ~4 weeks prior to imaging.

B.4 Discussion

In conclusion, the presence of self-assembled $(PPG)_{10}$ sheets was confirmed using AFM. Overall, using $(PPG)_{10}$ sequence enantiomers as minimal helical exemplars, the geometric prediction of preferred left-/right-handed helical pairings appears to hold true. However, even this system is complicated by chemical detail. In a series of experiments pairing $(_DP_DPG)_{10}$ with other right-handed triple-helical peptides such as $(POG)_{10}$, where

O is hydroxyproline, no interactions were observed. Hydroxylation of the proline side chain may sterically prevent close association of left- and right-handed species, increase the desolvation barrier required to allow tight ridges-in-grooves packing, or even modify the degree of supercoiling of the triple-helix such that different pitch of left- and right-handed species could preclude a coherent, columnar interaction. Ironically, this general rule of helix–helix association may best apply in special cases where complete mirror symmetry exists.

B.5 Acknowledgments

This work was published in the Journal of the American Chemical Society. This work was supported by NIH DP2 OD-006478 and NIH R01 GM-089949. We thank Hiroshi Matsui for access to AFM. This work was done in close collaboration with Fei Xu, I. John Khan, Avanish S. Parmar, Teresita Silva, N. Sanjeeva Murthy, and Vikas Nanda.

References

1. Woolfson, D.N. & Mahmoud, Z.N. More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials. *Chem Soc Rev* **39**, 3464-3479 (2010).
2. Wen, Y. & Collier, J.H. Supramolecular peptide vaccines: tuning adaptive immunity. *Curr Opin Immunol* **35**, 73-79 (2015).
3. Ravi, J., Bella, A., Correia, A.J., Lamarre, B. & Ryadnov, M.G. Supramolecular amphipathicity for probing antimicrobial propensity of host defence peptides. *Phys Chem Chem Phys* **17**, 15608-15614 (2015).
4. Mehrban, N. et al. Assessing cellular response to functionalized alpha-helical peptide hydrogels. *Adv Healthc Mater* **3**, 1387-1391 (2014).
5. Webber, M.J., Berns, E.J. & Stupp, S.I. Supramolecular Nanofibers of Peptide Amphiphiles for Medicine. *Isrl J Chem* **53**, 530-554 (2013).
6. Adler-Abramovich, L. & Gazit, E. The physical properties of supramolecular peptide assemblies: from building block association to technological applications. *Chem Soc Rev* **43**, 6881-6893 (2014).
7. Ingolfsson, H.I. et al. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip Rev Comput Mol Sci* **4**, 225-248 (2014).

