

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Rebecca A. Bartlett

Date

Engineering Advanced Self-Assembling Protein Biomaterials through the Extrapolation and Functionalization of Peptide-Based Systems

By

Rebecca A. Bartlett
Doctor of Philosophy

Chemistry

Dr. Vincent P. Conticello
Advisor

Dr. Khalid Salaita
Committee Member

Dr. Emily Weinert
Committee Member

Accepted:

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

Date

Engineering Advanced Self-Assembling Protein Biomaterials through the Extrapolation and Functionalization of Peptide-Based Systems

By

Rebecca A. Bartlett
B.S., University of North Georgia, 2011
M.S., Emory University, 2016

Advisor: Vincent P. Conticello, Ph.D.

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

Doctor of Philosophy
in Chemistry

2018

Abstract

Engineering Advanced Self-Assembling Protein Biomaterials through the Extrapolation and Functionalization of Peptide-Based Systems

By Rebecca A. Bartlett

Chemically synthesized peptides represent useful building blocks for engineering self-assembling biomaterials because their rapid production permits high-throughput investigation of numerous peptide sequences in a brief time period. However, chemical synthesis limits the length and complexity of the generated peptides, which subsequently limits the potential applications. Despite these restrictions, useful information from successful self-assembling peptides can be extended to the development of more intricate systems through the use of genetic and protein engineering using *E. coli* as a host organism.

A previously published α -helical peptide was designed to assemble intermolecularly into a two-dimensional nanoporous framework where the thickness of the nanosheets was determined by the length of the peptide. Utilizing molecular cloning techniques, a system was designed for concatemerization of the α -helical peptide to generate nanosheets with greater thickness and improved thermodynamic and mechanical stability. This would support the application of these materials in therapeutic drug delivery or through incorporation into nanodevices.

In an unrelated investigation, HEAT-based concatemers were produced via genetic and protein engineering techniques for the development of thermodynamically stable, high aspect ratio nanotubes with two functionally distinct surfaces. Hypervariable residues on the concave surface are amenable to modification for application purposes. Meanwhile, the exterior surfaces of HEAT-based protein nanotubes were rapidly functionalized with the mCherry fluorescent protein through the formation of a covalent isopeptide bond by the SpyTag/SpyCatcher ligation system. Functionalization of the optimized HEAT-based protein nanotubes displaying the SpyTag peptide can promptly result from the introduction of the SpyCatcher protein fused to any arbitrary functional protein. The asymmetrical and interchangeable functionalization of these nanotubes results in a highly controllable system for various applications.

Engineering Advanced Self-Assembling Protein Biomaterials through the Extrapolation and Functionalization of Peptide-Based Systems

By

Rebecca A. Bartlett
B.S., University of North Georgia, 2011
M.S., Emory University, 2016

Advisor: Vincent P. Conticello, Ph.D.

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

Doctor of Philosophy
in Chemistry

2018

Acknowledgements

I would first like to thank my advisor, Dr. Vince Conticello, for his consistent support and patience throughout my graduate career. He is such a kind and wonderful mentor, and his immeasurable knowledge about science has always inspired me. I am honored to have worked as a graduate student in his lab and to have had the opportunity to learn from him. The skills that I have developed under his advisement will surely help me succeed throughout my career. Thank you also to my committee members, Dr. Emily Weinert and Dr. Khalid Salaita, for their helpful guidance. I am grateful for all of the advice and direction they provided as it helped me excel in my research. They took valuable time out of their days to sit down with me and talk about my research, career, and life goals. There are few other programs where faculty care so much about the education of the students as they do at Emory.

A huge thanks also goes to my mentors at North Georgia, Dr. Holly Carpenter, Dr. Brad Herbert, and Dr. Michael McGinnis. They were always there for me and took the time to help me with any problem. I had a family in the chemistry department at North Georgia and that still does not fully express the immense gratitude I have for these individuals. Without the support and encouragement of my mentors at North Georgia, I would not have become a research scientist. Thank you for being such powerful and inspiring role models.

Thanks to my lab members Charles Modlin, Spencer Hughes, Shengyuan Wang, Dr. Andrea Merg, Avi Tuachi, and Gavin Touponse. They contributed to my wonderful experience in graduate school. I am particularly grateful to Charles for his words of encouragement and support as we tackled every major stepping stone of the graduate experience. Without his assistance and humor, graduate school would have been less enjoyable. Spencer and Charles always provided an

exciting and enjoyable workspace with their comical murals and photoshopped “Conticello family” portraits. I will miss those. I greatly appreciate the help of current and past lab members, Charles Modlin, Spencer Hughes, Shengyuan Wang, Dr. Andrea Merg, Dr. I-Lin Wu, Dr. Chunfu Xu, Dr. Tao Jiang, and Dr. Elizabeth Magnotti, with biophysical characterization techniques as I transitioned from working with *E. coli* for protein expression into the realm of peptide and protein assemblies. They were all enthusiastic to train me on a new instrument or help me work through a particularly difficult problem.

Thank you to my collaborators, who without their help, many aspects of my research would not be possible. Dr. Xiaobing Zuo from Argonne National Lab aided with small-angle X-ray scattering analysis of peptide and protein assemblies. The Emory University Robert P. Apkarian Integrated Electron Microscopy Core staff, Hong Yi, Jeannette Taylor, and Art McCanna, was extremely helpful with training and assistance on the transmission electron microscope. Dr. Elizabeth Wright assisted with the acquisition of electron cryomicroscopy images. Dr. Fred Strobel aided with mass spectrometry for some particularly recalcitrant proteins. Dr. Tim Lian provided the atomic force microscope for collecting height measurements, and Dr. Stefan Lutz provided the fluorimeter for fluorescence measurements. The assistance and training from these individuals were imperative for the completion of my research.

Thank you to all of my amazing friends who have supported me, encouraged me, and listened to me through all of the highs and lows of graduate school. To Nick, Sarah, Wes, Haley, Sam, Sara, Christina, Jacob, Theo, Ashlynn, Manny, Kate, Thomas, Carolina, Renato, Ram, Joseph, Luis, Nick, Anna, and many others, thank you all. Without you, this would have been impossible.

Although she was a previous member of the Conticello family, my undergraduate research mentor at North Georgia, and a previous employer, Dr. Holly Carpenter deserves an especially dedicated thank you. I have immeasurable gratitude for everything she has done for me and I am indebted to her for life. She has been a consistent and strong support system for me over the past 10 years, and she became my dearest friend. The advice and guidance she provided over the past decade has been invaluable, and she has always ensured that I followed my goals to the fullest extent. I will remember our talks forever, and I look forward to many more. Thank you for always being there for me!

A special thanks to my parents, Helen and Duane. Without your unwavering love, support, and encouragement, I would not be where I am today. You have supported my decisions and taught me to never give up, work hard, and follow my goals with determination. There are no words that can express the level of appreciation I have for everything you have done for me over the years. Thank you also to my sister and my brother. I am forever grateful to the support you have shown me over the years. We share a special understanding that will keep us strong forever. You are both special to me in so many ways. Thank you for almost 30 years of trust, support, and love.

Lastly, to the person I thank most of all, my fiancé and best friend, Skylar Hyatt. You have consistently provided me with encouragement, support, patience, and comfort since the day we met. The strong foundation of our relationship and love has provided me with strength through the darkest moments of graduate school and life. You are my true counterpart, and I am so excited for the life we will share together. I dedicate this dissertation to Skylar and our baby beagle, Buddy.

Table of Contents

Chapter 1. Self-Assembly of α-Helical Proteins into Thermodynamically Stable Two-Dimensional Nanosheets	1
1.1 Introduction	2
1.2 Two-Dimensional Peptide Assemblies	2
1.3 3FD-IL and 3FD-LL α -Helical Peptide Nanosheets	5
1.4 Genetic Engineering of Biomaterials	12
1.5 Conclusions	15
1.6 References	16
Chapter 2. Development of a System for the Generation of 3FD-LL Concatemers	27
2.1 Introduction	28
2.2 Results and Discussion	31
2.2.1 Genetic Engineering of 3FD-LL Concatemer Sequences	31
Design of Adapter and Insertion Genes	31
Construction of the Expression Vector	36
Generation of the 3FD-LL Monomer Gene Sequence	40
Concatemerization of 3FD-LL Sequence	42
2.2.2 Generation of the 3FD-LL Concatemer Proteins	45
Expression of 3FD-LL Concatemers	45
Cell Lysis	48
Protein Solubilization	48
Inclusion Body Purification	52
Purification of the 3FD-LL Concatemers	54

2.3 Conclusions	56
2.4 Experimental	57
2.4.1 Materials	57
2.4.2 General Methods	58
2.4.3 Genetic Engineering of 3FD-LL Concatemer Sequences	60
2.4.4 Production of the 3FD-LL Concatemer Proteins	70
2.4.5 Tables	79
2.5 References	82
Chapter 3. Production of 3FD-LL Concatemer Nanosheet Assemblies	83
3.1 Introduction	84
3.2 Results and Discussion	86
3.2.1 Initial Assembly and Optimization of 3FD-LL Concatemers	86
Assembly of HisPur™ 3FD-LL Concatemers	86
Assembly of HisPur™ 3FD-LL Concatemers with TFE	102
3.2.2 Characterization of HPLC Purified 3FD-LL Trimer Nanosheets	121
Circular Dichroism Spectropolarimetry	123
Transmission Electron Microscopy	127
Atomic Force Microscopy	129
3.3 Conclusions	132
3.4 Experimental	134
3.4.1 Materials	134
3.4.2 General Methods	135
3.4.3 Initial Assembly and Optimization of 3FD-LL Concatemers	136

3.4.4	Assembly and Characterization of HPLC Purified 3FD-LL Trimer	138
3.4.5	Tables	142
3.5	References	143
Chapter 4.	Structurally Defined Helical Nanotubes for the Construction of Functionally Asymmetric Hybrid Nanomaterials	145
4.1	Introduction	146
4.2	Tandem Repeat Proteins (TRPs)	148
4.3	HEAT_R1 Peptide Assemblies	153
4.4	HEAT_6R Protein Assemblies	158
4.5	SpyTag/SpyCatcher Ligation System	162
4.6	Conclusions	168
4.7	References	169
Chapter 5.	Design and Production of SpyTag_HEAT and mCherry_SpyCatcher	174
5.1	Introduction	175
5.2	Results and Discussion	179
5.2.1	Generation of ST_HEAT	179
	Sequence Design	179
	Protein Expression and Purification	181
5.2.2	Generation of mCh_SC	185
	Sequence Design	185
	Protein Expression and Purification	189
5.2.3	Initial Verification of Linking Between ST_HEAT and mCh_SC	197
	Linking of ST_HEAT Lysate with mCh_SC Lysate	197

Linking of HisPur™ Eluents of ST_HEAT and mCh_SC	199
5.3 Conclusions	201
5.4 Experimental	202
5.4.1 Materials	202
5.4.2 General Methods	203
5.4.3 SpyTag_HEAT Production	204
5.4.4 mCherry_SpyCatcher Production	209
5.4.5 Verification of Linking Between ST_HEAT and mCh_SC	216
5.4.6 Tables	218
5.5 References	220
Chapter 6. Characterization of ST_HEAT Nanotubes and Functionalization with	
mCherry_SpyCatcher	223
6.1 Introduction	224
6.2 Results and Discussion	226
6.2.1 Assembly and Characterization of SpyTag_HEAT Nanotubes	226
Low Concentration HisPur™ ST_HEAT Assemblies	226
High Concentration HisPur™ ST_HEAT Assemblies	232
High Concentration Assemblies of HPLC Purified ST_HEAT	235
Low Concentration Assemblies of HPLC Purified ST_HEAT	239
6.2.2 Functionalization of ST_HEAT Nanotubes with mCh_SC Protein	244
Functionalization of HisPur™ ST_HEAT with mCh_SC	245
Functionalization of HPLC Purified ST_HEAT with mCh_SC	248
Optimizing Amount of mCh_SC for Maximum Coverage	254

Ultracentrifugation for Removal of Excess mCh_SC	256
Centrifugal Filtration for Removal of Excess mCh_SC	263
Fluorimetry Analysis	268
Correlative Light Electron Microscopy (CLEM)	272
6.3 Conclusions	274
6.4 Experimental	277
6.4.1 Materials	277
6.4.2 General Methods	278
6.4.3 Assembly and Characterization of SpyTag_HEAT Nanotubes	279
6.4.4 Functionalization of ST_HEAT Nanotubes with mCh_SC	284
6.4.5 Tables	295
6.5 References	296
Appendix A. Sequences	302
Appendix B. Mass Spectra	313

List of Figures

Chapter 1. Self-Assembly of α -Helical Proteins into Thermodynamically Stable Two-Dimensional Nanosheets

Figure 1.1 3FD-IL and 3FD-LL amino acid sequences and helical wheels	7
Figure 1.2 3FD-IL and 3FD-LL CD spectropolarimetry analysis	9
Figure 1.3 3FD-IL and 3FD-LL TEM analysis	10
Figure 1.4 3FD-IL and 3FD-LL AFM analysis	11
Figure 1.5 Proposed hexagonal honeycomb packing of 3FD-IL	12

Chapter 2. Development of a System for the Generation of 3FD-LL Concatemers

Figure 2.1 Schematic for concatemerization of DNA cassettes	30
Figure 2.2 Gene sequence for 3FD-LL helix adapter	33
Figure 2.3 Gene sequence for 3FD-LL helix monomer insert	34
Figure 2.4 Plasmid maps for pBB12 and pBB13	35
Figure 2.5 Plasmid map for pQE-80L vector from QIAGEN	37
Figure 2.6 Plasmid map for pIL5 expression vector	38
Figure 2.7 Schematic for formation of pBB14	39
Figure 2.8 Schematic for formation of pBB15	41
Figure 2.9 Schematic for formation of pBB17	43
Figure 2.10 Schematic for formation of pBB20, pBB21, and pBB22	44
Figure 2.11 SDS-PAGE gel of expression results for the 3FD-LL concatemers	46
Figure 2.12 Western blot of expression results for the 3FD-LL concatemers	47
Figure 2.13 SDS-PAGE gel of lysis results for the 3FD-LL concatemers	49
Figure 2.14 SDS-PAGE gel of urea solubilization of the 3FD-LL concatemers	50

Figure 2.15 SDS-PAGE gel of GHCl solubilization of the 3FD-LL concatemers	51
Figure 2.16 SDS-PAGE gel of inclusion body washes of the concatemer samples	53
Figure 2.17 SDS-PAGE gel and western blot of concatemer HisPur™ purification	55
Figure 2.18 Agarose gel of plasmid pIL5	62
Figure 2.19 Agarose gel of plasmid pBB15, pBB17, pBB20, pBB21, and pBB22	69
Chapter 3. Production of 3FD-LL Concatemer Nanosheet Assemblies	
Figure 3.1 TEM analysis of 3FD-LL monomer in MOPS buffer, pH 7.0	88
Figure 3.2 TEM analysis of 3FD-LL monomer in TAPS buffer, pH 8.5	89
Figure 3.3 TEM analysis of 3FD-LL dimer in MOPS buffer, pH 7.0	90
Figure 3.4 TEM analysis of 3FD-LL dimer in TAPS buffer, pH 8.5	91
Figure 3.5 TEM analysis of 3FD-LL trimer in MOPS buffer, pH 7.0	92
Figure 3.6 TEM analysis of 3FD-LL trimer in TAPS buffer, pH 8.5	93
Figure 3.7 TEM analysis of 3FD-LL tetramer in MOPS buffer, pH 7.0	94
Figure 3.8 TEM analysis of 3FD-LL tetramer in TAPS buffer, pH 8.5	95
Figure 3.9 TEM analysis of 3FD-LL pentamer in MOPS buffer, pH 7.0	96
Figure 3.10 TEM analysis of 3FD-LL pentamer in TAPS buffer, pH 8.5	97
Figure 3.11 CD analysis of 3FD-LL monomer	99
Figure 3.12 CD analysis of 3FD-LL dimer	99
Figure 3.13 CD analysis of 3FD-LL trimer	100
Figure 3.14 CD analysis of 3FD-LL tetramer	100
Figure 3.15 CD analysis of 3FD-LL pentamer	101
Figure 3.16 TEM analysis of unannealed 3FD-LL trimer in TFE; MOPS, pH 7.0	104
Figure 3.17 TEM analysis of unannealed 3FD-LL trimer in TFE; TAPS, pH 8.5	105

Figure 3.18	TEM analysis of 3FD-LL monomer assembled in TFE; MOPS, pH 7.0	108
Figure 3.19	TEM analysis of 3FD-LL monomer assembled in TFE; TAPS, pH 8.5	109
Figure 3.20	TEM analysis of 3FD-LL dimer assembled in TFE; MOPS, pH 7.0	110
Figure 3.21	TEM analysis of 3FD-LL dimer assembled in TFE; TAPS, pH 8.5	111
Figure 3.22	TEM analysis of 3FD-LL trimer assembled in TFE; MOPS, pH 7.0	112
Figure 3.23	TEM analysis of 3FD-LL trimer assembled in TFE; TAPS, pH 8.5	113
Figure 3.24	TEM analysis of 3FD-LL tetramer assembled in TFE; MOPS, pH 7.0	114
Figure 3.25	TEM analysis of 3FD-LL tetramer assembled in TFE; TAPS, pH 8.5	115
Figure 3.26	TEM analysis of 3FD-LL pentamer assembled in TFE; MOPS, pH 7.0	116
Figure 3.27	TEM analysis of 3FD-LL pentamer assembled in TFE; TAPS, pH 8.5	117
Figure 3.28	CD analysis of 3FD-LL monomer assembled in TFE	119
Figure 3.29	CD analysis of 3FD-LL dimer assembled in TFE	119
Figure 3.30	CD analysis of 3FD-LL trimer assembled in TFE	120
Figure 3.31	CD analysis of 3FD-LL tetramer assembled in TFE	120
Figure 3.32	CD analysis of 3FD-LL pentamer assembled in TFE	121
Figure 3.33	CD analysis of HPLC purified 3FD-LL trimer assembled in TFE	123
Figure 3.34	CD melting curve for HPLC purified 3FD-LL trimer	124
Figure 3.35	CD analysis of HPLC pure 3FD-LL trimer before and after melt	126
Figure 3.36	TEM analysis of nascent HPLC pure 3FD-LL trimer assemblies	127
Figure 3.37	TEM analysis of annealed HPLC pure 3FD-LL trimer assemblies	128
Figure 3.38	AFM analysis of HPLC purified 3FD-LL trimer assemblies	131

Chapter 4. Structurally Defined Helical Nanotubes for the Construction of Functionally Asymmetric Hybrid Nanomaterials

Figure 4.1	Structure of LRV protein based on crystal structure	150
Figure 4.2	Crystal structure of α Rep- <i>n</i> 4-1 tetramer	152
Figure 4.3	CD analysis of HEAT_R1 peptide assemblies	154
Figure 4.4	TEM analysis of HEAT_R1 peptide assemblies	155
Figure 4.5	Reconstruction of HEAT_R1 nanotubes	157
Figure 4.6	CD analysis of HEAT_6R peptide	160
Figure 4.7	TEM analysis of HEAT_6R peptide assemblies	161
Figure 4.8	Schematic for the BIND system immobilization of proteins	165
Figure 4.9	Crystal structure and reaction mechanism of SpyTag/SpyCatcher	166
Chapter 5. Design and Production of SpyTag_HEAT and mCherry_SpyCatcher		
Figure 5.1	Structure of mCherry protein based on crystal structure	178
Figure 5.2	Designed amino acid sequence for ST_HEAT	180
Figure 5.3	Plasmid map encoding the ST_HEAT_8R sequence	181
Figure 5.4	SDS-PAGE gel of expression and lysis results for ST_HEAT	182
Figure 5.5	SDS-PAGE gel of urea solubilization of ST_HEAT	183
Figure 5.6	SDS-PAGE gel of HisPur TM purification of ST_HEAT	184
Figure 5.7	TEV cleavage recognition sequence	186
Figure 5.8	Designed amino acid sequence for mCh_SC	187
Figure 5.9	Plasmid map encoding the mCh_SC sequence	188
Figure 5.10	Media after the expression of the mCh_SC protein	189
Figure 5.11	SDS-PAGE gel of expression and lysis results for mCh_SC	190
Figure 5.12	Supernatant after the cell lysis of mCh_SC samples	191
Figure 5.13	SDS-PAGE gel of HisPur TM purification of mCh_SC	192

Figure 5.14	SDS-PAGE gel of mCh_SC before and after TEV cleavage	193
Figure 5.15	Images of mCh_SC color and fluorescence at pH 4.5	194
Figure 5.16	BSA standard curve utilized for concentration calculation	196
Figure 5.17	SDS-PAGE gel of mCh-HEAT formation in cell lysate	198
Figure 5.18	SDS-PAGE gel of mCh-HEAT formation in HisPur™ eluent	200
Chapter 6.	Characterization of ST_HEAT Nanotubes and Functionalization with mCherry_SpyCatcher	
Figure 6.1	CD analysis of HisPur™ ST_HEAT at low concentration	228
Figure 6.2	TEM analysis of HisPur™ ST_HEAT at low concentration	229
Figure 6.3	TEM analysis of HisPur™ ST_HEAT with EDTA	231
Figure 6.4	CD analysis of HisPur™ ST_HEAT at high concentration	233
Figure 6.5	TEM analysis of HisPur™ ST_HEAT at high concentration	234
Figure 6.6	CD analysis of HPLC ST_HEAT at high concentration	237
Figure 6.7	TEM analysis of HPLC ST_HEAT at high concentration	238
Figure 6.8	TEM analysis of HPLC ST_HEAT at low concentration	240
Figure 6.9	ImageJ measurements of ST_HEAT nanotube diameters	241
Figure 6.10	SAXS scattering plot for ST_HEAT nanotubes	242
Figure 6.11	SAXS modified Guinier plot for ST_HEAT nanotubes	243
Figure 6.12	TEM analysis of nascent HisPur™ ST_HEAT with mCh_SC	247
Figure 6.13	TEM analysis of annealed HisPur™ ST_HEAT with mCh_SC	249
Figure 6.14	TEM analysis of nascent HPLC ST_HEAT with mCh_SC	250
Figure 6.15	TEM analysis of annealed HPLC ST_HEAT with mCh_SC	251
Figure 6.16	SDS-PAGE gel of mCh-HEAT ligation results	253

Figure 6.17 TEM analysis of ST_HEAT with varying mCh_SC concentrations	255
Figure 6.18 SDS-PAGE gel of ultracentrifuge (30k x g) results	257
Figure 6.19 TEM analysis of ultracentrifuge (30k x g) results	259
Figure 6.20 SDS-PAGE gel of ultracentrifuge (72.7k x g) results	260
Figure 6.21 TEM analysis of ultracentrifuge (72.7k x g) results	262
Figure 6.22 TEM analysis of mCh-HEAT after centrifugal filtration (50 kDa)	264
Figure 6.23 TEM analysis of mCh-HEAT after centrifugal filtration (100 kDa)	266
Figure 6.24 SDS-PAGE gel of samples after centrifugal filtration (100 kDa)	267
Figure 6.25 Fluorimetry data for mCh-HEAT samples	269
Figure 6.26 Fluorimetry data for various mCh_SC samples	271
Figure 6.27 Cryo-EM analysis of mCh-HEAT nanotubes	273

Appendix B. Mass Spectra

Figure B.1 MALDI of HPLC purified 3FD-LL Trimer	314
Figure B.2 MALDI of HisPur™ 3FD-LL Monomer	315
Figure B.3 MALDI of HisPur™ 3FD-LL Dimer	316
Figure B.4 MALDI of HisPur™ 3FD-LL Tetramer	317
Figure B.5 MALDI of HisPur™ 3FD-LL Pentamer	318
Figure B.6 ESI mass spectra of TEV cleavage product of mCh_SC	319
Figure B.7 MALDI of HPLC purified ST_HEAT	320

List of Tables

Chapter 2. Development of a System for the Generation of 3FD-LL Concatemers

Table 2.1 Plasmids used in Chapter 2 79

Table 2.2 *E. coli* strains used in Chapter 2 80

Table 2.3 Primers used in Chapter 2 for DNA sequencing 80

Table 2.4 Protein sequences used in Chapter 2 81

Chapter 3. Production of 3FD-LL Concatemer Nanosheet Assemblies

Table 3.1 Protein sequences used in Chapter 3 142

Chapter 5. Design and Production of SpyTag_HEAT and mCherry_SpyCatcher

Table 5.1 Plasmids used in Chapter 5 218

Table 5.2 *E. coli* strains used in Chapter 5 218

Table 5.3 Protein sequences used in Chapter 5 219

Chapter 6. Characterization of ST_HEAT Nanotubes and Functionalization with mCherry_SpyCatcher

Table 6.1 Protein sequences used in Chapter 6 295

Chapter 1

Self-Assembly of α -Helical Proteins into Thermodynamically Stable Two-Dimensional Nanosheets

1.1 Introduction

Nanoarchitectonics is a rising field centered on the development of dynamic functional materials through selective tuning at the molecular level to control the subsequent material properties.¹⁻³ Biological molecules, such as nucleic acids,⁴⁻⁷ proteins,⁸⁻⁹ and structurally related foldamers,¹⁰ exemplify appealing candidates to extend these principles for construction of functional biomaterials. Variations in the amino acid sequence of peptides and proteins result in well understood modifications in the intra and intermolecular interactions, which make them important candidates for controlled self-assembly into highly ordered functional materials. One-dimensional peptide and protein nanomaterials are frequently found in natural systems in organelle and tissue components such as cytoskeleton frameworks and collagen fibers.¹¹⁻¹³ Rational design of synthetic peptides and proteins has resulted in the creation of many one-dimensional assemblies from structural motifs including β -strands,¹⁴⁻²³ β -hairpins,²⁴⁻²⁷ α -helical coiled-coils,²⁸⁻⁴⁰ and collagen triple helices.⁴¹⁻⁵⁰

1.2 Two-Dimensional Peptide Assemblies

Two-dimensional peptide and protein assemblies are scarcely seen in nature although they display potential for applications in nanoarchitectonics as their extended planar shapes can straightforwardly be incorporated into current nanoscale devices.⁵¹ Nanosheets are two-dimensional assemblies that could be utilized as scaffolds and functionalized at their surface for applications in filtration, energy storage, electronics, and immobilized enzymatic reactions. Collagen-mimetic peptides,⁵²⁻⁵⁶ β -sheet peptides,⁵⁷⁻⁵⁸ peptoids,⁵⁹⁻⁶² and globular proteins⁶³⁻⁶⁶ have been successfully designed for the fabrication of two-dimensional biomaterials; however, they present limitations that reduce their controllability.

Synthetic collagen-mimetic peptides have been utilized for the design of biomaterials based on the properties of native collagen. Native collagen strands form left-handed triple helices with a polyproline-II secondary structure.⁶⁷ Tripeptide repeats of Xaa-Yaa-Gly compose collagen sequences, where Xaa and Yaa are typically proline and the unnatural amino acid (4R)-hydroxyproline, respectively.⁶⁸ Modifications to the tripeptide repeat have resulted in the formation of two-dimensional assemblies exhibiting metal-ion coordination,⁶⁹ hydrophobic interactions,⁷⁰ ionic interactions,⁵²⁻⁵⁴ di-sulfide bonding interactions,^{44, 49-50} a mixture of left- and right-handed triple helices,⁵⁶ and the incorporation of aromatic amino acids.^{46, 69} Although many alterations can be made to the Xaa and Yaa positions to adjust the properties of the biomaterials, the sequence modifications are bound by the three amino acid repeat, which restricts sequence flexibility. Additionally, thermal denaturation of collagen-mimetic nanosheets display melting temperatures between 28 °C and 76 °C indicating that most of the materials exhibit relatively low thermodynamic stability.⁵²⁻⁵⁴ Two-dimensional biomaterials with higher thermostability would need to be considered for incorporation in nanoscale devices.

Due to their similarities to peptides, peptoids have also been investigated for two-dimensional material design. Peptoids, or N-substituted glycines, are formed from non-natural and achiral building blocks based on amino acids where the R-group side chain is relocated from the alpha carbon to the backbone nitrogen. Similar to peptides, the secondary and tertiary structures are directed by the primary sequence and driven by the hydrophobic effect. As a result of the relocation of the side chain, hydrogen is no longer present on the nitrogen atoms in the backbone. Though the absence of hydrogen bonding between the amide and carbonyl groups in the peptoid backbone grants simplicity to the design, this comes at the cost of decreased control over the resulting structure.^{62, 71} While synthesis of sequence specific peptoids is possible via a solid-phase

submonomer method, the achirality and subsequent minimized control over the structures compared to peptides decrease the practicality of using these species for stable two-dimensional materials.

Although significant progress had been made towards the fabrication of two-dimensional materials from peptides, proteins, and structurally related foldamers with advances in interfaces between the building blocks, limitations in sequence flexibility, structural controllability, and thermodynamic stability decrease their potential use in functional devices. While structural analysis of native proteins resulted in a thorough understanding of interhelical packing interactions,⁷²⁻⁷³ the use of synthetic α -helical coiled-coil peptides as a major building block for two-dimensional nanosheets had not been considered as a likely candidate primarily because the crossing angles between helices prevent tight interhelical packing. The left-handed superhelical twist found in traditional coiled-coil peptides results from a crossing angle of 20° , which causes a curvature between neighboring helices with limited contact that would not readily promote extended two-dimensional assemblies.⁷⁴⁻⁷⁵

Crystallography of materials composed of straight α -helices have been found to pack into layers within lattices.⁷⁶⁻⁸² Although the presence of these layers indicates a potential for two-dimensional assemblies, the short peptides used to form these layers are thermodynamically unstable unless helix-promoters (e.g., trifluoroethanol or unnatural helicogenic amino acids) are present in solution or in the protein design. Additionally, the specificity of interactions between the lateral helices in these layers are not well characterized. Nonetheless, conformationally stable assemblies have since been discovered through computational design by minimizing the superhelical twist between trimeric α -helical bundles.⁸³ A combination of a left-handed coiled coil with a right-handed coiled coil results in a cancelation of the two opposing supercoil twists.

Reducing the superhelical twist supports the potential for using straight α -helices for developing stable two-dimensional assemblies, especially if the canonical 18 amino acid repeat is maintained.⁸⁴

1.3 3FD-IL and 3FD-LL α -Helical Peptide Nanosheets

More recently, our lab published the design and self-assembly of a synthetic peptide, **3FD-IL**, which comprised two 18 amino acid repeats with 3-fold screw symmetry.⁸⁵ Biophysical analysis supported that the peptide folded into straight α -helix subunits that assembled intermolecularly to form two-dimensional nanosheets with nanoscale channels permeating the sheets. Furthermore, AFM data indicated that the thickness of the sheets was determined by the length of the peptide, which indicated that helices were aligning perpendicular to the surface of the nanosheet. Applications such as therapeutic delivery could take advantage of these nanosheets because of their extended two-dimensional porous scaffolds.

Left-handed coiled-coil peptides pack together through hydrophobic interactions between the residues in the *a* and *d* positions of the heptad.^{75, 86} Isoleucine, which was present in the *a* position in **3FD-IL**, is a β -branched amino acid. Leucine, which is not a β -branched amino acid, could allow for tighter packing and fewer steric hindrances between neighboring helices. The peptide **3FD-LL** was designed to contain leucine residues at all hydrophobic positions to investigate the effect of changing the hydrophobic residues on nanosheet formation.⁸⁷ The synthetic peptide **3FD-LL** was found to self-assemble into nanosheets with greater thermodynamic stability than those formed from **3FD-IL**.

The peptides **3FD-LL** and **3FD-IL** were designed based on sequence constraints for left- and right-handed coil-coiled peptides. The core 18 amino acid repeat of both peptides was

fundamentally composed of the seven amino acid heptad of a left-handed coiled-coil peptide followed by the eleven amino acid hendecad of a right-handed coiled-coil peptide.⁷⁴⁻⁷⁵ As mentioned previously, left-handed coiled-coil peptides undergo helical interactions when hydrophobic residues are placed in the *a* and *d* positions.^{75, 86} Therefore, leucine residues were assigned at both the *a* and *d* positions of the heptad section for the **3FD-LL** peptide. Isoleucine and leucine residues were assigned at the *a* and *d* positions of the heptad section for **3FD-IL** peptide. On the other hand, the residues at positions *a*, *d*, *e*, and *h* of the hendecad are responsible for helical packing of right-handed coil-coils. Alanine residues are highly conserved at the *a* and *h* positions since their short side chains support tight packing between neighboring helices.⁸⁸ Meanwhile, the *d* and *e* positions do not directly face towards the neighboring helix and are solvent facing positions. Therefore, these positions grant more flexibility in the size and property of the side chains. In order to increase solubility and introduce electrostatic interactions, the charged residues glutamate and lysine were appointed to these two positions. These six positions (the *a* and *d* of the heptad and the *a*, *d*, *e*, and *h* positions of the hendecad) come together to form a single interacting face of the straight α -helix with the positions corresponding to the *a*, *d*, *h*, *k*, *l*, and *o* positions of the 18 amino acid repeats.

In order to generate a 3-fold design in which the straight α -helix would promote interactions with three neighboring peptides, the six amino acid face was staggered through the 18 amino acid repeat to create three identical faces that were angularly offset by -120° .⁸⁴ To extend the length of the peptide and interacting faces, a second 18 amino acid repeat was added to create the core 36 amino acids. The charged amino acids in the second 18 amino acid repeat were assigned to be complementary to the charged amino acids in the first repeat. Furthermore, the sequence repeat was shifted to remove the most hydrophobic amino acid from the edge of the peptide to

increase helix stability. Lastly, the *N*- and *C*- termini of the 36 amino acid core were capped with glutamate and lysine residues, respectively, to counteract the macro-dipole of the α -helix.⁸⁹ As a result, the 38 amino acid sequences were Ac-E(ALEKLA)₃(ALKELA)₃K-NH₂ for the **3FD-LL** peptide and Ac-E(ALEKIA)₃(ALKEIA)₃K-NH₂ for the **3FD-IL** peptide.^{85, 87} The sequences are also displayed along with their helical wheel representations in **Figure 1.1**.

Microwave assisted solid-phase peptide synthesis and subsequent purification using reverse-phase high performance liquid chromatography (HPLC) were used to prepare the **3FD-**

3FD-IL: Ac-E**ALEKIAALEKIAALEKIAALKEIAALKEIAALKEIAK**-NH₂
3FD-LL: Ac-E**ALEKLAALEKLAALEKLAALKELAALKELAALKELA**K-NH₂

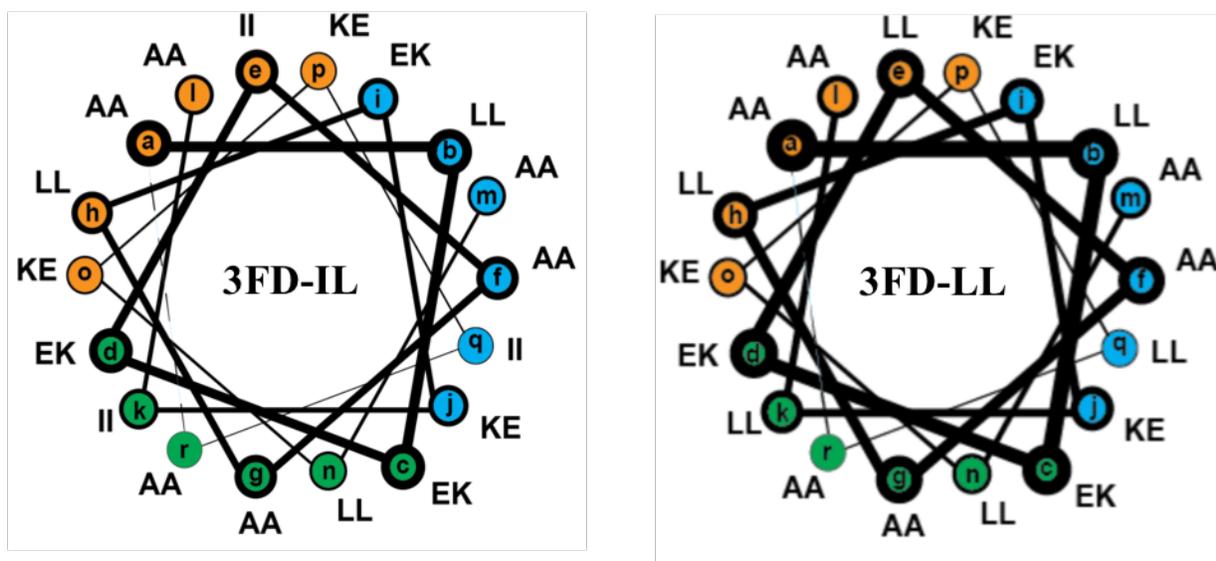


Figure 1.1 Amino acid sequences and helical wheels⁸⁷ for the 3FD-IL and 3FD-LL peptides indicating the register of residues within the 18 amino acid repeats to produce the three-fold alpha-helical design.

LL and **3FD-IL** peptides in high purity. Optimization of assembly conditions indicated that solubilization of a 10 mg/mL peptide concentration in 10 mM TAPS buffer (pH 8.5) followed by thermal annealing resulted in nanosheet formation for both peptides.^{85, 87}

Circular dichroism spectropolarimetry confirmed the presence of α -helical secondary structure of the **3FD-LL** and **3FD-IL** peptide assemblies. (**Figure 1.2 a-b**). CD melting curves were collected for the thermally annealed samples (**Figure 1.2 c-d**) and first derivatives of the CD melting curves displayed melting transitions (T_m) at 84 °C and 90 °C for the **3FD-IL** and **3FD-LL** peptide assemblies, respectively.^{85, 87} The higher T_m observed for the **3FD-LL** assemblies compared to the thermally annealed solutions of **3FD-IL** indicated that **3FD-LL** assemblies displayed improved thermodynamic stability. This could be in part due to the tighter packing allowed by the substitution of isoleucine to leucine.

The assemblies formed by the **3FD-IL** and **3FD-LL** peptides were analyzed using transmission electron microscopy (TEM). TEM analysis showed the presence of two-dimensional nanosheets for the nascent and thermally annealed solutions of both peptides (**Figure 1.3**).^{85, 87} The nanosheets exhibited sizes of hundreds of nanometers. Many sheets displayed sharp edges, but there was evidence of broken sheets, which may have been a result of the preparation process for TEM. This fragmenting was also present for the **3FD-IL** sheets indicating that there was room for improvement in the durability of these two-dimensional assemblies.

The average thickness of the sheet assemblies was determined using tapping mode atomic force microscopy (AFM). The **3FD-IL** nanosheets displayed a thickness around 6.2 nm, while the **3FD-LL** nanosheets exhibited a thickness of approximately 5.85 nm (**Figure 1.4**).^{85, 87} Assuming all 38 amino acids in each peptide adopted an α -helical conformation, the theoretical length of both

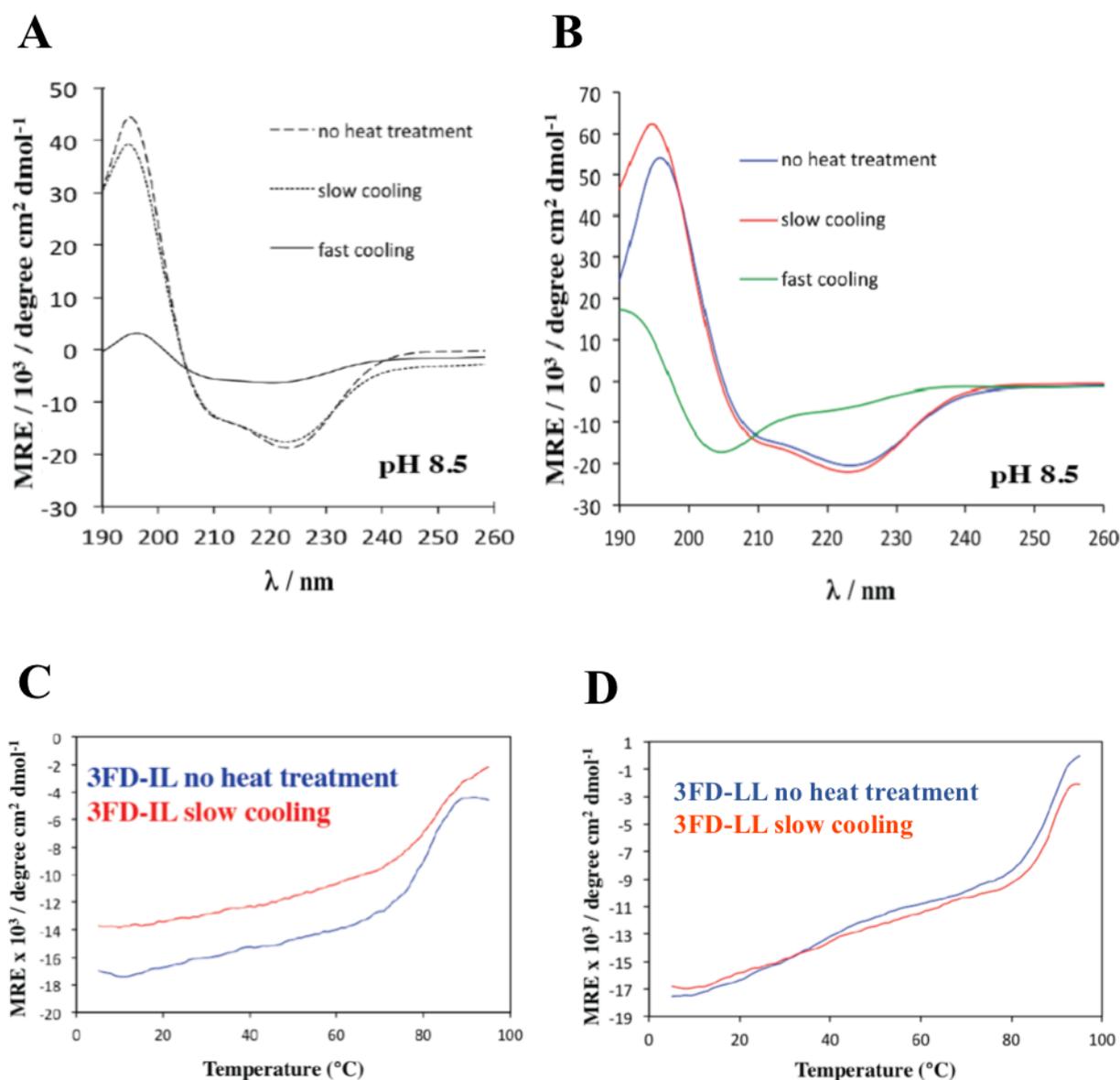


Figure 1.2 Circular dichroism spectroscopy of nascent (no heat treatment) and thermally annealed (slow cooling) 3FD-IL and 3FD-LL peptides assembled in 10 mM TAPS buffer (pH 8.5). Standard CD spectra for the (A) 3FD-IL and (B) 3FD-LL peptides. CD melting curves of the (C) 3FD-IL and (D) 3FD-LL peptides by monitoring the molar residue ellipticity at 222 nm from 5 °C to 95 °C.⁸⁷

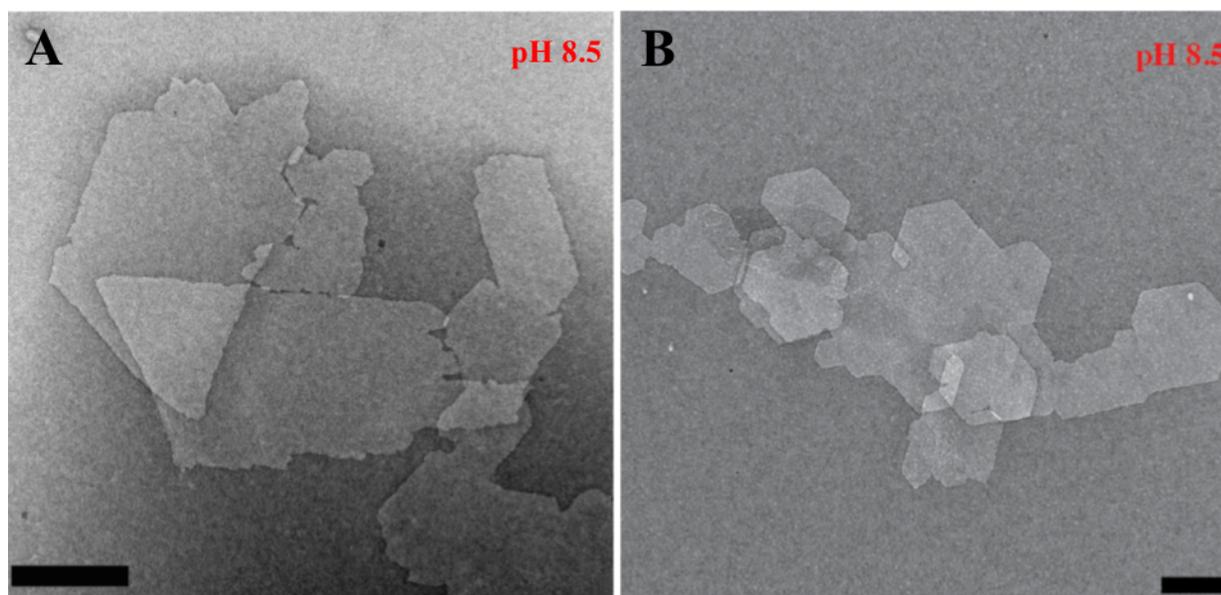


Figure 1.3 Transmission electron microscopy images of thermally annealed (A) 3FD-IL and (B) 3FD-LL peptides assembled in 10 mM TAPS buffer (pH 8.5). Scale bars = 200 nm.⁸⁷

peptides would be 5.7 nm owing to the 1.5 Å rise per residue. This supports the hypothesis that the helices were assembling perpendicularly with respect to the face of the sheet and that the sheet thickness was determined by the length of the α -helical peptide. Similar packing arrangements were documented for protein-based nanosheets derived from collagen-mimetic peptides where the triple helix oriented perpendicular to the surface of the nanosheets.⁵²⁻⁵⁴

In addition to verification of nanosheet formation by the designed straight α -helical peptides, extensive analysis of the **3FD-IL** peptide by small-angle x-ray scattering (SAXS) powder diffraction, electron cryomicroscopy (Cryo-EM), and scanning transmission electron microscopy (STEM) indicated arrangement of the helices based on hexagonal packing into a honeycomb lattice

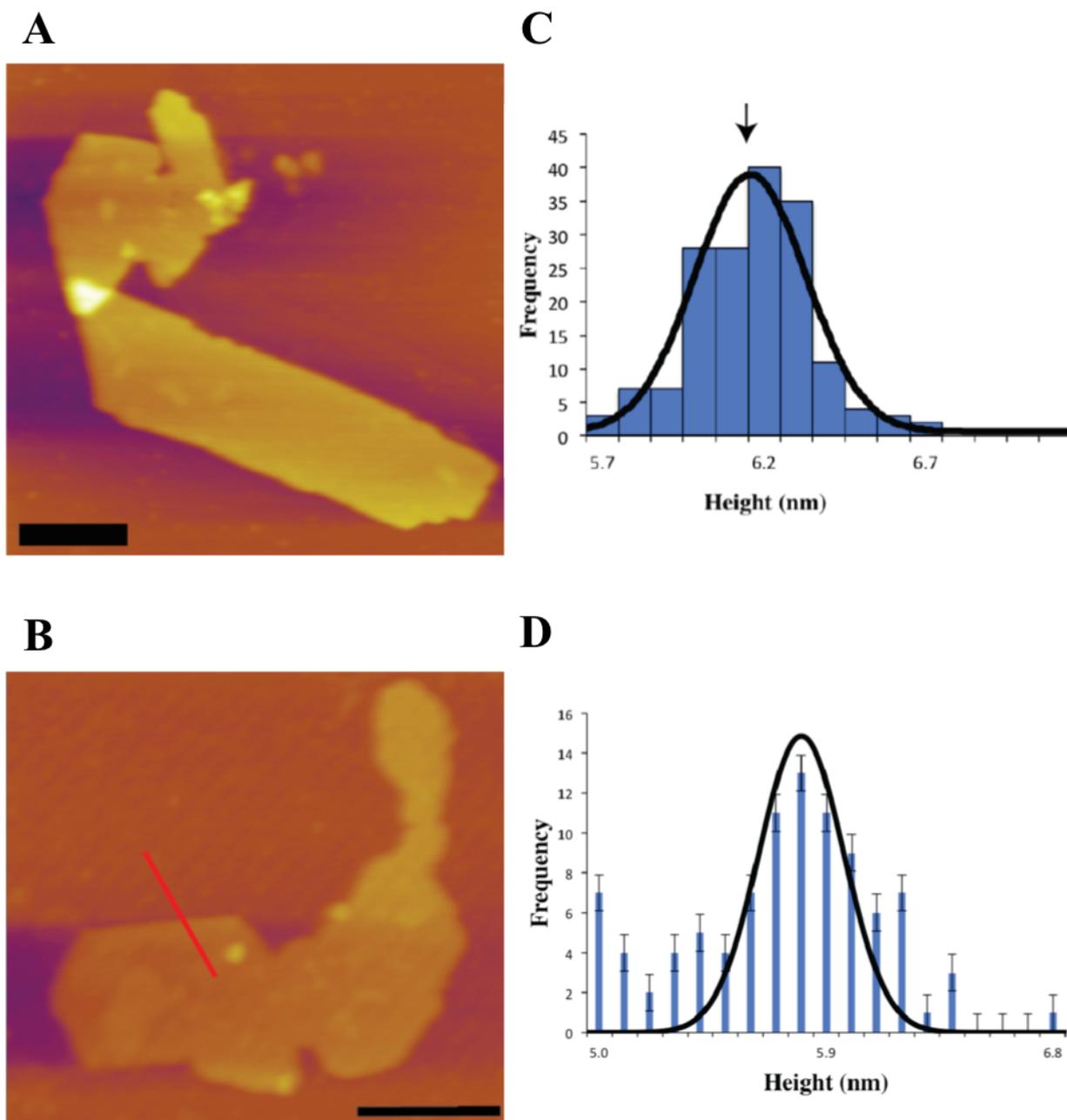


Figure 1.4 Atomic force microscopy of the 3FD-IL and 3FD-LL peptides assembled in 10 mM TAPS buffer (pH 8.5). AFM images of thermally annealed assemblies of the (A) 3FD-IL and (B) 3FD-LL peptides. Histograms of height measurements of thermally annealed assemblies of the (C) 3FD-IL and (D) 3FD-LL peptides. Scale bars = 200 nm.⁸⁷

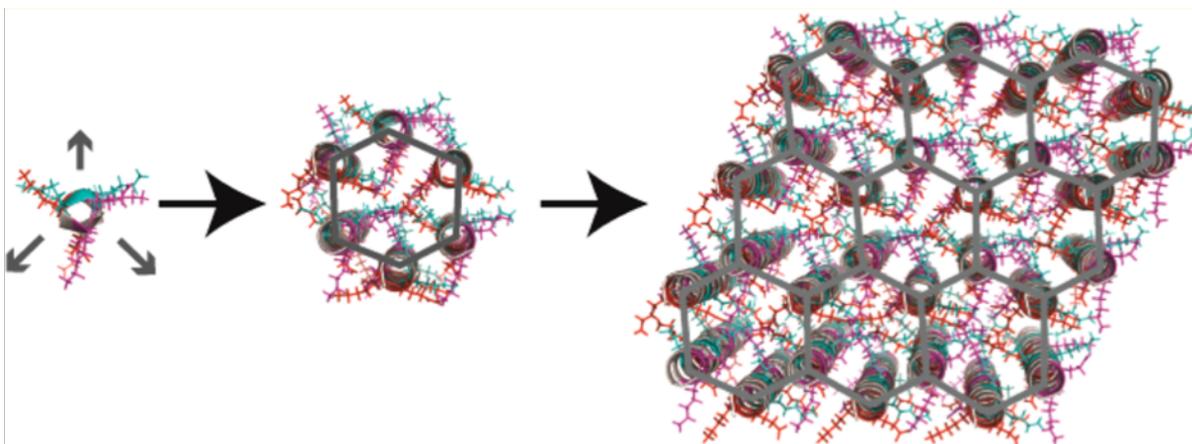


Figure 1.5 Proposed packing of the 3FD-IL helices depicting the 3-fold symmetry within the hexagonal honeycomb model during nanosheet formation.⁸⁵

with nanoscale channels penetrating the two-dimensional sheets, as show in the schematic in **Figure 1.5**.^{85, 87} The porous nature of the **3FD-IL** and **3FD-LL** nanosheets further expands the potential applications of biological nanosheets towards charge transfer or molecular encapsulation.

1.4 Genetic Engineering of Biomaterials

The **3FD-IL** peptide was the first published α -helical peptide to assemble into a two-dimensional nanosheet. The **3FD-LL** peptide containing leucine residues at all hydrophobic interfaces developed into similar nanosheets with improved thermodynamic stability. The nanosheets formed by these peptides displayed a high degree of internal order with nanopores permeating the nanosheets. The presence of these nanopores grant the two-dimensional materials various potential applications involving the encapsulation or transfer of ions and small molecules. The role and magnitude of these applications could be dependent on the depth or length of the

pore, which would be predetermined by the thickness of the nanosheet. Since the α -helices arrange perpendicular to the surface of the sheet, increasing the length of the peptide would likely result in an increase in the thickness of the nanosheet and lead to improved functionality of the nanomaterial.

Furthermore, fragmentation was frequently visualized in TEM images of the **3FD-IL** and **3FD-LL** nanosheets indicating that the nanosheets were possibly breaking during the TEM preparation process by mechanical shearing. If the nanosheets were to be used in functional devices, the robustness of the nanosheets would have to undergo improvements. The nanosheets are held together by the hydrophobic and electrostatic interactions between neighboring helices. It is possible that increasing the number of interactions between helices could increase the durability and thermodynamic stability of the two-dimensional nanosheets. Increasing the number of local interactions between helices would not be possible since all amino acids on each face are already participating in non-covalent hydrophobic and electrostatic interactions. However, if the length of the α -helical units were increased to facilitate the formation of longer nanopores, the increased number of interactions along the length of the α -helices would potentially result in improved thermodynamic and mechanical stability.

Up to this point, the **3FD-IL** and **3FD-LL** peptides were created by solid-phase peptide synthesis, which yielded a low quantity of peptide for a high cost. Peptide output was typically between 10-30 milligrams per synthesis and the measured quantity dropped by at least 50% after purification by reverse-phase high performance liquid chromatography to remove all peptide fragments. The low peptide quantity was sufficient for preliminary biophysical analysis of the peptides, but it is not feasible for large-scale application purposes. In addition, the limitations of peptide synthesis restrict the creation of polypeptide chains longer than 70 amino acids in length

as the yield decreases exponentially with length. Although chemical ligation could be used to couple two **3FD-IL** or **3FD-LL** peptides, the yield limitations would not be eradicated. Therefore, chemical peptide synthesis would be incapable of increasing the length of the **3FD-IL** or **3FD-LL** peptides while also improving yield for application purposes.

Genetic engineering and protein expression reveal immeasurable opportunities that are not conceivable by solid-phase peptide synthesis. A thorough understanding of molecular cloning techniques permit the construction of novel recombinant proteins displaying increased complexity while maintaining sequence control. Recombinant proteins can be specifically designed to support assembly and functionalization of a material. Biosynthesis and large-scale protein expression using *E. coli* as a host organism has been extensively applied for the production of peptides and proteins of various lengths. Recombinant protein engineering supports materials applications by significantly increasing protein yield over that which can be achieved by traditional peptide synthesis. Furthermore, long protein polymers can be produced on a large-scale with low batch-to-batch variation unlike in solid-phase peptide synthesis. Once expression procedures have been optimized, it is extremely affordable to produce a large quantity of protein as needed. Therefore, biological synthesis is an attractive avenue for the high-level production of extended **3FD-IL** and **3FD-LL** proteins.

1.5 Conclusions

The **3FD-IL** and **3FD-LL** peptides adopted straight α -helical conformations with 3-fold screw symmetry. These α -helices assembled to form two-dimensional nanosheets where helical units arranged perpendicular to the nanosheet surface with nanopores permeating the sheet. The nanosheets have a thickness that is defined by the length of a single α -helical peptide unit. In comparison to **3FD-IL**, the change in the hydrophobic interfaces to all leucine residues in **3FD-LL** produced a more robust, thermodynamically stable material. The ability of the **3FD-LL** peptides to develop into relatively stable two-dimensional structures lays the groundwork for the development of more advanced materials. However, fragmenting visualized through TEM analysis indicated that the mechanical stability of the assemblies could be improved.

Recombinant protein engineering at the molecular (DNA) level is a useful technique for generating large amounts of protein polymers with versatile control over sequence specificity. In order to expand the utility of these materials for potential applications, a system could be developed through genetic engineering to increase the length of the **3FD-LL** α -helical unit to acquire thicker nanosheets. This would likely improve the thermodynamic stability of the nanosheets and also increase the function of the sheets with regards to encapsulation and transportation of small molecules and ions. Additionally, biosynthesis through high-level expression would support the low-cost production of copious amounts of protein, which is necessary for materials applications.

1.6 References

1. Ariga, K.; Li, J.; Fei, J.; Ji, Q.; Hill, J. P., Nanoarchitectonics for Dynamic Functional Materials from Atomic-/Molecular-Level Manipulation to Macroscopic Action. *Advanced materials (Deerfield Beach, Fla.)* **2016**, *28* (6), 1251-86.
2. Govindaraju, T.; Avinash, M. B., Two-dimensional nanoarchitectonics: organic and hybrid materials. *Nanoscale* **2012**, *4* (20), 6102-17.
3. Avinash, M. B.; Govindaraju, T., Nanoarchitectonics of biomolecular assemblies for functional applications. *Nanoscale* **2014**, *6* (22), 13348-13369.
4. Wang, Z.-G.; Ding, B., DNA-Based Self-Assembly for Functional Nanomaterials. *Advanced Materials* **2013**, *25* (28), 3905-3914.
5. Zhang, X.; Xie, Y., Recent advances in free-standing two-dimensional crystals with atomic thickness: design, assembly and transfer strategies. *Chemical Society Reviews* **2013**, *42* (21), 8187-8199.
6. Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, *440*, 297.
7. Lin, C.; Liu, Y.; Rinker, S.; Yan, H., DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures. *ChemPhysChem* **2006**, *7* (8), 1641-1647.
8. De Santis, E.; Ryadnov, M. G., Peptide self-assembly for nanomaterials: the old new kid on the block. *Chemical Society Reviews* **2015**, *44* (22), 8288-8300.
9. Desai, M. S.; Lee, S.-W., Protein-based functional nanomaterial design for bioengineering applications. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **2015**, *7* (1), 69-97.

10. Robertson, E. J.; Battigelli, A.; Proulx, C.; Mannige, R. V.; Haxton, T. K.; Yun, L.; Whitelam, S.; Zuckermann, R. N., Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets. *Accounts of Chemical Research* **2016**, *49* (3), 379-389.
11. Egelman, E. H.; Francis, N.; DeRosier, D. J., F-actin is a helix with a random variable twist. *Nature* **1982**, *298* (5870), 131-5.
12. Galkin, V. E.; Orlova, A.; Vos, M. R.; Schroder, G. F.; Egelman, E. H., Near-atomic resolution for one state of F-actin. *Structure (London, England : 1993)* **2015**, *23* (1), 173-182.
13. Prockop, D. J.; Fertala, A., The collagen fibril: the almost crystalline structure. *J Struct Biol* **1998**, *122* (1-2), 111-8.
14. Aggeli, A.; Bell, M.; Boden, N.; Keen, J. N.; Knowles, P. F.; McLeish, T. C.; Pitkeathly, M.; Radford, S. E., Responsive gels formed by the spontaneous self-assembly of peptides into polymeric beta-sheet tapes. *Nature* **1997**, *386* (6622), 259-62.
15. Aggeli, A.; Bell, M.; Carrick, L. M.; Fishwick, C. W.; Harding, R.; Mawer, P. J.; Radford, S. E.; Strong, A. E.; Boden, N., pH as a trigger of peptide beta-sheet self-assembly and reversible switching between nematic and isotropic phases. *J Am Chem Soc* **2003**, *125* (32), 9619-28.
16. Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C.; Semenov, A. N.; Boden, N., Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta-sheet tapes, ribbons, fibrils, and fibers. *Proc Natl Acad Sci U S A* **2001**, *98* (21), 11857-62.
17. Bowerman, C. J.; Liyanage, W.; Federation, A. J.; Nilsson, B. L., Tuning beta-sheet peptide self-assembly and hydrogelation behavior by modification of sequence hydrophobicity and aromaticity. *Biomacromolecules* **2011**, *12* (7), 2735-45.

18. 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* **2007**, *129* (41), 12468-72.
19. Janek, K.; Behlke, J.; Zipper, J.; Fabian, H.; Georgalis, Y.; Beyermann, M.; Bienert, M.; Krause, E., Water-soluble beta-sheet models which self-assemble into fibrillar structures. *Biochemistry* **1999**, *38* (26), 8246-52.
20. Marini, D. M.; Hwang, W.; Lauffenburger, D. A.; Zhang, S.; Kamm, R. D., Left-Handed Helical Ribbon Intermediates in the Self-Assembly of a β -Sheet Peptide. *Nano Letters* **2002**, *2* (4), 295-299.
21. Matsumura, S.; Uemura, S.; Mihara, H., Fabrication of nanofibers with uniform morphology by self-assembly of designed peptides. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2004**, *10* (11), 2789-94.
22. Swanekamp, R. J.; DiMaio, J. T.; Bowerman, C. J.; Nilsson, B. L., Coassembly of enantiomeric amphipathic peptides into amyloid-inspired rippled beta-sheet fibrils. *J Am Chem Soc* **2012**, *134* (12), 5556-9.
23. Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A., Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc Natl Acad Sci U S A* **1993**, *90* (8), 3334-8.
24. Choo, D. W.; Schneider, J. P.; Graciani, N. R.; Kelly, J. W., Nucleated Antiparallel β -Sheet That Folds and Undergoes Self-Assembly: A Template Promoted Folding Strategy toward Controlled Molecular Architectures. *Macromolecules* **1996**, *29* (1), 355-366.
25. Nagarkar, R. P.; Hule, R. A.; Pochan, D. J.; Schneider, J. P., De novo design of strand-swapped beta-hairpin hydrogels. *J Am Chem Soc* **2008**, *130* (13), 4466-74.

26. Pochan, D. J.; Schneider, J. P.; Kretsinger, J.; Ozbas, B.; Rajagopal, K.; Haines, L., Thermally reversible hydrogels via intramolecular folding and consequent self-assembly of a de novo designed peptide. *J Am Chem Soc* **2003**, *125* (39), 11802-3.
27. Schneider, J. P.; Pochan, D. J.; Ozbas, B.; Rajagopal, K.; Pakstis, L.; Kretsinger, J., Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *J Am Chem Soc* **2002**, *124* (50), 15030-7.
28. Anzini, P.; Xu, C.; Hughes, S.; Magnotti, E.; Jiang, T.; Hemmingsen, L.; Demeler, B.; Conticello, V. P., Controlling self-assembly of a peptide-based material via metal-ion induced registry shift. *J Am Chem Soc* **2013**, *135* (28), 10278-81.
29. Dong, H.; Paramonov, S. E.; Hartgerink, J. D., Self-assembly of alpha-helical coiled coil nanofibers. *J Am Chem Soc* **2008**, *130* (41), 13691-5.
30. Dublin, S. N.; Conticello, V. P., Design of a selective metal ion switch for self-assembly of peptide-based fibrils. *J Am Chem Soc* **2008**, *130* (1), 49-51.
31. Kojima, S.; Kuriki, Y.; Yoshida, T.; Yazaki, K.; Miura, K.-i., Fibril Formation by an Amphipathic α-Helix-Forming Polypeptide Produced by Gene Engineering. *Proceedings of the Japan Academy, Series B* **1997**, *73* (1), 7-11.
32. Ogihara, N. L.; Ghirlanda, G.; Bryson, J. W.; Gingery, M.; DeGrado, W. F.; Eisenberg, D., Design of three-dimensional domain-swapped dimers and fibrous oligomers. *Proc Natl Acad Sci U S A* **2001**, *98* (4), 1404-9.
33. Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry* **2000**, *39* (30), 8728-34.

34. Papapostolou, D.; Smith, A. M.; Atkins, E. D.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N., Engineering nanoscale order into a designed protein fiber. *Proc Natl Acad Sci U S A* **2007**, *104* (26), 10853-8.
35. Potekhin, S. A.; Melnik, T. N.; Popov, V.; Lanina, N. F.; Vazina, A. A.; Rigler, P.; Verdini, A. S.; Corradin, G.; Kajava, A. V., De novo design of fibrils made of short alpha-helical coiled coil peptides. *Chem Biol* **2001**, *8* (11), 1025-32.
36. Ryadnov, M. G.; Woolfson, D. N., Introducing branches into a self-assembling peptide fiber. *Angew Chem Int Ed Engl* **2003**, *42* (26), 3021-3.
37. Ryadnov, M. G.; Woolfson, D. N., Engineering the morphology of a self-assembling protein fibre. *Nature materials* **2003**, *2* (5), 329-32.
38. Wagner, D. E.; Phillips, C. L.; Ali, W. M.; Nybakken, G. E.; Crawford, E. D.; Schwab, A. D.; Smith, W. F.; Fairman, R., Toward the development of peptide nanofilaments and nanoropes as smart materials. *Proc Natl Acad Sci U S A* **2005**, *102* (36), 12656-61.
39. Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P., Rational design of a nanoscale helical scaffold derived from self-assembly of a dimeric coiled coil motif. *Tetrahedron* **2004**, *60* (34), 7237-7246.
40. Zimenkov, Y.; Dublin, S. N.; Ni, R.; Tu, R. S.; Breedveld, V.; Apkarian, R. P.; Conticello, V. P., Rational design of a reversible pH-responsive switch for peptide self-assembly. *J Am Chem Soc* **2006**, *128* (21), 6770-1.
41. Cejas, M. A.; Kinney, W. A.; Chen, C.; Leo, G. C.; Tounge, B. A.; Vinter, J. G.; Joshi, P. P.; Maryanoff, B. E., Collagen-related peptides: self-assembly of short, single strands into a functional biomaterial of micrometer scale. *J Am Chem Soc* **2007**, *129* (8), 2202-3.

42. Cejas, M. A.; Kinney, W. A.; Chen, C.; Vinter, J. G.; Almond, H. R., Jr.; Balss, K. M.; Maryanoff, C. A.; Schmidt, U.; Breslav, M.; Mahan, A.; Lacy, E.; Maryanoff, B. E., Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions. *Proc Natl Acad Sci U S A* **2008**, *105* (25), 8513-8.
43. Kar, K.; Ibrar, S.; Nanda, V.; Getz, T. M.; Kunapuli, S. P.; Brodsky, B., Aromatic interactions promote self-association of collagen triple-helical peptides to higher-order structures. *Biochemistry* **2009**, *48* (33), 7959-68.
44. Koide, T.; Homma, D. L.; Asada, S.; Kitagawa, K., Self-complementary peptides for the formation of collagen-like triple helical supramolecules. *Bioorganic & medicinal chemistry letters* **2005**, *15* (23), 5230-3.
45. O'Leary, L. E.; Fallas, J. A.; Hartgerink, J. D., Positive and negative design leads to compositional control in AAB collagen heterotrimers. *J Am Chem Soc* **2011**, *133* (14), 5432-43.
46. Przybyla, D. E.; Rubert Pérez, 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.* **2013**, *135* (9), 3418-3422.
47. Rele, S.; Song, Y.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L., D-Periodic Collagen-Mimetic Microfibers. *J. Am. Chem. Soc.* **2007**, *129* (47), 14780-14787.
48. Xu, F.; Li, J.; Jain, V.; Tu, R. S.; Huang, Q.; Nanda, V., Compositional control of higher order assembly using synthetic collagen peptides. *J Am Chem Soc* **2012**, *134* (1), 47-50.
49. Kotch, F. W.; Raines, R. T., Self-assembly of synthetic collagen triple helices. *Proc Natl Acad Sci U S A* **2006**, *103* (9), 3028-33.
50. Yamazaki, C. M.; Asada, S.; Kitagawa, K.; Koide, T., Artificial collagen gels via self-assembly of de novo designed peptides. *Biopolymers* **2008**, *90* (6), 816-23.

51. Magnotti, E.; Conticello, V., Two-Dimensional Peptide and Protein Assemblies. In *Protein-based Engineered Nanostructures*, Cortajarena, A. L.; Grove, T. Z., Eds. Springer International Publishing: Cham, 2016; pp 29-60.
52. 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.* **2015**, *137* (24), 7793-7802.
53. Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P., Structurally Defined Nanoscale Sheets from Self-Assembly of Collagen-Mimetic Peptides. *J. Am. Chem. Soc.* **2014**, *136* (11), 4300-4308.
54. Jiang, T.; Xu, C.; Zuo, X.; Conticello, V. P., Structurally Homogeneous Nanosheets from Self-Assembly of a Collagen-Mimetic Peptide. *Angewandte Chemie International Edition* **2014**, *53* (32), 8367-8371.
55. Parmar, A. S.; James, J. K.; Grisham, D. R.; Pike, D. H.; Nanda, V., Dissecting Electrostatic Contributions to Folding and Self-Assembly Using Designed Multicomponent Peptide Systems. *J. Am. Chem. Soc.* **2016**, *138* (13), 4362-4367.
56. Xu, F.; Khan, I. J.; McGuinness, K.; Parmar, A. S.; Silva, T.; Murthy, N. S.; Nanda, V., Self-Assembly of Left- and Right-Handed Molecular Screws. *J. Am. Chem. Soc.* **2013**, *135* (50), 18762-18765.
57. Dai, B.; Li, D.; Xi, W.; Luo, F.; Zhang, X.; Zou, M.; Cao, M.; Hu, J.; Wang, W.; Wei, G.; Zhang, Y.; Liu, C., Tunable assembly of amyloid-forming peptides into nanosheets as a retrovirus carrier. *Proc Natl Acad Sci U S A* **2015**, *112* (10), 2996-3001.
58. Hamley, I. W.; Dehsorkhi, A.; Castelletto, V., Self-assembled arginine-coated peptide nanosheets in water. *Chemical communications (Cambridge, England)* **2013**, *49* (18), 1850-2.

59. Robertson, E. J.; Oliver, G. K.; Qian, M.; Proulx, C.; Zuckermann, R. N.; Richmond, G. L., Assembly and molecular order of two-dimensional peptoid nanosheets through the oil-water interface. *Proc Natl Acad Sci U S A* **2014**, *111* (37), 13284-9.
60. Kudirka, R.; Tran, H.; Sanii, B.; Nam, K. T.; Choi, P. H.; Venkateswaran, N.; Chen, R.; Whitelam, S.; Zuckermann, R. N., Folding of a single-chain, information-rich polypeptoid sequence into a highly ordered nanosheet. *Biopolymers* **2011**, *96* (5), 586-95.
61. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B. C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nature materials* **2010**, *9* (5), 454-60.
62. Sanii, B.; Kudirka, R.; Cho, A.; Venkateswaran, N.; Olivier, G. K.; Olson, A. M.; Tran, H.; Harada, R. M.; Tan, L.; Zuckermann, R. N., Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets. *J. Am. Chem. Soc.* **2011**, *133* (51), 20808-20815.
63. Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature chemistry* **2012**, *4* (5), 375-82.
64. Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, *533* (7603), 369-73.
65. Brodin, J. D.; Carr, J. R.; Sontz, P. A.; Tezcan, F. A., Exceptionally stable, redox-active supramolecular protein assemblies with emergent properties. *Proc Natl Acad Sci U S A* **2014**, *111* (8), 2897-902.

66. Matthaei, J. F.; DiMaio, F.; Richards, J. J.; Pozzo, L. D.; Baker, D.; Baneyx, F., Designing Two-Dimensional Protein Arrays through Fusion of Multimers and Interface Mutations. *Nano Letters* **2015**, *15* (8), 5235-5239.
67. Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M., Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science (New York, N.Y.)* **1994**, *266* (5182), 75-81.
68. Ramshaw, J. A. M.; Shah, N. K.; Brodsky, B., Gly-X-Y Tripeptide Frequencies in Collagen: A Context for Host-Guest Triple-Helical Peptides. *Journal of Structural Biology* **1998**, *122* (1), 86-91.
69. Przybyla, D. E.; Chmielewski, J., Metal-Triggered Collagen Peptide Disk Formation. *J. Am. Chem. Soc.* **2010**, *132* (23), 7866-7867.
70. McGuinness, K.; Khan, I. J.; Nanda, V., Morphological Diversity and Polymorphism of Self-Assembling Collagen Peptides Controlled by Length of Hydrophobic Domains. *ACS Nano* **2014**, *8* (12), 12514-12523.
71. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B.-C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nature materials* **2010**, *9*, 454.
72. Yu, S. M.; Conticello, V. P.; Zhang, G.; Kayser, C.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A., Smectic ordering in solutions and films of a rod-like polymer owing to monodispersity of chain length. *Nature* **1997**, *389* (6647), 167-70.
73. Yu, S. M.; Soto, C. M.; Tirrell, D. A., Nanometer-Scale Smectic Ordering of Genetically Engineered Rodlike Polymers: Synthesis and Characterization of Monodisperse Derivatives of Poly(γ -benzyl α ,l-glutamate). *J. Am. Chem. Soc.* **2000**, *122* (28), 6552-6559.

74. Woolfson, D. N.; Bartlett, G. J.; Bruning, M.; Thomson, A. R., New currency for old rope: from coiled-coil assemblies to alpha-helical barrels. *Current opinion in structural biology* **2012**, *22* (4), 432-41.
75. Lupas, A. N.; Gruber, M., The structure of alpha-helical coiled coils. *Advances in protein chemistry* **2005**, *70*, 37-78.
76. Patterson, W. R.; Anderson, D. H.; DeGrado, W. F.; Cascio, D.; Eisenberg, D., Centrosymmetric bilayers in the 0.75 Å resolution structure of a designed alpha-helical peptide, D,L-Alpha-1. *Protein science : a publication of the Protein Society* **1999**, *8* (7), 1410-22.
77. Karle, I. L., Folding, aggregation and molecular recognition in peptides. *Acta crystallographica. Section B, Structural science* **1992**, *48* (Pt 4), 341-56.
78. Bowie, J. U., Helix packing angle preferences. *Nature structural biology* **1997**, *4* (11), 915-7.
79. Chothia, C.; Levitt, M.; Richardson, D., Helix to helix packing in proteins. *J Mol Biol* **1981**, *145* (1), 215-50.
80. Walther, D.; Springer, C.; Cohen, F. E., Helix-helix packing angle preferences for finite helix axes. *Proteins* **1998**, *33* (4), 457-9.
81. Vasudev, P. G.; Shamala, N.; Balaram, P., Nucleation, Growth, and Form in Crystals of Peptide Helices. *The Journal of Physical Chemistry B* **2008**, *112* (4), 1308-1314.
82. Prive, G. G.; Anderson, D. H.; Wesson, L.; Cascio, D.; Eisenberg, D., Packed protein bilayers in the 0.90 Å resolution structure of a designed alpha helical bundle. *Protein science : a publication of the Protein Society* **1999**, *8* (7), 1400-9.

83. Huang, P. S.; Oberdorfer, G.; Xu, C.; Pei, X. Y.; Nannenga, B. L.; Rogers, J. M.; DiMaio, F.; Gonen, T.; Luisi, B.; Baker, D., High thermodynamic stability of parametrically designed helical bundles. *Science (New York, N.Y.)* **2014**, *346* (6208), 481-485.
84. Pauling, L.; Corey, R. B.; Branson, H. R., The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci U S A* **1951**, *37* (4), 205-11.
85. Magnotti, E. L.; Hughes, S. A.; Dillard, R. S.; Wang, S.; Hough, L.; Karumbamkandathil, A.; Lian, T.; Wall, J. S.; Zuo, X.; Wright, E. R.; Conticello, V. P., Self-Assembly of an α -Helical Peptide into a Crystalline Two-Dimensional Nanoporous Framework. *J. Am. Chem. Soc.* **2016**, *138* (50), 16274-16282.
86. Fletcher, J. M.; Boyle, A. L.; Bruning, M.; Bartlett, G. J.; Vincent, T. L.; Zaccai, N. R.; Armstrong, C. T.; Bromley, E. H. C.; Booth, P. J.; Brady, R. L.; Thomson, A. R.; Woolfson, D. N., A Basis Set of de Novo Coiled-Coil Peptide Oligomers for Rational Protein Design and Synthetic Biology. *ACS Synthetic Biology* **2012**, *1* (6), 240-250.
87. Magnotti, E. L. Self-assembly of porous α -helical nanosheets. Emory University, 2016.
88. Lupas, A., Coiled coils: new structures and new functions. *Trends in biochemical sciences* **1996**, *21* (10), 375-82.
89. Hol, W. G., The role of the alpha-helix dipole in protein function and structure. *Progress in biophysics and molecular biology* **1985**, *45* (3), 149-95.

Chapter 2

Development of a System for the Generation of 3FD-LL Concatemers

2.1 Introduction

Although the **3FD-LL** peptide formed a straight α -helix that assembled into defined two-dimensional porous nanosheets, the awareness of weaknesses represents an opportunity for improvement of the system prior to application of the materials. Fragmentation of the nanosheets was apparent during biophysical analysis, which could potentially be mitigated through an increase in sheet thickness. In addition, an increase in sheet thickness would result in deeper pores for encapsulation of small molecules. Lastly, an increase in sheet thickness may be necessary for applications in nanoarchitectonics through incorporation of the nanosheets into nanodevices.

Since the **3FD-LL** α -helical units arrange perpendicular to the sheet surface, the sheet thickness is directly related to the length of the peptide.¹⁻² Due to limitations in solid-phase peptide synthesis, longer peptides cannot be created without a detrimental decrease in yield. Additionally, higher peptide yields are imperative for the application of materials, and the yields previously measured from the chemical synthesis of the **3FD-LL** peptide was only sufficient for initial biophysical analysis. Therefore, other avenues would have to be investigated for the generation of extended **3FD-LL** peptides.

Genetic engineering can be effectively utilized to expand the potential of peptide research in the direction of protein application and functionalization. In order to approach a more applicable design for the **3FD-LL** peptide, a system was designed using molecular cloning techniques to extend the α -helical unit of the **3FD-LL** sequence using *E. coli* as a host. This would lead to the creation of thicker nanosheets with improved thermodynamic and mechanical stability by increasing the number of interactions between neighboring helices. The system was based on the

production of concatemers, or long continuous molecules that contain multiple copies of the same sequence linked in a series, of the **3FD-LL** sequence.

Genetic engineering techniques can allow for the biological synthesis of target proteins through the control of the sequence at the DNA level and subsequently the amino acid sequence. A general strategy for the biosynthesis of proteins includes synthesis of a DNA coding sequence, insertion of the gene into an expression vector, and expression of the protein using a bacterial host. In order to create concatemer sequences, a DNA monomer must be formed with non-palindromic, cohesive ends that can link in a head-to-tail fashion. The DNA concatemers must then be inserted into an expression vector that can be transformed into an *E. coli* expression host where expression of the target protein can be induced. A scheme representing the biosynthesis of repetitive peptides via the concatemerization of DNA cassettes is shown in **Figure 2.1**.

This technique was utilized in the design for the concatemerization of the **3FD-LL** sequence. A concatemer family was successfully formed through genetic engineering of the 3FD-LL sequence to form the 3FD-LL monomer, dimer, trimer, tetramer, and pentamer. Large-scale protein expression produced all five concatemers at high yields that were not previously achieved by solid-phase peptide synthesis. Finally, the *N*-terminal His-tags were effectively utilized for the purification of the 3FD-LL concatemers via affinity chromatography with HisPurTM cobalt resin.

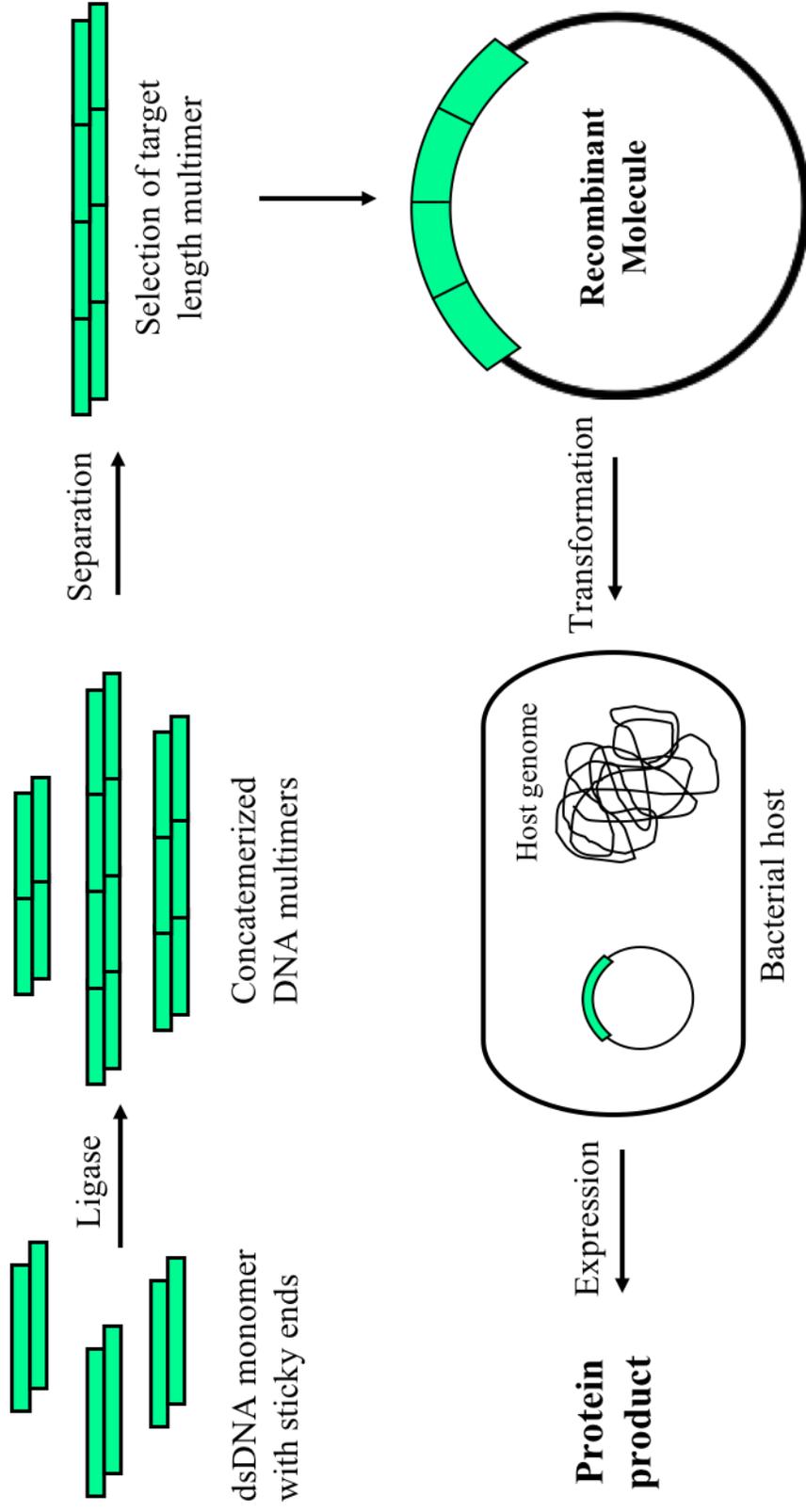


Figure 2.1 Schematic for the synthesis of repetitive polypeptides via concatenate DNA cassettes.

2.2 Results and Discussion

2.2.1 Genetic Engineering of 3FD-LL Concatemer Sequences

Design of Adapter and Insertion Genes

Initially, the **3FD-LL** peptide sequence was converted into a gene sequence that was codon optimized for expression using *E. coli* as a host. This gene sequence was used for the design of adapter and insert genes that would support the concatemerization of the **3FD-LL** sequence. The purpose of the 3FD-LL adapter gene was to enable the insertion of 3FD-LL helix monomers to produce the 3FD-LL concatemers. The 3FD-LL helix insert gene was designed for the isolation of a 3FD-LL helix monomer sequence for incorporation into the 3FD-LL adapter helix insertion site.

The simplest approach was to design the 3FD-LL adapter gene to encode for one **3FD-LL monomer** on the *C*-terminus with a helix insertion site preceding it. The helix insertion site of the adapter was designed to contain recognition sites for restriction enzymes that produce non-palindromic cohesive ends to accommodate the insertion of a 3FD-LL helix insert sequence. Type IIS restriction enzymes that cut at specific locations outside of their recognition sequence were considered, so the restriction sites, and consequent additional unwanted amino acids, would be completely removed from the sequence after digestion. One common restriction enzyme that met these criteria was *Bsa* I, which recognizes the asymmetric DNA sequence 5'-GGTCTC-3'. Placing two *Bsa* I recognition sites in the helix insertion site, with their cut sites located outside of their recognition sites, would allow the 3FD-LL helix sequence to be inserted into the adapter sequence. The *Bsa* I cut sites were designed to produce the overhangs for 5'-AGCC-3'. The nature of the design would also allow for the removal of the helix insertion site and the subsequent creation of the gene sequence for the **3FD-LL monomer** if no insert was introduced into the insertion site.

The 3FD-LL adapter gene was also designed to allow for transfer into another plasmid for expression. An adapter sequence was previously designed for the insertion of elastin sequences into **pIL5**,³ a modified **pQE-80L** plasmid. The elastin adapter could be cleaved from the **pIL5** vector through double digestion with the restriction endonucleases *EcoR* I and *HinD* III. By placing these two restriction sites on the terminal ends of the designed 3FD-LL adapter, the new adapter could be cloned in its entirety into the **pIL5** expression vector in place of the previous elastin adapter.

In the design of the adapter, it was important to ensure the 3FD-LL adapter plasmid had the genetic components necessary to direct transcription and translation of the genes encoding the 3FD-LL concatemer sequences. The ribosomal binding site for the **pIL5** vector was contained within the elastin adapter, which would be the sequence fragment removed from **pIL5** to enable the transfer of the 3FD-LL adapter. Therefore, a ribosomal binding site sequence was included at the *N*-terminus of the adapter to ensure expression of the synthetic gene constructs. A decahistidine tag was incorporated into the design at the *N*-terminus following the ribosomal binding site to facilitate the purification of the concatemers through affinity chromatography. **Figure 2.2** depicts the gene sequence of the 3FD-LL adapter.

EcoR I ribosomal binding site *Nco I*
 5'-GCTTAGAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCAT

 10x His Tag *BamH I* *Bsa I* *Bsa I*
 CATCATCATCATCACGGATCCGGCGAAGCCCGAGACCGTCGACGGTCTCCAGCC

 3FD-LL helix monomer
 CTGGAGAAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGGAGAAACTGGCAGC

 GCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAAAGAGCTGGCAA

HinD III
 AGTAATAAGCTTGGTCC-3'

Figure 2.2 Gene sequence for 3FD-LL helix adapter.

The 3FD-LL helix insert gene was designed to contain a codon optimized synthetic gene of the 3FD-LL helix sequence for incorporation into the 3FD-LL adapter helix insertion site. The 3FD-LL helix sequence was bordered by recognition sites for restriction endonucleases *BsmB I* and *Bbs I*, which also cut at designated locations outside of their recognition sites. This allowed for the specification of the overhangs to be complementary to the 5'-AGCC-3' cohesive ends formed from the *Bsa I* cut sites in the insertion site of the 3FD-LL adapter. Furthermore, this permitted the precise placement of the overhangs produced by the restriction enzymes to directly flank the 3FD-LL insert for the flawless formation of concatemers. The entire sequence was enclosed within the recognition sites for the *HinD III* and *BamH I* restriction enzymes for any necessary transfer of the gene into other cloning plasmids. **Figure 2.3** depicts the gene sequence of the 3FD-LL helix insert.

HinD III *Bbs I*

5'-GCATCAAGCTTGAAGACTTAGCCCTGGAGAACTGGCAGCGCTGGAAAAGC

3FD-LL helix monomer

TGGCAGCGCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGA

BsmB I *BamH I*

GCTGGCAGCGCTGAAAGAGCTGGCAGCCTGAGACGGGATCCACTAG-3'

Figure 2.3 Gene sequence for 3FD-LL helix monomer insert.

The designed 3FD-LL adapter and 3FD-LL helix insert genes were synthesized by ATUM (formerly DNA 2.0). The synthetic genes were cloned into the ATUM *E. coli* cloning vector pJ201, which contained a high copy number origin of replication (Ori_pUC) and kanamycin resistance. The 3FD-LL adapter plasmid was named **pBB12**, and the 3FD-LL helix insert plasmid was named **pBB13**. **Figure 2.4** depicts the plasmid maps of the synthetic 3FD-LL adapter and 3FD-LL helix insert genes within the pJ201 cloning vector.

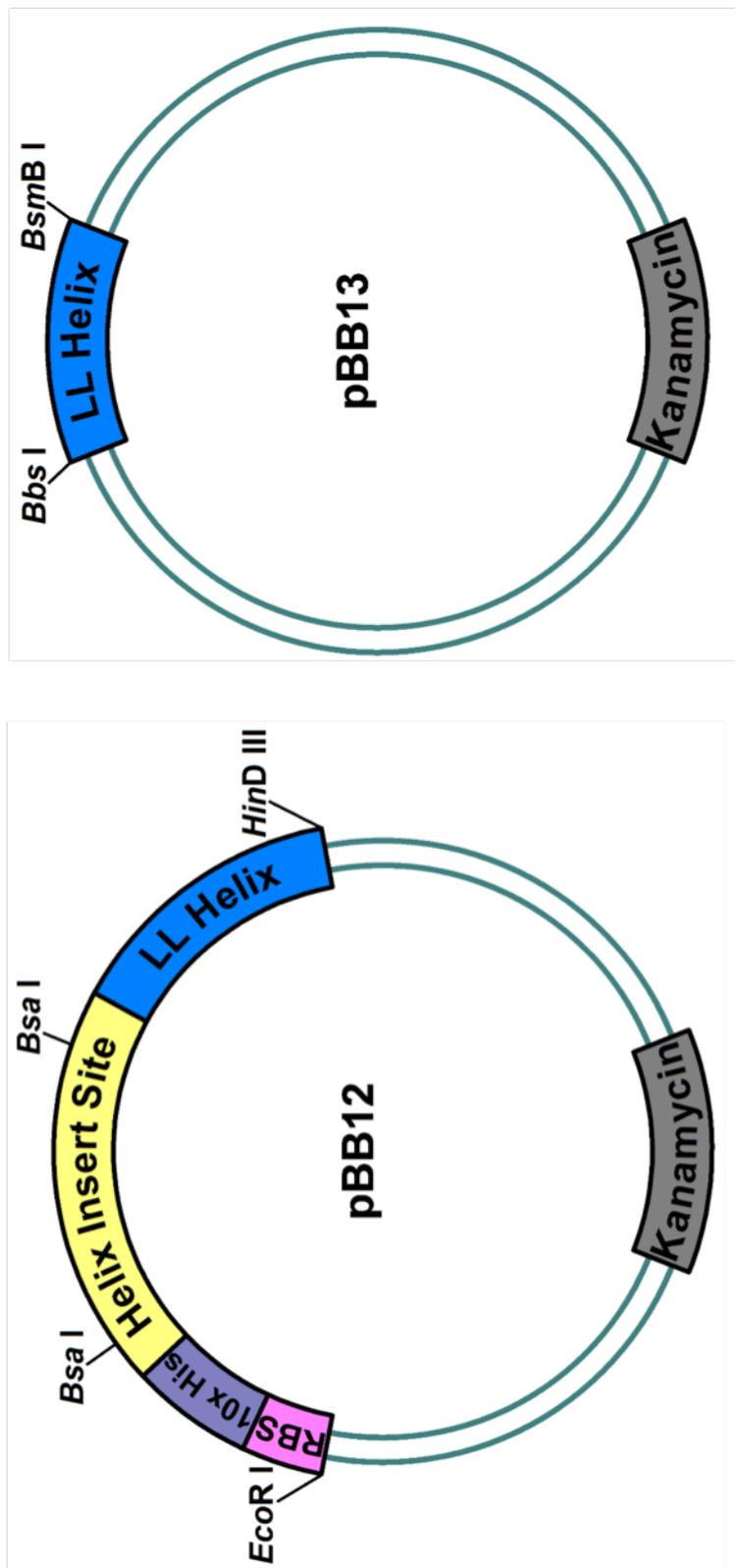
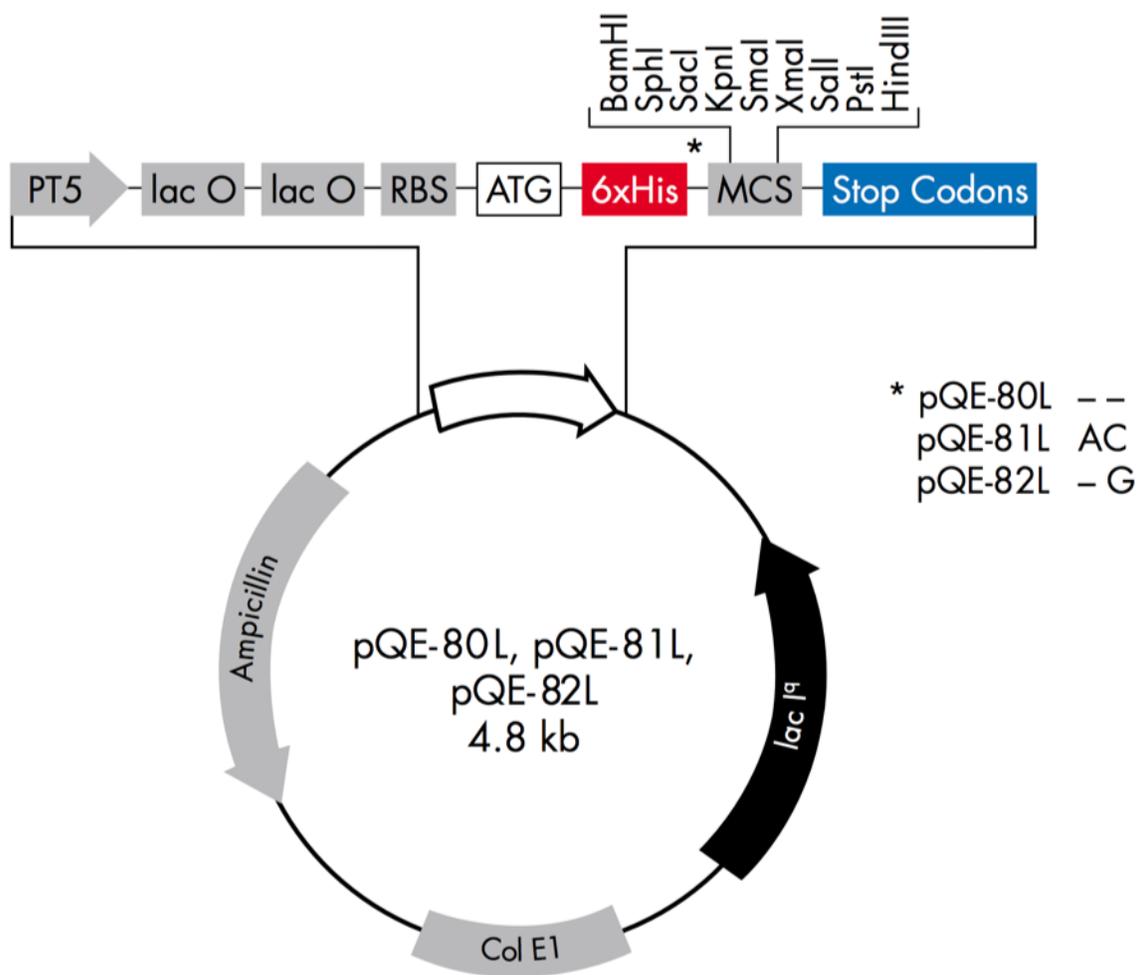


Figure 2.4 Plasmid maps for pBB12 and pBB13. Plasmid pBB12 (left) represents the synthetic 3FD-LL adapter gene contained within the ATUM *E. coli* cloning vector pJ201. Plasmid pBB13 (right) represents the synthetic 3FD-LL helix monomer insert gene contained within the ATUM *E. coli* cloning vector pJ201.

