# REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggesstions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any oenalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

| 1. REPORT DATE (DD-MM-YYYY) | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
|---|---|---|
| 01-03-2016 | Ph.D. Dissertation | - |

4. TITLE AND SUBTITLE
Self-assembling Protein Materials for Metal Nanoparticle Templation

5a. CONTRACT NUMBER
W911NF-11-1-0449

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Jasmin Hume

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Polytechnic University - Brooklyn
Office of Sponsored Research
Polytechnic Institute of New York University
Brooklyn, NY 11201 -3840

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
59617-LS.9

12. DISTRIBUTION AVAILIBILITY STATEMENT

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The fabrication of de novo proteins able to self-assemble on the nano- to meso- length scales is critical in the development of protein-based biomaterials in nanotechnology and medicine. Here we report the design and characterization of a protein-engineered coiledcoil that not only assembles into microfibers but also can bind hydrophobic small molecules. Under ambient conditions the protein forms fibers with nanoscale structure possessing large aspect ratios formed by bundles of ?-helical homopentameric assemblies, which further assemble into mesoscale fibers in the presence of curcumin through aggregation. Surprisingly, these biosynthesized fibers are

15. SUBJECT TERMS
protein-based biolaterials

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Jin Montclare |
| | | | | | 19b. TELEPHONE NUMBER: 646-997-3679 |

## Report Title

Self-assembling Protein Materials for Metal Nanoparticle Templation

## ABSTRACT

The fabrication of de novo proteins able to self-assemble on the nano- to meso- length scales is critical in the development of protein-based biomaterials in nanotechnology and medicine. Here we report the design and characterization of a protein-engineered coiledcoil that not only assembles into microfibers but also can bind hydrophobic small molecules. Under ambient conditions the protein forms fibers with nanoscale structure possessing large aspect ratios formed by bundles of ?-helical homopentameric assemblies, which further assemble into mesoscale fibers in the presence of curcumin through aggregation. Surprisingly, these biosynthesized fibers are able to form in conditions of remarkably low concentrations. Unlike previously designed coiled-coil fibers, these engineered protein microfibers can bind the small molecule curcumin throughout the assembly, serving as a depot for encapsulation and delivery of other
chemical agents within protein-based 3D microenvironments.

# Self-assembling Protein Materials for Metal Nanoparticle Templation

DISSERTATION

Submitted in Partial Fulfillment of
the Requirements for
the Degree of

DOCTOR OF PHILOSOPHY (Materials Chemistry)

at the

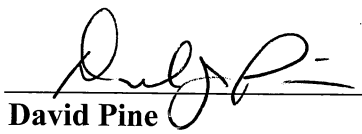
NEW YORK UNIVERSITY
POLYTECHNIC SCHOOL OF ENGINEERING

by

Jasmin Hume

May 2015

Approved:



**David Pine**
Department Head

MAY 14, 2015
Date

UMI Number: 3712449

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.




UMI 3712449

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code




ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Approved by the Guidance Committee:

Major: Materials Chemistry

**Jin Kim Montclare**
Associate Professor

February 9, 2015
Date

**Rastislav Levicky**
Associate Professor

February 9, 2015
Date

**Kalle Levon**
Professor

February 9, 2015
Date

**Cenk Sumen**
Adjunct Professor

February 9, 2015
Date

*Vita*

I grew up in Manhattan, New York after moving to the United States from Sweden at the age of two years. Until the age of thirteen I attended the United Nations International School in Manhattan. I attended Stuyvesant High School in Manhattan, a specialized high school for mathematics and science. I received my Bachelors in Engineering in Chemical Engineering with a minor in Biotechnology from McGill University in Montréal, Canada. During this time I pursued an undergraduate research position under the advisement of Professor Richard Leask, working on functionalized biomaterial surfaces for prosthetic heart valves and stents, sparking my interest in biomaterials development. After McGill I moved to Sweden to obtain my Masters in Science in Advanced Materials from Chalmers Tekniska Högskolan (Chalmers University of Technology) in Göteborg, Sweden. At Chalmers I performed my thesis research in the laboratory of Professor Bengt Kasemo in the Applied Physics department with Professor Sofia Svedhem as my advisor. My thesis was entitled "Thrombogenic Surface Modifications for Blood-Coagulation Studies" and involved aresearch collaboration with Linköping University in Sweden. After obtaining my Masters, I moved to Milan, Italy to work in the Advanced Systems Technology department doing research and development on a point-of-care DNA biosensor at STMicroelectronics in Agrate Brianza, Italy. I worked at STMicroelectronics as a research engineer forging collaborations with Politecnico di Milano in Italy and Université de Mons in Belgium over the course of three years.

After working at STMicroelectronics I returned to New York City to pursue my doctorate in Materials Chemistry in the Chemical and Biomolecular Engineering department at the NYU School of Engineering in Brooklyn, New York. My thesis work has been performed in the lab of Professor Jin Kim Montclare. Two years of my doctorate were supported through the NSF funded AMPS/CBRI program, where I taught at Community Partnership Charter School and the Mott Hall Bridges Academy. During my time at NYU I cofounded BenchPals, a startup focused on digitizing the laboratory environment. As a result, I became an active member of the entrepreneurial community at NYU and worked as a summer associate at Lux Capital in 2014.

I am one of four sisters and my parents are Dr. Rigmor Spång and Cameron Hume. My family is based in New York, Washington DC, and Switzerland. Some of my passions include traveling, skiing, baking, and spending time with friends, family, and my dog.

*Acknowledgements*

There are many people who I would like to sincerely thank for helping me complete my doctorate. Without the collective efforts of the people mentioned in this section I would surely have ended up on a different track, and so it gives me great satisfaction to briefly give my gratitude to these select individuals.

First of all, I express a great deal of appreciation for the members of the Montclare Lab: my lab mates, peers, and friends. Thank you Jennifer, Carlo, Susheel, Haresh, Ching Yao, Min, Kinjal, Liming, Joe, Rudy, Raymond, Jen Sun, and Michael for teaching me, laughing with me, and being willing to tackle almost any problem together. Thank you Susheel for taking my under your wings when I joined the lab and teaching me how to be a successful protein engineer. My lab mates showed me how to balance being a productive PhD student while making time for other important things in life, like friendship (and coffee). I have learned so much from you all and thank you from the bottom of my heart.