8. Frederix, P.W., Ulijn, R.V., Hunt, N.T. & Tuttle, T. Virtual Screening for Dipeptide Aggregation: Toward Predictive Tools for Peptide Self-Assembly. *J Phys Chem Lett* **2**, 2380-2384 (2011).
9. Frederix, P.W. et al. Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat Chem* **7**, 30-37 (2015).
10. Chothia, C. & Janin, J. Principles of protein-protein recognition. *Nature* **256**, 705-708 (1975).
11. Janin, J., Miller, S. & Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. *J Mol Biol* **204**, 155-164 (1988).
12. Miller, S., Janin, J., Lesk, A.M. & Chothia, C. Interior and surface of monomeric proteins. *J Mol Biol* **196**, 641-656 (1987).
13. Jones, S. & Thornton, J.M. Prediction of protein-protein interaction sites using patch analysis. *J Mol Biol* **272**, 133-143 (1997).
14. Miller, S., Lesk, A.M., Janin, J. & Chothia, C. The accessible surface area and stability of oligomeric proteins. *Nature* **328**, 834-836 (1987).
15. Bowie, J.U., Luthy, R. & Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164-170 (1991).
16. Murayama, M. Molecular mechanism of red cell "sickling". *Science* **153**, 145-149 (1966).
17. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C.M. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424**, 805-808 (2003).
18. Pawar, A.P. et al. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J Mol Biol* **350**, 379-392 (2005).
19. Parmar, A.S., Nunes, A.M., Baum, J. & Brodsky, B. A peptide study of the relationship between the collagen triple-helix and amyloid. *Biopolymers* **97**, 795-806 (2012).
20. Wang, W. & Hecht, M.H. Rationally designed mutations convert de novo amyloid-like fibrils into monomeric beta-sheet proteins. *Proc Natl Acad Sci U S A* **99**, 2760-2765 (2002).
21. Hilbich, C., Kisterswoike, B., Reed, J., Masters, C.L. & Beyreuther, K. Substitutions of Hydrophobic Amino-Acids Reduce the Amyloidogenicity of Alzheimers-Disease Beta-A4 Peptides. *J Mol Biol* **228**, 460-473 (1992).
22. Kim, W. & Hecht, M.H. Sequence determinants of enhanced amyloidogenicity of Alzheimer A beta 42 peptide relative to A beta 40. *Journal of Biological Chemistry* **280**, 35069-35076 (2005).
23. Dahiyat, B.I. & Mayo, S.L. De novo protein design: fully automated sequence selection. *Science* **278**, 82-87 (1997).
24. Shifman, J.M. & Mayo, S.L. Modulating calmodulin binding specificity through computational protein design. *J Mol Biol* **323**, 417-423 (2002).
25. Degrado, W.F. & Lear, J.D. Induction of Peptide Conformation at Apolar Water Interfaces .1. A Study with Model Peptides of Defined Hydrophobic Periodicity. *J Am Chem Soc* **107**, 7684-7689 (1985).
26. Ventura, S. et al. Conformational strain in the hydrophobic core and its implications for protein folding and design. *Nat Struct Biol* **9**, 485-493 (2002).

27. Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. & Hecht, M.H. Protein design by binary patterning of polar and nonpolar amino acids. *Science* **262**, 1680-1685 (1993).
28. Jacak, R., Leaver-Fay, A. & Kuhlman, B. Computational protein design with explicit consideration of surface hydrophobic patches. *Proteins-Structure Function and Bioinformatics* **80**, 825-838 (2012).
29. Trevino, S.R., Scholtz, J.M. & Pace, C.N. Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase Sa. *J Mol Biol* **366**, 449-460 (2007).
30. Holowka, E.P., Pochan, D.J. & Deming, T.J. Charged polypeptide vesicles with controllable diameter. *J Am Chem Soc* **127**, 12423-12428 (2005).
31. Nowak, A.P. et al. Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles. *Nature* **417**, 424-428 (2002).
32. Cui, H.G., Webber, M.J. & Stupp, S.I. Self-Assembly of Peptide Amphiphiles: From Molecules to Nanostructures to Biomaterials. *Biopolymers* **94**, 1-18 (2010).
33. Xu, H. et al. Hydrophobic-region-induced transitions in self-assembled peptide nanostructures. *Langmuir* **25**, 4115-4123 (2009).
34. Yu, Y.C. et al. Construction of biologically active protein molecular architecture using self-assembling peptide-amphiphiles. *Method Enzymol* **289**, 571-587 (1997).
35. Gore, T., Dori, Y., Talmon, Y., Tirrell, M. & Bianco-Peled, H. Self-assembly of model collagen peptide amphiphiles. *Langmuir* **17**, 5352-5360 (2001).
36. Vauthey, S., Santoso, S., Gong, H.Y., Watson, N. & Zhang, S.G. Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles. *Proc Natl Acad Sci U S A* **99**, 5355-5360 (2002).
37. Bakota, E.L., Sensoy, O., Ozgur, B., Sayar, M. & Hartgerink, J.D. Self-assembling multidomain peptide fibers with aromatic cores. *Biomacromolecules* **14**, 1370-1378 (2013).
38. Dong, H., Paramonov, S.E., Aulisa, L., Bakota, E.L. & Hartgerink, J.D. Self-assembly of multidomain peptides: balancing molecular frustration controls conformation and nanostructure. *J Am Chem Soc* **129**, 12468-12472 (2007).
39. Lakshmanan, A. & Hauser, C.A.E. Ultrasmall Peptides Self-Assemble into Diverse Nanostructures: Morphological Evaluation and Potential Implications. *Int J Mol Sci* **12**, 5736-5746 (2011).
40. Nam, K.T. et al. Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nature materials* **9**, 454-460 (2010).
41. Sanii, B. et al. Shaken, not stirred: collapsing a peptoid monolayer to produce free-floating, stable nanosheets. *J Am Chem Soc* **133**, 20808-20815 (2011).
42. Yue, K. & Dill, K.A. Forces of tertiary structural organization in globular proteins. *Proc Natl Acad Sci U S A* **92**, 146-150 (1995).
43. Yue, K. & Dill, K.A. Inverse protein folding problem: designing polymer sequences. *Proc Natl Acad Sci U S A* **89**, 4163-4167 (1992).
44. Agashe, V.R., Shastry, M.C. & Udgaonkar, J.B. Initial hydrophobic collapse in the folding of barstar. *Nature* **377**, 754-757 (1995).
45. Sosnick, T.R., Mayne, L. & Englander, S.W. Molecular collapse: the rate-limiting step in two-state cytochrome c folding. *Proteins* **24**, 413-426 (1996).