Construction of the Expression Vector

The **pQE-80L** *E. coli* expression plasmid possesses a mutation in the promoter of the *lacI* gene that facilitates the overproduction of the *lac* repressor within cells. This ensures the tight control of the basal level of transcription prior to induction with IPTG. The high copy number plasmid would support high-level production of the 3FD-LL concatemer proteins. This vector was a desirable host for the designed 3FD-LL adapter; however, the vector possesses the recognition sites for *Nco* I and *Bsa* I restriction endonucleases outside of the multiple cloning site that would impede the full functionality of the 3FD-LL adapter. **Figure 2.5** depicts the **pQE-80L** expression vector.

Previously, Dr. I-Lin Wu (formerly of the Conticello lab) engineered the commercially available **pQE-80L** plasmid to absolve it of these limitations.³ Initially, the Klenow-fill method was applied to remove the internal *Nco* I endonuclease restriction site. Next, an adapter sequence was designed with a C-terminal decahistidine tag and additional restriction sites to facilitate the cloning of elastin concatemers. The adapter, designated as the pQE-Adapter, was inserted into the **pQE-80L** vector in the multiple cloning site between the *Eco*R I and *Hin*D III endonuclease enzyme recognition sites. Finally, to support subsequent cloning of elastin concatemers, site-directed mutagenesis was utilized to introduce a single silent mutation to the *Bsa* I restriction site remaining outside of the multiple cloning site of the **pQE-80L** vector. The modified expression plasmid was named **pIL5**. With the absence of the *Nco* I and *Bsa* I restriction endonuclease sites outside of the multiple cloning site, the **pIL5** vector possessed the desired attributes for the 3FD-LL adapter system. **Figure 2.6** depicts the **pIL5** expression vector.



pQE-80 L

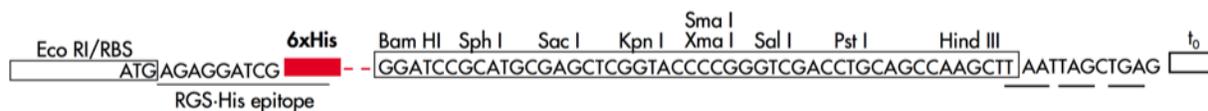


Figure 2.5 Plasmid map for the pQE-80L vector, from QIAGEN, displaying the restriction endonuclease cleavage sites within the multiple cloning site (MCS). Below the plasmid map is an expanded depiction of the MCS showing the MCS sequence.

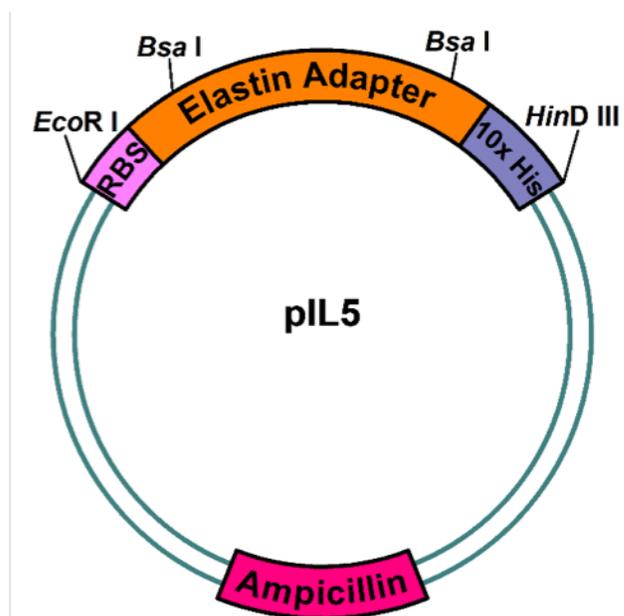


Figure 2.6 Plasmid map for pIL5 expression vector. Plasmid pIL5 was generated by site-directed mutagenesis of the pQE-80L plasmid to remove *Nco* I and *Bsa* I restriction sites in addition to the insertion of a PQE-adapter sequence into the multiple cloning site for concatemerization of elastin-mimetic sequences.

The 3FD-LL adapter was designed to contain all of the desired components present in the pQE-Adapter, such as the ribosomal binding site, a decahistidine tag, and a version of the insertion site that would support the concatemerization of the 3FD-LL helical sequence. That way, removal of the pQE-Adapter from the **pIL5** plasmid and subsequent insertion of the 3FD-LL adapter would not result in any deficiencies. The pQE-Adapter was removed from the **pIL5** expression plasmid through double digestion with *EcoR* I and *HinD* III restriction endonucleases. The synthetic 3FD-LL adapter gene was isolated from plasmid **pBB12** through restriction digestion with *EcoR* I and *HinD* III and cloned into its compatible sites in the plasmid **pIL5**. The transfer of the 3FD-LL adapter sequence into the **pIL5** plasmid resulted in the formation of the plasmid **pBB14**, which was verified through DNA sequencing. **Figure 2.7** depicts the **pBB14** expression vector.

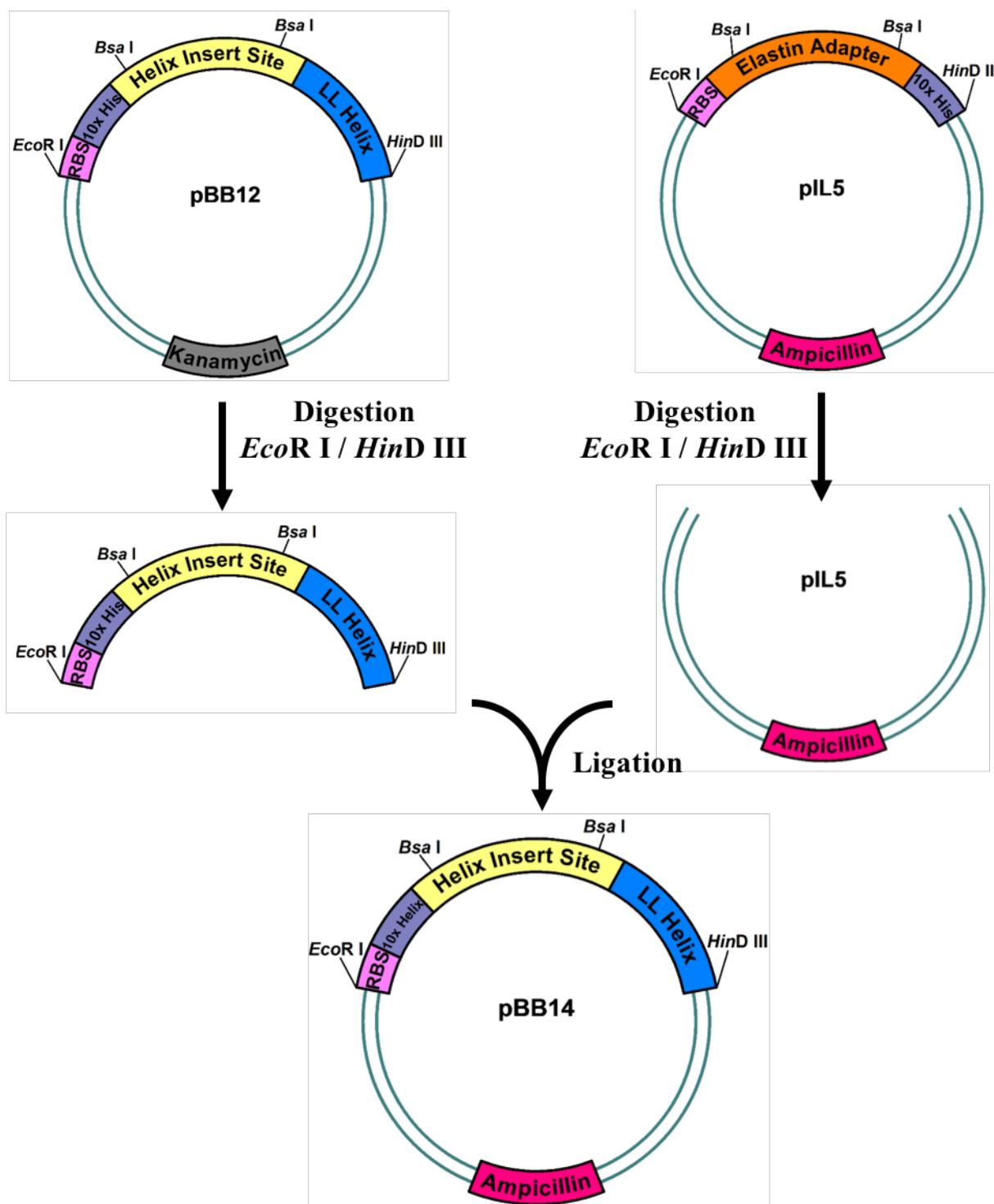


Figure 2.7 Schematic for formation of pBB14. Plasmids pBB12 and pIL5 were double digested with *EcoR* I and *HinD* III to isolate the 3FD-LL adapter and the open pIL5 vector. The 3FD-LL adapter sequence was ligated into the pIL5 vector to form plasmid pBB14.

Generation of the 3FD-LL Monomer Gene Sequence

As described earlier, the 3FD-LL adapter gene was designed to encode for one **3FD-LL monomer** on the C-terminus with the helix insertion site preceding it. The helix insertion site contained two *Bsa* I restriction endonuclease sites that when utilized would remove the helix insertion site and generate the complementary sticky ends 5'-AGCC-3'. These overhangs were necessary for cloning of the 3FD-LL concatemers into the expression vector. Furthermore, due to the nature of the *Bsa* I Type IIS restriction enzyme, the cut sites occur outside of the recognition site. The cut sites were designed at the termini of the helix insertion site, which would allow for the seamless introduction of the 3FD-LL concatemers without any unnecessary nucleotides left behind. However, if copies of the 3FD-LL insert gene were not introduced to the ligation reaction, the complementary overhangs of the adapter sequence would come back together. Normally, this would result in a false positive by recircularization of the plasmid, but in this case, it was the desired outcome as it would seamlessly remove the helix insertion site from the adapter sequence leaving an N-terminal decahistidine tag followed by the **3FD-LL monomer**.

The **3FD-LL monomer** gene sequence was produced through restriction digestion of the plasmid **pBB14** (3FD-LL adapter plasmid in expression vector) with *Bsa* I restriction endonuclease enzymes to form the complementary sticky ends that were then ligated together to remove the undesirable helix insertion segment. Transformation of the ligation products into Top10F' competent *E. coli* cells followed by standard plasmid isolation procedures were used to select potential candidates for sequencing. The **3FD-LL monomer** sequence (**pBB15**) was verified through DNA sequencing by GENEWIZ (South Plainfield, NJ). **Figure 2.8** depicts the **pBB15** expression vector and the coding sequence of the **3FD-LL monomer**.

Concatemerization of 3FD-LL Sequence

The 3FD-LL helix insert gene contained the recognition sites for the *Bbs* I and *BsmB* I restriction endonucleases at the periphery of the 3FD-LL helix monomer. When digested, the enzymes produce overhangs that were complementary to the overhangs generated by opening the helix insertion site of **pBB14** after digestion with *Bsa* I enzymes. Contrary to when the **3FD-LL monomer** was generated, it was imperative to prevent re-ligation of the adapter back together, or else the **3FD-LL monomer** would be formed, which was no longer the goal. In order to prevent re-ligation, Antarctic Phosphatase (AnP) was utilized to remove the 5'- and 3'- phosphates on the complementary sticky ends of the DNA. This greatly enhanced the chance of the 3FD-LL helix insert ligating into the 3FD-LL adapter plasmid.

Excision of the synthetic 3FD-LL helix insert gene was accomplished through sequential digestion of the **pBB13** plasmid with *BsmB* I and *Bbs* I restriction enzymes followed by isolation by agarose gel electrophoresis. The 3FD-LL helix insert was cloned into the 3FD-LL adapter of **pBB14** following *Bsa* I digestion and dephosphorylation of the helix insertion site. Potential 3FD-LL concatemers were selected and isolated through agarose gel electrophoresis and sent for DNA sequencing at GENEWIZ (South Plainfield, NJ). Sequencing verified the formation of plasmids encoding for the **3FD-LL dimer (pBB17)**, **3FD-LL trimer (pBB20)**, **3FD-LL tetramer (pBB21)**, and **3FD-LL pentamer (pBB22)**. The coding sequences of the LL concatemers are located in the appendix. **Figure 2.9** depicts the scheme for the creation of the 3FD-LL concatemers using the **3FD-LL dimer** as an example. The plasmid maps for the remaining concatemers are depicted in **Figure 2.10**.

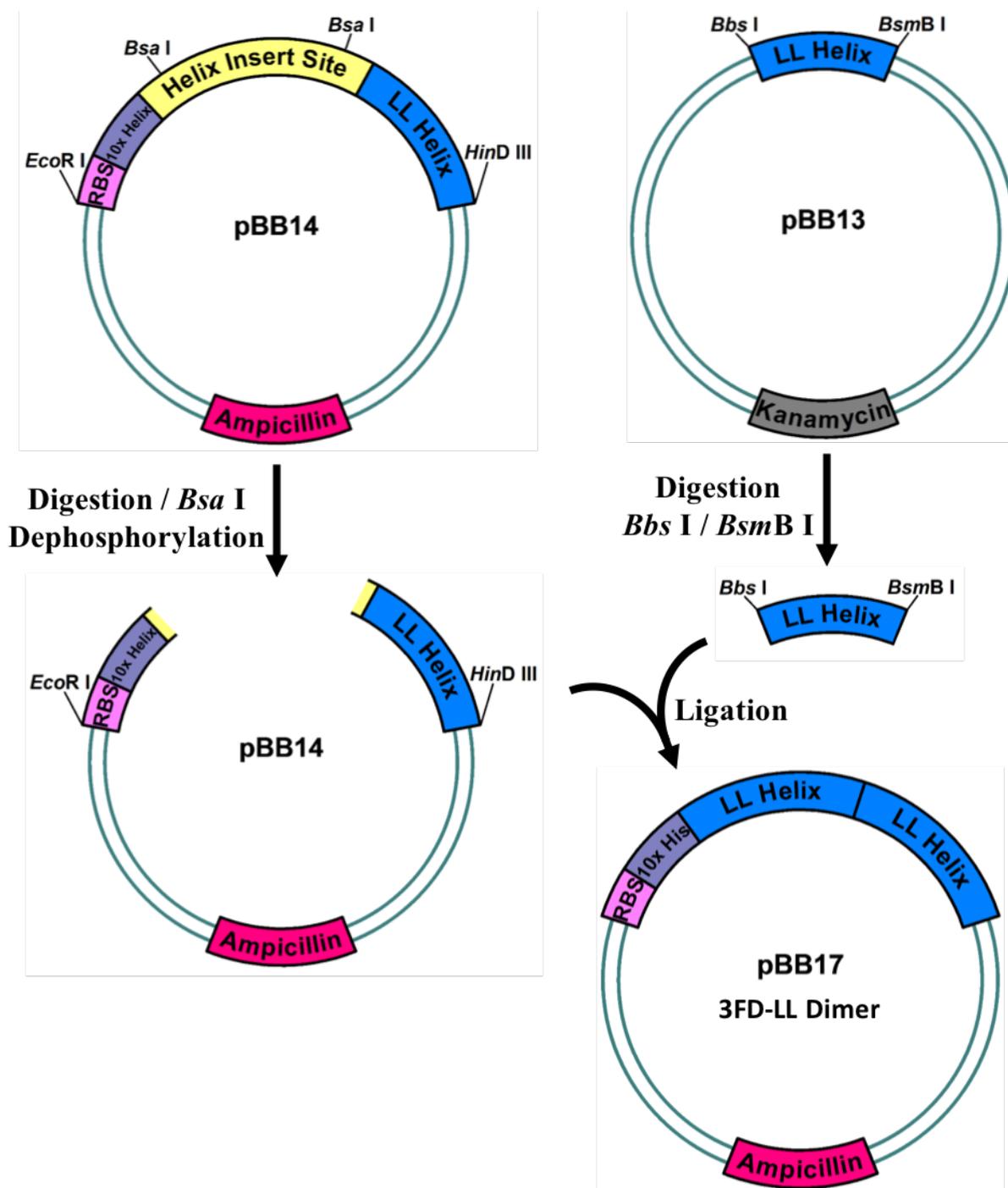


Figure 2.9 Schematic for formation of pBB17. Plasmid pBB14 was digested with *Bsa* I and dephosphorylated to isolate the 3FD-LL adapter vector. Plasmid pBB13 was sequentially digested with *BsmB* I and *Bbs* I to obtain the 3FD-LL helix insert. The 3FD-LL helix insert was ligated into the 3FD-LL adapter vector to form the 3FD-LL dimer plasmid, pBB17.

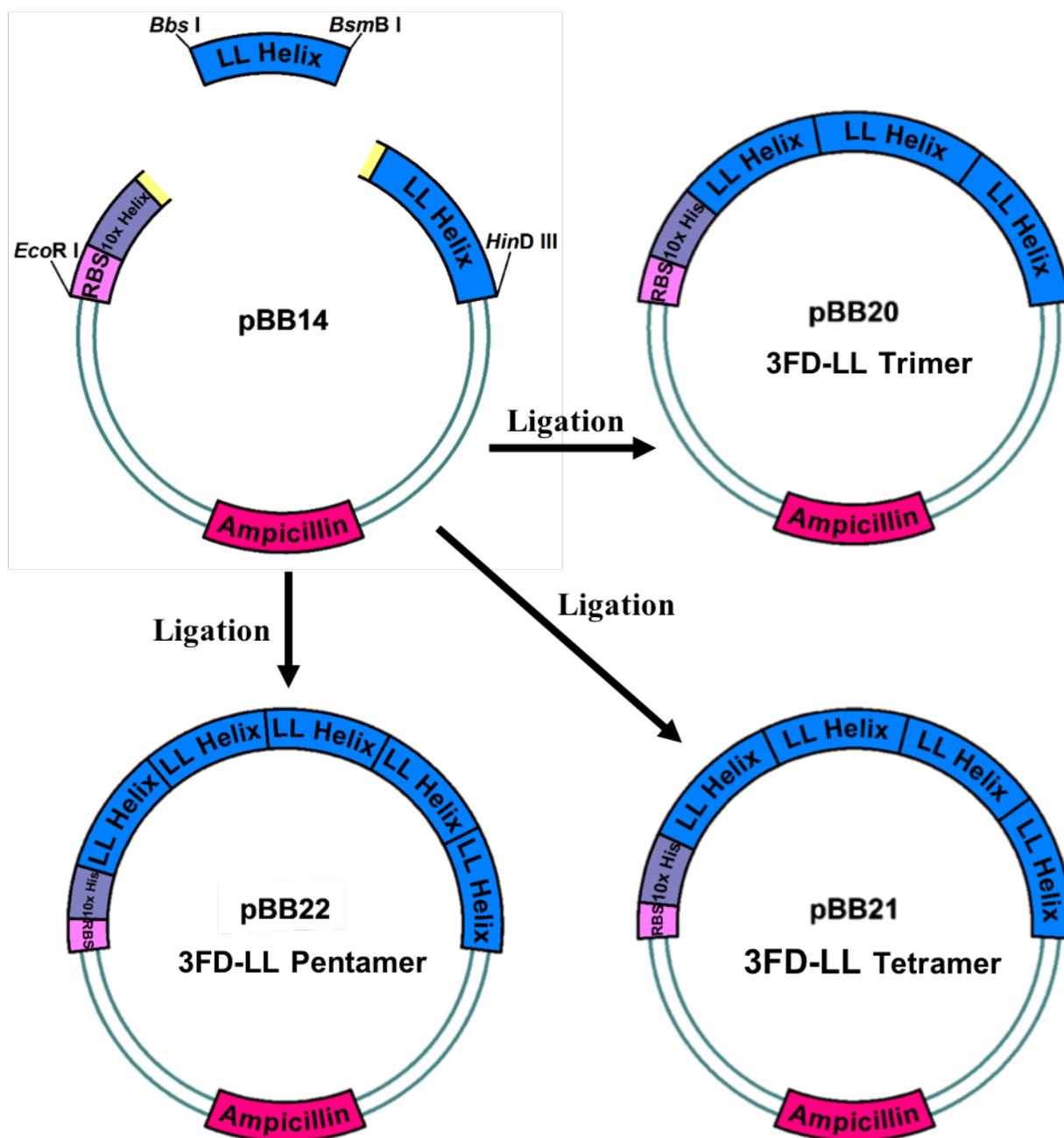


Figure 2.10 Schematic for formation of pBB20, pBB21, and pBB22. The digested and dephosphorylated 3FD-LL adapter plasmid (pBB14) was ligated with various concentrations of digested 3FD-LL helix insert to form the 3FD-LL trimer plasmid (pBB20), the 3FD-LL tetramer plasmid (pBB21), and 3FD-LL pentamer plasmid (pBB22).

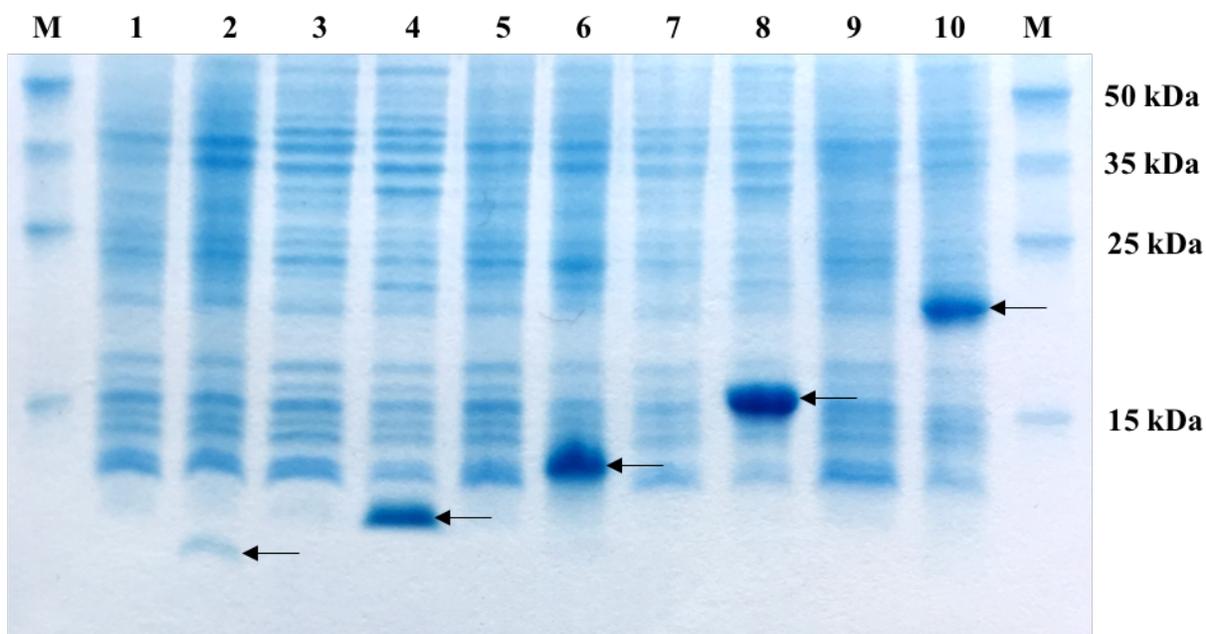
2.2.2 Generation of the 3FD-LL Concatemer Proteins

Expression of 3FD-LL Concatemers

The 3FD-LL concatemers were designed for expression using *E. coli* as a host organism. The target bacterial host for protein expression was the BL21 strain. The nutritionally rich Terrific broth (TB) medium was chosen as the culture media for expression of the 3FD-LL concatemers as it was recommended for expression of recombinant proteins in *E. coli* strains due its ability to support an extended growth phase as a result of its higher composition of glycerol, yeast extract, and peptone.⁴ Small-scale expressions of the five concatemers were carried out at 30 °C through *lac* operon function and induction by isopropyl thiogalactoside (IPTG).

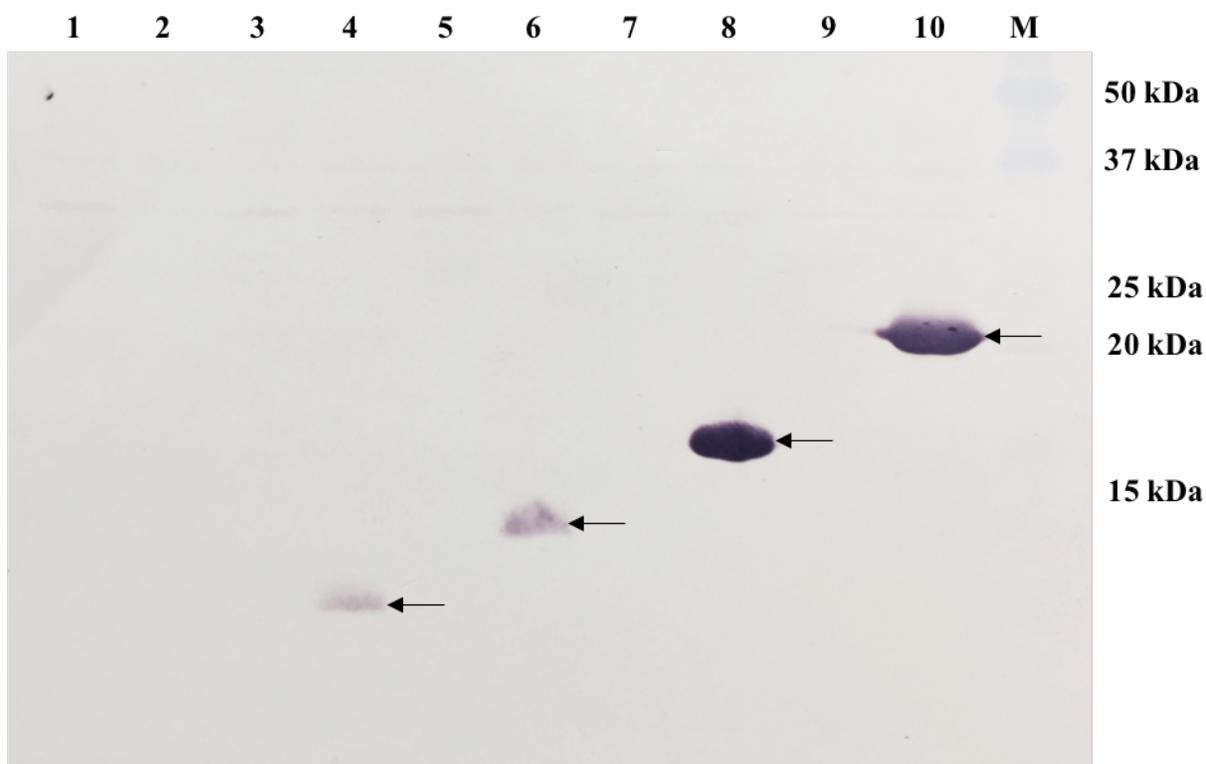
SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was used to verify protein production (**Figure 2.11**). Successful expression was confirmed for all five concatemer proteins; although, the band for the **3FD-LL monomer** was very faint. Small peptides are only partially focused in the stacking gel and they quickly migrate through the gel which results in weak, diffuse bands. Tricine gels were attempted for improved focusing and narrower separation of low molecular weight proteins,⁵ but this technique failed to increase visibility of the monomer band.

A western blot was performed to further verify successful protein expression of the concatemers (**Figure 2.12**). Although a nitrocellulose membrane with a small pore size of 0.2 µM was used, the **3FD-LL monomer** passed through the membrane and was not visualized. All of the larger concatemers were present on the western blot membrane and yielded dark purple precipitate from the reaction of alkaline phosphatase on the secondary antibodies with the NBT/BCIP developing reagent. Following verification of protein expression, large-scale expressions were conducted to produce large quantities of the proteins for purification.



- | | |
|---|---|
| M. Perfect protein marker (15-150 kDa) | 6. 3FD-LL trimer, 4 hrs after induction |
| 1. 3FD-LL monomer, before induction | 7. 3FD-LL tetramer, before induction |
| 2. 3FD-LL monomer, 4 hrs after induction | 8. 3FD-LL tetramer, 4 hrs after induction |
| 3. 3FD-LL dimer, before induction | 9. 3FD-LL pentamer, before induction |
| 4. 3FD-LL dimer, 4 hrs after induction | 10. 3FD-LL pentamer, 4 hrs after induction |
| 5. 3FD-LL trimer, before induction | M. Perfect protein marker (15-150 kDa) |

Figure 2.11 SDS-PAGE analysis indicating the results of protein expression for the 3FD-LL concatemers in the BL21 *E. coli* strain with IPTG induction. Distinctive protein expression bands were evident for all induced samples after 3-4 hours (lanes 2, 4, 6, 8, and 10). The 3FD-LL monomer band was likely faint due to its low molecular weight. MW: monomer = 5.66 kDa; dimer = 9.41 kDa; trimer = 13.2 kDa; tetramer = 16.9 kDa; pentamer = 20.7 kDa.



- | | |
|--|--|
| 1. 3FD-LL monomer, before induction | 7. 3FD-LL tetramer, before induction |
| 2. 3FD-LL monomer, 4 hrs after induction | 8. 3FD-LL tetramer, 4 hrs after induction |
| 3. 3FD-LL dimer, before induction | 9. 3FD-LL pentamer, before induction |
| 4. 3FD-LL dimer, 4 hrs after induction | 10. 3FD-LL pentamer, 4 hrs after induction |
| 5. 3FD-LL trimer, before induction | M. Perfect protein marker (15-150 kDa) |
| 6. 3FD-LL trimer, 4 hrs after induction | |

Figure 2.12 Western blot analysis indicating protein expression of the 3FD-LL concatemers in the BL21 *E. coli* strain with IPTG induction. Distinctive protein expression bands were evident 4 hours after induction (lanes 4, 6, 8, and 10). The 3FD-LL monomer likely diffused through the membrane too rapidly and did not appear. MW: monomer = 5.66 kDa; dimer = 9.41 kDa; trimer = 13.2 kDa; tetramer = 16.9 kDa; pentamer = 20.7 kDa.

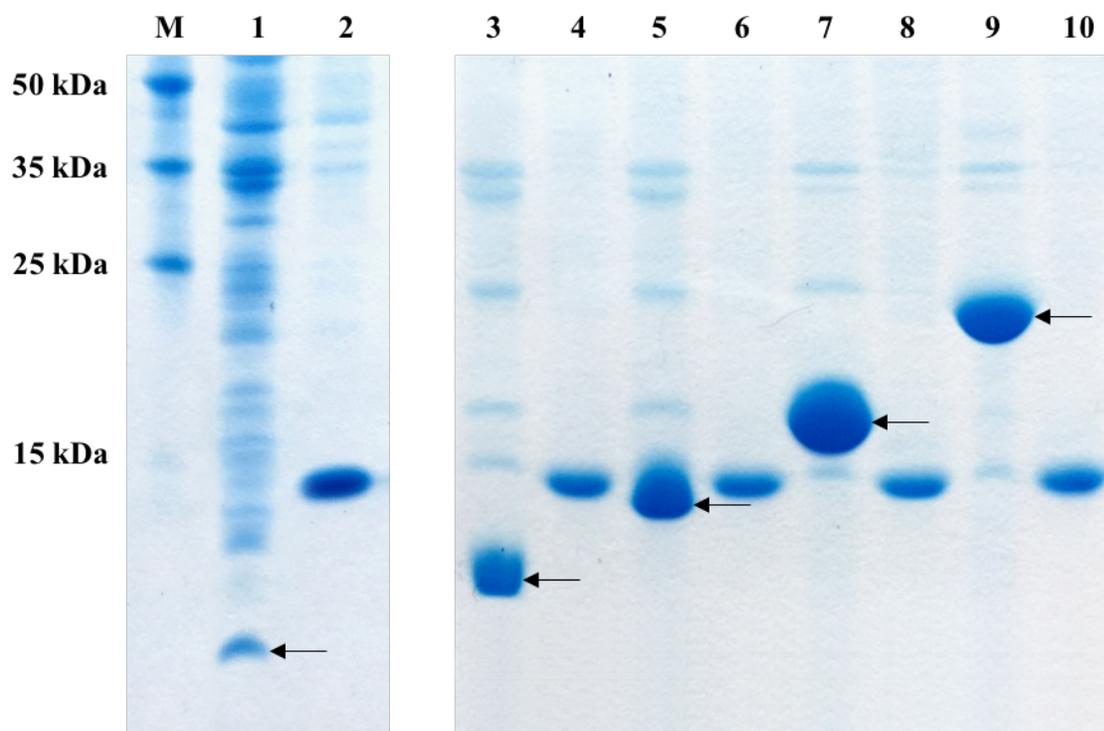
Cell Lysis

After the expression cells were harvested, they were subjected to multiple lysis methods including freeze/thaw cycles, enzymatic lysis, and sonication. The soluble and insoluble fractions were separated by centrifugation and analyzed using SDS-PAGE. A 16% SDS-PAGE gel indicated that the 3FD-LL concatemer proteins were insoluble as they were localized in the lysate pellet as seen in **Figure 2.13**. The dark band present around 14 kDa in all lysed supernatant samples was identified as lysozyme, which has a single polypeptide chain with a molecular weight of 14.3 kDa.

Protein Solubilization

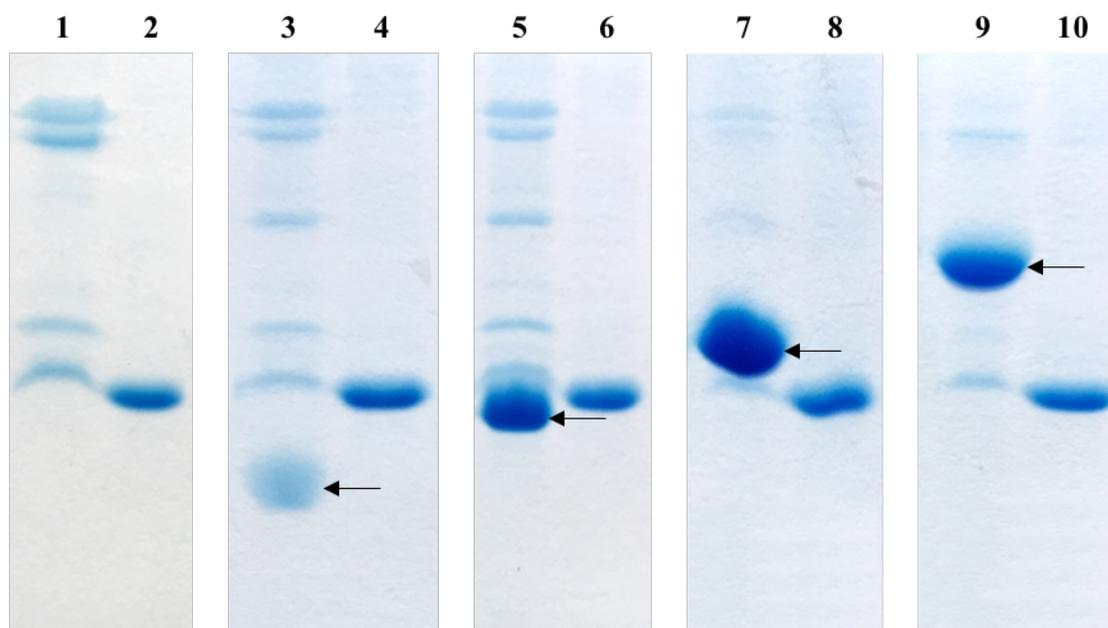
In order to facilitate purification on a cobalt resin column, the proteins had to be removed from the pellet and solubilized. Initial attempts to solubilize the proteins in 6 M urea were unsuccessful. Analysis using a 16% SDS-PAGE gel determined that all of the 3FD-LL concatemer proteins remained entirely in the insoluble phase after shaking in 6 M urea for a few days (**Figure 2.14**). This indicated that a harsher denaturant would be required to solubilize the 3FD-LL concatemers.

Guanidine hydrochloride (GHC1) is a strong protein denaturant capable of unfolding most proteins at high concentrations. The 3FD-LL concatemers in the lysed pellets were successfully solubilized under denaturing conditions in a 7 M guanidine hydrochloride buffer. SDS-PAGE analysis verified the isolation of the proteins in the GHC1 supernatant. The solubilization of the **3FD-LL tetramer** is depicted in **Figure 2.15**. Sodium dodecyl sulfate (SDS) is not soluble at high concentrations of GHC1, and the high ionic strength of the GHC1 buffer is problematic for electrophoresis. Therefore, a second sample of GHC1 supernatant was dialyzed in dH₂O for improved visualization during SDS-PAGE analysis.



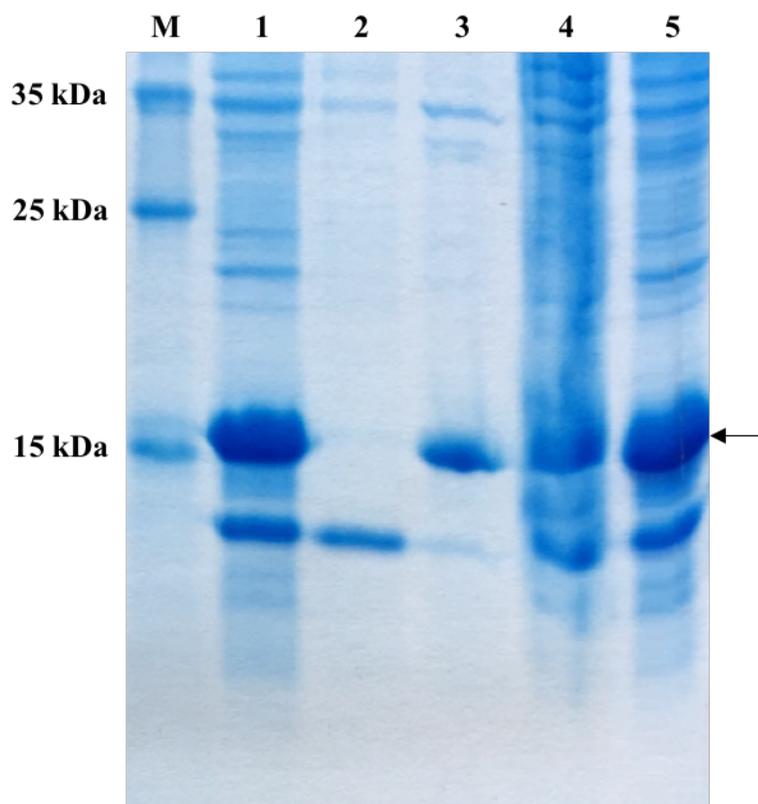
- | | |
|---|---|
| M. Perfect protein marker (15-150 kDa) | 6. 3FD-LL trimer, lysed supernatant |
| 1. 3FD-LL monomer, lysed pellet | 7. 3FD-LL tetramer, lysed pellet |
| 2. 3FD-LL monomer, lysed supernatant | 8. 3FD-LL tetramer, lysed supernatant |
| 3. 3FD-LL dimer, lysed pellet | 9. 3FD-LL pentamer, lysed pellet |
| 4. 3FD-LL dimer, lysed supernatant | 10. 3FD-LL pentamer, lysed supernatant |
| 5. 3FD-LL trimer, lysed pellet | |

Figure 2.13 SDS-PAGE analysis indicating the results of cell lysis for the 3FD-LL concatemers after expression. Target protein bands representing the 3FD-LL concatemers were present in the insoluble pellets after lysis (lanes 1, 3, 5, 7, and 9). The dark band present around 14.3 kDa in the lysed supernatant samples was lysozyme. MW: monomer = 5.66 kDa; dimer = 9.41 kDa; trimer = 13.2 kDa; tetramer = 16.9 kDa; pentamer = 20.7 kDa.



- | | |
|-------------------------------------|---------------------------------------|
| 1. 3FD-LL monomer, urea pellet | 6. 3FD-LL trimer, urea supernatant |
| 2. 3FD-LL monomer, urea supernatant | 7. 3FD-LL tetramer, urea pellet |
| 3. 3FD-LL dimer, urea pellet | 8. 3FD-LL tetramer, urea supernatant |
| 4. 3FD-LL dimer, urea supernatant | 9. 3FD-LL pentamer, urea pellet |
| 5. 3FD-LL trimer, urea pellet | 10. 3FD-LL pentamer, urea supernatant |

Figure 2.14 SDS-PAGE analysis indicating the solubilization results of the 3FD-LL concatemers in a 6 M urea lysis buffer. Target protein bands representing the 3FD-LL concatemers were present in the insoluble pellets after incubation in 6 M urea (lanes 3, 5, 7, and 9). The 3FD-LL monomer band (lane 1) could not be visualized on the gel and likely diffused out due to its low molecular weight. The dark band present around 14.3 kDa in the urea supernatant samples was lysozyme. MW: monomer = 5.66 kDa; dimer = 9.41 kDa; trimer = 13.2 kDa; tetramer = 16.9 kDa; pentamer = 20.7 kDa.



M. Perfect protein marker (15-150 kDa)

1. 3FD-LL tetramer, lysed pellet

2. 3FD-LL tetramer, lysed supernatant

3. 3FD-LL tetramer, GHCl pellet

4. 3FD-LL tetramer, GHCl supernatant

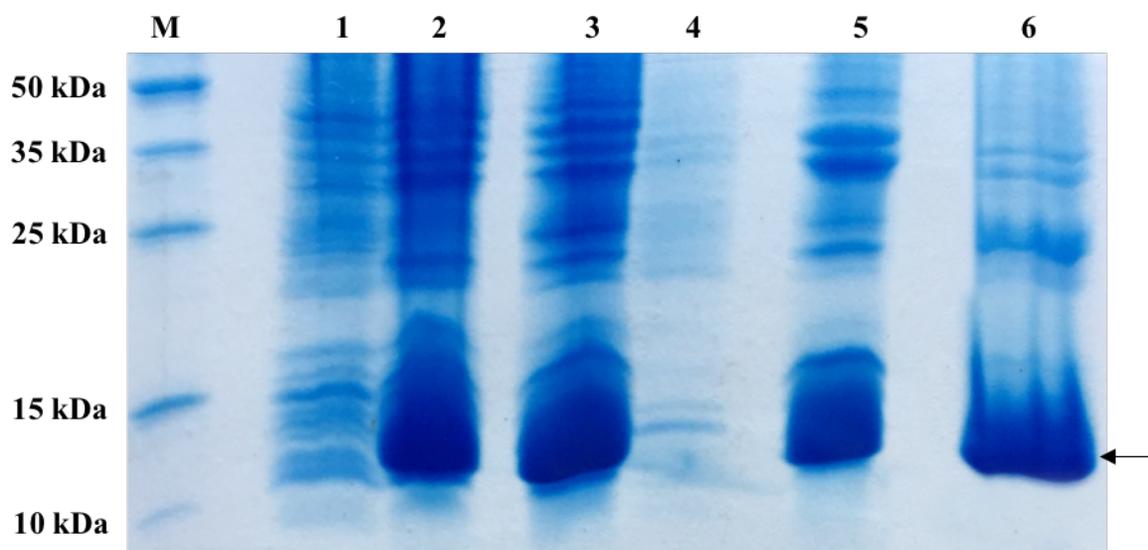
5. 3FD-LL tetramer, GHCl supernatant, dialyzed

Figure 2.15 SDS-PAGE analysis indicating the solubilization results of the 3FD-LL tetramer in 7 M guanidine hydrochloride (GHCl). The target protein bands representing the 3FD-LL tetramer were visualized in the soluble supernatants of the 7 M GHCl solution (lanes 4). The supernatant was dialyzed for improved analysis by SDS-PAGE (lane 5). Some 3FD-LL tetramer remained in the pellet after solubilization (lane 3). MW: tetramer = 16.9 kDa.

Inclusion Body Purification

Although the 3FD-LL concatemers were successfully solubilized in 7 M GdCl buffers, a vast majority of the other biomolecules in the lysed pellet were also pulled into the soluble fraction, which made the GdCl supernatant extremely viscous. While the concatemers were soluble and capable of purification on a cobalt resin column, the viscosity of the solution eventually made it so that gravity was insufficient for the flow of the GdCl supernatants through the column. In order to improve the protein purification process, the quantity of proteins solubilized from the cell lysate pellet was decreased. One option was to wash the lysed pellets with detergents to remove undesirable cellular debris and provide purer sources of the target proteins.

Triton X-100 has been successfully used for the removal of lipid and membrane-associated proteins from inclusion bodies.⁶ The lysed cell pellets were subjected to multiple rounds of washes with a lysis buffer containing 1% Triton X-100. Each wash was concluded by sonication followed by centrifugation to remove solubilized molecules. Triton X-100 was then removed by washing the pellets with a lysis buffer without the detergent. The washed pellets were analyzed by SDS-PAGE to determine if the Triton X-100 wash cycles improved the purity of the inclusion bodies. The purity of the **3FD-LL trimer** at various steps is depicted in **Figure 2.16**. SDS-PAGE indicated that the Triton X-100 detergent successfully removed some of the insoluble debris from the lysed pellets evident by the decreased quantity of bands in the washed inclusion body sample. Solubilization of the purified inclusion bodies in a 7 M GdCl buffer followed by isolation of the recombinant protein in the GdCl supernatant resulted in a solution with considerably lower viscosity, which would allow the purification process to occur by gravitational force alone.



M. Perfect protein marker (10-225 kDa)

1. 3FD-LL trimer, expression, before induction

2. 3FD-LL trimer, expression, 18 hrs after induction

3. 3FD-LL trimer, lysed pellet

4. 3FD-LL trimer, lysed supernatant

5. 3FD-LL trimer, washed inclusion bodies

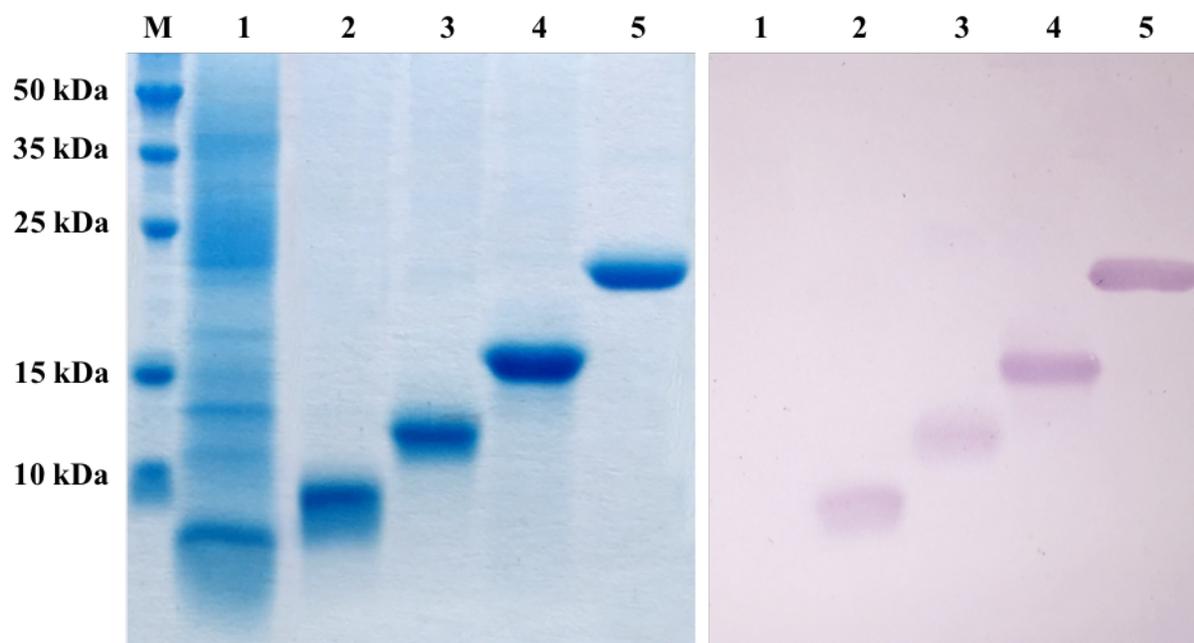
6. 3FD-LL trimer, GHI supernatant, dialyzed

Figure 2.16 SDS-PAGE analysis indicating the purification of the inclusion bodies within the 3FD-LL trimer lysed pellet. The target protein bands representing the 3FD-LL trimer were visualized in lanes 2, 3, 5, and 6. Fewer bands were visible in the washed inclusion bodies sample (lane 5) compared to the lysed pellet sample (lane 3) implying that undesirable cellular debris was successfully removed. MW: trimer = 13.2 kDa.

Purification of the 3FD-LL Concatemers

The decahistidine tags on the GHCl solubilized 3FD-LL concatemer proteins facilitated their purification on HisPur™ cobalt-charged resin columns with subsequent elution by imidazole. Dialysis of the purification eluents in ddH₂O was intended to remove GHCl, imidazole, and other small molecules from the solution. The proteins precipitated out of solution during dialysis, which was expected as the GHCl concentration decreased. The dialyzed proteins were lyophilized, which resulted in pure white powders, and protein yields were calculated. The protein yields from 1 L of expression culture were 82 mg for the **3FD-LL monomer**, 117 mg for the **3FD-LL dimer**, 157 mg for the **3FD-LL trimer**, 123 mg for the **3FD-LL tetramer**, and 139 mg for the **3FD-LL pentamer**.

SDS-PAGE and western blot analysis were utilized to determine relative purity and mass of the pure 3FD-LL concatemer proteins (**Figure 2.17**). A 16% SDS-PAGE gel confirmed that most of the concatemer proteins were highly pure, and the bands created were appropriate for the expected masses of the concatemer proteins. However, the **3FD-LL monomer** did not appear purified on the SDS-PAGE gel and it continued to elude visualization on the western blot likely due to its low molecular weight. Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry was used to verify the masses of the 3FD-LL concatemers including the **3FD-LL monomer**, which could not be easily verified by SDS-PAGE and western blot analysis.



- M.** Perfect protein marker (10-225 kDa) **3.** 3FD-LL trimer, HisPurTM, lyophilized
1. 3FD-LL monomer, HisPurTM, lyophilized **4.** 3FD-LL tetramer, HisPurTM, lyophilized
2. 3FD-LL dimer, HisPurTM, lyophilized **5.** 3FD-LL pentamer, HisPurTM, lyophilized

Figure 2.17 SDS-PAGE (left) and western blot (right) analysis depicting the purification results for the 3FD-LL concatemers by affinity chromatography using HisPurTM cobalt resin. Purified eluents were dialyzed against dH₂O and lyophilized to obtain powder. Purified protein was visualized for the 3FD-LL dimer (lane 2), 3FD-LL trimer (lane 3), 3FD-LL tetramer (lane 4), and 3FD-LL pentamer (lane 5). The 3FD-LL monomer was not successfully purified evident by many additional bands on the SDS-PAGE gel (lane 1). The 3FD-LL monomer did not appear on the western blot (lane 1). MW: monomer = 5.66 kDa; dimer = 9.41 kDa; trimer = 13.2 kDa; tetramer = 16.9 kDa; pentamer = 20.7 kDa.

2.3 Conclusions

Molecular cloning techniques were useful for the design and engineering of short concatemer protein polymers of the **3FD-LL** peptide sequence with defined sequence specificity. Through the use of these methods, control of the sequence at the DNA level results in control over the amino acid sequence. Control over sequence specificity is imperative for the design of proteins for materials applications since a single amino acid change can result in detrimental structural changes.

A system was designed to generate longer helical units by concatemerization of the **3FD-LL** sequence through the insertion of copies of the sequence into an adapter. A series of plasmids were generated that encoded the **3FD-LL monomer** and four larger concatemer sequences, and gene sequencing verified that sequence control was maintained throughout the process. The corresponding protein polymers were successfully expressed using *E. coli* as a host organism. Solubilization of the recalcitrant proteins was a particularly strenuous task; however, a protocol was eventually developed and optimized that successfully solubilized the concatemers prior to purification. The recombinant 3FD-LL concatemer proteins were purified from the endogenous proteins of the bacterial host using immobilized metal affinity chromatography using HisPur™ cobalt resin.

2.4 Experimental

2.4.1 Materials

All chemical reagents were purchased from either Thermo Fisher Scientific, Inc. (Waltham, MA) or MilliporeSigma (Burlington, MA) unless otherwise noted. T4 DNA ligase, T4 kinase, and restriction endonucleases (*EcoR* I-HF, *HinD* III-HF, *Bsa* I, *Bbs* I, and *BsmB* I) were purchased from New England Biolabs, Inc. (Ipswich, MA). The plasmid pIL5, an expression vector modified from pQE-80L, was synthesized by Dr. I-Lin Wu in the Conticello lab at Emory University³. The plasmid pQE-80L was obtained from QIAGEN, Inc. (Valencia, CA). Synthetic single-stranded PAGE-purified oligonucleotides were purchased from either Integrated DNA Technologies (Coralville, IA) or Sigma-Genosys, Inc. (The Woodlands, TX). Synthesis of plasmids containing codon-optimized genes encoding the 3FD-LL adapter and the 3FD-LL helix insert was conducted by ATUM, formerly known as DNA2.0 (Newark, CA). The Top10F' chemically competent *E. coli* strain was obtained from Invitrogen Corp. (Carlsbad, CA), and the BL21 (DE3) chemically competent *E. coli* strain was purchased from New England Biolabs, Inc. (Ipswich, MA). The QIAprep-spin miniprep kit was purchased from QIAGEN, Inc. (Valencia, CA). The Zymoclean™ Gel DNA Recovery Kit and DNA Clean & Concentrator™ kit were obtained from Zymo Research Corporation (Irvine, CA). The 6x DNA loading dye and O'GeneRuler™ 1 kb Plus DNA ladder for agarose gels were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

Luria-Bertani broth and agar powder were purchased from MilliporeSigma (Burlington, MA), and Terrific Broth (TB) powder was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). Kanamycin monosulfate was purchased from VWR International, LLC

(Suwanee, GA). Ampicillin was purchased from Research Products International Corp. (Prospect, IL). Isopropyl- β -D-thiogalactopyranoside (IPTG) and ampicillin was purchased from Research Products International Corp. (Prospect, IL). Benzonase[®] nuclease and protease-inhibitor cocktail (EDTA-free) were purchased from MilliporeSigma (Burlington, MA). Lysozyme from chicken egg white was purchased from Research Products International Corp. (Prospect, IL). HisPur[™] cobalt resin was obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). SnakeSkin[™] Dialysis Tubing with a MWCO of 3.5 kDa or 10 kDa MWCO and Slide-A-Lyzer[™] G2 Dialysis Cassettes with a 2 kDa MWCO were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

The Perfect Protein Marker (10-225 kDa or 15-150 kDa) was purchased from MilliporeSigma (Burlington, MA) and the Precision Plus Protein[™] Dual Color Standard was purchased from Bio-Rad Laboratories, Inc. (Hercules, CA). The His-tag AP Western Reagents Kit, including the Goat Anti-Mouse IgG AP Conjugate (H + L) secondary antibody, was purchased from MilliporeSigma (Burlington, MA). The anti His-tag monoclonal antibody was purchased from Clontech, Inc. (Mountain View, CA). The 1-step NBT/BCIP reagent mixture was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). Vacuum filters with a 0.2 μ m polyethersulfone (PES) membrane were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

2.4.2 General Methods

Basic molecular biology procedures were adapted from a standard molecular cloning manual⁷ or the protocol supplied by the manufacturer unless otherwise described in detail. All Reagents intended for use with bacteria, DNA, or recombinant proteins were sterilized by either

syringe filtration through a 0.2 μm cellulose membrane, vacuum filtration through a 0.2 μm polyethersulfone (PES) membrane, or by autoclaving. Synthetic plasmids were transformed into chemically competent Top10F' or BL21 *E. coli* strains. *E. coli* strains were grown at 37 °C in Luria-Bertani (LB) medium containing appropriate antibiotic with shaking at 200 rpm unless otherwise stated. A QIAprep-spin miniprep kit or a DNA Clean & ConcentratorTM kit was used for isolation of plasmid DNA. Plasmids were digested with restriction enzymes and separated by DNA agarose gel electrophoresis (0.75%-2% agarose). Agarose gels were run at 120 V for approximately 30 minutes. All enzyme reactions were conducted in the reagent buffers provided by the manufacturer unless otherwise noted. The ZymocleanTM Gel DNA Recovery Kit was used for isolation and purification of the desired fragments from enzyme digestions. Dephosphorylations were carried out for 2 hours at 37 °C in 30 μL total volumes with 1 μL of Antarctic Phosphatase and 1X phosphatase buffer. Gene fragments were ligated together at 16 °C overnight in 10 μL volumes with 0.5 μL T4DNA Ligase (200 units) and 1X T4DNA Ligase buffer. Ligation products were transformed into chemically competent Top10F' cells and recovered in 400 μL SOC rich media at 37 °C for 1 hour. The transformed cells were plated onto LB agar containing the appropriate antibiotic and incubated at 37 °C for 16 hours. Confirmation of desired gene sequences was determined by automated DNA sequencing analysis at GENEWIZ LLC (South Plainfield, NJ) using the appropriate forward and reverse oligonucleotide primers.

Protein electrophoresis was conducted using 16% SDS polyacrylamide gels on a Mini-PROTEIN 3 cell electrophoresis system from Bio-Rad Laboratories, Inc. (Hercules, CA). The buffer tank was filled with SDS run buffer (25 mM tris, 250 mM glycine, 0.1% SDS, pH 8.3). The Perfect Protein Marker was used as a protein standard for SDS-PAGE analysis. An initial sample volume of 10 μL was added to each well with any necessary concentration adjustments made

thereafter. Gels were run at 150 V for 1-1.5 hours depending on the desired level of separation. Gels were stained in coomassie overnight and then destained in a methanol-acetic acid buffer. Presence of desired proteins were further verified by western blotting using the His-tag AP Western Reagents Kit. The Precision Plus Protein™ Dual Color Standard was used as the protein standard for western blot analysis. The proteins were transferred onto a 0.2 μM nitrocellulose membrane. Proteins that contained a polyhistidine tag were labelled with a mouse anti-His-tag IgG monoclonal primary antibody followed by a goat anti-mouse IgG AP conjugate (H + L) secondary antibody. Analysis after incubation with the 1-step NBT/BCIP reagent mixture allowed for chemiluminescent visualization of the his-tagged recombinant proteins. All proteins were purified using immobilized metal affinity chromatography (IMAC) on HisPur™ cobalt resin followed by dialysis to remove imidazole.

2.4.3 Genetic Engineering of 3FD-LL Concatemer Sequences

Construction of the Expression Plasmid

A 3FD-LL adapter gene comprising an N-terminal decahistidine tag, a helix insertion site, and a C-terminal codon optimized DNA sequence of the 3FD-LL sequence, E(ALEKLA)₃(ALKELA)₃K, was ordered from ATUM (formerly DNA 2.0) in the pJ201 vector. The original 3FD-LL adapter plasmid was named **pBB12**. The lyophilized powder containing the plasmid was resuspended in 50 μL of distilled, deionized water (ddH₂O) upon arrival. The modified **pQE-80L** expression vector, **pIL5**, was synthesized by Dr. I-Lin Wu, and a sample of the plasmid was directed for use in the 3FD-LL system.

Plasmid DNA (1 μL) for the 3FD-LL adapter (**pBB12**) and the modified expression vector (**pIL5**) were transformed into chemically competent Top10F' *E. coli* cells. The cells were

recovered after heat shock in 400 μL SOC rich media for 30 minutes at 37 $^{\circ}\text{C}$. Aliquots of 50 μL of the transformed suspensions were spread onto LB agar plates and incubated for 14 hours at 37 $^{\circ}\text{C}$. The cells containing plasmid **pBB12** were spread on plates containing kanamycin (50 $\mu\text{g}/\text{mL}$) for antibiotic selection. The cells containing plasmid **pIL5** were spread on plates containing ampicillin (100 $\mu\text{g}/\text{mL}$) for antibiotic selection.

Twelve colonies of each plasmid were selected from the plates containing the Top 10 F' strain. The selected colonies were used to inoculate culture tubes containing 5 mL LB media supplemented with the appropriate antibiotic (50 $\mu\text{g}/\text{mL}$ kanamycin for cells containing **pBB12** and 100 $\mu\text{g}/\text{mL}$ ampicillin for cells containing **pIL5**). The cultures were grown at 37 $^{\circ}\text{C}$ overnight on a rotator. Two cultures of each plasmid were prepared for long term storage within the Top10F' *E. coli* strain. These frozen stocks were prepared by combining 200 μL of 80% glycerol with 800 μL of the culture. The frozen stocks of the Top10F' strains containing the **pBB12** or **pIL5** plasmid were stored at -80 $^{\circ}\text{C}$ until needed. To make secondary stocks of the **pBB12** and **pIL5** plasmid DNA, a QIAprep-spin miniprep kit (QIAGEN, Inc.) was used to isolate the plasmid DNA of each plasmid in the remaining cultures of Top10F' cells. The isolated DNA was recovered in 50 μL of EB buffer (10 mM Tris-HCl, pH 8.5) and stored in solution at -30 $^{\circ}\text{C}$ when not in use.

The synthetic 3FD-LL adapter gene was isolated through restriction digestion of the plasmid **pBB12** with *EcoR* I and *HinD* III endonucleases. A volume of 50 μL plasmid DNA was incubated with 3 μL of each restriction enzyme and the appropriate buffer with a total reaction volume of 100 μL . After incubation at 37 $^{\circ}\text{C}$ for 4 hours, the digestion products were separated by 1.5% DNA agarose gel electrophoresis. The small band, which contained the 3FD-LL adapter gene, was excised from the gel using a ZymocleanTM Gel DNA Recovery Kit and eluted into 10 μL of ddH₂O.

The elastin adapter gene was removed from the **pIL5** expression vector through restriction digestion of the plasmid with *EcoR* I and *HinD* III endonucleases. A volume of 50 μ L plasmid DNA was incubated with 3 μ L of each restriction enzyme and the appropriate buffer with a total reaction volume of 100 μ L. After incubation at 37 °C for 4 hours, the digestion products were separated by 0.75% DNA agarose gel electrophoresis (**Figure 2.18**). The large band, which contained the **pIL5** expression vector without the elastin adapter, was excised from the gel using a Zymoclean™ Gel DNA Recovery Kit and eluted into 10 μ L of ddH₂O.

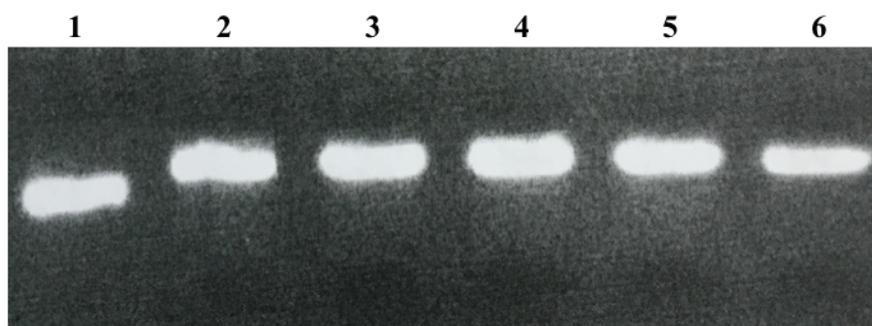


Figure 2.18 Agarose gel of plasmid pIL5. Lane 1 depicts the undigested pIL5 plasmid. Lanes 2-5 display the pIL5 vector after digestion with *EcoR* I and *HinD* III restriction endonucleases. The bright band in lanes 2-6 was excised. Size: pIL5 = 4779 bp.

The 3FD-LL adapter gene was cloned into its compatible sites in the **pIL5** expression vector through ligation of the isolated digestion products. The digested 3FD-LL adaptor gene and **pIL5** vector were combined together at molar ratios of 3:1 and 5:1. The gene fragments were incubated overnight at 16 °C in 10 μ L reaction volumes with 0.5 μ L T4DNA Ligase (200 units) and 1X T4DNA Ligase buffer. A 1 μ L aliquot of each ligation product was transformed into

chemically competent Top10F' cells. The cells were recovered in 400 μ L SOC rich media at 37 $^{\circ}$ C for 1 hour, plated onto LB agar containing ampicillin (100 μ g/mL), and incubated at 37 $^{\circ}$ C for 16 hours. Six overnight cultures containing ampicillin (100 μ g/mL) were prepared from each plate for isolation of the plasmid DNA using a QIAprep-spin miniprep kit (QIAGEN, Inc.). The isolated plasmid DNA was analyzed using 0.75% DNA agarose gel electrophoresis to screen for sequencing. The **pIL5** plasmid was run on the gel as a control. The selected ligations were sent for automated DNA sequencing analysis at GENEWIZ LLC (South Plainfield, NJ) to confirm the desired sequence. The pQE-F and pQE-R primers, which were designed to bind upstream and downstream of the multiple cloning site, were sent with the plasmid DNA for analysis.

The **pIL5** expression vector with the 3FD-LL adapter gene cloned into its multiple cloning site formed the plasmid **pBB14**. The plasmid sample with the confirmed sequence was transformed into chemically competent Top10F' cells and recovered in 400 μ L SOC rich media at 37 $^{\circ}$ C for 1 hour. The cells were plated onto LB agar containing ampicillin (100 μ g/mL) and incubated at 37 $^{\circ}$ C for 16 hours. Two overnight cultures were prepared to isolate of the plasmid DNA using a QIAprep-spin miniprep kit (QIAGEN, Inc.). Two overnight cultures were grown to generate frozen stocks of the plasmid inside of Top10F' cells by combining 200 μ L of 80% glycerol with 800 μ L of the culture.

Production of the 3FD-LL Monomer Expression Plasmid

An aliquot of 1 μ L of **pBB14** plasmid DNA (3FD-LL adapter inside the modified expression vector **pIL5**) was transformed into chemically competent Top10F' *E. coli* cells. The cells were recovered after heat shock in 400 μ L SOC rich media for 30 minutes at 37 $^{\circ}$ C. Aliquots of 50 μ L of the transformed suspensions were spread onto LB agar plates and incubated for 14

hours at 37 °C. The cells containing plasmid **pBB14** were spread on plates containing ampicillin (100 µg/mL) for antibiotic selection.

Ten colonies containing plasmid **pBB14** were selected from the plates containing the Top 10 F' strain. The selected colonies were used to inoculate culture tubes containing 5 mL LB media supplemented with 100 µg/mL ampicillin. The cultures were grown at 37 °C overnight on a rotator. A QIAprep-spin miniprep kit (QIAGEN, Inc.) was used to isolate the plasmid DNA. The isolated DNA was recovered in 50 µL of EB buffer (10 mM Tris-HCl, pH 8.5) and stored in solution at -30 °C when not in use.

The helix insertion site in the 3FD-LL adapter was opened through restriction digestion of the plasmid **pBB14** with *Bsa* I endonuclease to create complementary sticky ends. A volume of 50 µL plasmid DNA was incubated with 3 µL of *Bsa* I restriction enzyme and the appropriate buffer with a total reaction volume of 100 µL. After incubation at 37 °C overnight, the digestion products were separated by 0.75% DNA agarose gel electrophoresis. The large band, which contained the opened insertion site of the 3FD-LL adapter, was excised from the gel using a Zymoclean™ Gel DNA Recovery Kit and eluted into 8.5 µL of ddH₂O.

The complementary sticky ends of the adapter gene were ligated together to close the plasmid without any insert. The digested 3FD-LL adaptor of **pBB14** was incubated overnight at 16 °C in a 10 µL reaction volume with 0.5 µL T4DNA Ligase (200 units) and 1X T4DNA Ligase buffer. A 1 µL aliquot of the ligation product was transformed into the chemically competent Top10F' strain. The cells were recovered in 400 µL SOC rich media at 37 °C for 1 hour, plated onto LB agar containing ampicillin (100 µg/mL), and incubated at 37 °C for 16 hours. Six overnight cultures were prepared from each plate for isolation of the plasmid DNA using a

QIAprep-spin miniprep kit (QIAGEN, Inc.). The isolated plasmid DNA was analyzed using 0.75% DNA agarose gel electrophoresis to screen for sequencing. The **pBB14** plasmid was run on the gel as a control. The selected ligations were sent for automated DNA sequencing analysis at GENEWIZ LLC (South Plainfield, NJ) to confirm the desired sequence. The pQE-F and pQE-R primers were sent with the plasmid DNA for analysis.

With the helix insertion site removed from the 3FD-LL adapter, the **3FD-LL monomer** in the expression vector was formed and named **pBB15**. The plasmid sample with the confirmed sequence was transformed into chemically competent Top10F' cells and recovered in 400 μ L SOC rich media at 37 °C for 1 hour. The cells were plated onto LB agar containing ampicillin (100 μ g/mL) and incubated at 37 °C for 16 hours. Two overnight cultures were prepared to isolate of the plasmid DNA using a QIAprep-spin miniprep kit (QIAGEN, Inc.). Two overnight cultures were grown to generate frozen stocks of the plasmid inside of Top10F' cells by combining 200 μ L of 80% glycerol with 800 μ L of the culture.

Concatemerization of 3FD-LL Gene Sequence

A 3FD-LL helix insert gene containing the codon optimized synthetic DNA monomer encoding the 3FD-LL sequence, E(ALEKLA)₃(ALKELA)₃K, was ordered from ATUM (formerly DNA 2.0) in the pJ201 vector for insertion into the 3FD-LL adapter insertion site. The 3FD-LL helix insert plasmid was named **pBB13**. The lyophilized powder containing the plasmid was resuspended in 50 μ L of distilled, deionized water (ddH₂O) upon arrival. Aliquots of 1 μ L were used in transformations of the chemically competent Top10F' *E. coli* strain. The cells were recovered after heat shock in 400 μ L SOC rich media for 30 minutes at 37 °C. Aliquots of 50 μ L

of the transformed suspensions were spread onto LB agar plates containing kanamycin (50 µg/mL) for antibiotic selection. The plates were incubated for 14 hours at 37 °C.

Twelve colonies were selected from the plate containing the Top 10 F' strain. The selected colonies were used to inoculate culture tubes containing 5 mL LB media supplemented with kanamycin (50 µg/mL). The cultures were grown at 37 °C overnight on a rotator. Two of the cultures containing the Top10F' strain were prepared for long term storage of the plasmid contained within the *E. coli* strain. These frozen stocks were prepared by combining 200 µL of 80% glycerol with 800 µL of the culture. The frozen stocks of the Top10F' strains containing the **pBB13** plasmid were stored at -80 °C until needed. To make secondary stocks of the **pBB13** plasmid DNA, a QIAprep-spin miniprep kit (QIAGEN, Inc.) was used to isolate the plasmid DNA of each of the remaining cultures of Top10F' cells. The isolated DNA was recovered in 50 µL of EB buffer (10 mM Tris-HCl, pH 8.5) and stored in solution at -30 °C when not in use.

The synthetic 3FD-LL helix insert gene was isolated from the **pBB13** plasmid through sequential restriction digestion with *Bbs* I and *BsmB* I endonucleases. A volume of 50 µL plasmid DNA was incubated with 4 µL of *Bbs* I restriction enzyme and the appropriate buffer with a total reaction volume of 100 µL. After incubation at 37 °C overnight, the digestion product was separated by 1.5% DNA agarose gel electrophoresis. The only band, which contained the digested **pBB13** plasmid, was excised from the gel using a Zymoclean™ Gel DNA Recovery Kit and eluted into 34 µL of ddH₂O. The isolated digestion product was incubated with 2 µL of *BsmB* I restriction enzyme and the appropriate buffer with a total reaction volume of 40 µL. After incubation at 55 °C for 5 hours, the digestion products were separated by 0.75% DNA agarose gel electrophoresis.

The small band, which contained the 3FD-LL helix insert, was excised from the gel using a Zymoclean™ Gel DNA Recovery Kit and eluted into 10 µL of ddH₂O.

The helix insertion site in the 3FD-LL adapter was opened through restriction digestion of the plasmid **pBB14** with *Bsa* I endonuclease to create complementary sticky ends. A volume of 50 µL plasmid DNA was incubated with 3 µL of *Bsa* I restriction enzyme and the appropriate buffer with a total reaction volume of 100 µL. After incubation at 37 °C overnight, the digestion products were separated by 0.75% DNA agarose gel electrophoresis. The large band, which contained the opened insertion site of the 3FD-LL adapter, was excised from the gel using a Zymoclean™ Gel DNA Recovery Kit and eluted into 26 µL of ddH₂O. The digested **pBB14** DNA was dephosphorylated to prevent ligation of the plasmid with itself. The isolated digestion product was incubated with 1 µL of Antarctic Phosphatase and 1X phosphatase buffer for 2 hours at 37 °C. A DNA Clean & Concentrator™ kit was used to isolate the dephosphorylated **pBB14** digestion, and the gene fragment digested plasmid was eluted into 10 µL ddH₂O.

The 3FD-LL helix insert gene was cloned into the digested 3FD-LL adapter insertion site to form the **3FD-LL concatemers** through ligation of the isolated digestion products. The digested 3FD-LL helix insert and 3FD-LL adaptor plasmid were combined together in 3:1, 5:1, and 7:1 molar ratios. The gene fragments were incubated overnight at 16 °C in 10 µL reaction volumes with 0.5 µL T4DNA Ligase (200 units) and 1X T4DNA Ligase buffer. A 1 µL aliquot of each ligation product was transformed into chemically competent Top10F' cells. The cells were recovered in 400 µL SOC rich media at 37 °C for 1 hour, plated onto LB agar containing ampicillin (100 µg/mL), and incubated at 37 °C for 16 hours. Twelve overnight cultures containing ampicillin were prepared from each plate for isolation of the plasmid DNA using a QIAprep-spin miniprep

kit (QIAGEN, Inc.). The isolated plasmid DNA was analyzed using 0.75% DNA agarose gel electrophoresis to screen for sequencing. The **pBB14** plasmid was run on the gel as a control. The selected ligations were sent for automated DNA sequencing analysis at GENEWIZ LLC (South Plainfield, NJ) to confirm the desired sequence. A pQE-80 forward primer was sent with the plasmid DNA for analysis.

Various **3FD-LL concatemers** within the **pIL5** expression vector were created by this method. Plasmids containing 2, 3, 4, and 5 3FD-LL helical units were created. The **3FD-LL dimer** in the expression vector was named **pBB17**, the **3FD-LL trimer** in the expression vector was named **pBB20**, the **3FD-LL tetramer** in the expression vector was named **pBB21**, and the **3FD-LL pentamer** in the expression vector was named **pBB22**. **Figure 2.19** displays the sequence verified plasmids for the 3FD-LL concatemers using DNA agarose gel electrophoresis.

The plasmid DNA samples with the confirmed sequences of the **3FD-LL concatemer** genes were transformed into chemically competent Top10F' cells and recovered in 400 μ L SOC rich media at 37 °C for 1 hour. The cells were plated onto LB agar containing ampicillin (100 μ g/mL) and incubated at 37 °C for 16 hours. Two overnight cultures were prepared for each plasmid to isolate of the DNA using a QIAprep-spin miniprep kit (QIAGEN, Inc.). Two overnight cultures were grown for each plasmid to generate frozen stocks of the plasmid inside of Top10F' cells by combining 200 μ L of 80% glycerol with 800 μ L of the culture.

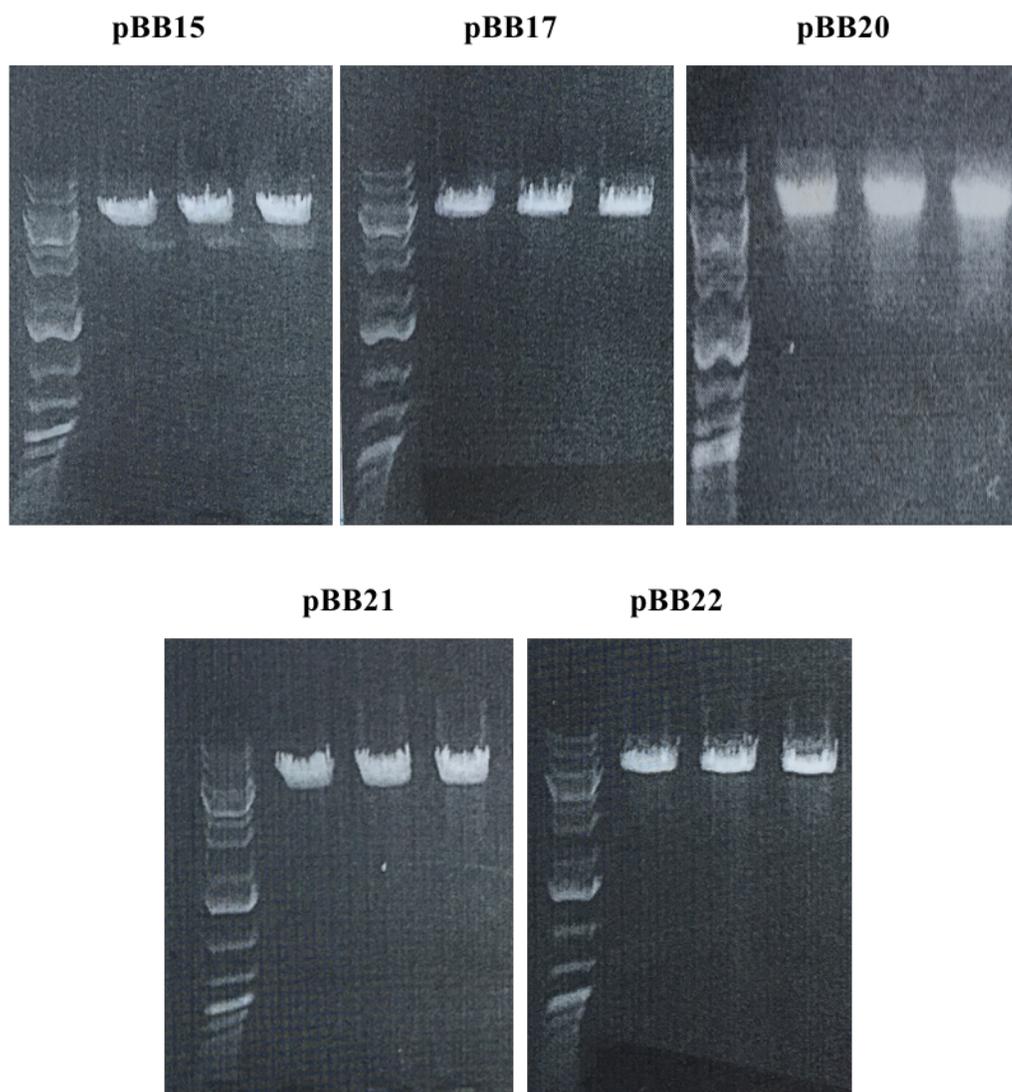


Figure 2.19 DNA agarose gel electrophoresis of sequence verified plasmids pBB15 (top left), pBB17 (top middle), pBB20 (top right), pBB21 (bottom left), and pBB22 (bottom right). The O'GeneRuler™ 1 kb Plus DNA ladder was added to the first lane of each gel. Each plasmid was run in triplicate. Plasmid size: pBB15 = 4842 bp; pBB17 = 4950 bp; pBB20 = 5058 bp; pBB21 = 5166 bp; pBB22 = 5274 bp.

2.4.4 Production of the 3FD-LL Concatemer Proteins

Gene and Bacterial Preparation

The genes encoding the **3FD-LL concatemers** were designed to be codon optimized for expression in *E. coli*. The plasmid DNA for each of the genes was previously isolated and purified using a QIAprep-spin miniprep kit (QIAGEN, Inc.) and stored at -20 °C. Aliquots of 1 µL of the miniprep plasmid DNA (**pBB15**, **pBB17**, **pBB20**, **pBB21**, and **pBB22**) were used in transformations of the chemically competent *E. coli* strain BL21 (DE3). The cells were recovered after heat shock in 400 µL SOC rich media for 30 minutes at 37 °C. Aliquots of 50 µL of the transformed suspensions were spread onto LB agar plates containing ampicillin (100 µg/mL) for antibiotic selection. The plates were incubated for 14 hours at 37 °C.

Two colonies were selected for each plasmid from the plates containing the BL21 strain. The selected colonies were used to inoculate culture tubes containing 5 mL LB media supplemented with ampicillin (100 µg/mL). The cultures were grown at 37 °C overnight on a rotator. The cultures containing the BL21 strain were prepared for long term storage of each plasmid contained within the *E. coli* strain. These frozen stocks were prepared by combining 200 µL of 80% glycerol with 800 µL of the culture. Frozen stocks of the BL21 strain containing the plasmids for the five **3FD-LL concatemers** were stored at -80 °C until needed.

Small-Scale Expression

Each plasmid of the **3FD-LL concatemers** was transformed into the BL21 (DE3) strain and plated onto LB agar plates containing ampicillin (100 µg/mL). Single colonies of BL21 strain containing the **3FD-LL concatemer** plasmids were used to inoculate 5 mL of LB broth

supplemented with ampicillin (100 µg/mL). The cell cultures were grown overnight on a rotator at 37 °C. Volumes of 500 µL of the overnight cultures were transferred to 250 mL Erlenmeyer flasks containing 50 mL of Terrific Broth (TB) media supplemented with ampicillin (100 µg/mL).

The small culture flasks shook at 200 rpm and 37 °C for approximately 3 hours until cell growth reached log phase growth with an OD₆₀₀ of ~0.6-0.8. The OD₆₀₀ was monitored hourly. Once log phase growth was achieved, expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG). For each 50 mL volume, 25 µL of 1 M IPTG was added for a final IPTG concentration of 0.5 mM. The expression cultures were incubated at 30 °C with shaking for 18 hours. Aliquots were removed from the expression culture throughout the process for SDS-PAGE analysis (see below). Eighteen hours after induction, the cells in the expression cultures were harvested by centrifugation at 4,000 x g for 20 minutes (4 °C). The cell pellets were resuspended in 5 mL lysis buffer (50 mM Tris, 100 mM NaCl, pH 8.0) and stored at -80 °C.

One mL aliquots were removed from each flask before induction and 1, 4, and 18 hours after induction. The OD₆₀₀ was determined for each aliquot and the aliquot volumes were normalized based on their OD so that the number of cells in each aliquot were equivalent. The normalized aliquot volumes were centrifuged for 5 minutes at 4,000 x g to pellet the cells. The supernatants were discarded, the pellets were resuspended in 50 µL ddH₂O, and the expression samples were stored at -30 °C. Aliquoted expression samples were later prepared for SDS-PAGE analysis by mixing the 50 µL samples with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. The expression samples were analyzed by 16% SDS-PAGE and western blotting according to the protocols described in the general methods section.

Large-Scale Expression

The plasmids of the **3FD-LL concatemers** were transformed into the BL21 (DE3) strain and plated onto LB agar plates containing ampicillin (100 µg/mL). Single colonies of BL21 strain containing individual **3FD-LL concatemer** plasmids were used to inoculate 5 mL cultures of LB broth supplemented with ampicillin (100 µg/mL). The cell cultures were grown overnight on a rotator at 37 °C. Five mL of the overnight cultures were transferred to 2,800 mL Erlenmeyer flasks containing 500 mL of Terrific Broth (TB) media supplemented with 100 µg/mL ampicillin. Four flasks were used for a total expression culture volume of 2000 mL (2L).

The large culture flasks shook at 200 rpm and 37 °C for approximately 3 hours until cell growth reached log phase growth with an OD₆₀₀ of ~0.6-0.8. The OD₆₀₀ was monitored hourly. Once log phase growth was achieved, expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG). For each 500 mL flask, 250 µL of 1 M IPTG was added for a final IPTG concentration of 0.5 mM. The expression cultures were incubated at 30 °C with shaking for 18 hours. Aliquots were removed from the expression culture throughout the process for SDS-PAGE analysis (see below). Eighteen hours after induction, the cells in the expression cultures were harvested by centrifugation at 4,000 x g for 20 minutes (4 °C). The supernatant was discarded, and the cell pellets were stored at -80 °C.

One mL aliquots were removed from each flask before induction and 1, 4, and 18 hours after induction. The OD₆₀₀ was determined for each aliquot and the aliquot volumes were normalized based on their OD so that the number of cells in each aliquot would be equivalent. The normalized aliquot volumes were centrifuged for 5 minutes at 4,000 x g to pellet the cells. The supernatants were discarded, the pellets were resuspended in 50 µL ddH₂O, and the expression

samples were stored at -30 °C. Aliquoted expression samples were later prepared for SDS-PAGE analysis by mixing the 50 µL samples with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. The expression samples were run on 16% SDS-PAGE gels according to the protocol described in the general methods section. Samples were also analyzed by western blot.

Cell Lysis

The cell pellets from large-scale expressions of the **3FD-LL concatemers** in the BL21 strain were previously frozen at -80 °C. The cell pellets were resuspended in 100 mL lysis buffer (50 mM Tris, 100 mM NaCl, pH 8.0) and lysis of the cell pellets was initiated by three freeze/thaw cycles (-80 °C; 25 °C). After the third thaw cycle, EDTA-free protease inhibitor cocktail (1X), lysozyme (1 mg/mL), benzonase nuclease (25 units/mL), and MgCl₂ (1 mM) were added to the lysate. The lysates were incubated at 30 °C for 1 hour with shaking at 200 rpm and then the temperature was dropped to 4 °C for incubation overnight. The cell lysates were sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. Aliquots of 50 µL were collected for SDS-PAGE analysis (see below).

The cell lysate aliquots were centrifuged at 10,000 rpm for 5 minutes. The aliquot supernatants were transferred into clean microcentrifuge tubes and the pellets were resuspended in 50 µL of ddH₂O. The lysed supernatant and lysed pellet samples were prepared for SDS-PAGE analysis by mixing the 50 µL samples with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. Analysis by 16% SDS-PAGE

concluded that all five of the **3FD-LL concatemer** proteins were located in the lysed pellet (insoluble fraction).

A sample of lysozyme was prepared as a control for SDS-PAGE analysis. The lysozyme control was prepared for SDS-PAGE analysis by mixing 20 μL of 50 mg/mL lysozyme stock with 20 μL of dH_2O to make 40 μL of a 25 mg/mL lysozyme solution. The lysozyme solution was combined with 40 μL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 $^\circ\text{C}$ for 5 minutes.

Solubilization with Urea

The cell lysate samples were analyzed by SDS-PAGE, which determined the **3FD-LL concatemers** were insoluble in the cell lysate solution. Urea was added to the cell lysate for a final concentration of 6 M. The 6 M urea solution was allowed to shake at 250 rpm for 24 hours at 4 $^\circ\text{C}$. Aliquots of 50 μL were collected for SDS-PAGE analysis (see below). The remaining urea solution was centrifuged at 15,000 rpm for 30 minutes (4 $^\circ\text{C}$). The urea supernatants were separated from the urea pellets until analysis by SDS-PAGE determined the location of the **3FD-LL concatemer** proteins.

Meanwhile, the urea aliquots were centrifuged at 15,000 rpm for 5 minutes. The aliquot supernatants were transferred into clean microcentrifuge tubes and the pellets were resuspended in 50 μL of ddH_2O . The urea supernatant and urea pellet samples were prepared for SDS-PAGE analysis by mixing the 50 μL samples with 50 μL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 $^\circ\text{C}$ for 5 minutes. Analysis by 16% SDS-PAGE

concluded that the **3FD-LL concatemer** proteins were insoluble in a 6 M urea solution and entirely located in the pellets.

Inclusion Body Washes

The cell pellets from large-scale expressions of the **3FD-LL concatemers** in the BL21 strain were previously frozen at -80 °C. The cell pellets were resuspended in 100 mL lysis buffer A (50 mM Tris, 100 mM NaCl, 5% glycerol, 0.1 mM DTT, 0.1 mM EDTA, pH 8.0) and lysis of the cell pellets was initiated by three freeze/thaw cycles (-80 °C; 25 °C). After the third thaw cycle, EDTA-free protease inhibitor cocktail (1X), lysozyme (1 mg/mL), benzonase nuclease (25 units/mL), and MgCl₂ (1 mM) were added to the lysate. The lysates were incubated at 30 °C for 1 hour with shaking at 200 rpm and then the temperature was dropped to 4 °C for incubation overnight. The cell lysates were sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. Aliquots of 50 µL were collected for SDS-PAGE analysis (see below).

A 1 mL volume of Triton X-100 was added to the cell lysate for a final concentration of 1%. The solution was allowed to shake for 1 hour at 4 °C. The Triton X-100 solution was centrifuged at 12,000 rpm for 20 minutes (4 °C). The inclusion bodies in the pellet were resuspended in lysis buffer B (50 mM Tris, 100 mM NaCl, 5% glycerol, 0.1 mM DTT, 1% Triton X-100, pH 8.0) and the solution shook for 2 hours at 4 °C. The cell lysates were sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. The solution was centrifuged at 15,000 rpm for 20 minutes (4 °C).

The inclusion bodies in the pellet were then resuspended in lysis buffer C (50 mM Tris, 100 mM NaCl, 5% glycerol, 0.1 mM DTT, pH 8.0) and the solution shook for 1 hour at 4 °C. The

cell lysates were sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. And then centrifuged at 12,000 rpm for 20 minutes (4 °C).

The cell lysate and washed inclusion body aliquots were centrifuged at 10,000 rpm for 5 minutes. The aliquot supernatants were transferred into clean microcentrifuge tubes and the pellets were resuspended in 50 μ L of ddH₂O. The supernatant and pellet samples were prepared for SDS-PAGE analysis by mixing the 50 μ L samples with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. Analysis by 16% SDS-PAGE concluded that all five of the **3FD-LL concatemer** proteins were located in the pellets of the cell lysate and the washed inclusion bodies.

Solubilization with Guanidine Hydrochloride

After expression and cell lysis, cell lysates were centrifuged at 10,000 rpm for 20 minutes (4 °C) and the lysed supernatants were discarded. The insoluble **3FD-LL concatemer** proteins in the cell lysate pellets were solubilized in guanidine hydrochloride to allow for purification on a HisPurTM cobalt column. In later experiments, the washed inclusion body pellets were solubilized in guanidine hydrochloride. In order to do this, the pellets were resuspended in 200 mL total volume of GHCl solubilization buffer (7 M guanidine hydrochloride, 50 mM Tris, pH 8.0). The insoluble pellets were allowed to shake at 250 rpm for 3 days at 4 °C. Aliquots of 50 μ L were collected for SDS-PAGE analysis (see below). The remaining GHCl solutions were centrifuged at 15,000 rpm for 30 minutes (4 °C). The GHCl supernatants were separated from the GHCl pellets until analysis by SDS-PAGE determined the location of the **3FD-LL concatemer** proteins.