I would never have been able to accomplish the same level of high quality research over the course of my PhD were it not for the tireless efforts of my bright and energetic mentees Jen Sun, Rudy, Raymond, and Michael. Jen, my first undergraduate mentee, worked side by side with me to generate protein fibers and travel up to the City College of New York to learn confocal microscopy and image fluorescent fibers. I will never forget turning off the lights in the microscopy room and visualizing glowing protein fibers with Jen for the first time, truly feeling the sense of scientific discovery. Over the years Jen and I not only became a great team in the lab, but also formed a friendship I hold very dear. Rudy, who worked with me while pursuing his Masters degree, was an instrumental member of the research team and I thank him very much for his dedication. Hard working and thoughtful, Rudy always proposed new directions for our project and boldly took on the challenge of doing an oral presentation at a NYAS conference with great success. Thank you to Raymond and Michael, who joined the project in the final stages and impressed me by quickly picking up new techniques and analyzing data with the expertise of students beyond their years.

Thank you to my advisor and mentor Professor Jin Kim Montclare. My time under Jin's tutelage has allowed me to better myself as a scientist and professional as a result of her constant guidance and support. The integrity she has for the research we do in our lab is admirable – tireless questioning of the data and encouragement to better our experiments and move forward with sound scientific reasoning is the type of leadership I have been privileged enough to learn from. I value that Jin recognizes the importance of her students' participation in arenas outside of the lab that benefit their education, and I thank her for fully supporting me in enriching endeavors such as entrepreneurship or teaching outside the lab as well. Jin has not only supported me in these activities, but constantly leads by example. She has also created a wonderful culture of diversity, intellectual curiosity, and rigorous research in our lab, teaching me that the world of scientific research is a collaborative one. Beyond science, I have drawn much inspiration from Jin as an educator, entrepreneur, and woman in STEM. Her ability to simultaneously devote herself to basic research, academic and entrepreneurial pursuits, and her family is something I strive to one day emulate. In moments of inspiration Jin is fond of quoting Sir Isaac Newton – "If I have seen further it is by standing on the shoulders of giants." For me, Jin and my lab mates have indeed been my giants.

Thank you to the professors that have dutifully served on my PhD committee for the past five years; Professor Rastislav Levicky, Professor Cenk Sumen, and Professor Kalle Levon. Their technical expertise has allowed me to take this work into directions I would not have thought of on my own and your guidance on many occasions has been instrumental in helping me stay focused on what is truly important, both in science and as a young professional. I thank the scientists I have had the fortune of collaborating with at the Brookhaven National Laboratory and at the City College of New York. Dr. Kim Kisslinger allowed me to master transmission electron microscopy and helped me ask the right questions, Dr. Dmytro Nykypanchuk assisted with atomic force microscopy, and Professor Lane Gilchrist showed me how incredible confocal microscopy can be. These people were not obligated to be so helpful and kind, and yet were consistently ready to help a young scientist by sharing their knowledge with me. Their enthusiasm for discovery has been infectious, and I thank them all very much.

I spent two years as a teaching fellow in the AMPS/CBRI program as a part of the NSF GK12 fellowship. During this time I received guidance on becoming a better educator from Professor Vikram Kapila, Professor Magued Iskander, and Ben Esner. Many thanks to Susan Hermon, who helped not only the fellows but also the teachers, professors, and students we worked with as well. Thank you to my peers in the fellowship program, who became my friends as we struggled through fellows meetings, Teach Engineering lesson plans, and many hours spent at robotics tournaments. Thank you Jeff, Eduardo, Ursula, Keeshan, Carlo, and Jennifer. GK124L! Thank you to the remarkable teachers at the Community Partnership Charter School, Ivonne Cintron and Maggie Cho, and William Principal at Mott Hall Bridges Academy. Their students are truly lucky to have them as their educators and role models and I consider myself fortunate to have worked alongside you in getting young scholars excited about STEM. The bright students at CPCS and MHBA that taught me the greatest lessons I learned while a teaching fellow, and I thank them and wish them continued strength along their journey in education. Teaching them, and learning from them, makes me proud to be a product of NYC public school system and hopeful for the future of our city.

Over the course of my doctorate I have engaged in many entrepreneurial activities, including cofounding a company with my friend and peer Ching Yao. Thank you Ching Yao for embarking upon that wild adventure with me. Thank you to Carlo and Joe being my partners in the early stages of an academic spinoff company, whom I wish much continued success. Also thank you to my wonderful teachers, mentors, and friends who helped me open my eyes to the world of entrepreneurship: Frank Rimalovski, Lindsay Marshall, Risa Cohn, and Adam Goulburn.

I have had no shortage of support inside classrooms and laboratories, which has been paralleled by the support of my friends and family outside of school. My best friends Anais, Terri, and Caitlin helped me push through the hard times and have been there to celebrate the triumphs. I thank you for somehow never doubting the fact that I would get my doctorate, even when I thought it was impossible. The good times with you have made the difficult ones bearable.

Thank you to my parents Cameron and Rigmor, who managed to instill in their daughters that no matter how many degrees we obtain we are never actually done learning. They pushed me without stressing me, and have always been there to support me through major life decisions. Thank you in particular to my mother, who allowed me to move back home with her and Parker when I began my doctorate and came (all the way) to Brooklyn on numerous occasions just to join me for dinner while I was waiting for my bacteria to grow. I would not have come this far had they not believed in me, and I thank them deeply for that.

My sisters Laura, Heather, and Ivy have been my partners in life, and I thank them for everything they have done to help me learn, cope, and love my entire life. When I joined the GK12 fellowship and engaged in discussions about the lack of women in STEM it struck me that I have three sisters in highly technical fields: conservation biology, obstetrics and gynecology, and architecture. I do not doubt that our constant and unwavering support for each other has given me the strength to accomplish most anything. My sisters are quick to help me see the light in tough situations and always put things in perspective. They helped me prepare for my defense by critiquing my presentation and of course vetting the outfit I chose to wear. Thank you Laura, Heather, and Ivy – I am who I am today because of the wonderful women you all are.

Thank you to my dog, Pastas, for being a wonderful stress reliever as my defense approached and for always having a tail wag and a kiss for me when I need it most. You managed to rescue me when I thought I was rescuing you.

Lastly, thank you to my amazing boyfriend Kevin. You have answered my tearful calls en route to Brownsville at 6:30 am in the sleet and rain, jumped on red-eye flights to New York every few weeks to spend time with me, taken care of me in California while I was thesis writing, video chatted to help prepare my thesis defense… the list goes on. Whenever I have needed you, you have always been there to cheer me on with unfaltering optimism. I promise to do my best to always be there for you the same way you have supported me these past few years.

Onwards, and upwards! Thank you all.

*Dedication*

I dedicate this work to the intelligent, driven, and courageous women that are my family. My grandmothers Margaret and Maj, mother Rigmor, and sisters Laura, Heather, and Ivy have provided me with the most incredible role models that a person can ask for. Their smarts, humor, kindness, and fearlessness have inspired me to never turn away from any challenge I deemed worthy.