46. Bella, J., Eaton, M., Brodsky, B. & Berman, H.M. Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science* **266**, 75-81 (1994).
47. Hulmes, D.J., Miller, A., Parry, D.A. & Woodhead-Galloway, J. Fundamental periodicities in the amino acid sequence of the collagen alpha1 chain. *Biochem Biophys Res Commun* **77**, 574-580 (1977).
48. Hulmes, D.J., Miller, A., Parry, D.A., Piez, K.A. & Woodhead-Galloway, J. Analysis of the primary structure of collagen for the origins of molecular packing. *J Mol Biol* **79**, 137-148 (1973).
49. Kar, K. et al. Aromatic interactions promote self-association of collagen triple-helical peptides to higher-order structures. *Biochemistry* **48**, 7959-7968 (2009).
50. Witten, T.A. & Sander, L.M. Diffusion-limited aggregation. *Phys. Rev. B* **27**, 5686-5697 (1983).
51. Witten, T.A. & Sander, L.M. Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.* **47**, 1400-1403 (1981).
52. Mu, Y. & Yu, M. Effects of hydrophobic interaction strength on the self-assembled structures of model peptides. *Soft matter* **10**, 4956-4965 (2014).
53. Narayanan, B., Gilmer, G.H., Tao, J., De Yoreo, J.J. & Ciobanu, C.C. Self-Assembly of Collagen on Flat Surfaces: The Interplay of Collagen−Collagen and Collagen−Substrate Interactions. *Langmuir* **30**, 1343-1350 (2014).
54. Lovelady, H.H., Shashidhara, S. & Matthews, W.G. Solvent specific persistence length of molecular type I collagen. *Biopolymers* **101**, 329-335 (2014).
55. Lomander, A., Hwang, W. & Zhang, S. Hierarchical self-assembly of a coiled-coil peptide into fractal structure. *Nano Lett* **5**, 1255-1260 (2005).
56. Han, T.H., Oh, J.K., Lee, G.J., Pyun, S.I. & Kim, S.O. Hierarchical assembly of diphenylalanine into dendritic nanoarchitectures. *Colloids Surf B Biointerfaces* **79**, 440-445 (2010).
57. Song, H. & Parkinson, J. Modelling the self-assembly of elastomeric proteins provides insights into the evolution of their domain architectures. *PLoS Comput Biol* **8**, e1002406 (2012).
58. Parkinson, J., Kadler, K.E. & Brass, A. Simple physical model of collagen fibrillogenesis based on diffusion limited aggregation. *J Mol Biol* **247**, 823-831 (1995).
59. Prockop, D.J. & Fertala, A. The collagen fibril: the almost crystalline structure. *J Struct Biol* **122**, 111-118 (1998).
60. Kadler, K.E., Hojima, Y. & Prockop, D.J. Collagen fibrils in vitro grow from pointed tips in the C- to N-terminal direction. *Biochem J* **268**, 339-343 (1990).
61. Persikov, A.V., Ramshaw, J.A. & Brodsky, B. Prediction of collagen stability from amino acid sequence. *J Biol Chem* **280**, 19343-19349 (2005).
62. Cejas, M.A. et al. Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions. *Proc Natl Acad Sci U S A* **105**, 8513-8518 (2008).
63. Kar, K., Wang, Y.H. & Brodsky, B. Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures. *Protein Sci* **17**, 1086-1095 (2008).
64. Xu, F. et al. Self-assembly of left- and right-handed molecular screws. *J Am Chem Soc* **135**, 18762-18765 (2013).