Meanwhile, the GHCl aliquots were centrifuged at 15,000 rpm for 5 minutes. The aliquot supernatants were transferred into clean microcentrifuge tubes and the pellets were resuspended in 50 μ L of ddH₂O. The GHCl supernatant and GHCl pellet samples were prepared for SDS-PAGE analysis by mixing the 50 μ L samples with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. Analysis by 16% SDS-PAGE concluded that the **3FD-LL concatemer** proteins were at least partially soluble in 7M GHCl and primarily located in the supernatant.

Purification

The presence of a decahistidine tag at the N-terminus of the **3FD-LL concatemer** proteins facilitated purification by immobilized metal affinity chromatography (IMAC). The soluble fractions recovered after protein solubilization in GHCl were each loaded directly onto 12 mL pre-equilibrated HisPurTM cobalt resin divided among four columns. The resin was then washed with 60 mL GHCl wash buffer (7 M guanidine hydrochloride, 50 mM Tris, 20 mM imidazole, pH 8.0). The target proteins were isolated by the addition of 60 mL elution buffer (7 M guanidine hydrochloride, 50 mM Tris, 250 mM imidazole, pH 7.0). A 50 μ L aliquot of the eluent was collected and prepared for SDS-PAGE analysis by mixing the 50 μ L samples with 50 μ L of 2X SDS samples loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. The pure eluent samples were run on 16% SDS-PAGE gels according to the protocol described in the general methods section. The samples were also analyzed by western blotting.

The purification eluents were loaded into a SnakeSkin™ dialysis tubing and dialyzed against ddH₂O for 5 days (4 L volume, switching buffer every 8 hours). The **3FD-LL monomer** and **3FD-LL dimer** were loaded into dialysis tubing with a 3,500 Da MWCO, while the **3FD-LL trimer**, **3FD-LL tetramer**, and **3FD-LL pentamer** were loaded into dialysis tubing with a 10,000 Da MWCO. After dialysis, the protein solutions were transferred from the dialysis cassettes to 50 mL conical tubes and stored at -80 °C. Once frozen, the tops of the conical tubes were replaced with Kimwipes™ and the tubes were placed in a lyophilizer until all that remained was the pure **3FD-LL concatemer** proteins in powder form. The lyophilized powder was weighed and stored at -30 °C. Protein yields were recorded as dry weight per unit volume of culture.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS)

MALDI-TOF experiments were performed on an Applied Biosystems® Voyager™ System 4700 mass spectrometer (Life Technologies Corporation; Carlsbad, CA) in the high mass positive ion linear mode. The matrix, α -Cyano-4-hydroxycinnamic acid (CHCA), was used at a concentration of 10 mg/mL in a mixture of 50% acetonitrile and 0.1% trifluoroacetic acid in deionized water. The **3FD-LL concatemer** protein solutions (1 mg/mL in ddH₂O) was mixed with the matrix solution in a 1:1 ratio. Two microliters of the mixture were spotted on a stainless-steel sample plate and dried under vacuum or air. Sample preparation was repeated in triplicate. Mass spectra were acquired from 1000 laser shots at an intensity of 7900.

2.4.5 Tables

Table 2.1 Plasmids used in Chapter 2.

Plasmids	Relevant Characteristics	Reference
pQE-80L	<i>E. coli</i> expression vector, Amp ^R	QIAGEN
pIL5	pQE Adapter in pQE-80L; removal of <i>Nco</i> I and <i>Bsa</i> I, Amp ^R	Reference ³
pBB12	3FD-LL adapter gene in pJ201, Kan ^R	Chapter 2
pBB13	3FD-LL helix insert gene in pJ201, Kan ^R	Chapter 2
pBB14	3FD-LL adapter in pIL5, Amp ^R	Chapter 2
pBB15	3FD-LL monomer in pIL5, Amp ^R	Chapter 2
pBB17	3FD-LL dimer in pIL5, Amp ^R	Chapter 2
pBB20	3FD-LL trimer in pIL5, Amp ^R	Chapter 2
pBB21	3FD-LL tetramer in pIL5, Amp ^R	Chapter 2
pBB22	3FD-LL pentamer in pIL5, Amp ^R	Chapter 2

Table 2.2 *E. coli* strains used in Chapter 2.

Strains	Genotype	Reference
Top10F'	F' [lacIq, Tn10(TetR)] mcrA Δ(mrr-hsdRMS- mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG	Invitrogen
BL21 (DE3)	<i>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ sBamHI</i> <i>ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5</i>	New England Biolabs

Table 2.3 Primers used in Chapter 2 for DNA sequencing.

Primers	Sequence (5'→3')
pQE-F	CCCGAAAAGTGCCACCTG
pQE-R	GGTCATTACTGGAGTCTTG
M13 (-20) Forward	TGTAAAACGACGGCCAGT
M13 Reverse	CAGGAAACAGCTATGAC

Table 2.4 Protein sequences used in Chapter 2.

Protein	Sequence	MW (Da) # of AA
3FD-LL monomer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₁ -K	5659.47 52
3FD-LL dimer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₂ -K	9414.02 88
3FD-LL trimer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₃ -K	13168.56 124
3FD-LL tetramer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₄ -K	16923.1 160
3FD-LL pentamer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₅ -K	20677.65 196

2.5 References

1. Magnotti, E. L. Self-assembly of porous α -helical nanosheets. Emory University, 2016.
2. Magnotti, E. L.; Hughes, S. A.; Dillard, R. S.; Wang, S.; Hough, L.; Karumbamkandathil, A.; Lian, T.; Wall, J. S.; Zuo, X.; Wright, E. R.; Conticello, V. P., Self-Assembly of an α -Helical Peptide into a Crystalline Two-Dimensional Nanoporous Framework. *J. Am. Chem. Soc.* **2016**, *138* (50), 16274-16282.
3. Wu, I. L. Multiple site-specific incorporation of non-canonical amino acids for novel biomaterials design. Emory University, 2013.
4. Tartof, K. D.; Hobbs, C. A., Improved media for growing plasmid and cosmid clones. *Bethesda Research Laboratories Focus* **1987**, *9*, 12-17.
5. Schägger, H.; von Jagow, G., Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry* **1987**, *166* (2), 368-379.
6. Palmer, I.; Wingfield, P. T., Preparation and Extraction of Insoluble (Inclusion-Body) Proteins from Escherichia coli. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **2004**, CHAPTER, Unit-6.3.
7. Sambrook, J.; Russell, D. W.; Sambrook, J.; Russell, D. W., *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press {a} , 10 Skyline Drive, Plainview, NY, 11803-2500, USA: 2001.

Chapter 3

Production of 3FD-LL Concatemer

Nanosheet Assemblies

3.1 Introduction

The **3FD-LL** model peptide readily adopted an α -helical secondary structure with 3-fold screw symmetry in aqueous buffer with 10 mM TAPS at pH 8.5. Two-dimensional nanosheets were formed from assembly of the α -helical units where the measured thickness of the nanosheets exhibited a close correspondence to the calculated length of the α -helix. This indicated that the α -helices aligned in a perpendicular orientation with respect to the nanosheet surface. Although nanosheets formed from the **3FD-IL** and **3FD-LL** model peptides displayed melting temperatures of 84 °C and 90 °C, respectively,¹⁻² the nanosheets exhibited fragmenting during biophysical analysis.

An extension of the **3DF-LL** α -helical unit was attempted to increase the thickness of the material for potential applications. A thicker nanosheet would be useful for incorporation into nanodevices, and it would provide deeper nanopores if utilized in molecular encapsulation or transportation applications. Additionally, increasing the length of the helical unit could improve the mechanical and thermodynamic stability of the material by increasing the number of interactions between neighboring α -helices.

Molecular cloning techniques were utilized in the previous chapter to accomplish the concatemerization of the **3DF-LL** gene sequence. In addition to the **3FD-LL monomer**, protein expression successfully biosynthesized a dimer, trimer, tetramer, and pentamer of the model **3FD-LL** peptide. Despite the recalcitrance of the proteins towards solubilization, the proteins were purified by immobilized metal affinity chromatography. Analysis of the assembly conditions and biophysical characteristics of the 3FD-LL concatemers would indicate similarities to the original **3FD-LL** model peptide and reveal improvements to the novel protein material.

The 3FD-LL concatemer proteins were assembled and significant attempts were made to optimize the assembly conditions. The **3FD-LL trimer**, which proved to be the most successful candidate, was further purified by reverse-phase high performance liquid chromatography. The HPLC purified **3FD-LL trimer** displayed a strong α -helical signature and assembled into highly defined nanosheets on the micrometer-scale. The nanosheets did not display a melting transition across the analyzed temperature range indicating that the **3FD-LL trimer** nanosheets possessed higher thermodynamic stability compared to the model 3FD-IL and **3FD-LL** peptides. Atomic force microscopy height measurements verified that the **3FD-LL trimer** produced nanosheets with greater thicknesses than the **3FD-LL** model peptide with an average height that corresponded closely with the expected length of the α -helical unit; however, a small fraction of the measurements indicated that the protein may be folding during assembly.

The advancements made to the **3FD-LL**-based materials supports the use of these extended two-dimensional biomaterials in a wider range of applications. Based on the porous nature of the model 3FD-IL peptide, concatemerization of the sequences could possibly improve functionality of the scaffolds for applications such as therapeutic delivery. However, more advanced structural analysis of the **3FD-LL trimer** nanosheets would be required for verification of the nanoporous structure.

3.2 Results and Discussion

3.2.1 Initial Assembly and Optimization of 3FD-LL Concatemers

Assembly of HisPurTM 3FD-LL Concatemers

The model **3FD-LL** and **3FD-IL** peptides assembled into sheet-like structures at a peptide concentration of 10 mg/mL in 10 mM TAPS buffer (pH 8.5),¹⁻² so assemblies of the HisPurTM 3FD-LL concatemers were analyzed under these conditions. However, to confirm this was the optimal buffer condition, the HisPurTM 3FD-LL concatemers were also assembled at a peptide concentration of 10 mg/mL under a wide pH range of 4-9. While attempting to dissolve the proteins, it was obvious that the solubility of the 3FD-LL concatemers was very low in water-based solvents, and thermal annealing failed to improve solubility. White precipitate was present under all buffers conditions, especially samples containing the **3FD-LL monomer**.

Transmission electron microscopy (TEM) was used to observe the local structural morphology of the 3FD-LL concatemer proteins. Nascent and thermally annealed assemblies of the HisPurTM 3FD-LL concatemers were analyzed at a protein concentration of 10 mg/mL in 10 mM buffer. The buffer conditions that were analyzed were 10 mM CHES (pH 9.0), 10 mM TAPS (pH 8.5), 10 mM TAPS (pH 8.0), 10 mM MOPS (pH 7.0), 10 mM MES (pH 6.0), 10 mM acetate (pH 5.0), and 10 mM acetate (pH 4.0). All nascent and thermally annealed samples did not exhibit any organized structures at low pH values. The most noteworthy assemblies were observed in 10 mM MOPS buffer (pH 7.0) and 10 mM TAPS buffer (pH 8.5).

The nascent samples of the 3FD-LL concatemers in both 10 mM MOPS buffer (pH 7.0) and 10 mM TAPS buffer (pH 8.5) were composed primarily of fibers or unaggregated peptide.

Fibers appeared as aggregates or as independent units. Aggregates of fibers were formed by lateral association, displayed in **Figure 3.3 a** and **Figure 3.4 a** by the **3FD-LL dimer**, or by fibers wrapping around each other in a rope-like fashion, as shown in **Figure 3.2 a** and **Figure 3.8 a** by the **3FD-LL monomer** and **3FD-LL tetramer**. No sheets were observed in the nascent samples for the **3FD-LL monomer** and **3FD-LL dimer** (**Figures 3.1-3.4 a**), and only a few small, poorly defined sheets were visualized in the nascent **3FD-LL tetramer** and **3FD-LL pentamer** assemblies (**Figures 3.7-3.10 a**). On the other hand, the **3FD-LL trimer** exhibited a variety of sheets in the MOPS and TAPS buffers, but they were only a small percentage of the total population of assemblies (**Figures 3.5-3.6 a**). As stated earlier, fibers were the predominant species in all of the nascent samples.

The thermally annealed samples of the **3FD-LL concatemers** were similar to the nascent samples in that they were largely comprised of fibers; however, the thermally annealed samples contained a slightly larger quantity of sheets than their respective nascent samples. A few small sheets were present in the thermally annealed **3FD-LL monomer**, **3FD-LL dimer**, **3FD-LL tetramer**, and **3FD-LL pentamer** samples (**Figures 3.1-3.4 b** and **3.7-3.10 b**). Many of these small sheets appeared to strongly associate with each other in addition to disorganized aggregated protein. The sheets in the **3FD-LL trimer** displayed edges with superior definition compared to the sheets present in the nascent sample indicating that the thermal annealing process improved the formation of sheets (**Figures 3.5-3.6 b**). For all proteins, more sheets were present in samples at pH 8.5 than the pH 7.0 samples. Although all samples contained at least a couple sheets, this did not compare to the thousands of fibers that were present in every sample. More optimization was needed to direct the proteins to favor the formation of sheets over fibers.

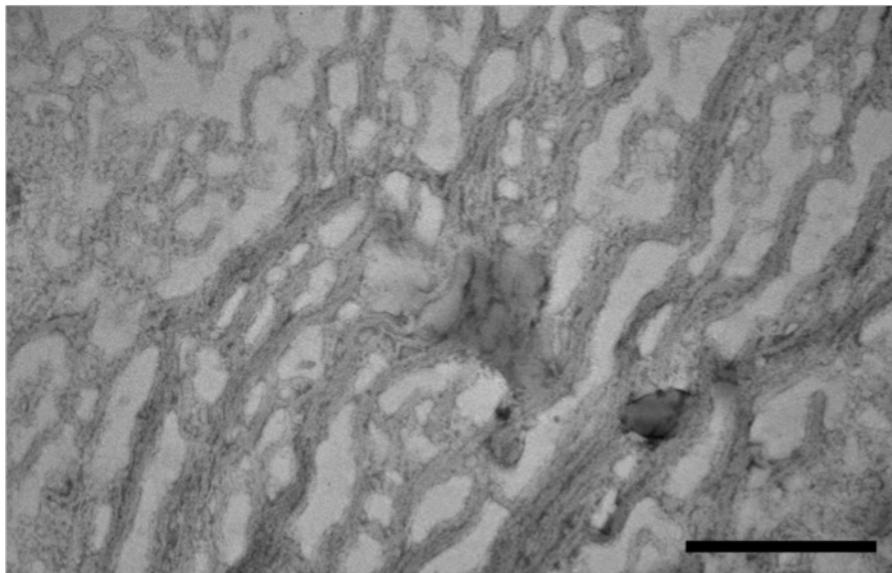
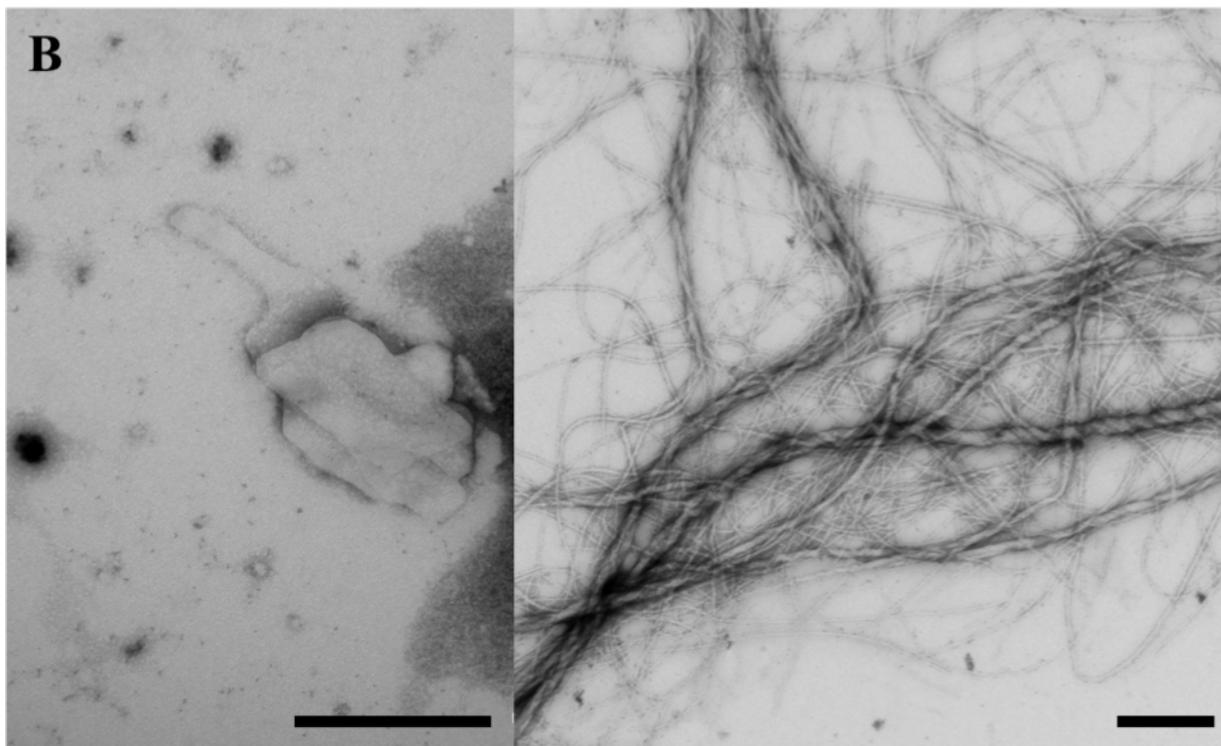
A**B**

Figure 3.1 Transmission electron microscopy images of 10 mg/mL HisPurTM 3FD-LL monomer assemblies in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

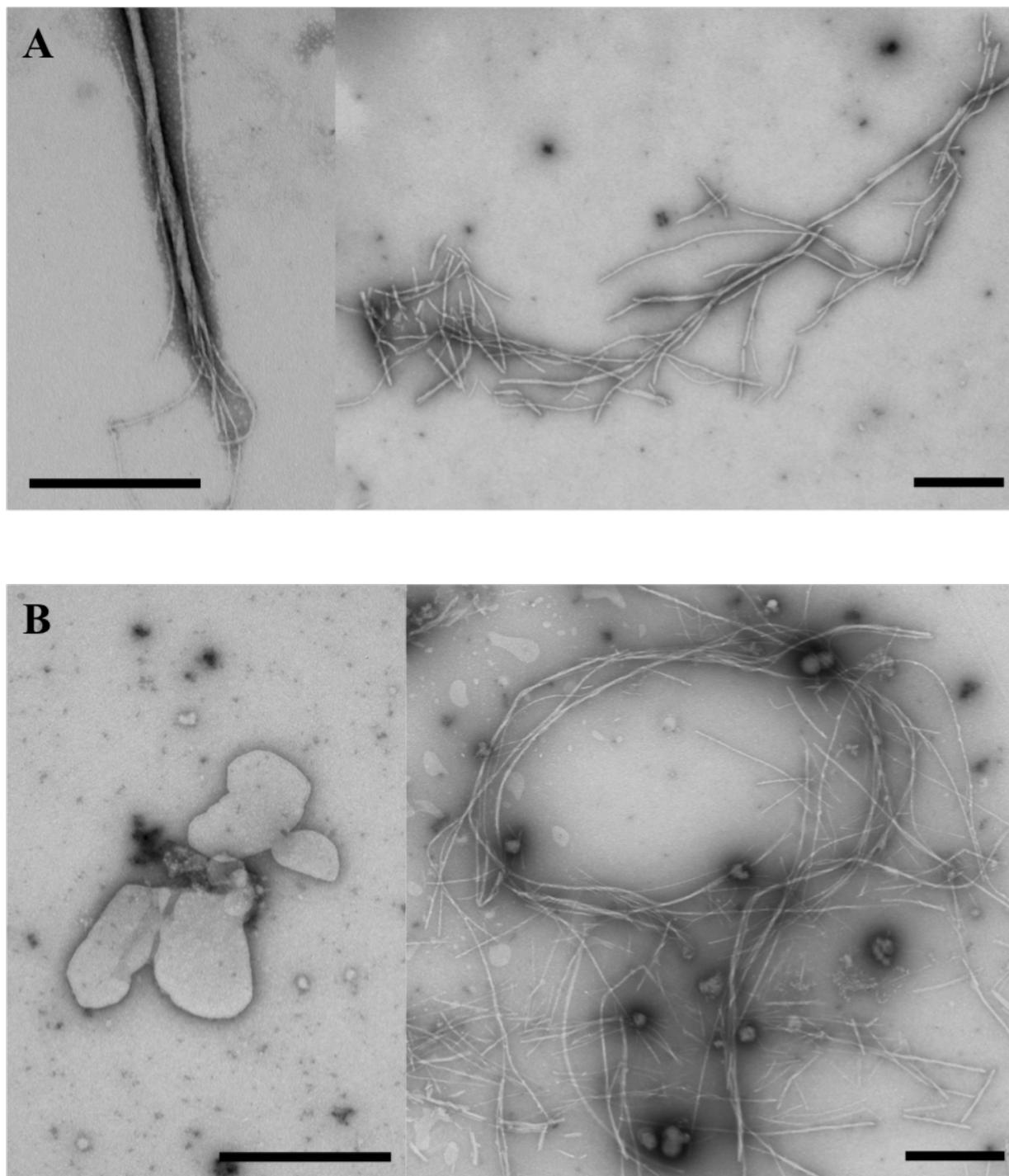


Figure 3.2 Transmission electron microscopy images of 10 mg/mL HisPur[™] 3FD-LL monomer assemblies in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

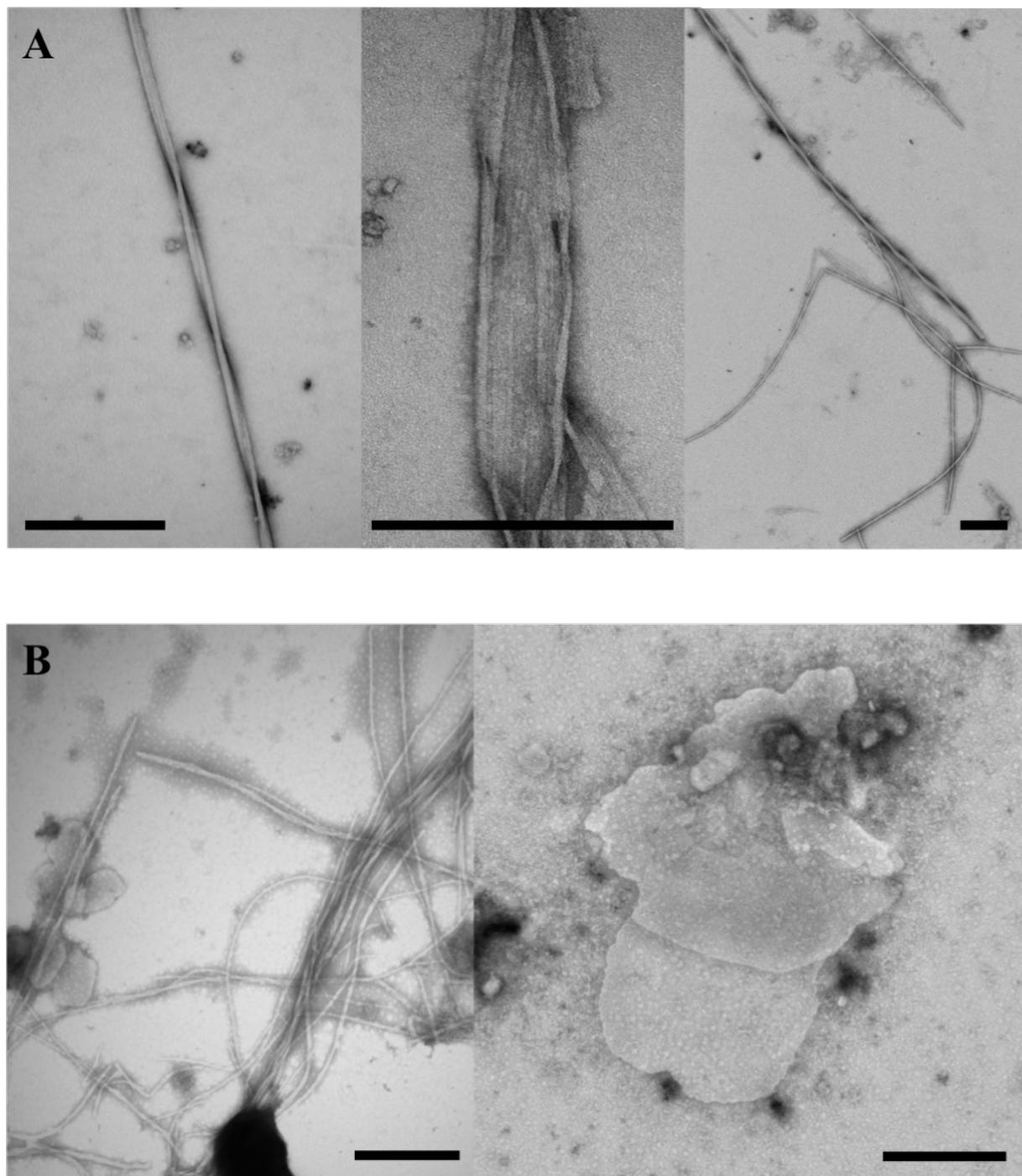


Figure 3.3 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL dimer assemblies in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

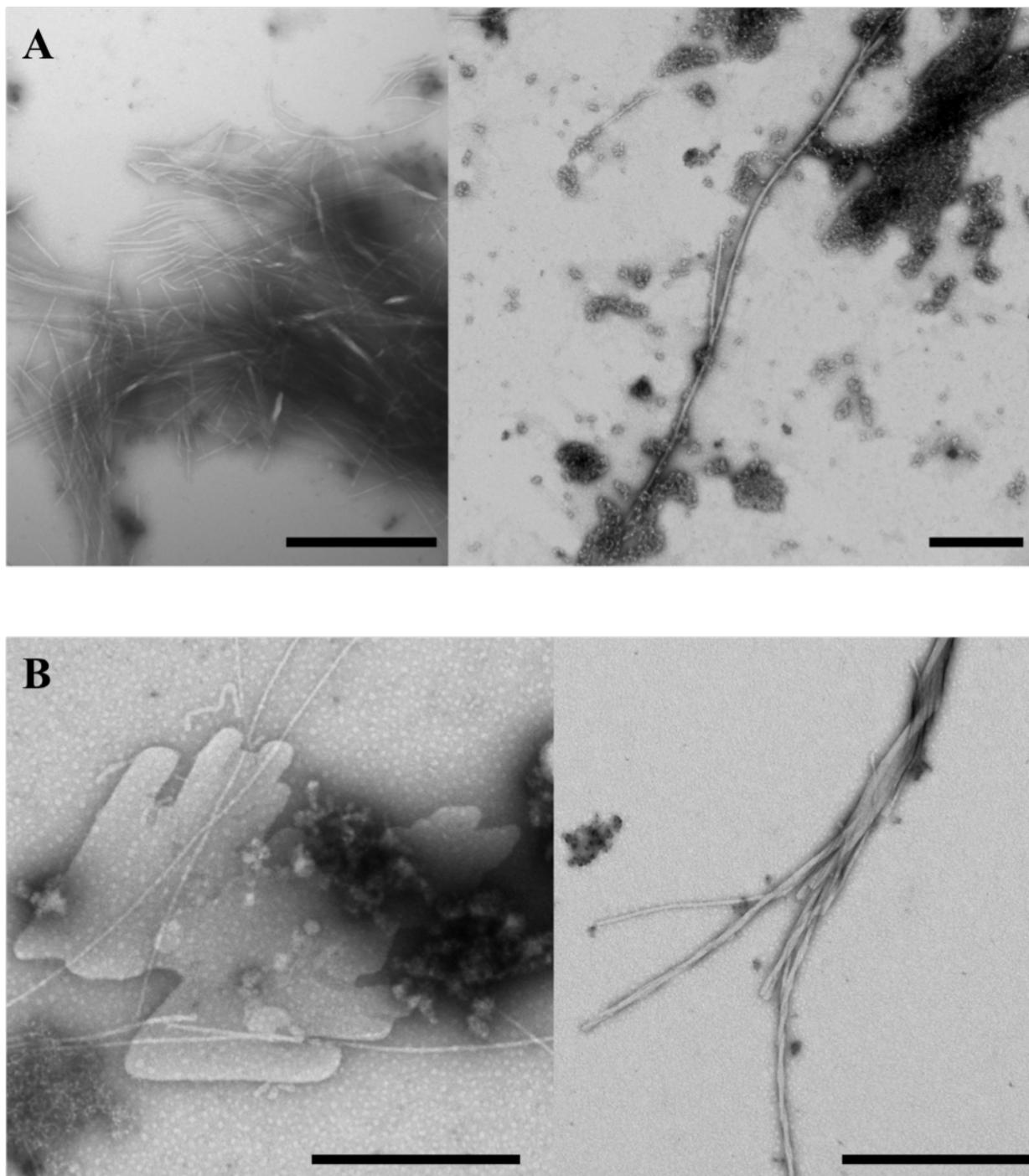


Figure 3.4 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL dimer assemblies in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

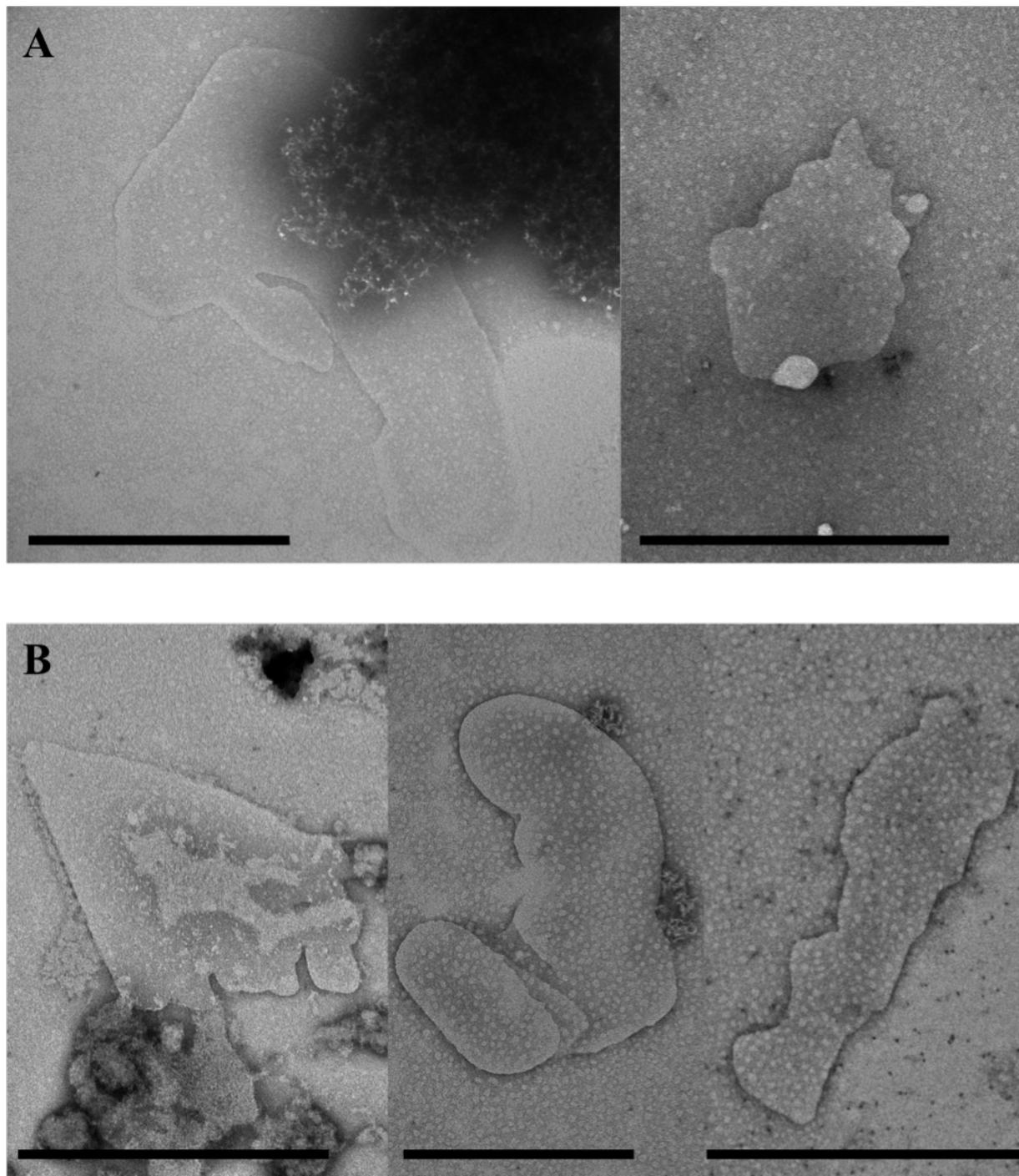


Figure 3.5 Transmission electron microscopy images of 10 mg/mL HisPur[™] 3FD-LL trimer assemblies in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

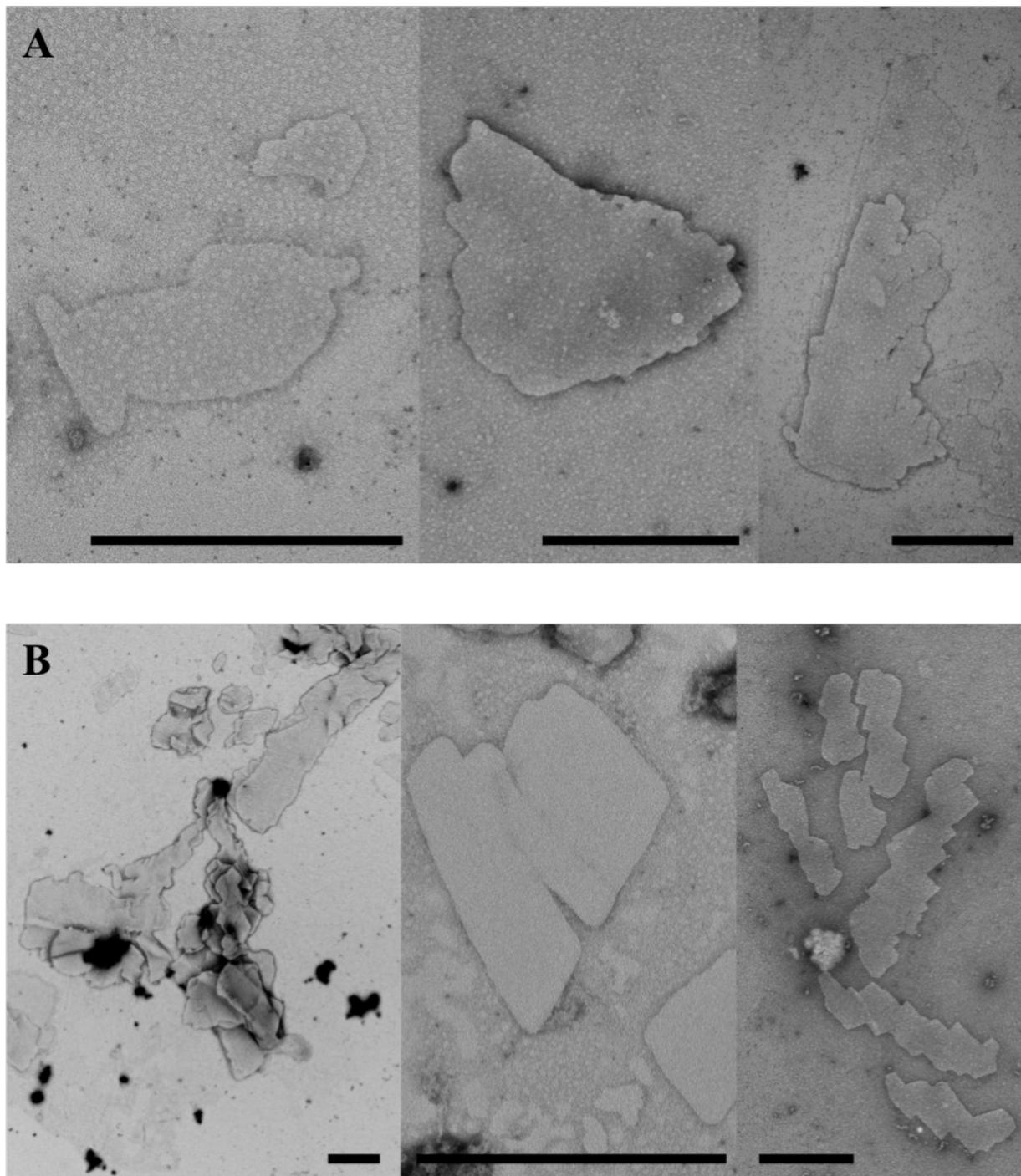


Figure 3.6 Transmission electron microscopy images of 10 mg/mL HisPurTM 3FD-LL trimer assemblies in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

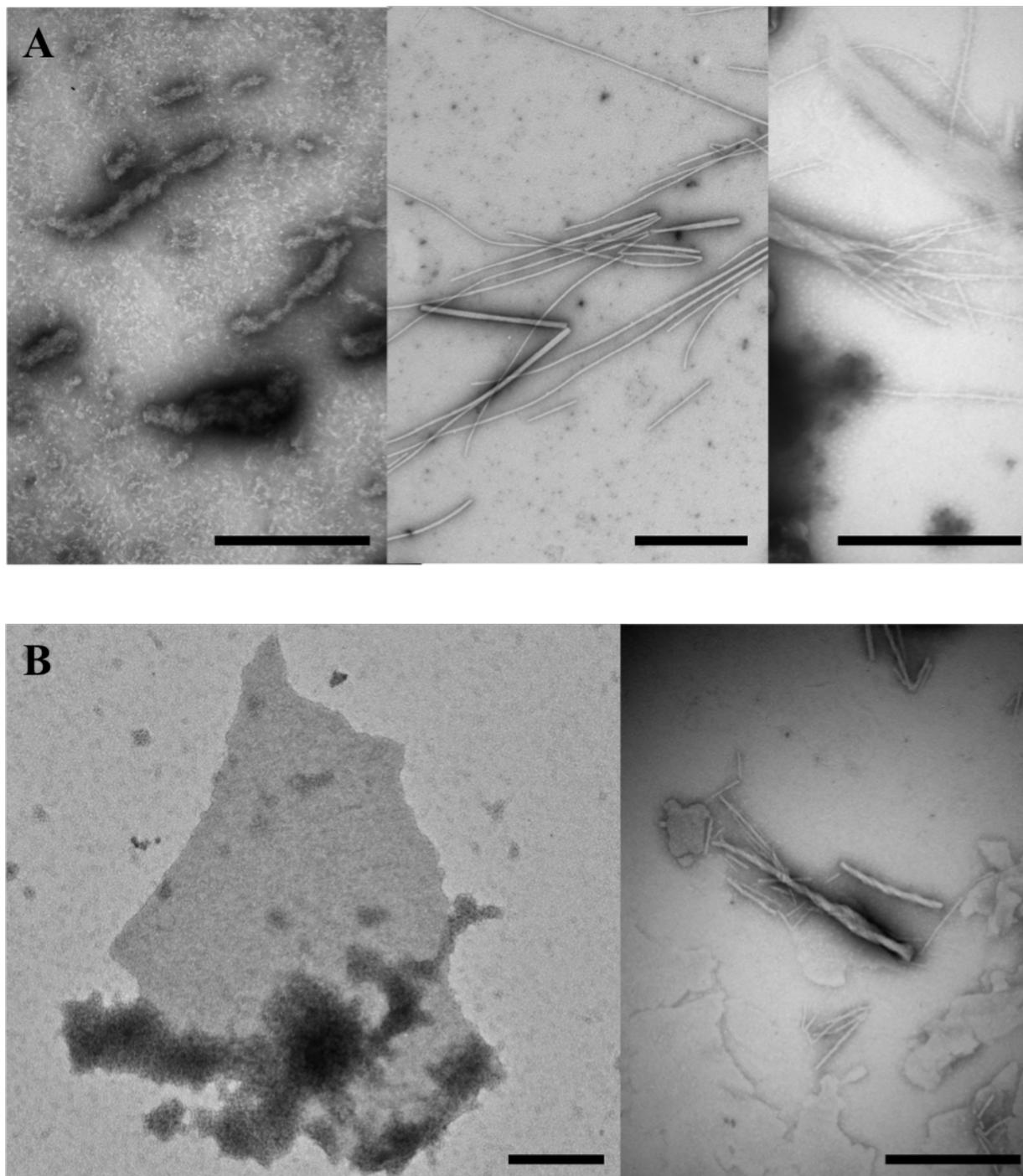


Figure 3.7 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL tetramer assemblies in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

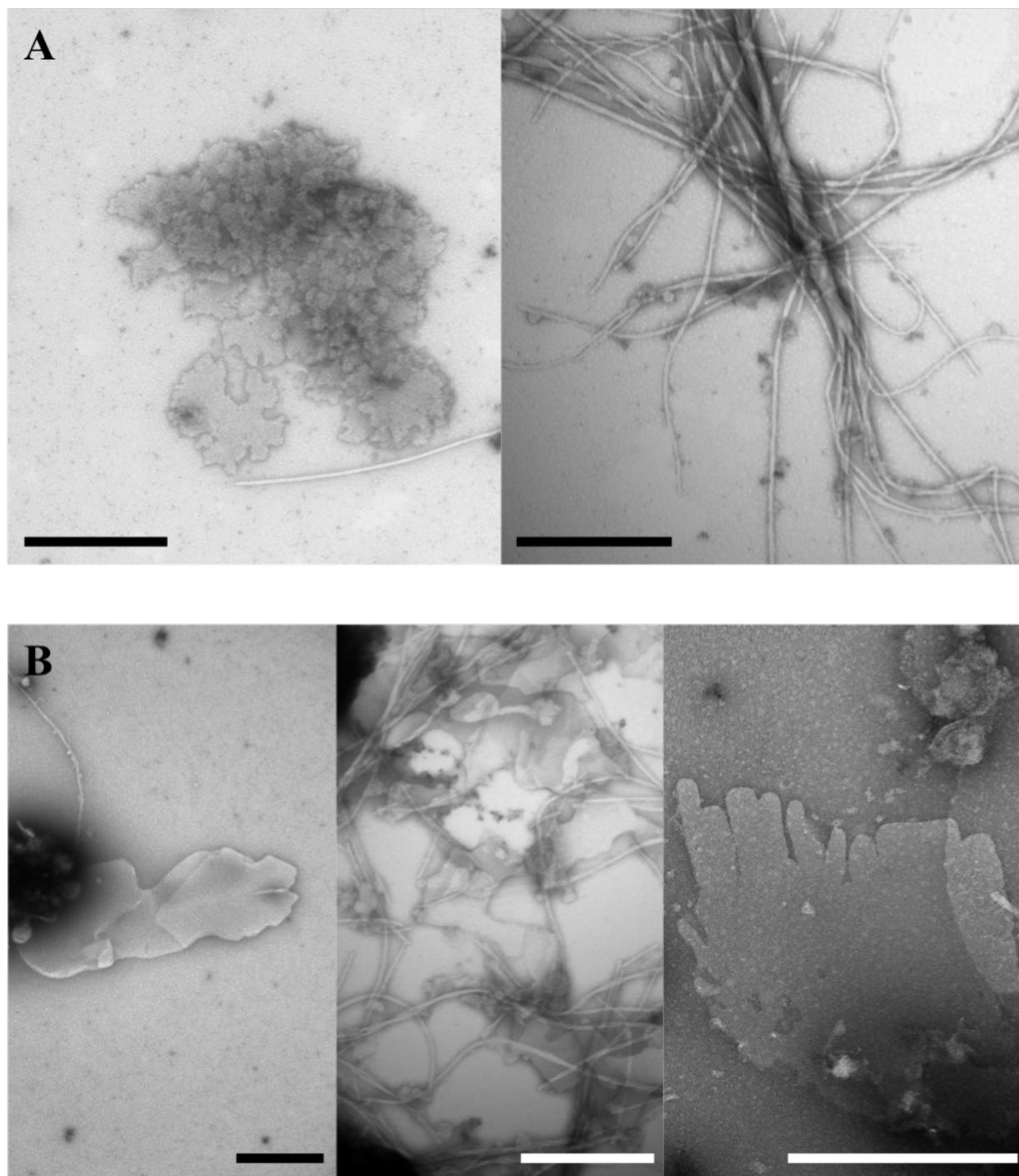


Figure 3.8 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL tetramer assemblies in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

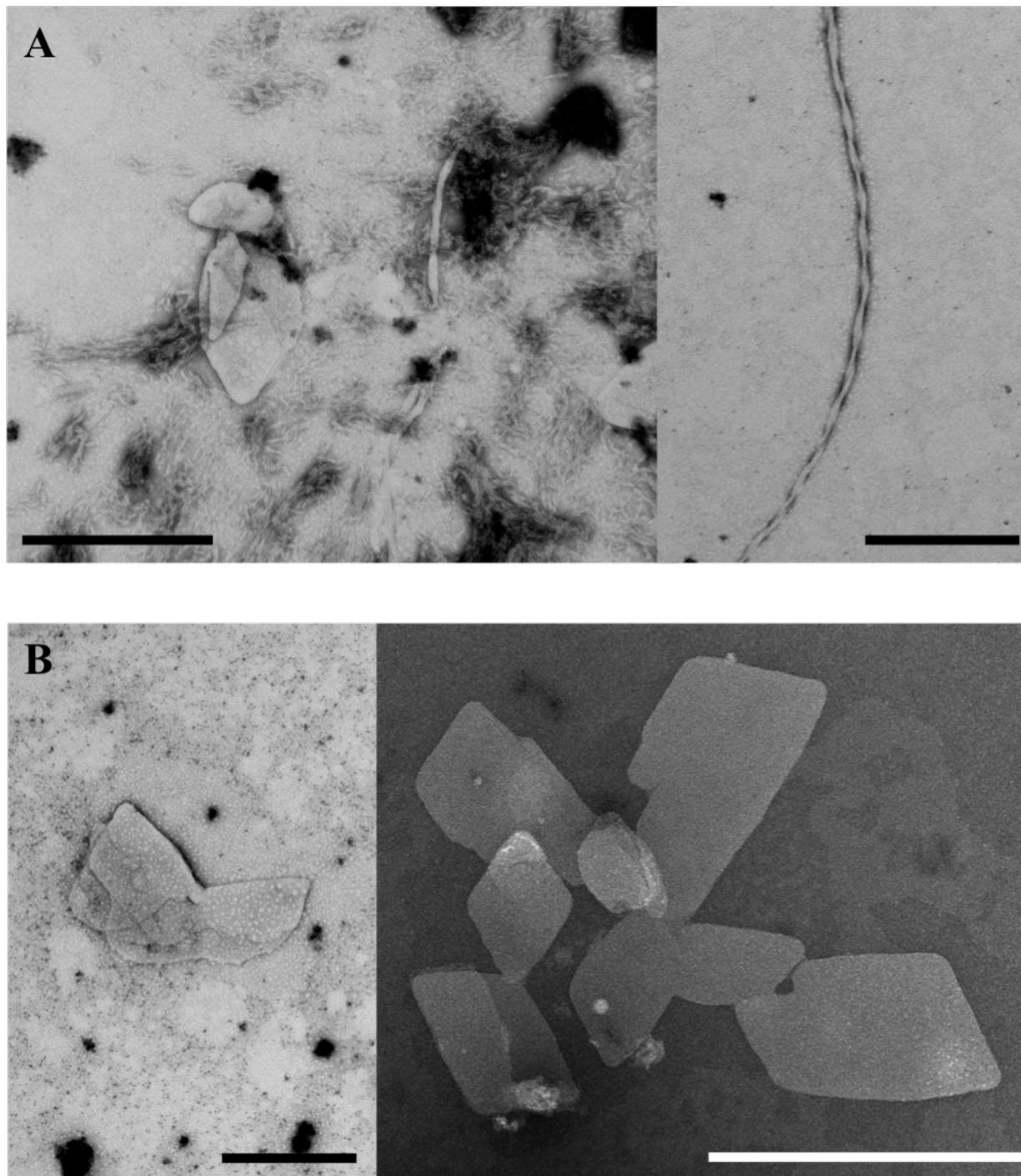


Figure 3.9 Transmission electron microscopy images of 10 mg/mL HisPur[™] 3FD-LL pentamer assemblies in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

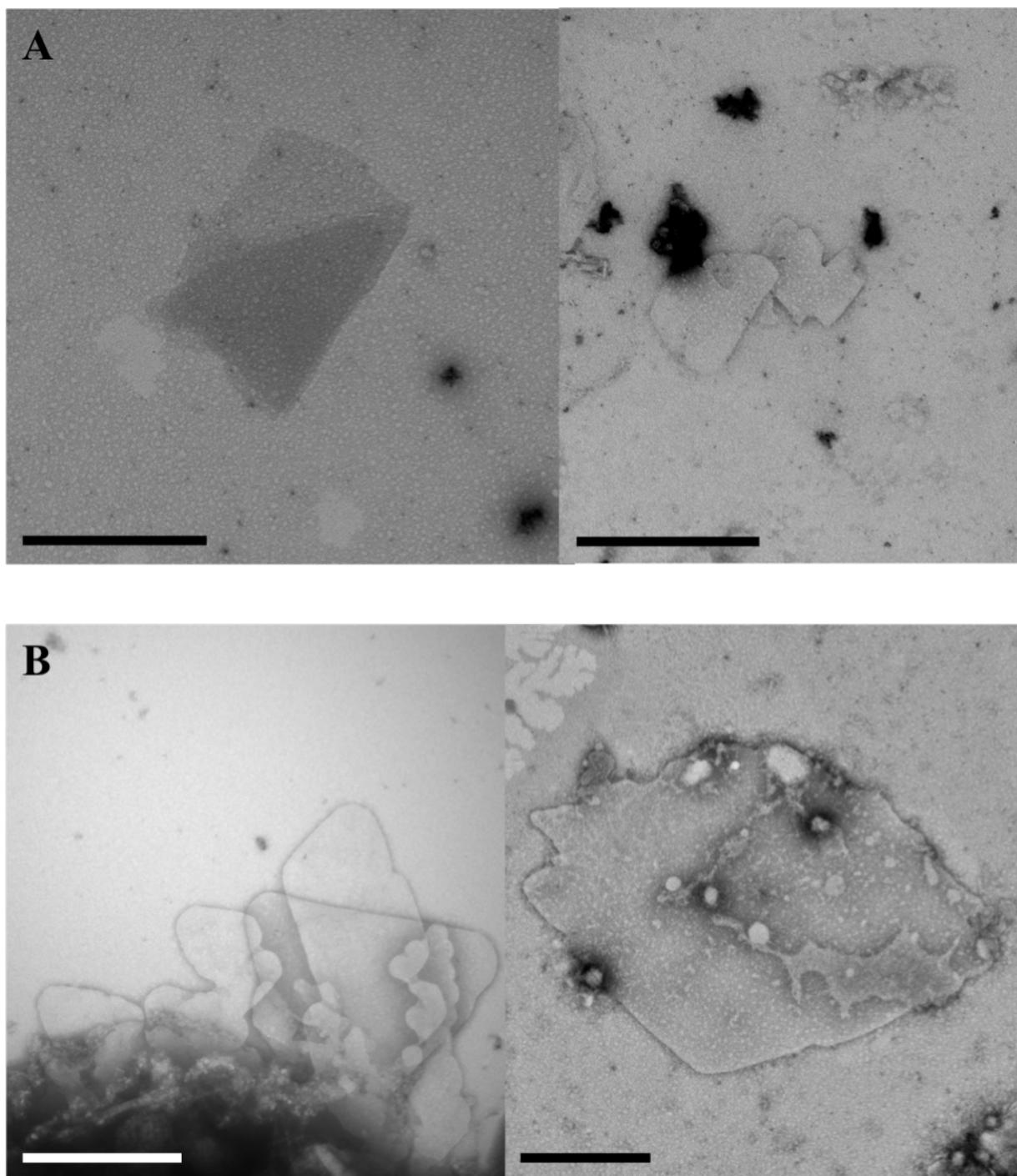


Figure 3.10 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL pentamer assemblies in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

Circular dichroism (CD) spectropolarimetry was utilized to determine the secondary structure of the proteins. Since sheet-like assemblies were primarily observed in 10 mM MOPS buffer (pH 7.0) and 10 mM TAPS buffer (pH 8.5), the protein samples under these buffer conditions were analyzed by circular dichroism. The nascent and thermally annealed HisPur™ 3FD-LL concatemer assemblies were analyzed at a protein concentration of 3 mg/mL.

The 3FD-LL concatemers were designed to be composed entirely of α -helical secondary structure, which should be evident by minima at 208 and 222 nm and a maximum at 195 nm. However, the CD spectra for all of the assemblies were extremely weak, which was likely a result of the low solubility of the protein in water-based buffers. CD analysis of nascent and thermally annealed **3FD-LL monomer**, **3FD-LL dimer**, and **3FD-LL tetramer** assemblies did not produce relevant spectra in either MOPS buffer or TAPS buffer (**Figures 3.11, 3.12, and 3.14**). The CD spectra for the nascent and thermally annealed **3FD-LL trimer** and **3FD-LL pentamer** assemblies displayed very weak minima around 225 nm, which could potentially represent a red-shifted 222 nm minima characteristic of the α -helical secondary structure (**Figures 3.13 and 3.15**). The red shift in addition to the absence of the maxima and minima at lower wavelengths was an expected outcome of the insolubility of the proteins. Scattering of left- and right-handed circularly polarized light is impacted to the greatest degree at lower wavelengths, which causes the signal intensity to be muted.³⁻⁷ Since solubility was a problem for all samples, a new method would have to be investigated for assembly and analysis of the 3FD-LL concatemer proteins.

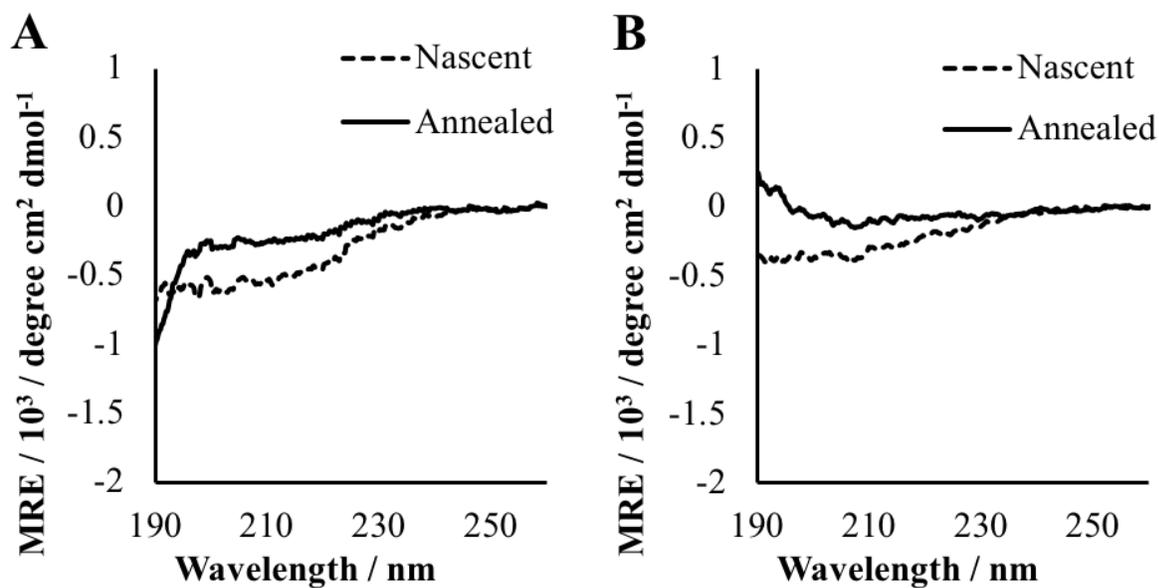


Figure 3.11 Circular dichroism spectra of nascent and thermally annealed assemblies of 3 mg/mL HisPur™ 3FD-LL monomer. (A) CD spectra of 3FD-LL monomer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL monomer in 10 mM TAPS, pH 8.5.

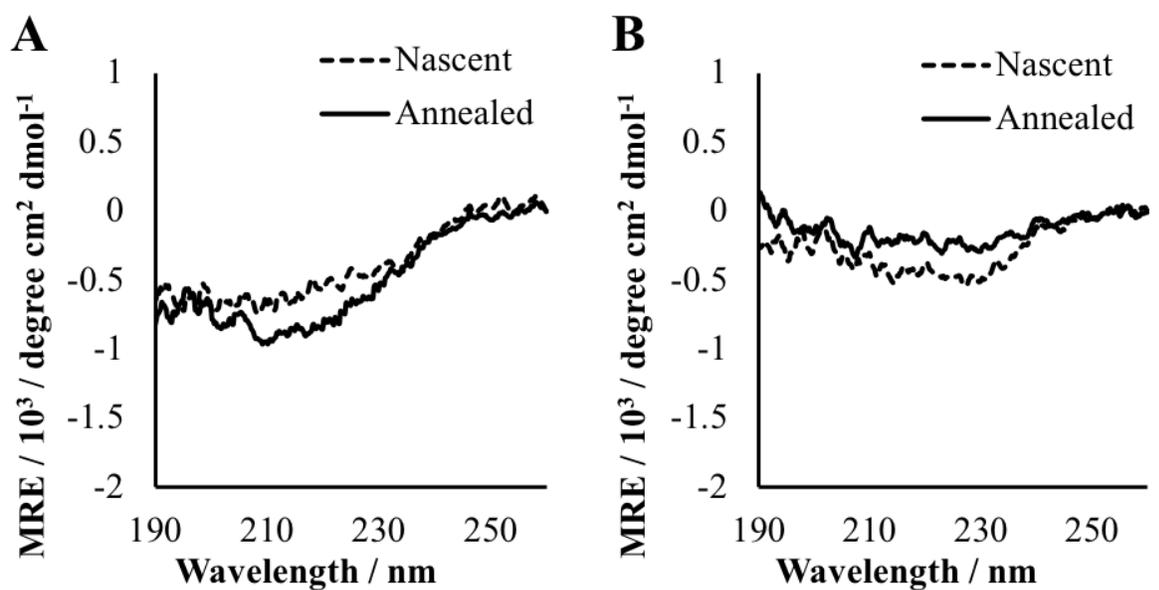


Figure 3.12 Circular dichroism spectra of nascent and thermally annealed assemblies of 3 mg/mL HisPur™ 3FD-LL dimer. (A) CD spectra of 3FD-LL dimer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL dimer in 10 mM TAPS, pH 8.5.

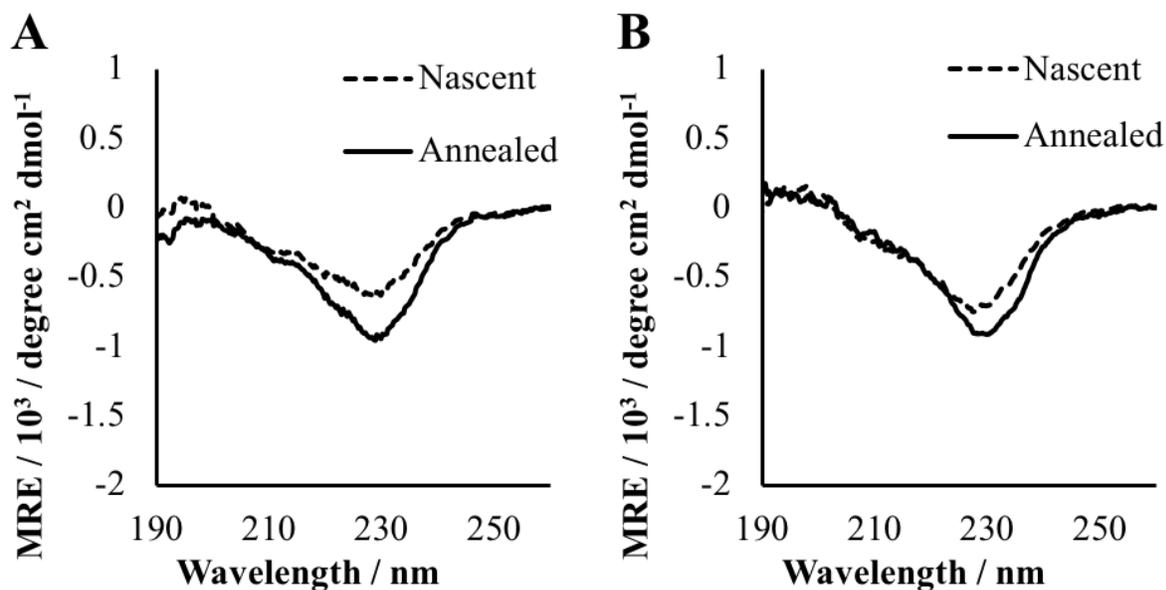


Figure 3.13 Circular dichroism spectra of nascent and thermally annealed assemblies of 3 mg/mL HisPur™ 3FD-LL trimer. (A) CD spectra of 3FD-LL trimer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL trimer in 10 mM TAPS, pH 8.5.

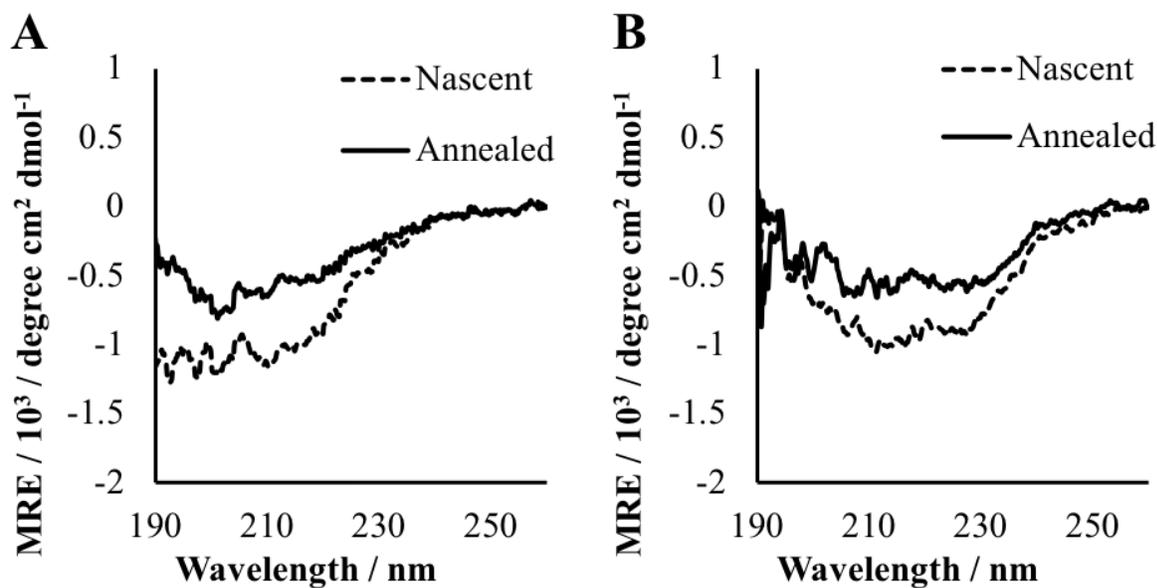


Figure 3.14 Circular dichroism spectra of nascent and thermally annealed assemblies of 3 mg/mL HisPur™ 3FD-LL tetramer. (A) CD spectra of 3FD-LL tetramer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL tetramer in 10 mM TAPS, pH 8.5.

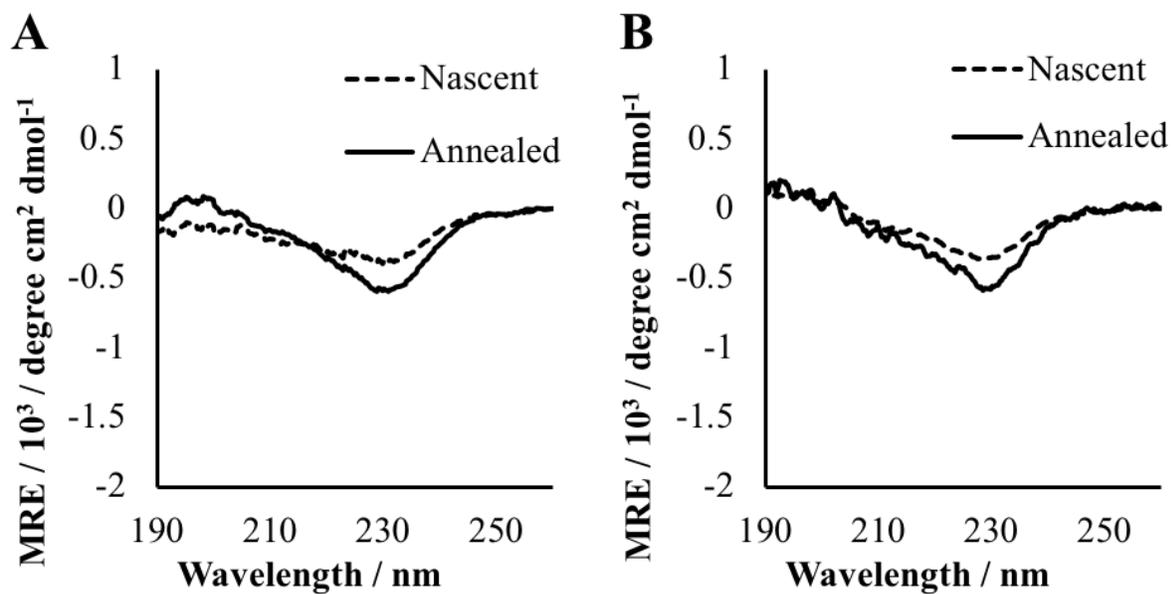


Figure 3.15 Circular dichroism spectra of nascent and thermally annealed assemblies of 3 mg/mL HisPur™ 3FD-LL pentamer. (A) CD spectra of 3FD-LL pentamer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL pentamer in 10 mM TAPS, pH 8.5.

Assembly of HisPurTM 3FD-LL Concatemers with TFE

To address the insolubility issues of the 3FD-LL concatemers, 2,2,2-trifluoroethanol (TFE) was considered as an organic cosolvent. TFE has been used for decades to increase the solubility of peptides for NMR and CD experiments due to its ability to break apart protein aggregates.⁸⁻⁹ In addition to its ability to facilitate solubility of recalcitrant peptides, the fluorinated alcohol is also capable of inducing the helicity of secondary structures within peptides.¹⁰⁻¹² Furthermore, TFE induces the stabilization of α -helical structures by slowing the formation of assemblies and therefore preventing kinetic trapping.¹⁰ As a cosolvent, TFE undergoes aggregation to form a matrix around the peptides, which consequently displaces water, removes intermolecular hydrogen-bonding partners, and consequently promotes local interactions between nearby residues. The matrix further stabilizes the α -helical conformation by providing a dielectric environment more closely resembling that of the interior of the peptide since the dielectric constant of TFE is one-third that of water.¹³ Unlike many other organic solvents, TFE only weakly interacts with nonpolar residues, so hydrophobic interactions are not disrupted. Therefore, using TFE as a cosolvent for the 3FD-LL concatemer assemblies would not only improve solubility of the proteins, but it would also help induce and stabilize the α -helical secondary structure.

The **3FD-LL trimer** displayed the most defined and numerous sheets in all prior experiments, so this concatemer protein was selected for the initial TFE experiments. To assess the potential of using TFE as a cosolvent, the **3FD-LL trimer** was dissolved in 100% TFE in order to break up any protein aggregates, and then the desired buffer was added to give a final concentration of 50% (v/v) TFE/buffer. The samples were allowed to incubate at room temperature with the cap open to allow the slow evaporation of TFE from the solution. Although the **3FD-LL trimer** exhibited significantly greater solubility in the 50% TFE/buffer solution and after the TFE

had evaporated, there was still white precipitate present in all samples. Initially, the samples continued to incubate at room temperature until analysis with no thermal annealing process.

The unannealed HisPurTM **3FD-LL trimer** assembled with TFE as a cosolvent were analyzed using transmission electron microscopy at a protein concentration of 10 mg/mL in 10 mM MOPS buffer (pH 7.0) and 10 mM TAPS buffer (pH 8.5). The **3FD-LL trimer** samples for both buffer conditions revealed a substantial increase in the size, quantity, and quality of the formed nanosheets (**Figures 3.16 and 3.17**). Instead of only forming small sheets in the one to two hundred nanometer range, sheets were frequently visualized in the one to two micrometer range. In addition, the majority of the sheets had considerably more defined edges. On the downside, some of the sheets had visible cracks traversing them and sheets were stacking on top of each other. The **3FD-LL trimer** nanosheets were more prevalent when assembled in TAPS buffer compared to MOPS buffer. Similar to the assemblies formed without TFE, there were still fibers present, but the population of sheets was approaching that of the fibers. Assembling in TFE substantially improved the assembly of **3FD-LL trimer** nanosheets.

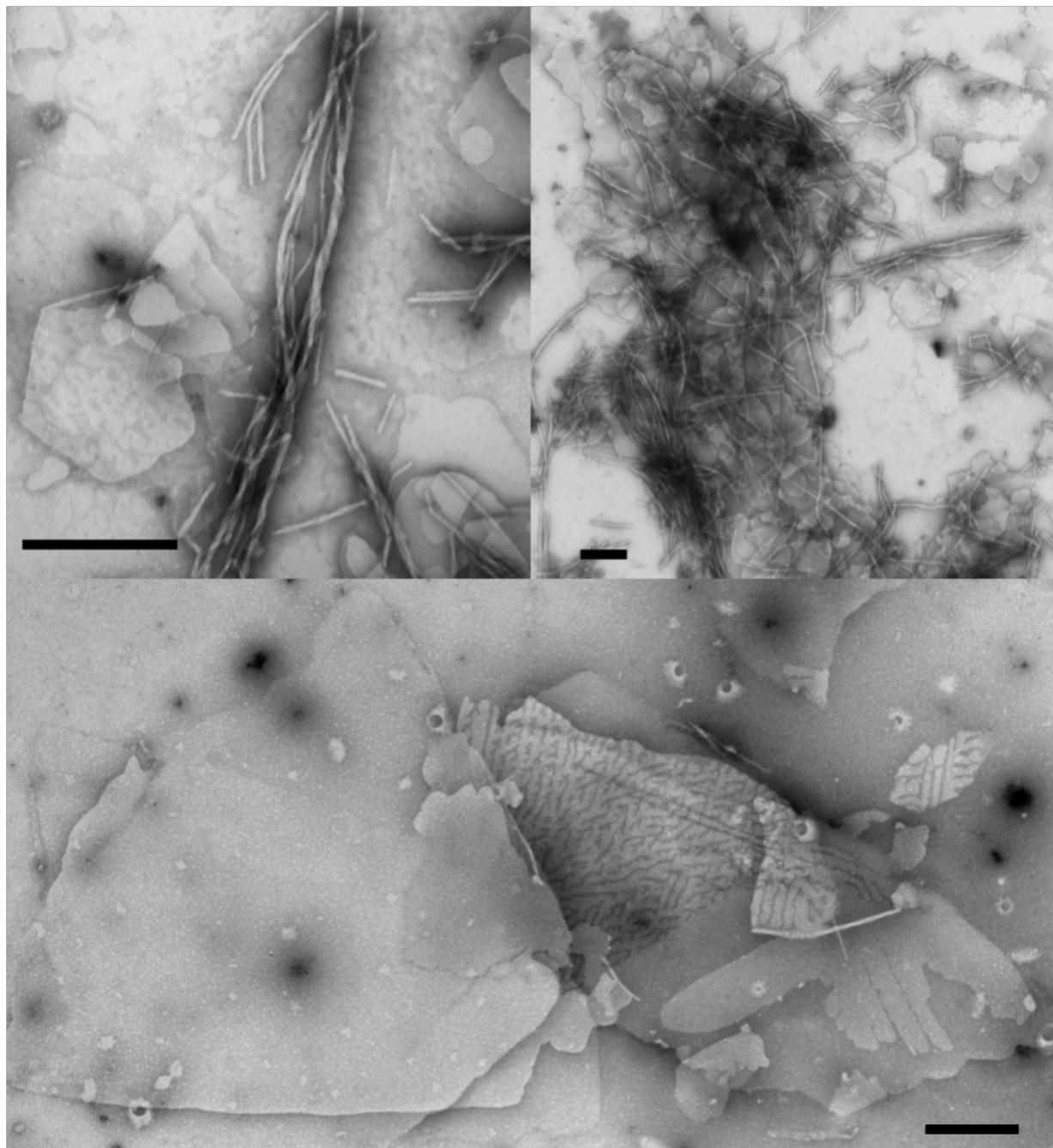


Figure 3.16 Transmission electron microscopy images of unannealed 10 mg/mL HisPur™ 3FD-LL trimer assembled with TFE in 10 mM MOPS buffer (pH 7.0). Scale bar = 500 nm.

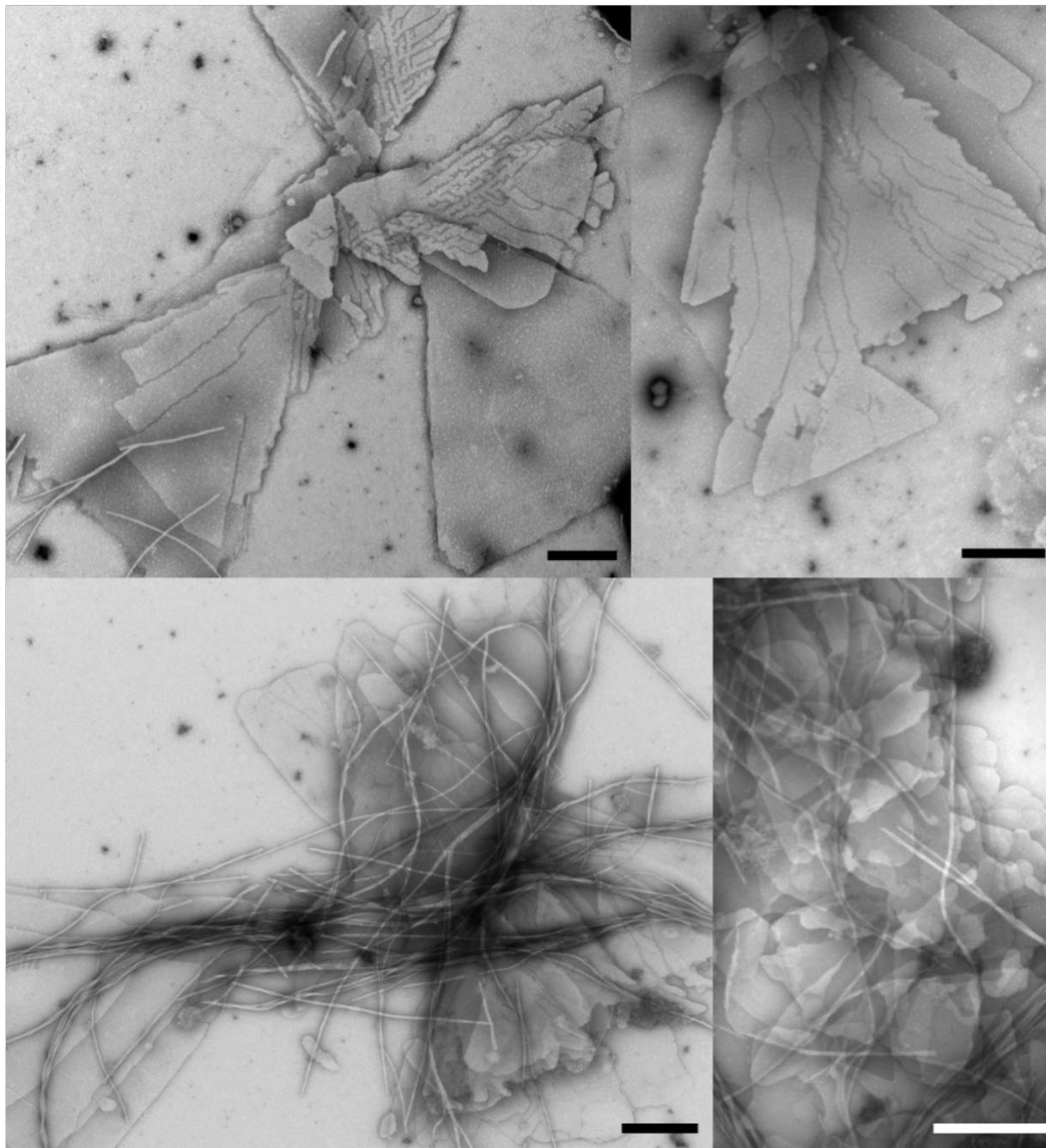


Figure 3.17 Transmission electron microscopy images of unannealed 10 mg/mL HisPur™ 3FD-LL trimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). Scale bar = 500 nm.

The use of TFE as a cosolvent effectively improved nanosheet formation for the **3FD-LL trimer**, so TFE was used for the assembly of all five concatemers. The HisPur™ 3FD-LL concatemers were assembled in TFE at a peptide concentration of 10 mg/mL in MOPS (pH 7.0) and TAPS (pH 8.5) buffers. In addition to the previous experiment, the samples were thermally annealed after the TFE evaporated from the samples with an intent to decrease the number of cracks formed in nanosheets by controlling the rate of assembly. The solubility of most of the 3FD-LL concatemers was significantly improved, but there was still some precipitate in all of the samples. The **3FD-LL monomer** was exceptionally insoluble even in a 100% TFE solution.

Transmission electron microscopy was used to analyze the nascent and thermally annealed HisPur™ 3FD-LL concatemers assembled in TFE. The proteins were assembled at a protein concentration of 10 mg/mL in MOPS (pH 7.0) and TAPS (pH 8.5) buffers. Similar to what was seen previously for the **3FD-LL trimer**, the nascent and thermally annealed samples assembled in TFE showed a significant increase in the quantity of nanosheets for all concatemer proteins. The nanosheets were considerably larger and all sheets were more defined than those visualized in the samples assembled without TFE. Although there were still fibers present in all of the samples, there was a smaller degree of unassembled peptide in solution, which indicates that TFE assisted in directing the proteins toward assembly.

Similar to the nanosheets observed in the unannealed **3FD-LL trimer**, many sheets formed in the nascent samples of the concatemers displayed cracks crossing through the sheets. Although some cracks were still observed in sheets of the annealed samples, there were more sheets without cracks indicating that the annealing processes may allow the protein helices to come together in a more controlled manner. If the assembly process occurs too rapidly, cracks could result from helix misalignment through staggering of helical units or the His-tag impeding proper alignment.

Furthermore, for all concatemer proteins, the sheets present in the TAPS buffer (pH 8.5) displayed fewer cracks compared to those located in the MOPS buffer (pH 7.0).

The thermally annealed **3FD-LL monomer** and **3FD-LL pentamer** samples exhibited a quantity of nanosheets that approached the quantity of fibers (**Figures 3.18, 3.19, 3.26, and 3.27**). However, there was a significant degree of protein aggregate present in the **3FD-LL monomer** sample, which may result from its notably low solubility. The **3FD-LL dimer** and **3FD-LL tetramer** continued to produce a vast amount of fibers in both nascent and thermally annealed samples (**Figures 3.20, 3.21, 3.24, and 3.25**). The thermally annealed **3FD-LL trimer** samples revealed a quantity of nanosheets that surpassed the quantity of fibers with only a handful of fibers present in the sample (**Figures 3.22 and 3.23**), which was a considerable improvement over the nascent and unannealed samples. Furthermore, the sheets in the thermally annealed **3FD-LL trimer** samples displayed edges with superior definition compared to the sheets present in the nascent sample. Based on TEM analysis of the TFE assembled samples, the **3FD-LL trimer** produced the most homogeneous sample of nanosheets and the thermal annealing process improved the formation of sheets. Lastly, the **3FD-LL trimer** consistently produced nanosheets in the one to three micrometer range, which were significantly larger than the sheets formed by the **3FD-LL** and **3FD-IL** model peptides.