Ni har skapat mig att bli kvinnan jag är idag, och jag älskar er med hela min hjärta.

ABSTRACT

# Self-assembling Protein Materials for Metal Nanoparticle Templation

by

Jasmin Hume

Advisor: Jin Kim Montclare, PhD

Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy (Materials Chemistry)

Jasmin Hume

May 2015

The fabrication of *de novo* proteins able to self-assemble on the nano- to meso- length scales is critical in the development of protein-based biomaterials in nanotechnology and medicine. Proteins that template and stabilize nanoparticles can serve as agents in biodetection and imaging, as well as serving in applications in catalysis and biosensing. Here we report the design and characterization of protein-engineered coiled-coils that not only assemble into microfibers but also can bind hydrophobic small molecules and template inorganic nanoparticles. Under ambient conditions proteins form fibers with nanoscale structure possessing large aspect ratios formed by bundles of $\alpha$-helical homopentameric assemblies, which further assemble into mesoscale fibers in the presence of curcumin through aggregation. The same protein materials have been used to template gold nanoparticles (AuNPs) and iron oxide nanoparticles (MNPs) *in situ*, leading to very distinct assemblies and properties that are largely dependent upon the protein secondary structure. Genetic engineering, as well as protein engineering techniques such as enzymatic site-specific cleavage, unnatural amino acid incorporation, and covalent binding of orthogonal peptides have been employed in this work to fully manipulate and control physical properties of self-assembling protein materials and direct the templation of metal nanoparticles thereupon.

# Table of Contents

1 Self-assembling protein fibers ..... 1

1.1 Abstract ..... 1

1.2 Introduction ..... 1
1.2.1 Bio-inspired self-assembling proteins ..... 1
1.2.2 α-helical coiled-coil proteins and their fibers ..... 5
1.2.3 Control of fiber assembly ..... 16
1.2.4 COMPcc ..... 25
1.2.5 Design methodology ..... 28
1.2.6 Applications ..... 35

1.3 Materials and methods ..... 37
1.3.1 Materials ..... 37
1.3.2 Genetic engineering ..... 37
1.3.3 Protein expression ..... 42
1.3.4 Conditions for curcumin binding ..... 44
1.3.5 Circular dichroism ..... 44
1.3.6 Nuclear magnetic resonance ..... 46
1.3.7 ATR-FTIR ..... 46
1.3.8 TEM ..... 47
1.3.9 Zeta potential ..... 47
1.3.10 Dynamic light scattering ..... 48
1.3.11 Scatchard binding analysis ..... 48
1.3.12 SEM ..... 49
1.3.13 AFM ..... 49
1.3.14 Fluorescence microscopy ..... 50
1.3.15 Confocal microscopy ..... 50

1.4 Results and discussion ..... 50
1.4.1 Genetic engineering ..... 50
1.4.2 Protein expression and purification ..... 57
1.4.3 Modeling ..... 59
1.4.4 Secondary structure and thermodynamic stability ..... 63
1.4.5 Fiber morphology ..... 77
1.4.6 Curcumin binding ..... 89

1.5 Conclusions ..... 106

2 Tunable Conformation-Dependent Protein·Gold Nanoparticle Nanocomposites ..... 108

2.1 Abstract ..... 108

2.2 Introduction ..... 109
2.2.1 Gold nanomaterials ..... 109
2.2.2 Biomineralization ..... 111
2.2.3 Impact of protein structure on assembly and templation ..... 118
2.2.4 Design methodology ..... 128
2.2.5 Applications ..... 129

2.3 Materials and methods ..... 131

2.3.1 Materials ..... 131
2.3.2 Expression, purification, and cleavage ..... 132
2.3.3 Gold binding ..... 134
2.3.4 Absorbance ..... 134
2.3.5 Circular dichroism ..... 134
2.3.6 ATR-FTIR ..... 134
2.3.7 TEM ..... 135
2.3.8 Cyclic voltammetry and impedance ..... 135
2.4 Results and discussion ..... 136
2.4.1 Histidine tag cleavage ..... 137
2.4.2 Absorption profiles ..... 138
2.4.3 Nanocomposite morphology and elemental characterization ..... 139
2.4.4 Secondary structure ..... 147
2.4.5 Triggered assembly ..... 150
2.4.6 Electrochemical Characterization of Protein·AuNP Assemblies ..... 155
2.5 Conclusions ..... 160
3 Magnetite templation on functionalized protein biomaterials ..... 162
3.1 Abstract ..... 162
3.2 Introduction ..... 162
3.2.1 Ferric oxide nanomaterials ..... 162
3.2.2 Biologically generated magnetic particles ..... 165
3.2.3 Applications of magnetic nanomaterials ..... 173
3.2.4 Design methodology ..... 179
3.3 Materials and methods ..... 188
3.3.1 Materials ..... 188
3.3.2 Synthesis of L-azidohomoalanine ..... 190
3.3.3 Expression with unnatural amino acid incorporation ..... 192
3.3.4 ImageQuant analysis ..... 195
3.3.5 Click chemistry ..... 195
3.3.6 Ferric oxide nanoparticle templation ..... 197
3.3.7 Optical microscopy ..... 198
3.3.8 Circular dichroism ..... 199
3.3.9 TEM and electron diffraction ..... 199
3.3.10 Magnetometry ..... 200
3.4 Results and discussion ..... 200
3.4.1 Protein expression via unnatural amino acid incorporation ..... 200
3.4.2 Click chemistry ..... 204
3.4.3 Circular dichroism ..... 213
3.4.4 TEM ..... 215
3.4.5 Optical microscopy ..... 223
3.4.6 Magnetometry ..... 228
3.5 Conclusions and future work ..... 230
4 References ..... 232
5 Appendices ..... 254
5.1 Publications ..... 254

5.1.1 Hume et al. Biomacromolecules 2014 ..... 254
5.1.2 Hume et al. Biomacromolecules 2015 ..... 262

5.2 Abbreviations ..... 270

5.3 Source of materials ..... 271

5.4 Plasmid maps and DNA sequences ..... 273
5.4.1 COMPcc$^{s}$ ..... 273
5.4.2 C ..... 275
5.4.3 Q ..... 279
5.4.4 L ..... 283

5.5 Additional information on proteins of interest ..... 287
5.5.1 DSAg and DSD ..... 287
5.5.2 PA ..... 288
5.5.3 α-amylase ..... 288
5.5.4 Sericin silk protein ..... 288
5.5.5 BP124NAB peptide ..... 289
5.5.6 Helical gelator ..... 290
5.5.7 Cty c ..... 290
5.5.8 Bovine serum albumin ..... 290