65. Jiang, T. et al. Structurally Defined Nanoscale Sheets from Self-Assembly of Collagen-Mimetic Peptides. *J Am Chem Soc* **136**, 4300-4308 (2014).
66. Jiang, T., Xu, C., Zuo, X. & Conticello, V.P. Structurally Homogeneous Nanosheets from Self-Assembly of a Collagen-Mimetic Peptide. *Angew Chem Int Ed Engl* (2014).
67. Brodsky, B. & Eikenberry, E.F. Characterization of fibrous forms of collagen. *Methods Enzymol* **82 Pt A**, 127-174 (1982).
68. Holmes, D.F., Chapman, J.A., Prockop, D.J. & Kadler, K.E. Growing tips of type I collagen fibrils formed in vitro are near-paraboloidal in shape, implying a reciprocal relationship between accretion and diameter. *Proc. Natl. Acad. Sci. U S A* **89**, 9855-9859 (1992).
69. Garbuzynskiy, S.O., Lobanov, M.Y. & Galzitskaya, O.V. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics* **26**, 326-332 (2010).
70. Hwang, E.S., Thiagarajan, G., Parmar, A.S. & Brodsky, B. Interruptions in the collagen repeating tripeptide pattern can promote supramolecular association. *Protein Sci* **19**, 1053-1064 (2010).
71. Klunk, W.E., Jacob, R.F. & Mason, R.P. Quantifying amyloid by congo red spectral shift assay. *Methods Enzymol* **309**, 285-305 (1999).
72. Xu, F. et al. Compositional control of higher order assembly using synthetic collagen peptides. *J Am Chem Soc* **134**, 47-50 (2012).
73. Engel, J., Chen, H.T., Prockop, D.J. & Klump, H. The triple helix in equilibrium with coil conversion of collagen-like polytripeptides in aqueous and nonaqueous solvents. Comparison of the thermodynamic parameters and the binding of water to (L-Pro-L-Pro-Gly)n and (L-Pro-L-Hyp-Gly)n. *Biopolymers* **16**, 601-622 (1977).
74. Aggeli, A. et al. Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta -sheet tapes, ribbons, fibrils, and fibers. *Proc Natl Acad Sci U S A* **98**, 11857-11862 (2001).
75. Theato, P., Sumerlin, B.S., O'Reilly, R.K. & Epps, T.H., 3rd Stimuli responsive materials. *Chem Soc Rev* **42**, 7055-7056 (2013).
76. Matson, J.B., Newcomb, C.J., Bitton, R. & Stupp, S.I. Nanostructure-templated control of drug release from peptide amphiphile nanofiber gels. *Soft matter* **8**, 3586-3595 (2012).
77. Burkoth, T.S. et al. Toward the synthesis of artificial proteins: the discovery of an amphiphilic helical peptoid assembly. *Chemistry & biology* **9**, 647-654 (2002).
78. Niesen, F.H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221 (2007).
79. Lee, S.G., Lee, J.Y. & Chmielewski, J. Investigation of pH-dependent collagen triple-helix formation. *Angew Chem Int Ed Engl* **47**, 8429-8432 (2008).
80. Chen, L., Cai, S., Lim, J., Lee, S.S. & Lee, S.G. Elucidating pH-dependent collagen triple helix formation through interstrand hydroxyproline-glutamic acid interactions. *Chembiochem* **16**, 407-410 (2015).
81. Xia, F., Nagrath, D. & Cramer, S.M. Effect of pH changes on water release values in hydrophobic interaction chromatographic systems. *J Chromatogr A* **1079**, 229-235 (2005).