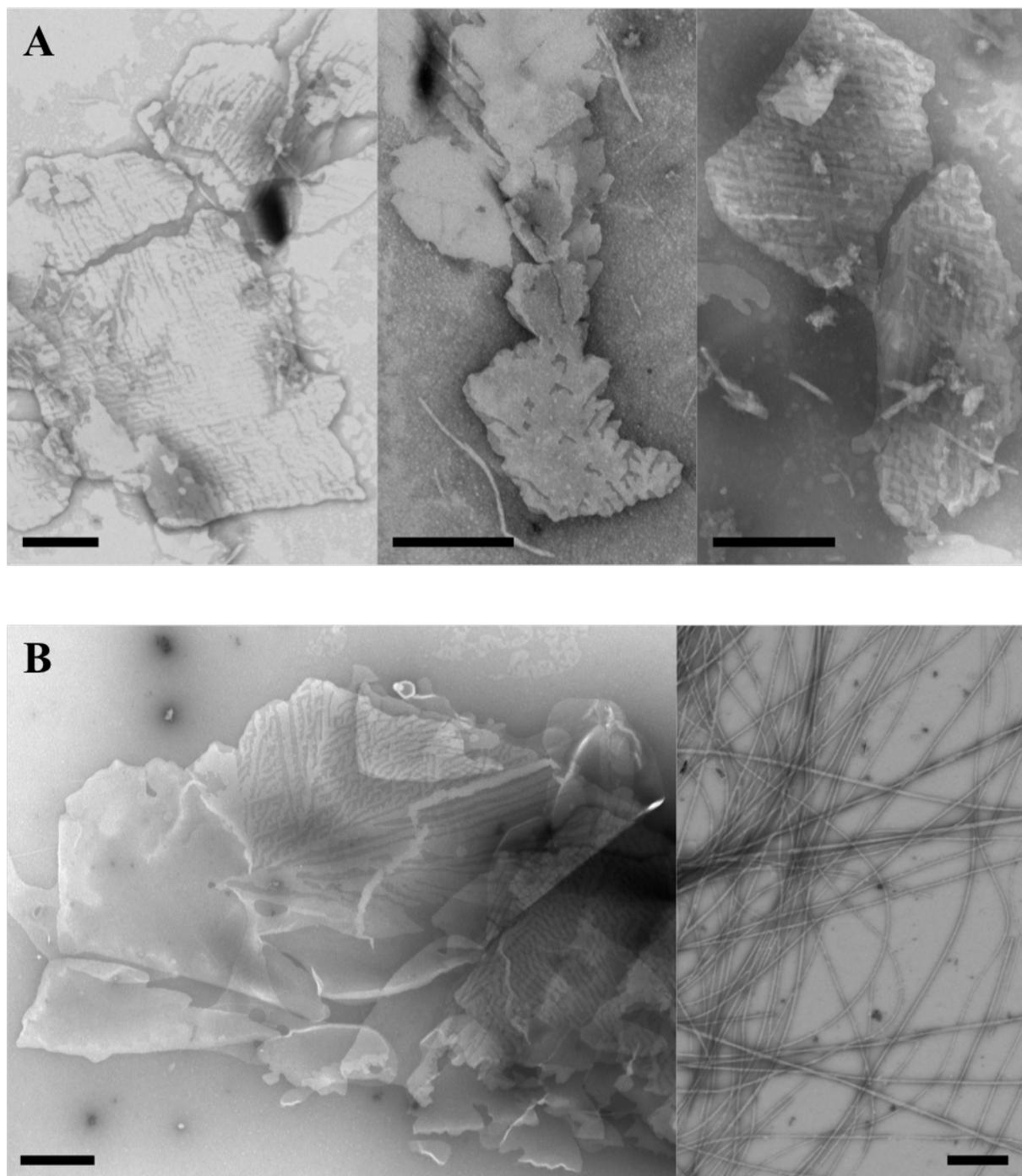


Figure 3.18 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL monomer assembled with TFE in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

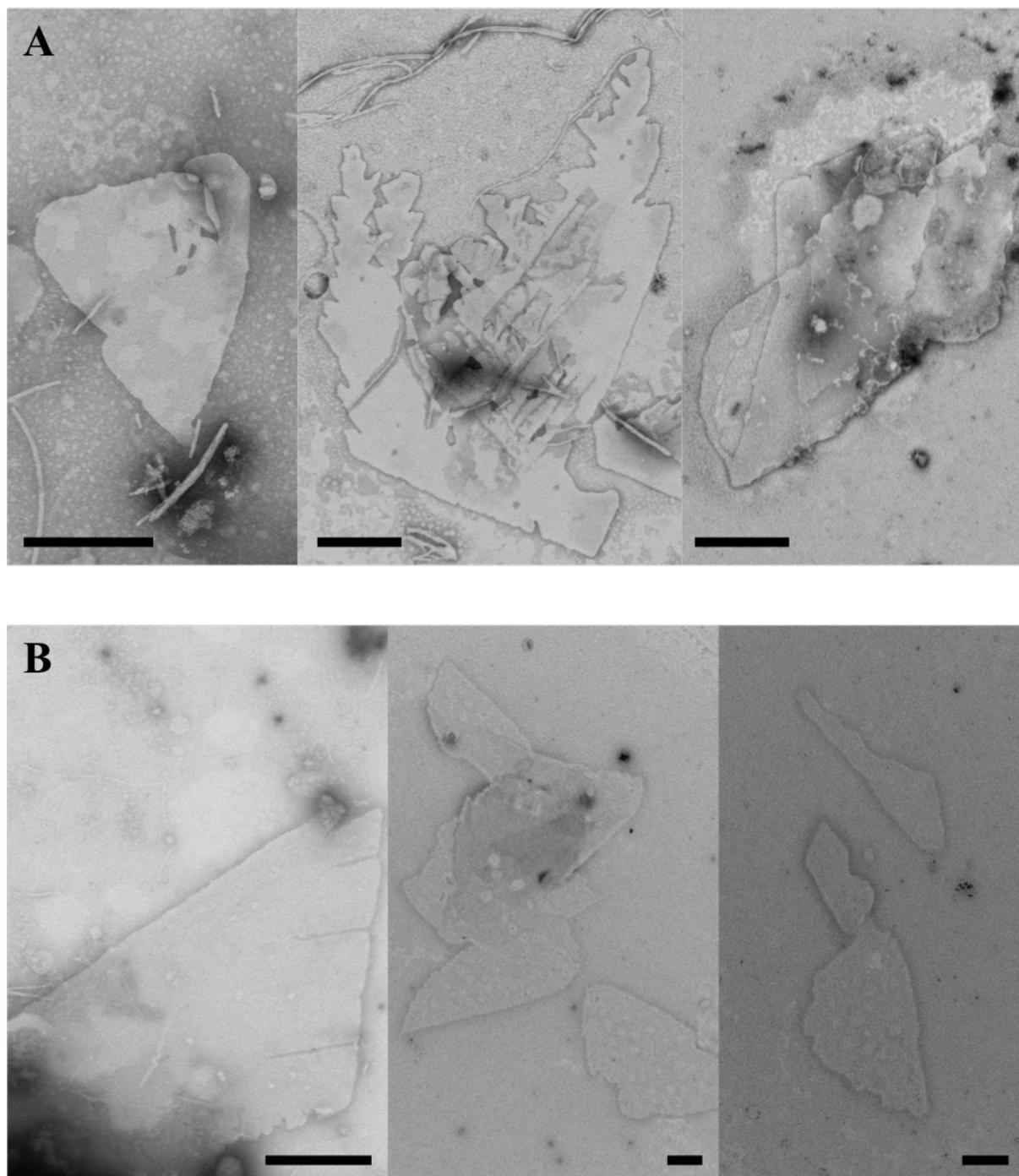


Figure 3.19 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL monomer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

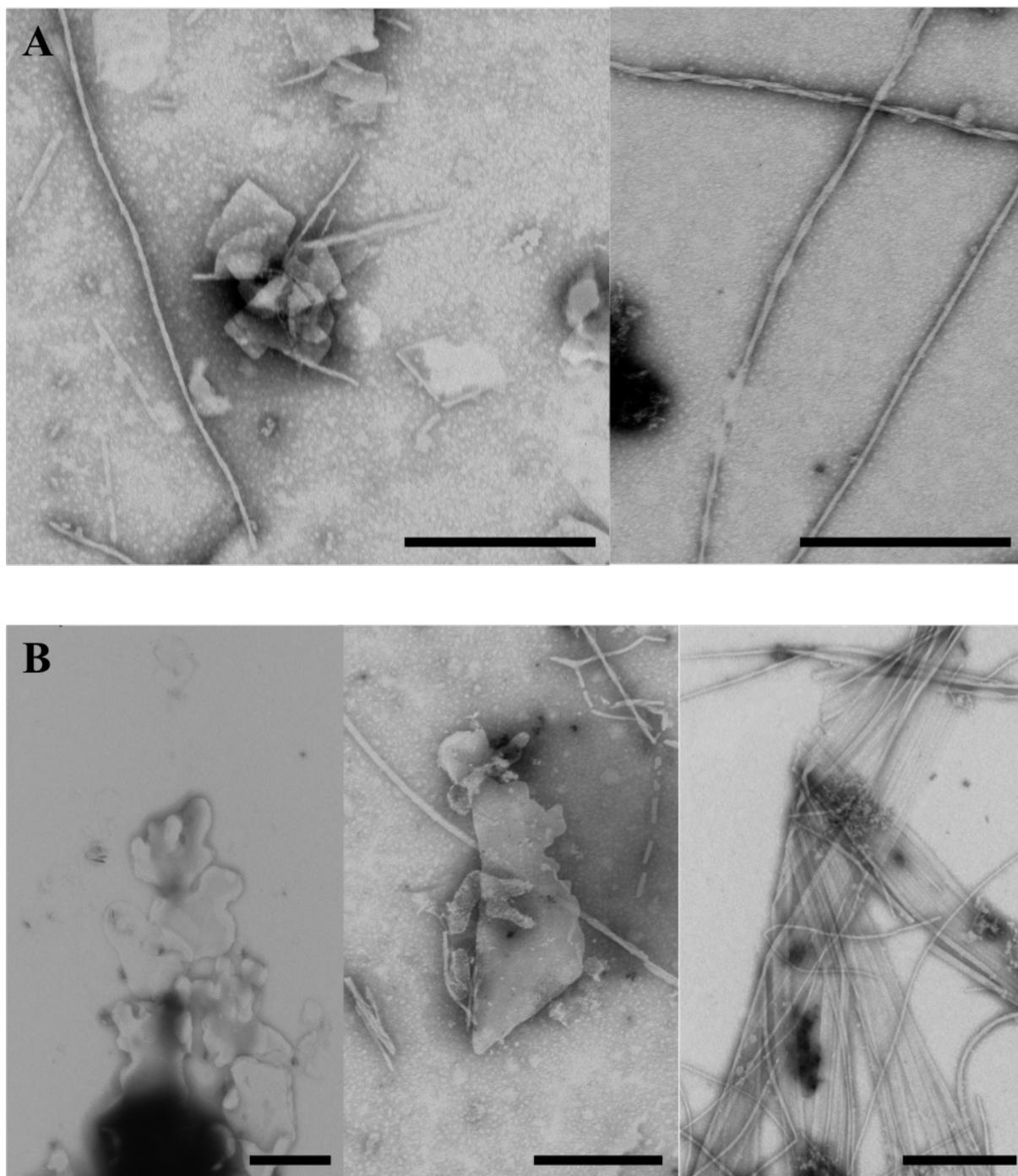


Figure 3.20 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL dimer assembled with TFE in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

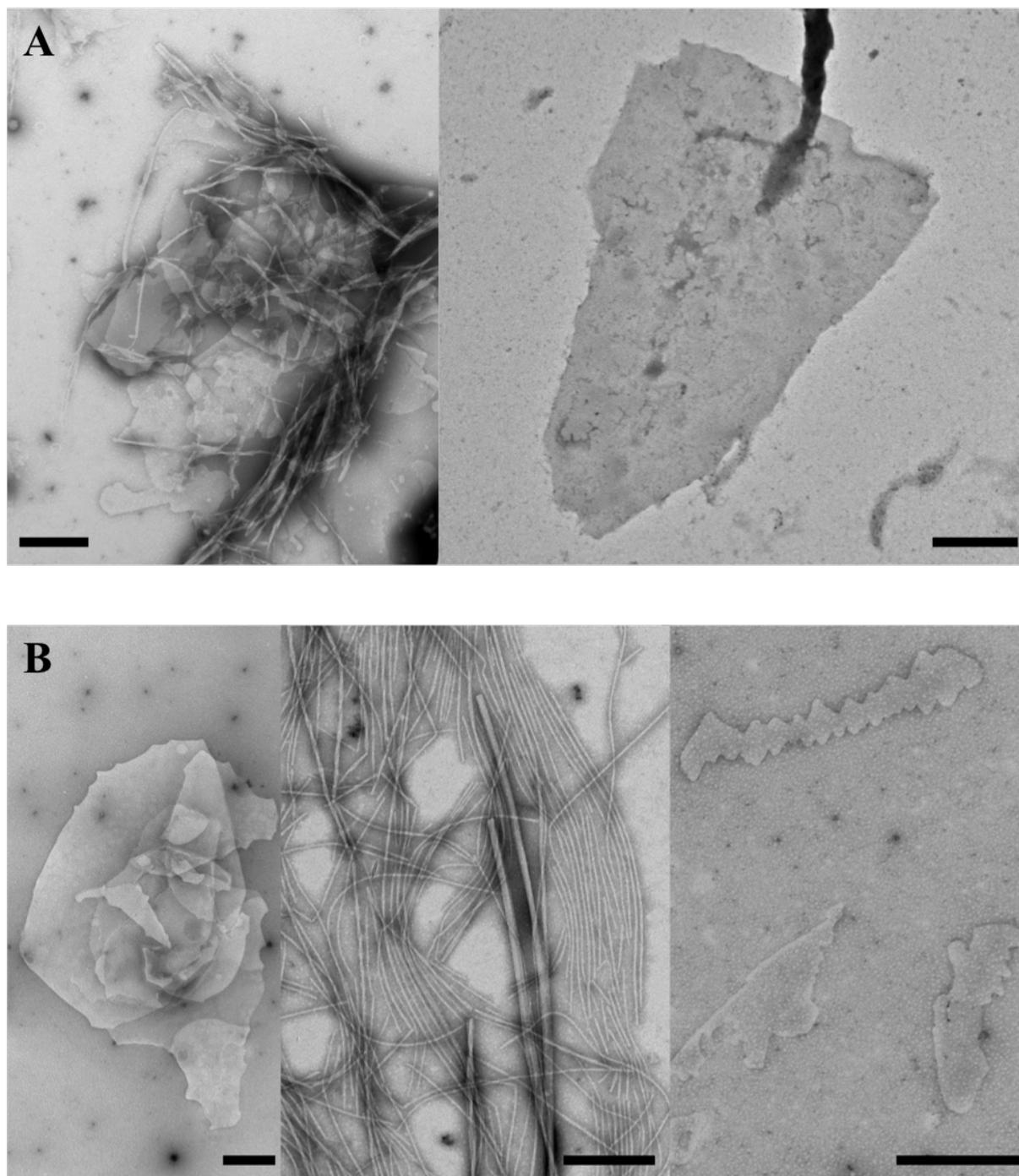


Figure 3.21 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL dimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

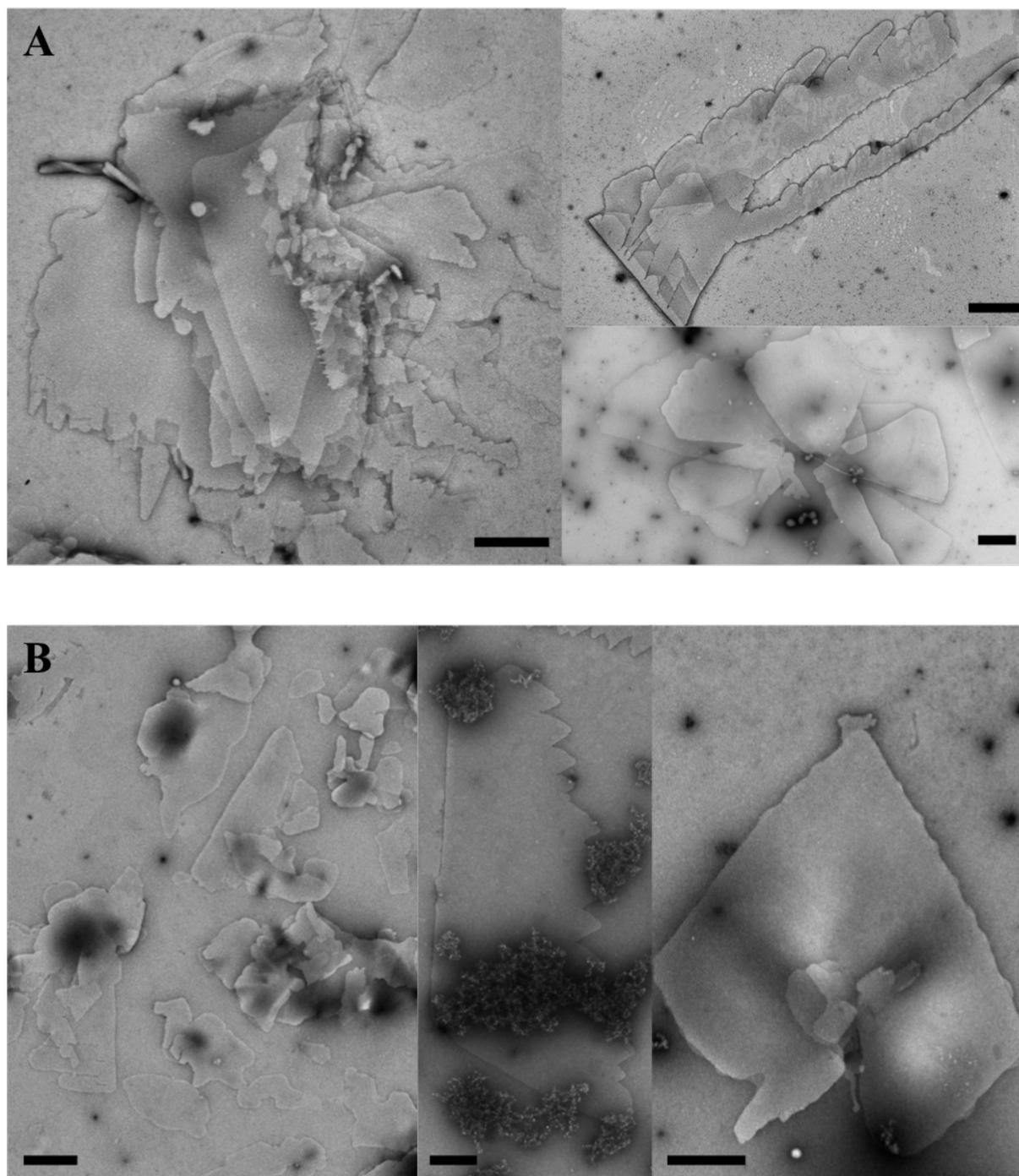


Figure 3.22 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL trimer assembled with TFE in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

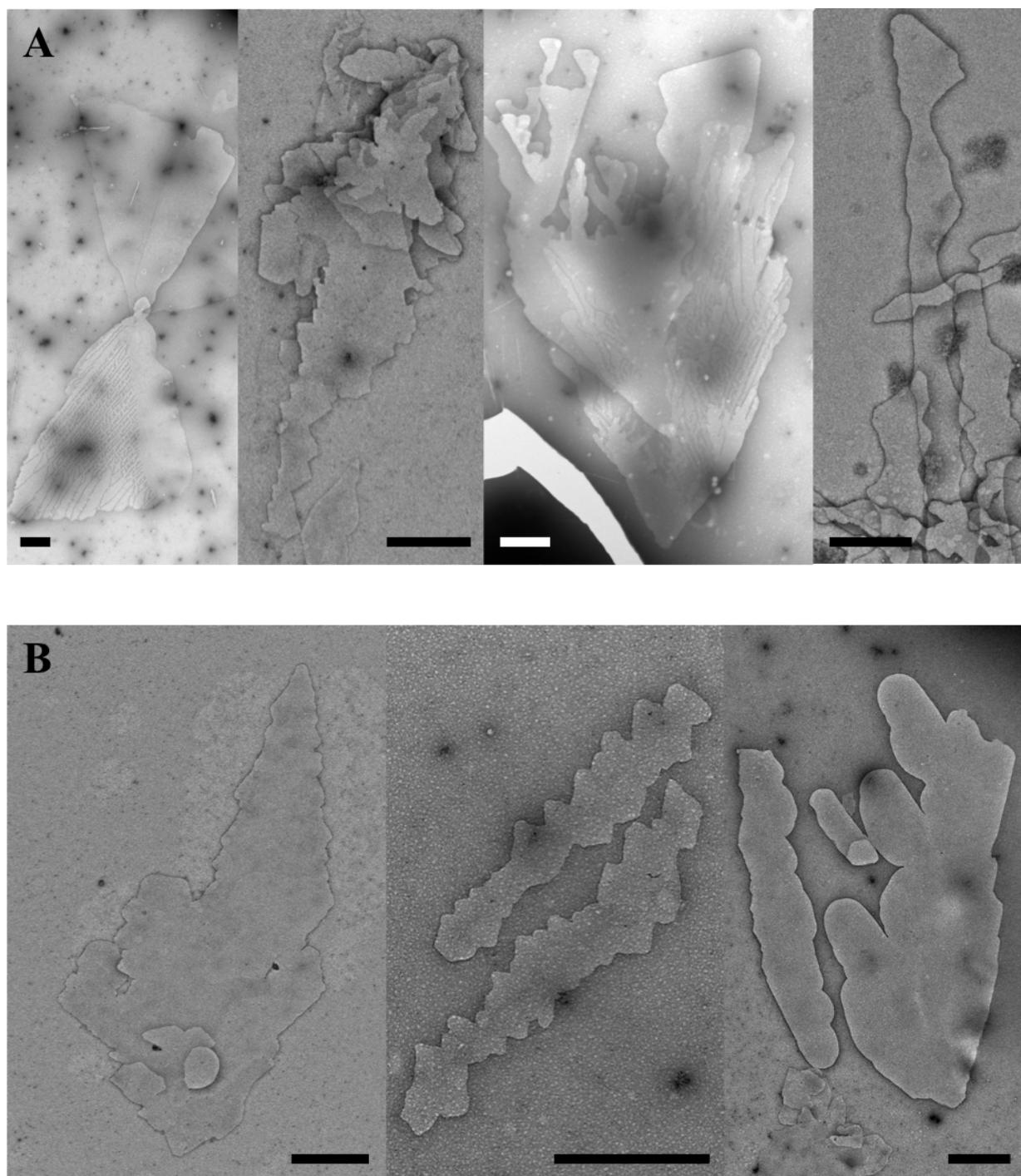


Figure 3.23 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL trimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

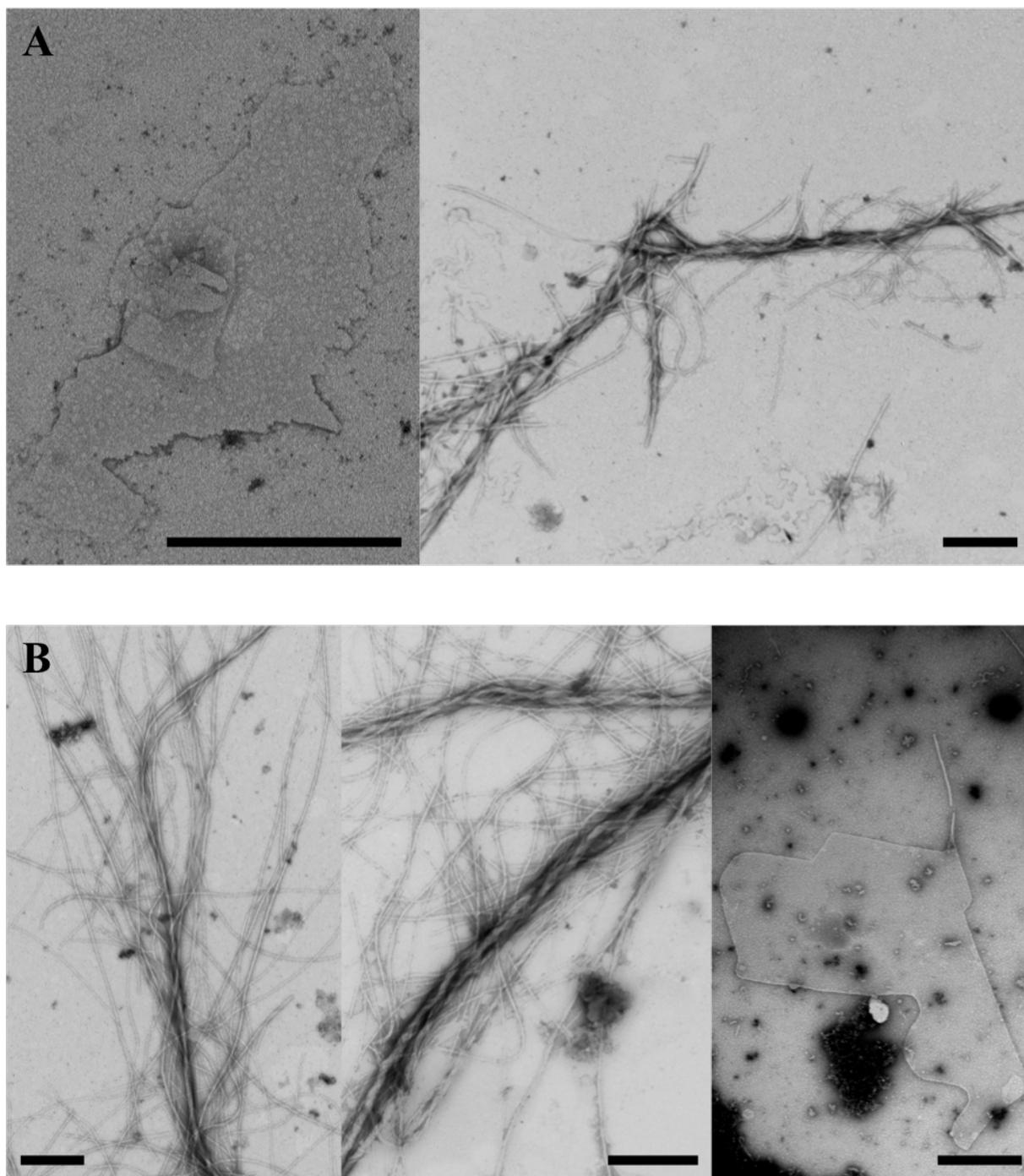


Figure 3.24 Transmission electron microscopy images of 10 mg/mL HisPur[™] 3FD-LL tetramer assembled with TFE in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

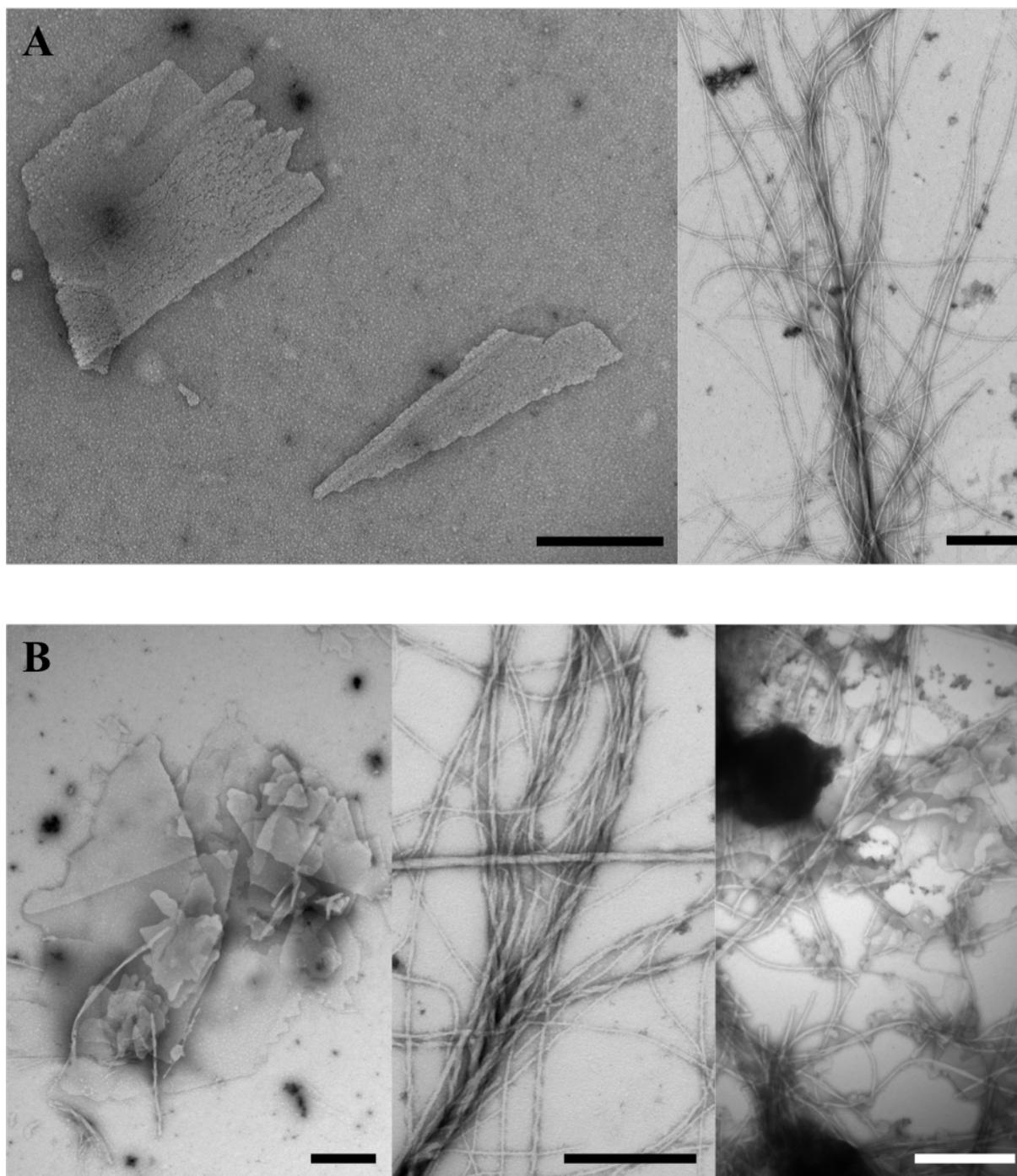


Figure 3.25 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL tetramer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

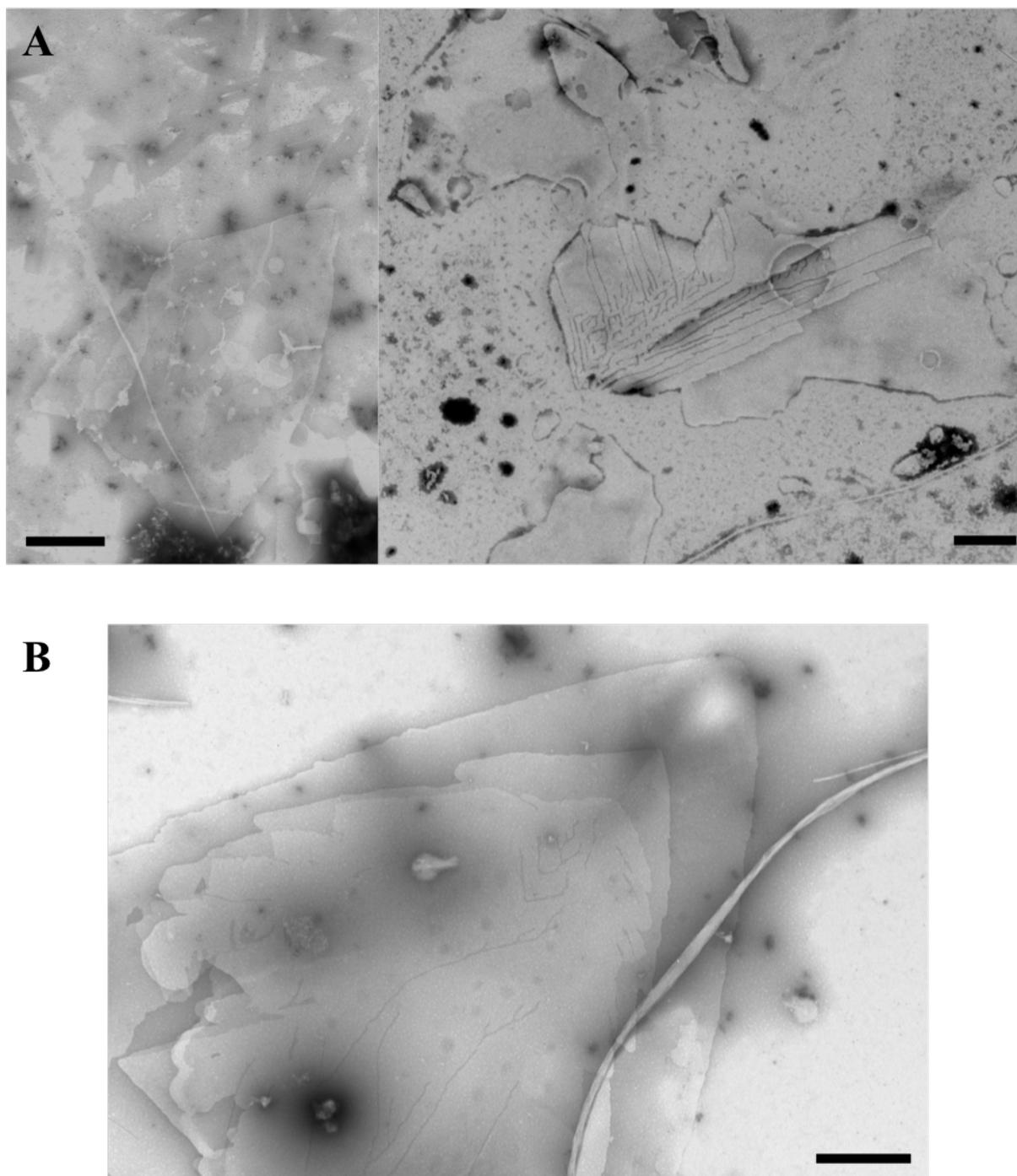


Figure 3.26 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL pentamer assembled with TFE in 10 mM MOPS buffer (pH 7.0). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

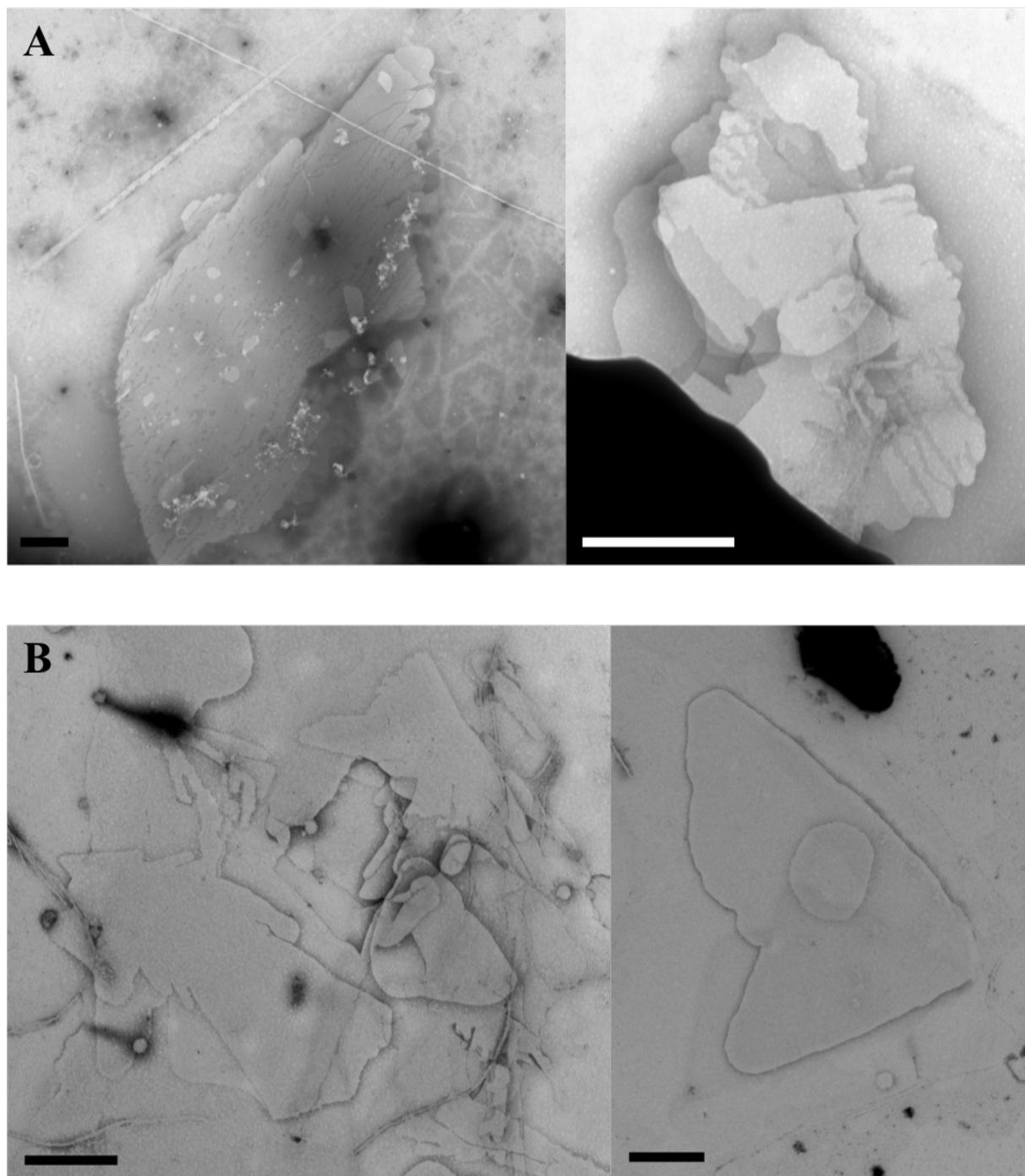


Figure 3.27 Transmission electron microscopy images of 10 mg/mL HisPur™ 3FD-LL pentamer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) TEM images of nascent assemblies. (B) TEM images of thermally annealed assemblies. Scale bar = 500 nm.

The nascent and thermally annealed HisPur™ 3FD-LL concatemer assemblies were also analyzed using circular dichroism spectropolarimetry at a protein concentration of 10 mg/mL in 10 mM TAPS buffer (pH 8.5) and 10 mM MOPS buffer (pH 7.0). As expected, the CD spectra of the 3FD-LL concatemers led to enhanced signal intensities compared to the samples assembled without TFE due to the improved solubility of most samples. The main exception was the **3FD-LL monomer**, which failed to produce relevant spectra owing to the especially low solubility of this protein (**Figure 3.28**). The CD spectra of the **3FD-LL dimer** and **3FD-LL tetramer** assemblies displayed very weak minima around 208 nm and 225 nm, and only some of the samples had a positive maximum around 195 nm (**Figures 3.29 and 3.31**). On the other hand, the **3FD-LL trimer** and **3FD-LL pentamer** assemblies exhibited weak α -helical signatures with maximum at 195 nm and minima around 208 and 222 nm (**Figures 3.30 and 3.32**). The **3FD-LL trimer** samples produced the strongest α -helical CD signatures with the greatest intensities.

Although solubility was improved by using TFE as a cosolvent, none of the LL concatemers demonstrated complete solubility in TFE. As a result, all samples displayed muted signal intensities at lower wavelengths as a result of light scattering.³⁻⁷ Thermally annealed samples resulted in spectra more closely resembling the characteristic α -helical signature with stronger intensities of maxima and minima, which indicated that thermal annealing may improve the α -helical secondary structure during assembly formation. Furthermore, the CD spectra acquired for the concatemers verified that the proteins exhibit a greater degree of α -helicity in TAPS buffer (pH 8.5) than MOPS buffer (pH 7.0). However, the signal intensities remained low, and the intensities were heavily muted in the lower wavelength range, even for the **3FD-LL trimer**, which produced the best signature. y

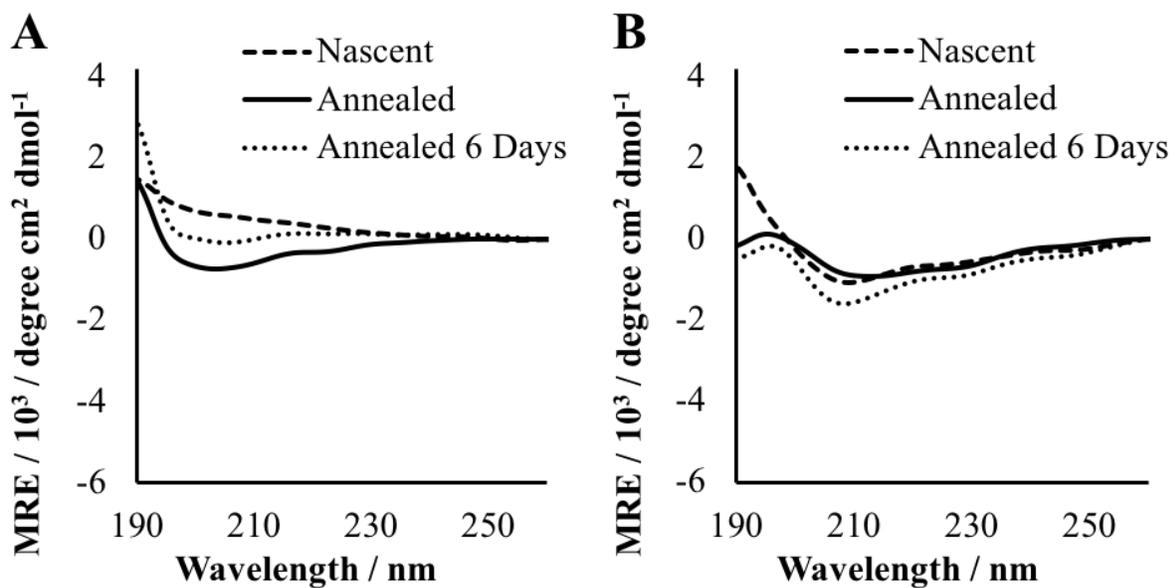


Figure 3.28 Circular dichroism spectra of 10 mg/mL nascent and thermally annealed HisPur™ 3FD-LL monomer assembled in TFE. (A) CD spectra of 3FD-LL monomer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL monomer in 10 mM TAPS, pH 8.5.

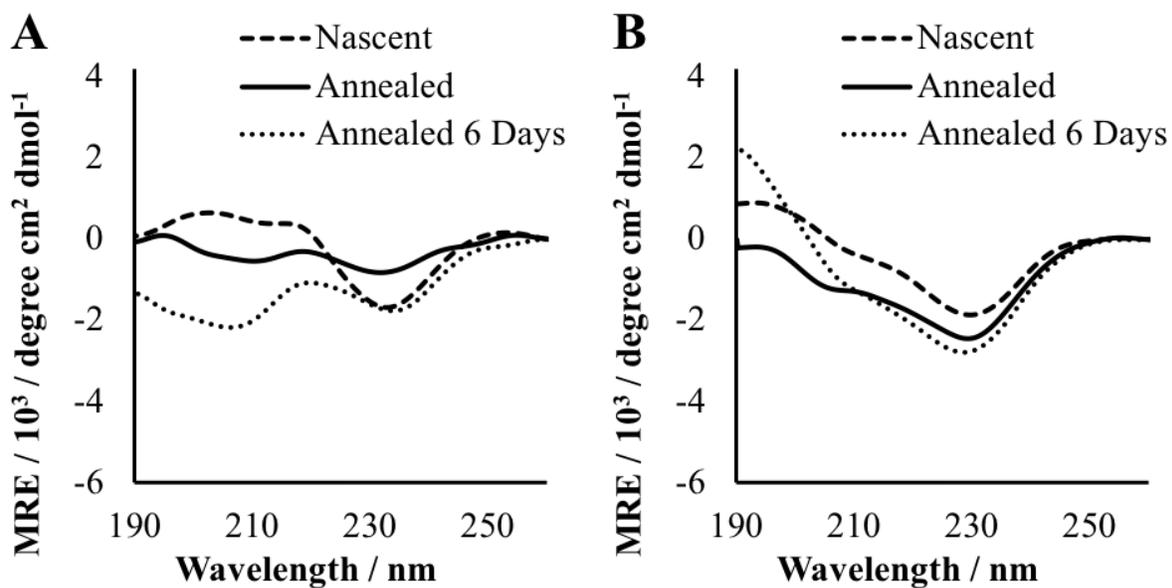


Figure 3.29 Circular dichroism spectra of 10 mg/mL nascent and thermally annealed HisPur™ 3FD-LL dimer assembled in TFE. (A) CD spectra of 3FD-LL dimer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL dimer in 10 mM TAPS, pH 8.5.

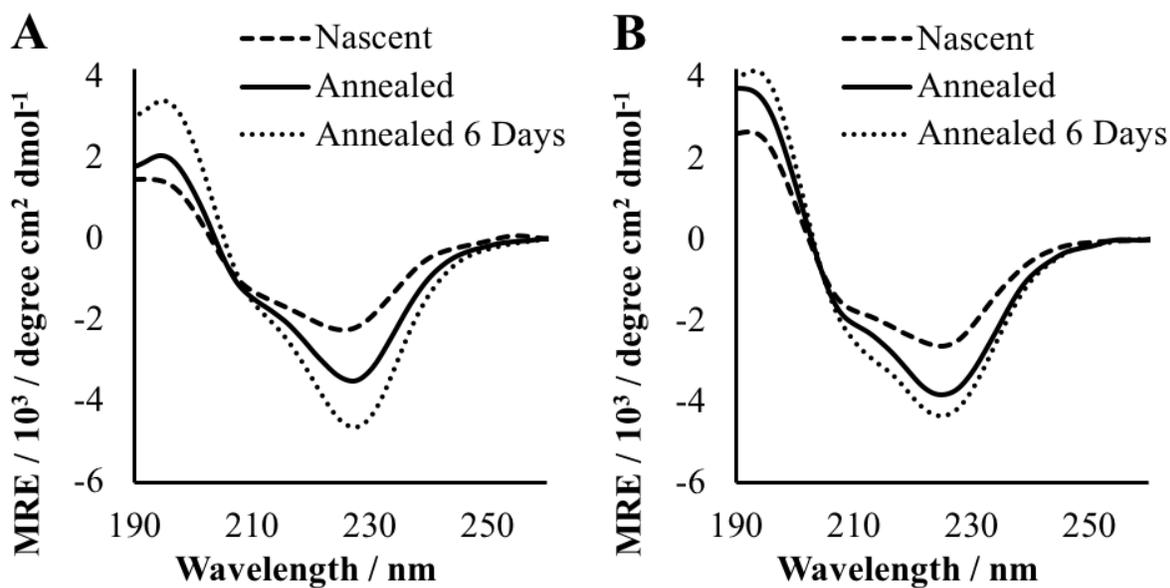


Figure 3.30 Circular dichroism spectra of 10 mg/mL nascent and thermally annealed HisPur[™] 3FD-LL trimer assembled in TFE. (A) CD spectra of 3FD-LL trimer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL trimer in 10 mM TAPS, pH 8.5.

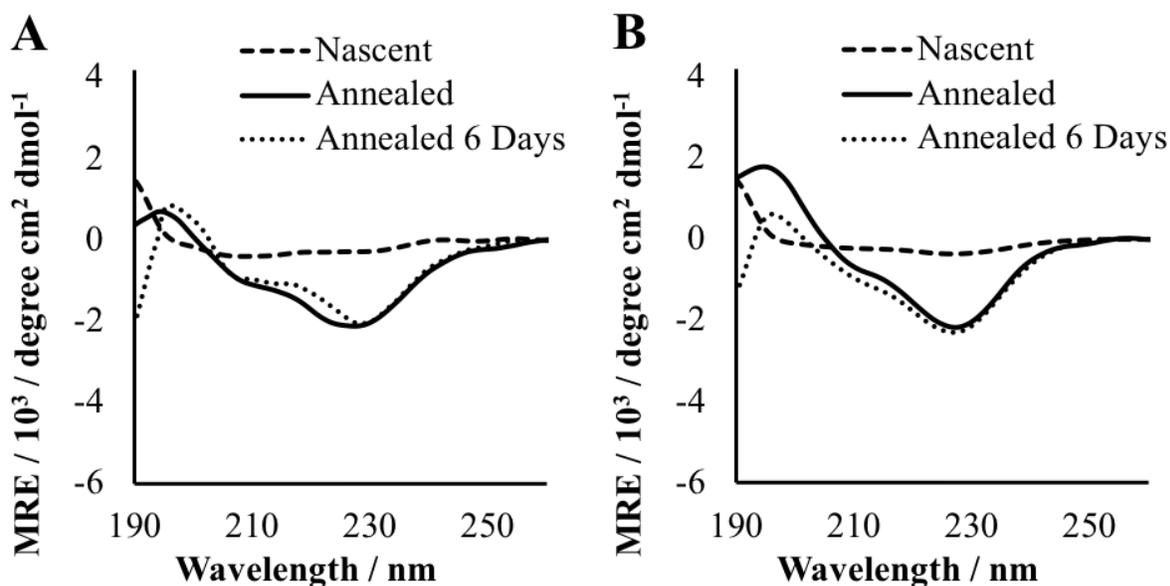


Figure 3.31 Circular dichroism spectra of 10 mg/mL nascent and thermally annealed HisPur[™] 3FD-LL tetramer assembled in TFE. (A) CD spectra of 3FD-LL tetramer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL tetramer in 10 mM TAPS, pH 8.5.

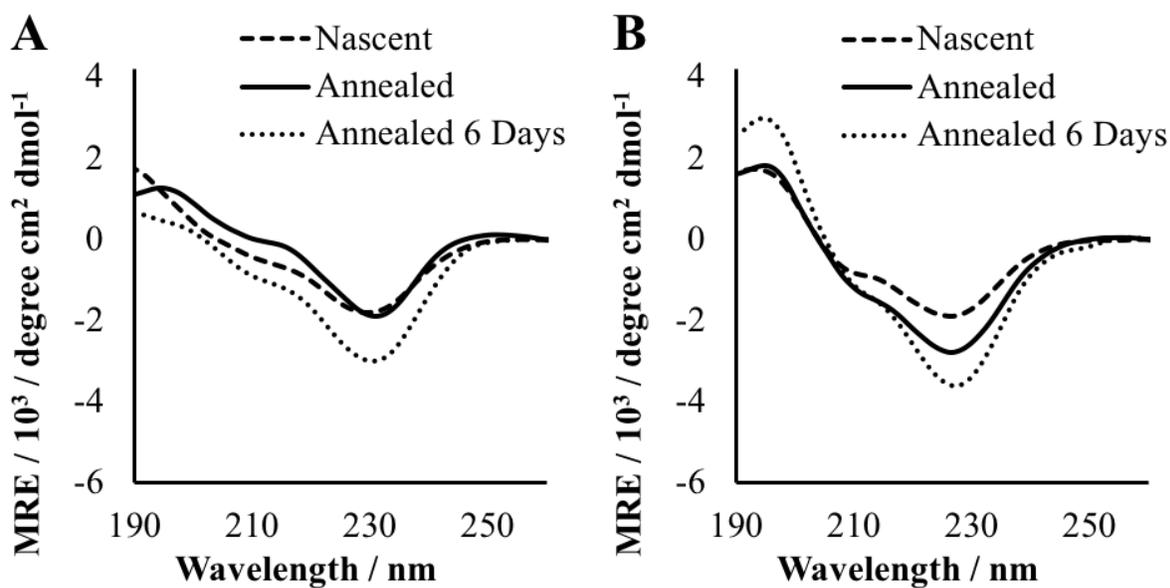


Figure 3.32 Circular dichroism spectra of 10 mg/mL nascent and thermally annealed HisPur™ 3FD-LL pentamer assembled in TFE. (A) CD spectra of 3FD-LL pentamer in 10 mM MOPS, pH 7.0. (B) CD spectra of 3FD-LL pentamer in 10 mM TAPS, pH 8.5.

3.2.2 Characterization of HPLC Purified 3FD-LL Trimer Nanosheets

TFE drastically improved the ability of the concatemers to produce nanosheets; however, the **3FD-LL trimer** was the most promising candidate as it formed the most defined two-dimensional assemblies in the most reproducible manner. In addition, the **3FD-LL trimer** exhibited the most characteristic α -helical signature with the greatest helicity when analyzed by CD spectropolarimetry. Although favorable, the peptide only formed a low quantity of sheets and the samples displayed a small degree of heterogeneity due to the presence of fibers. This may result from the low solubility of the peptide in water-based solvents even when assembled with TFE. It was apparent that further optimization was necessary to enhance the properties and assemblies of the **3FD-LL trimer**.

After the acquisition of promising data from CD and TEM analysis of the **3FD-LL trimer**, reverse-phase high performance liquid chromatography (HPLC) was utilized to further increase purity and solubility of the **3FD-LL trimer**. The protein was purified against a water-acetonitrile (+0.1% TFA) gradient and eluted at approximately 43.1% acetonitrile. Lyophilization of the HPLC purified **3FD-LL trimer** protein produced a fine white powder.

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry can be used to determine the exact molecular weight of the purified **3FD-LL trimer** protein. DNA sequencing of the plasmid ensured the DNA sequence was correct at appropriate genetic engineering stages. After expression, the molecular weight was approximated through analysis by SDS-PAGE, but the exact size was only estimated. Mass spectrometry allowed the determination of the exact molecular mass of the pure protein used for assemblies. The expected mass of the **3FD-LL trimer** was 13168.56 Da. MALDI indicated the HPLC purified protein had an approximate molecular weight of 13167.04 based on an average of 4 different scans.

The HPLC purified **3FD-LL trimer** was analyzed by assembling at a protein concentration of 10 mg/mL in 10 mM TAPS buffer at pH 8.5. The protein was dissolved in trifluoroethanol (TFE) prior to the addition of TAPS buffer in order to improve solubility and promote helix formation in a more controlled manner. The assemblies were perfectly clear with no indications of insoluble aggregates. Samples were thermally annealed by slow cooling from 90°C, which supported the formation of the 3FD-LL nanosheets previously published.¹ Characterization using circular dichroism (CD) and transmission electron microscopy (TEM) was performed on the samples to define structural characteristics of the assemblies.

Circular Dichroism Spectropolarimetry

Circular dichroism spectropolarimetry (CD) was used to analyze the nascent and thermally annealed HPLC purified **3FD-LL trimer** assemblies in 10 mM TAPS buffer (pH 8.5). Based on analysis of the HisPur™ proteins, the **3FD-LL trimer** was the most promising candidate due to its higher solubility and greater helicity in 10 mM TAPS buffer at pH 8.5. However, a decrease in signal intensity at lower wavelengths was observed due to low solubility of the peptides in aqueous buffer, which impacts light scattering with the highest hindrance in the low wavelength range.³⁻⁴ CD analysis of the HPLC purified **3FD-LL trimer** displayed a significantly improved CD spectra

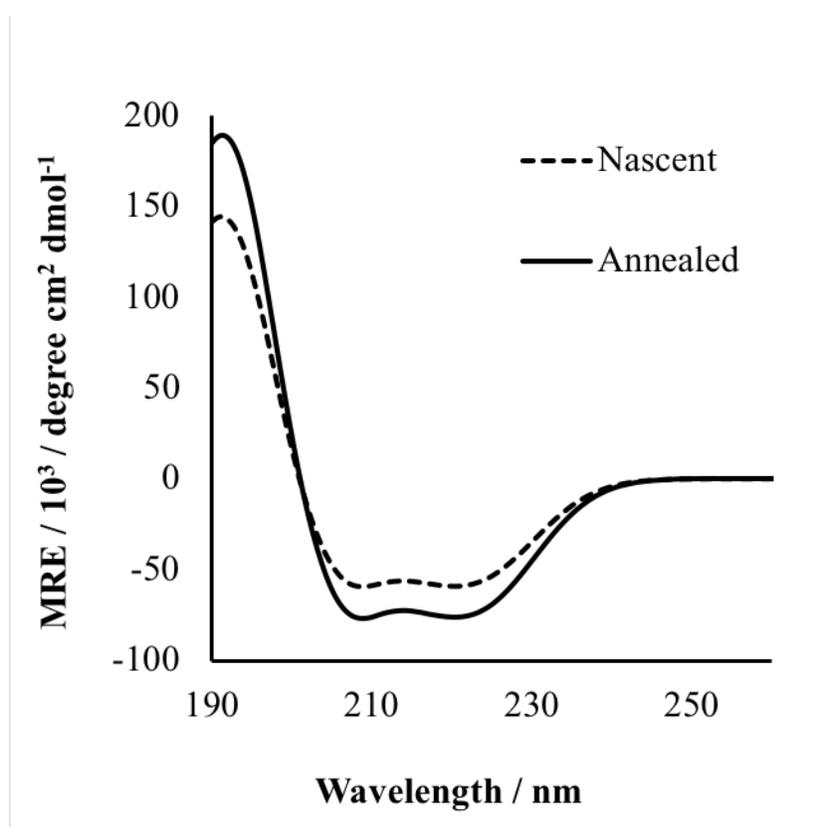


Figure 3.33 Circular dichroism spectra of 1 mg/mL nascent and thermally annealed HPLC purified 3FD-LL trimer assembled in TFE and 10 mM TAPS buffer (pH 8.5).

as transparency issues due to light scattering were minimized (**Figure 3.33**). CD spectra of the assemblies indicated the presence of α -helical secondary structure in the nascent samples along with the preservation and enhancement of the α -helical signature in thermally annealed samples.

Thermally annealed samples of the HPLC purified **3FD-LL trimer** in 10 mM TAPS buffer (pH 8.5) were analyzed by monitoring the ellipticity (MRE) at a wavelength of 222 nm from 5 °C to 98 °C. The resulting CD melting curve displayed diminishing intensity above 90 °C (**Figure 3.34**), and the first derivative indicated that the steepest segment of the sigmoidal curve was present

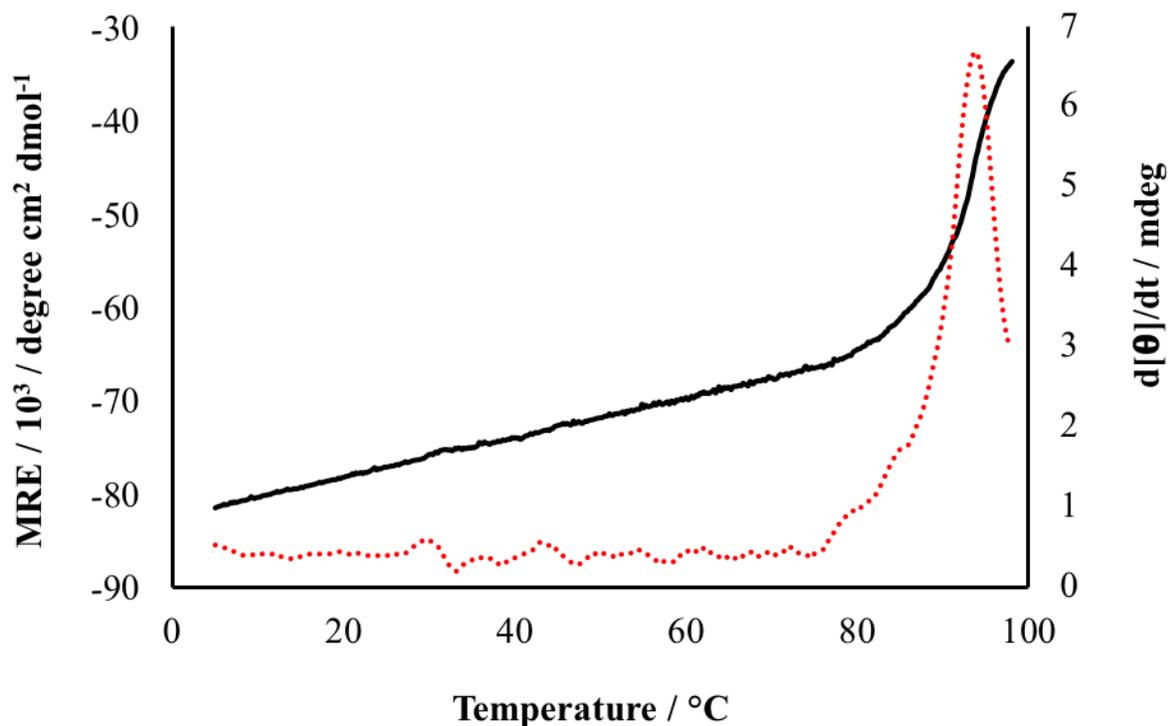


Figure 3.34 Circular dichroism melting curve of 1 mg/mL thermally annealed HPLC purified 3FD-LL trimer assembled in TFE and 10 mM TAPS buffer (pH 8.5). The molar residue ellipticity was monitored at 222 nm from 5 °C to 98 °C. The first derivative of the melting curve is displayed by the red dotted line.

at 94 °C. This would typically indicate the protein was unfolding with a melting temperature (T_m) of 94 °C. However, the atypical sharpness of the transition at the elevated temperature indicated that the decreased ellipticity was likely due to circumstances surrounding evaporation of the solvent. With the exception of the sharp ellipticity change observed just below boiling, there was only a slight decrease in ellipticity across the analyzed temperature range.

Standard CD spectra were collected before and after the melt. The CD spectrum collected after the melt at 98 °C displayed a weaker α -helical signature than the CD spectrum collected before the melt at 5 °C (**Figure 3.35**). This data correlates with the intensities seen at the start and end of the CD melting curve. However, the cuvette was inspected after the CD spectrum was collected at 98 °C and bubbles were observed throughout the sample. The bubbles were likely a product of evaporation of the solvent at extreme temperatures and would result in the significant decrease in signal intensity. The bubbles were quickly removed from path of the light source through gentle tapping, and a second CD spectrum was collected at 98 °C. The CD spectrum produced after the bubbles were moved was only slightly weaker than the CD spectrum collected at 5 °C before the melt. This supported the theory that the sharp transition observed in the melting curve was the result of solvent evaporation. This data indicated that the assemblies exhibited high thermodynamic stability across the analyzed temperature range with solvent evaporation occurring at extreme temperatures.

Previously published data indicated that the **3FD-IL** and **3FD-LL** model peptides displayed melting temperatures of 84 °C and 90 °C, respectively.¹⁻² The **3FD-LL trimer** displayed high thermodynamic stability across the analyzed temperature range with no cooperative melting transition, which indicated that the nanosheets formed by the **3FD-LL trimer** were more thermodynamically stable than the individual model peptides. This may result from the greater

number of interactions present between neighboring helices as a result of the longer length of the helical unit. In addition, the **3FD-LL trimer** formed significantly larger assemblies than the model peptides, which may also cause an increase in melting temperature.

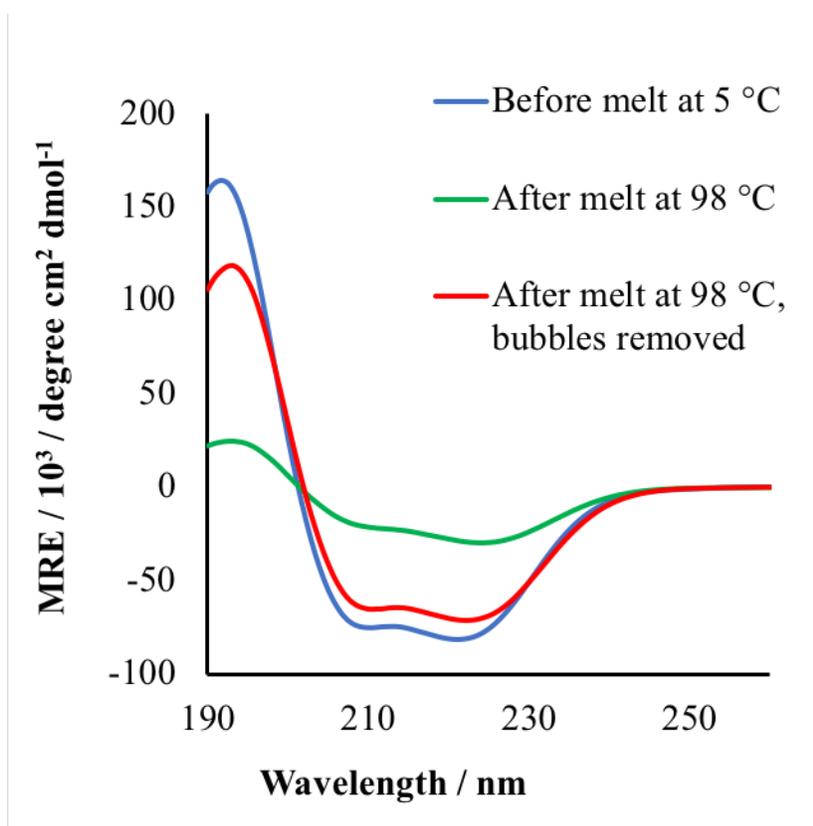


Figure 3.35 Circular dichroism spectra collected before and after CD melt for a 1 mg/mL thermally annealed sample of HPLC purified 3FD-LL trimer assembled in TFE and 10 mM TAPS buffer (pH 8.5). The CD spectra were collected at 5 °C before the CD melt was performed, at 98 °C immediately after the CD melt occurred, and at 98 °C after all bubbles were removed from the path of the light source.

Transmission Electron Microscopy

Transmission electron microscopy (TEM) was used to analyze the HPLC purified **3FD-LL trimer** assemblies in 10 mM TAPS buffer (pH 8.5). The nascent and thermally annealed samples exhibited sheet-like assemblies with improved definition compared to the assemblies produced from HisPur™ protein (**Figures 3.36 and 3.37**). The increased solubility of the HPLC purified **3FD-LL trimer** allowed the formation of a significantly greater quantity of nanosheets. Additionally, there were fewer visible aggregates implying that HPLC purification of the protein removed impurities from the protein sample. Thermally annealed solutions of the HPLC purified **3FD-LL trimer** led to highly reproducible nanosheets on a size scale of hundreds of nanometers

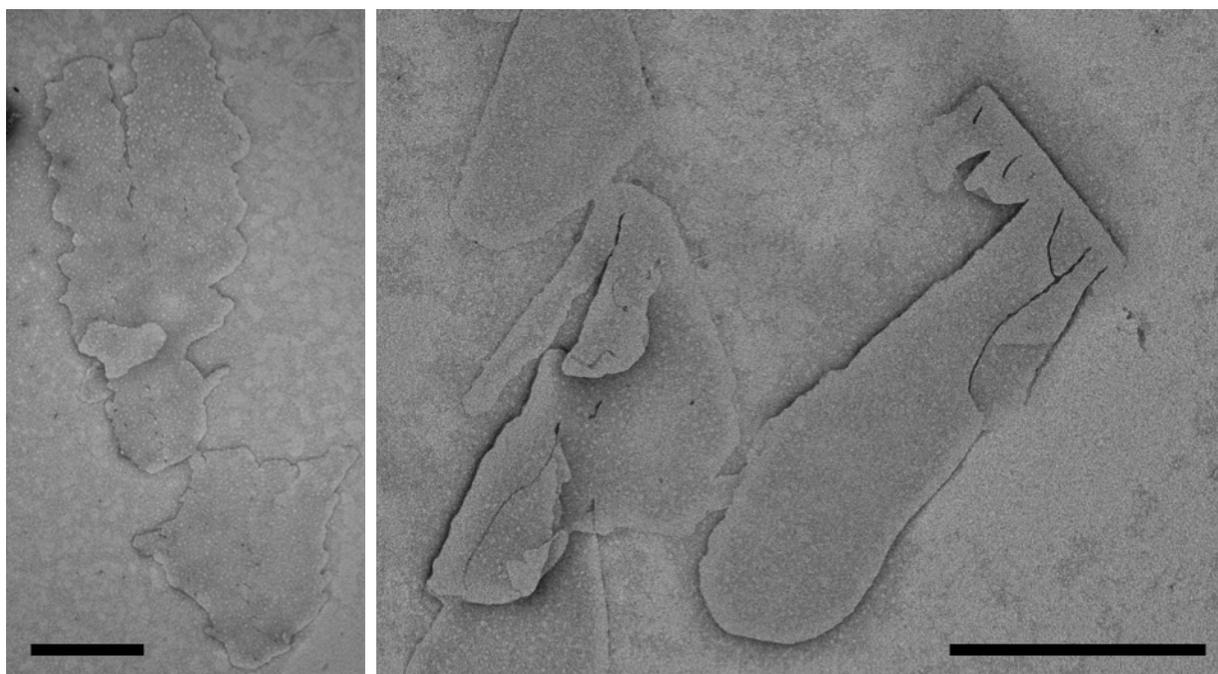


Figure 3.36 Transmission electron microscopy images of nascent 1 mg/mL HPLC purified 3FD-LL trimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). Scale bars = 500 nm.

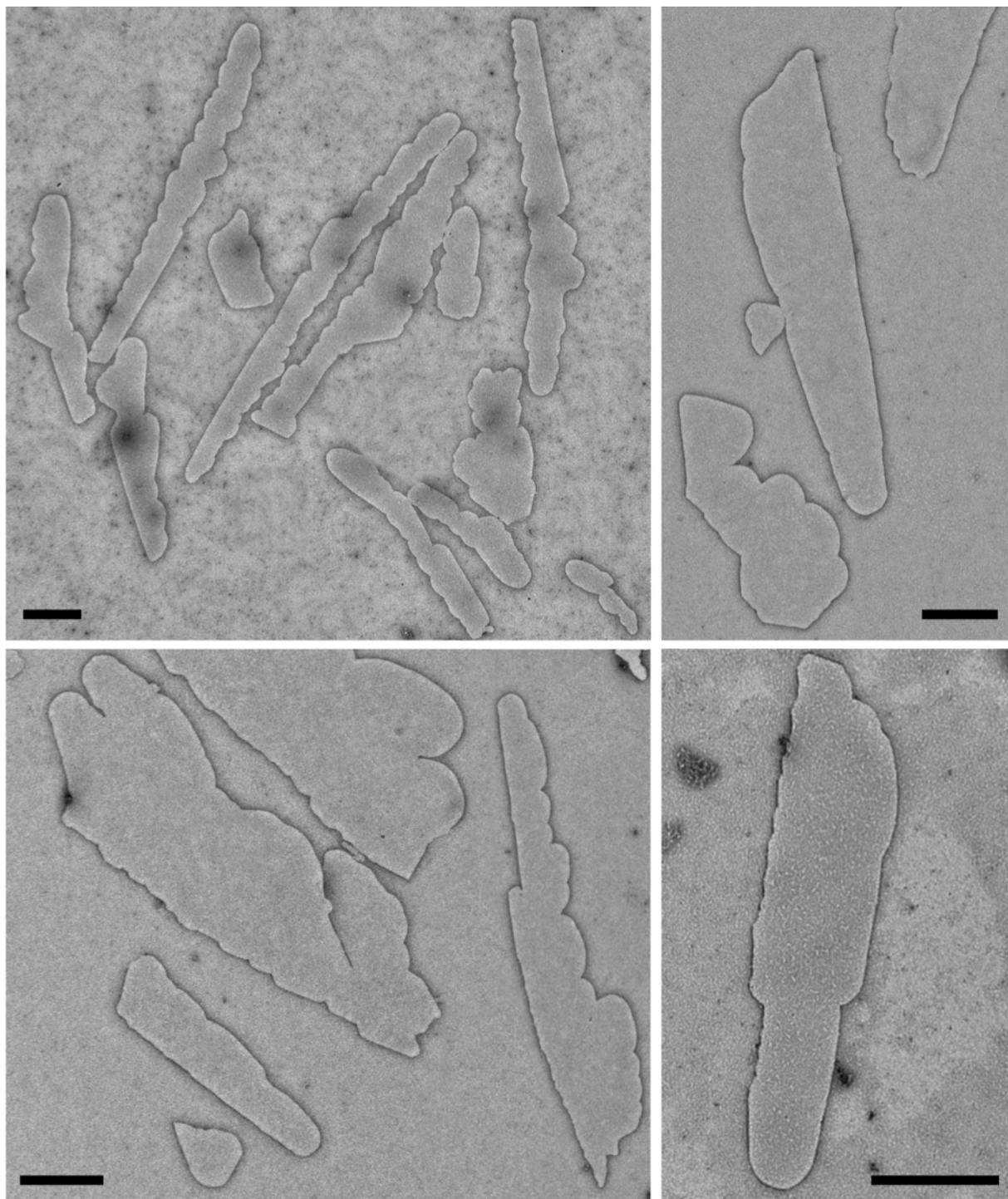


Figure 3.37 Transmission electron microscopy images of thermally annealed 1 mg/mL HPLC purified 3FD-LL trimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). Scale bars = 500 nm.

up to several micrometers in width with highly defined edges. Nascent samples of HPLC purified **3FD-LL trimer** occasionally showed the presence of smaller sheets with less defined edges and cracks traversing the sheets. Thermal annealing reduced the frequency of cracks in the nanosheets due to the controlled assembly process. The two-dimensional assemblies formed by the **3FD-LL trimer** were considerably larger than those produced by the **3FD-IL** and **3FD-LL** peptides.

Atomic Force Microscopy

To assess the thickness of the sheets, the **3FD-LL trimer** assemblies were analyzed using tapping mode atomic force microscopy (AFM). AFM measurements of single layer sheets of **3FD-IL** and **3FD-LL** afforded thicknesses near the length of the α -helix formed by the respective peptides, indicating the peptides packed perpendicularly with respect to the surface of the nanosheet.¹⁻² Assuming the entire α -helix segment of the **3FD-LL trimer** assembled into a uniform α -helical conformation, the theoretical length of the α -helix would be 16.5 nm (110 residues x 1.5 Å rise/residue). In addition to this theoretical thickness, the His-tag would protrude from both faces on the nanosheets since the 3FD-LL peptides were designed to pack antiparallel to each other. The morphology of the **3FD-LL trimer** nanosheets agreed with the morphology seen by TEM analysis (**Figure 3.38 a**). AFM height measurements of the **3FD-LL trimer** assemblies resulted in a few clusters of data (**Figure 3.38 b**).

The main cluster of height measurements, which accounted for approximately 75% of the heights collected, indicated an average height of 17.82 +/- 0.48 nm. The average thickness for this cluster represented the nanosheets formed when the whole α -helix lined up perpendicular to the face of the sheet. The His-tag coating the surfaces of the sheets was likely responsible for the 1.3 nm discrepancy between the theoretical and experimental thicknesses. Partially unfolded segments

of the α -helical conformation may have also contributed to some of the observed divergence. The wide distribution of thicknesses was likely a result of the various potential arrangements of the His-tag on the surface of the sheets.

While a majority of the nanosheets possessed a consistent thickness of approximately 17.82 nm, a couple nanosheets exhibited shorter heights in sections of the sheets. These thinner height changes resulted in three smaller clusters of height measurements with average thicknesses of 12.24 +/- 0.40 nm, 9.58 +/- 0.40 nm, and 6.83 +/- 0.40 nm. These averages likely resulted from the peptide folding in half or at the concatemer intersections. Half of the helical segment would produce a length of 8.25 nm. Folding at one or both of the concatemer intersections within the trimer would result in theoretical lengths of 10.95 nm and 5.55 nm, respectively. A similar size discrepancy of approximately 1.3 nm existed between the theoretical and experimental values, and therefore, was likely due to the His-tag protruding from the nanosheet surface.

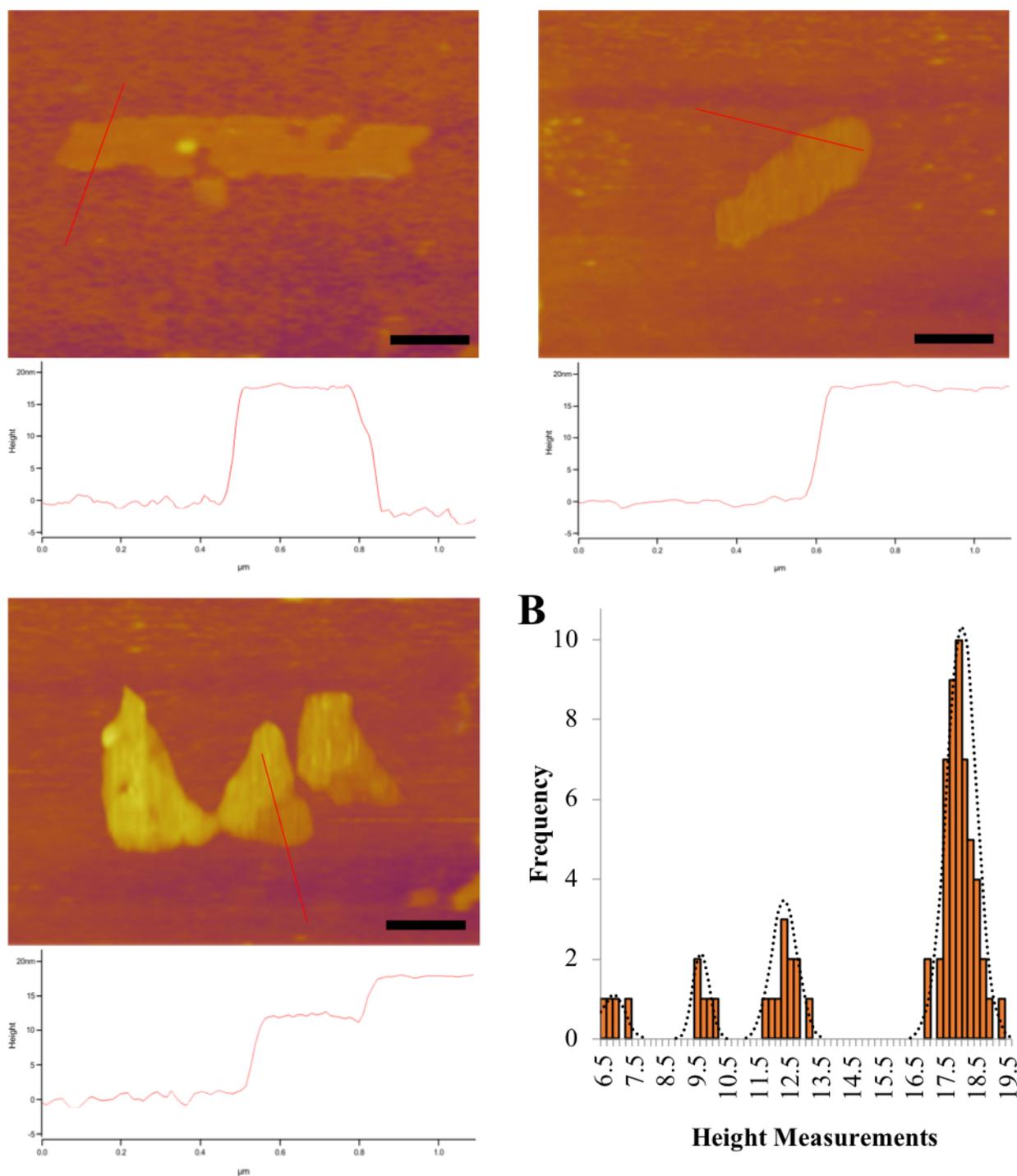


Figure 3.38. Atomic force microscopy of 0.05 mg/mL HPLC purified 3FD-LL trimer assembled with TFE in 10 mM TAPS buffer (pH 8.5). (A) AFM images of thermally annealed assemblies. AFM image in bottom left displays sheets with varying heights. (B) Histogram of height measurements of thermally annealed assemblies. Scale bars = 500 nm.

3.3 Conclusions

As indicated during the purification process of the biosynthesized proteins, solubility was very low for the recombinant proteins. A significant amount of effort was directed at discovering methods for solubilizing the proteins for assembly purposes. The 3FD-LL concatemer proteins were not soluble in aqueous buffer and it was apparent in the circular dichroism and transmission electron microscopy analysis. Resuspending the proteins in 2,2,2-trifluoroethanol (TFE) before addition of the buffer slightly increased the solubility of the proteins; however, even when assembled in TFE, all of the 3FD-LL concatemer proteins displayed a mixture of nanosheets, fibers, and protein aggregates.

Thermal annealing after protein assembly significantly improved the formation of nanosheets. Fragmentation was commonly visualized in nascent assemblies, which decreased considerably after the nanosheets were able to assemble in a controlled manner during slow-cooling. Additionally, thermal annealing led to a lower concentration of unassembled peptide and fibers, which would indicate that the thermal annealing process helped direct the proteins toward nanosheet formation. However, even when the HisPur™ recombinant proteins were assembled in TFE and thermally annealed, extreme heterogeneity of the protein assemblies was still present.

The most promising concatemer produced was the **3FD-LL trimer**, which displayed the strongest α -helical signature during CD analysis and assembled into the most defined nanosheets. Therefore, it was chosen for HPLC purification to remove impurities and promote the homogeneous assembly of nanosheets. The additional purification successfully led to the production of very large nanosheets on the micrometer-scale with highly defined edges. Fragmentation of the nanosheets was not evident during biophysical analysis of thermally annealed

assemblies indicating that the nanosheets possess greater mechanical stability compared to the **3FD-IL** and **3FD-LL** model peptides. Furthermore, the nanosheets did not display a melting transition across the analyzed temperature range indicating that the **3FD-LL trimer** nanosheets possess higher thermodynamic stability compared to the model peptides.

Atomic force microscopy analysis was utilized to collect height measurements of the nanosheets. The height measurements verified that the **3FD-LL trimer** produced nanosheets with greater thicknesses than the **3FD-LL** model peptide; however, the height measurements clustered into a couple height ranges. The majority of the collected heights displayed a height that corresponded closely with the expected length of the α -helical unit. This agreed with the results reported for the **3FD-IL** and **3FD-LL** model peptides, and the data supports the theory that the helices are arranging perpendicularly with respect to the nanosheet surface. However, some of the **3FD-LL trimer** nanosheets exhibited small sections where shorter heights were visualized. These step-wise changes in height resulted in the remaining three groupings of height measurements, which produced averages that correspond to approximately two-thirds, half, and one-third of the expected α -helical length. This could indicate that the protein is folding in half or at the intersections of the concatemer units. Additionally, the lack of structure within the His-tag, which coated both surfaces of the nanosheets, resulted in a wider range of height measurements.

The **3FD-LL trimer** successfully formed thicker nanosheets with increased thermodynamic stability compared to the **3FD-LL** model peptide. Despite the occasional undesirable folding of the α -helical peptide in segments of nanosheets, the advancements made to the 3FD-LL based materials support the use of the biomaterials in a wider range of applications. As stated earlier, thicker porous nanosheets could support improved molecular encapsulation. Further optimization of the assembly conditions would likely mitigate the folding of the α -helical unit.

3.4 Experimental

3.4.1 Materials

All chemical reagents were purchased from either Thermo Fisher Scientific, Inc. (Waltham, MA) or MilliporeSigma (Burlington, MA) unless otherwise noted. The BL21 (DE3) chemically competent *E. coli* strain was purchased from New England Biolabs, Inc. (Ipswich, MA). Luria-Bertani broth and agar powder were purchased from MilliporeSigma (Burlington, MA). The Perfect Protein Marker (10-225 kDa or 15-150 kDa) was purchased from MilliporeSigma (Burlington, MA) and the Precision Plus Protein™ Dual Color Standard was purchased from Bio-Rad Laboratories, Inc. (Hercules, CA).

Benzonase® nuclease and protease-inhibitor cocktail (EDTA-free) were purchased from MilliporeSigma (Burlington, MA). Lysozyme from chicken egg white was purchased from Research Products International Corp. (Prospect, IL). HisPur™ cobalt resin was obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). SnakeSkin™ Dialysis Tubing with a MWCO of 3.5 kDa or 10 kDa MWCO and Slide-A-Lyzer™ G2 Dialysis Cassettes with a 2 kDa MWCO were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

Methylamine tungstate (MAT) was purchased from Ted Pella, Inc. (Redding, CA). The 200-mesh carbon-coated copper grids were obtained from Electron Microscopy Sciences (Hatfield, PA). Vacuum filters with a 0.2 µm polyethersulfone (PES) membrane were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

3.4.2 General Methods

Basic molecular biology procedures were adapted from a standard molecular cloning manual¹⁴ or the protocol supplied by the manufacturer unless otherwise described in detail. All Reagents intended for use with bacteria, DNA, or recombinant proteins were sterilized by either syringe filtration through a 0.2 μm cellulose membrane, vacuum filtration through a 0.2 μm polyethersulfone (PES) membrane, or by autoclaving. All enzyme reactions were conducted in the reagent's buffers provided by the manufacturer unless otherwise noted. Synthetic plasmids were transformed into chemically competent Top10F' or BL21 *E. coli* strains. *E. coli* strains were grown at 37 °C in Lauria-Bertani (LB) medium containing appropriate antibiotic with shaking at 200 rpm unless otherwise stated.

Protein electrophoresis was conducted using 16% SDS polyacrylamide gels on a Mini-PROTEIN 3 cell electrophoresis system from Bio-Rad Laboratories, Inc. (Hercules, CA). The buffer tank was filled with SDS run buffer (25 mM tris, 250 mM glycine, 0.1% SDS, pH 8.3). The Perfect Protein Marker was used as a protein standard for SDS-PAGE analysis. An initial sample volume of 10 μL was added to each well with any necessary concentration adjustments made thereafter. Gels were run at 150 V for 1-1.5 hours depending on the desired level of separation. Gels were stained in coomassie overnight and then destained in a methanol-acetic acid buffer.

Plasmids containing codon-optimized genes of the 3FD-LL concatemers were genetically engineered in Chapter 1. Proteins were expressed using BL21 *E. coli* cells as a host organism and purified using immobilized metal affinity chromatography (IMAC) on HisPurTM cobalt resin followed by dialysis to remove imidazole. Additional purification of the **3FD-LL trimer** protein was achieved by reverse-phase high performance liquid chromatography (HPLC). Assemblies

were thermally annealed using a MJ MiniTM Gradient Thermal Cycler from Bio-Rad Laboratories, Inc. (Hercules, CA).

3.4.3 Initial Assembly and Optimization of 3FD-LL concatemers

Assembly of HisPurTM 3FD-LL Concatemers

The HisPurTM **3FD-LL concatemers** were dissolved in 10 mM buffer at a protein concentration of 10 mg/mL. The buffer conditions were 10 mM CHES (pH 9.0), 10 mM TAPS (pH 8.0), 10 mM MOPS (pH 7.0), 10 mM MES (pH 6.0), 10 mM acetate (pH 5.0), and 10 mM acetate (pH 4.0). The pH of the solution was adjusted to ensure accuracy. Samples were then thermally annealed (see below).

Assembly of HisPurTM 3FD-LL Concatemers with TFE

The HisPurTM **3FD-LL concatemers** were dissolved in trifluoroethanol (TFE) at a protein concentration of 10 mg/mL. An equal volume of 10 mM TAPS buffer (pH 8.5) or 10 mM MOPS buffer (pH 7.0) was added, and the TFE was allowed to evaporate from the solution. Samples were thermally annealed (see below) or incubated at ambient temperature until analysis.

Thermal Annealing of HisPurTM 3FD-LL Concatemers

The 10 mg/mL HisPurTM **3FD-LL concatemer** assemblies were transferred to PCR tubes and thermally annealed using a programmable MJ MiniTM Gradient Thermal Cycler. The protocol heated the protein samples at 90 °C for 10 minutes, and then the samples were cooled at a rate of 0.2 °C/min to a final temperature of 40 °C. Annealed protein samples were stored at 40 °C for two days then stored at ambient temperature until analysis.

Transmission Electron Microscopy (TEM)

The HisPurTM **3FD-LL concatemers** purified by affinity chromatography were analyzed at a protein concentration of 3 mg/mL or 10 mg/mL in 10 mM buffer. The buffer conditions were 10 mM CHES (pH 9.0), 10 mM TAPS (pH 8.5), 10 mM TAPS (pH 8.0), 10 mM MOPS (pH 7.0), 10 mM MES (pH 6.0), 10 mM acetate (pH 5.0), and 10 mM acetate (pH 4.0). The protein samples were either nascent, unannealed, or thermally annealed. Samples were either assembled with or without TFE. Protein assemblies were analyzed using 200 mesh carbon-coated copper grids (Electron Microscopy Sciences) with a 1% methylamine tungstate (MAT) solution as a stain. Samples were prepared by incubating 4 μ L of a protein solution on a grid for 1 minute before wicking excess sample away using filter paper. MAT stain was immediately added to the grid and the solution was allowed to incubate for 1 minute. Stain was wicked away with filter paper and the grid was left to dry under vacuum until imaging. TEM images were obtained on either a Hitachi 7700 TEM or a Hitachi H-7500 TEM. TEM analysis was conducted at accelerating voltages of 80 kV.

Circular Dichroism Spectropolarimetry (CD)

The HisPurTM **3FD-LL concatemers** were analyzed at a protein concentration of 3 mg/mL or 10 mg/mL in 10 mM TAPS buffer (pH 8.5) and 10 mM MOPS buffer (pH 7.0). Samples were either assembled with or without TFE. The samples were either nascent or thermally annealed. CD measurements were collected using a 0.10 mm quartz cell (Hellma Analytics) in a Jasco J-810 CD spectropolarimeter. Sample volumes of approximately 30 μ L were used to generate spectra from 260 to 190 nm at a resolution of 0.1 nm and a continuous scanning rate of 50 nm/min. The CD

spectra were accumulated in triplicate and averaged together. CD spectra were represented by the molar residue ellipticity (MRE) as a function of wavelength.

3.4.4 Assembly and Characterization of HPLC Purified 3FD-LL Trimer

Reverse-Phase High Performance Liquid Chromatography (HPLC)

A fraction of the lyophilized powder of the HisPur™ **3FD-LL trimer** (see previous chapter) was soaked in trifluoroacetic acid (TFA) for 3 hours. Cold ethyl ether (35 mL) was added and the precipitated protein was pelleted using centrifugation for 5 minutes at 4,000 x g (4 °C). The protein pellet was put through a series of ethyl ether washes (5 x 35 mL each) with centrifugation (4000 x g, 5 min, 4 °C), and the final pellet was desiccated. The dried protein pellet was dissolved in 50% acetonitrile in water with 0.1% trifluoroacetic acid (TFA) with vortexing. A Shimadzu HPLC was used to purify the **3FD-LL trimer** on a C18 column using a water-acetonitrile (+0.1% TFA) gradient. HPLC fractions were collected at the targeted time point and analyzed by MALDI-TOF-MS (see below). The collected HPLC fragments with the appropriate mass were transferred to a round bottom flask and a rotavap was used to remove all acetonitrile from the solution and decrease the volume. The remaining solution containing the HPLC pure **3FD-LL trimer** was then poured into 50 mL conical tubes and the tubes were stored at -80 °C. Once the solution was frozen, the tube caps were replaced with Kimwipes™ and the tubes were lyophilized until all that remained was a fine white powder. The powder was weighed and then stored at -30 °C.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS)

MALDI-TOF experiments were performed on an Applied Biosystems® Voyager™ System 4700 mass spectrometer (Life Technologies Corporation; Carlsbad, CA) in the high mass positive

ion linear mode. The matrix, α -Cyano-4-hydroxycinnamic acid (CHCA), was used at a concentration of 10 mg/mL in a mixture of 50% acetonitrile and 0.1% trifluoroacetic acid in deionized water. The **3FD-LL trimer** protein solution (1 mg/mL in ddH₂O or HPLC eluent in approximately 47.8% acetonitrile and 0.1% trifluoroacetic acid in deionized water) was mixed with the matrix solution in a 1:1 ratio. Two microliters of the mixture were spotted on a stainless-steel sample plate and dried under vacuum or air. Sample preparation was repeated in triplicate. Mass spectra were acquired from 1000 laser shots at an intensity of 7900.

Assembly and Thermal Annealing of HPLC Purified 3FD-LL Trimer

The HPLC purified **3FD-LL trimer** was dissolved in trifluoroethanol (TFE) at a protein concentration of 10 mg/mL. An equal volume of either 10 mM TAPS buffer (pH 8.5) or 10 mM TAPS buffer (pH 8.5) with 1 mM EDTA was added, and the TFE was allowed to evaporate from the solution. The sample was then transferred to a PCR tube and thermally annealed using a programmable MJ MiniTM Gradient Thermal Cycler. The protocol heated the protein sample at 90 °C for 10 minutes, and then the samples were cooled at a rate of 0.2 °C/min to a final temperature of 40 °C. The annealed protein samples were incubated at 40 °C for an additional 2 days and then stored at ambient temperature until analysis.

Circular Dichroism Spectropolarimetry (CD)

The HPLC purified **3FD-LL trimer** was diluted from the 10 mg/mL assembly concentration to a peptide concentration of 3 mg/mL in 10 mM TAPS buffer (pH 8.5) or 10 mM TAPS buffer (pH 8.5) with 1 mM EDTA. The samples were either nascent or thermally annealed. CD measurements were collected using a 0.10 mm quartz cell (Hellma Analytics) in a Jasco J-810

CD spectropolarimeter. Sample volumes of approximately 30 μL were used to generate spectra from 260 to 190 nm at a resolution of 0.1 nm and a continuous scanning rate of 50 nm/min. The CD spectra were accumulated in triplicate and averaged together. CD spectra were represented by the molar residue ellipticity (MRE) as a function of wavelength.

CD melting experiments of thermally annealed HPLC purified **3FD-LL trimer** assemblies were performed at 1 or 3 mg/mL in 10 mM TAPS buffer (pH 8.5) by monitoring the signal intensity at 222 nm as the protein samples were heated at a rate of 40 $^{\circ}\text{C}/\text{hour}$ from 5 to 98 $^{\circ}\text{C}$. Parafilm and Teflon were used to wrap the cuvette to minimize evaporation of the solvent at high temperatures. CD melting spectra were represented by the signal intensity as a function of temperature. The first derivatives of the melting spectra were used to acquire the melting temperatures. Standard CD spectra were collected before the melt at 5 $^{\circ}\text{C}$, after the melt at 98 $^{\circ}\text{C}$, and at 98 $^{\circ}\text{C}$ after bubbles were displaced from the path of the light source by gentle tapping.

Transmission Electron Microscopy (TEM)

HPLC purified protein solutions were diluted to 1 mg/mL in 10 mM TAPS buffer (pH 8.5) or 10 mM TAPS buffer (pH 8.5) with 1 mM EDTA. The protein samples were either nascent or thermally annealed. Protein assemblies were analyzed using 200 mesh carbon-coated copper grids with a 1% methylamine tungstate (MAT) solution as a stain. Samples were prepared by incubating 4 μL of a protein solution on a grid for 1 minute before wicking excess sample away using filter paper. MAT stain was immediately added to the grid and the solution was allowed to incubate for 1 minute. Stain was wicked away with filter paper and the samples were left to dry under vacuum until imaging. TEM images were obtained on either a Hitachi 7700 TEM or a Hitachi H-7500 TEM. TEM analysis was conducted at accelerating voltages of 80 kV.

Atomic Force Microscopy (AFM)

AFM samples were prepared from thermally annealed 10 mg/mL **3FD-LL trimer** assemblies in 10 mM TAPS buffer (pH 8.5) that were diluted to a protein concentration of 0.05 mg/mL. The mica plates were cleaved multiple times with scotch tape before sample deposition. Once diluted, the samples were deposited onto mica plates by placing a drop of solution on the mica. The sample was allowed to incubate for 1 minute before the sample was wicked away with filter paper. Samples were also prepared using TEM grids by depositing the 4 μ L of the sample solution onto a copper mesh grid and wicking the sample away with filter paper after 1 minute. Samples were washed with 4 μ L water and then allowed to dry under vacuum or air. Samples were all produced in triplicate. AFM samples were imaged using a MFP-3D-BIO atomic force microscope (Asylum Research) with silicon AFM tips (MikroMasch U.S.A.) applying a force constant of 5.4-16 N/m in tapping mode at a scan rate of 0.5 Hz with 512 points and lines. The set point was 650 mV and the integral gain was 5. IGOR Pro software was used to collect height measurements of the protein assemblies.

3.4.5 Tables

Table 3.1 Protein sequences used in Chapter 3.

Protein	Sequence	MW (Da) # of AA
3FD-LL monomer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₁ -K	5659.47 52
3FD-LL dimer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₂ -K	9414.02 88
3FD-LL trimer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₃ -K	13168.56 124
3FD-LL tetramer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₄ -K	16923.1 160
3FD-LL pentamer	GHHHHHHHHHHHGSGE-[(ALEKLA) ₃ (ALKELA) ₃] ₅ -K	20677.65 196

3.5 References

1. Magnotti, E. L.; Hughes, S. A.; Dillard, R. S.; Wang, S.; Hough, L.; Karumbamkandathil, A.; Lian, T.; Wall, J. S.; Zuo, X.; Wright, E. R.; Conticello, V. P., Self-Assembly of an α -Helical Peptide into a Crystalline Two-Dimensional Nanoporous Framework. *J. Am. Chem. Soc.* **2016**, *138* (50), 16274-16282.
2. Magnotti, E. L. Self-assembly of porous α -helical nanosheets. Emory University, 2016.
3. Dorman, B. P.; Maestre, M. F., Experimental differential light-scattering correction to the circular dichroism of bacteriophage T2. *Proc Natl Acad Sci U S A* **1973**, *70* (1), 255-9.
4. Gregory, R. P.; Raps, S., The differential scattering of circularly polarized light by chloroplasts and evaluation of their true circular dichroism. *The Biochemical journal* **1974**, *142* (2), 193-201.
5. Chakraborty, H.; Lentz, B. R., A Simple Method for Correction of CD Spectra Obtained from Membrane-Containing Samples. *Biochemistry* **2012**, *51* (5), 1005-1008.
6. Mao, D.; Wallace, B. A., Differential light scattering and absorption flattening optical effects are minimal in the circular dichroism spectra of small unilamellar vesicles. *Biochemistry* **1984**, *23* (12), 2667-73.
7. Wallace, B. A.; Mao, D., Circular dichroism analyses of membrane proteins: an examination of differential light scattering and absorption flattening effects in large membrane vesicles and membrane sheets. *Analytical biochemistry* **1984**, *142* (2), 317-28.
8. Tamburro, A. M.; Scatturin, A.; Rocchi, R.; Marchiori, F.; Borin, G.; Scoffone, E., Conformational-transitions of bovine pancreatic ribonuclease S-peptide. *FEBS Letters* **1968**, *1* (5), 298-300.

9. Goodman, M.; Naider, F.; Toniolo, C., Circular dichroism studies of isoleucine oligopeptides in solution. *Biopolymers* **1971**, *10* (9), 1719-1730.
10. Shiraki, K.; Nishikawa, K.; Goto, Y., Trifluoroethanol-induced stabilization of the alpha-helical structure of beta-lactoglobulin: implication for non-hierarchical protein folding. *J Mol Biol* **1995**, *245* (2), 180-94.
11. Luo, P.; Baldwin, R. L., Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* **1997**, *36* (27), 8413-21.
12. Myers, J. K.; Pace, C. N.; Scholtz, J. M., Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. *Protein science : a publication of the Protein Society* **1998**, *7* (2), 383-388.
13. Roccatano, D.; Colombo, G.; Fioroni, M.; Mark, A. E., Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proceedings of the National Academy of Sciences* **2002**, *99* (19), 12179.
14. Sambrook, J.; Russell, D. W.; Sambrook, J.; Russell, D. W., *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press {a} , 10 Skyline Drive, Plainview, NY, 11803-2500, USA: 2001.

Chapter 4

Structurally Defined Helical Nanotubes for the Construction of Functionally Asymmetric Hybrid Nanomaterials

4.1 Introduction

Compartmentalization is essential for proper function of biological systems. The barrier materials responsible for physical separation of two environments also govern the transmission of material, energy, and information across the interface. In order to control the direction of transmission, functional differentiation between the two sides of these interfaces is required. This functional asymmetry for the two surfaces allows for the generation of a gradient between the two separated environments that can be utilized as a thermodynamic driving force for chemical and physical processes. The design of structurally defined nanomaterials exhibiting biologically-inspired asymmetric functionalization would allow for the construction of artificial barrier materials and compartments. These materials would sanction applications that require directional processes or physical separation such as electron transfer, energy transduction, and small-molecule activation, encapsulation, and transport.

Helical protein assemblies formed from the lateral association of subunits have developed into the most structurally economical process for the fabrication of asymmetric interfaces.¹⁻² Helical symmetry of the supramolecular polymers results from the combination of a specific twist angle (φ) and a corresponding axial rise (z) repeated throughout the length of the assembly. Alterations in the individual helical unit can result in a wide variation of helical assemblies each of which is defined in terms of the number of subunits per helical turn, N ($= 2\pi/\varphi$), the helical pitch, P ($= Nz$), and the helical radius, r . Stable axial interactions between successive helical turns can result in the formation of a closed cylindrical structure; however, these interactions are not necessarily required by all applications. In either case, hollow cylinders are formed with a central channel, oriented parallel to the super-helical axis, traversing the structure. Biologically derived helical protein assemblies formed in this fashion display a wide range of functional attributes

making them attractive for applications such as cargo transport (conjugative pili, type 3 secretion system needle complex),³⁻⁴ controlled release and delivery (filamentous phage and viruses),⁵ locomotion (flagellum and archaellum),⁶⁻⁷ signal transduction and actuation (pyrin domains and sterile α -motifs associated with the inflammasome and the signalosome, respectively),⁸⁻⁹ and others.

In order to design helical protein assemblies, non-covalent interactions must be controlled based on a thorough understanding of the relationship between sequence and structure. Despite the rational design of peptides in this manner, structural studies of protein filaments have observed variations in twist angle and axial translation that result in a high degree of structural polymorphism and disorder. Additionally, slight variations in assembly conditions can result in an enormous variety of structurally distinct supramolecular structures. Even at a specific condition, a uniform peptide sequence can produce an assortment of distinct structural assemblies. Isolation of a single macromolecular species with identical characteristics can be challenging, but homogeneous structural assemblies are necessary for controlled application.

Furthermore, a single conservative mutation in a peptide sequence can lead to extremely divergent structural assemblies. Conticello and coworkers recently described two structurally different helical assemblies formed from similar sequences based on a coiled-coil motif. Variations in the sequences by one or two amino acids is sufficient for interchangeability between the two distinct structures.¹⁰ Similarly, Garcia-Seisdedos, et al. demonstrated that a broad variety of native *E. coli* proteins could be driven to assemble into filamentous structures after a few mutations were incorporated at surface interaction “hot spots,”¹¹ which suggested that evolution of certain native proteins occurs on the verge of supramolecular self-assembly with only a few “gate keeper” residues managing this transition.

4.2 Tandem Repeat Proteins (TRPs)

Although flexibility is occasionally a desirable attribute in rational design, unpredictable flexibility of quaternary structure yields perplexing complications to the reliable design of super-helical assemblies. The identification of native peptide and protein motifs that inherently correspond to quaternary structures may allow for the avoidance of this undesirable structural plasticity. One protein supergroup that displays immense potential for the design of peptide-based supramolecular assemblies are tandem repeat proteins (TRPs),¹²⁻¹⁶ which are comprised of repetitive structural motifs of twenty to fifty amino acids that exhibit various levels of sequence identity based on the protein family.¹⁷ The concatenated domains assemble through lateral association between repeat motifs into a highly diverse collection of extended interfaces ranging from super-helical folds¹⁸⁻²¹ to closed and compact conformations with internal pores.²² Since the internal geometry and local packing of the repetitive modular units dictates the overall architecture and subsequent purpose of the structures, TRPs have been found to occupy a diverse range of important biological roles such as scaffolding domains, enzyme catalysis, fibrous assembly, and selective substrate recognition and binding.^{12, 23-24}

Solenoid protein motifs associated with the extended supramolecular assemblies have been analyzed to define the consensus sequences of TRPs, including ankyrin, armadillo, HEAT (Huntingtin, Elongation Factor 3, Protein Phosphatase Subunit A, and Yeast Kinase Tor1), TPR (tetratricopeptide repeat), and LRR (leucine-rich) repeats.^{23, 25-27} Amino acids responsible for lateral association of repeat units are highly conserved, while hypervariability is observed for solvent exposed residues. The hypervariable residues are responsible for biological roles such as the selective binding of macromolecular substrates, and therefore, evolutionary selection of the variable residues can be utilized to direct and improve the application of the supramolecular

assemblies.²³ Computationally designed TRP sequences based on the consensus sequences have successfully resulted in the fabrication of synthetic solenoid proteins with a strong correlation between the assembled structure and computational model.²⁸⁻³⁰

Although most solenoidal TRPs found in nature display strong lateral interactions, weak axial interactions typically exist between successive superhelical turns resulting in open and extended structures. Additionally, many native TRPs have evolved to prevent polymerization through the display of *N*- and *C*-terminal caps, which block uncontrolled non-covalent interactions with other modular units. These naturally designed attributes are beneficial for the specific biological roles of the native TRPs; however, these qualities embody structural deficiencies that are not admirable for the fabrication of synthetic helical nanotubes.

In 1996, Rees and coworkers described the LRV repeat motif from the analysis of the crystal structure of a protein (PDB ID:1LRV) that originated from the nitrogen-fixing bacterium *A. vinelandii* (**Figure 4.1**).²⁶ The protein was composed of two structural domains; a short Fe₄-S₄ cluster and a concatemer of eight helical LRV repeat motifs (Pfam database, PF01816), which was categorized as a 24 amino acid helical hairpin of a 3₁₀ helix-loop-alpha helix. The LRV repeats in 1LRV laterally associated to form a curved surface with the 3₁₀-helix positioned on the exterior surface and the α -helix oriented at the interior surface. The curvature of the structure corresponded to a right-handed super-helix with an average twist angle of 15°. Based on these results collected from the crystal structure, Rees and coworkers predicted that the LRV domain could be utilized for the formation of a super-helical nanotube.