5.6 Experimental protocols ..... 291
5.6.1 Preparation of chemically competent cells ..... 291
5.6.2 Expression in XL1blue ..... 292
5.6.3 Expression in M15MA with incorporation of azidohomoalanine ..... 293
5.6.4 Purification ..... 295
5.6.5 Dialysis ..... 296
5.6.6 BCA, Enhanced Protocol ..... 297
5.6.7 Running SDS-Page gels ..... 298
5.6.8 Chemical synthesis of L-azidohomoalanine ..... 298

5.7 Reagent recipes ..... 300

# List of Figures

Figure 1.1. Protein assemblies and their biological representatives.[9] ..... 2

Figure 1.2. Transmission electron micrographs of designed two-component protein cages, protein components being T32-28 (a), T33-09 (b), T33-15 (c), T33-21 (d) and T33-28 (e). Insets represent the two class averages of the particles (top and bottom) for the actual particles (left) and computational models (right).[14] ..... 4

Figure 1.3. Mesoscale protein fibers. SEM of the byssal threads of *Tridacna maxima* (a),[20] TEM of collagen fibers (b),[21] SEM of natural spider silk (c),[22] and SEM of spun spider silk made recombinantly in mammalian cells (d).[23] Scale bar is 100 μm in (a), 50 nm in (b), and fibers are approximately 15 μm in width in (c) and 20 μm in (d). ..... 4

Figure 1.4. Dimeric coiled-coils. Helical wheels for parallel dimeric coiled-coil (a), and antiparallel dimeric coiled-coil (b),[31] where hydrophobic interactions occur between the *a* and *d* residues and ionic interactions between *g/e* (a) and *e/e or g/g* residues (b). Amphipathic assembly of a coiled-coil dimer (c).[32] ..... 6

Figure 1.5. Orthogonal views of several coiled-coil structures: parallel dimer (PDB 2ZTA) (a), antiparallel dimer (PDB 1HF9) (b), parallel trimer (PDB 1BB1) (c), and parallel pentamer (PDB 2GUV) (d).[38] .. 7

Figure 1.6: Protein fiber assembly by sticky-end coiled-coil dimers. (a) From left to right: single $\alpha$-helical protein, blunt-end protein assembly, and sticky-end protein assembly. (b) TEM micrograph of fiber formation due to sticky-end assembly.[34] ..... 8

Figure 1.7. Schematic representation of pentameric fibril created by assembly of staggered $\alpha$-helices (a) and the structural model of the same (b).[47] ..... 11

Figure 1.8. Helical coiled-coil assemblies bundle to form fibers. Sequence of collagen mimetic peptide which forms homotrimers (a),[21] whose fiber structure was solved using X-ray diffraction (b), quasihexagonal packing with 1.2 nm protofibrils, corresponding to the cross-section of the triple helix (c).[21] Structural model of the four-stranded nanofibrils of *Tridacna maxima* (d), nanofibril cross-section (e), and bundling of nanofibrils to form mesoscale fibers (f) through hexagonal packing (g).[20] ..... 12

Figure 1.9. Electrostatic assembly of dimeric coiled-coils by SAFs with electrostatic charge patterns of "+ + - -" (a)[40] and "+ - - +" (b, c),[48] where both assemblies take advantage of electrostatic interactions and sticky ends. Blue represents positively charged heptads, red represents negatively charged heptads. Sequence of the MagicWand (b) leads to efficient self-assembly (c).[48] ..... 14

Figure 1.10. 3D domain swapped DSAg protein designed to form fibrils (a), demonstrated by TEM under neutral conditions (pH 6.5) (b) and acidic conditions (pH 2.4) (c). Inset of (c) shows clear protofibril formation.[49] ..... 15

Figure 1.11. Self-assembly of CMA with respect to temperature cycling, demonstrating thermoreversibility of storage modulus *G'* between 4 and 37 °C (a), fiber formation before melting (b, d) and upon cooling (c, e).[73] ..... 21

Figure 1.12. Molecular structure of curcumin. ..... 22

Figure 1.13. Crystal structure of COMPcc depicts a homopentamer 73 Å in length (a), a hydrophobic pore 2-6 Å in width (viewed from the C-terminal end of the protein) (b), and a C-terminus with serine residues shown in black creating inter-chain disulfide bonds (c). The primary amino acid sequence of COMPcc$^s$ with respective heptad positions as well as the corresponding residue numbers is shown in (d). ..... 26

Figure 1.14. COMPcc complexed with two molecules of vitamin $D_3$, one in each binding pocket. The ion channel within COMPcc is divided in two cavities, with residue Q54 defining the separation.[94] ..... 27

Figure 1.15. Palmitic acid binding inside (a) and outside (b) the hydrophobic channel of COMPcc.[93] ..... 28

Figure 1.16. Helical wheel for C, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K43 and K49, in the

*f* and *e* positions, respectively) and arginine residues (R22 and R39, in the *f* and *b* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E23, E26, and E36 in the *g*, *c*, and *f* positions, respectively) and asparagine residues (D33 in the *c* position). Individual $\alpha$-helices self-assemble to form homomeric pentamers. ..... 30

Figure 1.17. Helical wheel for Q, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K18 and K24, in the *f* and *e* positions, respectively) and arginine residues (R45 and R49, in the *e* and *b* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E33, E36, and E46 in the *g*, *c*, and *f* positions, respectively) and asparagine residues (D43 in the *c* position). Individual $\alpha$-helices self-assemble to form homomeric pentamers. ..... 31

Figure 1.18. Sticky-ended fiber formed by longitudinal assembly of C (red) and Q (black). Slashes denote end/beginning of protein subunits. ..... 31

Figure 1.19. Longitudinally arranged pentameric assemblies comprised of 4 Q helices and 1 C helix (a, b) and 4 C helices and 1 Q helix (c, d), viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively. ..... 32

Figure 1.20. Cartoons of pentameric assemblies comprised of 2 Q helices and 3 C helices (a, b) and 3 Q helices and 2 C helices (c, d), viewed down the pentamer axis. In these configurations two distinct arrangements of the subunits are possible, which are referred to as A when C and Q are beside one another in the pentamers and B when they are interspersed. C and Q are represented in blue and orange, respectively. ..... 32

Figure 1.21. Longitudinally arranged pentameric assemblies comprised of 2 C helices and 3 Q helices in the A (a, b) and B (c, d) configurations, viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively. ..... 33

Figure 1.22. Longitudinally arranged pentameric assemblies comprised of 3 C helices and 2 Q helices in the A (a, b) and B (c, d) configurations, viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively. ..... 33

Figure 1.23. Helical wheel for L, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K29 and K35, in the *f* and *e* positions, respectively) and arginine residues (R25 and R43, in the *b* and *f* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E22, E30, and E47 in the *f*, *g*, and *c* positions, respectively) and asparagine residues (D19 in the *c* position). Individual $\alpha$-helices self-assemble to form homomeric pentamers. ..... 34