82. Emsley, J., Knight, C.G., Farndale, R.W., Barnes, M.J. & Liddington, R.C. Structural basis of collagen recognition by integrin alpha2beta1. *Cell* **101**, 47-56 (2000).
83. Pires, M.M., Lee, J., Ernenwein, D. & Chmielewski, J. Controlling the morphology of metal-promoted higher ordered assemblies of collagen peptides with varied core lengths. *Langmuir* **28**, 1993-1997 (2012).
84. Przybyla, D.E. & Chmielewski, J. Metal-triggered radial self-assembly of collagen peptide fibers. *J Am Chem Soc* **130**, 12610-12611 (2008).
85. Przybyla, D.E. & Chmielewski, J. Metal-triggered collagen peptide disk formation. *J Am Chem Soc* **132**, 7866-7867 (2010).
86. Przybyla, D.E., Rubert Perez, C.M., Gleaton, J., Nandwana, V. & Chmielewski, J. Hierarchical assembly of collagen peptide triple helices into curved disks and metal ion-promoted hollow spheres. *J Am Chem Soc* **135**, 3418-3422 (2013).
87. Rele, S. et al. D-periodic collagen-mimetic microfibers. *J Am Chem Soc* **129**, 14780-14787 (2007).
88. O'Leary, L.E., Fallas, J.A., Bakota, E.L., Kang, M.K. & Hartgerink, J.D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nat Chem* **3**, 821-828 (2011).
89. Haines-Butterick, L. et al. Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells. *Proc Natl Acad Sci U S A* **104**, 7791-7796 (2007).
90. Faruqui, N. et al. Differentially instructive extracellular protein micro-nets. *J Am Chem Soc* **136**, 7889-7898 (2014).
91. Savitzky, A. & Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**, 1627-1639 (1964).
92. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent Eur J Phys* **10**, 181-188 (2012).
93. Horcas, I. et al. WSXM: a software for scanning probe microscopy and a tool for nanotechnology. *Rev Sci Instrum* **78**, 013705 (2007).
94. MATLAB version 8.03.0 (R2014a). (The MathWorks Inc., Natick, Massachusetts; 2014).
95. McGuinness, K., Khan, I.J. & Nanda, V. Morphological diversity and polymorphism of self-assembling collagen peptides controlled by length of hydrophobic domains. *ACS Nano* **8**, 12514-12523 (2014).
96. Emsley, J., Knight, C.G., Farndale, R.W., Barnes, M.J. & Liddington, R.C. Structural Basis of Collagen Recognition by Integrin α2β1. *Cell* **101**, 47-56 (2000).
97. Helseth, D.L., Jr. & Veis, A. Collagen self-assembly in vitro. Differentiating specific telopeptide-dependent interactions using selective enzyme modification and the addition of free amino telopeptide. *J Biol Chem* **256**, 7118-7128 (1981).
98. Li, Y. et al. Targeting collagen strands by photo-triggered triple-helix hybridization. *Proc Natl Acad Sci U S A* **109**, 14767-14772 (2012).
99. Giasson, B.I., Murray, I.V., Trojanowski, J.Q. & Lee, V.M. A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J Biol Chem* **276**, 2380-2386 (2001).
100. Landreh, M. & Robinson, C.V. A new window into the molecular physiology of membrane proteins. *J Physiol* **593**, 355-362 (2015).