Minard and co-workers described a class of stable HEAT repeat proteins, classified as PBS (phycobilin synthase) HEAT-like repeats (SMART database, SM000567 and Pfam database,

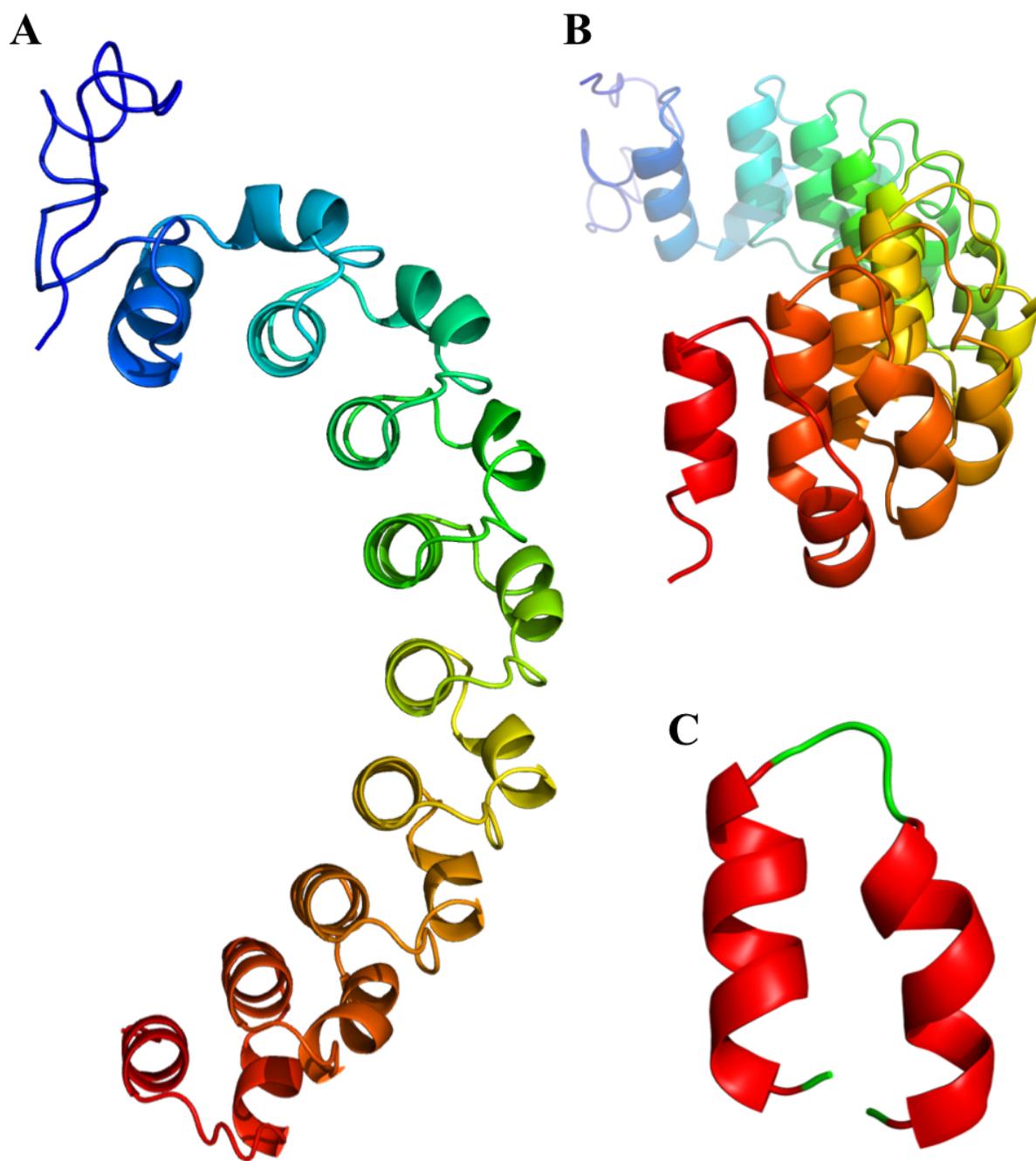


Figure 4.1 Structure of LRV protein based on the crystal structure (PDB code: 1LRV) displaying (A) the internal curvature, (B) the right-handed superhelical nature, and (C) the helical hairpin unit.

PF03130), derived from thermostable microbes.²⁵ The PBS HEAT-like repeat was based on a 31 amino acid helical hairpin motif of an alpha helix-loop-alpha helix. A consensus sequence was generated as GDERAVEPLIKALKDEDWYVRRRAAAEALGEI. Similar to the LRV repeat sequence, highly conserved residues were located at positions responsible for stabilizing the helical hairpin or forming non-covalent lateral interactions with neighboring repeats. The remaining positions were solvent exposed and displayed variability; however, the variability of the two helices of a single hairpin unit displayed significant asymmetry. The solvent facing residues on the surface of the first helix (3, 4, 7, and 11) were composed primarily of polar and helix-stabilizing residues, which limited their variability. The solvent facing residues on the surface of the second helix (18, 19, 22, 23, and 26) displayed hypervariability in residue composition. The hypervariable residues of the second helix could be modified to direct applications as needed.

A series of concatemers were designed from the consensus sequence with the presence of *N*- and *C*-terminal caps to form the α Rep library.²⁵ The crystal structure of the PBS HEAT-like α Rep tetramer was reported (PDB ID: 3LTJ) (**Figure 4.2**), which displayed many resemblances to the LRV repeat assemblies. Similar to the LRV repeat, the structures assembled to form a curved surface with the first helix on the exterior (convex) surface and the second highly variable helix on the internal (concave) surface. Residues that facilitated lateral association of the PBS HEAT-like repeat were highly conserved, whereas the solvent-exposed residues at the convex and concave surfaces were often variable. The structure exhibited a right-handed helical structure with a helical twist of approximately 20°, which indicated that helical packing was similar between the LRV and PBS HEAT-like repeat motifs. Further analysis of concatemers of the same class of HEAT-like repeats characterized their ability to bind various distinct protein partners.³¹ Since both

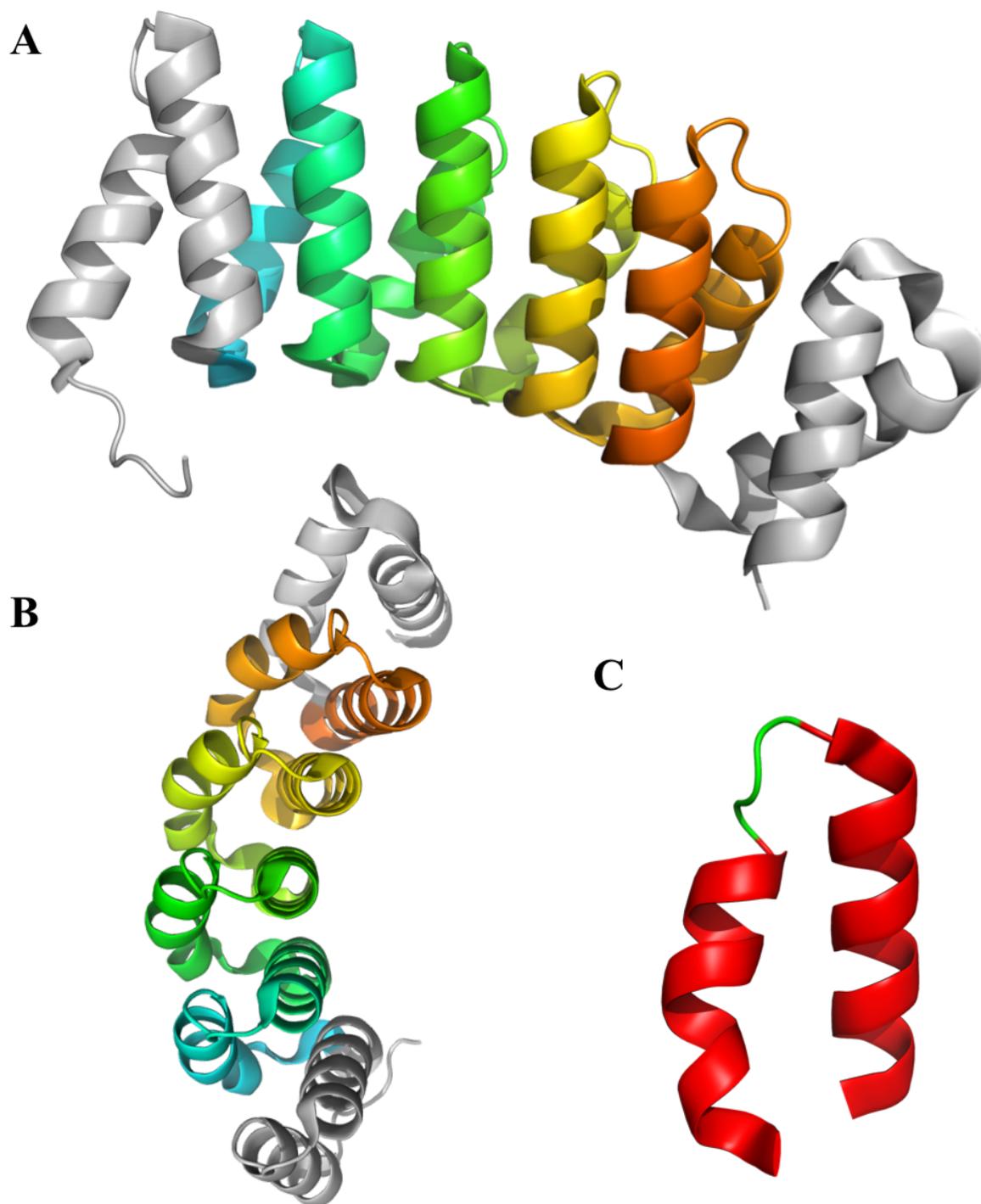


Figure 4.2 Crystal structure of α Rep-*n*4-1 tetramer (PDB ID: 3LTJ) based on the PBS HEAT-like repeat depicting (A) the right-handed superhelical nature, (B) the internal curvature, and (C) the helical hairpin unit. *N*- and *C*-terminal caps are displayed in grey.

the LRV and PBS HEAT-like repeats possessed an *N*- or *C*-terminal cap, neither sequence supported the formation of helical nanotubes. However, based on parallel sequence constraints and structural similarities to the LRV repeat, it is possible that the PBS HEAT-like repeat could be utilized for the formation of super-helical nanotubes.

4.3 HEAT_R1 Peptide Assemblies

Recently, Conticello and coworkers designed the peptide sequence, **HEAT_R1**, based on the consensus sequence of the PBS HEAT-like repeat motif in order to promote the fabrication of helical protein assemblies. A single 30 amino acid repeat unit was designed where charged groups were appointed to surface exposed sites to reduce repulsive electrostatic interactions with careful consideration to conserve the pI of the peptide in the near-neutral pH range. Residues imperative for lateral interactions between structural units were highly conserved in the tandem repeat proteins and remained unaltered in the designed protomer. As a result, the generated 30 amino acid sequence for the **HEAT_R1** peptide was Ac-DERAVEALIKALKDPDWYVRKAAAEALGRI-NH₂, which maintained the native structural features associated with the PBS HEAT-like repeat motif.

Chemical solid-phase peptide synthesis was used to prepare the **HEAT_R1** peptide with subsequent purification via reverse-phase HPLC. Circular dichroism (CD) spectropolarimetry of the peptide indicated a classical α -helical signature with a positive maximum at 193 nm and negative minima at 208 nm and 222 nm (**Figure 4.3 a**). Monitoring the thermal denaturation of the peptide at 222 nm from 25 °C to 90 °C failed to display a melting transition (**Figure 4.3 b**), which indicated that the thermodynamic stability of the peptide was high enough to prevent unfolding within the analyzed temperature range.

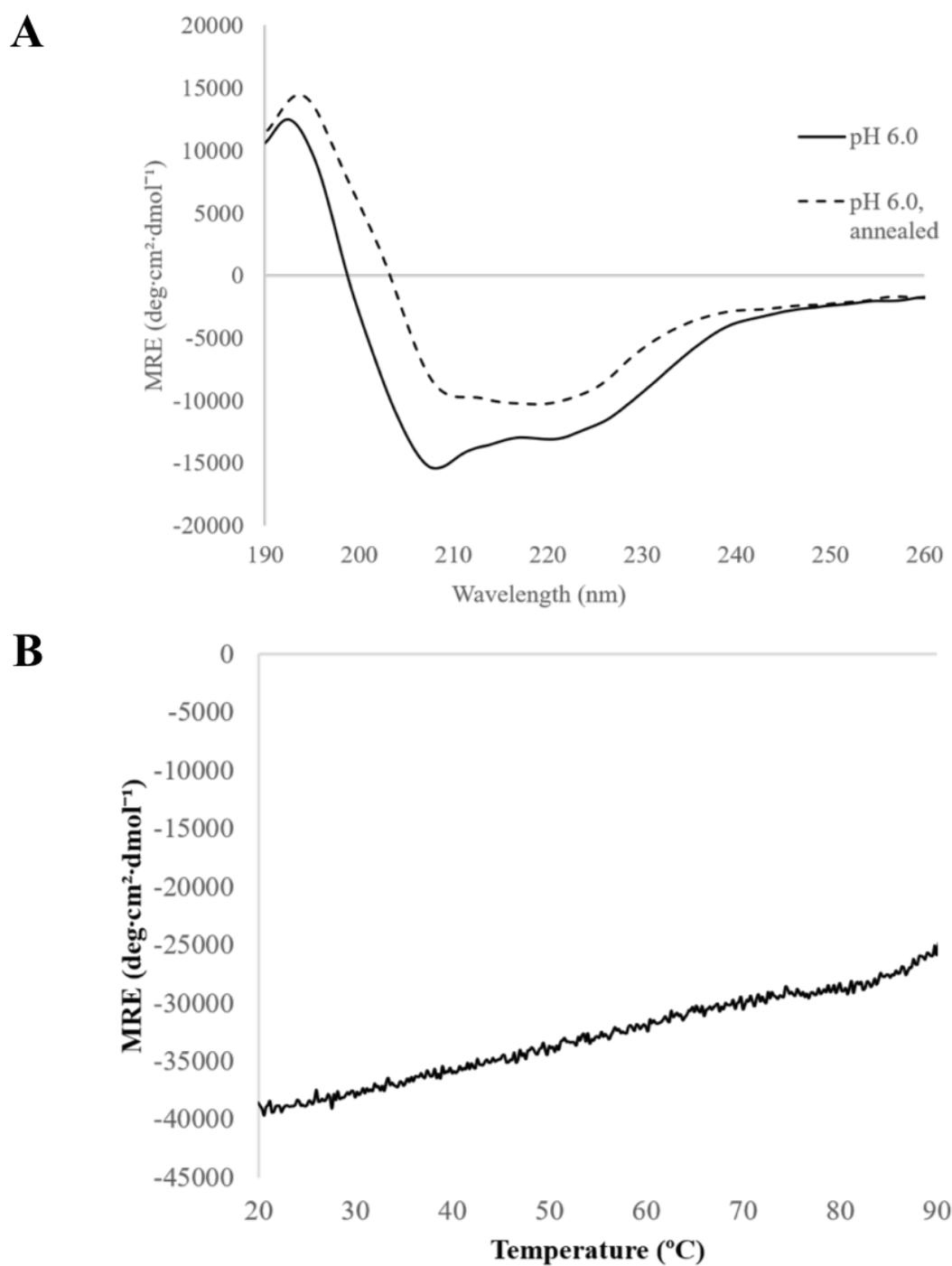


Figure 4.3 (A) Circular dichroism spectra of nascent and thermally annealed assemblies of HEAT_R1 in MES buffer (pH 6.0). (B) CD melting curve of thermally annealed HEAT_R1 assemblies. The molar residue ellipticity was monitored at 222 nm from 20 °C to 90 °C.

Transmission electron microscopy (TEM) of the assemblies of **HEAT_R1** in aqueous buffer presented high aspect-ratio filaments within the pH range from 5-8. Uniform diameters were observed for the filaments with measurements averaging 9.61 ± 0.67 nm. A variance in contrast between the middle and the edges of the structures, which resulted from the presence of negative stain in the inner lumen, indicated that the assemblies were forming nanotubes (**Figure 4.4**). Extremely short nanotube segments were visualized where the nanotubes laid down to present cross-sectional views with the inner lumen entirely exposed. Local unwinding of the structures at the terminal ends of the filaments provided additional evidence for nanotube formation, and the phenomenon may be a consequence of weak axial interactions between subunits of synthetic peptides. Linear power spectrum of the **HEAT_R1** assemblies resulted in a strong axial periodicity

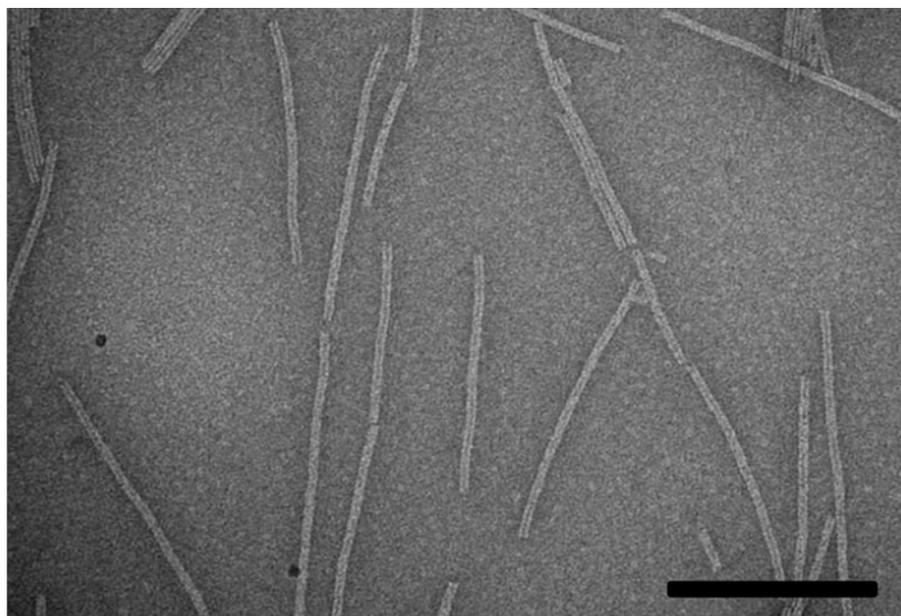


Figure 4.4 Transmission electron microscopy image of thermally annealed HEAT_R1 assemblies in MES buffer (pH 6.0). Scale bar = 200 nm.

at 30.7 Å, which was associated with the helical pitch of the assembly. A mass per length value was calculated to be 111 Da/Å, which gave rise to an estimation of 21 subunits per axial repeat.

Small- and wide-angle X-ray scattering (SAXS and WAXS) indicated the presence of cylindrical structures with uniform diameter. A strong diffraction peak corresponding to a distance of 30.8 Å agreed with the axial periodicity calculated from the linear powder spectrum. The radius of gyration calculated from the modified Guinier model of 34.456 ± 1.156 Å indicated a cross-sectional diameter of 9.7 nm, which was consistent with TEM measurements.

Electron cryomicroscopy (Cryo-EM) was used to generate three-dimensional reconstructions using iterative helical real space reconstruction (IHRSR) to a final resolution of 6.0 Å. The reconstruction of the **HEAT_R1** nanotubes revealed a right-handed helix with 10.3 asymmetric units per turn and a pitch of 31 Å (**Figure 4.5 a**). The asymmetric units comprised a dimer of units for a total of 20.7 HEAT-repeats per helical turn. The dimerization of two adjacent HEAT-repeats peptides occurred through a planar π -stacking interaction between *anti*-oriented tryptophan residues near the N-terminus of the concave helices (**Figure 4.5 b**). The units did not display a highly conserved interaction between a main chain carbonyl group, the Asp14 side chain, and the Arg20 side chain, which is presumed to be responsible for axial interactions between tandem repeats.²⁸ The strong interaction between the tryptophan residues may be responsible for the absence of this highly conserved interaction resulting in distortion of the structure, weakening of axial interactions, and the observed unwinding of filaments in EM images.

Despite the distortion in the asymmetrical units, reconstruction verified the expected orientation of the **HEAT_R1** units with their concave hypervariable α -helices lining the inner lumen while the conserved convex helices lined the exterior surface. The formation of large-pore

nanotubes with a highly functionalizable, hypervariable concave surface could support applications such as the controlled capture or release of specific molecules or modular substrate binding. Additionally, the thermophilic origin of the HEAT-repeat peptide promotes the thermodynamic stability of the assemblies.

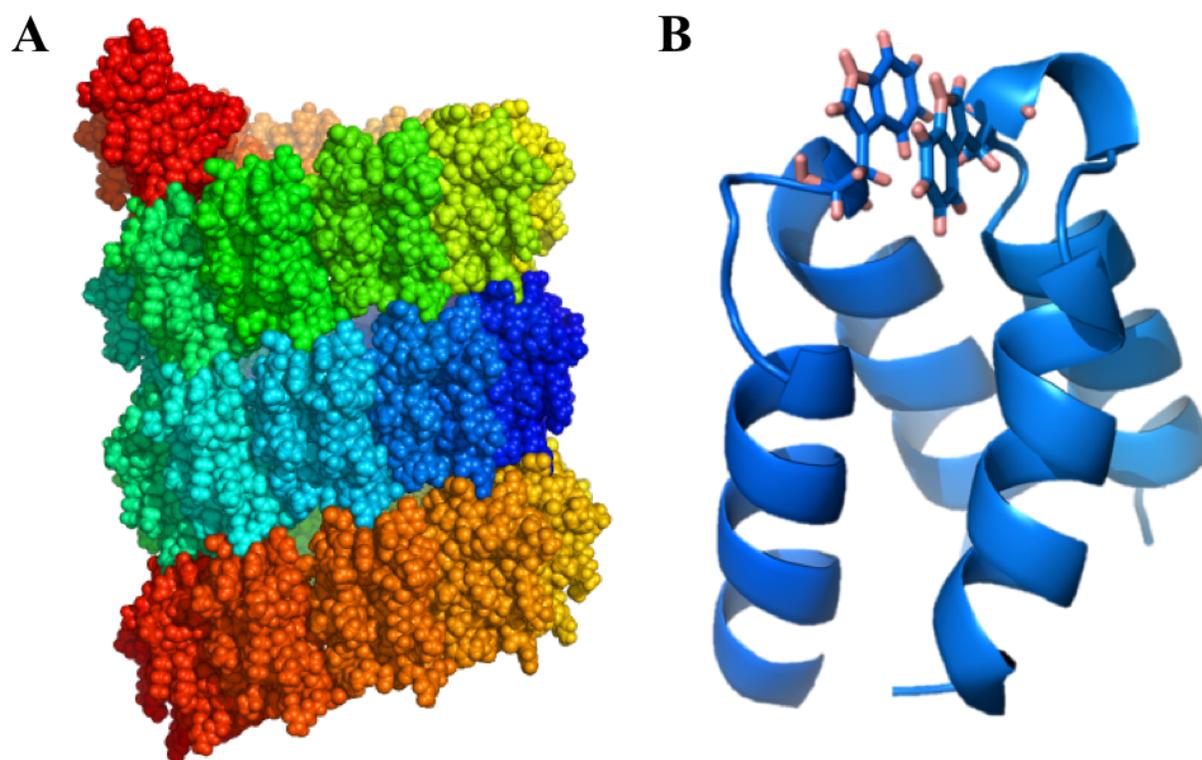


Figure 4.5 (A) Reconstruction of the HEAT_R1 nanotubes revealing a right-handed helix with 10.3 asymmetric units per turn and a pitch of approximately 31 Å. (B) The asymmetric unit of the HEAT_R1 peptide, which was formed from the dimerization of two adjacent HEAT-repeat peptides through a planar π -stacking interaction between anti-oriented tryptophan residues near the N-terminus of the concave helices.

4.4 HEAT_6R Protein Assemblies

A one-dimensional Ising model was used to analyze the folding of various tandem repeat proteins, which implied that the strength of lateral association and the number of repeat units correspond to the stability of the TRP.³²⁻³³ This indicates that concatemers of TRPs may display advantages over single repeat motifs by adopting structures that exhibit higher thermodynamic stabilities. To support this theory, Minard and coworkers described the thermal denaturation of various PBS HEAT-like repeat concatemers ($n = 1, 2, 3, 4,$ and 6) in the α Rep library. The melting transitions for the α Rep- $n1$ -a, α Rep- $n2$ -a, α Rep- $n3$ -a, and α Rep- $n4$ -a were all sigmoidal with melting temperatures (T_m) of $71.0\text{ }^\circ\text{C}$, $77.4\text{ }^\circ\text{C}$, $77.6\text{ }^\circ\text{C}$, and $81.7\text{ }^\circ\text{C}$, respectively.²⁵ The α Rep- $n6$ -a protein could not be unfolded within the analyzed range of $35\text{ }^\circ\text{C}$ to $90\text{ }^\circ\text{C}$. These findings support the data from the Ising model and show that longer concatemers of HEAT-like repeats display increased thermodynamic stability.

It is possible that single repeat units are capable of forming more defined nanotubes than TRP concatemers due to increased rotational freedom of the individual molecular units allowing tighter packing. However, the rotational freedom of the single repeat units may have also promoted and enhanced the distortion resulting from the tryptophan interaction. The covalent connection between neighboring repeats of concatemers may prevent this distortion and lead to the formation of closed cylindrical assemblies with structurally robust, thermally stable helical nanotubes.

Peptide synthesis is not feasible for the chemical synthesis of peptides longer than 70 amino acids in length as the yield decreases exponentially with length. Additionally, production of peptides by solid-phase peptide synthesis yield low quantities at a high cost. Initial biophysical analysis of the peptides requires a small quantity of peptide, but higher yields are required for

application of nanomaterials. Therefore, chemical peptide synthesis would be incapable of producing concatemers of HEAT-based peptides while also improving yield for application purposes.

Similar to methods utilized by Minard and co-workers, genetic engineering and protein expression can permit the construction of novel HEAT-based concatemers in high yields while maintaining sequence control. After optimization of the expression and purification protocols, large quantities can be produced at an affordable price. The **HEAT_R1** sequence was used for the design of a HEAT-based concatemer with six HEAT-like repeat motifs. The sequence contained two alternating HEAT-based sequences with minor sequence modifications from **HEAT_R1**. For one of the sequences, the tryptophan residue located at the *N*-terminus of the concave helix was moved one position further from the turn to mitigate the subunit distortion. Three hypervariable residues in the second sequence were adjusted to have complementary electrostatic interactions with the other HEAT-repeat motif in order to alleviate charge repulsion between the repeat units. A *C*-terminal hexahistidine tag was included to facilitate purification via immobilized metal affinity chromatography (IMAC). The final sequence for the HEAT-based hexamer protein, **HEAT_6R**, was M-(GDERAVEALIKALKDPDGWVRKAAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEI)₃-GSMHHHHHH.

A plasmid containing the coding sequence was transformed into the BL21 (DE3) *E. coli* strain and the **HEAT_6R** protein was expressed with high yield. The concatemer was solubilized in a urea buffer and sequentially purified by immobilized metal affinity chromatography (IMAC) using HisPur™ cobalt resin and reverse-phase HPLC using a water-acetonitrile gradient. Circular dichroism (CD) spectropolarimetry of the hexamer indicated a classical α -helical signature with a positive maximum at 193 nm and negative minima at 208 nm and 222 nm (**Figure 4.6 a**). The

protein exhibited the strongest helical signature at lower pH values with 10 mM acetate buffer (pH 5.0) representing optimal conditions. Similar to **HEAT_R1**, analysis of the helicity before and after thermal denaturation of the **HEAT_6R** protein from 25 °C to 90 °C indicated high thermodynamic stability of the protein (**Figure 4.6 b**). Both **HEAT_R1** and **HEAT_6R** displayed a slight decrease in helical intensity after thermal denaturation, but neither displayed a melting transition. **HEAT_6R** was expected to maintain the high thermodynamic stability since an increase in the number of tandem repeats has been reported to correspond to an increase in stability.

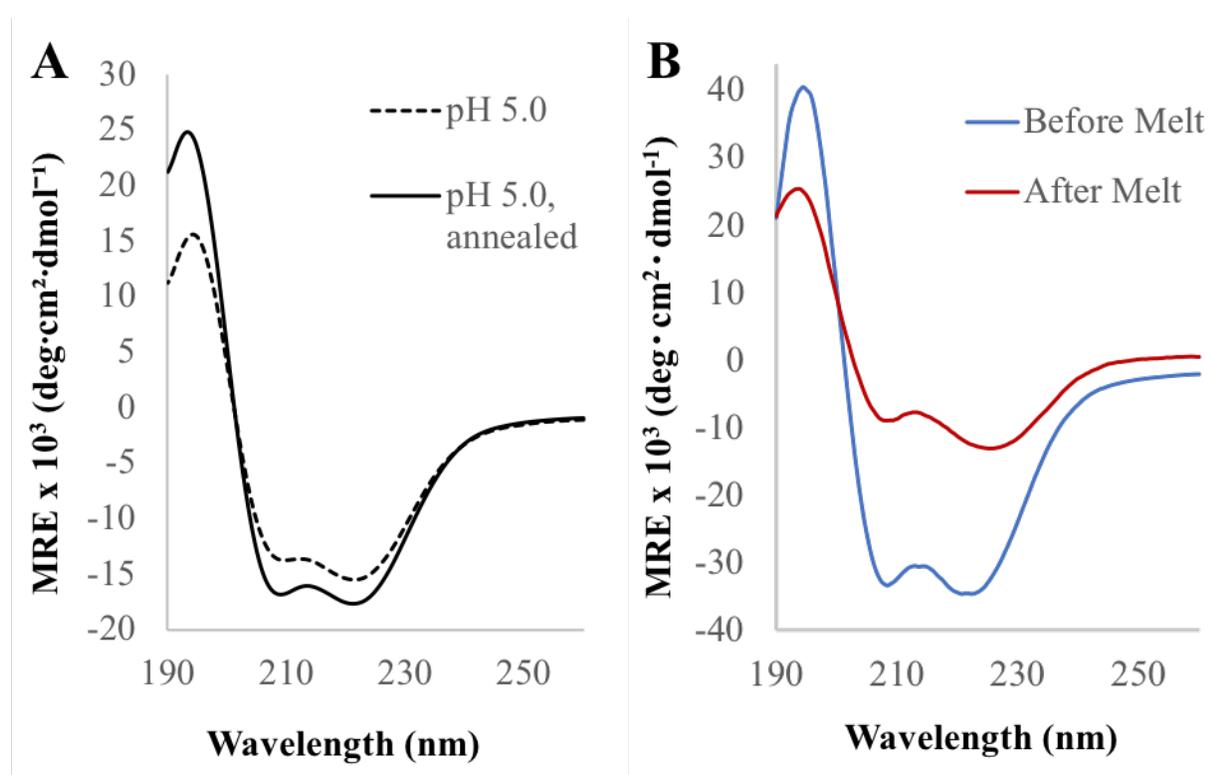


Figure 4.6 (A) Circular dichroism spectra of nascent and thermally annealed assemblies of HEAT_6R in acetate buffer (pH 5.0). (B) Standard CD melting spectra of HEAT_6R assemblies collected before and after thermal denaturation from 20 °C to 90 °C.

Transmission electron microscopy (TEM) of the assemblies of **HEAT_6R** in aqueous buffer presented high aspect-ratio filaments similar to **HEAT_R1**. Differential contrast between the interior lumen and exterior walls indicated the structures were nanotubes (**Figure 4.7**). Contrary to **HEAT_R1**, local unwinding of the structures was not visualized at the terminal ends of the filaments, which indicated that concatemerization of the HEAT-based sequence improved the axial interactions between successive turns of the superhelix. This improvement may result from a decrease in the distortion of the repeat motif due to either the shifting of the tryptophan in the first sequence or the presence of the covalent linkage restricting the distortion.

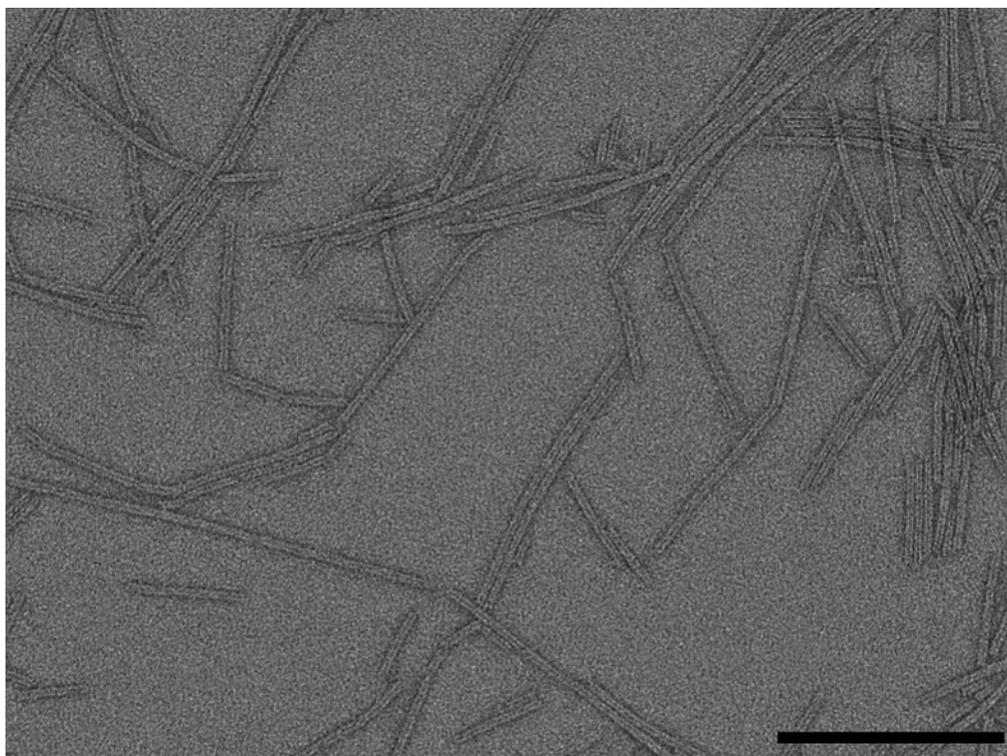


Figure 4.7 Transmission electron microscopy image of thermally annealed HEAT_6R assemblies in acetate buffer (pH 5.0). Scale bar = 200 nm.

Removal of this distortion could allow for the highly conserved interaction between a main chain carbonyl group, the Asp14 side chain, and the Arg20 side chain, which is presumed to be responsible for axial interactions between tandem repeats.²⁸

Small- and wide-angle X-ray scattering (SAXS and WAXS) indicated a monodisperse sample of cylindrical structures. The radius of gyration calculated from the modified Guinier model of $35.965 \pm 1.675 \text{ \AA}$ indicated a cross-sectional diameter of approximately 10.2 nm. This diameter for the **HEAT_6R** nanotubes is similar to the experimental diameter of 9.7 nm for the **HEAT_R1** nanotubes.

These initial biophysical results for **HEAT_6R** indicate that the HEAT-based hexamer is capable of forming closed cylindrical assemblies with structurally robust, thermally stable helical nanotubes. The HEAT-repeat motif displays remarkable tolerance with respect to modifications of the hypervariable residue positions. These hypervariable residues are potentially functional positions that can be assigned with high precision through sequence specificity. The asymmetry in the hypervariability of the convex exterior and concave interior surfaces provides asymmetry to the functionalization of the two surfaces. The HEAT-repeat derived concatemers present a promising building block for the fabrication of helical assemblies and barrier materials with structural and functional asymmetry.

4.5 SpyTag/SpyCatcher Ligation System

The asymmetry between the interior and exterior surfaces of the nanotube promotes the potential for functional differentiation between the two interfaces. The hypervariable residues on the concave surface can be modified without constraints to alter the function of the pore. However, the convex surface exhibits conserved residues that are required for proper assembly of the

nanotubes. Modifications to these amino acids would decrease the stability of the structure or hinder nanotube formation. This obstacle prevents functionalization of the exterior surface of the nanotubes via residue modifications to the convex helices of HEAT-based peptides.

Molecular cloning techniques can be utilized to support the functionalization of the external nanotube surface without altering the conserved convex helix. Fusion of HEAT-based concatemers with small proteins or peptides would potentially allow the formation of HEAT-based nanotubes displaying the functional proteins on the exterior surface exhibiting properties such as substrate adhesion and nanoparticle biotemplating. A gene encoding a fusion protein optimized for expression using *E. coli* as a host organism would allow for the mass production of a functional protein fused to a HEAT-based protein. Optimization of purification and assembly conditions would be necessary prior to verification and characterization of the nanotubes and the functional protein. The biosynthesis and optimization process would have to be performed for every functional protein that is desired for display on the nanotubes since nanotube formation could be affected by the fusion protein.

To further increase the utility of these nanotubes, the display of several proteins on the exterior of the tubes would result in a multifunctional biomaterial. A nanotube system that possesses two differential surfaces each with their own controllable variability would result in a biomaterial with powerful applications. However, since the optimization of protein expression, purification, and assembly of a biomaterial is a time intensive process, the production of various HEAT-based fusion proteins loses its appeal. Additionally, unforeseen complications could arise from the combination of various HEAT-based fusion proteins. The complexity of the assembly solution could be detrimental to nanotube formation. Moreover, functional fusions may be limited to small proteins to prevent hinderance of nanotube formation. Therefore, the biosynthesis of

numerous HEAT-based fusion proteins would not be the optimal approach for the presentation of various proteins to form multifunctional nanotubes.

The ultimate goal would be to produce a simplified system capable of withstanding a boundless potential for functionalization of HEAT-based nanotubes with minimal time and monetary investment. In order to do so, a system would have to be developed where a single HEAT-based fusion protein is biosynthesized and optimized to form nanotubes displaying a short peptide whose function supports the specific attachment of any desired protein. This system would allow for the one-time optimization of HEAT-based fusion protein nanotubes with the ability to display a vast variety of other functional proteins.

In 2014, Joshi and coworkers developed a novel display system called Biofilm-Integrated Nanofiber Display (BIND) where they were able to functionalize biofilms through the expression and secretion of amyloid peptides (CsgA), which form biofilm “curli” fibers, fused to small functional peptides.³⁴ This technique formed a biofilm with functional peptides displayed on the surface and throughout the biofilm. For protein immobilization, they utilized a SpyCatcher/SpyTag linking system that was first developed in 2011 by Howarth and coworkers.³⁵⁻³⁷ The short SpyTag peptide, which was fused to CsgA, was capable of rapidly forming a covalent bond with the SpyCatcher protein, which was fused to GFP (**Figure 4.8**).³⁴ Shortly after that, they published the ability of the BIND system to display enzymes, such as α -amylase.³⁸ Based on its application by the Joshi group, the SpyTag/SpyCatcher linking system appeared to possess attributes that would support the design of a system where a multitude of proteins could be displayed on the surface of the HEAT-based nanotubes.

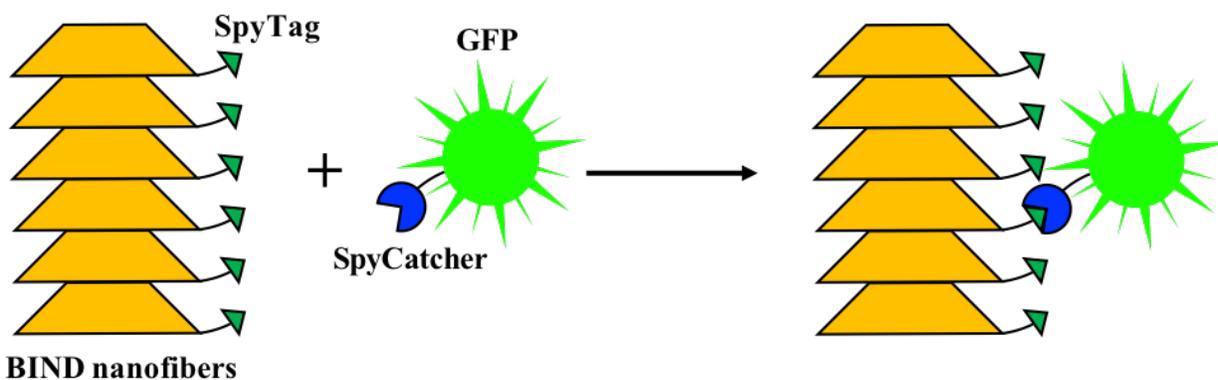


Figure 4.8 Schematic of the biofilm-integrated nanofiber display (BIND) system for the immobilization of proteins. The system utilized the isopeptide bond formed by the SpyTag/SpyCatcher ligation partners to covalently attach proteins, in this case the green fluorescent protein (GFP), to the surface of the biofilm nanofibers.

Howarth and coworkers developed the SpyTag/SpyCatcher protein pair as a peptide tag. Analysis of the fibronectin binding protein (FbaB) of the Gram-positive bacteria *Streptococcus pyogenes* lead to the discovery of a natural intramolecular isopeptide bond in the second immunoglobulin-like collagen adhesion domain (CnaB2).³⁵ Two adjacent residues Lys31 and Asp117, located on two neighboring β -strands, spontaneously form an irreversible isopeptide bond (**Figure 4.9**). The CnaB2 domain was split and rationally modified to form two separate proteins, which they named SpyTag and SpyCatcher, that maintained the ability to spontaneously form a covalent linkage.³⁶ SpyTag was a 13 amino acid peptide tag that was at the C-terminus of the CnaB2 sequence, and SpyCatcher was a 116 amino acid protein from the N-terminal end of the sequence. Structural analysis resulted in the elucidation of the crystal structure for the SpyTag/SpyCatcher complex (PDB: 4MLI), which is shown in **Figure 4.9**.³⁹

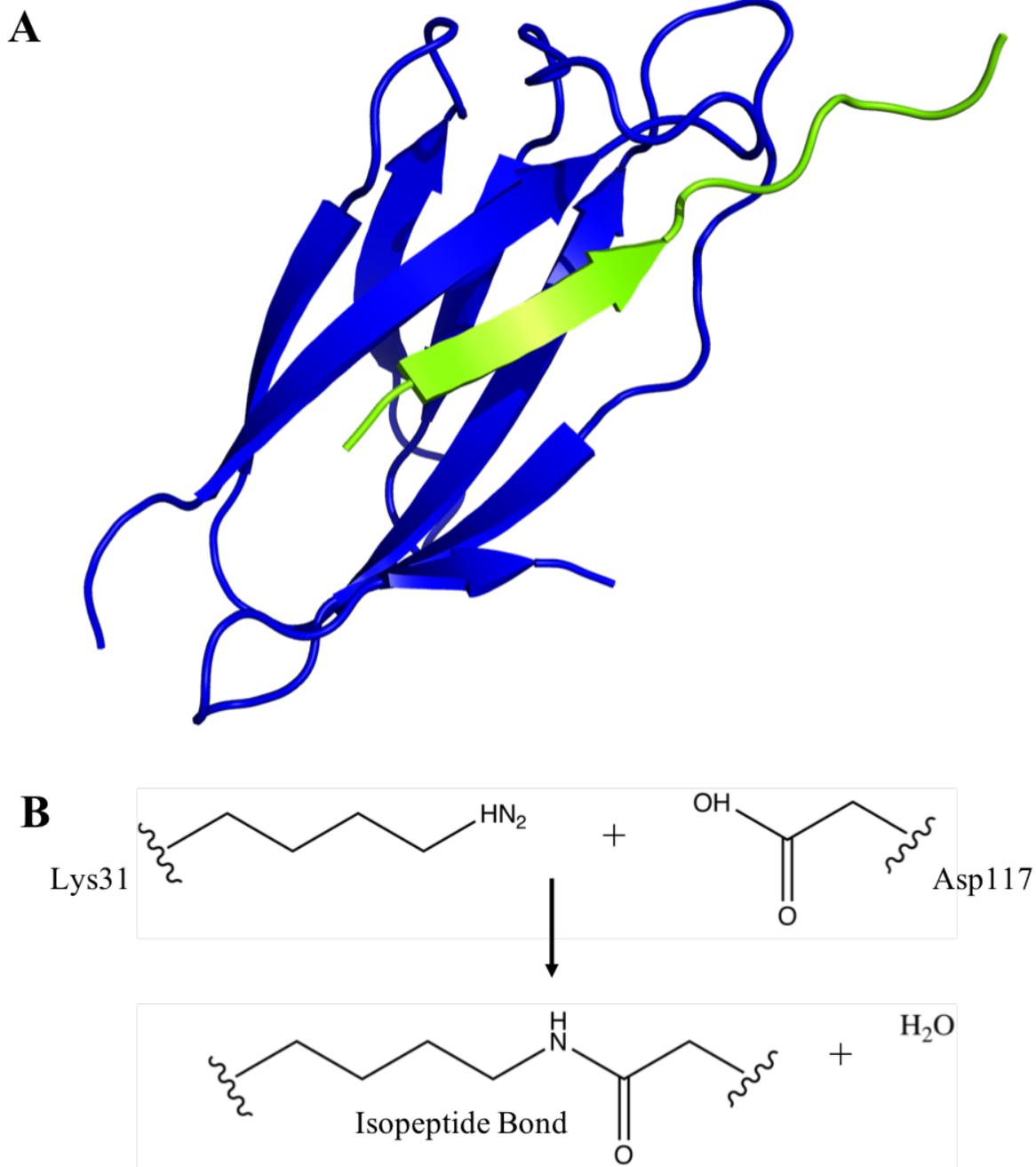


Figure 4.9 (A) Crystal structure of SpyTag/SpyCatcher complex (PDB ID: 4MLI) with the SpyCatcher protein and SpyTag peptide depicted in blue and green, respectively. (B) Reaction mechanism of the Lys31 of SpyTag with the Asp117 of SpyCatcher to form the covalent isopeptide bond.

Analysis of the isopeptide bond formation determined that the link formed rapidly upon mixing with a half-time of 74 s at a protein concentration of 10 μM .³⁶ The reaction process occurred in high yield amidst diverse conditions and was insensitive to changes in temperature, pH, and buffer composition. Once formed, the bond would withstand boiling in SDS and display high thermodynamic and mechanical stability. Single-molecule dynamic force spectroscopy indicated that forces exceeding 1 nN, the force required to break covalent bonds, were required to separate the two protein partners.³⁶ The SpyTag/SpyCatcher system is advantageous over other enzyme ligation systems since no exogenous enzymes need to be added or removed from the solution and the rapid spontaneous reaction is site-selective.

The short SpyTag peptide sequence, AHIVMVDAYKPTK, can be fused to the *N*-terminus, *C*-terminus, or internal section of a protein sequence without impeding the ligation process.^{34, 36, 40} SpyCatcher is a 15 kDa protein with the amino acid sequence GAMVDTLSGLSSEQGQSGD MTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGQKTFV ETAAPDGYEVATAITFTVNEQGQVTVNGKATKGD AHI. Original analysis indicated that the SpyCatcher protein retains its ability to covalently link to the SpyTag peptide when it is genetically fused to the *C*-terminus of a protein or enzyme.^{34, 36, 38, 40} However, there have been a handful of cases where the SpyCatcher protein was able to maintain function after fusion to the *N*-terminus of a protein.⁴¹

In addition to the display of functional proteins on BIND-based biofilm fibers, the SpyTag/SpyCatcher ligation system has been utilized for the display of bioactive peptides, such as the cell adhesive sequence RGD, MMP-1 cleavable sequences, and leukemia inhibitory factor (LIF), within recombinant elastin-like protein hydrogel networks.⁴¹ Utilization of the SpyTag/SpyCatcher ligation system on the exterior surface of the HEAT-based nanotubes would

allow for the precise localization or spatial arrangement of proteins and enzymes, the display of epitopes for immunomodulatory biomaterials, or the targeted nanofiber binding to macromolecular or cellular partners.

4.6 Conclusions

The production of barrier materials with asymmetrically functionalized surfaces allows for the generation of gradients for energy production in addition to the transportation of information in the form of molecules. Extended one-dimensional helical nanotubes are promising due to the presence of the nanopore, but controllability of the structure can be difficult since single sequence mutations can lead to a disordered structure. HEAT-based repeats are an attractive building block for the formation of helical nanotubes because the hypervariability of certain residues allows for tailored functionalization of the nanotube channel without affecting the nanotube architecture. Concatemers of HEAT-based repeats have been reported to display improved thermodynamic stability compared to individual HEAT-like motifs despite the decreased rotational freedom of the structure. The utilization of the SpyTag/SpyCatcher ligation system would allow for functionalization of the exterior conserved surface of the HEAT-based nanotube assemblies to develop materials with two surfaces that are distinct in structure and function.

4.7 References

1. Sachse, C., Single-particle based helical reconstruction-how to make the most of real and Fourier space. *AIMS Biophys.* **2015**, *2* (2), 219-244.
2. Egelman, E. H., Three-dimensional reconstruction of helical polymers. *Arch. Biochem. Biophys.* **2015**, *581*, 54-58.
3. Costa, T. R. D.; Ilangovan, A.; Ukleja, M.; Redzej, A.; Santini, J. M.; Smith, T. K.; Egelman, E. H.; Waksman, G., Structure of the Bacterial Sex F Pilus Reveals an Assembly of a Stoichiometric Protein-Phospholipid Complex. *Cell* **2016**, *166* (6), 1436-+.
4. Loquet, A.; Sgourakis, N. G.; Gupta, R.; Giller, K.; Riedel, D.; Goosmann, C.; Griesinger, C.; Kolbe, M.; Baker, D.; Becker, S.; Lange, A., Atomic model of the type III secretion system needle. *Nature* **2012**, *486* (7402), 276-+.
5. Stubbs, G.; Kendall, A., Helical Viruses. In *Viral Molecular Machines*, Rossmann, M. G.; Rao, V. B., Eds. Springer-Verlag Berlin: Berlin, 2012; Vol. 726, pp 631-658.
6. Lu, A.; Magupalli, V. G.; Ruan, J.; Yin, Q.; Atianand, M. K.; Vos, M. R.; Schroder, G. F.; Fitzgerald, K. A.; Wu, H.; Egelman, E. H., Unified Polymerization Mechanism for the Assembly of ASC-Dependent Inflammasomes. *Cell* **2014**, *156* (6), 1193-1206.
7. Poweleit, N.; Ge, P.; Nguyen, H. H.; Loo, R. R. O.; Gunsalus, R. P.; Zhou, Z. H., CryoEM structure of the *Methanospirillum hungatei* archaeellum reveals structural features distinct from the bacterial flagellum and type IV pilus. *Nat. Microbiol* **2017**, *2* (3), 11.
8. Wang, F.; Burrage, A. M.; Postel, S.; Clark, R. E.; Orlova, A.; Sundberg, E. J.; Kearns, D. B.; Egelman, E. H., A structural model of flagellar filament switching across multiple bacterial species. *Nature communications* **2017**, *8*, 13.

9. Knight, M. J.; Leettola, C.; Gingery, M.; Li, H.; Bowie, J. U., A human sterile alpha motif domain polymerizome. *Protein Sci.* **2011**, *20* (10), 1697-1706.
10. Egelman, E. H.; Xu, C.; DiMaio, F.; Magnotti, E.; Modlin, C.; Yu, X.; Wright, E.; Baker, D.; Conticello, V. P., Structural Plasticity of Helical Nanotubes Based on Coiled-Coil Assemblies. *Structure (London, England : 1993)* **2015**, *23* (2), 280-289.
11. Garcia-Seisdedos, H.; Empereur-Mot, C.; Elad, N.; Levy, E. D., Proteins evolve on the edge of supramolecular self-assembly. *Nature* **2017**, *548* (7666), 244-+.
12. Kajava, A. V., Tandem repeats in proteins: From sequence to structure. *Journal of Structural Biology* **2012**, *179* (3), 279-288.
13. Andrade, M. A.; Bork, P., HEAT REPEATS IN THE HUNTINGTONS-DISEASE PROTEIN. *Nature Genetics* **1995**, *11* (2), 115-116.
14. Doyle, L.; Hallinan, J.; Bolduc, J.; Parmeggiani, F.; Baker, D.; Stoddard, B. L.; Bradley, P., Rational design of alpha-helical tandem repeat proteins with closed architectures. *Nature* **2015**, *528* (7583), 585-+.
15. Groves, M. R.; Hanlon, N.; Turowski, P.; Hemmings, B. A.; Barford, D., The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* **1999**, *96* (1), 99-110.
16. Kobe, B.; Gleichmann, T.; Horne, J.; Jennings, I. G.; Scotney, P. D.; Teh, T., Turn up the HEAT. *Structure with Folding & Design* **1999**, *7* (5), R91-R97.
17. Kobe, B.; Kajava, A. V., When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends in biochemical sciences* **2000**, *25* (10), 509-515.

18. Barkan, A.; Rojas, M.; Fujii, S.; Yap, A.; Chong, Y. S.; Bond, C. S.; Small, I., A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS genetics* **2012**, *8* (8), e1002910.
19. Deng, D.; Yan, C.; Pan, X.; Mahfouz, M.; Wang, J.; Zhu, J. K.; Shi, Y.; Yan, N., Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science (New York, N.Y.)* **2012**, *335* (6069), 720-3.
20. Mak, A. N.; Bradley, P.; Cernadas, R. A.; Bogdanove, A. J.; Stoddard, B. L., The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science (New York, N.Y.)* **2012**, *335* (6069), 716-9.
21. Wang, X.; McLachlan, J.; Zamore, P. D.; Hall, T. M., Modular recognition of RNA by a human pumilio-homology domain. *Cell* **2002**, *110* (4), 501-12.
22. Wierenga, R. K., The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Letters* **2001**, *492* (3), 193-198.
23. Pluckthun, A., Designed Ankyrin Repeat Proteins (DARPin): Binding Proteins for Research, Diagnostics, and Therapy. In *Annual Review of Pharmacology and Toxicology, Vol 55*, Insel, P. A., Ed. 2015; Vol. 55, pp 489-511.
24. Marcotte, E. M.; Pellegrini, M.; Yeates, T. O.; Eisenberg, D., A census of protein repeats. *J Mol Biol* **1999**, *293* (1), 151-60.
25. Urvoas, A.; Guellouz, A.; Valerio-Lepiniec, M.; Graille, M.; Durand, D.; Desravines, D. C.; van Tilbeurgh, H.; Desmadril, M.; Minard, P., Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (alphaRep) based on thermostable HEAT-like repeats. *J Mol Biol* **2010**, *404* (2), 307-27.

26. Peters, J. W.; Stowell, H. B.; Rees, D. C., A leucine-rich repeat variant with a novel repetitive protein structural motif. *Nature structural biology* **1996**, *3* (12), 991-994.
27. Peifer, M.; Berg, S.; Reynolds, A. B., A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell* **1994**, *76* (5), 789-91.
28. Parmeggiani, F.; Huang, P. S.; Vorobiev, S.; Xiao, R.; Park, K.; Caprari, S.; Su, M.; Seetharaman, J.; Mao, L.; Janjua, H.; Montelione, G. T.; Hunt, J.; Baker, D., A General Computational Approach for Repeat Protein Design. *J. Mol. Biol.* **2015**, *427* (2), 563-575.
29. Voet, A. R. D.; Noguchi, H.; Addy, C.; Simoncini, D.; Terada, D.; Unzai, S.; Park, S.-Y.; Zhang, K. Y. J.; Tame, J. R. H., Computational design of a self-assembling symmetrical β -propeller protein. *Proceedings of the National Academy of Sciences* **2014**, *111* (42), 15102.
30. Rämisch, S.; Weininger, U.; Martinsson, J.; Akke, M.; André, I., Computational design of a leucine-rich repeat protein with a predefined geometry. *Proceedings of the National Academy of Sciences* **2014**, *111* (50), 17875.
31. Guellouz, A.; Valerio-Lepiniec, M.; Urvoas, A.; Chevrel, A.; Graille, M.; Fourati-Kammoun, Z.; Desmadril, M.; van Tilbeurgh, H.; Minard, P., Selection of specific protein binders for pre-defined targets from an optimized library of artificial helicoidal repeat proteins (alphaRep). *PLoS One* **2013**, *8* (8), e71512.
32. Kajander, T.; Cortajarena, A. L.; Main, E. R.; Mochrie, S. G.; Regan, L., A new folding paradigm for repeat proteins. *J Am Chem Soc* **2005**, *127* (29), 10188-90.
33. Wetzel, S. K.; Settanni, G.; Kenig, M.; Binz, H. K.; Pluckthun, A., Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J Mol Biol* **2008**, *376* (1), 241-57.

34. Nguyen, P. Q.; Botyanszki, Z.; Tay, P. K.; Joshi, N. S., Programmable biofilm-based materials from engineered curli nanofibres. *Nature communications* **2014**, *5*, 4945.
35. Zakeri, B.; Howarth, M., Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting. *J. Am. Chem. Soc.* **2010**, *132* (13), 4526-+.
36. Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M., Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A* **2012**, *109* (12), E690-7.
37. Kang, H. J.; Coulibaly, F.; Clow, F.; Proft, T.; Baker, E. N., Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science (New York, N.Y.)* **2007**, *318* (5856), 1625-1628.
38. Botyanszki, Z.; Tay, P. K.; Nguyen, P. Q.; Nussbaumer, M. G.; Joshi, N. S., Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. *Biotechnology and bioengineering* **2015**, *112* (10), 2016-24.
39. Li, L.; Fierer, J. O.; Rapoport, T. A.; Howarth, M., Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *J. Mol. Biol.* **2014**, *426* (2), 309-317.
40. Moon, H.; Bae, Y.; Kim, H.; Kang, S., Plug-and-playable fluorescent cell imaging modular toolkits using the bacterial superglue, SpyTag/SpyCatcher. *Chem. Commun.* **2016**, *52* (97), 14051-14054.
41. Sun, F.; Zhang, W. B.; Mahdavi, A.; Arnold, F. H.; Tirrell, D. A., Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111* (31), 11269-11274.

Chapter 5

Design and Production of SpyTag_HEAT and mCherry_SpyCatcher

5.1 Introduction

HEAT (Huntingtin, Elongation Factor 3, Protein Phosphatase Subunit A, and Yeast Kinase Tor1) repeats are attractive for the construction of extended assemblies through the repetitive modular assembly of helical hairpin motifs. Amino acids responsible for lateral association of repeat units are highly conserved, while hypervariability is observed for solvent exposed residues, which are amenable for modifications to direct the application of the nanotubes.¹ The **HEAT_R1** sequence, which was derived from the PBS HEAT-like repeat consensus sequence,² was used for the design of the HEAT-based concatemer protein, **HEAT_6R**, which formed highly stable nanotubes with hypervariable amino acids presented on the inside channel.

Although the internal pore can be functionalized for applications, the external surface of the nanopore is highly conserved. To develop a system with two asymmetric surfaces with distinct structure and function, the SpyTag/SpyCatcher ligation system³⁻⁵ could be used to decorate the exterior surfaces of nanotubes after the HEAT-based nanotubes have been optimized. The small size of the SpyTag peptide was less likely to interrupt the formation of the nanotubes, and therefore, a fusion protein of the SpyTag peptide and a HEAT-based concatemer was designed. The SpyCatcher could then be fused to a variety of functional proteins for rapid functionalization of the nanotubes. Furthermore, a combination of SpyCatcher fusion proteins could be utilized for the creation of multifunctional tubes without interfering with the nanotube structure.

Joshi and coworkers utilized the SpyCatcher/SpyTag linking system for the immobilization of GFP to the biofilm networks prior to advancing to enzymes.⁶⁻⁷ Fluorescent proteins are commonly used for the immobilization assessment of novel systems because it is extremely simple to verify its presence. Similar to Joshi *et al.*, verification of fluorescent protein immobilization

would be desirable before progression to more complex systems. Since most fluorescent proteins are relatively sizable, the impact of linking large proteins to the external surface of the HEAT-based nanotubes could be analyzed.

Initially, a fluorescent protein would need to be selected that is appropriate for the estimated assembly conditions. Initially, superfolder green fluorescent protein (superfolderGFP)⁸ was considered for the fluorescent protein role. However, despite the pH stability of **HEAT_R1** assemblies between pH 5-8, the **HEAT_6R** structures displayed decreased stability at increased pH values with low helicity in solutions above pH 6. One downside of the commonly used superfolderGFP is that conformational changes negatively affect the chromophore as the conditions drop below a pH of 6 and the fluorescent protein very rapidly loses fluorescence.⁹ Engineering has led to the creation of a tandem dimer pH stable GFP (pH-tdGFP), whose fluorescence is stable as low as 5.5.¹⁰ However, pH 5 appeared to be the optimal pH value for the assembly of the **HEAT_6R** concatemer proteins and the utilized fluorescent protein would need to be stable under those conditions.

Red fluorescent proteins typically display lower pKa values than green fluorescent proteins; however, they originally existed as tetramers or dimers like the tetrameric red fluorescent protein drFP583 (DsRed) from *Discosoma* sea anemones.¹¹ Generation of the first truly monomeric red fluorescent protein (mRFP1) was only possible after 33 mutations to the DsRed sequence.¹² Unfortunately, the fluorescence of the protein was seriously diminished compared to DsRed, and the monomer was extremely sensitive to photobleaching. mFruits were designed as second-generation red fluorescent proteins (mRFPs) by Tsien and coworkers through directed evolution.¹³ Although the mFruits all displayed improved brightness and fluorescence, mCherry

and mStrawberry were the most useful with brightness levels of ~50% and ~75% that of EGFP, respectively.

The mCherry fluorescent protein (PDB code: 2H5Q) is photostable and resistant to photobleaching,¹⁴ and the protein has a pKa of <4.5.¹⁵ As a result, the mCherry fluorescent protein has been utilized in over 200 references for *N*- and *C*-terminal fusions for bacterial expression, FRET donor/acceptor pairs, cellular and subcellular imaging, *in vivo* imaging in animals and plants, and protein dynamics and half-life. The mCherry chromophore is composed of three amino acids (Met66, Tyr67, and Gly68) located in the internal helix passing through the β -barrel scaffold (**Figure 5.1**).¹⁶ The surrounding β -barrel produces a microenvironment for the chromophore-forming tripeptide and shields the chromophore from the exterior solution. The mCherry red fluorescent protein has an excitation maximum of 587 nm and an emission maximum of 610 nm. The relative quantum yield is 0.22 and the extinction coefficient is 72,000 M⁻¹cm⁻¹, which gives rise to a brightness of 15,840. The mCherry fluorescent protein is stable under the acidic conditions targeted for this system and was selected for fusion to SpyCatcher.

Plasmids for the two fusion proteins, **SpyTag_HEAT** and **mCherry_SpyCatcher**, were designed and transformed into *E. coli* cells where protein expression was induced. The proteins were successfully purified by affinity chromatography utilizing the His-tags. The ligation capabilities of SpyTag and SpyCatcher were verified throughout the production process to ensure that fusion to the HEAT-based concatemer and mCherry fluorescent protein did not hinder the formation of a covalent linkage.

5.2 Results and Discussion

5.2.1 Generation of ST_HEAT

Sequence Design

In order to develop a system that would allow the ligation of functional proteins to the exterior of pre-assembled HEAT-based tubes, a sequence was designed that comprised a HEAT-based concatemer sequence fused to the SpyTag linking unit. We decided to use a sequence that was based off of the **HEAT_6R** protein; however, we decided that it would be more beneficial to use a HEAT-based concatemer with eight repeats (**HEAT_8R**) in order to space out the SpyTag fusion protein along the outside of the assembled tube. The electron cryomicroscopy reconstruction of **HEAT_R1** indicated that approximately 21 repeat motifs were required to make one full turn of the super-helical assembly. Based on this data, we hypothesized that roughly 2.6 eight repeat concatemers would compose a single turn of the super-helical structure, which would ensure that the SpyTag linkers are staggered across the exterior in successive turns. Due to a correlation between the number of tandem repeats and the stability of the structure,¹⁷⁻¹⁸ the eight repeat concatemer would likely result in a highly stable supramolecular structure. The HEAT_8R sequence was identical to the HEAT_6R sequence, with the exception of the number of repeats. Since the SpyTag linking unit could be linked to either the *N*-terminus or *C*-terminus of a protein and still maintain its linking ability,^{4, 6-7, 19-20} the SpyTag peptide was placed at the *N*-terminus of the designed sequence. A *C*-terminal hexahistadine tag was included to facilitate protein purification using metal affinity chromatography. The designed fusion protein was given the name **ST_HEAT**. **Figure 5.2** depicts the amino acid sequence of **ST_HEAT**.

MGAHIVMVDAYKPTKTSGGGSGGGASGDERAVEALIKALKDPDGWVRKAAAEAL
 GRIGDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALIKALKDPDGWVRK
 AAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALIKALKDP
 DGWVRKAAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALI
 KALKDPDGWVRKAAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGSM
 HHHHHH

Figure 5.2 Amino acid sequence for the ST_HEAT protein. Red letters correspond to the *N*-terminal SpyTag peptide sequence. The blue lettering represents the 8 HEAT-based hairpin repeats. Green letters are the *C*-terminal hexahistadine tag. Molecular weight = 30069.92 Da.

The designed amino acid sequence of **ST_HEAT** was converted to an *E. coli* codon optimized gene sequence and the gene was synthesized by ATUM. The synthetic gene was cloned into the ATUM *E. coli* expression vector pD441-SR, which contained a T5 promoter (*lacI*), a high copy number origin of replication (Ori_pUC), a strong ribosomal binding site, and kanamycin resistance. This high copy number vector would allow for high level production of the designed **ST_HEAT** protein while the repressor allele, *lacI_q*, would ensure tight control over the basal level of transcription prior to induction. **Figure 5.3** depicts the plasmid map of the synthetic gene of **ST_HEAT** within the pD441-SR plasmid vector.

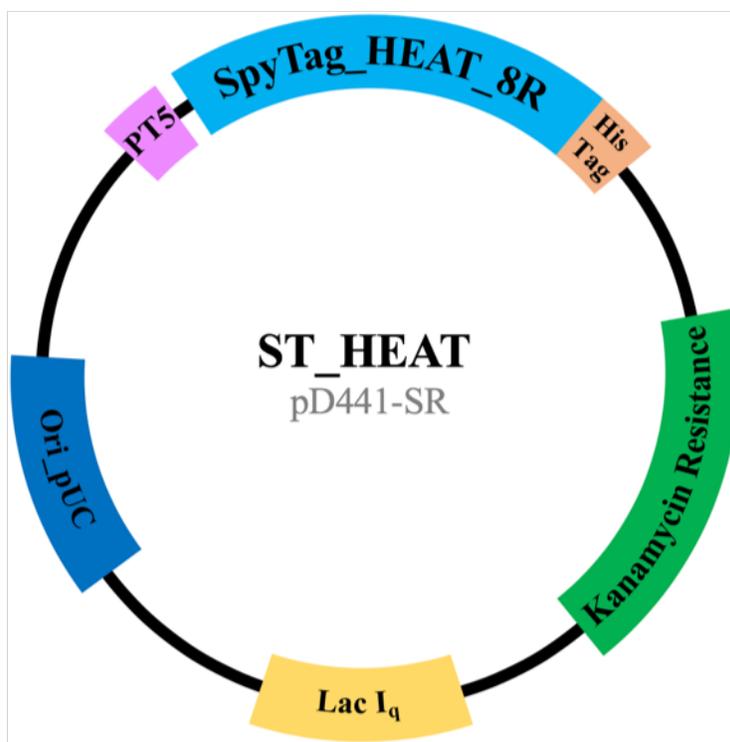
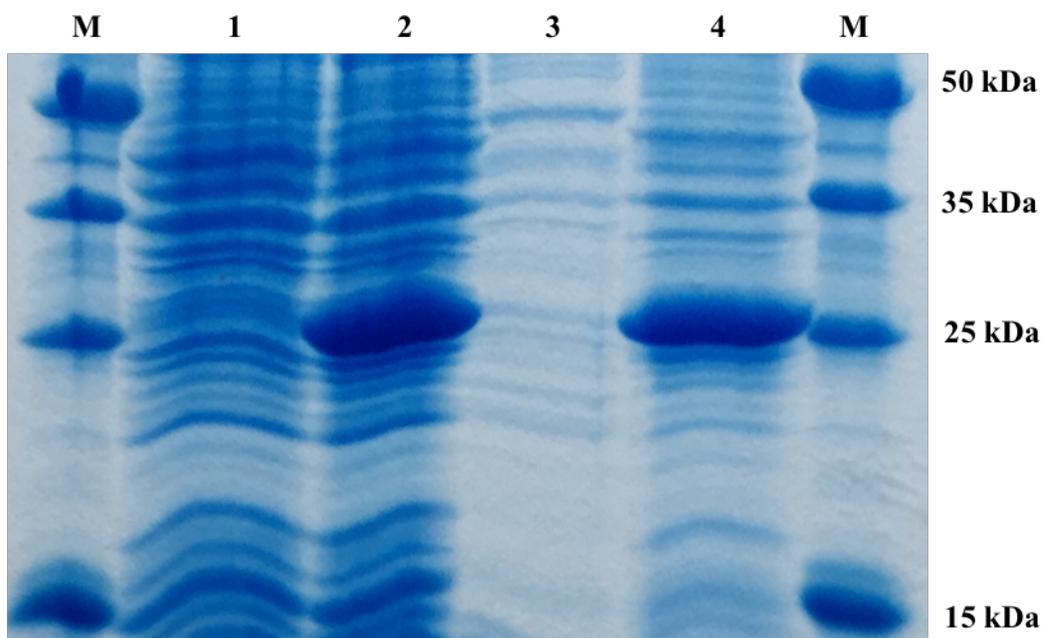


Figure 5.3 Plasmid map for ST_HEAT_8R from ATUM. Plasmid ST_HEAT represents the synthetic ST_HEAT gene contained within the ATUM *E. coli* cloning vector pD441-SR.

Protein Expression and Purification

The ST_HEAT protein was designed for expression using *E. coli* as a host. The BL21 strain was the target bacterial host for protein expression, while the Top10F' strain was used for the production and isolation of plasmid DNA. Large-scale expression of the ST_HEAT protein in BL21 cells was carried out in LB media at 30 °C. A significant amount of protein production after induction with IPTG was confirmed by SDS-PAGE analysis (**Figure 5.4**). After the cells were harvested, they were lysed using multiple methods including freeze/thaw cycles, enzymatic lysis, and sonication. The soluble and insoluble fractions were separated by

centrifugation and analyzed using SDS-PAGE. A 16% SDS-PAGE gel indicated that the **ST_HEAT** protein was insoluble as it was localized in the lysate pellet as seen in **Figure 5.4**.



M. Perfect protein marker (10-225 kDa)

3. ST_HEAT, lysed supernatant

1. ST_HEAT, expression, before induction

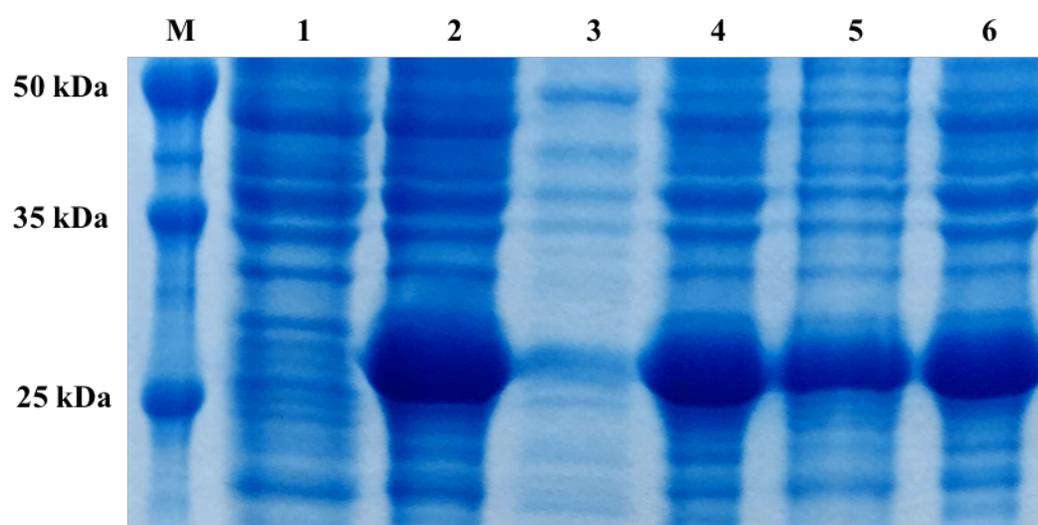
4. ST_HEAT, lysed pellet

2. ST_HEAT, expression 15 hrs after induction

M. Perfect protein marker (10-225 kDa)

Figure 5.4 SDS-PAGE analysis indicating the results of protein expression and cell lysis for the **ST_HEAT** protein. A distinctive protein expression band was evident for the induced sample (lane 2). The **ST_HEAT** protein was present in the insoluble pellet after cellular lysis (lane 4). MW: **ST_HEAT** = 30.1 kDa.

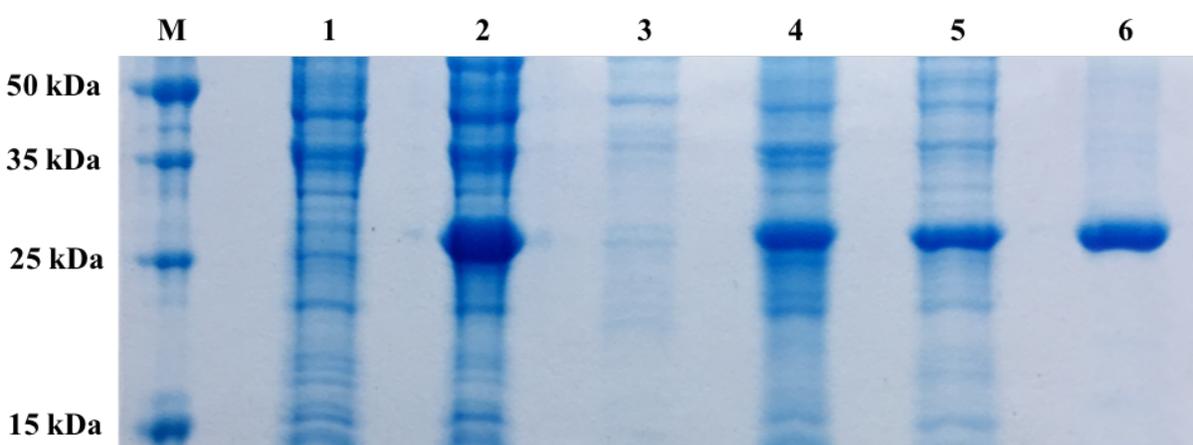
In order to facilitate purification on a cobalt resin column, the protein was removed from the pellet and solubilized. Initial attempts to solubilize the protein in 6 M urea was moderately successful. Analysis using a 16% SDS-PAGE gel determined that a relevant portion of the **ST_HEAT** protein was recovered in the soluble phase after shaking in 6 M urea (**Figure 5.5**). This was not surprising as it agreed with the results seen during purification of the **HEAT_6R** protein.



- | | |
|---|-------------------------------------|
| M. Perfect protein marker (10-225 kDa) | 4. ST_HEAT, lysed pellet |
| 1. ST_HEAT, expression, before induction | 5. ST_HEAT, urea pellet |
| 2. ST_HEAT, expression, 15 hrs after induction | 6. ST_HEAT, urea supernatant |
| 3. ST_HEAT, lysed supernatant | |

Figure 5.5 SDS-PAGE analysis indicating the solubilization results of the ST_HEAT protein in 6 M urea. The target protein band representing the ST_HEAT protein was visualized in the soluble supernatant of the 6 M urea solution (lanes 6). Some ST_HEAT protein remained in the pellet after solubilization (lane 5). MW: ST_HEAT = 30.1 kDa.

The hexahistidine tag on the urea solubilized **ST_HEAT** protein facilitated purification on a cobalt resin column and subsequent elution by imidazole. SDS-PAGE analysis was utilized to determine relative purity and mass of the **ST_HEAT** protein within the eluent. A 16% SDS-PAGE confirmed that the **ST_HEAT** eluent was highly pure, and the band created by the eluent was appropriate for the expected mass of the **ST_HEAT** protein (**Figure 5.6**).



M. Perfect protein marker (10-225 kDa)

4. **ST_HEAT**, lysed pellet

1. **ST_HEAT**, expression, before induction

5. **ST_HEAT**, urea supernatant

2. **ST_HEAT**, expression, 15 hrs after induction

6. **ST_HEAT**, HisPur™ eluent

3. **ST_HEAT**, lysed supernatant

Figure 5.6 SDS-PAGE analysis depicting the purification results for the **ST_HEAT** protein by affinity chromatography using HisPur™ cobalt resin. The purified eluent was dialyzed against dH₂O and lyophilized to obtain powder. The pure **ST_HEAT** protein is visualized in lane 6. MW: **ST_HEAT** = 30.1 kDa.

Dialysis of the purification eluent in ddH₂O was intended to remove urea, imidazole, and other small molecules from the solution. After a couple days of dialysis, the protein precipitated out of solution, which was expected as the urea concentration decreased. The dialyzed protein was lyophilized to isolate the protein as a white powder. The protein yield was 178 mg / 2 L expression. A higher protein yield would have likely been obtained if all of the protein was solubilized in urea. However, the recovered amount was sufficient for assembly and characterization of the **ST_HEAT** nanotubes (see next chapter).

5.2.2 Generation of mCh_SC

Sequence Design

The second designed sequence contained the mCherry fluorescent protein, which acted as the functional protein, and SpyCatcher, which was the second half of the protein ligation system. The mCherry fluorescent protein was selected because of the desired assembly conditions that our lab previously optimized for the **HEAT_R1** peptide and **HEAT_6R** protein. **HEAT_R1** and **HEAT_6R** assembled the best under acidic conditions in the pH range of 4-6. The mCherry fluorescent protein has a low pKa of less than 4.5,¹⁵ and is consequently stable under the acidic conditions that were investigated for this system. Although there are a few references reporting the success of fusion proteins with SpyCatcher on the *N*-terminus,²⁰ most cases in the literature indicate that the *C*-terminus of SpyCatcher should remain unaltered and exposed to prevent hinderance of the ligation process.^{4, 6-7, 19} Therefore, the design placed the SpyCatcher protein at the *C*-terminus of the fusion protein to construct mCherry_SpyCatcher.

To allow for protein purification using metal affinity chromatography, the inclusion of a *N*-terminal His-tag was required. However, it would be preferable to be able to remove the His-

tag after protein purification. The tobacco etch virus (TEV) protease has been engineered to act as a useful molecular engineering tool with applications in removing fusion tags from recombinant proteins.²¹ The TEV protease is a highly specific cysteine protease with a very strict cleavage recognition site directed by the seven amino acid sequence Glu-Asn-Leu-Tyr-Phe-Gln↓Gly (**Figure 5.7**). Since cleavage occurs before the glycine, unnecessary amino acids are not afflicted on a recombinant protein that immediately follows the recognition site. TEV protease and the cleaved fusion tag can be captured based on the properties of the cleaved fusion tag, which in this case was the His-tag.

The ATUM *E. coli* expression vector PD441-NHT contains an *N*-terminal 6x His-tag followed by a TEV protease cleavage site immediately before the multiple cloning site. We selected this vector to house the designed gene, **mCh_SC**, which was intended for insertion immediately after the glycine of the TEV cleavage recognition site. **Figure 5.8** depicts the amino acid sequence of **mCh_SC** before TEV cleavage.

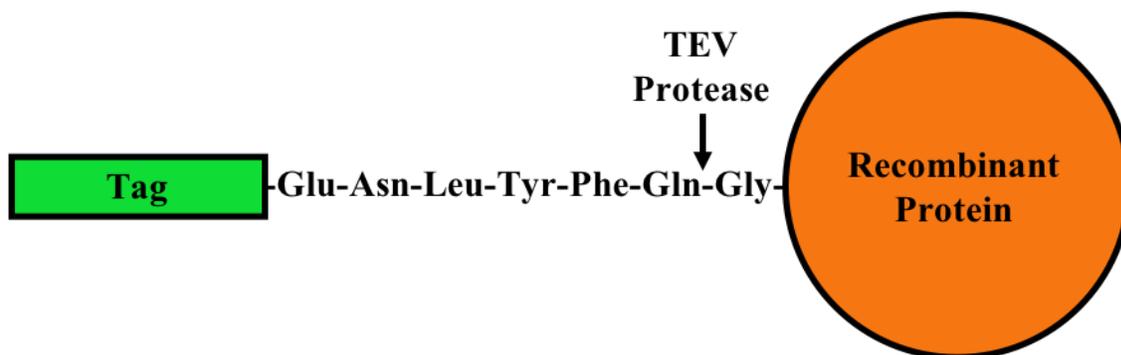


Figure 5.7 TEV cleavage recognition sequence and cleavage site.

MKHHHHHGTSENLYFQGMSSGLVPRGSHMVSKGEEDNMAIIEFMRFKVHMEGS
 VNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPA
 DIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPDGP
 VMQKKTMGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVQ
 LPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKNSGGGLVAGGSG
 GSGGGTGGGSGGGTSGAMVDTL SGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGK
 ELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTV
 NEQGQVTVNGKATKGDAHI

Figure 5.8 Amino acid sequence for the mCh_SC protein before TEV cleavage. Red letters correspond to the mCherry protein sequence. The blue lettering represents the C-terminal SpyCatcher sequence. Green letters indicate the N-terminal hexahistadine tag. The TEV cleavage recognition site is shown in purple. Molecular weight: before TEV cleavage = 44298.7 Da; after TEV cleavage = 42406.74 Da.

The designed amino acid sequence of **mCh_SC** was converted to an *E. coli* codon optimized gene sequence and the gene was synthesized by ATUM. The synthetic gene was cloned into the ATUM *E. coli* expression vector pD441-NHT, which contained a T5 promoter (*lacI*), a high copy number origin of replication (*Ori_pUC*), a strong ribosomal binding site, and kanamycin resistance in addition to the N-terminal His-tag and TEV cleavage site. Similar to the **ST_HEAT** plasmid, this vector would guarantee tight control over pre-induced transcription by the *lacI_q* repressor allele while promoting a high level of protein production by the high copy number origin

of replication. **Figure 5.9** depicts the plasmid map of the synthetic gene of **mCh_SC** within the pD441-NHT plasmid vector.

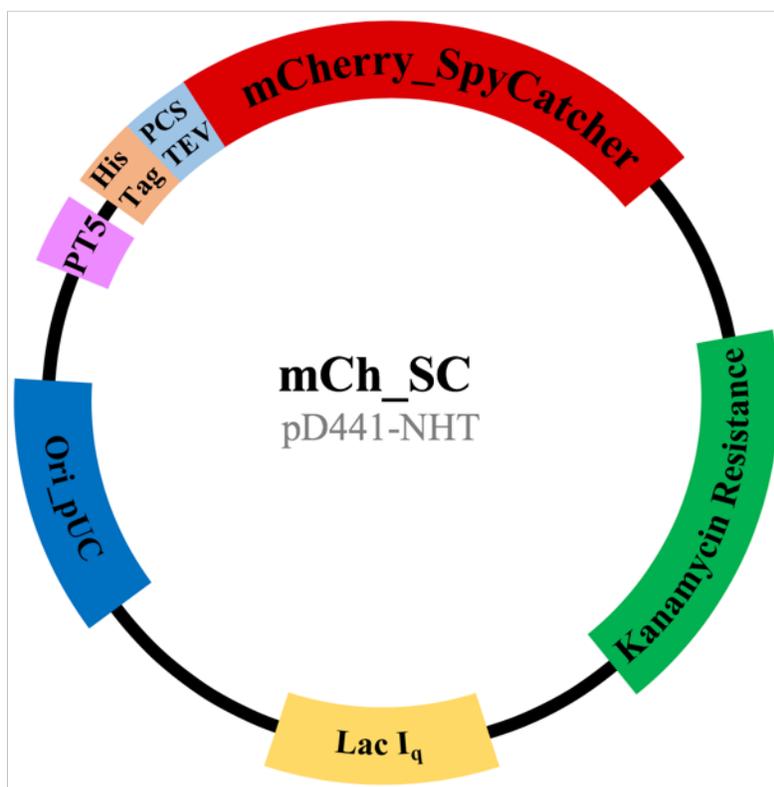


Figure 5.9 Plasmid map for mCh_SC from ATUM. The plasmid represents the synthetic mCh_SC gene contained within the ATUM *E. coli* cloning vector pD441-NHT.

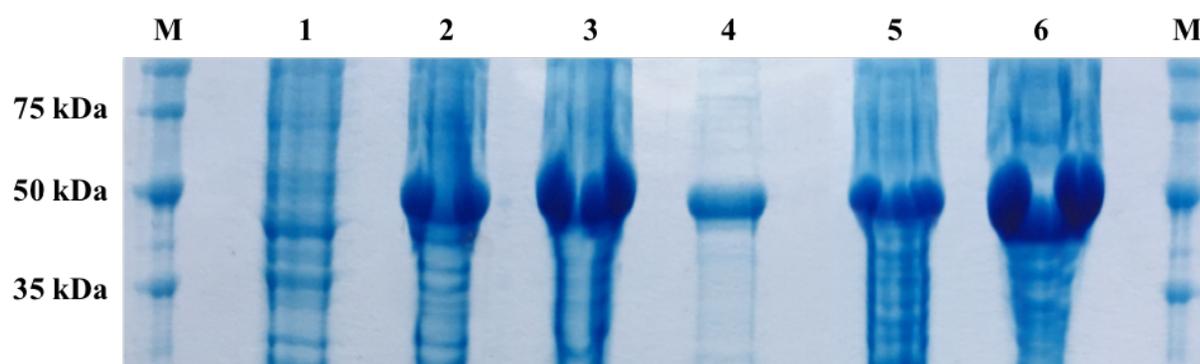
Protein Expression and Purification

The **mCh_SC** protein was designed for expression using *E. coli* as a host. The isolation and purification of plasmid DNA was carried out in the Top10F' strain while the BL21 strain was the intended bacterial host for over-expression of the protein. Large-scale expression of the **mCh_SC** plasmid in BL21 cells promoted high-level expression under the tight control of the inducible, phage T5 promoter. It was obvious that protein expression was successful due to the observable color change of the Terrific Broth media into a dark red (**Figure 5.10**).



Figure 5.10 Media after the expression of the **mCh_SC** protein before the cells were harvested. The solution displayed a deep red color, which indicated successful high yield production of **mCh_SC**.

After harvesting the cells from the expression culture, a significant amount of dark red remained in the Terrific Broth expression media. This could suggest that some of the **mCh_SC** protein was being secreted from the cells. Analysis by SDS-PAGE confirmed that the **mCh_SC** fusion protein was present in all expression aliquots including the expression media after the cells were harvested. (**Figure 5.11**). Although it is not always the case, extracellular auto-secretion of sfGFP and mCherry into culture media has been reported in multiple *E. coli* strains including BL21



- | | |
|---|---|
| M. Perfect protein marker (10-225 kDa) | 4. mCh_SC, expression media |
| 1. mCh_SC, expression, before induction | 5. mCh_SC, lysed pellet |
| 2. mCh_SC, expression, 1.5 hrs after induction | 6. mCh_SC, lysed supernatant |
| 3. mCh_SC, expression, 4 hrs after induction | M. Perfect protein marker (10-225 kDa) |

Figure 5.11 SDS-PAGE analysis indicating the results of protein expression and cell lysis for the mCh_SC protein samples. Distinctive protein expression bands were evident for the induced samples (lane 2 and 3). The mCh_SC protein was still present in the expression media after the cells were harvested (lane 4). The mCh_SC protein was primarily in the soluble supernatant after cellular lysis (lane 6). MW: mCh_SC = 44.3 kDa.

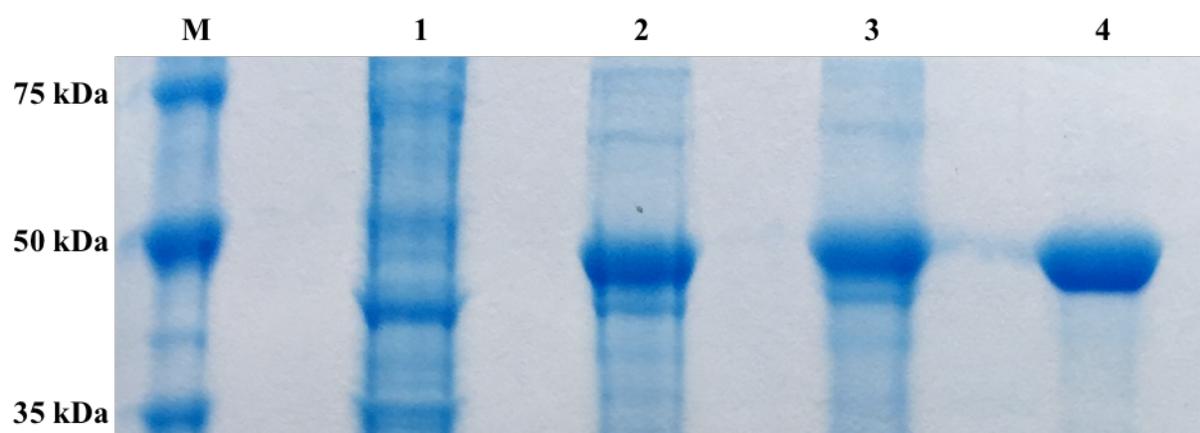
(DE3) when the fluorescent proteins were over-expressed.²² The secretion was observed for the individual fluorescent proteins in addition to fusions with recombinant proteins at their C-termini.

The cells harvested from the expression culture were lysed using multiple methods including freeze/thaw cycles, enzymatic lysis, and sonication. Centrifugation was applied to the lysate to separate the soluble and insoluble fractions. The lysate supernatant exhibited an intense purple color (**Figure 5.12**), which was indicative of successful **mCh_SC** production and very high protein yields. A 12% SDS-PAGE gel confirmed that the majority of the **mCh_SC** protein was soluble as it was localized in the lysate supernatant as seen in **Figure 5.11**.



Figure 5.12 Supernatant after cell lysis and centrifugation of cells containing the expressed recombinant protein, **mCh_SC**. The lysed supernatant solution displayed an intense purple color, which indicated the presence of the **mCh_SC** protein in the soluble phase.

The *N*-terminal His-tag was exploited for the purification of **mCh_SC** lysed supernatant by metal affinity chromatography with cobalt resin. The Terrific Broth expression media that contained a significant amount of **mCh_SC** was also purified. The eluent exhibited a deep purple color confirming the presence of the **mCh_SC** protein. The mass of **mCh_SC** and the relative purity of the eluent was analyzed using SDS-PAGE. A 12% SDS-PAGE gel confirmed that the **mCh_SC** eluent was reasonably pure, and the band created by the eluent was appropriate for the expected mass of the **mCh_SC** protein (**Figure 5.13**).



M. Perfect protein marker (10-225 kDa)

3. mCh_SC, lysed supernatant

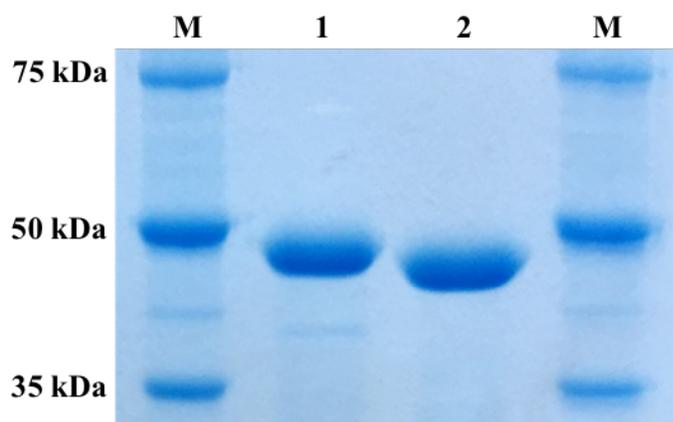
1. mCh_SC, expression, before induction

4. mCh_SC, HisPur™ eluent

2. mCh_SC, expression, 4 hrs after induction

Figure 5.13 SDS-PAGE analysis depicting the purification results for the **mCh_SC** protein by affinity chromatography using HisPur™ cobalt resin. The purified eluent was dialyzed against dH₂O and lyophilized to obtain powder. The pure **mCh_SC** protein is visualized in lane 4. MW: **mCh_SC** = 44.3 kDa.

Once the pure protein was acquired, the worth of the His-tag was exhausted. In order to remove the His-tag, we exploited the enzymatic capabilities of the TEV protease for detachment of the His-tag by utilizing the TEV protease cleavage recognition site contained within the sequence. After incubation of the pure **mCh_SC** with the TEV protease, the TEV protease enzyme and attached His-tag were captured from solution by metal affinity chromatography. The TEV-cleaved pure **mCh_SC** was analyzed by SDS-PAGE to ensure the removal of the His-tag and TEV protease cleavage site (**Figure 5.14**). ESI mass spectrometry verified the cleavage and transfer of the His-tag and TEV cleavage site from **mCh_SC** to the TEV protease enzyme.



M. Perfect protein marker (10-225 kDa)

2. mCh_SC, Pure, TEV cleaved

1. mCh_SC, HisPur™ eluent

M. Perfect protein marker (10-225 kDa)

Figure 5.14 SDS-PAGE analysis comparing the HisPur™ mCh_SC protein samples before and after TEV protease cleavage to remove the His-tag. The HisPur™ mCh_SC protein is displayed in lane 1. The slightly smaller TEV cleaved mCh_SC is located in lane 2. MW: mCh_SC = 44.3 kDa; cleaved mCh_SC = 42.4 kDa

The ligation partner, **ST_HEAT**, was based off of the sequences for **HEAT_R1** and **HEAT_6R**. These peptides exhibited the most successful assemblies under acidic conditions, so assuming that **ST_HEAT** would behave similarly, **mCh_SC** would have to exhibit fluorescence at a low pH range. A small portion of the cleaved **mCh_SC** underwent stepwise dialysis from pH 8.0 to pH 3.5 dropping incrementally by a pH of 0.5. The **mCh_SC** protein solution exhibited strong fluorescence from pH 4.5-8. The fluorescence of **mCh_SC** decreased slightly at pH 4.0 and then dropped rapidly as the pH was decreased to 3.5. This sudden decrease in fluorescence was accompanied by the precipitation of the protein as it crashed out of solution. Therefore, the fluorescence and solubility of **mCh_SC** is strong at and above pH 4.5 which indicates that **mCh_SC** is suitable for ligation to the **ST_HEAT** protein under these conditions. **Figure 5.15** displays the TEV cleaved HisPurTM **mCh_SC** protein solution at pH 4.5 along with its fluorescence under UV light.