Figure 1.24. DNA gel of amplified fragments for C, L1, L2, Q1, and Q2 on a 2 % agarose gel. Each fragment is located at its expected base pair length. ..... 51

Figure 1.25. PCR assembled L (a) and Q (b) DNA on a 2 % agarose gel. The expected length of these fragments is 142 bp. ..... 52

Figure 1.26. PCR amplified L and Q DNA on a 2 % agarose gel. The expected length of these fragments is 142 bp. ..... 53

Figure 1.27. Restriction DNA gel for PheRS** plasmid and restricted C fragment on a 2 % agarose gel. ..... 53

Figure 1.28. Restriction DNA gel for PheRS**/C DNA of uncut (U) and double cut (D) samples on a 1 % agarose gel. ..... 54

Figure 1.29. Restriction of C, L, and Q DNA. The expected length of these fragments is 142 bp on a 2 % agarose gel. ..... 55

Figure 1.30. Restriction DNA gel for PheRS**/insert DNA for six samples of C, Q, and L. C1-5, L5, L6, and Q2-6 were ligated and transformed. ..... 55

Figure 1.31. + and –φ ladder ligation samples. Solid band in the + control samples proves successful ligation was achieved. ..... 56

Figure 1.32. 12 % SDS-PAGE protein gels for pre/post expression of six samples of C, L, and Q. Lanes in (a): 1. Ladder, 2. C1 pre, 3. C1 post, 4. C2 pre, 5. C2 post, 6. C3 pre, 7. C3 post, 8. C4 pre, 9. C4

post, 10. C6 pre. (b): 1. Ladder, 2. L1 pre, 3. L1 post, 4. L2 pre, 5. L2 post, 6. L3 pre, 7. L3 post, 8. L4 pre, 9. L4 post, 10. L5 pre. (c): 1. Ladder, 2. Q1 pre, 3. Q1 post, 4. Q2 pre, 5. Q2 post, 6. Q3 pre, 7. Q3 post, 8. Q4 pre, 9. Q4 post, 10. C6 post. (d): 1. Ladder, 2. C5 pre, 3. C5 post, 4. L6 pre, 5. L6 post, 6. Q6 pre, 7. Q6 post, 8. Q5 pre, 9. Q5 post, 10. L5 post. ..... 56

Figure 1.33. 12 % SDS-PAGE protein gels for pre/post expression of C and Q (a), and L and wt (b). Strong bands in the post expression lanes can be observed for all proteins, where the MW of C, Q, and L is 6.3 kDa and the MW of wt is 6.9 kDa. ..... 58

Figure 1.34. 12 % SDS-PAGE protein gels for purification of C (a), L (b), Q (c), and wt (d) by increasing imidazole gradient. Pure protein was typically eluted with imidazole concentrations of 100-500 mM (lanes 6-9). Lanes correspond to the following in gels (a) – (d): 1: ladder, 2: supernatant, 3: flowthrough, 4: wash 1, 5: wash 2, 6: elution 1, 7: elution 2, 8: elution 3, 9: elution 4, 10: elution 5, 11: elution 6, 12: elution 7, 13: elution 8, 14: elution 9. ..... 58

Figure 1.35. Helical subunits of wt (a, b), C (c, d), and Q (e, f) viewed along (a, c, e) and down (b, d, f) the helical axis. ..... 60

Figure 1.36. Pentameric assemblies of wt (a, b), C (c, d), and Q (e, f) viewed along (a, c, e) and down (b, d, f) the pentamer axis. ..... 61

Figure 1.37. Surface charge representation of wt (a and b), C (c and d), and Q (e and f) pentamers under acidic conditions viewed along (a, c, e) and down (b, d, f) the pentamer axis. Positive blue patches are attributed to solvent-exposed lysine and arginine residues. Negative red patches are attributed to solvent-exposed glutamate and aspartate residues. ..... 62

Figure 1.38. Secondary structure of proteins in 50 mM PB at 25 °C at pH 4 (a), 8 (b), and 10 (c). Circular dichroism variable wavelength scans of wt (solid line), C (dashed line), Q (dotted line), and L (dashed, dotted line) at pH 4 (a), 8 (b), and 10 (c). Data is averaged from at least two replicates and was obtained with 10 µM protein concentration. ..... 63

Figure 1.39. $\alpha$-helical (black), $\beta$-sheet (grey), and random coil (white) content for wt (a), C (b), Q (c), and L (d) at pH 4, 8, and 10 calculated from wavelength scans in Figure 1.38. ..... 65

Figure 1.40. ATR-FTIR of 10 µM Q at pH 4 (a), pH 8 (b), and pH 10 (c) measured at room temperature. 66

Figure 1.41. Secondary structure represented by circular dichroism variable wavelength scans of mixtures of C and Q of proteins at pH 4 (a), pH 8 (b), and pH 10 (c). Data is averaged from at least two replicates and was obtained with 50 mM PB and 10 µM protein concentration at 25 °C. ..... 67

Figure 1.42. Ratio of $\theta_{222}$ to $\theta_{min}$ for mixtures of C and Q at various pH conditions. ..... 68

Figure 1.43. Secondary structure represented by circular dichroism variable wavelength scans of mixtures of C (black), Q (grey), and an equimolar mixture of the two (red) at pH 8 (a), and pH 10 (b) in the presence (dashed lines) and absence (solid lines) of 25 v/v % TFE at 25 °C. Proteins in the absence of TFE were in a 10 mM PB solution and at a concentration of 13.3 µM. Scans in the presence of TFE were in a 10 mM PB with 25 v/v % TFE at a concentration of 10 µM. Ratio of $\theta_{222}$ to $\theta_{min}$ for C, Q, and the equimolar mixture at pH 8 and 10 in the presence and absence of TFE (c). ..... 69

Figure 1.44. Thermal melt scans at 222 nm for wt (a), Q (b), and C (c) at pH 8. Melts were run in the forward, 20 to 85 °C (solid lines), and reverse, 85 to 20 °C (dashed lines), directions to confirm reversibility. Complete reversibility allowed for the van't Hoff analysis to be performed. Thermal melt scans at 222 nm for wt (a) and Q (b) at pH 8 at different scan speeds. Scan speeds of 1 (solid lines), 2 (dashed lines), and 5 (dotted lines) °C/min were run in the forward direction for both proteins. ..... 70

Figure 1.45. Melt reversibility for 20 µM C (a) and Q (b) at 50 mM PB pH 8. ..... 71