101. Siebert, C.A. et al. Molecular architecture of photosynthetic membranes in Rhodobacter sphaeroides: the role of PufX. *Embo J* **23**, 690-700 (2004).
102. Killian, J.A. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim Biophys Acta* **1376**, 401-415 (1998).
103. Song, W.S. & Yoon, S.I. Crystal structure of FliC flagellin from Pseudomonas aeruginosa and its implication in TLR5 binding and formation of the flagellar filament. *Biochem Biophys Res Commun* **444**, 109-115 (2014).
104. Greenfield, N.J. et al. Solution NMR structure of the junction between tropomyosin molecules: implications for actin binding and regulation. *J Mol Biol* **364**, 80-96 (2006).
105. Bruns, R.R., Hulmes, D.J., Therrien, S.F. & Gross, J. Procollagen segment-long-spacing crystallites: their role in collagen fibrillogenesis. *Proc Natl Acad Sci U S A* **76**, 313-317 (1979).
106. Kotha, R.R. & Chmielewski, J. Controlling the morphology of metal-triggered collagen peptide assemblies through ligand alteration. *Biopolymers* **104**, 379-383 (2015).
107. Nobeli, I., Favia, A.D. & Thornton, J.M. Protein promiscuity and its implications for biotechnology. *Nat Biotechnol* **27**, 157-167 (2009).
108. Schreiber, G. & Keating, A.E. Protein binding specificity versus promiscuity. *Curr Opin Struct Biol* **21**, 50-61 (2011).
109. Martin, J. Beauty is in the eye of the beholder: proteins can recognize binding sites of homologous proteins in more than one way. *PLoS Comput Biol* **6**, e1000821 (2010).
110. Ma, C.D., Wang, C., Acevedo-Velez, C., Gellman, S.H. & Abbott, N.L. Modulation of hydrophobic interactions by proximally immobilized ions. *Nature* **517**, 347-350 (2015).
111. Pires, M.M. & Chmielewski, J. Self-assembly of collagen peptides into microflorettes via metal coordination. *J Am Chem Soc* **131**, 2706-2712 (2009).
112. Jiang, T., Vail, O.A., Jiang, Z., Zuo, X. & Conticello, V.P. Rational Design of Multilayer Collagen Nanosheets with Compositional and Structural Control. *J Am Chem Soc* **137**, 7793-7802 (2015).
113. Stapleton, J.A., Whitehead, T.A. & Nanda, V. Computational redesign of the lipid-facing surface of the outer membrane protein OmpA. *Proc Natl Acad Sci U S A* **112**, 9632-9637 (2015).
114. Okuyama, K. et al. Unique side chain conformation of a Leu residue in a triple-helical structure. *Biopolymers* **86**, 212-221 (2007).
115. Guinier, A. & Fournet, G. Small-angle scattering of X-rays. (Wiley, New York,; 1955).
116. Glatter, O. & Kratky, O. Small angle x-ray scattering. (Academic Press, London ; New York; 1982).
117. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. & Svergun, D.I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277-1282 (2003).
118. Microsoft Microsoft Excel. (2010).
119. Woodhead Galloway, J. Closest Packing of Multistranded Coiled-Coils a Possible Application to Collagen. *Acta crystallogr. B. Struct Sci, Cryst Eng Mat* **32** (1976).

120. Straley, J.P. Theory of Piezoelectricity in the nematic liquid crystals, and of the cholesteric ordering. *Phys Rev A* **14** (1976).
121. Chothia, C., Levitt, M. & Richardson, D. Helix to helix packing in proteins. *J Mol Biol* **145**, 215-250 (1981).
122. Nanda, V. & DeGrado, W.F. Computational design of heterochiral peptides against a helical target. *J Am Chem Soc* **128**, 809-816 (2006).
123. Boncheva, M., Bruzewicz, D.A. & Whitesides, G.M. Formation of chiral, three-dimensional aggregates by self-assembly of helical components. *Langmuir* **19**, 6066-6071 (2003).
124. Berisio, R., Vitagliano, L., Mazzarella, L. & Zagari, A. Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3). *Protein Sci* **11**, 262-270 (2002).