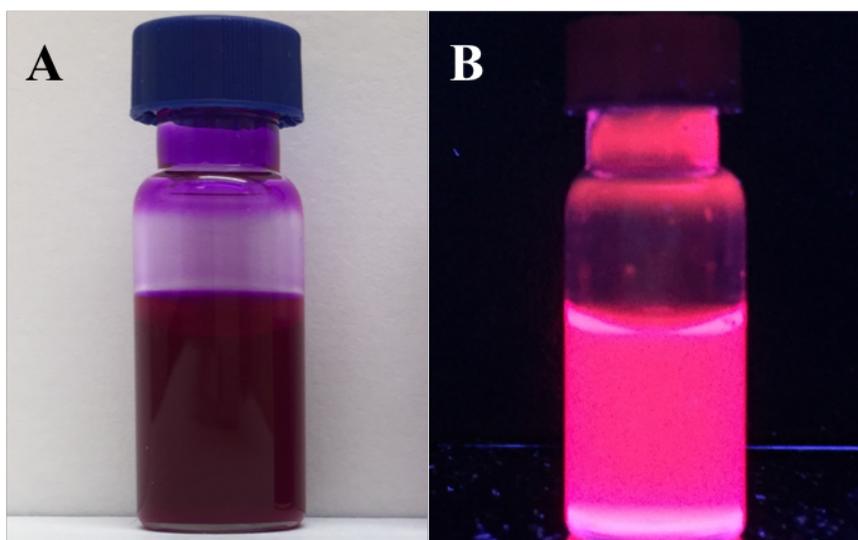


Figure 5.15 (A) TEV cleaved HisPurTM **mCh_SC** protein solution at pH 4.5. (B) TEV cleaved HisPurTM **mCh_SC** protein solution at pH 4.5 under UV light.

After the bulk solution of TEV cleaved pure **mCh_SC** was concentrated and dialyzed, the concentration was determined using a Bradford protein assay (BPA) kit. The concentrated **mCh_SC** sample was likely too concentrated even with the recommended dilutions, so a 2% **mCh_SC** solution was also prepared. Serial dilutions were made of both the stock 100% **mCh_SC** solution and the 2% **mCh_SC** solution. A BSA standard curve was prepared by plotting the blank-corrected absorbances of the BSA standards vs. the concentrations of the BSA standards, and the best linear trendline and equation were produced. The blank-corrected absorbances of the **mCh_SC** dilutions within the concentration range of the linear portion of the BSA standard curve were used to determine the concentration of the original concentrated **mCh_SC** solution.

The most dilute sample from the 100% **mCh_SC** stock did not fall within the range of the linear trendline as expected. All dilutions prepared from the 2% **mCh_SC** stock fell within the range of the linear trendline except for the most dilute sample, which did not exhibit any significant absorbance (**Figure 5.16**). The two **mCh_SC** samples that fell in the middle of the trendline were used to calculate the concentration of the 2% **mCh_SC** sample, which was in turn used to determine the concentration of the original concentrated stock of **mCh_SC**. The calculated concentration of the 2% **mCh_SC** solution was approximately 18-20 μM . The original concentrated stock of the **mCh_SC** solution, which was used for all future experiments, was approximately 1 mM. The protein yield was 254 mg (or 6 μmoles) from 1 L of expression culture.

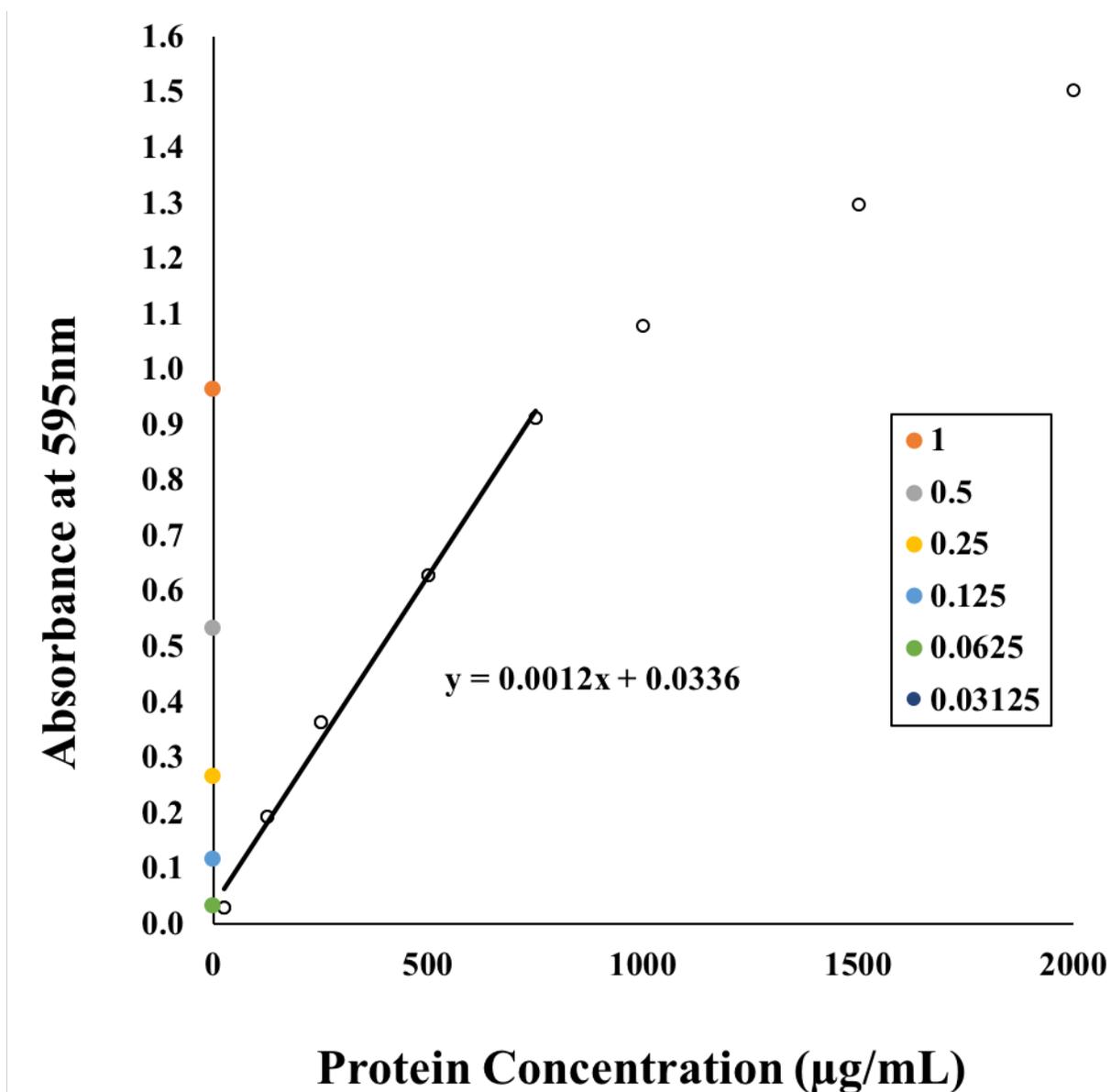


Figure 5.16 BSA standard curve produced from the absorbances of the BSA standards vs. the concentrations of the BSA standards. A trendline was generated in the linear section of the curve and an equation was generated. The absorbances of the mCh_SC dilution samples are plotted on the y-axis. They are labelled in the figure legend by their dilution factor from the 2% mCh_SC stock.

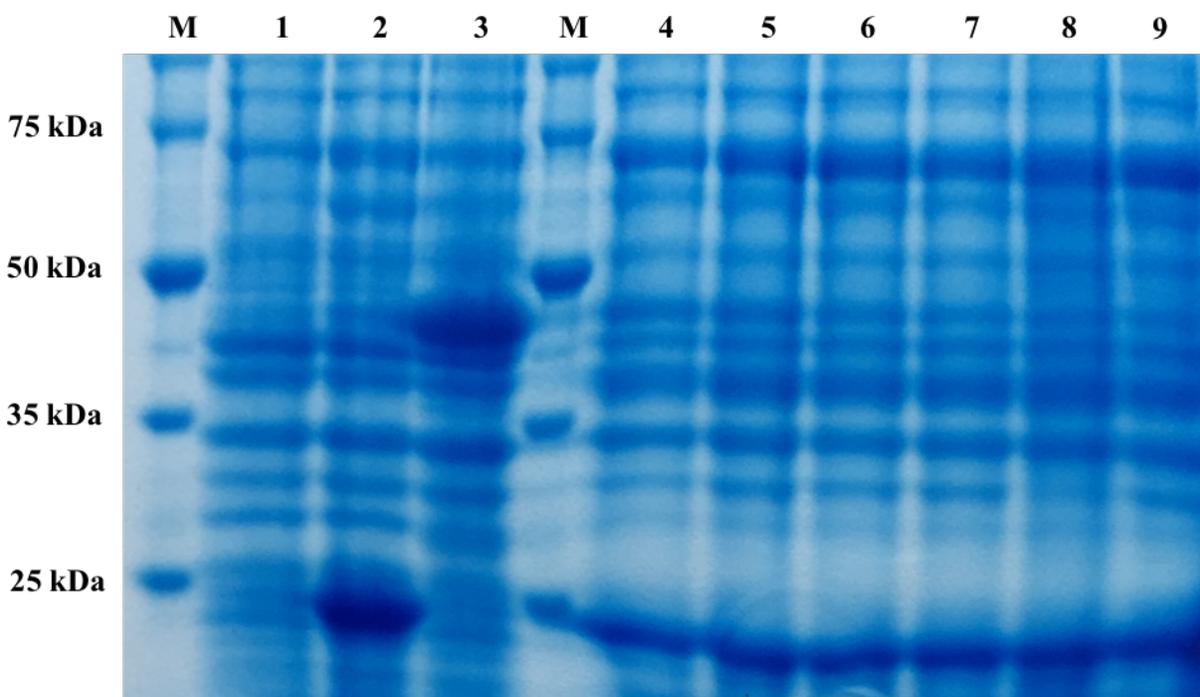
5.2.3 Initial Verification of Linking Between ST_HEAT and mCh_SC

During the production of the ST_HEAT and mCh_SC protein, samples were combined at various stages to analyze the ligation ability of SpyCatcher and SpyTag while fused to the 8 repeat HEAT protein and the mCherry fluorescent protein. The covalent link between two components is resistant to boiling in sodium dodecyl sulfate (SDS) once it is formed, but the presence of SDS inhibits any further reaction between the two ligation components.⁴ This allows for accurate analysis of the ligation process over time through the use of SDS-PAGE as long as time points are immediately combined with the SDS loading dye.

Linking of ST_HEAT Lysate with mCh_SC Lysate

Aliquots were collected after cell lysis to assess the ligation capabilities of the SpyCatcher and SpyTag segments of the mCh_SC and ST_HEAT proteins. The aliquots were collected before the soluble and insoluble fractions were separated. Each protein was in a slurry composed of the lysis buffer (50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0), the components added for chemical lysis (EDTA-free protease inhibitor cocktail, 1 mg/mL lysozyme, 25 units/mL benzonase nuclease, and 1 mM MgCl₂), the broken-down cellular components, and a large variety of other cellular proteins. Suffice it to say, the proteins were in a very complex solution.

The complex lysate solutions containing the mCh_SC and ST_HEAT proteins were combined in a 1:1 (v/v) ratio, and aliquots were removed starting 1 minute after integration for analysis. Time points were collected 1 minute, 5 minutes, 10 minutes, 15 minutes, 30 minutes, 45 minutes, and 60 minutes after combination. The aliquots were collected and immediately prepared for SDS-PAGE analysis to halt the ligation process. The presence of mCh-HEAT (the ligated product of mCh_SC with ST_HEAT) was visible in the solution as early as 1 minute after mixing



M. Perfect protein marker (10-225 kDa)

1. ST_HEAT, expression, before induction

2. ST_HEAT, expression, 15 hrs after induction

3. mCh_SC, expression, 4 hrs after induction

M. Perfect protein marker (10-225 kDa)

4. Lysate linking, 1 min after mixing

5. Lysate linking, 10 min after mixing

6. Lysate linking, 30 min after mixing

7. Lysate linking, 1 hrs after mixing

8. Lysate linking, 5 hrs after mixing

9. Lysate linking, 24 hrs after mixing

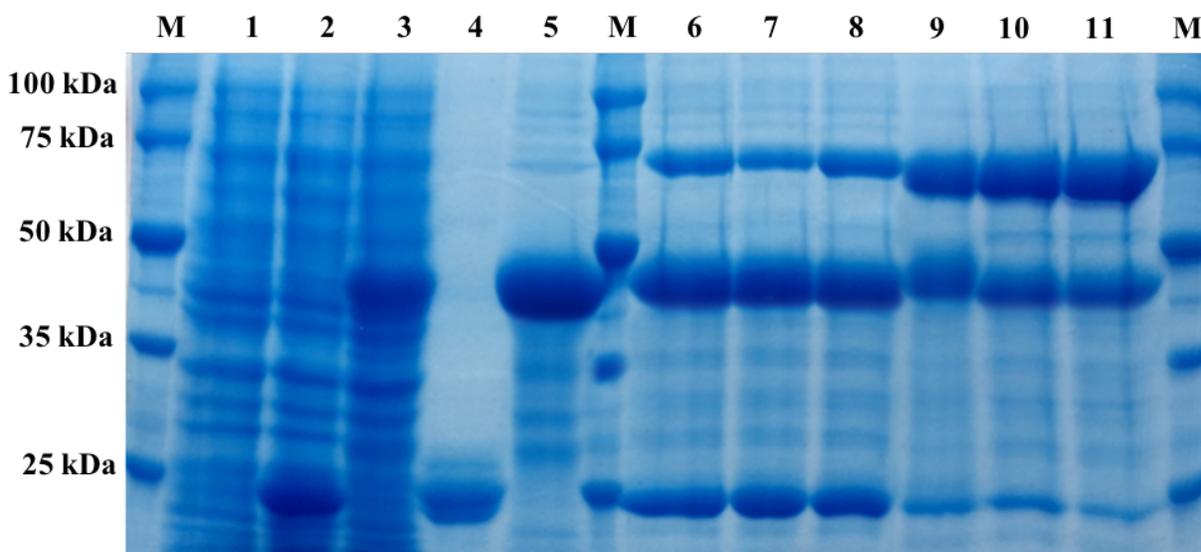
Figure 5.17 SDS-PAGE analysis depicting the SpyTag/SpyCatcher ligation results from combining the cell lysate solutions containing ST_HEAT and mCh_SC proteins. Aliquots were collected at various time points after cell lysate solutions were combined. The ligation product, mCh-HEAT, was visualized as early as 1 min after combining. Expression samples for ST_HEAT and mCh_SC are displayed for reference in lanes 2 and 3, respectively. MW: ST_HEAT = 30.1 kDa; pre-cleaved mCh_SC = 44.3 kDa; mCh-HEAT = 74.4 kDa.

the solutions (**Figure 5.17**). This rapid process occurred until all **mCh_SC** protein (the limiting reagent of the mixture) was completely removed from solution.

Linking of ST_HEAT HisPurTM Eluent with mCh_SC HisPurTM Eluent

The ligation capabilities of the **mCh_SC** and **ST_HEAT** fusion proteins were further assessed using aliquots collected after purification using metal affinity chromatography on HisPurTM cobalt columns. The aliquots were collected as the proteins were eluting from the HisPurTM resin. The **mCh_SC** protein was in a solution containing the non-denaturing elution buffer (50 mM NaH₂PO₄, 100 mM NaCl, 250 mM imidazole, pH 8.0), while the **ST_HEAT** protein was in a solution containing the 6 M urea denaturing elution buffer (6 M urea, 50 mM NaH₂PO₄, 100 mM NaCl, 250 mM imidazole, pH 8.0). Although these eluent solutions were less complex than the lysate solutions, urea was present in the **ST_HEAT** solution.

The HisPurTM eluent solutions containing the **mCh_SC** and **ST_HEAT** fusion proteins were combined in a 1:1 (v/v) ratio, and aliquots were removed starting 2 minutes after integration for analysis. Time points were collected 2 minutes, 10 minutes, 30 minutes, 5 hours, 9 hours, and 24 hours after combination. The aliquots were collected and immediately prepared for SDS-PAGE analysis to terminate the ligation process. The ligated product, **mCh-HEAT**, was visible in the solution as early as 2 minutes after mixing the solutions (**Figure 5.18**). This time, the process occurred until all **ST_HEAT** protein, which was the limiting reagent of the mixture, was utilized.



M. Perfect protein marker (10-225 kDa)

1. ST_HEAT, expression, before induction

2. ST_HEAT, expression, 15 hrs after induction

3. mCh_SC, expression, 4 hrs after induction

4. ST_HEAT, HisPur™ eluent

5. mCh_SC, HisPur™ eluent

M. Perfect protein marker (10-225 kDa)

6. Eluent linking, 2 min after mixing

7. Eluent linking, 10 min after mixing

8. Eluent linking, 30 min after mixing

9. Eluent linking, 5 hrs after mixing

10. Eluent linking, 9 hrs after mixing

11. Eluent linking, 24 hrs after mixing

M. Perfect protein marker (10-225 kDa)

Figure 5.18 SDS-PAGE analysis depicting the SpyTag/SpyCatcher ligation results from combining the HisPur™ eluents containing ST_HEAT and mCh_SC proteins. Aliquots were collected at various time points after solutions were combined. The ligation product, mCh-HEAT, was visualized as early as 2 min after combining. Expression and HisPur™ eluent samples for ST_HEAT and mCh_SC are displayed for reference in lanes 2-5. MW: ST_HEAT = 30.1 kDa; pre-cleaved mCh_SC = 44.3 kDa; mCh-HEAT = 74.4 kDa.

5.3 Conclusions

A system was designed for the production of nanotubes displaying asymmetric surfaces on the exterior convex surface and the interior concave pore. The HEAT-based concatemer composed of with 8 HEAT-based helical hairpins was designed and fused to the *C*-terminus of the SpyTag peptide to form **ST_HEAT**. The mCherry fluorescent protein was fused to the *N*-terminus of the SpyCatcher protein to form **mCh_SC**. The two proteins were successfully expressed using *E. coli* as a host organism and purified via immobilized metal affinity chromatography.

Solutions containing the **ST_HEAT** and **mCh_SC** proteins were combined at various stages throughout the production process to verify that two proteins could covalently link together to form **mCh-HEAT**. The SpyTag/SpyCatcher ligation system allowed for the rapid linking of the two proteins in the highly heterogeneous cell lysate, which verified the specificity of the linkage. The purified protein eluents were also combined to verify linkage, and remarkably, the link rapidly formed even though the solution contained 3 M urea. Optimization of the assembly conditions for **ST_HEAT** nanotubes followed by ligation experiments could lead to highly defined nanotubes with asymmetric functionalization between the interior and exterior surfaces.

5.4 Experimental

5.4.1 Materials

All chemical reagents were purchased from either Thermo Fisher Scientific, Inc. (Waltham, MA) or MilliporeSigma (Burlington, MA) unless otherwise noted. Synthesis of plasmids containing codon-optimized genes encoding mCherry_SpyCatcher and SpyTag_HEAT8R was conducted by ATUM, formerly known as DNA2.0 (Newark, CA). The Top10F' chemically competent *E. coli* strain was obtained from Invitrogen Corp. (Carlsbad, CA), and the BL21 (DE3) chemically competent *E. coli* strain was purchased from New England Biolabs, Inc. (Ipswich, MA). The QIAprep-spin miniprep kit was purchased from QIAGEN, Inc. (Valencia, CA). Luria-Bertani broth and agar powder were purchased from MilliporeSigma (Burlington, MA), and Terrific Broth (TB) powder was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). Kanamycin monosulfate was purchased from VWR International, LLC (Suwanee, GA). Isopropyl- β -D-thiogalactopyranoside (IPTG) was purchased from Research Products International Corp. (Prospect, IL). The Perfect Protein Marker (10-225 kDa) was purchased from MilliporeSigma (Burlington, MA).

Benzonase[®] nuclease and protease-inhibitor cocktail (EDTA-free) were purchased from MilliporeSigma (Burlington, MA). Lysozyme from chicken egg white was purchased from Research Products International Corp. (Prospect, IL). HisPur[™] cobalt resin was obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). TEV Protease was purchased from MilliporeSigma (Burlington, MA). SnakeSkin[™] Dialysis Tubing with a 10 kDa MWCO and Slide-A-Lyzer[™] G2 dialysis cassettes with a 20 kDa MWCO were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). Amicon[®] ultra centrifugal filters (10 kDa NMWL, 30 kDa NMWL, 50 kDa

NMWL, and 100 kDa NMWL) for sample volumes of 15 mL and 0.5 mL were purchased from MilliporeSigma (Burlington, MA). The Coomassie (Bradford) Protein Assay Kit was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). Vacuum filters with a 0.2 μ m polyethersulfone (PES) membrane were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

5.4.2 General Methods

Basic molecular biology procedures were adapted from a standard molecular cloning manual²³ or the protocol supplied by the manufacturer unless otherwise described in detail. All Reagents intended for use with bacteria, DNA, or recombinant proteins were sterilized by either syringe filtration through a 0.2 μ m cellulose membrane, vacuum filtration through a 0.2 μ m polyethersulfone (PES) membrane, or by autoclaving. All enzyme reactions were conducted in the reagent buffers provided by the manufacturer unless otherwise noted.

Synthetic plasmids were ordered from ATUM, resuspended in water upon arrival, and transformed into chemically competent Top10F' or BL21 *E. coli* strains. *E. coli* strains were grown at 37 °C in Lauria-Bertani (LB) medium containing appropriate antibiotic with shaking at 200 rpm unless otherwise stated. A QIAprep-spin miniprep kit was used for isolation of plasmid DNA. All OD values were determined using a UltroSpec 3000 UV/Visible spectrophotometer at 600 nm with a 1 cm cuvette. All proteins were purified using immobilized metal affinity chromatography (IMAC) on HisPur™ cobalt resin followed by dialysis to remove imidazole.

Protein electrophoresis was conducted using 12 or 16% SDS polyacrylamide gels on a Mini-PROTEAN 3 cell electrophoresis system from Bio-Rad Laboratories, Inc. (Hercules, CA).

The buffer tank was filled with SDS run buffer (25 mM tris, 250 mM glycine, 0.1% SDS, pH 8.3). The Perfect Protein Marker was used as a protein standard for SDS-PAGE analysis. An initial sample volume of 10 μ L was added to each well with any necessary concentration adjustments made thereafter. Gels were run at 150 V for 1-1.5 hours depending on the desired level of separation. Gels were stained in coomassie overnight and then destained in a methanol-acetic acid buffer.

5.4.3 SpyTag_HEAT Production

Gene and Bacterial Preparation

A gene encoding a fusion protein of SpyTag and HEAT_8R was codon optimized for expression in *E. coli* and synthesized by ATUM (formerly DNA 2.0). The lyophilized powder was resuspended in 50 μ L of distilled, deionized water (ddH₂O) upon arrival. Aliquots of 1 μ L were used in transformations of the chemically competent *E. coli* strains BL21 (DE3) and Top10F'. The cells were recovered after heat shock in 400 μ L SOC rich media for 30 minutes at 37 °C. Aliquots of 25 μ L and 50 μ L of the transformed suspensions were spread onto LB agar plates containing kanamycin (50 μ g/mL) for antibiotic selection. The plates were incubated for 14 hours at 37 °C.

Six colonies were selected from the plate containing the Top 10 F' strain and two colonies were selected from the plate containing the BL21 strain. The selected colonies were used to inoculate eight culture tubes containing 5 mL LB media supplemented with kanamycin (50 μ g/mL). The cultures were grown at 37 °C overnight on a rotator. Two of the cultures containing the Top10F' strain in addition to the two cultures containing the BL21 strain were prepared for long term storage of the plasmid contained within the *E. coli* strain. These frozen stocks were

prepared by combining 200 μ L of 80% glycerol with 800 μ L of the culture. The frozen stocks of the BL21 and Top10F' strains containing the plasmid **ST_HEAT** were stored at -80 °C until needed. To make secondary stocks of the **ST_HEAT** plasmid DNA, a QIAprep-spin miniprep kit (QIAGEN, Inc.) was used to isolate the plasmid DNA of each of the remaining cultures of Top10F' cells. The isolated DNA was recovered in 50 μ L of EB buffer (10 mM Tris-HCl, pH 8.5) and stored in solution at -30 °C when not in use.

Large-Scale Expression

The plasmid **ST_HEAT** was transformed into BL21 (DE3) strain and plated onto LB agar plates containing kanamycin (50 μ g/mL). Single colonies of BL21 strain containing the **ST_HEAT** plasmid were used to inoculate 5 mL cultures of LB broth supplemented with kanamycin (50 μ g/mL). The cell cultures were grown overnight on a rotator at 37 °C. Five mL of the overnight cultures were transferred to 2,800 mL Erlenmeyer flasks containing 500 mL of Lauria-Bertani (LB) media supplemented with 50 μ g/mL kanamycin. Four flasks were used for a total expression culture volume of 2000 mL (2L).

The large culture flasks shook at 200 rpm and 37 °C for approximately 3 hours until cell growth reached log phase growth with an OD₆₀₀ of ~0.6-0.8. The OD₆₀₀ was monitored hourly. Once log phase growth was achieved, expression was induced by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG). For each 500 mL flask, 250 μ L of 1 M IPTG was added for a final IPTG concentration of 0.5 mM. The expression cultures were incubated at 30 °C with shaking for 15 hours. Aliquots were removed from the expression culture throughout the process for SDS-PAGE analysis (see below). Fifteen hours after induction, the cells in the expression cultures were

harvested by centrifugation at 4,000 x g for 20 minutes (4 °C). The cell pellets were resuspended in 100 mL lysis buffer (50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0) and stored at -80 °C.

One mL aliquots were removed from each flask before induction and 1.5, 3 and 15 hours after induction. The OD₆₀₀ was determined for each aliquot and the aliquot volumes were normalized based on their OD so that the number of cells in each aliquot would be equivalent. The normalized aliquot volumes were centrifuged for 5 minutes at 4,000 x g to pellet the cells. The supernatants were discarded, the pellets were resuspended in 50 µL ddH₂O, and the expression samples were stored at -30 °C. Aliquoted expression samples were later prepared for SDS-PAGE analysis by mixing the 50 µL samples with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. The expression samples were run on a 16% SDS-PAGE gel according to the protocol described in the general methods section.

Cell Lysis

The cell pellets from a large-scale expression of **ST_HEAT** in the BL21 strain were previously frozen in 100 mL lysis buffer. Lysis of the cell pellets was initiated by three freeze/thaw cycles (-80 °C; 25 °C). After the third thaw cycle, EDTA-free protease inhibitor cocktail (1X), lysozyme (1 mg/mL), benzonase nuclease (25 units/mL), and MgCl₂ (1 mM) were added to the lysate. The lysate was incubated at 30 °C for 1 hour with shaking at 200 rpm and then the temperature was dropped to 4 °C for incubation overnight. The cell lysate was sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. A 50 µL aliquot was collected for SDS-PAGE analysis (see below). An aliquot of 3 mL was removed after lysis for experiments to verify linking between **ST_HEAT** and **mCh_SC** (see section below on linking of **ST_HEAT** lysate with **mCh_SC**

lysate). The remaining cell lysate was centrifuged at 10,000 rpm for 20 minutes (4 °C). The lysed supernatant was separated from the lysed pellet until analysis by SDS-PAGE determined the location of the **ST_HEAT** fusion protein.

Meanwhile, the cell lysate aliquot was centrifuged at 10,000 rpm for 5 minutes. The aliquot supernatant was transferred into a clean microcentrifuge tube and the pellet was resuspended in 50 μ L of ddH₂O. The lysed supernatant and lysed pellet samples were prepared for SDS-PAGE analysis by mixing the 50 μ L samples with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. Analysis by 16% SDS-PAGE concluded that the **ST_HEAT** fusion protein was primarily located in the lysed pellet (insoluble fraction).

Protein Solubilization

The insoluble **ST_HEAT** protein in the cell lysate pellet was solubilized to allow for purification on a HisPur™ cobalt column. In order to do this, the pellets were resuspended in 120 mL total volume of urea solubilization buffer (6 M urea, 50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0). The insoluble pellets were allowed to shake at 250 rpm for 3 days at 4 °C. A 50 μ L aliquot was collected for SDS-PAGE analysis (see below). The remaining urea solution was centrifuged at 12,000 rpm for 20 minutes (4 °C). The urea supernatant was separated from the urea pellet until analysis by SDS-PAGE determined the location of the **ST_HEAT** fusion protein.

Meanwhile, the urea aliquot was centrifuged at 12,000 rpm for 5 minutes. The aliquot supernatant was transferred into a clean microcentrifuge tube and the pellet was resuspended in 50 μ L of ddH₂O. The urea supernatant and urea pellet samples were prepared for SDS-PAGE analysis

by mixing the 50 μ L samples with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 $^{\circ}$ C for 5 minutes. Analysis by 16% SDS-PAGE concluded that the **ST_HEAT** fusion protein was at least partially soluble in 6M urea and primarily located in the supernatant.

Purification

The presence of a hexahistadine tag at the C-terminus of the fusion protein **ST_HEAT** facilitated purification by immobilized metal affinity chromatography (IMAC). The soluble fraction recovered after protein solubilization in urea was loaded directly onto 12 mL pre-equilibrated HisPurTM cobalt resin divided among four columns. The resin was then washed with 60 mL urea wash buffer (6 M urea, 50 mM NaH₂PO₄, 100 mM NaCl, 20 mM imidazole, pH 8.0). The target protein was isolated by the addition of 60 mL elution buffer (6 M urea, 50 mM NaH₂PO₄, 100 mM NaCl, 250 mM imidazole, pH 8.0). An aliquot of 100 μ L of HisPurTM eluent was removed after purification for experiments to verify linking between **ST_HEAT** and **mCh_SC** (see section below on linking of **ST_HEAT** HisPurTM eluent with **mCh_SC** HisPurTM eluent). A 50 μ L aliquot of the eluent was collected and prepared for SDS-PAGE analysis by mixing the 50 μ L sample with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 $^{\circ}$ C for 5 minutes. The pure eluent sample was run on a 16% SDS-PAGE gel according to the protocol described in the general methods section.

The purification eluent was loaded into a Slide-A-LyzerTM dialysis cassette (20,000 MWCO) and dialyzed against ddH₂O for 4 days (4 L volume, switching buffer every 8 hours).

After dialysis, the protein solution was transferred from the dialysis cassette to a 50 mL conical tubes and stored at -80 °C. Once frozen, the tops of the conical tubes were replaced with Kimwipes™ and the tubes were placed in a lyophilizer until all that remained was the pure ST_HEAT protein in powder form. The lyophilized powder was weighed and stored at -30 °C.

5.4.4 mCherry_SpyCatcher Production

Gene and Bacterial Preparation

A gene encoding a fusion protein of mCherry red fluorescent protein and SpyCatcher was codon optimized for expression in *E. coli* and synthesized by ATUM (formerly DNA 2.0). The lyophilized powder was resuspended in 50 µL of distilled, deionized water (ddH₂O) upon arrival. Aliquots of 1 µL were used in transformations of the chemically competent *E. coli* strains BL21 (DE3) and Top10F'. The cells were recovered after heat shock in 400 µL SOC rich media for 30 minutes at 37 °C. Aliquots of 50 µL of the transformed suspensions were spread onto LB agar plates containing kanamycin (50 µg/mL) for antibiotic selection. The plates were incubated for 14 hours at 37 °C.

Six colonies were selected from the plate containing the Top 10 F' strain and two colonies were selected from the plate containing the BL21 strain. The selected colonies were used to inoculate eight culture tubes containing 5 mL LB media supplemented with kanamycin (50 µg/mL). The cultures were grown at 37 °C overnight on a rotator. Two of the cultures containing the Top10F' strain in addition to the two cultures containing the BL21 strain were prepared for long term storage of the plasmid contained within the *E. coli* strain. These frozen stocks were prepared by combining 200 µL of 80% glycerol with 800 µL of the culture. The frozen stocks of the BL21 and Top10F' strains containing the plasmid mCh_SC were stored at -80 °C until needed.

To make secondary stocks of the **mCh_SC** plasmid DNA, a QIAprep-spin miniprep kit (QIAGEN, Inc.) was used to isolate the plasmid DNA of each of the remaining cultures of Top10F' cells. The isolated DNA was recovered in 50 μ L of EB buffer (10 mM Tris-HCl, pH 8.5) and stored in solution at -30 °C when not in use.

Large-Scale Expression

mCh_SC plasmid was transformed into BL21 (DE3) strain and plated onto LB agar plates containing kanamycin (50 μ g/mL). Single colonies of BL21 strain containing the **mCh_SC** plasmid were used to inoculate 5 mL of LB broth supplemented with kanamycin (50 μ g/mL). The cell cultures were grown overnight on a rotator at 37 °C. Five mL of the overnight cultures were transferred to 2,800 mL Erlenmeyer flasks containing 500 mL of Terrific Broth media supplemented with 50 μ g/mL kanamycin. Four flasks were used for a total expression culture volume of 2000 mL (2L).

The large culture flasks shook at 200 rpm and 37 °C for 2-3 hours until cell growth reached log phase growth with an OD₆₀₀ of ~0.6-0.8. The OD₆₀₀ was monitored hourly. Once log phase growth was achieved, expression was induced by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG). For each 500 mL flask, 250 μ L of 1 M IPTG was added for a final IPTG concentration of 0.5 mM. The expression cultures were incubated at 30 °C with shaking for 4 hours. Aliquots were removed from the expression culture throughout the process for SDS-PAGE analysis (see below). Four hours after induction, the cells in the expression cultures were harvested by centrifugation at 4,000 x g for 20 minutes (4 °C). The supernatants were transferred into a clean 2,800 mL Erlenmeyer flask since there appeared to be a large amount of **mCh_SC** present in the media. A 50 μ L aliquot of the expression media was collected for SDS-PAGE

analysis. The cell pellets were resuspended in 100 mL lysis buffer (50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0) and stored at -80 °C.

One mL aliquots were removed from each flask before induction and each hour after induction. The OD₆₀₀ was determined for each aliquot and the aliquot volumes were normalized based on their OD so that the number of cells in each aliquot would be equivalent. The normalized aliquot volumes were centrifuged for 5 minutes at 4,000 x g to pellet the cells. The supernatants were discarded, the pellets were resuspended in 50 µL ddH₂O, and the expression samples were stored at -30 °C. Aliquoted expression samples were later prepared for SDS-PAGE analysis by mixing the 50 µL samples with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the samples at 100 °C for 5 minutes.

Cell Lysis

The cell pellets from a large-scale expression of **mCh_SC** in the BL21 strain were previously frozen in 100 mL lysis buffer. Lysis of the cell pellets was initiated by three freeze/thaw cycles (-80 °C; 25 °C). After the third thaw cycle, EDTA-free protease inhibitor cocktail (1X), lysozyme (1 mg/mL), benzonase nuclease (25 units/mL), and MgCl₂ (1 mM) were added to the lysate. The lysate was incubated at 30 °C for 1 hour with shaking at 200 rpm and then the temperature was dropped to 4 °C for incubation overnight. The cell lysate was sonicated (9 seconds on; 9 seconds off) at 4 °C for 15 minutes. A 50 µL aliquot was collected for SDS-PAGE analysis (see below). A 3 mL aliquot was removed after lysis for experiments to verify linking between **mCh_SC** and **ST_HEAT** (see section below on linking of **ST_HEAT** lysate with **mCh_SC** lysate). The remaining cell lysate was centrifuged at 15,000 rpm for 20 minutes (4 °C). The lysed

supernatant was separated from the lysed pellet until analysis by SDS-PAGE determined the location of the **mCh_SC** fusion protein.

Meanwhile, the cell lysate aliquot was centrifuged at 15,000 rpm for 5 minutes. The aliquot supernatant was transferred into a clean microcentrifuge tube and the pellet was resuspended in 50 μ L of ddH₂O. The lysed supernatant and lysed pellet samples were prepared for SDS-PAGE analysis by mixing the 50 μ L samples with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the samples at 100 °C for 5 minutes. Analysis by 12% SDS-PAGE concluded that the **mCh_SC** fusion protein was primarily located in the lysed supernatant (soluble fraction) with some remaining in the lysed pellet (insoluble fraction).

Purification

The presence of a hexahistadine tag at the N-terminus of the fusion protein **mCh_SC** facilitated purification by immobilized metal affinity chromatography (IMAC). The soluble fraction recovered after lysis was loaded directly onto 12 mL pre-equilibrated HisPur™ cobalt resin distributed across four columns. The resin was then washed with 60 mL wash buffer (50 mM NaH₂PO₄, 100 mM NaCl, 20 mM imidazole, pH 8.0). The target protein was isolated by the addition of 40 mL elution buffer (50 mM NaH₂PO₄, 100 mM NaCl, 250 mM imidazole, pH 8.0). An aliquot of 100 μ L of HisPur™ eluent was removed after purification for experiments to verify linking between **mCh_SC** and **ST_HEAT** (see section below on linking of **ST_HEAT** HisPur™ eluent with **mCh_SC** HisPur™ eluent). The remaining eluent was loaded into a Slide-A-Lyzer™ dialysis cassette (20,000 MWCO) and dialyzed against 50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0 for 12 hours (4 L volume, switching buffer every 4 hours). After dialysis, the protein solution was

transferred from the dialysis cassette to a 50 mL conical tube and stored at 4 °C. A 50 µL aliquot of the eluent was collected and prepared for SDS-PAGE analysis by mixing the 50 µL sample with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the sample at 100 °C for 5 minutes.

As noted previously, the expression media present after harvesting the cells contained a significant amount of **mCh_SC**. The expression media was centrifuged at 15,000 rpm for 20 minutes to ensure the removal of all cells. The supernatant was saved for purification. The expression media shook with 20 mL of HisPur™ cobalt resin for 1 hour and the resin was added to four columns. The loaded resin was washed with 100 mL wash buffer (50 mM NaH₂PO₄, 100 mM NaCl, 20 mM imidazole, pH 8.0). The target protein was isolated by the addition of 40 mL elution buffer (50 mM NaH₂PO₄, 100 mM NaCl, 250 mM imidazole, pH 8.0). The eluent was loaded into a Slide-A-Lyzer™ dialysis cassette (20,000 MWCO) and dialyzed against 50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0 for 12 hours (4 L volume, switching buffer every 4 hours). After dialysis, the protein solution was transferred from the dialysis cassette to a 50 mL conical tube and stored at 4 °C.

TEV Protease Cleavage

A TEV protease cleavage site was included to facilitate the removal of the His-tag from the **mCh_SC** fusion protein after purification. TEV protease was added to half of the purified **mCh_SC** protein to a final concentration of 1:100 (10,000 units TEV protease for every 100 mg of target protein). The solution was transferred to a Slide-A-Lyzer™ dialysis cassette (20,000 MWCO) and dialyzed against 50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0 at 4 °C for 3 days (4 L

volume). A stir bar was used to agitate the dialysis buffer, and the dialysis buffer was switched out every 8 hours.

To remove the cleaved His-tag-TEV protease unit, the dialyzed protein was removed from the dialysis cassette and loaded onto a column containing pre-equilibrated HisPur™ cobalt resin. Flow through was collected to obtain the TEV-cleaved pure **mCh_SC**. The cleaved His-tag and the TEV protease enzyme remained attached to the cobalt resin. Dialysis buffer (50 mM NaH₂PO₄, 100 mM NaCl, pH 8.0) was added until all **mCh_SC** left the column. A total volume of 50 mL of eluent was collected. A 50 µL aliquot of the eluent was collected and prepared for SDS-PAGE analysis by mixing the 50 µL sample with 50 µL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β-mercaptoethanol) and boiling the sample at 100 °C for 5 minutes.

A small portion of the cleaved **mCh_SC** was dialyzed using a Slide-A-Lyzer™ dialysis cassette (20,000 MWCO) using step-wise dialysis from pH 8.0 to pH 3.5. The pH was dropped incrementally by 0.5 while the solubility and fluorescence were monitored. The remaining cleaved pure **mCh_SC** protein was filtered through a 0.2 µm filter and concentrated using Amicon® Ultra-15 centrifugal filters with a 30 kDa NMWL. The volume was decreased from 50 mL down to approximately 6 mL of concentrated eluent. The concentrated **mCh_SC** was dialyzed using a Slide-A-Lyzer™ dialysis cassette (20,000 MWCO) using step-wise dialysis to 10 mM acetate pH 5.0. After dialysis, the protein solution was removed from the dialysis cassette and stored at 4 °C.

Concentration Calculation

A Coomassie (Bradford) Protein Assay Kit from Thermo Fischer Scientific was used to determine the concentration of the **mCh_SC** solution. Diluted albumin (BSA) standards were

prepared with a final BSA concentration range of 0-2000 $\mu\text{g}/\text{mL}$ using 10 mM Acetate pH 4.5 as a diluent. Dilutions were prepared using a sample that was a 100% solution of the concentrated **mCh_SC** and a sample that was a 2% solution of the concentrated **mCh_SC**. The 100% **mCh_SC** and 2% **mCh_SC** samples were each diluted using serial dilution to a dilution factor of 1, 2, 4, 8, 16, and 32. A 30 μL volume of each standard and unknown sample was combined with 1.5 mL coomassie (G-250) reagent. The samples sat at room temperature for 10 minutes to equilibrate. The UV/Visible spectrophotometer was zeroed with ddH₂O in a 1 cm cuvette at 595 nm. The absorbances of all standards and unknown samples were then measured at 595 nm in triplicate.

The BSA standards and unknown samples were blank-corrected by subtracting the average of the absorbances of the 0 $\mu\text{g}/\text{mL}$ BSA standard (containing only ddH₂O and coomassie) from the average of the absorbances of all standards and unknown samples. A BSA standard curve was prepared by graphing the blank-corrected BSA standard absorbances vs. the concentration of the BSA standards in $\mu\text{g}/\text{mL}$. A trendline and equation were produced using the points that allowed for the best linear trendline. The blank-corrected absorbances of the **mCh_SC** dilutions were plotted against the BSA standard curve.

The **mCh_SC** dilutions within the concentration range of the linear portion of the BSA standard curve were used to determine the concentration of the original concentrated **mCh_SC** solution. The linear trendline equation was used to calculate the concentrations of the unknown dilutions in $\mu\text{g}/\text{mL}$. The concentrations were multiplied by the dilution factors to determine the concentration of the original stock **mCh_SC**. The samples that originated from the 2% stock concentration were adjusted to reflect the concentration of the 100% stock solution. The concentrations were then converted to mg/mL by dividing by 1000 and then to molarity (M) by

dividing by the molecular weight of **mCh_SC** (42406.74 g/mol). The final concentration of the stock **mCh_SC** solution was recorded in μM after multiplying by 10^6 .

After the concentration was determined, 50 μL volumes of pure concentrated **mCh_SC** in 10 mM Acetate pH 5.0 was aliquoted in 1.5 mL microcentrifuge tubes. The protein was stored at $-80\text{ }^\circ\text{C}$ until needed.

5.4.5 Verification of Linking Between ST_HEAT and mCh_SC

Linking of ST_HEAT Lysate with mCh_SC Lysate

The 3 mL aliquot of **ST_HEAT** lysate was combined with the 3 mL aliquot of **mCh_SC** lysate. The solution was incubated at room temperature with shaking for 1 hour. Aliquots of 100 μL were removed throughout the 1 hour linking process. The time points collected were 1 minute, 5 minutes, 10 minutes, 15 minutes, 30 minutes, 45 minutes, and 60 minutes after combination of the two samples. The aliquots were collected and immediately prepared for SDS-PAGE analysis by mixing the 100 μL sample with 100 μL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at $100\text{ }^\circ\text{C}$ for 5 minutes.

Linking of ST_HEAT HisPurTM Eluent with mCh_SC HisPurTM Eluent

The 100 μL of **ST_HEAT** HisPurTM eluent was combined with the 100 μL of **mCh_SC** HisPurTM eluent. The solution was incubated at room temperature with shaking for 24 hours. Aliquots of 15 μL were removed throughout the 24-hour linking process. The time points collected were 2 minutes, 10 minutes, 30 minutes, 5 hours, 9 hours, and 24 hours after combination of the two samples. The aliquots were collected and immediately prepared for SDS-PAGE analysis by

mixing the 15 μ L sample with 15 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 $^{\circ}$ C for 5 minutes.

5.4.6 Tables

Table 5.1 Plasmids utilized in Chapter 5.

Plasmids	Relevant Characteristics	Reference
ST_HEAT	ST_HEAT gene in ATUM expression vector pD411-SR, Kan ^R	Chapter 5
mCh_SC	mCh_SC gene in ATUM expression vector pD411-NHT, Kan ^R	Chapter 5

Table 5.2 *E. coli* strains utilized in Chapter 5.

Strains	Genotype	Reference
Top10F'	F' [lacIq, Tn10(TetR)] mcrA Δ(mrr-hsdRMS- mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG	Invitrogen
BL21 (DE3)	<i>fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ sBamHIo ΔEcoRI- B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5</i>	New England Biolabs

Table 5.3 Protein sequences utilized in Chapter 5.

Protein	Sequence	MW (Da) # of AA
ST_HEAT	GAHIVMVDAYKPTKTSGGGSGGGAS[GDERAVEALIKALKDPDGWVRKAAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEI] ₄ GSMHHHHHH	30069.92 282
mCh_SC	MKHHHHHHGTSENLYFQGMSSGLVPRGSHMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTA KLVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTFNPSDGPVMQKKTMTGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKNSGGGLVAGGSGGGSGGGTGGGSGGGTSGAMVDTL SGLSSEQQQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQQQVTVNGKATKGAHI	44298.7 408
mCh_SC (After TEV Cleavage)	GMSSGLVPRGSHMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTA KLVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTFNPSDGPVMQKKTMTGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKNSGGGLVAGGSGGGSGGGTGGGSGGGTSGAMVDTL SGLSSEQQQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQQQVTVNGKATKGAHI	42406.74 392

5.5 References

1. Pluckthun, A., Designed Ankyrin Repeat Proteins (DARPin): Binding Proteins for Research, Diagnostics, and Therapy. In *Annual Review of Pharmacology and Toxicology, Vol 55*, Insel, P. A., Ed. 2015; Vol. 55, pp 489-511.
2. Urvoas, A.; Guellouz, A.; Valerio-Lepiniec, M.; Graille, M.; Durand, D.; Desravines, D. C.; van Tilbeurgh, H.; Desmadril, M.; Minard, P., Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (alphaRep) based on thermostable HEAT-like repeats. *J Mol Biol* **2010**, *404* (2), 307-27.
3. Zakeri, B.; Howarth, M., Spontaneous Intermolecular Amide Bond Formation between Side Chains for Irreversible Peptide Targeting. *J. Am. Chem. Soc.* **2010**, *132* (13), 4526-+.
4. Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M., Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A* **2012**, *109* (12), E690-7.
5. Kang, H. J.; Coulibaly, F.; Clow, F.; Proft, T.; Baker, E. N., Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science (New York, N.Y.)* **2007**, *318* (5856), 1625-1628.
6. Nguyen, P. Q.; Botyanszki, Z.; Tay, P. K.; Joshi, N. S., Programmable biofilm-based materials from engineered curli nanofibres. *Nature communications* **2014**, *5*, 4945.
7. Botyanszki, Z.; Tay, P. K.; Nguyen, P. Q.; Nussbaumer, M. G.; Joshi, N. S., Engineered catalytic biofilms: Site-specific enzyme immobilization onto E. coli curli nanofibers. *Biotechnology and bioengineering* **2015**, *112* (10), 2016-24.
8. Pedelacq, J. D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S., Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **2006**, *24* (1), 79-88.

9. Campbell, T. N.; Choy, F., *The effect of pH on green fluorescent protein: A brief review*. 2001; Vol. 2, p 1-4.
10. Roberts, T. M.; Rudolf, F.; Meyer, A.; Pellaux, R.; Whitehead, E.; Panke, S.; Held, M., Identification and Characterisation of a pH-stable GFP. *Scientific Reports* **2016**, *6*, 28166.
11. Baird, G. S.; Zacharias, D. A.; Tsien, R. Y., Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc Natl Acad Sci U S A* **2000**, *97* (22), 11984-9.
12. Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y., A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A* **2002**, *99* (12), 7877-82.
13. Shu, X.; Shaner, N. C.; Yarbrough, C. A.; Tsien, R. Y.; Remington, S. J., Novel chromophores and buried charges control color in mFruits. *Biochemistry* **2006**, *45* (32), 9639-47.
14. Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y., Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp red fluorescent protein. *Nat. Biotechnol.* **2004**, *22* (12), 1567-1572.
15. Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y., A guide to choosing fluorescent proteins. *Nature Methods* **2005**, *2* (12), 905-909.
16. Stepanenko, O. V.; Stepanenko, O. V.; Kuznetsova, I. M.; Verkhusha, V. V.; Turoverov, K. K., Beta-Barrel Scaffold of Fluorescent Proteins: Folding, Stability and Role in Chromophore Formation. *International review of cell and molecular biology* **2013**, *302*, 221-278.
17. Kajander, T.; Cortajarena, A. L.; Main, E. R.; Mochrie, S. G.; Regan, L., A new folding paradigm for repeat proteins. *J Am Chem Soc* **2005**, *127* (29), 10188-90.

18. Wetzel, S. K.; Settanni, G.; Kenig, M.; Binz, H. K.; Pluckthun, A., Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J Mol Biol* **2008**, *376* (1), 241-57.
19. Moon, H.; Bae, Y.; Kim, H.; Kang, S., Plug-and-playable fluorescent cell imaging modular toolkits using the bacterial superglue, SpyTag/SpyCatcher. *Chem. Commun.* **2016**, *52* (97), 14051-14054.
20. Sun, F.; Zhang, W. B.; Mahdavi, A.; Arnold, F. H.; Tirrell, D. A., Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111* (31), 11269-11274.
21. Miladi, B.; El Marjou, A.; Boeuf, G.; Bouallagui, H.; Dufour, F.; Di Martino, P.; Elm'selmi, A., Oriented immobilization of the tobacco etch virus protease for the cleavage of fusion proteins. *Journal of biotechnology* **2012**, *158* (3), 97-103.
22. Zhang, Z.; Tang, R.; Zhu, D.; Wang, W.; Yi, L.; Ma, L., Non-peptide guided auto-secretion of recombinant proteins by super-folder green fluorescent protein in Escherichia coli. *Scientific Reports* **2017**, *7*, 6990.
23. Sambrook, J.; Russell, D. W.; Sambrook, J.; Russell, D. W., *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press {a} , 10 Skyline Drive, Plainview, NY, 11803-2500, USA: 2001.

Chapter 6

Characterization of ST_HEAT Nanotubes and Functionalization with mCherry_SpyCatcher

6.1 Introduction

The HEAT-repeat motif displays remarkable tolerance with respect to modifications of the hypervariable residue positions.¹ These hypervariable residues are potentially functional positions that can be assigned with high precision through sequence specificity. The asymmetry in the hypervariability of the convex exterior and concave interior surfaces provide asymmetry to the functionalization of the two surfaces. The **HEAT_R1** sequence, which was based on the consensus sequence of the PBS HEAT-like repeat motif,² was used for the design of the HEAT-based concatemer protein, **HEAT_6R**, and both formed highly stable nanotubes with hypervariable amino acids presented on the inside channel. The reconstruction of the **HEAT_R1** nanotubes revealed a right-handed helix with 10.3 asymmetric dimer units per turn (total of 20.7 HEAT-repeats per helical turn) with a pitch of 31 Å. Additionally, initial biophysical results for **HEAT_6R** indicated that the HEAT-based hexamer was capable of forming closed cylindrical assemblies with structurally robust, thermally stable helical nanotubes. The HEAT-repeat derived concatemers present a promising building block for the fabrication of helical assemblies and barrier materials with structural and functional asymmetry.

The **ST_HEAT** sequence was identical to the **HEAT_6R** sequence, with the exception of the number of repeats (8 repeats instead of 6) and the addition of an *N*-terminal SpyTag linking unit. Due to a correlation between the number of tandem repeats and the stability of the structure,³⁻⁴ the eight repeat concatemer would likely result in highly stable supramolecular structures. **ST_HEAT** and its the ligation partner, **mCh_SC**, composed of a fusion of the mCherry fluorescent protein and the SpyCatcher protein, were successfully expressed and purified in the previous chapter.

Optimization of the assembly conditions for **ST_HEAT** is required for structural characterization of the nanotubes. Biophysical analysis would illuminate the structural similarities and differences between the **ST_HEAT** protein, the **HEAT_6R** protein, and the **HEAT_R1** peptide. Although the optimization and thorough biophysical characterization of protein assemblies can be a lengthy process, the design of the system would support modular functionalization with a wide variety of proteins. Once optimized, the nanotubes would be ready for functionalization by potentially any fusion protein that contains the SpyCatcher ligation partner. Verification of **mCh_SC** functionalization of **ST_HEAT** nanotubes would indicate utility of these nanotubes with asymmetric surfaces.

Assembly and characterization of the **ST_HEAT** protein afforded high aspect-ratio nanotubes of uniform diameter with SpyTag protruding from the exterior surface. Introduction of the **mCh_SC** ligation partner resulted in the production of nanotubes with tiny white projections protruding from the surface, which were hypothesized to be mCherry protein fused to the exterior surface of the nanotubes. Florescence analysis of the conjugated nanotubes verified the presence of mCherry on the exterior surfaces of the nanotubes.

The stable **ST_HEAT** nanotubes provide a promising substrate for the asymmetric functionalization of the two surfaces due to a hypervariable concave channel and the presence of the SpyTag peptide on the convex surface. The use of the SpyTag/SpyCatcher ligation system allows for the functionalization of the convex surface of the optimized **ST_HEAT** nanotubes with arbitrary proteins fused to the SpyCatcher ligation partner. Additionally, the system prevents the need to optimize HEAT-based nanosheets each time a novel functional protein is fused to the surface. Only one HEAT-based fusion product, **ST_HEAT**, is required to support the attachment of a potentially limitless range functional proteins.

6.2 Results and Discussion

6.2.1 Assembly and Characterization of SpyTag_HEAT Nanotubes

Low Concentration HisPurTM ST_HEAT Assemblies

Initial assemblies of the **ST_HEAT** nanotubes were based off of the conditions that **HEAT_R1** was subjected to in order to discern optimal buffer conditions. The first round of assemblies was at a lower peptide concentration of 3 mg/mL using the HisPurTM **ST_HEAT** protein that was purified by affinity chromatography (see previous chapter). The peptide was assembled in 10 mM acetate (pH 5.0), 10 mM MES (pH 6.0), 10 mM MOPS (pH 7.0), and 10 mM TAPS (pH 8.0). While attempting to dissolve the peptide in the buffers, it was obvious that the solubility of the peptide could present an issue under some conditions. White precipitate was present in all buffers, with the most aggressive precipitation occurring in the buffers with higher pH values. It appeared as though the solubility of **ST_HEAT** decreased as it passed its pI value of pH 5.5 and then remained insoluble as the solution became more basic.

After thermal annealing of HisPurTM **ST_HEAT** in the various buffers, solubility of **ST_HEAT** appeared to increase in acetate buffer (pH 5.0), but precipitate remained in the buffers above pH 6.0. The actual peptide concentrations in the assemblies were calculated to be 77 μ M in acetate buffer (pH 5.0), 75 μ M in MES buffer (pH 6.0), 82 μ M in MOPS buffer (pH 7.0), and 45.94 μ M in TAPS buffer (pH 8.0). The peptide in the TAPS solution heavily precipitated making it difficult to collect a representative sample for concentration analysis, which resulted in the significantly lower concentration for that sample.

Circular dichroism spectropolarimetry (CD) was utilized to determine information about the secondary structure of the **ST_HEAT** assemblies. Analysis was performed on the nascent and thermally annealed 3 mg/mL HisPur™ **ST_HEAT** assemblies in 10 mM acetate (pH 5.0), 10 mM MES (pH 6.0), 10 mM MOPS (pH 7.0), and 10 mM TAPS (pH 8.0) buffers. The CD spectra for the assembly in acetate buffer (pH 5.0) exhibited a weak α -helical signature with a maximum at 195 nm and minima around 208 and 222 nm (**Figure 6.1 a**). The thermally annealed sample revealed a higher intensity indicating that thermal annealing may promote the α -helical secondary structure in **ST_HEAT** assemblies. The low peptide concentration was the cause of the weak signal intensity. The assemblies in the 10 mM MES (pH 6.0), 10 mM MOPS (pH 7.0), and 10 mM TAPS (pH 8.0) buffers did not produce relevant CD spectra, as you can see in **Figure 6.1 b-d**. This is the result of scattering due to the insolubility of the peptide under these conditions.

Transmission electron microscopy (TEM) was applied to analyze the **ST_HEAT** assemblies. Based on the **HEAT_R1** peptide and **HEAT_6R** protein assemblies, we anticipated that the **ST_HEAT** assemblies would form nanotubes of 9.7 nm in diameter. TEM analysis was performed on the nascent and thermally annealed 3 mg/mL HisPur™ **ST_HEAT** assemblies in 10 mM acetate (pH 5.0), 10 mM MES (pH 6.0), 10 mM MOPS (pH 7.0), and 10 mM TAPS (pH 8.0) buffers. Similar to CD analysis, TEM confirmed that there was little useful information that could be gathered from the assemblies in buffers above pH 6.0. The samples contained a mixture of unassembled peptide or large undefined protein aggregates. On the other hand, the assemblies formed in acetate buffer (pH 5.0) confirmed nanotube formation. The thermally annealed sample displayed highly defined nanotubes with a visible inner lumen, which is indicative of a pore traversing the tube (**Figure 6.2 b**). The nascent assembly was composed of peptide that was grouping into preformed tubes with less definition (**Figure 6.2 a**).

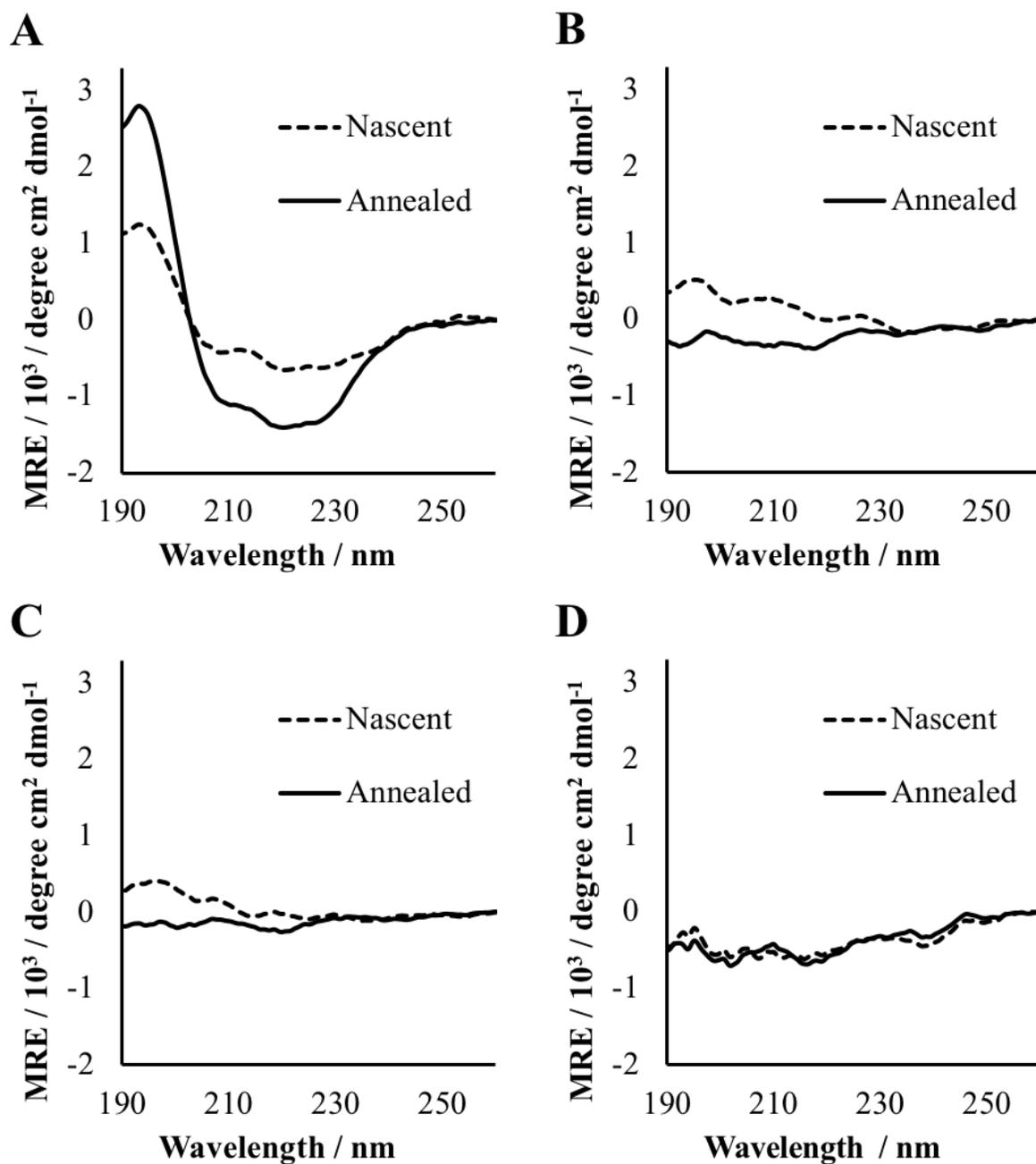


Figure 6.1 Circular dichroism spectra of nascent and thermally annealed assemblies of HisPur[™] ST_HEAT. (A) CD spectra of 77 μM ST_HEAT in 10 mM acetate, pH 5.0. (B) CD spectra for 75 μM ST_HEAT in 10 mM MES, pH 6.0. (C) CD spectra for 82 μM ST_HEAT in 10 mM MOPS, pH 7.0. (D) CD spectra for 46 μM ST_HEAT in 10 mM TAPS, pH 8.0.

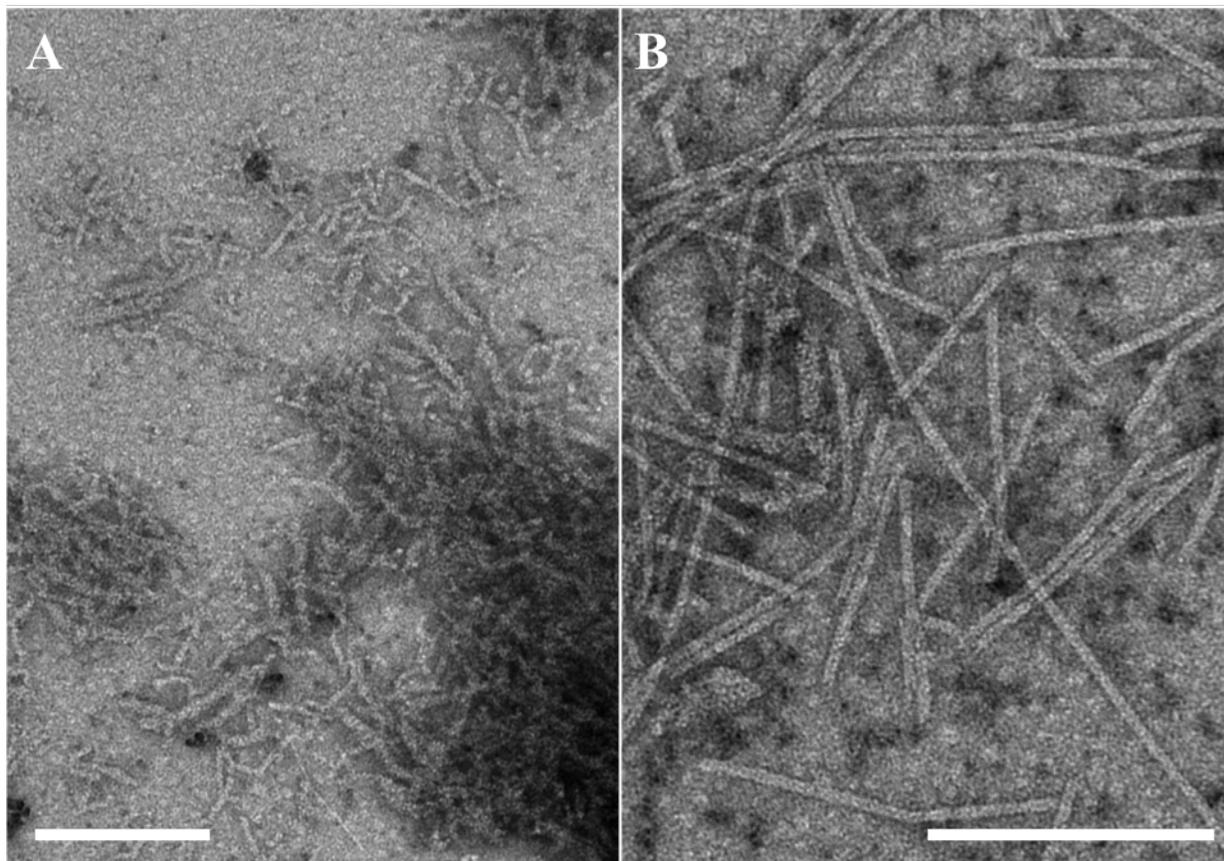


Figure 6.2 Transmission electron microscopy images of 77 μM HisPurTM ST_HEAT assemblies in 10 mM acetate, pH 5.0. (A) TEM image of nascent assemblies. (B) TEM image of thermally annealed assemblies. Scale bar = 200 nm.

The ligation process of SpyCatcher and SpyTag can readily occur between pH 4-8, so the low pH range for ST_HEAT assembly would not hinder linking. The mCh_SC fluorescent protein does not begin to decrease in solubility and fluorescence until the solution drops below pH 4.0, so assembling ST_HEAT at pH 5.0 appeared as though it would not be an issue.

Although the assemblies in acetate buffer primarily produced nanotubes, there were some large aggregates present as well. These could be formed by impurities or misassembled protein.

Metal ions may interact with charged amino acids in the helical structure either blocking the designed electrostatic interactions between peptides or inducing incorrect interactions. In addition, metal ions have been known to cause the aggregation of his-tagged proteins. Cobalt ions can potentially leach off of the HisPurTM metal affinity resin contributing metal ions to the solution. Ethylenediaminetetraacetic acid (EDTA) is a chelator that sequesters metal ions in solution, which consequently reduces their activity, and it is commonly added to protein solutions to hinder metal ion interactions.

To rule out metal ion induced protein aggregation, we assembled the **ST_HEAT** protein in the presence of ethylenediaminetetraacetic acid (EDTA). The addition of EDTA would divert metal ions away from the protein and direct assembly according to the designed hydrophobic and electrostatic interactions between helices. The HisPurTM **ST_HEAT** protein was assembled at a lower peptide concentration of 3 mg/mL in 10 mM acetate buffer (pH 5.0) containing 1 mM EDTA. An identical protein solution was prepared without EDTA for comparison. The protein concentrations of the samples were adjusted to 60 μ M.

The nascent and thermally annealed 60 μ M HisPurTM **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0) and 10 mM acetate, 1 mM EDTA buffer (pH 5.0) were analyzed by TEM to determine if metal ions were affecting the protein assemblies. TEM analysis indicated that there was no significant difference between the nascent and thermally annealed samples prepared in acetate buffer with and without EDTA (**Figure 6.3**). The formed nanotubes displayed similar characteristics and the presence of aggregates were comparable between the two samples. This indicates that aggregation in the samples was not a result of metal ions interfering with designed electrostatic interactions or inducing His-tag aggregation.

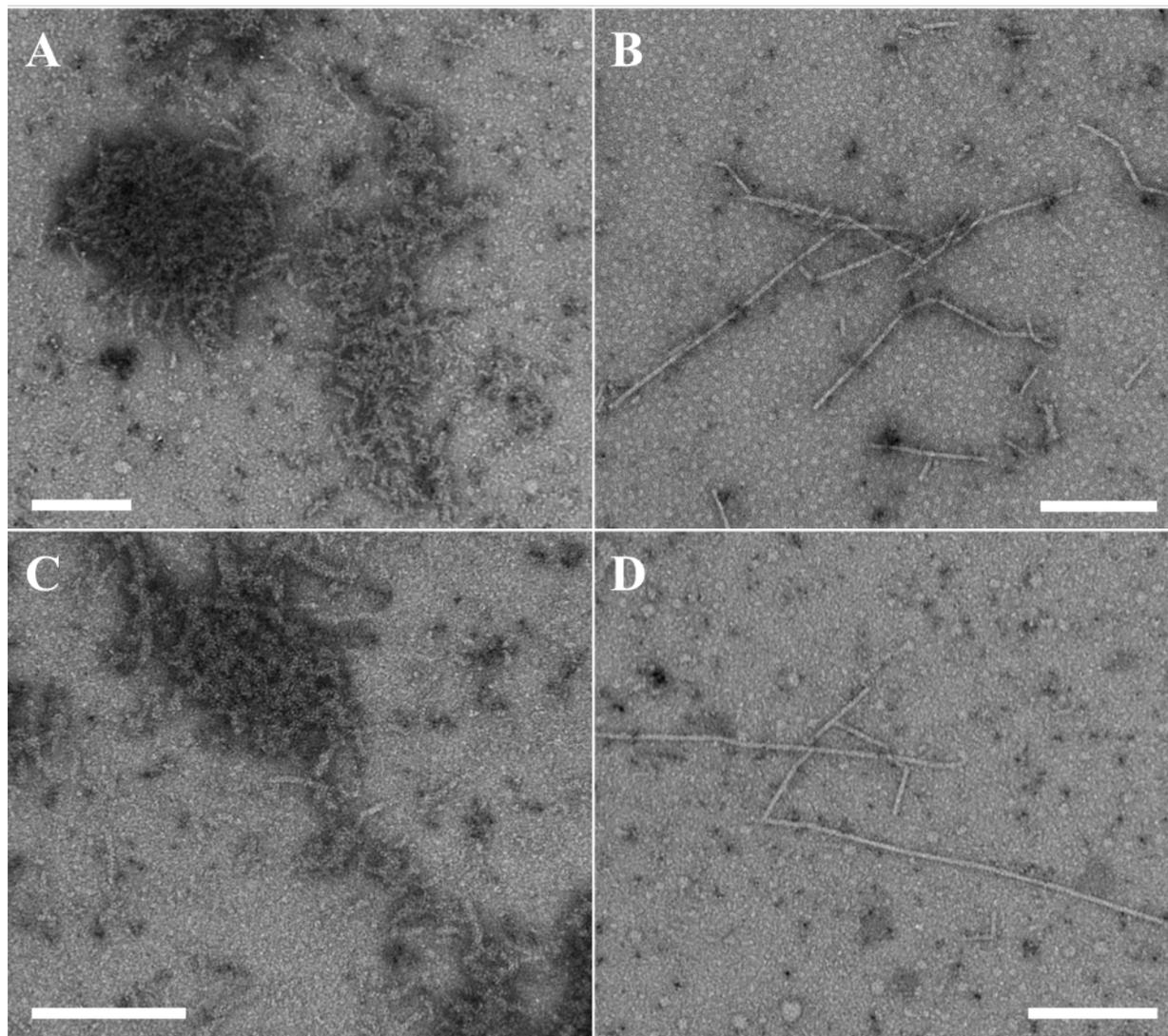


Figure 6.3 Transmission electron microscopy images of 60 μM HisPur[™] ST_HEAT assemblies. (A) TEM image of nascent assemblies in 10 mM acetate, pH 5.0. (B) TEM image of thermally annealed assemblies in 10 mM acetate, pH 5.0. (C) TEM image of nascent assemblies in 10 mM acetate, 1 mM EDTA, pH 5.0. (D) TEM image of thermally annealed assemblies in 10 mM acetate, 1 mM EDTA, pH 5.0. Scale bar = 200 nm.

High Concentration HisPurTM ST_HEAT Assemblies

The CD spectra of the nascent and thermally annealed **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0) exhibited a weak α -helical signal when assembled at 3 mg/mL. This was due to a low protein concentration, which was calculated to be 77 μ M. A 9 mg/mL sample of HisPurTM **ST_HEAT** was assembled in 10 mM acetate buffer (pH 5.0) to acquire CD spectra with more relevant intensities. In addition, we wanted to analyze the effect of concentration on assembly formation. The calculated concentration of the sample was 200 μ M. The thermally annealed solution was opaque, which may be due to the high protein concentration saturating solution.

Circular dichroism spectropolarimetry (CD) was used to analyze the nascent and thermally annealed 200 μ M HisPurTM **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0). As expected, the CD spectra of the 200 μ M nascent and thermally annealed assemblies revealed much stronger α -helical signatures compared to the low concentration assemblies analyzed at 77 μ M (**Figure 6.4**). For comparison, the CD spectra of the 77 μ M samples are shown on the same scale as the 200 μ M samples. Although the signal intensity was superior for the samples assembled at higher concentration, the CD spectra exhibited a decrease in signal intensity at lower wavelengths. This is an expected consequence of the solution opacity causing differential light scattering of left- and right-circularly polarized light, which is wavelength-dependent and displays the greatest distortion in the low wavelength range.⁵⁻⁹ The root cause behind the cloudiness of the solution would need to be further investigated to mitigate this effect.

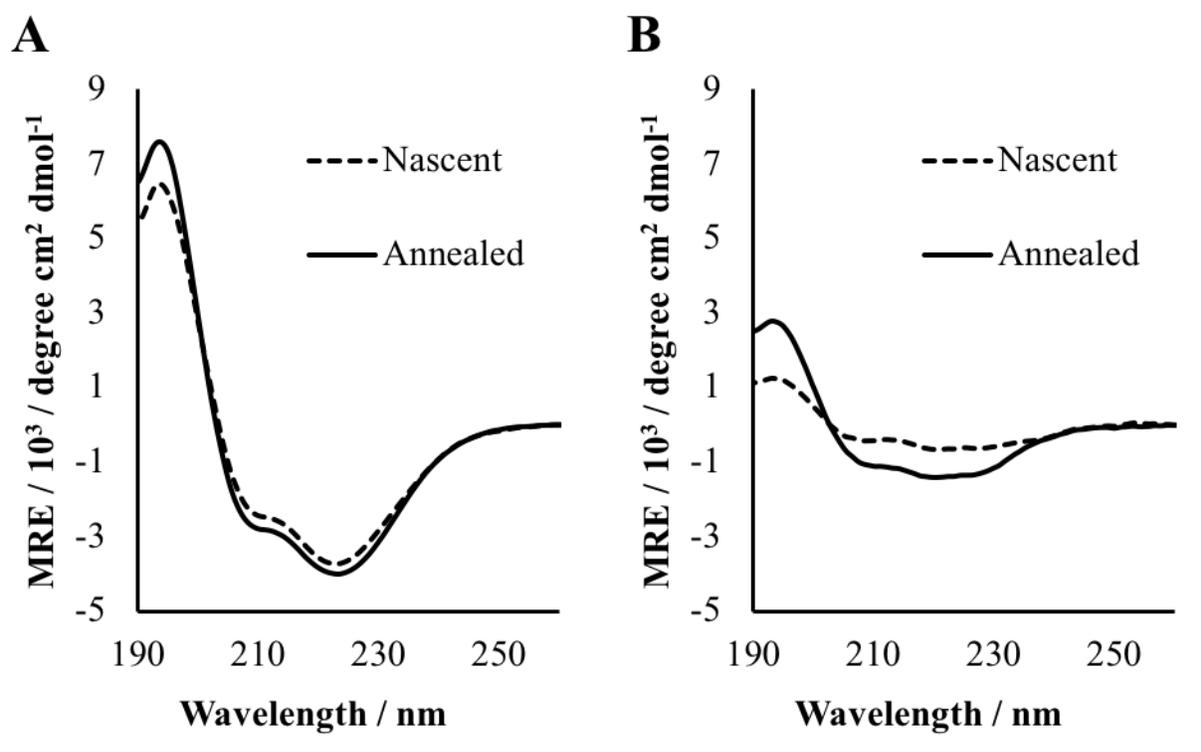


Figure 6.4 Circular dichroism spectra of nascent and thermally annealed assemblies of HisPur™ ST_HEAT. (A) CD spectra of 200 μM ST_HEAT in 10 mM acetate, pH 5.0. (B) CD spectra of 77 μM ST_HEAT in 10 mM acetate, pH 5.0, are shown for comparison.

Transmission electron microscopy (TEM) images were collected to analyze the nascent and thermally annealed 200 μM HisPur™ ST_HEAT assemblies in 10 mM acetate buffer (pH 5.0). Similar to the lower concentration assemblies, the thermally annealed sample displayed highly defined nanotubes (**Figure 6.5 b**). However, unlike the lower concentration sample, the nascent sample at higher concentration contained fully formed nanotubes (**Figure 6.5 a**). The higher concentration of protein in solution allowed the formation of tubes without thermal annealing. Therefore, the assembly process was concentration dependent, although CD analysis indicated that the tubes formed after thermal annealing possess a greater magnitude of the α -helical secondary structure.

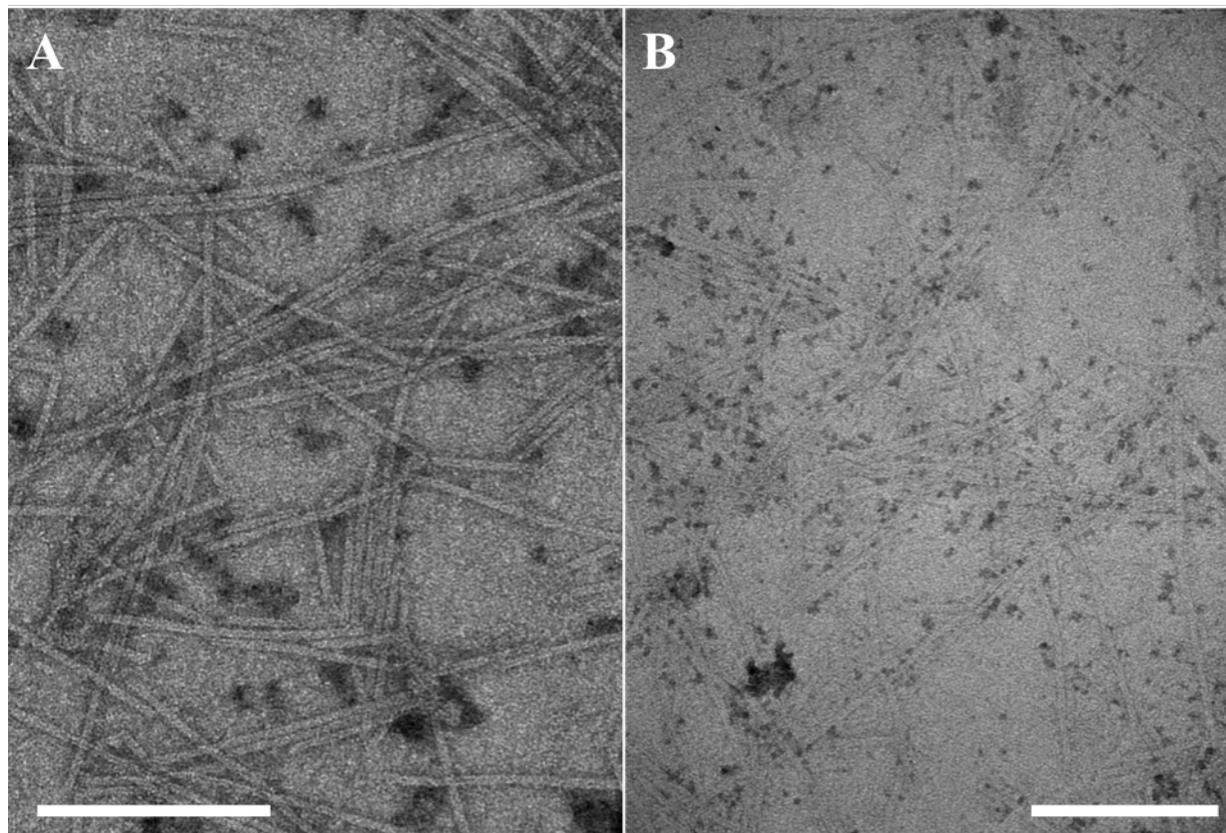


Figure 6.5 Transmission electron microscopy images of 200 μM HisPur[™] ST_HEAT assemblies in 10 mM acetate, pH 5.0. (A) TEM image of nascent assemblies. (B) TEM image of thermally annealed assemblies. Scale bar = 200 nm.

Although analysis of HisPur[™] ST_HEAT assemblies was promising, it brought up a few concerns. First, the mass of ST_HEAT weighed for the assemblies was much higher than the actual protein mass calculated after assembly. For example, the samples assembled from 9 mg/mL of lyophilized powder yielded a protein concentration of 200 μM or 6 mg/mL. This would indicate that only 67% of the weight was protein mass in solution. Most of this discrepancy would result from the low solubility of the peptide, but it could also be due to impurities. Secondly, thermally annealed assemblies at higher peptide concentration exhibited a minimal degree of opacity. This

could be due to saturation of the solution with a high concentration of protein, which was already displaying some characteristics of insolubility at low concentration. Lastly, even at low concentration, TEM analysis indicated the occasional presence of large aggregates. Assemblies including EDTA were used to rule out metal induced aggregation. All three of these concerns suggest that the HisPur™ **ST_HEAT** was not completely pure and that low solubility was still an issue. Although SDS-PAGE analysis of the HisPur™ **ST_HEAT** eluent (see previous chapter) indicated a relatively pure protein, small molecules may be left over contributing to the overall mass of the powder and promoting insolubility and aggregation.

High Concentration Assemblies of HPLC Purified ST_HEAT

Reverse-phase high performance liquid chromatography (HPLC) was utilized to purify the HisPur™ **ST_HEAT** protein. Further purification by HPLC would improve the purity of the assemblies leading to fewer aggregations. The **ST_HEAT** protein was purified against a water-acetonitrile (+0.1% TFA) gradient and eluted at 47.8% acetonitrile. The HPLC purified **ST_HEAT** was lyophilized to produce a fine white powder.

Initially, the HPLC purified **ST_HEAT** was analyzed by assembling at a protein concentration of 5 mg/mL in 10 mM acetate (pH 5.0). The calculated concentration was significantly higher (600 μ M) than the concentrations seen for HisPur™ protein samples. The assemblies were also perfectly clear with no indications of insoluble aggregates.

Circular dichroism spectropolarimetry (CD) was used to analyze the nascent and thermally annealed 80 μ M HPLC **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0). (**Figure 6.6 a**) The CD spectra of the 80 μ M nascent and thermally annealed HPLC pure assemblies exhibited significantly stronger α -helical signatures than the 200 μ M HisPur™ assemblies despite the lower

protein concentration. This was a result of increased solubility of the **ST_HEAT** protein after HPLC purification due to the removal of impurities. This was confirmed by the absence of signal intensity distortion at lower wavelengths.

The thermally annealed sample of 80 μM HPLC pure **ST_HEAT** protein in 10 mM acetate buffer (pH 5.0) was further analyzed by monitoring the ellipticity at a wavelength of 222 nm from 5 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$. A melting curve was generated, and it failed to indicate a melting transition (T_m) within that temperature range (**Figure 6.6 b**). The ellipticity only showed a minimal decrease as the temperature increased. Standard CD spectra were collected before and after the melt, and the spectra from the two temperature points were similar (**Figure 6.6 c**). This data indicated that the nanotubes were forming thermodynamically stable nanotubes within the temperature range analyzed, which agreed with the data collected for the **HEAT_R1** peptide and **HEAT_6R** protein.

Transmission electron microscopy (TEM) was used to analyze the 80 μM HPLC purified **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0). The thermally annealed sample exhibited nanotubes similar to the assemblies produced from HisPurTM protein (**Figure 6.7 a**). On a positive note, there were no visible aggregates in the sample implying that HPLC purification of the protein removed the underlying source of the undesirable aggregation. However, the density of the tubes was very high leading to a huge forest of tubular assemblies across the whole surface of the grid. Further optimization was attempted by diluting the 5 mg/mL sample to 20 μM , 5 μM , and 1 μM concentrations for imaging. Dilution of the sample failed to produce TEM images with a low density of defined tubes, as you can see in the TEM image of the 5 μM HPLC purified **ST_HEAT** assemblies (**Figure 6.7 b**).

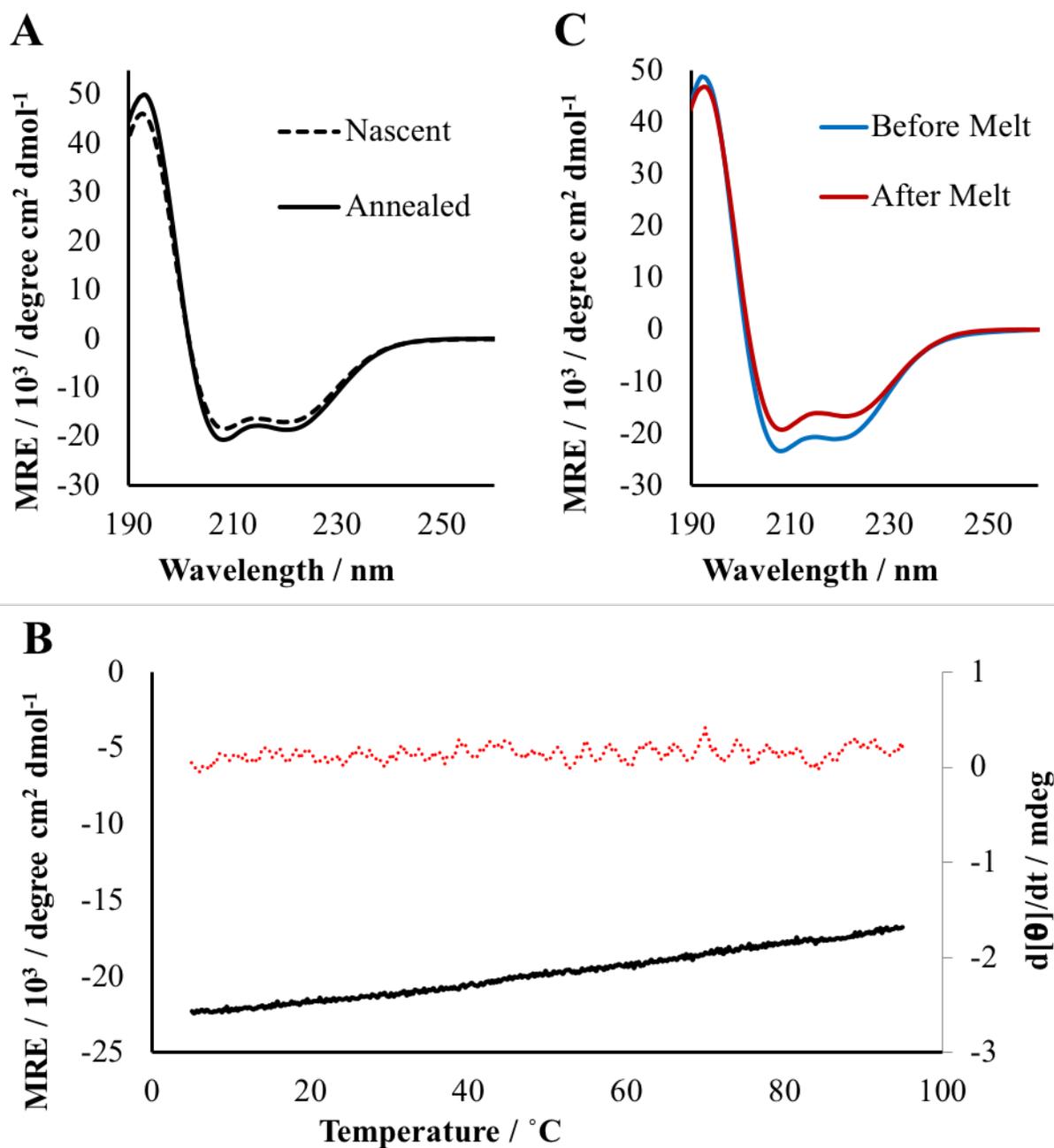


Figure 6.6 Circular dichroism analysis of 80 μM HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0. (A) CD spectra of nascent and thermally annealed assemblies (B) CD melting curve of thermally annealed assemblies. The molar residue ellipticity was monitored at 222 nm from 5 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$. The first derivative is displayed by the red dotted line. (C) CD spectra collected before and after thermal denaturation at 5 $^{\circ}\text{C}$ and 95 $^{\circ}\text{C}$, respectively.

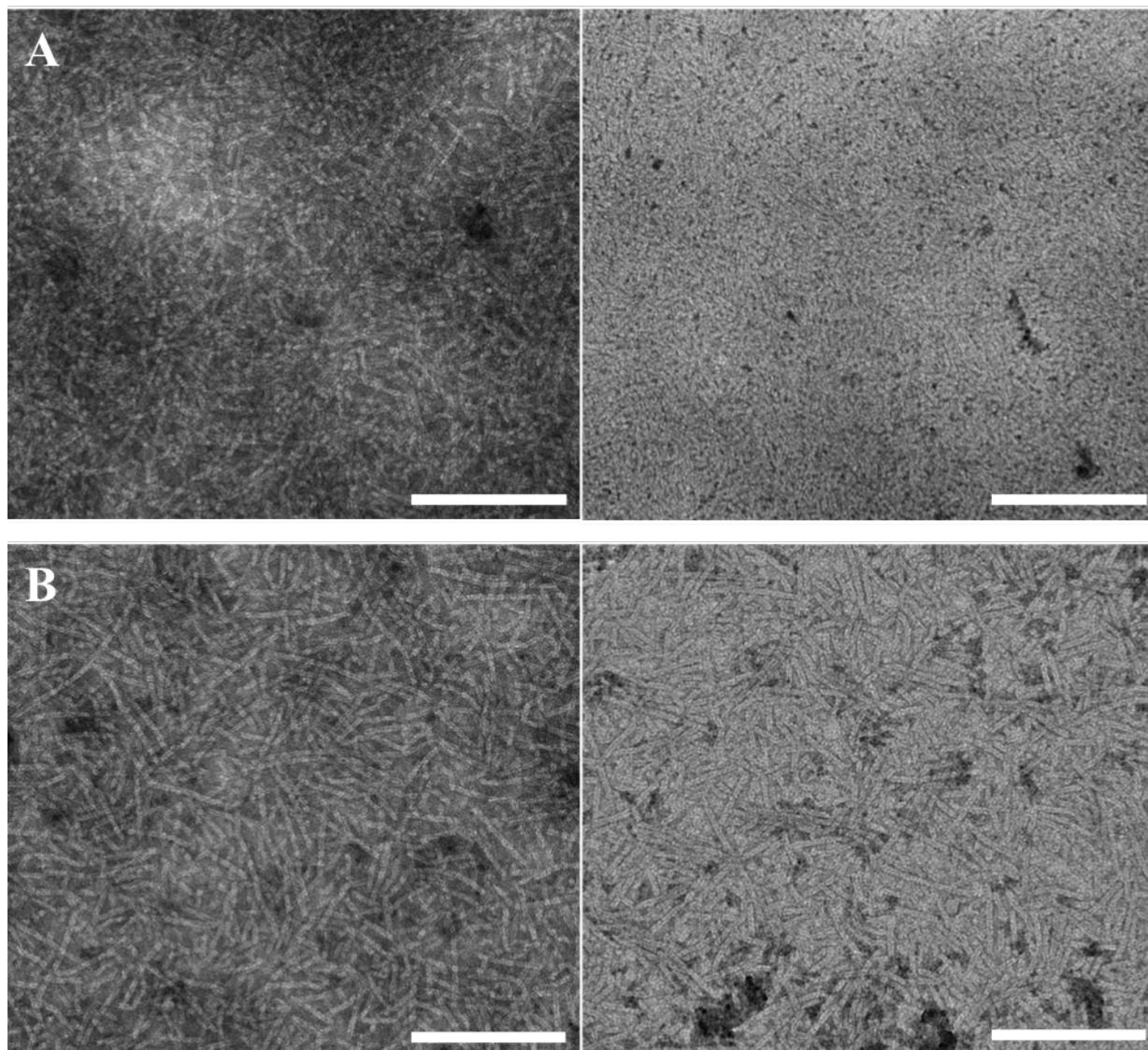


Figure 6.7 Transmission electron microscopy images of HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0. Samples were diluted from an assembly of 5 mg/mL peptide. (A) TEM image of 80 μ M thermally annealed assemblies. (B) TEM image of 5 μ M thermally annealed assemblies. Scale bar = 200 nm.