Figure 1.46. Thermal melt isotherms for 10 µM L at 50 mM PB pH 4, where the gradient in the MRE signal at 222 nm was not large enough over this temperature range for determination of thermodynamic properties. Wavelength scans were performed at 25 °C (black), 45 °C (red), 55 °C (blue), 65 °C (green), and 85 °C (yellow). ..... 71

Figure 1.47. Thermal melt scans at 222 nm for wt (a), Q (b), and C (c) at pH 4 (solid lines), pH 8 (dotted lines), and pH 10 (dashed lines). All scans represent an average of two trials. ..... 72

Figure 1.48. SEM images of an equimolar mixture of C and Q at in 50 mM PB pH 8. Scale bars in (c) and (e) represent 300 and 100 nm, respectively, in (a), (c), and (f) represent 1 μm, and in (b) and (d) represent 200 nm. ..... 78

Figure 1.49. Equimolar mixture of C and Q at 10 mM PB pH 8 on a mica surface at room temperature. Heat map indicates topographical height of features on the mica surface, where a fiber on the surface is displayed in white in (a), corresponding to a height of about 3 nm. The 3 dimensional rendering of topographical scan of equimolar mixture of C and Q on mica surface is shown in (b). ..... 79

Figure 1.50. Equimolar mixture of C and Q at 10 mM PB pH 8. 3 dimensional rendering of topographical scan of equimolar mixture of C and Q on mica surface shown in (a). Fiber with adjacent protofibril can be seen in the scans in (a) and (b). Heat map indicates topographical height of features on the mica surface, where a fiber on the surface is displayed in white in (b). Line scan of the fiber is shown in (c). ..... 80

Figure 1.51. Transmission electron micrographs of wt fibers, 10 μM, pH 8, 10 mM PB at room temperature. Scale bar in (a) represents 5 μm and in (b) and (c) represents 200 nm. ..... 81

Figure 1.52. Transmission electron micrographs of sheets formed by equimolar mixture of C and Q, 10 μM, pH 8, 10 mM PB at room temperature. Scale bar in (a) represents 50 nm and in (b) represents 20 nm. ..... 82

Figure 1.53. Transmission electron micrographs of equimolar mixtures of C and Q, 20 μM, 10 mM PB pH 8 in the presence of 25 v/v % TFE at room temperature. Distinct fibers could be seen, with protofibrils. Scale bars represent 0.5 μm in both (a) and (b). ..... 82

Figure 1.54. Transmission electron micrographs of equimolar mixtures of C and Q, 5 μM, 10 mM PB at various pH conditions: pH 4 (a)-(c), pH 8 (d)-(f), and pH 10 (g)-(i) at room temperature. Distinct fibers could be seen at pH 4. Distinct fibers were not seen at pH 8, although striations were observed that had a regularity of approximately 3 nm. Distinct fibers were seen at pH 10, and striations were observed that had a regularity of approximately 3 nm. Scale bars represent 0.5 μm in (a), 100 nm in (b), 50 nm in (c), (f), and (i), 200 nm in (d) and (h), 20 nm in (e), and 2 μm in (g). ..... 83

Figure 1.55. Transmission electron micrographs of equimolar mixture of C and Q, 20 μM, 50 mM PB pH 8 at room temperature. Distinct fibers could be seen, with visible protofibrils. Scale bars represent 1 μm in (a), 500 nm in (b), 200 nm in (c), and 0.5 μm in (d). ..... 84

Figure 1.56. Transmission electron micrographs of C (a)-(d) and Q (e)-(f), 20 μM, 10 mM PB pH 8 at room temperature. Distinct fibers could be seen, with protofibrils for C whereas striations could be seen in large sheet-like aggregates for Q under these conditions. Scale bars represent 20 nm in (a), 100 nm in (b) and (c), 200 nm in (d), 50 nm in (e), and 1 μm in (f). ..... 86

Figure 1.57. Microscopy and modeling of protein fibers. (a) Transmission electron micrograph of Q fiber, 10 μM, pH 4. (b) Schematic representation of Q fiber assembly with staggered positive (red) and negative (blue) regions of the pentamer. ..... 87

Figure 1.58. Histogram of Q pH 4, 50 mM PB protofibril diameters ($n$ = 210) (a) and fiber diameters ($n$ = 14) (b). ..... 88

Figure 1.59. Transmission electron micrographs of protein fibers and sheets at room temperature. Transmission electron micrographs Q protein fibers at pH 4, 50 mM PB (a) – (o). A fibrous network is formed by wt (p) at pH 8, 10 mM PB. L (q) forms sheet-like assemblies and larger aggregates at pH 4, 50 mM PB conditions. Scale bars in (a), (d), (e), (l), (o) represent 50 nm, in (b), (c), (i), and (n) represent 100 nm, (m) and (p) represent 200 nm, (f), (h), (k), and (q) represent 0.5 μm, (j) represents 2 μm, and (g) represents 5 μm. ..... 88

Figure 1.60. Absorption spectra of aqueous solutions containing 20 μM protein with varying concentrations of curcumin, from 0 to 160 μM at pH 8. Absorption spectra of wt and curcumin are presented in (a), C and curcumin in (b), and Q and curcumin in (c). Black lines are the protein in the absence of curcumin, red is 20 μM curcumin, green is 40 μM, purple is 60 μM, blue is 80 μM, yellow is 100 μM, green is 120 μM, and grey is 160 μM. ..... 90

Figure 1.61. Linear plot for 1/ΔA vs. inverse curcumin (ccm) concentration. ..... 91

Figure 1.62. Linear plot for 1/r vs. inverse curcumin (ccm) concentration. ..... 91

Figure 1.63. Helicity of wt and Q increase as a result of increasing curcumin concentration, pH 4 50 mM PB at room temperature. Circular dichroism wavelength scans of wt (a) and Q (b) with varying molar ratios of curcumin. Legend represents curcumin concentration in μM. Inserts show relationship between negative MRE at 222 nm of the mixtures with increasing curcumin/protein molar ratio, displaying near linear fits over the examined range. ..... 93

Figure 1.64. Fourier transform self-deconvoluted spectra of Q in 50 mM PB pH 4 in the presence of curcumin at a 5:1 molar ratio. Spectrum represents the average of two trials. ..... 94

Figure 1.65. Fluorescence microscopy of an equimolar mixture of C and Q with embedded curcumin at 50 mM PB pH 4 (a), pH 8 (b), and pH 10 (c). Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Scale bars omitted due to software limitations. Insets in (a) and (b) show fiber networks which were observed, and inset of (c) shows a protein fiber at pH 10 displaying particularly bright spots along the fiber. ..... 95

Figure 1.66. Fluorescence microscopy of curcumin-embedded C (a) and Q (b) protein fibers at pH 8 at a curcumin to protein molar ratio of 5:1. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Scale bars omitted due to software limitations. ..... 95

Figure 1.67. Confocal microscopy of an equimolar mixture of C and Q with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Fiber widths in (a) range from 18 - 21 μm, with the fiber in (b) possessing a width of 15.2 μm, (c) 12.8 μm, and (d) 7.4 μm.Scale bars represent 200 μm (a), 50 μm (b-c), and 100 μm (d). ..... 96

Figure 1.68. Histogram of fiber diameters in the presence of 100 μM curcumin as viewed in confocal microscopy for fibers of C (black), Q (white), and an equimolar mixture of C and Q (grey). ..... 97

Figure 1.69. Confocal images of C (a-c), Q (d-f), and wt (g-i) with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. ..... 98

Figure 1.70. Orthogonal views of wt with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Yellow crosshairs pinpoint the location of the orthogonal slices within the fiber. The presence of a hollow cavity can be seen in these orthogonal views. ..... 100

Figure 1.71. Confocal microscopy fluorescence images of several Q protein fibers in the presence of 50 μM curcumin at pH 4, 50 mM PB at room temperature. Scale bars in (d), (k), and (l) represent 10 μm, (a)-(c) and (e)-(j) represent 20 μm, and (b) and (l)-(o) represent 30 μm. Panels (p)-(r) originate from a single z-slice (119 of 227) within the XYZ data set from panel (i). The image line profiles of the confocal (red trace) and interference contrast (black trace) resulting from the yellow sampling line are shown in panel (r), indicating the boundaries of the fiber and the curcumin fluorescence are coincident. ..... 101

Figure 1.72. (a) Reconstruction of 3D confocal XYZ data of 10 μM Q protein in the presence of 50 μM curcumin (1:5 molar ratio of protein:curcumin) at room temperature. (b) 3D representation of the same Q fiber showing XZ and YZ orthogonal views and an oblique slice cross section above the 3D bounding box. Scale bars in (a) and (b) represent 20 and 10 μm, respectively. ..... 102

Figure 1.73. Transmission electron micrographs of 10 μM Q in the presence of 50 μM curcumin (molar ratio of 5:1 curcumin:protein) at pH 4, 50 mM PB at room temperature. Large aggregates were seen. Scale bars represent 0.5 μm in main image and 100 nm in inset. ..... 103

Figure 1.74. 1D $^{1}H$ nuclear magnetic resonance scans for wt (a) and Q (b) in the absence (blue curves) and presence (red curves) of 100 μM curcumin (5:1 molar ratio of curcumin:protein). All protein concentrations were kept constant at 20 μM. Buffer conditions were 50 mM PB pH 4 with 1% (v/v) methanol and 1 % (v/v) $D_2O$. These protein fibers, both with and without curcumin, are not completely soluble but not crystalline and the methods commonly used to determine protein structure, e.g. nuclear magnetic resonance, provide a limited scope of data to interpret. ..... 103

Figure 1.75. Aggregation of Q protein fibers due to increasing curcumin concentration. (a) Zeta potential as a function of curcumin concentration for Q (white bars) and L (grey bars). (b) Dynamic light scattering measurements of count rate for Q (solid line) as a function of curcumin concentration.

Count rate was not obtained for L as signal at all curcumin concentrations was too low for detection. Absorbance measured at 420 nm as a function of curcumin concentration for Q (dashed line) at room temperature. Error bars in figures (a) and (b) represent an average of three trials..... 104

Figure 1.76. Aggregation of wt as a result of curcumin. (a) Zeta potential increases slightly as a function of increasing curcumin concentration for wt. (b) Count rate (solid line) and absorbance (dashed lines) at 420 nm of wt (black) and L (red) versus curcumin concentration. Absorbance measured at 420 nm as a function of curcumin concentration. As curcumin/protein molar ratio increases the absorbance increases correspondingly. Error bars in figures (a) and (b) represent an average of three trials. ..... 105

Figure 1.77. Schematic for fiber assembly. Schematic representation of fiber assembly of Q on various length scales. Amino acids form α-helices which assemble to form pentamers with approximate diameters of 3 nm. Protofibrils are formed by assembly of pentameric subunits and bundle together to create protein fibers ranging from tens to hundreds of nm in diameter, and upon addition of curcumin mesofibers with diameters on the micrometer scale are generated. ..... 106

Figure 2.1. UV-vis absorption of gold nanoparticles of various sizes and shapes (a).[137] UV-vis and TEM characterization of AuNPs synthesized chemically in various solvents: 100 % $D_2O$ (b), 50 % $D_2O/H_2O$ (c), and 100 % $H_2O$ (d).[138] ..... 110

Figure 2.2. Biomineralization of Au in *C. metallidurans*. TEM micrograph of a *C. metallidurans* cell containing a AuNP in the periplasmic space (a), SEM micrographs (b, c) and corresponding EDAX analysis (d) of biofilm containing mineralized Au.[145] ..... 112

Figure 2.3. Setup for measuring *in situ* conductivity of a biofilm composed of *Geobacter sulfurreducens*, grown over two gold electrodes separated by a non-conductive gap (a). TEM micrograph of biofilm and pili nanofilaments grown on the electrode surface (b) and the conductivity of the wild type biofilm compared to a control of buffer and PilA-deficient mutant strain (c).[146] ..... 113

Figure 2.4. Phospholipid tubules plated with nickel and aligned with a bar magnet (magnification of 120x) (a).[155] Conductive DNA nanowires constructed with AgNPs shown schematically between two gold electrodes (b) and AFM images of the nanowires (c).[156] ..... 114

Figure 2.5. Procedure of gold toning (a) used to obtain conductive amyloid fibers deposited on $Si_3N_4$ viewed by AFM (b) and TEM (c), where diameters of coated fibers were enlarged from 50 nm to 100 nm by increasing the silver enhancement time from 3 minutes (c, left) to 5 minutes (c, right).[149].... 117

Figure 2.6. Schematic of the peptide·AuNP nanotube complex (a) and TEM images of AuNPs templated outside and inside the nanotube at pH 8.0 (b).[160] AuNPs templated with peptides A3 (c) and Flg-A3 (d), where scale bars indicate 50 nm.[161] ..... 118

Figure 2.7. UV-vis spectra of citrate-stabilized AuNPs in the presence of varying volume fractions of native vs. denatured $\alpha$-amylase (a). TEM of AuNPs with 100 % structured $\alpha$-amylase (b), 50 % structured/denatured $\alpha$-amylase (c), and 100 % denatured $\alpha$-amylase (d).[164] ..... 119

Figure 2.8. Stern-Volmer quenching constant ($K_{SV}$) of BSA by AuNPs of different sizes (a) and $K_{SV}$ as a function of AuNP surface area (b).[177] ..... 121