Low Concentration Assemblies of HPLC Purified ST_HEAT

After exhausting many avenues to produce disperse tubular nanostructures from dilutions of the 5 mg/mL assembly condition, a sample was assembled at a very low protein concentration of 0.33 mg/mL HPLC purified **ST_HEAT**, which translated into a concentration of 5 μ M **ST_HEAT** protein. Transmission electron microscopy (TEM) images were collected to analyze the nanotubes when assembled at a low protein concentration (5 μ M). The 5 μ M thermally annealed HPLC purified sample exhibited nanotubes that were significantly more disperse (**Figure 6.8**) than the 5 μ M sample that was diluted from the 5 mg/mL assembly. Assembling at a higher protein concentration and diluting to a low concentration only produced heavily intertwined networks of **ST_HEAT** nanotubes that remained clustered after dilution. Assembling at a low protein concentration produced nanotubes that were more disperse and prepared for **mCh_SC** ligation.

ImageJ software was utilized for diameter measurements of HisPurTM and HPLC purified **ST_HEAT** nanotubes from various TEM images. Measurements were collected from TEM images of 77 μ M HisPurTM **ST_HEAT**, 200 μ M HisPurTM **ST_HEAT**, and 5 μ M HPLC **ST_HEAT** to get a fair distribution among various assembly conditions and the data was plotted on a histogram (**Figure 6.9**). The average diameter acquired from a sample size of 160 measurements was 7.955 ± 0.391 nm. There was no significant difference between the diameters of nanotubes assembled from HisPurTM protein and HPLC purified protein. The approximate diameter of 8.0 nm for the **ST_HEAT** nanotubes was reasonable as the value was only slightly smaller than the average diameter of 9.7 nm for the **HEAT_R1** nanotubes.

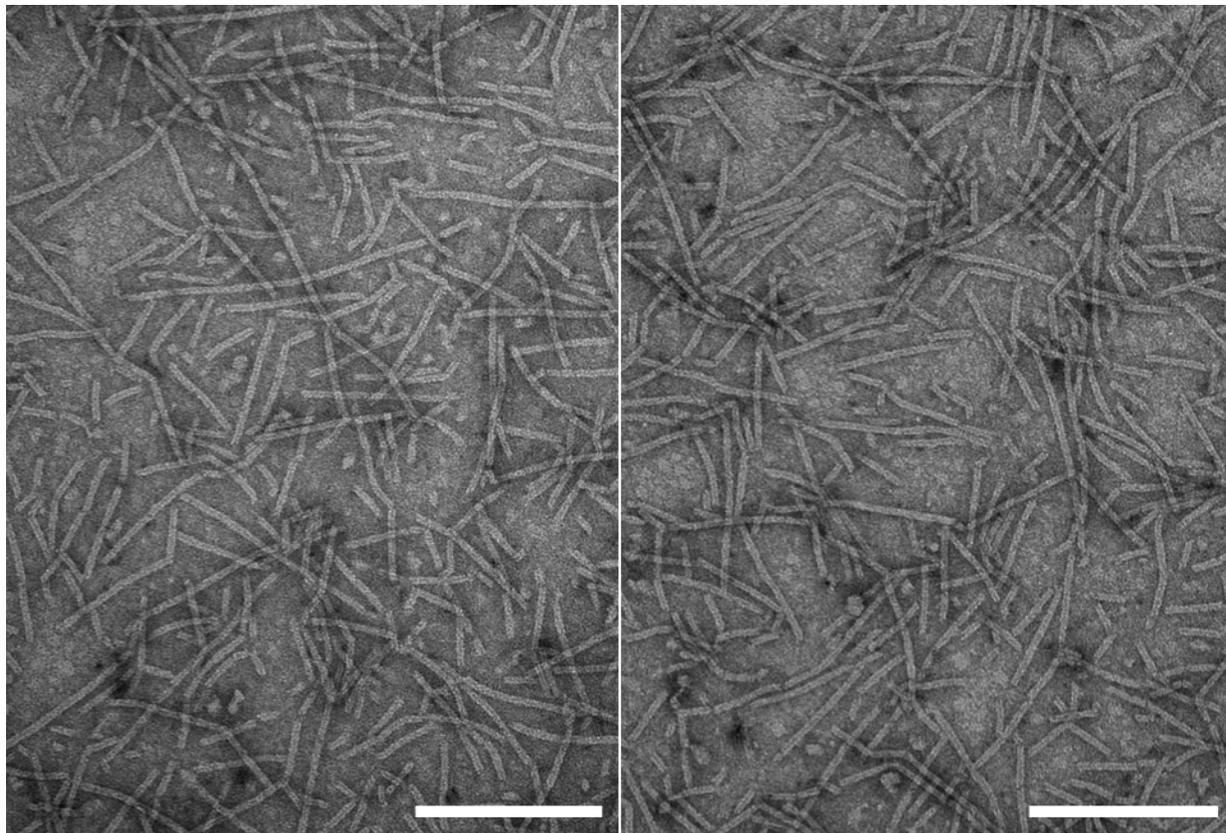


Figure 6.8 Transmission electron microscopy images of thermally annealed 5 μM HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0. Sample was assembled at 5 μM so no dilution was necessary. Scale bar = 200 nm.

Small- and wide-angle X-ray scattering (SAXS/WAXS) was employed to qualitatively analyze the structure of the **ST_HEAT** assemblies in solution. Raw data collected from scattering of the beam was used to determine the size and shape of the assemblies by analyzing the resulting scattering and Guinier plots. The scattering plot indicated the presence of cylindrical structures of uniform diameter. Although there were no diffraction peaks in the high q region, which would be used to determine the distances between crystalline planes, the scattering plot did contain an oscillation with a deep minimum at 0.11 \AA^{-1} (**Figure 6.10**). This indicated that the sample was monodisperse and well-ordered with identical particles of the same size. Using the equation

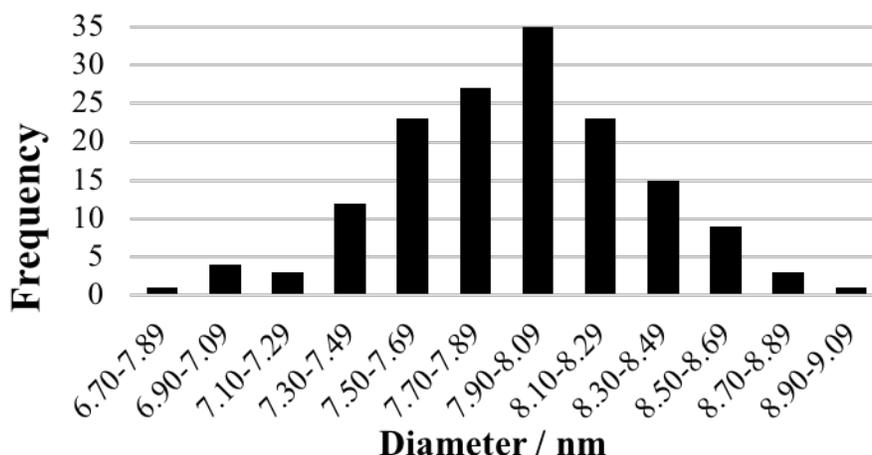


Figure 6.9 Histogram of ImageJ measurements collected from TEM images of thermally annealed HisPurTM and HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0. The experimental average diameter is 7.955 ± 0.391 nm from a sample size of 160.

$d = 2\pi/q_{peak}$ and inserting the oscillation minimum of 0.11 for q_{peak} , a length of 57.12 Å was calculated. Further structural analysis would be required to assign this measurement.

The SAXS data for ST_HEAT was fit in the low q region to a modified Guinier model for rod-like forms to describe the size of the species. A straight-line fitting of a linear segment of the Guinier plot was used to generate a cross-sectional radius of gyration, R_c (**Figure 6.11**). The R_c value for the ST_HEAT nanotubes was 27.248 ± 0.22135 Å. Using the equation $R_c^2 = (R_o^2 + R_i^2)/2$, the outside and internal radii were determined. The ST_HEAT nanotubes exhibited a calculated internal radius (R_i) of 20.53 Å and an outside radius (R_o) of 38.55 Å. These translated to internal and external diameters of 4.10 nm and 7.71 nm, respectively. The external nanotube diameter of 7.71 nm agreed closely with the diameter measured from TEM images. The diameter was slightly smaller than the external diameter acquired for the HEAT_R1 nanotubes, which was around 9.7 nm.

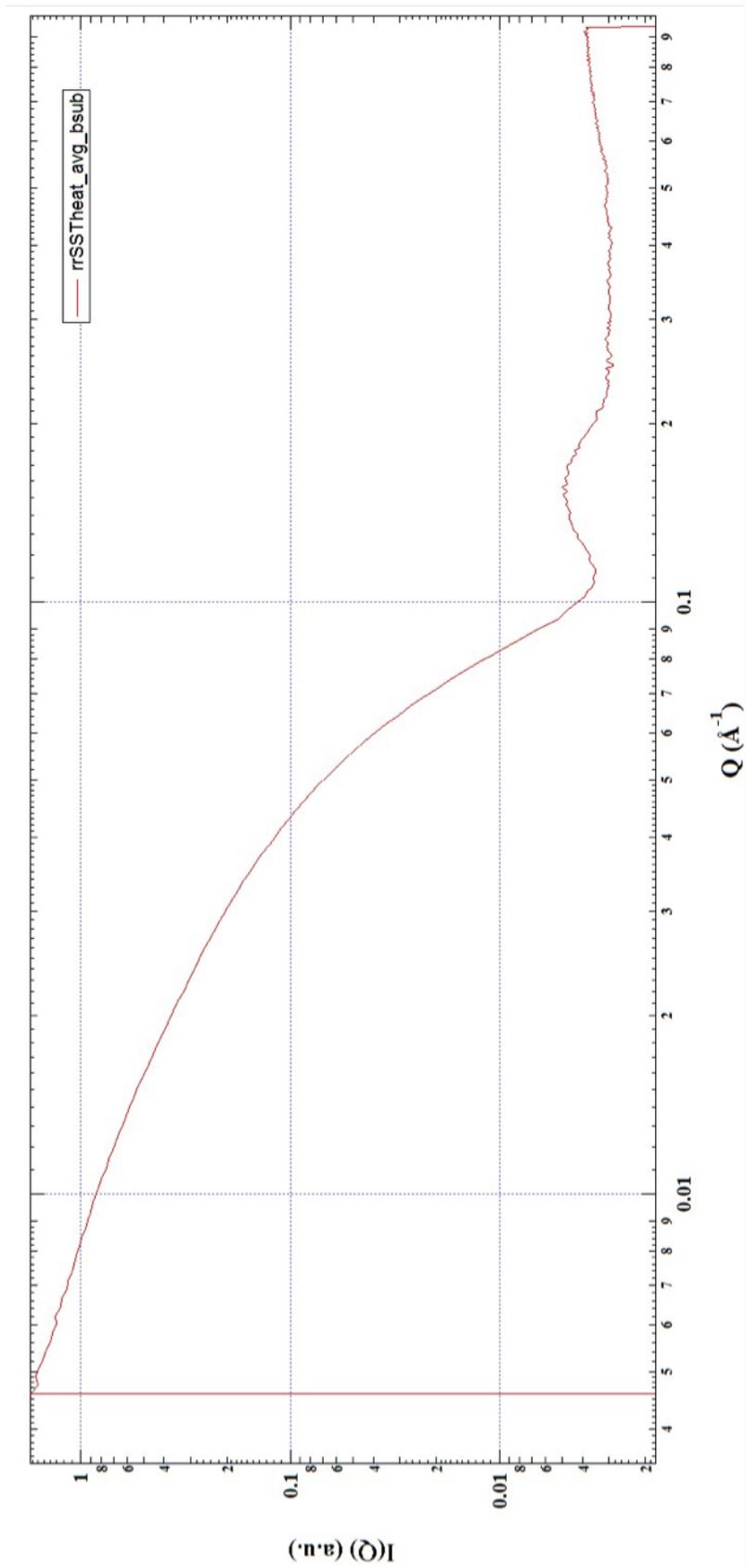


Figure 6.10 Small-angle X-ray scattering curve of thermally annealed 5 mg/mL HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0.

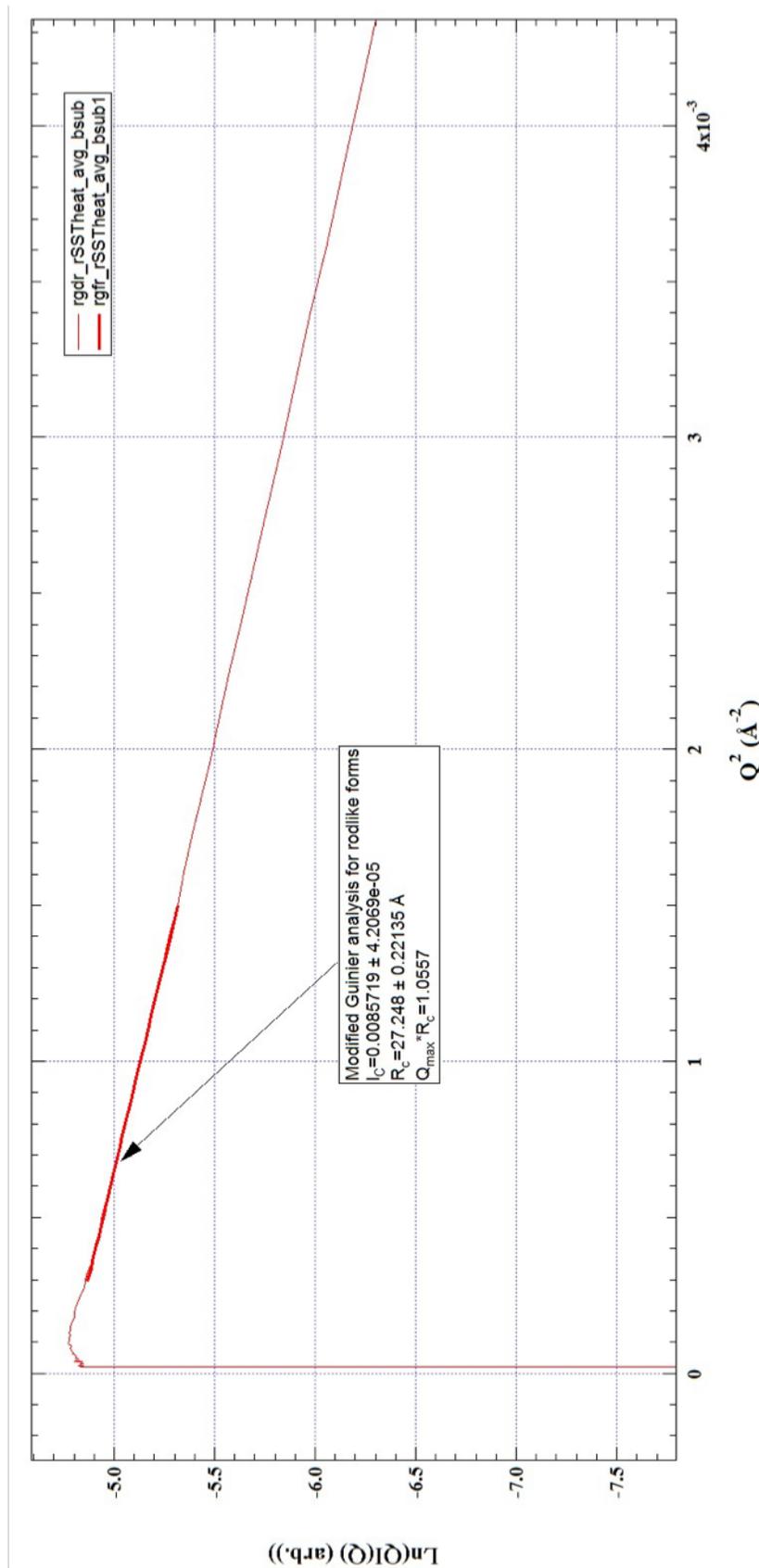


Figure 6.11 Modified Guinier plot from small-angle X-ray scattering of thermally annealed 5 mg/mL HPLC purified ST_HEAT assemblies in 10 mM acetate, pH 5.0. The radius of gyration, R_g , value was $27.248 \pm 0.22135 \text{ \AA}$.

6.2.2 Functionalization of ST_HEAT Nanotubes with mCh_SC Protein

During expression and purification of the **mCh_SC** and **ST_HEAT** proteins, linking experiments were conducted, and the formation of a covalent linkage between the two proteins was confirmed by SDS-PAGE analysis. Once tube formation by HisPur™ **ST_HEAT** was verified, more thorough analysis was performed to assess the feasibility of functionalizing the tubular assemblies with the **mCh_SC** fusion protein. Therefore, functionalization experiments started during the **ST_HEAT** nanotube optimization process, so initial tests were conducted with HisPur™ **ST_HEAT** or high concentration HPLC purified **ST_HEAT** assemblies. Regardless, valuable information was gained from these preliminary experiments.

First, we wanted to determine when the **mCh_SC** protein should be linked to the **ST_HEAT** protein. The SpyCatcher-SpyTag ligation process can occur under an immense range of conditions (pH, buffer) and the covalent linkage formed is resistant to high temperature, etc.¹⁰ Therefore, the two purified proteins could be linked before or after thermal annealing without the fear of heat breaking the linkage. If the **mCh_SC** and **ST_HEAT** proteins were linked before thermal annealing, the linking process would occur before the formation of tubes, which would prevent the linking process from disturbing the tubes after they have formed; however, linking before thermal annealing could also make it more difficult for the HEAT segment to arrange into well-defined tubes. If **ST_HEAT** nanotubes were thermally annealed first, then **mCh_SC** protein would link to the exterior of fully formed tubes. Though, the linking process could potentially disrupt on the structure of the tubular assemblies once they have formed. Both methods were analyzed to establish a suitable protocol for the formation of organized **ST_HEAT** nanotubes functionalized with **mCh_SC** protein.

Functionalization of HisPurTM ST_HEAT with mCh_SC

The first linking experiment was conducted using HisPurTM ST_HEAT and mCh_SC. Two sets of conditions were set up to assess whether mCh_SC should be linked to ST_HEAT before or after tube formation by thermal annealing.

One obstacle that had to be considered was that HisPurTM ST_HEAT was not entirely soluble in nascent solutions of 10 mM acetate buffer (pH 5.0). The ligation process would not likely occur if ST_HEAT remained in the precipitate. Previous successful linking experiments of HisPurTM eluent of ST_HEAT with mCh_SC verified that the reaction buffer, which contained 3 M urea, did not impede the ligation process and mCh_SC remained both soluble and fluorescent. Therefore, in order to facilitate linking before thermal annealing, HisPurTM ST_HEAT was resuspended in 2 M urea, 10 mM acetate buffer (pH 5.0) and diluted to a protein concentration of 80 μ M. Once completely soluble, mCh_SC was added to the solution to acquire a final concentration of 10% (v/v) mCh_SC, which was arbitrarily chosen, and the sample incubated at ambient temperature with shaking for 18 hours. The solution remained red after linking in the urea solution, which indicated that urea was not denaturing mCh_SC.

Transmission electron microscopy was used to visually analyze the ligation product of mCh_SC and ST_HEAT in urea. Fully formed tubes were not anticipated since the solution composed 1.8 M urea; however, if tubes were forming, we would expect them to have a “fuzzy” nature since mCherry would be projecting from the exterior surface of the HEAT-based nanotubes. As shown in **Figure 6.12 a**, the sample appeared to have come together into a network that could be a consequence of loose associations between HEAT subunits of ST_HEAT and ligated mCh-HEAT.

Since the covalent link had been made, the urea could be slowly removed from solution to allow for more appropriate assembly of the nanotubes. The ligated solution was dialyzed in a step-wise manner to remove the urea from the solution. During dialysis, a red precipitate formed and settled on the bottom of the tube with some red still visible in solution. The dialyzed sample was gently mixed and analyzed by TEM. Tube-like formations were visible after dialysis, which was supported by slow removal of urea from solution; however, they had an appearance similar to the pre-formed tubes previously seen in nascent assemblies of HisPur™ **ST_HEAT** (**Figure 6.12 b**).

Thermal annealing would be required for the formation of the best quality tubular assemblies. After the dialyzed assemblies were annealed, there were two important observations about the sample. First, the amount of precipitate in the sample was greatly decreased by thermal annealing, which was expected based on previously reported data. Secondly, the sample no longer exhibited any red coloring or fluorescence under UV-light. Linking of **mCh_SC** to **ST_HEAT** before thermal annealing was not limited by the SpyCatcher-SpyTag ligation process. The covalent link remains stable even at 100 °C. However, **mCh_SC** cannot withstand the thermal annealing process as evident by the loss of all fluorescence. The denaturation of the protein was obvious by the transition of the color from red before thermal annealing to almost clear after thermal annealing. It was apparent that if **mCh_SC** was ligated to nascent **ST_HEAT**, the sample would have to remain unannealed.

A second sample was prepared concurrently with the previous sample to analyze the functionalization of **ST_HEAT** nanotubes with **mCh_SC** after the nanotubes were thermally annealed. Since the precipitate present after **ST_HEAT** assembly usually decreases drastically with thermal annealing, there was no need to assemble the sample with urea resulting in a less

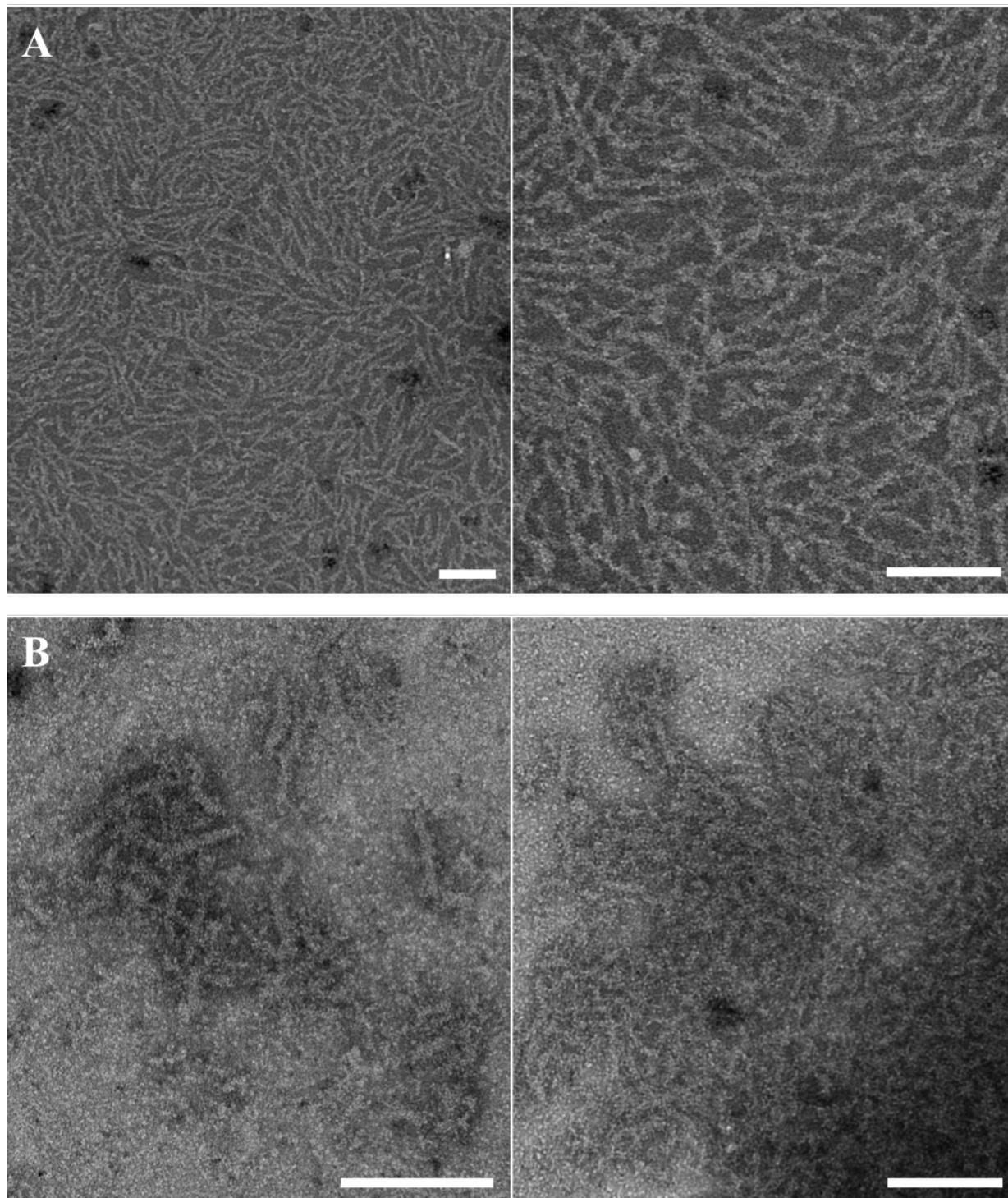


Figure 6.12 Transmission electron microscopy images of a solution of nascent 80 μM HisPur[™] ST_HEAT with 10% (v/v) mCh_SC. (A) Images after 1 day of incubation. (B) Images after dialysis to remove urea. Scale bar = 200 nm.

complicated protocol. HisPur™ **ST_HEAT** was assembled in 10 mM acetate buffer (pH 5.0) and the protein concentration was adjusted to 80 μ M. The sample was thermally annealed and analyzed by transmission electron microscopy to facilitate a comparison of the nanotubes before and after functionalization (**Figure 6.13 a**). The nanotubes appeared similar to those in previous TEM images of thermally annealed HisPur™ **ST_HEAT** nanotubes.

The thermally annealed **ST_HEAT** nanotubes were functionalized with 10% (v/v) **mCh_SC**. The ligation process was allowed to occur for 18 hours with gentle shaking at ambient temperature. TEM analysis was conducted on the functionalized nanotubes and compared to the nanotubes before the ligation process. The acquisition of high-resolution images was made difficult by the inherent “fuzzy” nature of the tubes, but the functionalized tubes were undoubtedly present and made evident by nanotubes covered in tiny protrusions (**Figure 6.13 b**). Based on the simplicity of the protocol and formation of the unmistakably functionalized nanotubes, the method for linking **mCh_SC** to thermally annealed **ST_HEAT** nanotubes was highly promising.

Functionalization of HPLC Purified ST_HEAT with mCh_SC

Once it was discovered that superior nanotubes were formed from HPLC purified **ST_HEAT**, further functionalization experiments were conducted. Since the HPLC purified **ST_HEAT** was assembling into highly dense forests of nanotubes, it was decided to functionalize the nanotubes in 1:2 ratio of **ST_HEAT** to **mCh_SC** fluorescent protein to ensure maximum functionalization. This correlates to a final concentration of 13.8% (v/v) **mCh_SC**.

Before ruling it out, a second attempt was made for the functionalization of nascent **ST_HEAT** with **mCh_SC** since HPLC purified **ST_HEAT** was completely soluble in 10 mM acetate buffer (pH 5.0) and nanotubes formed in nascent samples. The previous attempt revealed

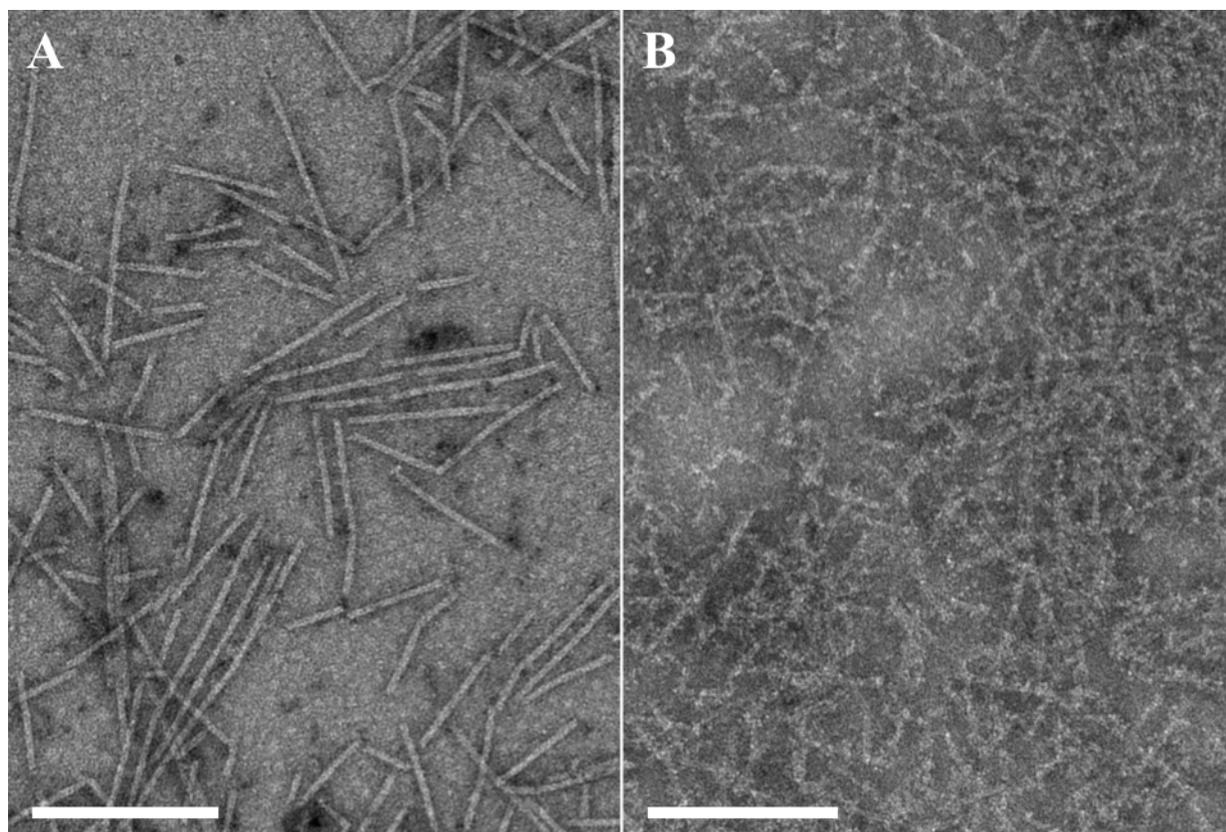


Figure 6.13 Transmission electron microscopy images of (A) a solution of thermally annealed 80 μM HisPur[™] ST_HEAT and (B) the same sample 1 day after the addition of 10% (v/v) mCh_SC. Scale bar = 200 nm.

that **mCh_SC** would denature and lose fluorescence during the thermal annealing process, so the sample would have to remain unannealed. To assess the functionalization of unannealed nanotubes, HPLC purified **ST_HEAT** was assembled in 10 mM acetate buffer (pH 5.0), diluted to a protein concentration of 80 μM , and linked with 13.8% **mCh_SC**. The linked sample was subjected to TEM analysis after incubating at ambient temperature for 18 hours and 4 days. Although the functionalized tubes that formed 18 hours after linking were superior to those seen in the previous unannealed sample, the homogeneity of the sample did not sustain itself after 4

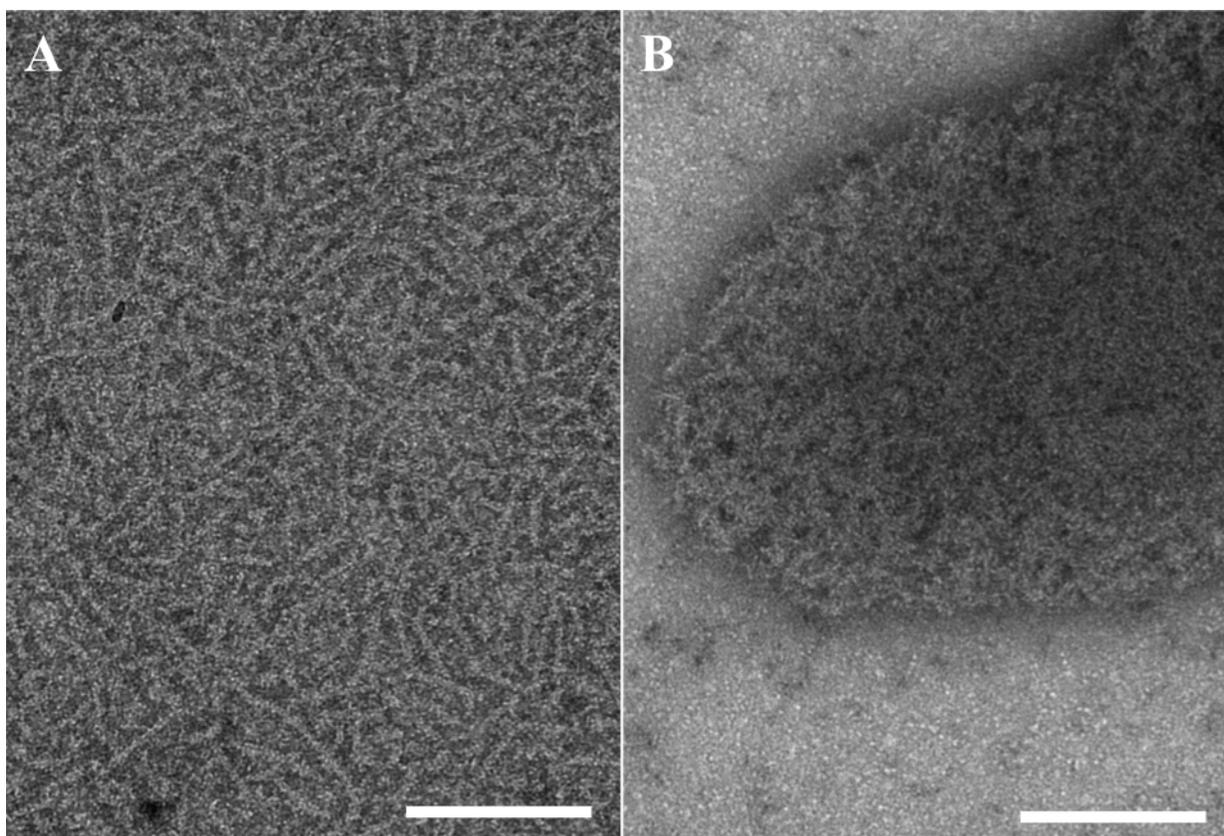


Figure 6.14 Transmission electron microscopy images of a solution of nascent 80 μM HPLC purified ST_HEAT with 13.8% (v/v) mCh_SC in 10 mM acetate buffer (pH5.0). (A) Image after 1 day of incubation. (B) Image after 4 days of incubation. Samples were diluted to 5 μM for analysis. Scale bar = 200 nm.

days (**Figure 6.14**). While the red color of the mCherry fluorescent protein within the sample did not diminish over the course of the 4 days, the nanotubes began to aggregate into either small clusters or large protein aggregates.

Simultaneously, a second sample was prepared to analyze the functionalization of thermally annealed HPLC purified ST_HEAT with mCh_SC. HPLC purified ST_HEAT was

assembled in 10 mM acetate buffer (pH 5.0), diluted to a protein concentration of 80 μM , and then thermally annealed. The sample was linked with 13.8% (v/v) **mCh_SC** and analyzed by TEM after incubation at ambient temperature for 18 hours and 4 days. TEM analysis obtained images of nanotubes that were visibly functionalized with **mCh_SC** (**Figure 6.15**). The functionalized nanotubes appeared more fully formed and defined than the nascent ligation sample, supporting the theory that thermal annealing would lead to the formation of superior nanotubes. In addition,

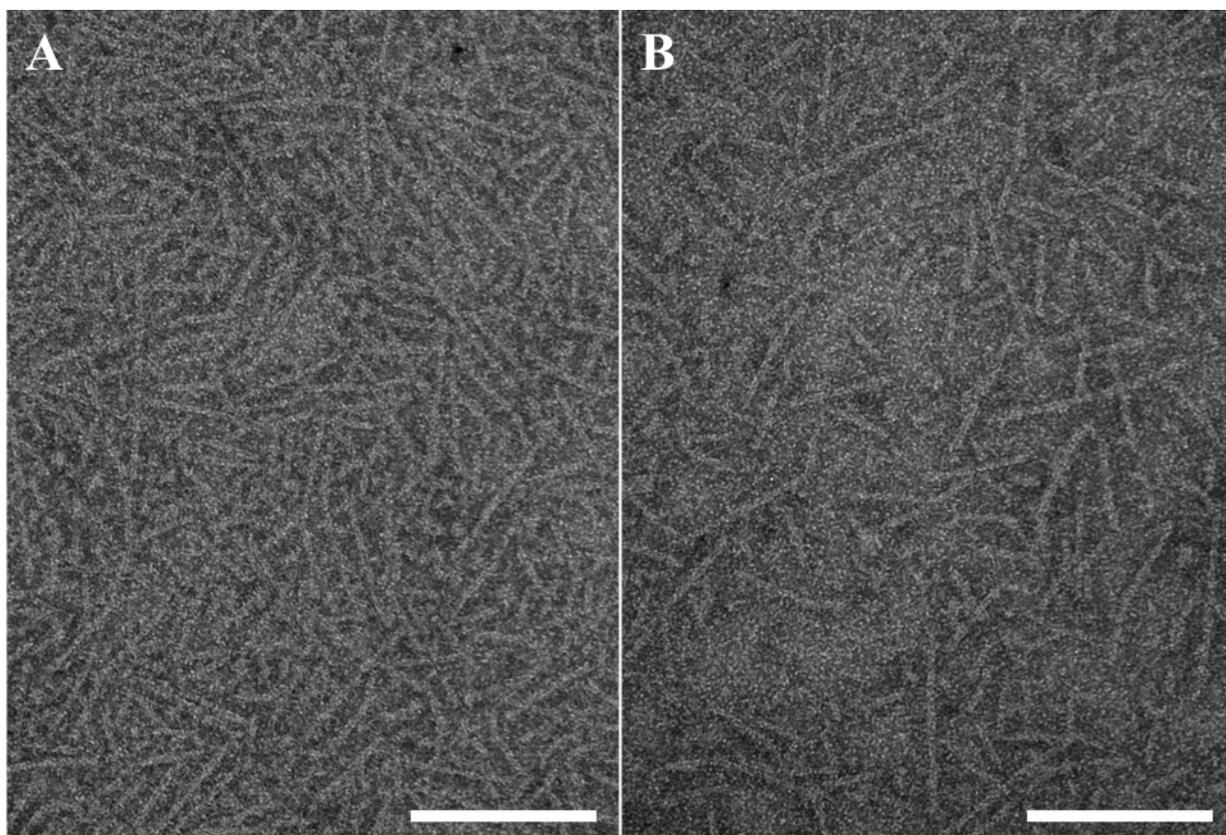
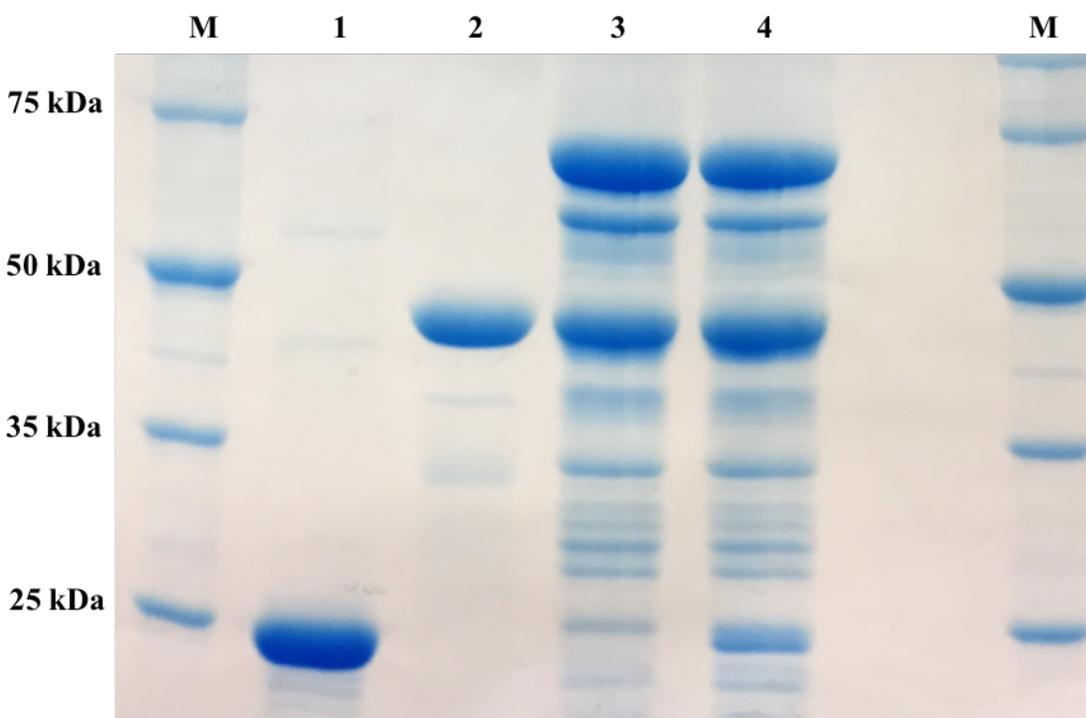


Figure 6.15 Transmission electron microscopy images of a solution of thermally annealed 80 μM HPLC purified ST_HEAT with 13.8% (v/v) mCh_SC in 10 mM acetate buffer (pH5.0). (A) Image after 1 day of incubation. (B) Image after 4 days of incubation. Samples were diluted to 5 μM for analysis. Scale bar = 200 nm.

the nanotubes did not aggregate after 4 days unlike the nascent ligation sample. Furthermore, the sample assembled with HPLC purified **ST_HEAT** formed tubes with higher definition along the edges compared to the samples formed previously from HisPur™ protein.

Samples were collected 18 hours after linking of the nascent or thermally annealed HPLC purified **ST_HEAT** with **mCh_SC** and analyzed using SDS-PAGE on a 10% gel (**Figure 6.16**). The unannealed functionalized nanotube sample, which can be seen in lane 3, did not appear significantly different from the thermally annealed functionalized nanotube sample, which is displayed in lane 4. The only notable difference was a small amount of unlinked **ST_HEAT** in the thermally annealed **ST_HEAT**, which was due to the dense forest of tubes concealing a small fraction of SpyTag ligation sites. Both samples had a significant amount of linked **mCh-HEAT** appearing around 72 kDa and an expected equimolar amount of **mCh_SC** around 42 kDa as a result of the excess from the 1:2 ratio.

Based on the information gathered from the SDS-PAGE gel, the simple linking mechanics were unaffected by whether **mCh_SC** was ligated to unannealed or thermally annealed **ST_HEAT**. However, even though the fluorescent protein could link with the unannealed nanotubes, the nanotubes displayed superior definition after thermal annealing. The nanotubes were linked with an excess amount of **mCh_SC** to ensure maximum functionalization, but the excess unbound **mCh_SC** resulted in difficult visualization of the functionalized **mCh-HEAT** nanotubes using TEM. Additionally, excess **mCh_SC** could cause complications for other analysis methods. At that time, it was important to manage the excess **mCh_SC** by determining the amount of **mCh_SC** required for maximum functionalization and develop methods for removal of any unbound excess.



- M.** Perfect protein marker (10-225 kDa) **3.** mCh-HEAT, nascent ST_HEAT + mCh_SC
1. ST_HEAT, HisPur™ eluent **4.** mCh-HEAT, annealed ST_HEAT + mCh_SC
2. mCh_SC, HisPur™ eluent **M.** Perfect protein marker (10-225 kDa)

Figure 6.16 SDS-PAGE analysis depicting the SpyTag/SpyCatcher ligation results from combining 80 μ M HPLC purified ST_HEAT with 13.8% (v/v) mCh_SC proteins. The sample in lane 3 was linked with a nascent ST_HEAT sample, while the sample in lane 4 was produced from a thermally annealed sample of ST_HEAT. The ligation product, mCh-HEAT, was visualized in both ligation samples. HisPur™ ST_HEAT and mCh_SC are displayed for reference in lanes 1 and 2. MW: ST_HEAT = 30.1 kDa; mCh_SC = 42.4 kDa; mCh-HEAT = 72.4 kDa.

Optimizing Amount of mCh_SC for Maximum Coverage

One obvious consequence of linking a higher concentration of **mCh_SC** was the difficulty to distinguish the functionalized nanotubes from the excess **mCh_SC**. To determine the amount of fluorescent protein required for maximum functionalization of the nanotubes, HPLC purified **ST_HEAT** was assembled in 10 mM acetate buffer (5.0), diluted to a protein concentration of 80 μ M, thermally annealed, and linked with various concentrations of **mCh_SC**. Ratios of 1:4, 1:3, 1:2, 1:1, and 2:1 of **ST_HEAT** to **mCh_SC** were prepared, which corresponded to 24.2%, 19.4%, 13.8%, 7.4%, and 3.9% (v/v) **mCh_SC**, respectively. After ligation of the samples for 18 hours, the samples were analyzed by TEM.

Nanotubes were present in all samples, and they were heavily functionalized in all except for the nanotubes in the 2:1 ratio of **ST_HEAT** to **mCh_SC**, which were not clearly functionalized (**Figure 6.17 e**). The samples that contained ratios of 1:3 and 1:4 **ST_HEAT** to **mCh_SC** contained a very large excess of unbound fluorescent protein, which made it very difficult to differentiate between bound and unbound fluorescent protein (**Figures 6.17 a-b**). Analysis of the 1:1 and 1:2 **ST_HEAT** to **mCh_SC** revealed that both samples appeared to have excess **mCh_SC** in the sample; however, the 1:2 nanotubes appeared to have a larger amount of **mCh_SC** bound (**Figures 6.17 c-d**). This could indicate that **mCh_SC** was bound to residual unassembled **ST_HEAT**, which could explain the unassembled peptide in the 1:1 ratio sample. Since the tubes looked more functionalized in the 2:1 sample, it was decided to stick with that ratio and determine methods for removal of excess **mCh_SC**.

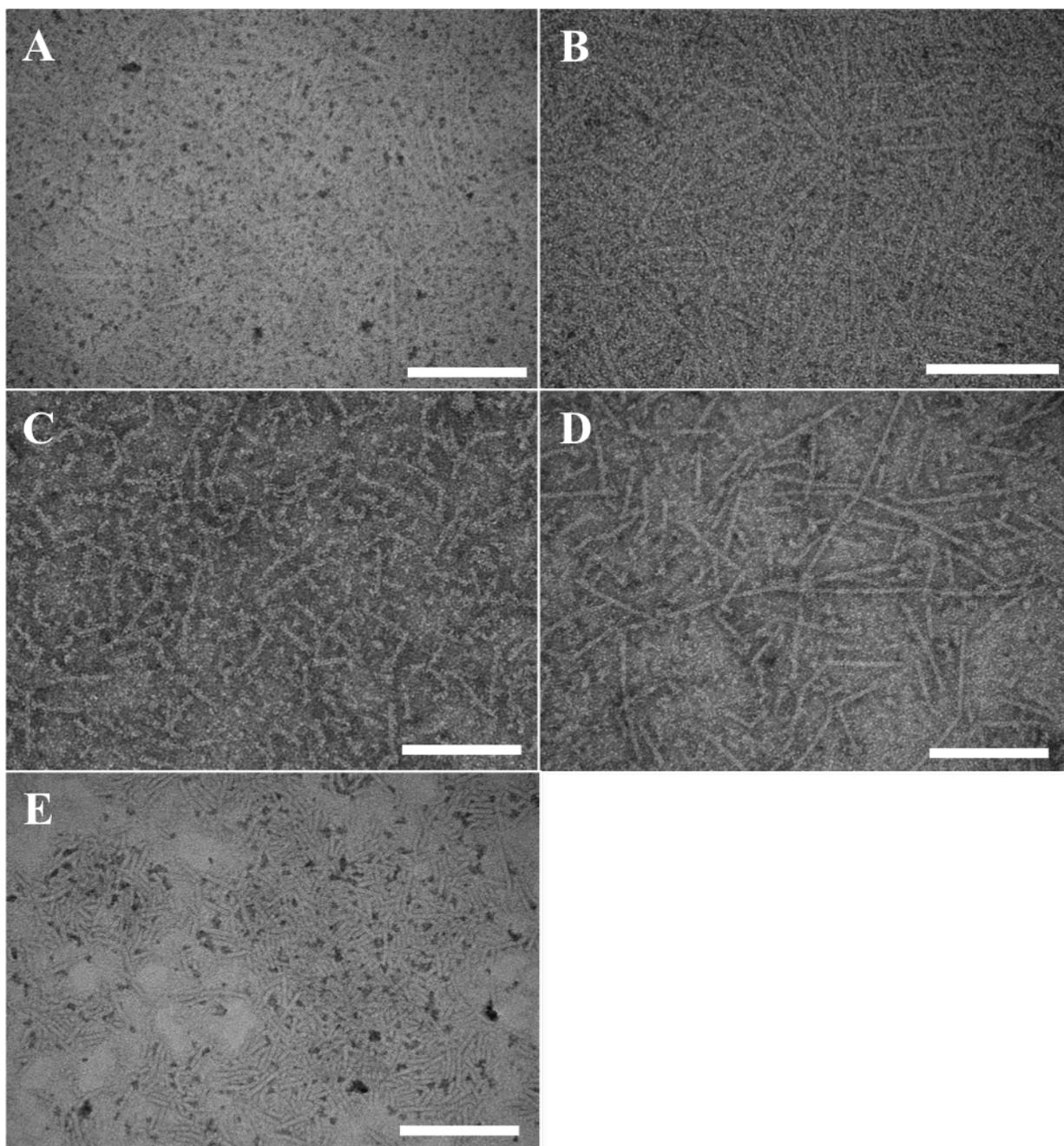


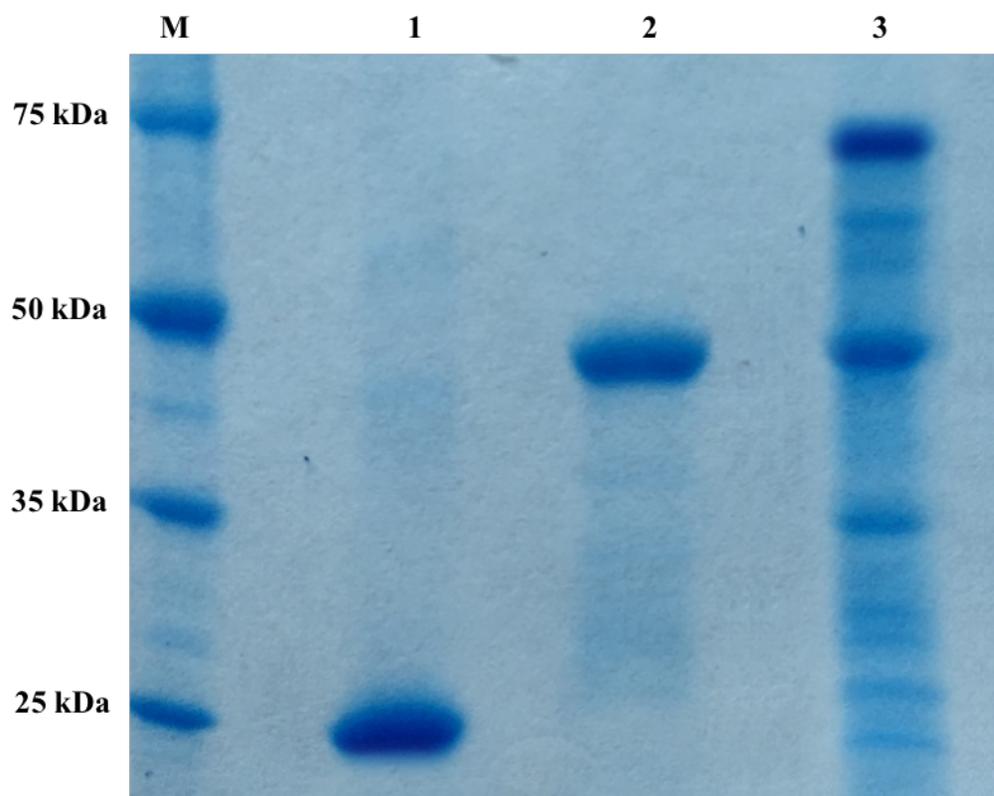
Figure 6.17 Transmission electron microscopy images of thermally annealed 80 μM HPLC purified ST_HEAT with varying concentrations of mCh_SC in 10 mM acetate buffer (pH5.0). (A) 1:4 ratio of ST_HEAT to mCh_SC. (B) 1:3 ratio of ST_HEAT to mCh_SC. (C) 1:2 ratio of ST_HEAT to mCh_SC. (D) 1:1 ratio of ST_HEAT to mCh_SC. (E) 2:1 ratio of ST_HEAT to mCh_SC. Samples were diluted to 5 μM for analysis. Scale bar = 200 nm.

Ultracentrifugation for Removal of Excess mCh_SC

Preparative ultracentrifugation was utilized to remove excess **mCh_SC** from functionalized nanotube samples. The goal was to pellet the functionalized nanotubes and any excess unbound fluorescent protein would remain in the supernatant. This method was tested on an 80 μM thermally annealed HPLC purified **ST_HEAT** sample linked with 13.8% (v/v) **mCh_SC** at a molar ratio of 1:2, respectively. An 80 μM thermally annealed HPLC purified **ST_HEAT** sample and a 13.8% (v/v) **mCh_SC** sample were prepared as controls. The sample volumes were adjusted up to 1 mL, which was the lowest volume allowed for the ultracentrifuge.

The samples were centrifuged using a Beckman Coulter Optima™ XE-90 Preparative Ultracentrifuge with a Type 45 Ti fixed-angle rotor. The configuration of the ultracentrifuge, rotor, tube adapters, and sample volume designated the maximum speed of rotation at 25,000 rpm. Initially, the samples were subjected to ultracentrifugation at 30,000 x g (16,100 rpm) at 4 °C for 2 hours. After carefully removing the liquid from the tubes, there did not appear to be pellets in any of the tubes. However, the ultracentrifuge tubes were allowed to shake with 10 mM acetate buffer (pH 5.0) for resuspension of any molecules that may have pelleted.

The supernatants of the ultracentrifuge samples were analyzed by SDS-PAGE on a 10% gel to verify the presence of the molecular species (**Figure 6.18**). Analysis of the supernatants indicated that the functionalized **mCh-HEAT** nanotubes were present in the supernatant after ultracentrifugation at 30,000 x g for 2 hours. Excess **mCh_SC** was also present in the supernatant of the ligation sample. The control samples of the **ST_HEAT** nanotubes and the **mCh_SC** fluorescent protein confirmed that the two species remained in the supernatant after ultracentrifugation.



M. Perfect protein marker (10-225 kDa)

1. 80 μ M HPLC purified ST_HEAT control, ultracentrifuge supernatant

2. 13.8% (v/v) mCh_SC control, ultracentrifuge supernatant

3. mCh-HEAT, annealed ST_HEAT + mCh_SC, ultracentrifuge supernatant

Figure 6.18 SDS-PAGE analysis of supernatants recovered after ultracentrifugation (30,000 x g) to separate mCh_SC from the nanotubes. Lane 1 and 2 depict the thermally annealed 80 μ M HPLC purified ST_HEAT control and the 13.8% (v/v) mCh_SC control, respectively. Lane 3 represents the sample of ST_HEAT nanotubes mixed with mCh_SC. ST_HEAT, mCh_SC, and mCh-HEAT are present in the supernatant after centrifugation. MW: ST_HEAT = 30.1 kDa; mCh_SC = 42.4 kDa; mCh-HEAT = 72.4 kDa.

Transmission electron microscopy was used to analyze the ultracentrifugation samples. TEM analysis was conducted on the supernatant sample for the **mCh-HEAT** functionalized nanotubes in addition to the potentially resuspended ultracentrifugation pellet. The supernatant for the **ST_HEAT** nanotubes was visualized by TEM as a control. The linked **mCh-HEAT** supernatant displayed visibly functionalized tubes in addition to unbound peptide, which was assumed to be excess unbound **mCh_SC** (**Figure 6.19 a**). The pellet of the linked **mCh-HEAT** sample only exhibited disorganized or unassembled peptide with no nanotubes (**Figure 6.19 b**). The TEM image of the supernatant for the **ST_HEAT** control showed nanotubes comparable to the tubes previously imaged. Based on this data, in addition to the information gathered from SDS-PAGE analysis, the **mCh-HEAT**, **mCh_SC**, and **ST_HEAT** proteins remained in the supernatant after ultracentrifugation at 30,000 x g for 2 hours, and therefore, it failed to separate the excess **mCh_SC** from the **mCh-HEAT** nanotubes.

The ultracentrifuge supernatants containing **mCh_SC**, **ST_HEAT**, and **mCh-HEAT** were subjected to a second round of ultracentrifugation at 72,700 x g (25,000 rpm) at 4 °C for 2 hours. After the spin cycle, red coloring visibly grouped near the bottom of the tubes for the **mCh-HEAT** and **mCh_SC** samples, but no pellets formed. Aliquots were carefully removed from the center of the red fractions for SDS-PAGE analysis (**Figure 6.20**). The samples were returned to the ultracentrifuge for an additional 2-hour spin cycle at 72,700 x g (25,000 rpm). The second spin cycle did not change the appearance of the samples. There were still clusters of red near the bottom of the **mCh-HEAT** and **mCh_SC** samples and no visible pellet. The red and clear fractions of the supernatant were carefully removed and kept separate for analysis. The empty ultracentrifuge tubes were allowed to shake with 10 mM acetate to resuspend any pelleted molecules.

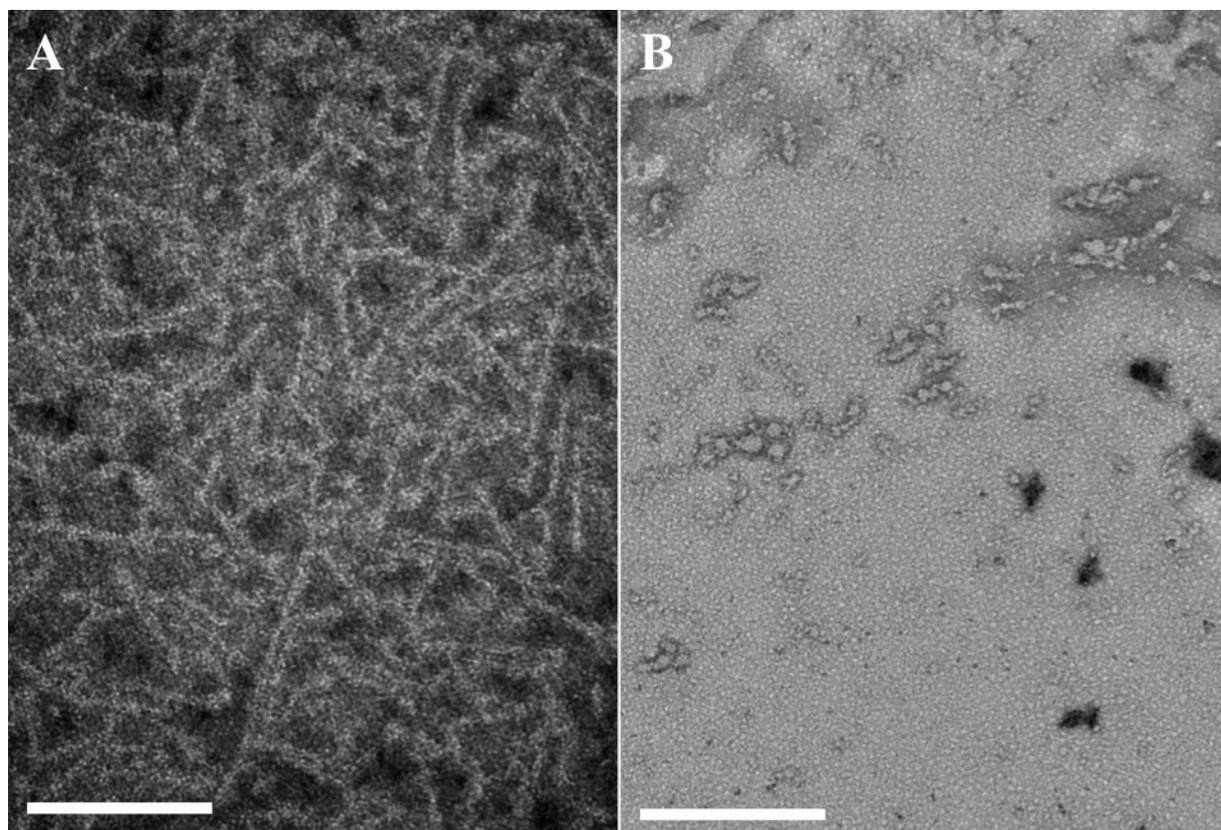
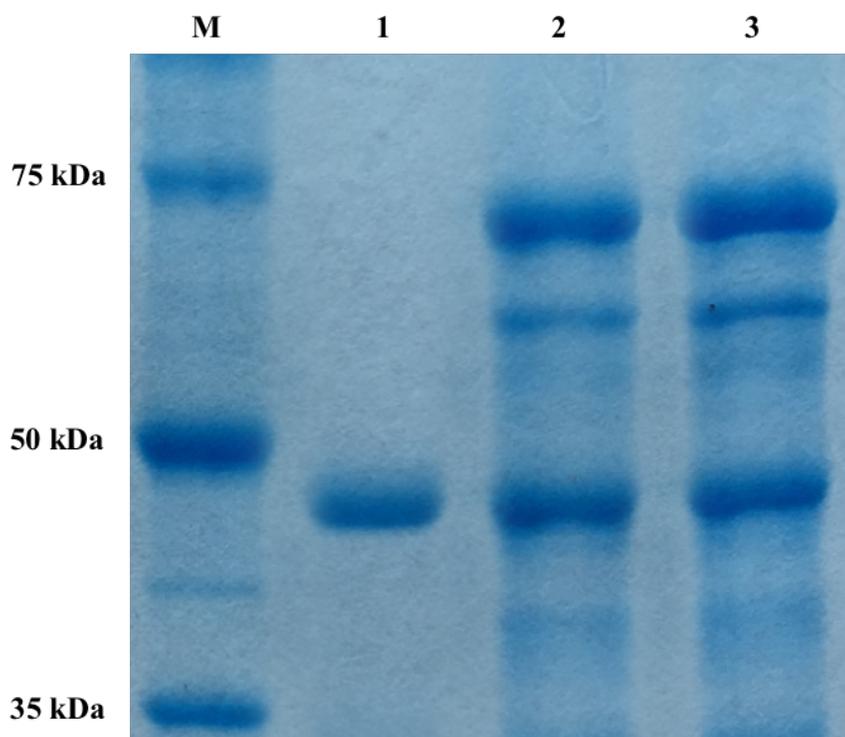


Figure 6.19 Transmission electron microscopy images of samples recovered after ultracentrifugation ($30,000 \times g$) to separate mCh_SC from the nanotubes. (A) Supernatant recovered from mCh-HEAT sample (thermally annealed $80 \mu\text{M}$ HPLC purified ST_HEAT with 13.8% (v/v) mCh_SC in 10 mM acetate buffer, pH5.0). (B) Pellet recovered from mCh-HEAT sample. Samples were diluted to $10 \mu\text{M}$ for analysis. Scale bar = 200 nm.

SDS-PAGE analysis on a 10% gel was used to analyze the composition of the red fraction of the ultracentrifugation supernatant. The **mCh_SC** sample was compared to the **mCh-HEAT** samples removed after 2 and 4 hours of ultracentrifugation. The **mCh-HEAT** samples showed the presence of both **mCh-HEAT** around 72 kDa and excess **mCh_SC** around 42 kDa, which indicated that ultracentrifugation at $72,700 \times g$ resulted in the migration of both proteins to the



M. Perfect protein marker (10-225 kDa)

1. 13.8% (v/v) mCh_SC control, ultracentrifuge supernatant

2. mCh-HEAT, annealed ST_HEAT + mCh_SC, ultracentrifuge supernatant, 2 hrs

3. mCh-HEAT, annealed ST_HEAT + mCh_SC, ultracentrifuge supernatant, 4 hrs

Figure 6.20 SDS-PAGE analysis of supernatants recovered after ultracentrifugation (72,700 x g) to separate mCh_SC from the nanotubes. Lane 1 represents the sample from the 13.8% (v/v) mCh_SC control. Lane 2 and 3 represent the samples from mCh-HEAT tube removed after 2 and 4 hours of ultracentrifugation, respectively. mCh_SC and mCh-HEAT are both present in the supernatant verifying that mCh_SC was not separated from mCh-HEAT nanotubes. MW: ST_HEAT = 30.1 kDa; mCh_SC = 42.4 kDa; mCh-HEAT = 72.4 kDa.

lower red fraction of the supernatant. The **mCh_SC** sample verified that the fluorescent protein separated into the red supernatant fraction.

The ultracentrifugation samples were analyzed by transmission electron microscopy to determine the presence of the **mCh-HEAT**, **ST_HEAT**, and **mCh_SC** proteins within the various fractions. TEM analysis of the red supernatant fraction removed after 4 hours of centrifugation from the **mCh-HEAT** sample confirmed the presence of the **mCh-HEAT** functionalized nanotubes (**Figure 6.21 a**). There was also a significant amount of unassembled peptide, which was assumed to be excess unbound **mCh_SC**. The presence of both proteins in this fraction implied that they were migrating at a similar rate and separation would be difficult. There was not a significant amount of tubes in the clear supernatant fraction and the resuspended pellet for the **mCh-HEAT** sample. Only a small amount of disorganized peptide assemblies or unaggregated peptide was observed. TEM analysis of a sample collected from the **ST_HEAT** tube near the bottom of the supernatant after 4 hours of centrifugation verified the presence of nanotubes (**Figure 6.21 b**).

SDS-PAGE and TEM analysis of the samples after ultracentrifugation indicated that the nanotubes failed to pellet when centrifuged at 72,700 x g (25,000 rpm) for 4 hours. As this was the maximum speed of rotation for the configuration of the instrument, it would be difficult to pellet the nanotubes. In addition, analysis of the supernatant fractions implied that all three proteins of interest were present in the lower half of the tube after the first 2 hours, although none of the proteins formed a pellet. Since both **mCh_SC** and **mCh-HEAT** migrated to a similar location in the sample, ultracentrifugation at 72,700 x g failed to remove excess unbound **mCh_SC** from ligated samples. It was decided that different avenues would be investigated for the removal of excess unbound **mCh_SC** protein.

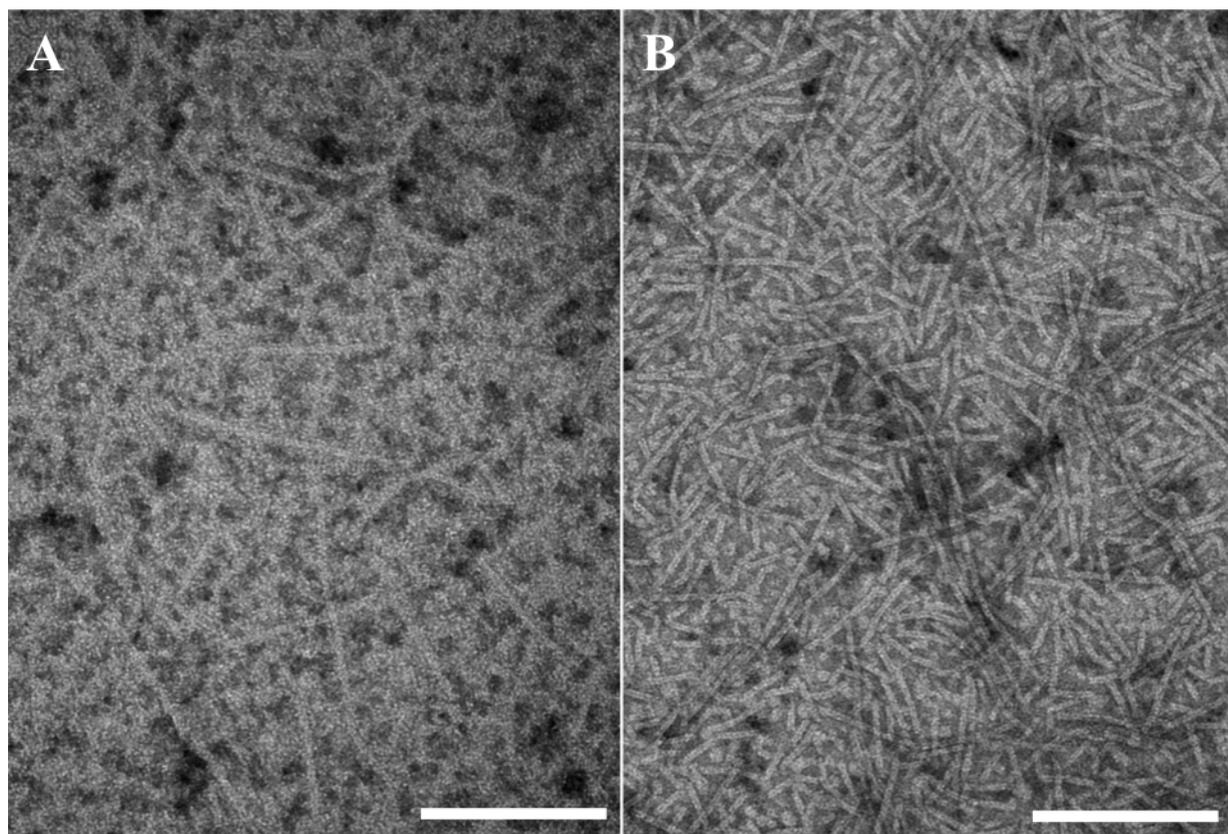


Figure 6.21 Transmission electron microscopy images of samples recovered after ultracentrifugation ($72,700 \times g$) to separate mCh_SC from the nanotubes. (A) Supernatant recovered from mCh-HEAT sample (thermally annealed $80 \mu\text{M}$ HPLC purified ST_HEAT with 13.8% (v/v) mCh_SC in 10 mM acetate buffer, pH5.0). (B) Supernatant recovered from ST_HEAT sample. Samples were diluted to $10 \mu\text{M}$ for analysis. Scale bar = 200 nm.

Centrifugal Filtration for Removal of Excess mCh_SC

Since ultracentrifugation failed to separate the functionalized nanotubes from the excess fluorescent protein, centrifugal filters with a defined nominal molecular weight limit (NMWL) were utilized. Samples of 5 μ M thermally annealed HPLC purified **ST_HEAT** were linked with 2% (v/v) **mCh_SC** for 18 hours. A 5 μ M thermally annealed HPLC purified **ST_HEAT** sample and a 2% (v/v) **mCh_SC** sample were prepared as controls. Initially, samples were processed using centrifugal filters with a 50 kDa NMWL. Any unbound **mCh_SC** protein, which had a molecular weight of 42 kDa, would be permitted to move through the filter, while formed nanotubes would remain in the upper filter unit. Samples were washed until no red coloring remained in the upper filter unit of the **mCh_SC** control sample. The **mCh-HEAT** sample retained a significant amount of red coloring after filtration.

TEM analysis was used to analyze the **mCh-HEAT** and **ST_HEAT** samples before and after filtration. Dilutions of the 5 μ M samples were prepared for improved visualization of functionalized nanotubes. The **ST_HEAT** samples appeared normal before and after filtration, indicating that the HEAT-based nanotubes were not being destroyed by the filtration process. Although it appeared as though there was a decrease in the background of the **mCh-HEAT** samples after filtration, there was still a significant amount of unassembled peptide (**Figures 6.22**).

Unlinked **mCh_SC** was removed; however, the **mCh_SC** that linked with unassembled **ST_HEAT** would not be able to pass through the filter. Therefore, centrifugal filters with a NMWL of 100 kDa were used to allow the passage of linked, unassembled **mCh-HEAT**, which had a molecular weight of 72 kDa. In addition, thermally annealed **ST_HEAT** samples were filtered prior to ligation with **mCh_SC** to remove unassembled **ST_HEAT** and subsequently

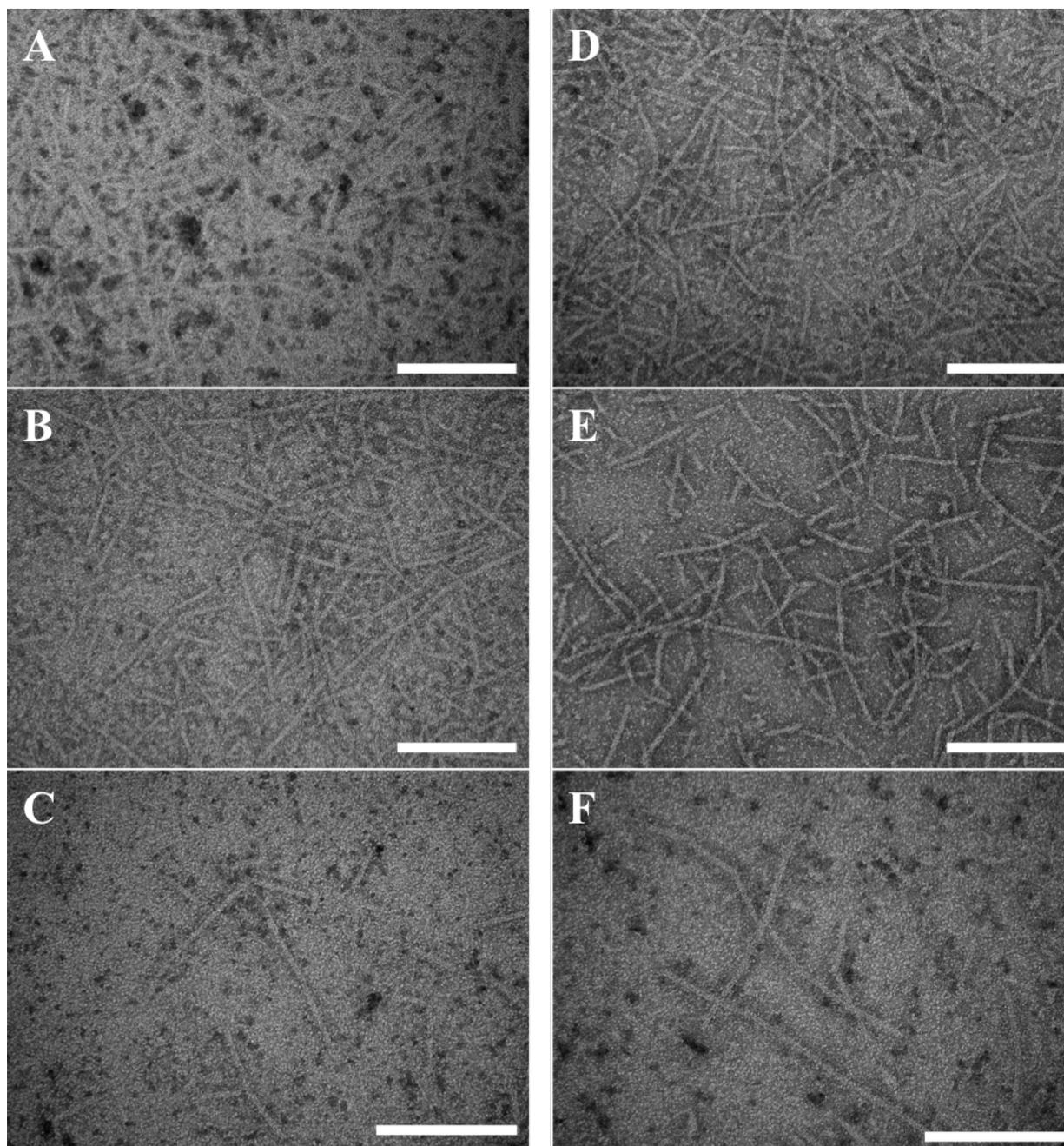


Figure 6.22 Transmission electron microscopy images of (A) a solution of thermally annealed 5 μM HPLC purified ST_HEAT with 5% (v/v) mCh_SC in 10 mM acetate buffer (pH5.0) with dilutions to 1 μM (B) and 0.2 μM (C) for analysis. (D) Images of the previous solution after a centrifuge filter (50 kDa NMWL) was utilized for mCh_SC removal with dilutions to 1 μM (E) and 0.2 μM (F) for analysis. Scale bar = 200 nm.

decrease the amount of unassembled **mCh-HEAT**. Thermally annealed **ST_HEAT** samples were analyzed after filtration, and the TEM images showed a very clean sample of nanotubes. Only a small amount of unassembled protein was observed (**Figure 6.23 a-c**). The linked and filtered **mCh-HEAT** samples showed a significantly lower amount of background compared to the samples filtered with 50 kDa NMWL filters (**Figure 6.23 d-f**). The nanotubes displayed **mCh_SC** on the exterior surfaces resulting in a fuzzy nature.

The linked **mCh-HEAT** samples were also analyzed using SDS-PAGE on a 10% gel (**Figure 6.24**). Before the **mCh-HEAT** sample was subjected to filtration, a large quantity of residual **mCh_SC** protein remained in the solution. On the other hand, after filtration, only a very weak band was present around 42 kDa, indicating that most of the **mCh_SC** was successfully removed from the solution. Additionally, the small decrease in size of the **mCh-HEAT** band at 72 kDa was due to the removal of unassembled **mCh-HEAT**. The centrifugal filters with a 100 kDa NMWL facilitated the removal of a significant amount of excess **mCh_SC** from the samples by supporting the passage of **mCh_SC** and unassembled **mCh-HEAT** through the filter. Furthermore, reacquired samples from the filtration process displayed fluorescence indicating the **mCh_SC** remained attached to the nanotubes.

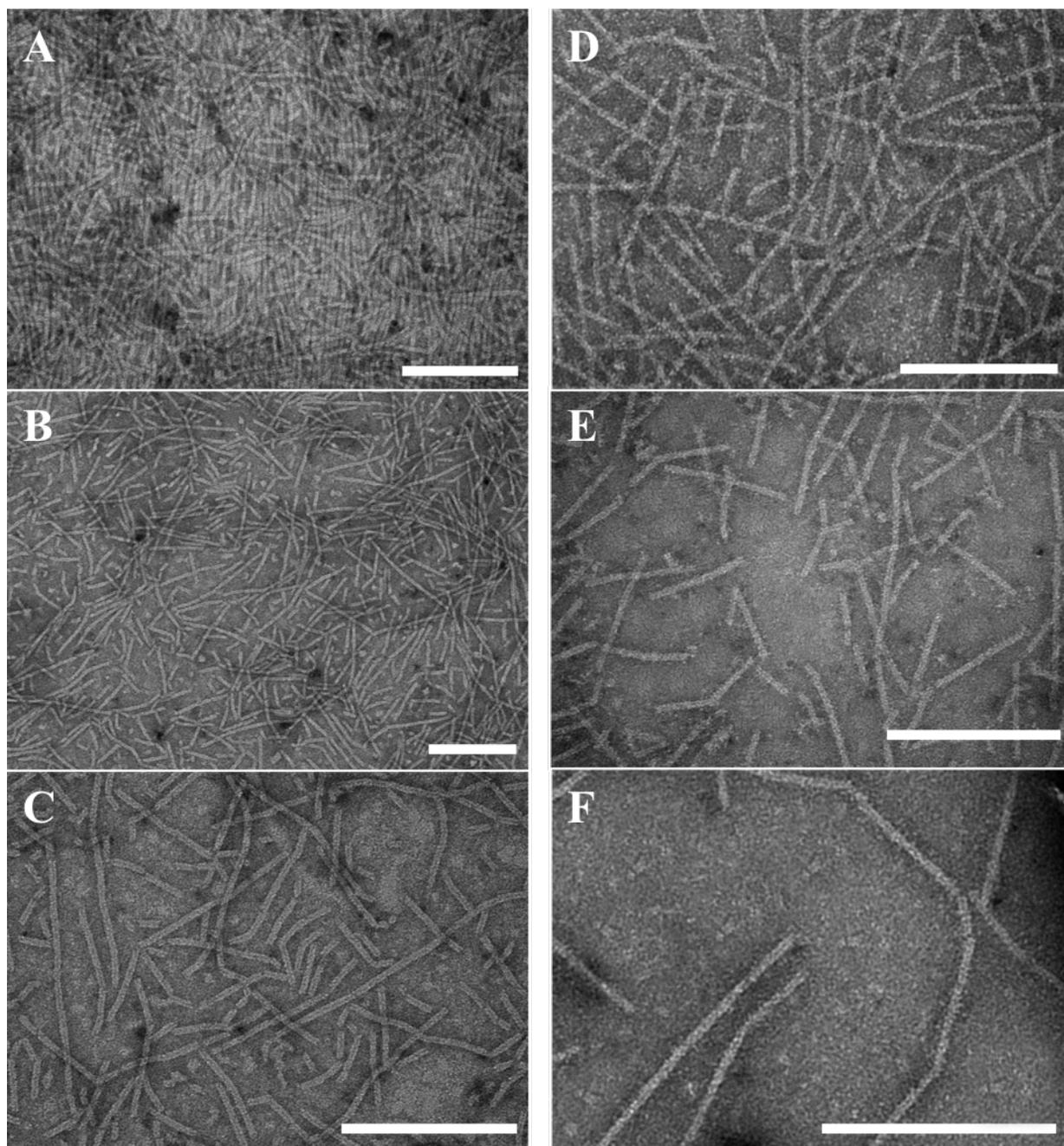
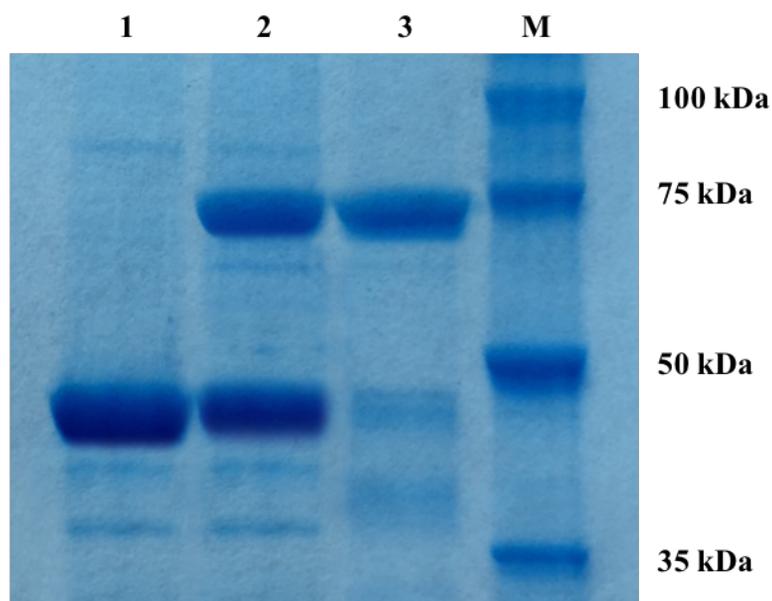


Figure 6.23 Transmission electron microscopy images of (A) a filtered (100 kDa NMWL) solution of thermally annealed 5 μM HPLC purified ST_HEAT, with dilutions to 1 μM (B) and 0.2 μM (C), and (D) images of the previous solution after linked with 2% (v/v) mCh_SC and filtered (100 kDa NMWL) and diluted to 1 μM , 0.2 μM (E), and 0.05 μM (F) for analysis. Scale bar = 200 nm.



1. mCh_SC control

2. mCh-HEAT, annealed ST_HEAT + mCh_SC

3. mCh-HEAT, annealed ST_HEAT + mCh_SC, filtered (100 kDa NMWL)

M. Perfect protein marker (10-225 kDa)

Figure 6.24 SDS-PAGE analysis indicating the result of filtering an mCh_HEAT sample with a centrifuge filter (100 kDa NMWL) to remove excess mCh_SC. Lane 1 represents the mCh_SC control. Lane 2 depicts a sample of mCh-HEAT (5 μ M HPLC purified ST_HEAT + 2% (v/v) mCh_SC). Lane 3 displays the filtered mCh-HEAT sample, where the majority of the excess mCh_SC has been removed. MW: mCh_SC = 42.4 kDa; mCh-HEAT = 72.4 kDa.

Fluorimetry Analysis

Once excess unbound **mCh_SC** was removed, fluorimetry was utilized to gather information about the fluorescence of the functionalized nanotubes. Thermally annealed samples of 5 μM HPLC purified **ST_HEAT** were prepared and filtered. The samples were then linked with 2 % (v/v) **mCh_SC** and subjected to a second round of filtration. Fluorescence spectroscopy was performed on the samples using a Fluoromax-3 fluorimeter with a xenon lamp. A 5 μM thermally annealed **ST_HEAT** sample was prepared as a negative control, and a sample containing 2% (v/v) **mCh_SC**, which corresponds 20 μM **mCh_SC**, was prepared as a positive control.

Fluorescence intensities of the samples were measured across an emission wavelength scan of 550 to 750 nm at a fixed excitation wavelength of 587 nm. The resulting fluorescence emission spectrum for the **mCh-HEAT** sample displayed a maximum intensity of 6.018 million arbitrary units at 610 nm (**Figure 6.25**), which matches the emission wavelength for the mCherry protein. The emission spectrum for the 2% (v/v) **mCh_SC** solution produced a similar yet stronger emission spectrum, which was expected due to the higher concentration of mCherry fusion protein in the sample. The **ST_HEAT** negative control sample produced an almost flat emission spectrum with the exception of a small peak at 587 nm. The fluorimeter detector was located at a 90° angle to the excitation light in order to decrease the quantity of transmitted incident light that was detected; however, excitation light was scattered by particles in the sample, which resulted in the presence of a peak at the excitation wavelength.

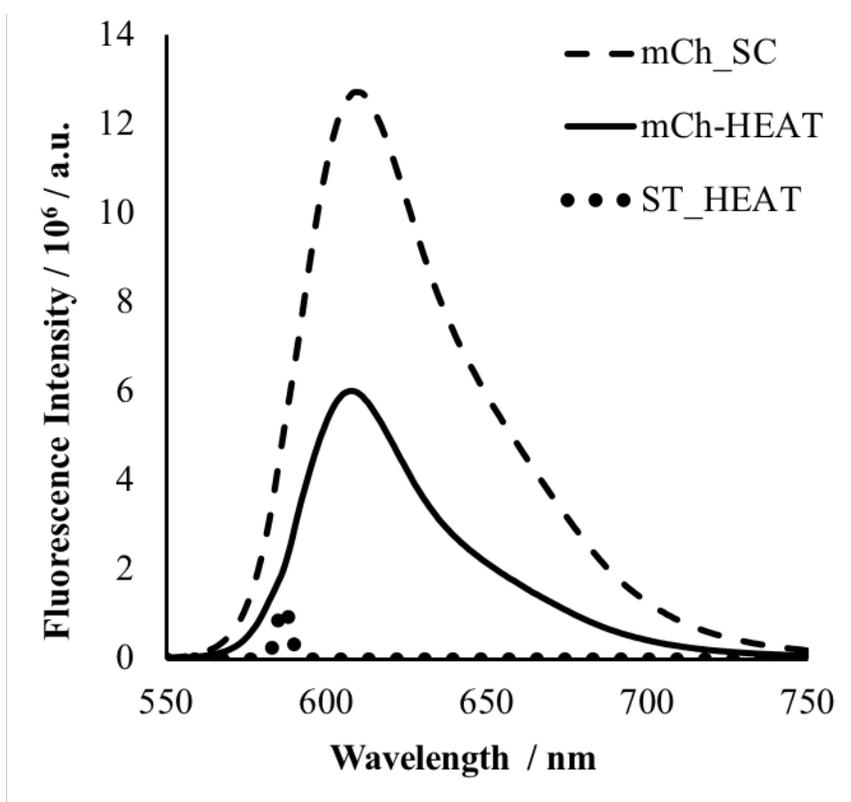


Figure 6.25 Fluorimetry data for mCh_HEAT formed from a 5 μM sample of thermally annealed HPLC purified ST_HEAT linked with 2% (v/v) mCh_SC. A 5 μM sample of thermally annealed ST_HEAT and a sample of 2% (v/v) mCh_SC were used as controls. The ST_HEAT negative control was subtracted from mCh-HEAT sample.

As described previously, 5 μM thermally annealed ST_HEAT samples were linked with 2% (v/v) mCh_SC samples to ensure maximum functionalization. Although a 2% (v/v) sample of mCh_SC was used as a positive control, this represented 20 μM of mCh_SC. If every ST_HEAT molecule in the sample assembled into nanotube structures and every SpyTag ligated with a SpyCatcher on mCh_SC, then the maximum amount of mCh-HEAT in the sample after filtration would equal 5 μM . However, unassembled ST_HEAT was removed by filtration after thermal

annealing, which was one factor that would decrease the total potential quantity of **mCh-HEAT** in the final sample. Therefore, it was expected that less than 5 μM of **mCh-HEAT** would be present in the sample during analysis.

To assess the amount of mCherry present in the **mCh-HEAT** samples, **mCh_SC** samples were prepared at various concentrations to develop a standard curve. Samples were prepared with 2%, 1.5%, 1%, 0.8%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, and 0.1% (v/v) **mCh_SC**, which correspond to 20 μM , 15 μM , 10 μM , 8 μM , 6 μM , 5 μM , 4 μM , 3 μM , 2 μM , and 1 μM **mCh_SC** protein. Fluorescence spectra were produced for all samples within an emission wavelength range of 550 to 750 nm at an excitation wavelength of 587 nm (**Figure 6.26 a**). All spectra exhibited a maximum at 610 nm. A second smaller peak appeared around 587 nm in the samples with lower concentration, which was due to scattered light passing through the sample to the detector with more ease in less concentrated samples. The maximum fluorescence intensities for all **mCh_SC** samples were plotted against the **mCh_SC** concentration to form a standard curve (**Figure 6.26 b**). A trendline equation for the linear segment of the curve determined that the approximate quantity of mCherry present in the **mCh-HEAT** sample was 4.44 μM . The expected fluorescence intensity for a sample containing 5 μM of **mCh_SC** was 6.782 million arbitrary units, while the **mCh-HEAT** sample produced a fluorescence intensity of 6.018 million arbitrary units.

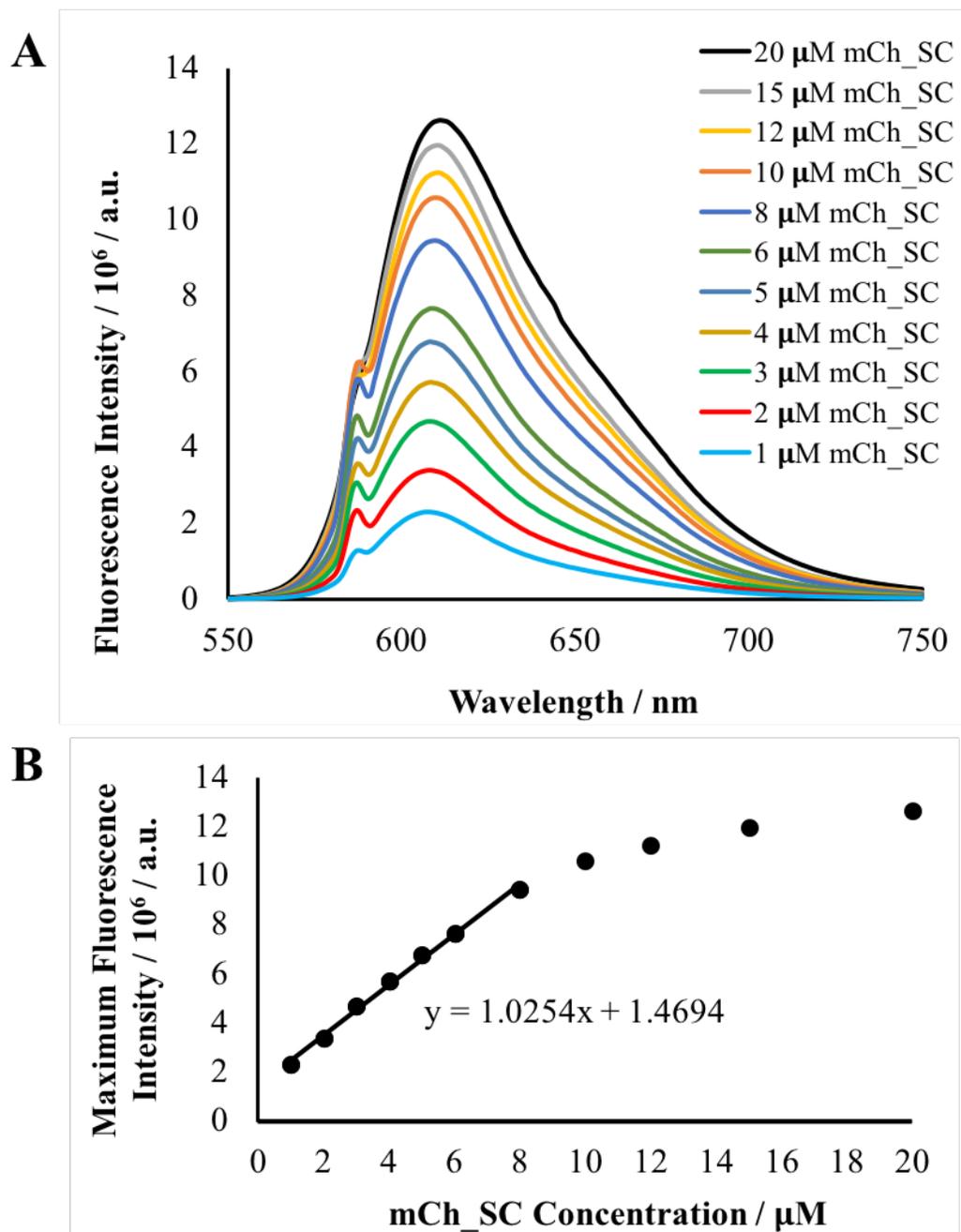


Figure 6.26 (A) Fluorimetry data for various concentrations of mCh_{SC} in 10 mM acetate pH 5.0. (B) Standard curve of maximum fluorescence intensities produced by various concentrations of mCh_{SC} in 10 mM acetate pH 5.0. Samples contained 1 μM , 2 μM , 3 μM , 4 μM , 5 μM , 6 μM , 8 μM , 10 μM , 15 μM , and 20 μM mCh_{SC}. A trendline and the corresponding trendline equation were produced for the linear segment of the data.

Correlative Light Electron Microscopy (CLEM)

Correlative light electron microscopy allows the dual examination of a sample using two typically distinct microscopy platforms. Fluorescence microscopy (FM) allows for wide field imaging of whole systems, while electron cryomicroscopy (cryo-EM) provides a high-resolution analysis of the assemblies of interest. The two microscopy techniques provide two complementary yet unique sets of data that provide a significant amount of information about a system. An integrated CLEM system can be utilized to subject a sample to an optical light path and an electron beam at the same time to automatically generate an overlay of the two images.¹¹⁻¹² Images from the two techniques can also be collected sequentially for a given region and an overlay can be performed at a later time. The complementary data from this technique would allow for imaging of the **mCh-HEAT** nanotubes with verification of **mCh_SC** functionalization.

Correlative light electron microscopy was applied to thermally annealed samples of 5 μM **ST_HEAT** that were functionalized with 2% (v/v) **mCh_SC** and filtered as previously described to remove excess **mCh_SC** protein. TEM and fluorimetry analysis were conducted on the sample to verify the presence of nanotubes and fluorescence prior to CLEM analysis. CLEM analysis was conducted by the Emory University Robert P. Apkarian Integrated Electron Microscopy Core service instructor, Spencer Hughes. Cryo-EM images shown in **Figure 6.27** successfully displayed **ST_HEAT** nanotubes functionalized on the surface with **mCh_SC** fluorescent protein. Although **mCh_SC** fluorescence was visualized under the secondary fluorescence microscope, an overlay of the two images was not possible due to a sample loading error. Further CLEM studies will be conducted to acquire an overlay of the two images.

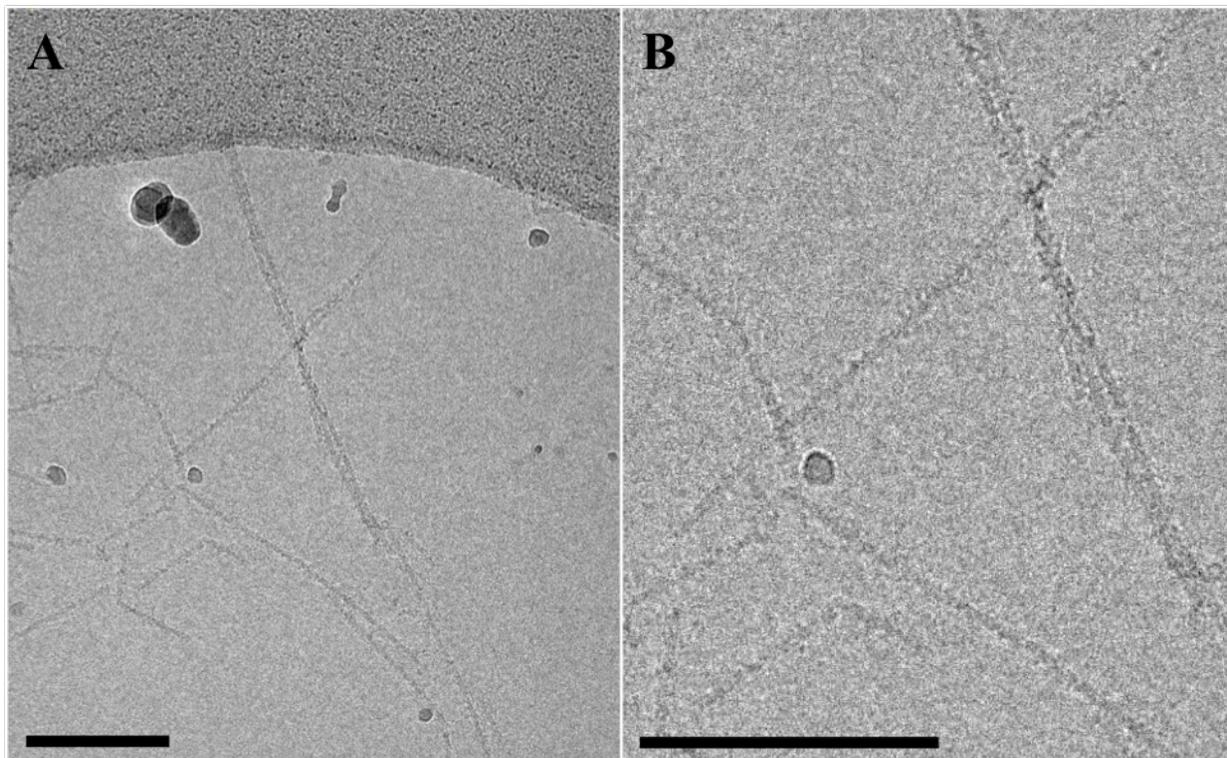


Figure 6.27 Electron cryo-microscopy images of (A) a thermally annealed solution of 5 μM ST_HEAT linked with 2% (v/v) mCh_SC (filtered). (B) A zoomed in region to highlight the presence of mCh_SC on the ST_HEAT nanotubes. Scale bar = 200 nm.

6.3 Conclusion

Optimization of assembly conditions for the HPLC purified **ST_HEAT** peptide resulted in the formation of high aspect ratio nanotubes with a diameter of 7.71 nm and a distinct inner lumen. The nanotubes displayed a strong alpha helical signature, and thermal denaturation studies failed to unfold the nanotubes. The stable **ST_HEAT** nanotubes provide a promising substrate for the asymmetric functionalization of the two surfaces due to a hypervariable concave channel and the presence of the SpyTag peptide tag on the convex surface.

Ligation studies resulted in a protocol that allowed for the functionalization of the **ST_HEAT** nanotubes with the **mCh_SC** fluorescent protein. Further analysis devised a method for the removal of excess **mCh_SC** protein in addition to unassembled **ST_HEAT** and **mCh-HEAT**. The absence of these individual molecules allows for an accurate analysis of fully formed functionalized nanotubes. Although correlative light electron microscopy could not be obtained due to a sample loading error, electron cryomicroscopy resulted in images that displayed the presence of nanoscale projections on the surface of the nanotubes, which was presumed to be the **mCh_SC** fluorescence protein. Additionally, after the nanotubes went through multiple washes in a centrifuge filter to remove all excess **mCh_SC**, the sample continued to display a high level of fluorescence. The absence of excess **mCh_SC** in the sample was verified by SDS-PAGE analysis. Finally, repetitive wash cycles in the centrifugal filters failed to shear the nanotubes, which indicates a high level of mechanical stability. These features advocate the production of highly stable functional nanotubes.

The use of the SpyTag/SpyCatcher ligation system allows for the additional functionalization of the convex surface of the optimized **ST_HEAT** nanotubes with novel arbitrary

proteins fused to the SpyCatcher ligation partner. This rapid reaction can specifically form the irreversible covalent isopeptide bond in highly complex solutions including a wide pH range and at extreme temperatures. This system prevents the need to optimize HEAT-based nanosheets each time a novel functional protein is fused to the surface. Only one HEAT-based fusion product, **ST_HEAT**, is required to support the attachment of a potentially limitless range functional proteins.

Functional differentiation between the two interfaces of a barrier material supports the separation and transportation of information, energy, and material in the form of small molecules, such as ions, and building blocks or macromolecules, such as DNA and protein. Biologically-inspired asymmetric functionalization of structurally defined nanomaterials would support the construction of artificial barrier materials and compartments, which would permit applications that require directional processes or physical separation such as electron transfer, energy transduction, and small-molecule activation, encapsulation, and transport.

The optimized protein-based **ST_HEAT** nanotubes are a stable material that authorizes the functional differentiation of the two surfaces in a controllable manner. Variation of the residues on the concave surface of the nanotubes can direct the functional properties and result in the encapsulation and transport of molecules. Conticello and coworkers are currently investigating the potential of a **HEAT_R1** variant displaying arginine residues in multiple hypervariable positions for the capture of DNA within the pore. Meanwhile, the exterior surface of the **ST_HEAT** nanotubes can be coated with any functional protein via the SpyTag/SpyCatcher ligation system to produce two functionally distinct surfaces. Additionally, the SpyTag/Catcher system was designed to support the attachment of various proteins to the external surface concurrently resulting in multifunctional nanotubes.

Recently, significant interest has expanded the utilization of biomaterials to modulate immune responses with a shift from the development of biocompatible ‘immune-evasive’ biomaterials towards the fabrication of smart ‘immune-interactive’ biomaterials in order to exploit the power of the immune system and direct the inflammatory response towards healing and regeneration.¹³ These immunomodulatory biomaterials include a wide range of potential therapeutic applications such as altering immune reactions to biomedical implants to promote tissue regeneration over fibrosis or eliciting a desired immune response against non-infectious diseases and conditions induced by biomaterial-based vaccines.¹⁴

Historically, vaccines were typically developed to combat infectious diseases;¹⁵⁻¹⁶ however, immunologically active biomaterials have recently been investigated to expand protection against a wide range of non-infectious diseases and conditions including cancer,¹⁷⁻²⁰ inflammation,²¹⁻²⁵ autoimmunity,²⁶⁻²⁹ wounds,³⁰ and others.³¹⁻³² Among other methods, the immune system can be governed through the controlled display of T cell and B cell epitopes, which can regulate the duration and magnitude of immune responses and antibody titers. Collier and coworkers have reported the abilities of self-adjuvating nanofiber vaccines for the elucidation of B cell responses to displayed peptide epitopes while avoiding the undesirable inflammatory response.³³⁻³⁴ These **ST_HEAT** nanotubes have potential for expanding this novel field through the encapsulation and controlled release of small molecules responsible for immune regulation within the inner channel in addition to the display of various epitopes on the external surface.

6.4 Experimental

6.4.1 Materials

All chemical reagents were purchased from either Thermo Fisher Scientific, Inc. (Waltham, MA) or MilliporeSigma (Burlington, MA) unless otherwise noted. Synthesis of plasmids containing codon-optimized genes encoding mCherry_SpyCatcher and SpyTag_HEAT8R was conducted by ATUM, formerly known as DNA2.0 (Newark, CA). The BL21 (DE3) chemically competent *E. coli* strain were purchased from New England Biolabs, Inc. (Ipswich, MA). Luria-Bertani broth and agar powder were purchased from MilliporeSigma (Burlington, MA), and Terrific Broth (TB) powder was purchased from Thermo Fisher Scientific, Inc. (Waltham, MA). Kanamycin monosulfate was purchased from VWR International, LLC (Suwanee, GA). Isopropyl- β -D-thiogalactopyranoside (IPTG) was purchased from Research Products International Corp. (Prospect, IL). The Perfect Protein Marker (10-225 kDa) was purchased from MilliporeSigma (Burlington, MA).

Benzonase[®] nuclease and protease-inhibitor cocktail (EDTA-free) were purchased from MilliporeSigma (Burlington, MA). Lysozyme from chicken egg white was purchased from Research Products International Corp. (Prospect, IL). HisPur[™] cobalt resin was obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). TEV Protease was purchased from MilliporeSigma (Burlington, MA). SnakeSkin[™] Dialysis Tubing with a 10 kDa MWCO and Slide-A-Lyzer[™] G2 Dialysis Cassettes with a 20 kDa MWCO were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA). Amicon[®] ultra centrifugal filters (10 kDa NMWL, 30 kDa NMWL, 50 kDa NMWL, and 100 kDa NMWL) for sample volumes of 15 mL and 0.5 mL were purchased from MilliporeSigma (Burlington, MA).

Methylamine tungstate (MAT) was purchased from Ted Pella, Inc. (Redding, CA). The 200-mesh carbon-coated copper grids were obtained from Electron Microscopy Sciences (Hatfield, PA). Vacuum filters with a 0.2 μm polyethersulfone (PES) membrane were obtained from Thermo Fisher Scientific, Inc. (Waltham, MA).

6.4.2 General Methods

Basic molecular biology procedures were adapted from a standard molecular cloning manual³⁵ or the protocol supplied by the manufacturer unless otherwise described in detail. All Reagents intended for use with bacteria, DNA, or recombinant proteins were sterilized by either syringe filtration through a 0.2 μm cellulose membrane, vacuum filtration through a 0.2 μm polyethersulfone (PES) membrane, or by autoclaving. All enzyme reactions were conducted in the reagent's buffers provided by the manufacturer unless otherwise noted.

All proteins were purified using immobilized metal affinity chromatography (IMAC) on HisPurTM cobalt resin followed by dialysis to remove imidazole. The **mCh_SC** protein was cleaved by TEV protease, concentrated using Amicon[®] Ultra-15 centrifugal filters (30 kDa NMWL), and dialyzed using a Slide-A-LyzerTM dialysis cassette (20,000 MWCO) to 10 mM acetate pH 5.0. Additional purification of the **ST_HEAT** protein was achieved by reverse-phase high performance liquid chromatography (HPLC).

Protein electrophoresis was conducted using 10, 12, or 16% SDS polyacrylamide gels on a Mini-PROTEAN 3 cell electrophoresis system from Bio-Rad Laboratories, Inc. (Hercules, CA). The buffer tank was filled with SDS run buffer (25 mM tris, 250 mM glycine, 0.1% SDS, pH 8.3). The Perfect Protein Marker was used as a protein standard for SDS-PAGE analysis. An initial sample volume of 10 μL was added to each well with any necessary concentration adjustments

made thereafter. Gels were run at 150 V for 1-1.5 hours depending on the desired level of separation. Gels were stained in coomassie overnight and then destained in a methanol-acetic acid buffer.

Assemblies were thermally annealed using a MJ MiniTM Gradient Thermal Cycler from Bio-Rad Laboratories, Inc. (Hercules, CA). All protein concentration measurements were determined using an UltroSpec 3000 UV/Visible spectrophotometer at 280 nm with a 1 cm cuvette. Fluorimetry measurements were collected with a 1 cm quartz cuvette using a Fluoromax-3 fluorimeter with a xenon lamp from HORIBA Jobin Yvon Inc. (Edison, NJ).

6.4.3 Assembly and Characterization of SpyTag_HEAT Nanotubes

Assembly of HisPurTM ST_HEAT

The HisPurTM ST_HEAT was dissolved in ddH₂O at a concentration of 6 mg/mL or 18 mg/mL. The solution was divided into the number of desired buffer conditions. An equal volume of 20 mM buffer was added to the tubes to make a 3 mg/mL or 9 mg/mL protein concentration in 10 mM buffer. The buffer conditions were 20 mM TAPS (pH 8.0), 20 mM MOPS (pH 7.0), 20 mM MES (pH 6.0), 20 mM acetate (pH 5.0), and 20 mM acetate, 1 mM EDTA (pH 5.0).

Reverse-Phase High Performance Liquid Chromatography (HPLC)

A fraction of the HisPurTM ST_HEAT lyophilized powder (52 mg; see previous chapter) was dissolved in 50% acetonitrile in water with 0.1% trifluoroacetic acid (TFA) with added help from vortexing. A Shimadzu HPLC was used to purify the ST_HEAT protein on a C18 column using a water-acetonitrile (+0.1% TFA) gradient. HPLC fractions were collected at the targeted time points. The collected HPLC fragments were transferred to a round bottom flask and a rotavap

was used to remove all acetonitrile from the solution and decrease the volume. The remaining solution containing the HPLC pure **ST_HEAT** was poured into 50 mL conical tubes and stored at -80 °C. Once the solution was frozen, the tube caps were replaced with Kimwipes™ and the tubes were lyophilized until all that remained was a fine white powder. The powder was weighed and stored at -30 °C.

Assembly of HPLC Purified ST_HEAT

Initially, the HPLC purified **ST_HEAT** assemblies were dissolved in 10 mM acetate buffer (pH 5.0) at a peptide concentration of 5 mg/mL followed by dilution to the desired peptide concentration (see concentration protocol). Eventually, the samples were assembled at a peptide concentration of 0.33 mg/mL in 10 mM acetate buffer (pH 5.0) to directly produce a 5 μM peptide solution. The protein concentration of the 5 μM assemblies was confirmed and adjusted if necessary (see concentration protocol).

Concentration

The protein concentrations of **ST_HEAT** assemblies (HisPur™ and HPLC purified) were determined by UV/Visible spectrophotometry. The samples were prepared by combining 30 μL of nascent protein sample with 270 μL of 6 M guanidine hydrochloride (GHCl) for a ratio of 1:9. Buffer samples were prepared by combining 30 μL of the appropriate buffer (10 mM) with 270 μL of 6 M GHCl. The protein and buffer samples were allowed to equilibrate for 30 minutes at room temperature. The absorbances were determined at 280 nm using 70 μL of sample per run. The appropriate buffer sample was used to zero the absorbance signal four times. Then the protein sample corresponding to the buffer blank was analyzed. The absorbance of the protein sample was

collected four times. The process was repeated for all protein samples with their appropriate buffer blank. The cuvette was thoroughly washed with water and acetone between samples.

An average of the four collected absorbances for each protein sample was used to calculate the protein concentration. Using the Beer-Lambert law equation, $A = \epsilon bc$, the absorbances (A) were converted into concentration (c). The extinction coefficient (ϵ) for **ST_HEAT** is $46800 \text{ M}^{-1} \text{ cm}^{-1}$ and the pathlength (b) across the cuvette was 1 cm. The concentrations were recorded in μM . When desired, the concentrations were adjusted by the addition of 10 mM buffer.

Thermal Annealing

Half of the assembled **ST_HEAT** (HisPurTM and HPLC purified) volume was transferred to PCR tubes (100 μL per tube) and thermally annealed using a programmable MJ MiniTM Gradient Thermal Cycler. The protocol heated the protein samples at 90 °C for 10 minutes, and then the samples were cooled at a rate of 0.2 °C/min to a final temperature of 25 °C. The annealed protein samples were stored at ambient temperature until analysis.

Circular Dichroism Spectropolarimetry (CD)

The HisPurTM **ST_HEAT** assemblies were analyzed at their 3 mg/mL or 9 mg/mL peptide assembly concentrations in 10 mM buffer. The buffer conditions were TAPS (pH 8.0), MOPS (pH 7.0), MES (pH 6.0), and acetate (pH 5.0). The HPLC purified **ST_HEAT** assemblies were analyzed at a protein concentration of 80 μM in 10 mM acetate buffer (pH 5.0). The samples were either nascent or thermally annealed. CD measurements were collected using a 0.10 mm quartz cell (Hellma Analytics) in a Jasco J-810 CD spectropolarimeter. Sample volumes of 30 μL were used to generate spectra from 260 to 190 nm at a resolution of 0.1 nm and a continuous scanning

rate of 50 nm/min. The CD spectra were accumulated in triplicate and averaged together. CD spectra were represented by the mean residue ellipticity (MRE) as a function of wavelength. The extinction coefficient for **ST_HEAT** is $46800 \text{ M}^{-1}\text{cm}^{-1}$.

CD melting experiments of thermally annealed HPLC purified **ST_HEAT** samples were performed at 80 μM in 10 mM acetate buffer (pH 5.0) by monitoring the signal intensity at 222 nm as the protein samples were heated at a rate of 40 $^{\circ}\text{C}/\text{hour}$ from 5 to 95 $^{\circ}\text{C}$. CD melting spectra were represented by the signal intensity as a function of temperature. The first derivatives of the melting spectra were used to acquire the melting temperatures. Standard CD spectra were collected before the melt at 5 $^{\circ}\text{C}$ and after the melt at 95 $^{\circ}\text{C}$.

Transmission Electron Microscopy (TEM)

The HisPurTM **ST_HEAT** assemblies were analyzed at their 3 mg/mL or 9 mg/mL peptide assembly concentrations in 10 mM buffer. The buffer conditions were TAPS (pH 8.0), MOPS (pH 7.0), MES (pH 6.0), acetate (pH 5.0), and acetate with 0.5 mM EDTA (pH 5.0). The HPLC purified **ST_HEAT** assemblies were analyzed at their 80 μM or 5 μM assembly concentrations in 10 mM acetate buffer (pH 5.0) and with or without **mCh_SC**. The protein samples were either nascent or thermally annealed. Protein assemblies were analyzed using 200 mesh carbon-coated copper grids (Electron Microscopy Sciences) with a 2% methylamine tungstate (MAT) solution as a stain. Samples were prepared by incubating 4 μL of a protein solution on a grid for 1 minute before wicking excess sample away using filter paper. MAT stain was immediately added to the grid and the solution was allowed to incubate for 1 minute. Stain was wicked away with filter paper and the samples were left to dry under vacuum until imaging. TEM images were obtained on a Hitachi 7700 TEM with accelerating voltages of 80 kV.

ImageJ Diameter Measurements

ImageJ software was used to measure the diameter of HisPur™ and HPLC purified ST_HEAT nanotubes in 10 mM acetate buffer (pH 5.0). Measurements were collected from TEM images of 77 μM HisPur™ ST_HEAT, 200 μM HisPur™ ST_HEAT, and 5 μM HPLC ST_HEAT to get a fair distribution among various assembly conditions. A sample size of 160 was gathered from different nanotubes. The diameters were recoded, grouped, and plotted on a histogram. An average and standard deviation were calculated from the data.

Small- and Wide-Angle X-Ray Scattering (SAXS/WAXS)

Synchrotron SAXS/WAXS data was acquired at the 12-ID-B beamline of Advanced Photon Sources at Argonne National Laboratory, using previously published methods.³⁶⁻³⁸ A solution of HPLC purified ST_HEAT was assembled in 10 mM acetate buffer (pH 5.0) to a peptide concentration of 5 mg/mL. The assembly solution was dialyzed against 10 mM acetate buffer (pH 5.0) for 2 days to remove trifluoroacetic acid remaining in the sample from HPLC purification. The dialyzed sample was thermally annealed and then incubated at room temperature for 2 days until SAXS analysis. A sample of the 10 mM acetate dialysis buffer (pH 5.0) was utilized as a buffer sample for background subtraction. SAXS/WAXS measurements were performed by subjecting each protein and buffer sample to 20 images at 25 °C using a quartz capillary flow cell (1.5 mm diameter) to prevent radiation damage. Utilizing a software package at beamline 12-ID-B, the 2-D scattering images were transformed through solid angle correction and azimuthal averaging into 1-D curves with normalization of the transmitted X-ray beam intensity. The 1-D curves of each sample were averaged, and the averaged background curve from the buffer sample was subtracted from the averaged peptide sample.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS)

MALDI-TOF experiments were performed on an Applied Biosystems® Voyager™ System 4700 mass spectrometer (Life Technologies Corporation; Carlsbad, CA) in the high mass positive ion linear mode. The matrix, sinapinic acid (3-(4-hydroxy-3, 5-dimethoxyphenyl)prop-2-enoic acid), was used at a concentration of 20 mg/mL in a mixture of 50% acetonitrile and 0.1% trifluoroacetic acid in deionized water. HisPur™ ST_HEAT protein solutions (1 mg/mL in a solution of 50% acetonitrile and 0.1% trifluoroacetic acid in deionized water) were mixed with the matrix solution in a 1:10 ratio. The ST_HEAT HPLC eluent (in a solution of 47.8% acetonitrile and 0.1% trifluoroacetic acid in deionized water) was mixed with the matrix solution in a 1:10 ratio. Two microliters of the protein and matrix mixtures were spotted on a stainless-steel sample plate and dried under vacuum or air. Sample preparation was repeated in triplicate. Mass spectra were acquired from 1000 laser shots at an intensity of 7900. Four MALDI spectra were collected for the ST_HEAT protein under various conditions and the molar masses were averaged together to produce an experimental average molar mass of 30072.6 Da. The theoretical molar mass of ST_HEAT was 30069.92 Da.

6.4.4 Functionalization of SpyTag_HEAT Nanotubes with mCherry_SpyCatcher

Ligation of Nascent HisPur™ ST_HEAT with mCh_SC

The HisPur™ ST_HEAT protein was dissolved in 2 M urea, 10 mM acetate buffer (pH 5.0) at a protein concentration of 3 mg/mL followed by dilution to the desired protein concentration of 80 μM. Refer to the concentration protocol above for determining and adjusting protein concentration.

A 300 μL volume of the nascent 80 μM HisPurTM **ST_HEAT** solution was combined with 33.33 μL of a 1 mM **mCh_SC** protein solution to acquire a 10% (v/v) **mCh_SC** final concentration. The solution contained 72 μM **ST_HEAT** and 100 μM **mCh_SC** at a ratio of 1:1.4. The solution was incubated at room temperature for 18 hours on a benchtop tube rotator with a rotisserie platform exhibiting a 360° rotation at 22 rpm. An aliquot of the linked sample was removed for TEM analysis (see below).

The solution was dialyzed using a Slide-A-LyzerTM dialysis cassette (20,000 MWCO) in a 4 L volume of dialysis buffer. The dialysis buffer urea concentration was dropped by 0.5 M in a step-wise manner every 4 hours until all urea was removed. The ligation solution continued to dialyze in 10 mM acetate buffer (pH 5.0) for 2 more cycles to ensure complete removal of urea from the solution. The ligation sample was removed from the dialysis cassette and transferred to a 1.5 mL microcentrifuge tube. An aliquot of the dialyzed linked sample was removed for TEM analysis (see below).

The dialyzed ligation sample of HisPurTM **ST_HEAT** with 10% **mCh_SC** was transferred to PCR tubes (100 μL per tube) and thermally annealed using the previously described protocol. The annealed protein samples were stored at ambient temperature. An aliquot of the sample that was linked before annealing was removed for TEM analysis (see below).

Ligation of Nascent HPLC Purified ST_HEAT with mCh_SC

The HPLC purified **ST_HEAT** protein was dissolved in 10 mM acetate buffer (pH 5.0) at a protein concentration of 5 mg/mL followed by dilution to the desired protein concentration of 80 μM . Refer to the concentration protocol above for determining and adjusting protein concentration.

A 200 μL volume of the nascent 80 μM HPLC purified **ST_HEAT** solution was combined with 32.02 μL of a 1 mM **mCh_SC** protein solution to acquire a 13.8 % (v/v) **mCh_SC** final concentration. The solution contained 69 μM **ST_HEAT** and 138 μM **mCh_SC** at a ratio of 1:2. The solution was incubated at room temperature without agitation until samples were collected for analysis. An aliquot of 15 μL was collected from the ligation sample after 18 hours of incubation and immediately prepared for SDS-PAGE analysis by mixing the 15 μL sample with 15 μL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 $^{\circ}\text{C}$ for 5 minutes. The sample was analyzed on a 10% SDS-PAGE gel.

The ligation sample was not thermally annealed. The sample continued to incubate at room temperature until analysis. Aliquots of the unannealed ligation sample were removed after 1 day of incubation and after 4 days of incubation for TEM analysis (see below).

Ligation of Thermally Annealed HisPurTM ST_HEAT with mCh_SC

The HisPurTM **ST_HEAT** protein was dissolved in 10 mM acetate buffer (pH 5.0) at a protein concentration of 3 mg/mL followed by dilution to the desired protein concentration of 80 μM . Refer to the concentration protocol above for determining and adjusting protein concentration. The sample of assembled 80 μM HisPurTM **ST_HEAT** was transferred to PCR tubes (100 μL per tube) and thermally annealed using the previously described protocol. An aliquot of the annealed **ST_HEAT** sample was removed for TEM analysis (see below).

A 300 μL volume of the annealed 80 μM HisPurTM **ST_HEAT** solution was combined with 33.33 μL of a 1 mM **mCh_SC** protein solution to acquire a 10% (v/v) **mCh_SC** final

concentration. The solution contained 72 μM **ST_HEAT** and 100 μM **mCh_SC** at a ratio of 1:1.4. The solution was incubated at room temperature for 18 hours on a benchtop tube rotator with a rotisserie platform exhibiting a 360 ° rotation at 22 rpm. An aliquot of the linked sample was removed for TEM analysis (see below).

Ligation of Thermally Annealed HPLC Purified ST_HEAT with mCh_SC

Initially, the HPLC purified **ST_HEAT** protein was dissolved in 10 mM acetate buffer (pH 5.0) at a protein concentration of 5 mg/mL followed by dilution to the desired protein concentration of 80 μM . Refer to the concentration protocol above for determining and adjusting protein concentration. Eventually, the HPLC purified **ST_HEAT** samples were assembled at a protein concentration of 0.33 mg/mL in 10 mM acetate buffer (pH 5.0) to directly produce a 5 μM protein solution or at a protein concentration of 2.0 mg/mL in 10 mM acetate buffer (pH 5.0) to directly produce a 30 μM protein solution. The protein concentration of the 5 μM and 30 μM assemblies was confirmed and adjusted if necessary (see concentration protocol). The samples of assembled HPLC purified **ST_HEAT** were transferred to PCR tubes (100 μL per tube) and thermally annealed using the previously described protocol. Aliquots of the annealed **ST_HEAT** samples were removed for TEM analysis (see below).

Volumes of the annealed HPLC purified **ST_HEAT** solutions were combined with appropriate volumes of 1 mM **mCh_SC** protein solution to acquire assemblies with varying ratios of **ST_HEAT** to **mCh_SC**. Annealed 80 μM **ST_HEAT** solutions were combined with 1 mM **mCh_SC** to produce solutions with 24.2% (v/v), 19.4% (v/v), 13.8% (v/v), 7.4% (v/v), and 3.9% (v/v) **mCh_SC**, which correlates to molar ratios of 1:4, 1:3, 1:2, 1:1, and 2:1 of **ST_HEAT** to **mCh_SC**. Annealed 30 μM **ST_HEAT** solutions were combined with 1 mM **mCh_SC** to produce

solutions with 5% (v/v) **mCh_SC**, which correlates to a molar ratio of 1:1.75 of **ST_HEAT** to **mCh_SC**. Annealed 5 μM **ST_HEAT** solutions were combined with 1 mM **mCh_SC** to produce solutions with 2% (v/v) and 5% (v/v) **mCh_SC**, which correlates to molar ratios of 1:4 and 1:10 of **ST_HEAT** to **mCh_SC**. The solutions were incubated at room temperature without agitation until samples were collected for analysis. Aliquots of the ligation sample were removed during incubation for TEM analysis (see below).

An aliquot of 15 μL was collected from the 80 μM **ST_HEAT** (HPLC; annealed) + 13.8% (v/v) **mCh_SC** ligation sample after 18 hours of incubation. The sample was immediately prepared for SDS-PAGE analysis by mixing the 15 μL sample with 15 μL of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 $^{\circ}\text{C}$ for 5 minutes.

Transmission Electron Microscopy (TEM)

The HisPurTM **ST_HEAT** samples were analyzed at their 80 μM concentration in 10 mM acetate buffer (pH 5.0). Some of the HisPurTM **ST_HEAT** samples were linked with 10% (v/v) **mCh_SC**. The 80 μM HPLC purified **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0) were diluted to a 5 μM **ST_HEAT** concentration for analysis unless otherwise noted. Some of the 80 μM HPLC purified **ST_HEAT** assemblies were linked with 24.2% (v/v), 19.4% (v/v), 13.8% (v/v), 7.4% (v/v), or 3.9% (v/v) **mCh_SC**. The 5 μM HPLC purified **ST_HEAT** assemblies in 10 mM acetate buffer (pH 5.0) were either imaged at their 5 μM concentration or diluted to a 1 μM or 0.2 μM **ST_HEAT** concentration. Some of the 5 μM HPLC purified **ST_HEAT** assemblies were linked with 5% (v/v) or 2% (v/v) **mCh_SC**. The protein samples were either nascent or thermally annealed before linking with **mCh_SC**.

Protein assemblies were analyzed using 200 mesh carbon-coated copper grids (Electron Microscopy Sciences) with a 2% methylamine tungstate (MAT) solution as a stain. Samples were prepared by incubating 4 μL of a protein solution on a grid for 1 minute before wicking excess sample away using filter paper. MAT stain was immediately added to the grid and the solution was allowed to incubate for 1 minute. Stain was wicked away with filter paper and the samples were left to dry under vacuum until imaging. TEM images were obtained on a Hitachi 7700 TEM with accelerating voltages of 80 kV.

Ultracentrifugation for Removal of Excess mCh_SC

Preparative ultracentrifugation was used as an attempt to remove excess **mCh_SC** from ligations of thermally annealed **ST_HEAT** (HPLC) with **mCh_SC**. The centrifuge was a Beckman Coulter Optima™ XE-90 Preparative Ultracentrifuge with a Type 45 Ti fixed-angle titanium rotor, which has a maximum speed of rotation of 45,000 rpm (235,000 x g). Delrin™ adapters were used for holding smaller 13.5 mL Thickwall tubes, which decreases the maximum speed of rotation to 30,000 rpm since the tubes were run without caps. Furthermore, the sample volume for the experiment was 1 mL, which was the minimum volume for the instrument in this configuration. Running at low volumes reduces the maximum speed of rotation even further, so the maximum speed of rotation was set at 25,000 rpm for these samples.

A thermally annealed 80 μM **ST_HEAT** (HPLC) sample in 10 mM acetate buffer (pH 5.0) was ligated to 13.8% (v/v) **mCh_SC** for a molar ratio of 1:2 (described in detail above). The sample was allowed to incubate at room temperature for 18 hours. A thermally annealed sample of 80 μM **ST_HEAT** (HPLC) in 10 mM acetate buffer (pH 5.0) was prepared as a control. A 13.8% **mCh_SC** solution was used as a second control. The initial sample volume of the ligation

sample and the two controls was 150 μ L. All three samples were diluted to a volume of 1 mL to reach the minimum volume for the ultracentrifuge setup.

The ligation sample and the two control samples sustained ultracentrifugation at 30,000 x g (16,100 rpm) at 4 °C for 2 hours. After the spin, all of the supernatant was carefully removed from the tubes and transferred to clean 1.5 mL centrifuge tubes. A 150 μ L volume of 10 mM acetate buffer (pH 5.0) was added to the ultracentrifuge tubes. The tubes were agitated for 1 week to resuspend any pelleted molecules.

Samples of were collected from each of the supernatants and resuspended pellets for analysis by TEM. An aliquot of 50 μ L was collected from each of the supernatants for SDS-PAGE analysis. The samples were immediately prepared for SDS-PAGE analysis by mixing the 50 μ L sample with 50 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 °C for 5 minutes. Analysis of the samples on a 10% SDS-PAGE gel confirmed that **ST_HEAT**, **mCh_SC**, and **mCh-HEAT** were all present in the supernatants.

The supernatants of the ligation sample and the two control samples were transferred into clean ultracentrifuge tubes. The supernatants were subjected to a second round of ultracentrifugation at 72,700 x g (25,000 rpm), which was the maximum possible speed of rotation for these samples. The samples spun at 4 °C for 2 hours. After the spin, the supernatants of the ligation sample and the **mCh_SC** sample separated into a clear fraction on the top and a pink fraction near the bottom. A 15 μ L aliquot was collected from the pink section of the supernatants for the ligation sample and the **mCh_SC** sample (see below).

The ultracentrifuge tubes were put back into the ultracentrifuge and the samples were allowed to spin at 72,700 x g (25,000 rpm) for an additional 2 hours. After the spin, the ligation sample and the **mCh_SC** sample were still only separated into two fractions with no visible pellet. Samples were collected from each of the supernatants near the bottom of the tube for TEM analysis. A 15 μ L aliquot of 15 μ L of the pink fraction at the bottom of supernatant was collected for the ligation sample for SDS-PAGE analysis (see below). The remaining supernatant was carefully removed in two fractions for the ligation sample and the **mCh_SC** sample. All of the supernatant was removed from the **ST_HEAT** tube. A 150 μ L volume of 10 mM acetate buffer (pH 5.0) was added to the ultracentrifuge tubes, and the tubes were agitated for 1 week to resuspend any pelleted molecules. Samples were collected from each of the resuspended pellets for TEM analysis.

The samples collected for SDS-PAGE analysis were immediately prepared by mixing the 15 μ L sample with 15 μ L of 2X SDS sample loading buffer (100 mM Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.2% (w/v) bromophenol blue, 10% (v/v) β -mercaptoethanol) and boiling the sample at 100 $^{\circ}$ C for 5 minutes. The samples were analyzed on a 10% SDS-PAGE gel.

Filtration using Centrifugation Filters for Removal of Excess mCh_SC

Centrifuge filters were used to remove **mCh_SC** from ligated samples. Initially, a thermally annealed 5 μ M **ST_HEAT** (HPLC) sample in 10 mM acetate buffer (pH 5.0) was ligated to 2% (v/v) **mCh_SC** and allowed to incubate for 18 hours (described in detail above). The initial sample volume of the ligation sample was 500 μ L. An aliquot was collected for TEM analysis.

To carry out the centrifuge filter protocol, the remaining sample volume was added to an Amicon® Ultra-0.5 centrifugal filter with a 50 kDa NMWL that was pre-equilibrated with 10 mM acetate buffer (pH 5.0). The filter was centrifuged at 14,000 x g for 10 minutes, and the flow through was transferred to a separate tube. A 500 µL volume of 10 mM acetate buffer (pH 5.0) was added to the centrifuge filter and then the filter was centrifuged at 14,000 x g for 10 minutes. This wash step was repeated five times while transferring the flow through to a separate tube after each wash. Once the buffer has flowed through on the last wash cycle, the inner filter was inverted into a clean centrifuge tube and centrifuged at 1,000 x g for 3 minutes. Buffer (10 mM acetate, pH 5.0) was added up to the initial volume to generate the initial concentration. An aliquot was collected for TEM analysis.

To improve results, an amended method was established. A thermally annealed 5 µM **ST_HEAT** (HPLC) sample in 10 mM acetate buffer (pH 5.0) was subjected to the centrifuge filter protocol above before the linking process; however, the annealed **ST_HEAT** sample was filtered using an Amicon® Ultra-0.5 centrifugal filter with a 100 kDa NMWL. An aliquot was collected of the filtered **ST_HEAT** sample for TEM analysis. The filtered **ST_HEAT** sample was then ligated to 2% (v/v) **mCh_SC** and allowed to incubate for 18 hours (described in detail above). After ligation, the sample was exposed to a second round of filtering using a 100 kDa NMWL centrifugal filter. The sample was washed an additional five rounds before the sample was removed from the filter. An aliquot was collected for TEM analysis of the filtered ligation sample.

Fluorimetry

Thermally annealed **ST_HEAT** (HPLC) assemblies with a protein concentration of 5 µM in 10 mM acetate buffer (pH 5.0) were filtered before the linking process with Amicon® Ultra-0.5

centrifugal filters (100 kDa NMWL). The filtered 5 μ M **ST_HEAT** samples were ligated to 2% (v/v) **mCh_SC**. After an 18-hour ligation period, the samples were filtered again using a 100 kDa NMWL centrifugal filter (described in detail above). A sample of thermally annealed 5 μ M **ST_HEAT** in 10 mM acetate buffer (pH 5.0) was prepared using the same methods for use as a negative control. A 2% **mCh_SC** sample in 10 mM acetate buffer (pH 5.0) was used as positive control. A range of **mCh_SC** samples with varying concentrations were analyzed to develop a standard curve. Samples were prepared with 2%, 1.5%, 1%, 0.8%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, and 0.1% (v/v) **mCh_SC**, which corresponds to 20 μ M, 15 μ M, 10 μ M, 8 μ M, 6 μ M, 5 μ M, 4 μ M, 3 μ M, 2 μ M, and 1 μ M **mCh_SC**.

Fluorescence data was collected using a Fluoromax-3 fluorimeter with a xenon lamp. Sample volumes of 50 μ L were analyzed in triplicate in a quartz cuvette. The excitation wavelength was set to 587 nm. The scan was set up to analyze the determined emission wavelength range at an increment of 1 nm and integration time of 0.1 s. After optimizing the conditions, the scans were run through an emission range of 550 nm to 750 nm with the excitation and emission slits (amount of light allowed to or from the sample) set at 3 nm. Fluorescence data was produced from an average of 3 scans and the intensity was plotted as a function of wavelength for data analysis. The **ST_HEAT** intensity was subtracted from the **mCh-HEAT** intensity to remove the excitation peak. The maximum fluorescence intensities of the **mCh_SC** samples were plotted against the concentration of **mCh_SC** to produce a standard curve. A trendline was formed from the linear segment of the curve and the resulting equation was used to calculate the approximate quantity of mCherry present in the **mCh-HEAT** sample.

Correlative Light Electron Microscopy (CLEM)

Correlative light electron microscopy was applied to thermally annealed samples of **ST_HEAT** that were functionalized with **mCh_SC**. A thermally annealed **ST_HEAT** (HPLC) assembly with a protein concentration of 5 μM in 10 mM acetate buffer (pH 5.0) was filtered before the linking process with Amicon[®] Ultra-0.5 centrifugal filters (100 kDa NMWL). The filtered 5 μM **ST_HEAT** sample was ligated to 2% (v/v) **mCh_SC** and then incubated for 18 hours at room temperature. The sample was filtered a second time using a 100 kDa NMWL centrifugal filter (described in detail above). TEM and fluorimetry analysis were conducted on the sample to verify the presence of nanotubes and fluorescence prior to CLEM analysis.

Electron cryomicroscopy (cryo-EM) samples were prepared by the Emory University Robert P. Apkarian Integrated Electron Microscopy Core service instructor, Spencer Hughes. Samples were diluted 1:10 in buffer. Cryo-EM specimens were prepared by applying aliquots (4 μL) of the diluted solution of nanotubes to glow-discharged 200 mesh copper Quantifoil grids (Quantifoil Micro Tools GMBH; Großlöbichau, Germany) and plunge freezing in liquid ethane using a Vitrobot Mark III (FEI, Hillsboro, Oregon). Cryo-EM data was collected using a JEOL JEM-2200FS 200 kV FEG-TEM with an in-column Omega energy filter (slit width 20 eV). Images were acquired on a Direct Electron, LP DE-20 direct electron detector (San Diego, CA). Contrast was enhanced using a hole-free phase plate for phase imaging of cryoimmobilized specimens.

6.4.5 Tables

Table 6.1 Protein sequences utilized in Chapter 6.

Protein	Sequence	MW (Da) # of AA
ST_HEAT	GAHIVMVDAYKPTKTSGGGSGGGAS[GDERAVEALIKALK DPDGWVRKAAAEALGRIGDERAVEALIKALKDPDWFVRE AAAKALGEI] ₄ GSMHHHHHH	30069.92 282
mCh_SC (After TEV Cleavage)	GMSSGLVPRGSHMVSKGEEDNMAIIEFMRFKVHMEGSV NGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSP QFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDG GVVTVTQDSSLQDGEFIYKVKLRGTFNPSDGPVMQKKTMG WEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTY KAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHS TGGMDELYKKNSSGGLVAGGSGGGSGGGTGGGSGGGTSG AMVDTLSGLSSEQQSGDMTIEEDSATHIKFSKRDEDGKEL AGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETA APDGYEVATAITFTVNEQQQVTVNGKATKGAHI	42406.74 392

6.5 References

1. Pluckthun, A., Designed Ankyrin Repeat Proteins (DARPin): Binding Proteins for Research, Diagnostics, and Therapy. In *Annual Review of Pharmacology and Toxicology, Vol 55*, Insel, P. A., Ed. 2015; Vol. 55, pp 489-511.
2. Urvoas, A.; Guellouz, A.; Valerio-Lepiniec, M.; Graille, M.; Durand, D.; Desravines, D. C.; van Tilbeurgh, H.; Desmadril, M.; Minard, P., Design, production and molecular structure of a new family of artificial alpha-helicoidal repeat proteins (alphaRep) based on thermostable HEAT-like repeats. *J Mol Biol* **2010**, *404* (2), 307-27.
3. Kajander, T.; Cortajarena, A. L.; Main, E. R.; Mochrie, S. G.; Regan, L., A new folding paradigm for repeat proteins. *J Am Chem Soc* **2005**, *127* (29), 10188-90.
4. Wetzel, S. K.; Settanni, G.; Kenig, M.; Binz, H. K.; Pluckthun, A., Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J Mol Biol* **2008**, *376* (1), 241-57.
5. Dorman, B. P.; Maestre, M. F., Experimental differential light-scattering correction to the circular dichroism of bacteriophage T2. *Proc Natl Acad Sci U S A* **1973**, *70* (1), 255-9.
6. Gregory, R. P.; Raps, S., The differential scattering of circularly polarized light by chloroplasts and evaluation of their true circular dichroism. *The Biochemical journal* **1974**, *142* (2), 193-201.
7. Chakraborty, H.; Lentz, B. R., A Simple Method for Correction of CD Spectra Obtained from Membrane-Containing Samples. *Biochemistry* **2012**, *51* (5), 1005-1008.
8. Mao, D.; Wallace, B. A., Differential light scattering and absorption flattening optical effects are minimal in the circular dichroism spectra of small unilamellar vesicles. *Biochemistry* **1984**, *23* (12), 2667-73.

9. Wallace, B. A.; Mao, D., Circular dichroism analyses of membrane proteins: an examination of differential light scattering and absorption flattening effects in large membrane vesicles and membrane sheets. *Analytical biochemistry* **1984**, *142* (2), 317-28.
10. Zakeri, B.; Fierer, J. O.; Celik, E.; Chittock, E. C.; Schwarz-Linek, U.; Moy, V. T.; Howarth, M., Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A* **2012**, *109* (12), E690-7.
11. Voortman, L. M., Integration Without Compromise. *Microscopy Today* **2014**, *22* (6), 30-35.
12. de Boer, P.; Hoogenboom, J. P.; Giepmans, B. N., Correlated light and electron microscopy: ultrastructure lights up! *Nat Methods* **2015**, *12* (6), 503-13.
13. Vishwakarma, A.; Bhise, N. S.; Evangelista, M. B.; Rouwkema, J.; Dokmeci, M. R.; Ghaemmaghami, A. M.; Vrana, N. E.; Khademhosseini, A., Engineering Immunomodulatory Biomaterials To Tune the Inflammatory Response. *Trends in Biotechnology* **2016**, *34* (6), 470-482.
14. Kelly, S. H.; Shores, L. S.; Votaw, N. L.; Collier, J. H., Biomaterial strategies for generating therapeutic immune responses. *Advanced drug delivery reviews* **2017**, *114*, 3-18.
15. van Riet, E.; Aina, A.; Suzuki, T.; Kersten, G.; Hasegawa, H., Combatting infectious diseases; nanotechnology as a platform for rational vaccine design. *Advanced drug delivery reviews* **2014**, *74*, 28-34.
16. Fan, Y.; Moon, J. J., Particulate delivery systems for vaccination against bioterrorism agents and emerging infectious pathogens. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **2017**, *9* (1), e1403.

17. Koshy, S. T.; Mooney, D. J., Biomaterials for enhancing anti-cancer immunity. *Current Opinion in Biotechnology* **2016**, *40*, 1-8.
18. Mehta, N. K.; Moynihan, K. D.; Irvine, D. J., Engineering New Approaches to Cancer Vaccines. *Cancer Immunology Research* **2015**, *3* (8), 836.
19. Melief, C. J. M.; van Hall, T.; Arens, R.; Ossendorp, F.; van der Burg, S. H., Therapeutic cancer vaccines. *The Journal of Clinical Investigation* **2015**, *125* (9), 3401-3412.
20. Pradhan, P.; Qin, H.; Leleux, J. A.; Gwak, D.; Sakamaki, I.; Kwak, L. W.; Roy, K., The effect of combined IL10 siRNA and CpG ODN as pathogen-mimicking microparticles on Th1/Th2 cytokine balance in dendritic cells and protective immunity against B cell lymphoma. *Biomaterials* **2014**, *35* (21), 5491-5504.
21. Agarwal, R.; Volkmer, T. M.; Wang, P.; Lee, L. A.; Wang, Q.; García, A. J., Synthesis of self-assembled IL-1Ra-presenting nanoparticles for the treatment of osteoarthritis. *Journal of Biomedical Materials Research Part A* **2016**, *104* (3), 595-599.
22. Singh, A.; Agarwal, R.; Diaz-Ruiz, C. A.; Willett, N. J.; Wang, P.; Lee, L. A.; Wang, Q.; Guldborg, R. E.; García, A. J., Nanoengineered Particles for Enhanced Intra-Articular Retention and Delivery of Proteins. *Advanced Healthcare Materials* **2014**, *3* (10), 1562-1567.
23. Whitmire, R. E.; Scott Wilson, D.; Singh, A.; Levenston, M. E.; Murthy, N.; García, A. J., Self-assembling nanoparticles for intra-articular delivery of anti-inflammatory proteins. *Biomaterials* **2012**, *33* (30), 7665-7675.
24. Kimmerling, K. A.; Furman, B. D.; Mangiapani, D. S.; Moverman, M. A.; Sinclair, S. M.; Huebner, J. L.; Chilkoti, A.; Kraus, V. B.; Setton, L. A.; Guilak, F.; Olson, S. A., SUSTAINED INTRA-ARTICULAR DELIVERY OF IL-1RA FROM A THERMALLY-RESPONSIVE

ELASTIN-LIKE POLYPEPTIDE AS A THERAPY FOR POST-TRAUMATIC ARTHRITIS.

European cells & materials **2015**, *29*, 124-140.

25. Pompano, R. R.; Chen, J.; Verbus, E. A.; Han, H.; Fridman, A.; McNeely, T.; Collier, J. H.; Chong, A. S., Titrating T-Cell Epitopes within Self-Assembled Vaccines Optimizes CD4+ Helper T Cell and Antibody Outputs. *Advanced Healthcare Materials* **2014**, *3* (11), 1898-1908.
26. Lewis, J. S.; Dolgova, N. V.; Zhang, Y.; Xia, C. Q.; Wasserfall, C. H.; Atkinson, M. A.; Clare-Salzler, M. J.; Keselowsky, B. G., A combination dual-sized microparticle system modulates dendritic cells and prevents type 1 diabetes in prediabetic NOD mice. *Clinical Immunology* **2015**, *160* (1), 90-102.
27. Yoon, Y. M.; Lewis, J. S.; Carstens, M. R.; Campbell-Thompson, M.; Wasserfall, C. H.; Atkinson, M. A.; Keselowsky, B. G., A combination hydrogel microparticle-based vaccine prevents type 1 diabetes in non-obese diabetic mice. *Scientific Reports* **2015**, *5*, 13155.
28. Röhn, T. A.; Jennings, G. T.; Hernandez, M.; Grest, P.; Beck, M.; Zou, Y.; Kopf, M.; Bachmann, M. F., Vaccination against IL-17 suppresses autoimmune arthritis and encephalomyelitis. *European Journal of Immunology* **2006**, *36* (11), 2857-2867.
29. Sonderegger, I.; Röhn, T. A.; Kurrer, M. O.; Iezzi, G.; Zou, Y.; Kastelein, R. A.; Bachmann, M. F.; Kopf, M., Neutralization of IL-17 by active vaccination inhibits IL-23-dependent autoimmune myocarditis. *European Journal of Immunology* **2006**, *36* (11), 2849-2856.
30. Vigneswaran, Y.; Han, H.; De Loera, R.; Wen, Y.; Zhang, X.; Sun, T.; Mora-Solano, C.; Collier, J. H., Winner of the Student Award in the Hospital Intern Category, 10th World Biomaterials Congress, May 17–22, 2016, Montreal QC, Canada: Peptide biomaterials raising adaptive immune responses in wound healing contexts. *Journal of Biomedical Materials Research Part A* **2016**, *104* (8), 1853-1862.

31. Sicari, B. M.; Rubin, J. P.; Dearth, C. L.; Wolf, M. T.; Ambrosio, F.; Boninger, M.; Turner, N. J.; Weber, D. J.; Simpson, T. W.; Wyse, A.; Brown, E. H. P.; Dziki, J. L.; Fisher, L. E.; Brown, S.; Badylak, S. F., An Acellular Biologic Scaffold Promotes Skeletal Muscle Formation in Mice and Humans with Volumetric Muscle Loss. *Science Translational Medicine* **2014**, *6* (234), 234ra58.
32. Badylak, S. F.; Dziki, J. L.; Sicari, B. M.; Ambrosio, F.; Boninger, M. L., Mechanisms by which acellular biologic scaffolds promote functional skeletal muscle restoration. *Biomaterials* **2016**, *103*, 128-136.
33. Chen, J.; Pompano, R. R.; Santiago, F. W.; Maillat, L.; Sciammas, R.; Sun, T.; Han, H.; Topham, D. J.; Chong, A. S.; Collier, J. H., The use of self-adjuvanting nanofiber vaccines to elicit high-affinity B cell responses to peptide antigens without inflammation. *Biomaterials* **2013**, *34* (34), 8776-85.
34. Wu, Y.; Norberg, P. K.; Reap, E. A.; Congdon, K. L.; Fries, C. N.; Kelly, S. H.; Sampson, J. H.; Conticello, V. P.; Collier, J. H., A Supramolecular Vaccine Platform Based on α -Helical Peptide Nanofibers. *ACS Biomaterials Science & Engineering* **2017**, *3* (12), 3128-3132.
35. Sambrook, J.; Russell, D. W.; Sambrook, J.; Russell, D. W., *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press {a} , 10 Skyline Drive, Plainview, NY, 11803-2500, USA: 2001.
36. 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.* **2015**, *137* (24), 7793-7802.

37. Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P., Structurally Defined Nanoscale Sheets from Self-Assembly of Collagen-Mimetic Peptides. *J. Am. Chem. Soc.* **2014**, *136* (11), 4300-4308.
38. Jiang, T.; Xu, C.; Zuo, X.; Conticello, V. P., Structurally Homogeneous Nanosheets from Self-Assembly of a Collagen-Mimetic Peptide. *Angewandte Chemie International Edition* **2014**, *53* (32), 8367-8371.

Appendix A

Sequences

A.1 Gene Sequence for 3FD-LL Helix Insert (located in pBB13)

GCATCAAGCTTGAAGACTTAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCA
GCGCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAG
CGCTGAAAGAGCTGGCAGCCTGAGACGGGATCCACTAG

A.2 Gene Sequence for 3FD-LL Adapter (located in pBB12 and pBB14)

GCTTAGAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATC
ATCATCATCACGGATCCGGCGAAGCCCGAGACCGTCGACGGTCTCCAGCCCTGGAG
AACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGGAGAACTGGCAGCGCTGAAAG
AACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAAAGAGCTGGCAAAGTAATAAGCT
TGGTCC

Adapter expresses to the 61 amino acid protein:

MGHHHHHHHHHHH-GSGEA-RDRRRSP-

ALEKLALEKLALEKLAALKELAALKELAALKELA-K

A.3 Gene Sequence for 3FD-LL Monomer (located in pBB15)

3FD-LL adapter with insertion site removed

GAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATCATCAT
CATCACGGATCCGGCGAAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGC
GCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCG
CTGAAAGAGCTGGCAAAGTAATAAGCTT

A.4 Amino Acid Sequence for 3FD-LL Monomer

GHHHHHHHHHHHGSGEALEKLAALEKLAALEKLAALKELAALKELAALKELAK

53 amino acids

5659.47 Da

pI = 8.22

A.5 Gene Sequence for 3FD-LL Dimer (located in pBB17)

3FD-LL adapter with one 3FD-LL helix inserted

GAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATCATCATCAT
CATCACGGATCCGGCGAAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGC
GCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCG
CTGAAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCT
GGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTG
AAAGAGCTGGCAAAGTAATAAGCTT

A.6 Amino Acid Sequence for 3FD-LL Dimer

GHHHHHHHHHHGSGEALEKLAALEKLAALEKLAALKELAALKELAALKELAALKELA
ALEKLAALEKLAALKELAALKELAALKELAK

88 amino acids

9414.02 Da

pI = 8.16

A.7 Gene Sequence for 3FD-LL Trimer (located in pBB20)

3FD-LL adapter with two 3FD-LL helices inserted

GAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATCATCAT
CATCACGGATCCGGCGAAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGC
GCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCG
CTGAAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCT
GGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTG
AAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGG
AGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAA
AGAGCTGGCAAAGTAATAAGCTT

A.8 Amino Acid Sequence for 3FD-LL Trimer

GHHHHHHHHHHGSGEALEKLAALEKLAALEKLAALKELAALKELAALKELA
ALEKLAALEKLAALKELAALKELAALKELAALKELAALKELAALKELA
LAALKELAK

124 amino acids

13168.56 Da

pI = 8.11

A.9 Gene Sequence for 3FD-LL Tetramer (located in pBB21)

3FD-LL adapter with three 3FD-LL helices inserted

GAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATCATCAT
 CATCACGGATCCGGCGAAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGC
 GCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCG
 CTGAAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCT
 GGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTG
 AAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGG
 AGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAA
 AGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGGAG
 AACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAAAG
 AGCTGGCAAAGTAATAAGCTT

A.10 Amino Acid Sequence for 3FD-LL Tetramer

GHHHHHHHHHHGSGEALEKLAALKLAALKLAALKELAALKELAALKELAALKLA
 ALEKLAALKLAALKELAALKELAALKELAALKLAALKLAALKLAALKELAALKE
 LAALKELAALKLAALKLAALKLAALKELAALKELAALKELAK

160 amino acids

16923.1 Da

pI = 8.08

A.11 Gene Sequence for 3FD-LL Pentamer (located in pBB22)

3FD-LL adapter with four 3FD-LL helices inserted

GAATTCATTAAAGAGGAGAAATTAACCATGGGCCATCATCATCATCATCATCAT
 CATCACGGATCCGGCGAAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGC
 GCTGGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCG
 CTGAAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCT
 GGAGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTG
 AAAGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGG
 AGAACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAA
 AGAGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGGAG
 AACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAAAG
 AGCTGGCAGCCCTGGAGAACTGGCAGCGCTGGAAAAGCTGGCAGCGCTGGAGAA
 ACTGGCAGCGCTGAAAGAACTGGCGGCACTGAAAGAGCTGGCAGCGCTGAAAGAG
 CTGGCAAAGTAATAAGCTT

A.12 Amino Acid Sequence for 3FD-LL Pentamer

GHHHHHHHHHGSGEALEKLAALEKLAALEKLAALKELAALKELAALKELAALKELA
 ALEKLAALEKLAALKELAALKELAALKELAALKELAALKELAALKELAALKELA
 LAALKELAALKELAALKELAALKELAALKELAALKELAALKELAALKELAALKELA
 EKLAALKELAALKELAALKELAK

196 amino acids

20677.65 Da

pI = 8.05

A.13 Gene Sequence for SpyTag-HEAT-8R-HisTag

ATGGGTGCACATATTGTGATGGTCGATGCGTATAAACCGACGAAAAGCTAGCGGCGG
CGGCAGCGGTGGCGGCGCGAGCGGGGATGAACGGGCTGTAGAAGCATTGATTAAG
CTCTGAAAGACCCCGATGGATGGGTAAGGAAAGCAGCTGCAGAGGCACTTGGTAGG
ATAGGAGACGAGCGTGCTGTCGAAGCACTGATCAAAGCACTCAAAGATCCTGATTG
GTTCGTAAGAGAGGCTGCAGCAAAGGCTCTTGGGGAAATCGGAGATGAACGCGCAG
TCGAGGCTCTAATTAAGGCATTGAAGGACCCCTGATGGTTGGGTTAGAAAAGCCGCA
GCTGAGGCTTTAGGAAGAATAGGGGACGAACGTGCAGTAGAGGCGTTAATTAAGC
GTTGAAAGATCCGGACTGGTTTGTACGTGAAGCTGCCGCTAAGGCTTTGGGAGAAAT
TGGTGACGAAAGAGCAGTGGAGGCCCTCATTAAGCACTAAAAGATCCCGACGGGT
GGGTCAGAAAGGCAGCAGCAGAAGCTCTAGGTAGAATTGGAGATGAGAGAGCGGT
TGAAGCCTTGATAAAAGCATTAAAAGACCCGGATTGGTTTGTAGGGAAGCAGCCG
CGAAAGCTCTCGGTGAAATAGGTGATGAGCGAGCTGTGGAAGCCCTAATAAAGGCG
CTTAAAGACCCAGATGGCTGGGTACGGAAGGCCGCTGCTGAAGCATTAGGTCGTAT
TGGCGATGAAAGGGCAGTTGAGGCATTAATAAAAGCTCTTAAAGATCCAGACTGGT
TCGTTAGAGAAGCGGCGGCTAAAGCATTGGGTGAGATAGGATCCATGCATCACCAC
CATCACCATTAATAA

A.14 Gene Sequence for HisTag-mCherry-SpyCatcher

ATGAAACATCACCACCACCACCATGGCACATCTGAAAACCTTGTATTTCCAGGGCATG
TCATCTGGTCTGGTCCCTCGCGGTTCTCACATGGTTTCTAAAGGTGAAGAGGATAAC
ATGGCTATCATCAAAGAATTCATGCGCTTCAAAGTCCACATGGAGGGTTCCGTGAAT
GGTCACGAATTTGAGATTGAGGGCGAGGGTGAGGGTCGTCCGTACGAAGGCACGCA
AACCGCTAAGCTGAAAGTCACCAAGGGTGGTCCGCTGCCGTTTCGCGTGGGACATTCT
GAGCCCTCAGTTTATGTACGGTAGCAAAGCGTACGTTAAGCATCCGGCAGACATCCC
GGACTATCTGAAACTGTCGTTTCCGGAAGGCTTTAAATGGGAGCGTGTCATGAACTT
TGAGGACGGCGGCGTGGTGACCGTACTCAGGACTCCAGCCTGCAAGACGGCGAGT
TTATCTATAAAGTTAAGTTGCGTGGTACCAACTTCCAAGCGATGGTCCGGTTATGC
AGAAAAAGACCATGGGTTGGGAAGCATCCAGCGAGCGTATGTATCCAGAGGACGGT
GCGCTGAAGGGTGAGATTAACAGCGCTTGAAGCTGAAGGATGGCGGTCACTACGA
CGCCGAAGTTAAGACCACCTATAAAGCGAAAAAACCGGTGCAGTTGCCGGGTGCGT
ACAATGTTAACATTAAACTGGATATTACCAGCCATAACGAAGATTACACCATCGTAG
AGCAATATGAACGTGCAGAGGGCCGCGACTCGACGGGTGGCATGGACGAACTGTAC
AAAAAGAATAGCGGTGGCGGCCTGGTGGCCGGTGGTAGCGGTGGCGGTTCTGGTGG
CGGCACTGGTGGTGGCTCCGGTGGTGGCACGAGCGGTGCGATGGTCGATACGTTGA
GCGGCCTGAGCAGCGAGCAAGGTCAATCCGGTGACATGACCATTGAAGAGGACAGC
GCCACCCATATCAAATTCAGCAAGCGTGACGAAGATGGCAAAGAACTGGCCGGTGC
GACGATGGAACTGCGTGATAGCAGCGGTAAGACTATCAGCACCTGGATTAGCGACG
GTCAGGTCAAAGATTTCTACCTGTATCCGGGTAAATATACTTTCGTGGAAACCGCAG
CACCGGATGGTTATGAAGTCGCAACCGCGATTACGTTTACGGTGAACGAACAGGGT
CAGGTGACGGTTAATGGTAAGGCTACGAAAGGTGACGCGCATATCTAA

A.15 Amino Acid Sequence for SpyTag-HEAT-8R-HisTag

MGAHIVMVDAYKPTKTSGGGSGGGASGDERAVEALIKALKDPDGWVRKAAAEALGRI
GDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALIKALKDPDGWVRKAAAE
ALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALIKALKDPDGWVRK
AAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGDERAVEALIKALKDPDG
WVRKAAAEALGRIGDERAVEALIKALKDPDWFVREAAAKALGEIGSMHHHHHH

282 amino acids

30069.92 Da

pI = 5.5

A.16 Amino Acid Sequence for HisTag-mCherry-SpyCatcher (Before TEV Cleavage)

MKHHHHHHGHTSENLYFQGMSSGLVPRGSHMVSKGEEDNMAIIKEFMRFKVHMEGSVN
GHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY
LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPDGPVMQKKT
MGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVQLPGAYNVNI
KLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKNSGGGLVAGGSGGGSGGGTGGGS
GGGTSGAMVDTLSGLSSEQQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSG
KTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQQQVTVNGKATKG
DAHI

408 amino acids

44298.7 Da

pI = 5.55

A.17 Amino Acid Sequence for HisTag-mCherry-SpyCatcher (After TEV Cleavage)

MKHHHHHHGTSENLYFQGMSSGLVPRGSHMVSKGEEDNMAIIKEFMRFKVHMEGSVN
GHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY
LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPDGPVMQKKT
MGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVQLPGAYNVNI
KLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKNSGGGLVAGGSGGGSGGGTGGGS
GGGTSGAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSG
KTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQQQVTVNGKATKG
DAHI

392 amino acids

42406.74 Da

pI = 5.04

Appendix B

Mass Spectra

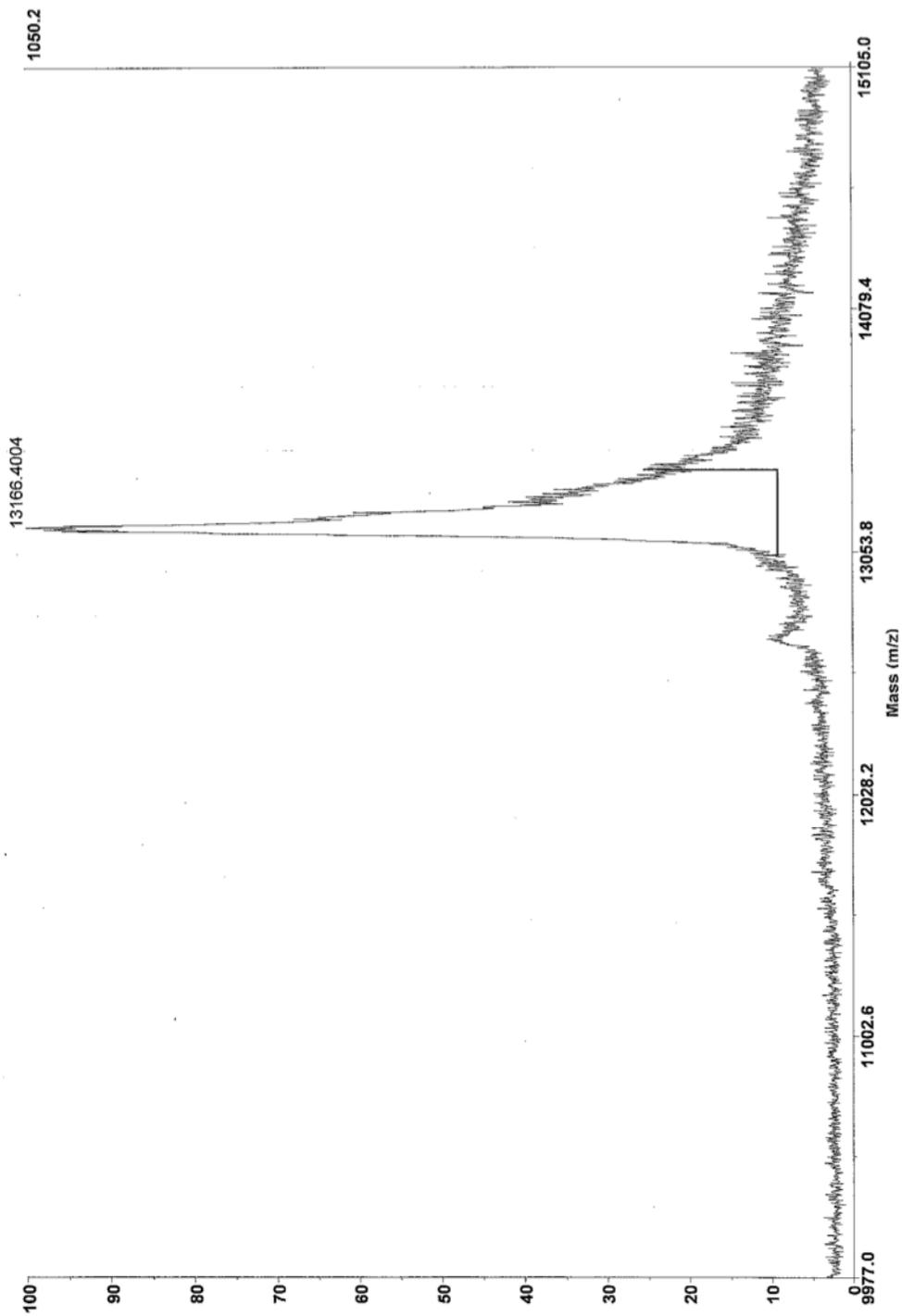


Figure B.1 MALDI-TOF mass spectrum of HPLC purified 3FD-LL trimer. The experimental average molar mass of the 3FD-LL trimer was 13167.04 Da from 4 separate MALDI spectra. The theoretical molar mass of the 3FD-LL trimer was 13168.56 Da.

4700 Reflector Spec #1[BP = 5662.4, 135]

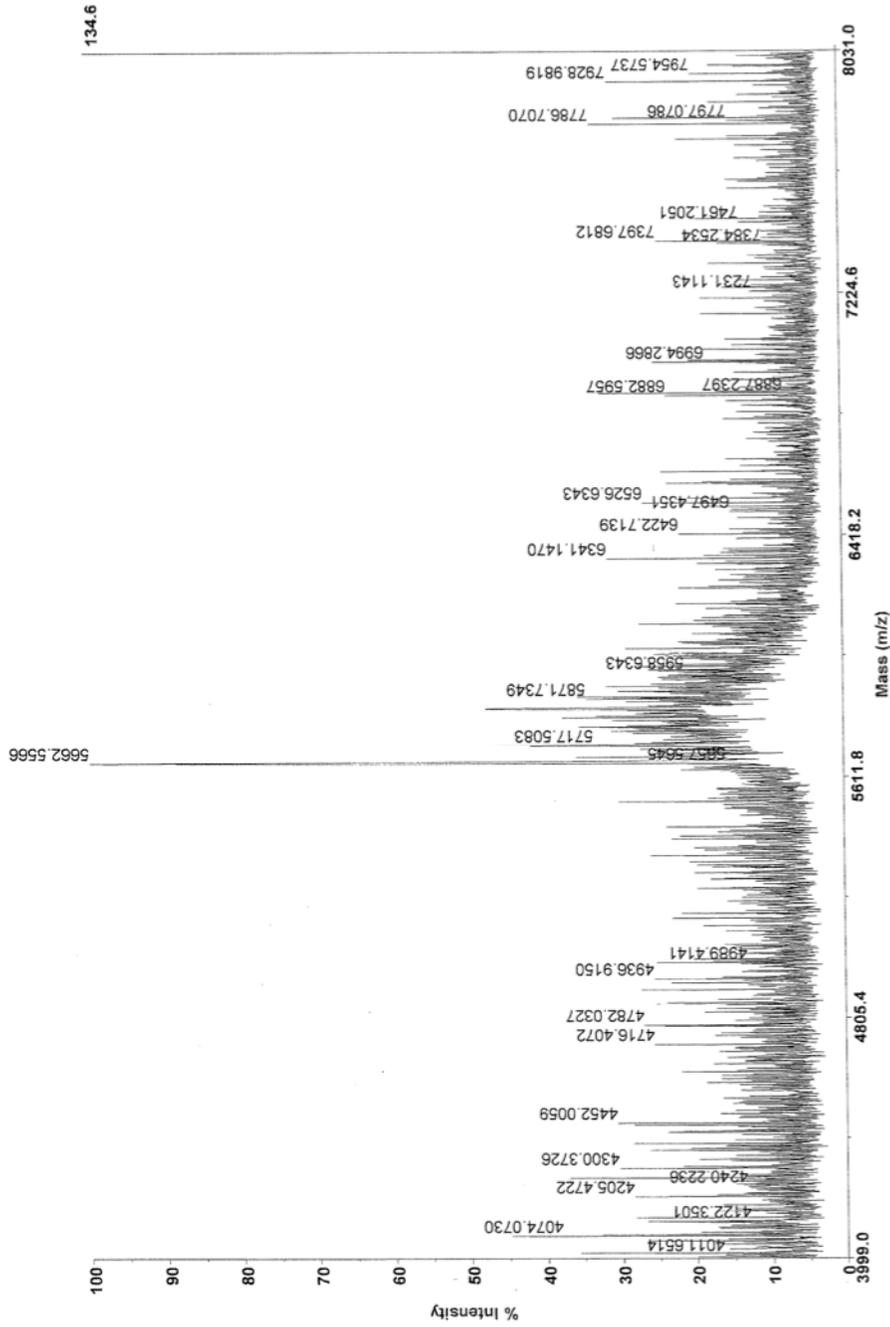


Figure B.2 MALDI-TOF mass spectrum of the HisPur™ 3FD-LL monomer. The experimental average molar mass of the 3FD-LL monomer was 5660.74 Da from 5 separate MALDI spectra. The theoretical molar mass of the 3FD-LL monomer was 5659.47 Da.

4700 Reflector Spec #1 [BP = 9416.9, 135]

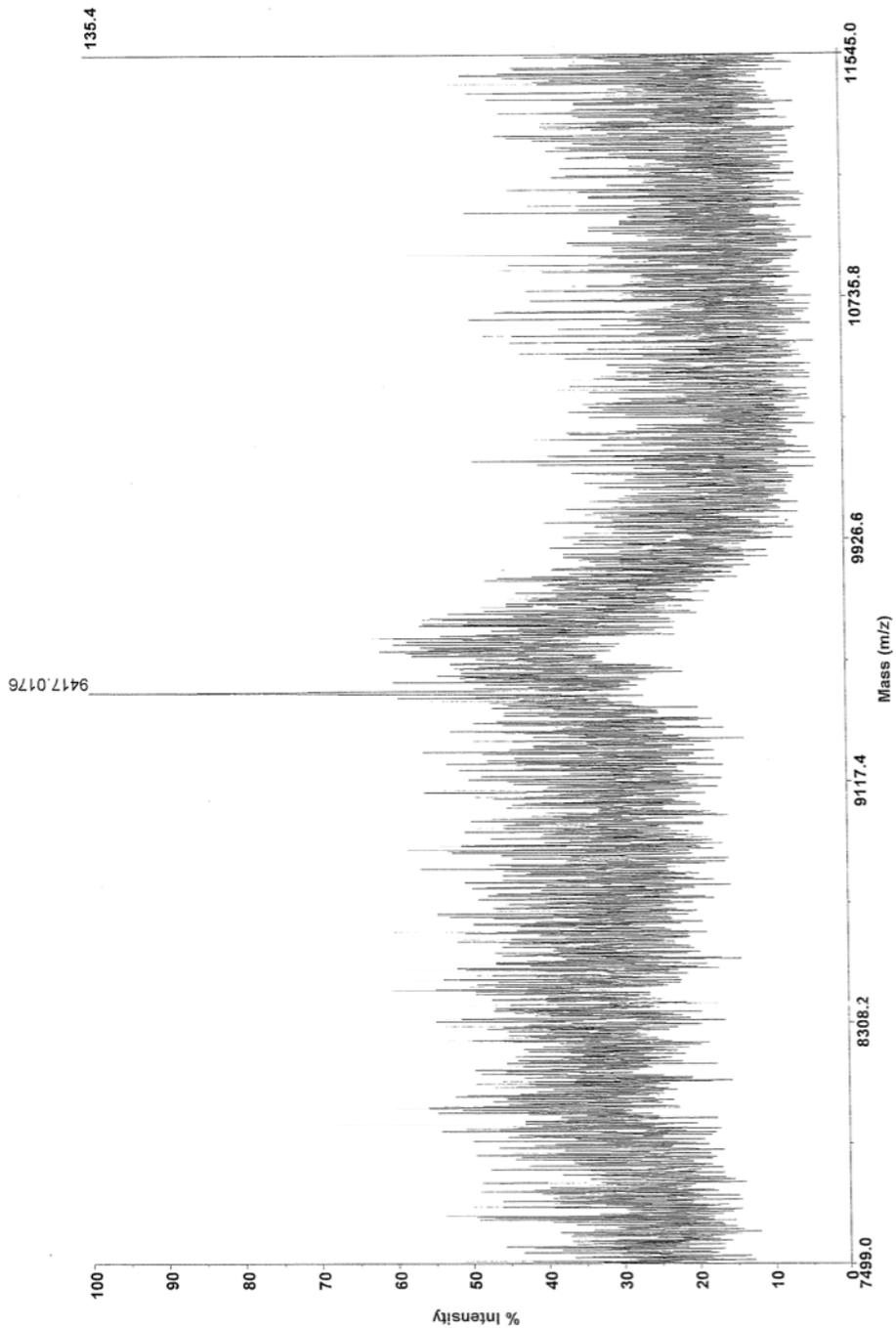


Figure B.3 MALDI-TOF mass spectrum of the HisPur™ 3FD-LL dimer. The experimental average molar mass of the 3FD-LL dimer was 9415.12 Da from 4 separate MALDI spectra. The theoretical molar mass of the 3FD-LL dimer was 9414.02 Da.

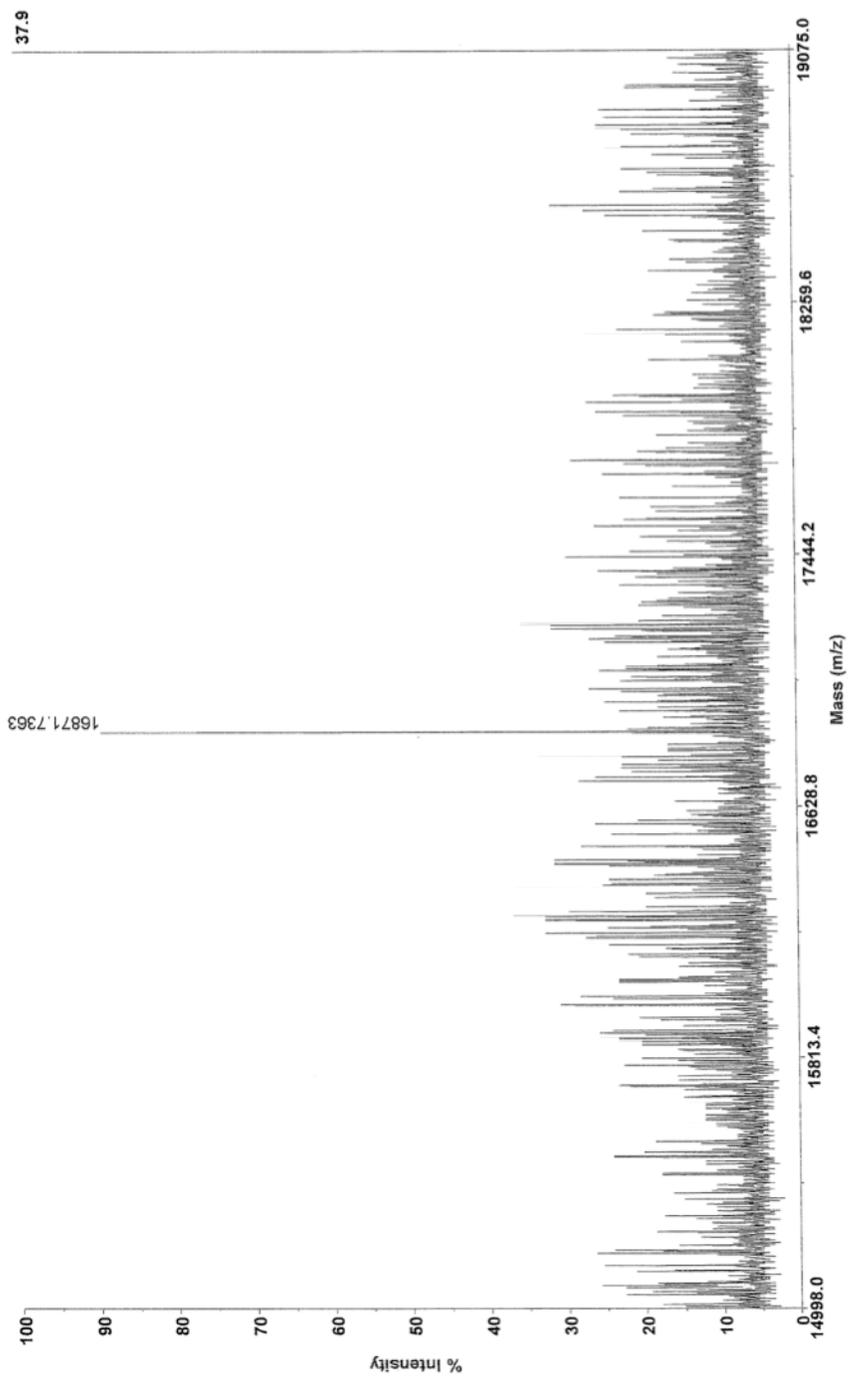


Figure B.4 MALDI-TOF mass spectrum of the HisPur™ 3FD-LL tetramer. The experimental average molar mass of the 3FD-LL tetramer was 16918.54 Da from 5 separate MALDI spectra. The theoretical molar mass of the 3FD-LL tetramer was 16923.1 Da.

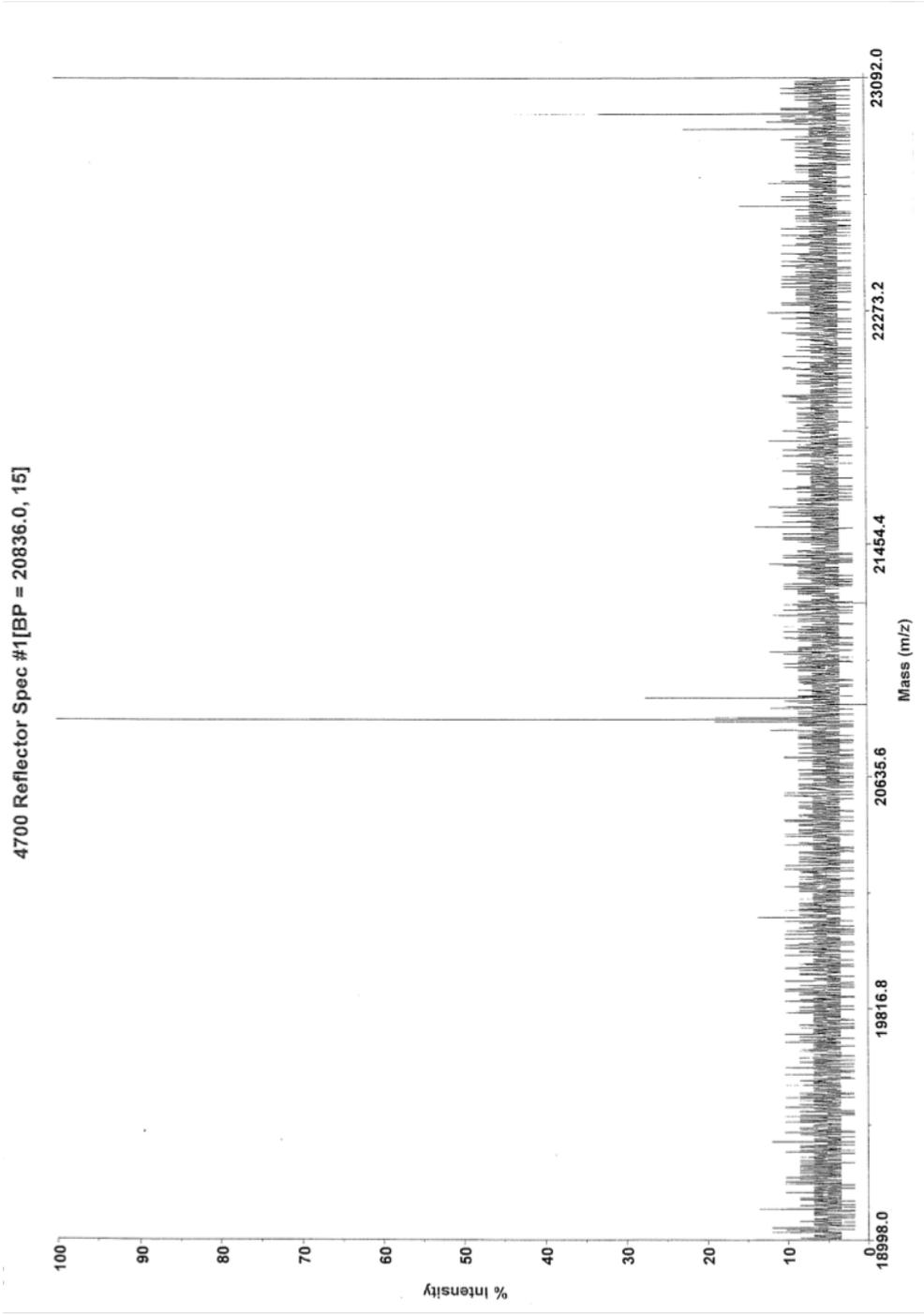


Figure B.5 MALDI-TOF mass spectrum of the HisPur™ 3FD-LL pentamer. The experimental average molar mass of the 3FD-LL pentamer was 20828.19 Da from 4 separate MALDI spectra. The theoretical molar mass of the 3FD-LL pentamer was 20677.65 Da. The theoretical corresponds to the protein sequence with the *N*-terminal methionine attached.

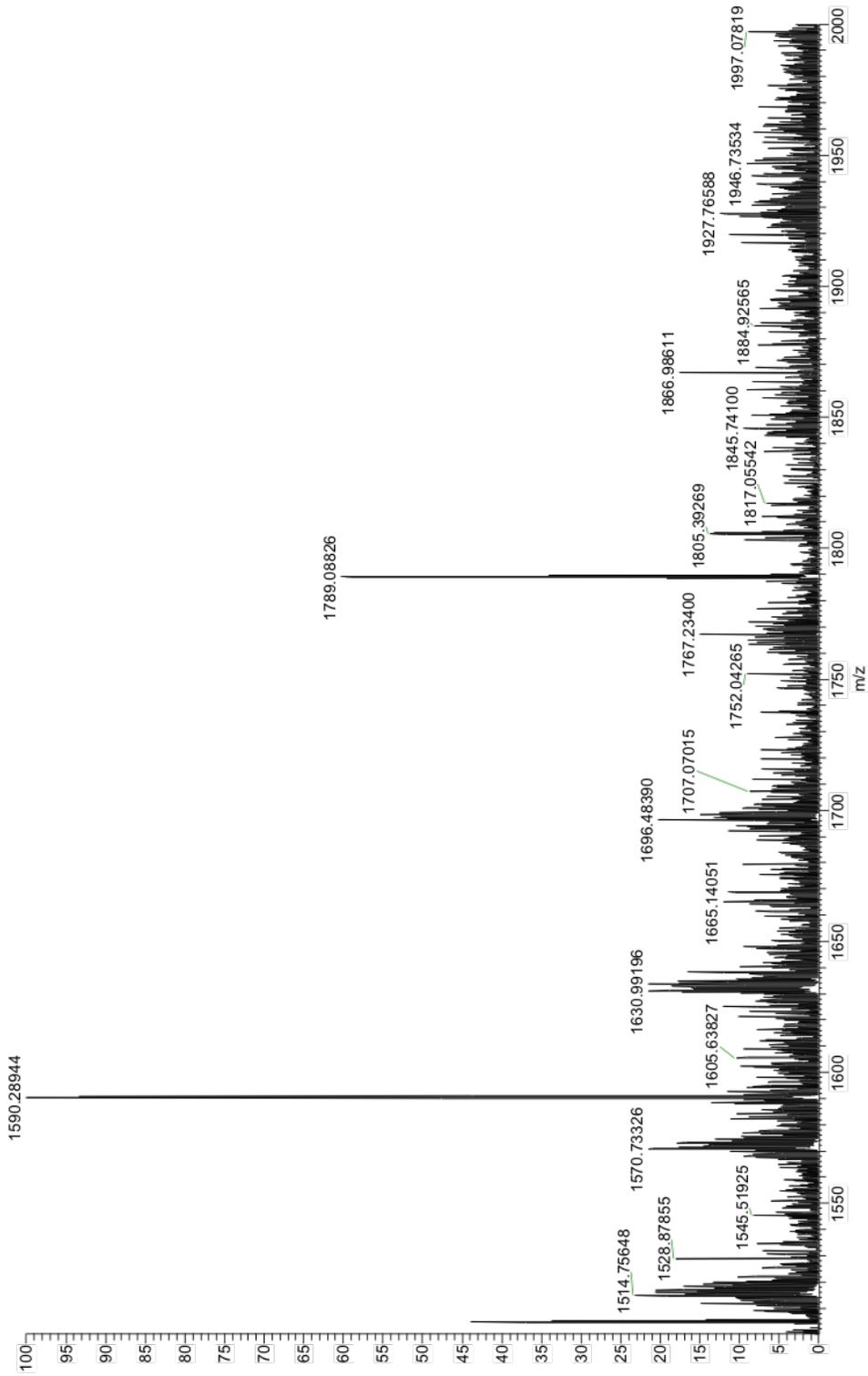


Figure B.6 ESI mass spectrometry of the mCh_{SC} protein after cleavage with TEV protease. The mass spectra verifies the molecular weight for the mCh_{SC} protein without the N-terminal His-tag and TEV cleavage site. The presence of addition bands correspond to the expected molecular weight for the TEV protease enzyme in addition to the cleaved amino acids.

4700 Linear Spec #1[BP = 30072.4, 79]

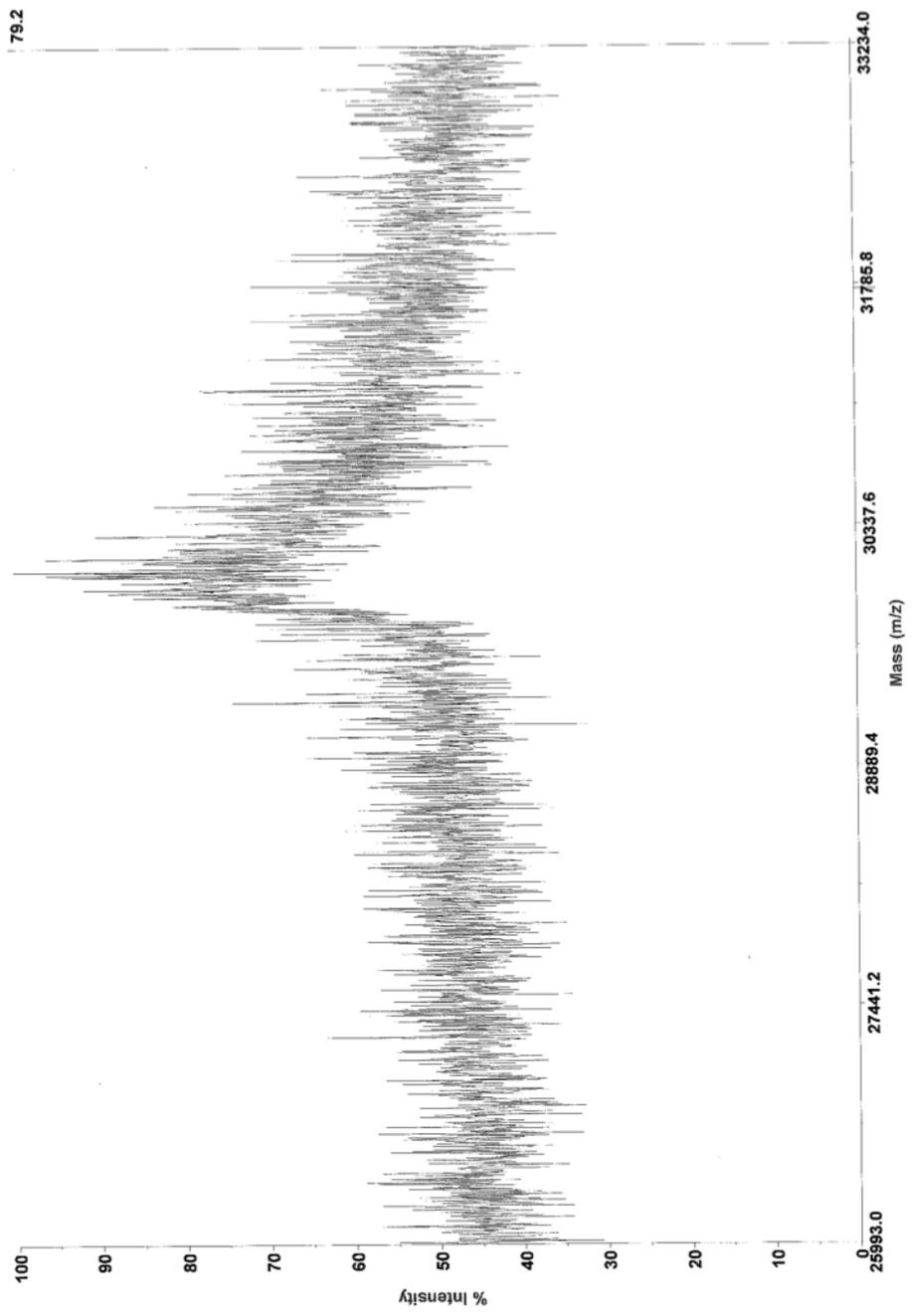


Figure B.7 MALDI-TOF mass spectrum of HPLC purified ST_HEAT. The experimental average molar mass of the ST_HEAT protein was 30072.6 Da from 4 separate MALDI spectra. The theoretical molar mass of the ST_HEAT protein was 30069.92 Da.