Figure 2.9. Experimentally determined crystallographic packing of 20 nm AuNPs within the lysozyme protein crystal.[178] ..... 122

Figure 2.10. Electron tomography (a, b) of left-handed double helices of $\beta$-sheet $C_{12}$-$PEP_{Au}$ decorated with AuNPs and a schematic rendering of their assembly (c).[166] ..... 123

Figure 2.11. UV-vis spectra of E5 (a) and Flg-A3 (b) in the presence and absence of AuNPs.[162] ..... 124

Figure 2.12. Dependence of hydrodynamic radius of apoCyt c·AuNP complexes at a 10:1 (a, b) and 20:1 (c, d) molar ratio between pH 7.4 and pH 10.8. TEM images of the complexes show smaller particles at a molar ratio of 20:1 (d) compared to large aggregates at 10:1.[170] ..... 126

Figure 2.13. CD wavelength scans of MG/GA·AuNP nanocomposites in TFE (a, top) and $dH_2O$ (a, bottom). AFM scans (b, c left) and particle height analysis from cross-section evaluation of AFM scans (b, c right) of assemblies in TFE (b) and $dH_2O$ (c).[169] ..... 128

Figure 2.14. Amperometric measurements of a glucose biosensor with glucose oxidase-functionalized AuNPs on a multiwalled carbon nanotube (MWNT) electrode. *i-t* curves obtained at various concentrations of glucose (a) were translated into a calibration curve (b) for the biosensor.[191] ..... 130

Figure 2.15. Schematic illustration of protein and nanoparticle assembly upon gold templation. Protein units are illustrated in orange, his tags in purple, and AuNPs as grey spheres. (a) $\alpha$-helical proteins C and Q with his tags form ordered pentamer assemblies, and upon gold templation exhibit a decrease in helical content and template AuNPs forming large aggregates of disordered AuNP-embedded protein. Photo demonstrates the formation of precipitates in C·AuNP 24 h post templation. (b) Cleaved C and Q with his tags removed are unstructured, and upon gold templation exhibit an increase in helical content, surrounding AuNPs and stabilizing them in solution. Photo demonstrates characteristic purple color of protein-stabilized Cx·AuNP 24 h post templation. ..... 137

Figure 2.16. SDS-PAGE gel showing cleavage of histidine tags for both C and Q. Lanes (from left to right) are ladder, C protein (6.31 kDa), C with cleavage buffer (6.31 kDa), C after incubation with Factor Xa (4.45 kDa), Q protein (6.31 kDa), Q with cleavage buffer (6.31 kDa), and Q after incubation with Factor Xa (4.45 kDa). ..... 138

Figure 2.17. Absorbance of protein-AuNP complexes at 520 nm, where precipitation of C and Q after 1 d results in a dramatic decrease in absorbance and Cx and Qx samples have absorbances that stabilize after 2 days and cease to decrease in value. ..... 138

Figure 2.18. Photos of protein templated with AuNPs, taken immediately after templation. C (a) and Q (c) have a dark purple hue, similar to that seen in the absence of any protein (e). Cleaved proteins Cx (b) and Qx (d), on the other hand, have a pinkish hue that is maintained over a period of 8 days. ..... 139

Figure 2.19. Transmission electron micrographs aggregates formed by C (a), fibers of Q (b), and sheet-like structures seen from cleaved proteins Cx (c) and Qx (d). Scale bars are 1 µm in (a) and (b), 200 nm in (c) and 0.5 µm in (d). ..... 140

Figure 2.20. Transmission electron micrographs for BSA·AuNP in 50 mM PB pH 8 at various magnifications. Scale bars are 200 nm in (a) and 50 nm in (b). ..... 141

Figure 2.21. Transmission electron micrographs for wt·AuNP in 50 mM PB pH 8 at various magnifications. Scale bars are 0.5 µm in (a), 5 µm in (b), 100 nm in (c), and 20 nm in (d). ..... 142

Figure 2.22. Histogram of AuNP diameters templated on BSA (grey) and wt (black). ..... 142

Figure 2.23. Transmission electron micrographs for C·AuNP (a), Q·AuNP (b), Cx·AuNP (c), Qx·AuNP (d), and AuNP in the absence of protein (in phosphate buffer) (e). Scale bars are 50 nm in (a), (c), (d), and (e) and 100 nm in (b). ..... 143

Figure 2.24. Histogram of nanoparticle sizes shows that cleaved proteins template smaller AuNPs than 6-His tagged proteins. ..... 144

Figure 2.25. Transmission electron micrographs and elemental maps for C (a), Cx (b), Q (c), and Qx (d) with templated AuNPs. Area outlined in red was selected for data acquisition for elemental maps in Figure 2.26. Scale bars in TEM micrographs are 50 nm in (b), 100 nm in (a), (c), and (d). ..... 145

Figure 2.26. Elemental maps of C (a), Cx (b), Q (c), and Qx (d) with templated AuNPs. Maps resulted from the area outlined in red in corresponding TEM micrographs in Figure 2.25. Maps display location of signals resulting from C (purple), N (pink), oxygen (lime green), Na (rust), P (neon green), and Au (orange). Scale bars in EDAX maps are 100 nm in (a), (b), and (d) and 200 nm in (c). ..... 145

Figure 2.27. EDS spectra from C·AuNP (a), Q·AuNP (b), Cx·AuNP (c), and Qx·AuNP (d). ..... 146

Figure 2.28. Secondary structure via circular dichroism, 50 mM PB pH 8 10 µM protein concentration. Wavelength scans of 6x-His and cleaved proteins with C, Cx in (a) and Q, Qx in (b). Wavelength scans of 6x-His and cleaved proteins with templated AuNPs with C·AuNP, Cx·AuNP in (c) and Q·AuNP, Qx·AuNP in (d). ..... 148

Figure 2.29. Fourier transform self-deconvoluted spectra of proteins in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials. ..... 149

Figure 2.30. Fourier transform self-deconvoluted spectra of proteins after gold templation in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials. ..... 149

Figure 2.31. Secondary structure via circular dichroism, pH 8 10 µM protein concentration. Wavelength scans of 6x-His and cleaved proteins in the presence of 25 v/v % TFE, with C, Cx in (a) and Q, Qx in (b). Wavelength scans of 6x-His and cleaved proteins with templated AuNPs in the presence of 25 v/v % TFE, with C·AuNP, Cx·AuNP in (c) and Q·AuNP, Qx·AuNP in (d). ..... 150

Figure 2.32. $\alpha$-helical (black), $\beta$-sheet (grey), and random coil (light grey) content for C and Cx in the presence and absence of 25 v/v % TFE and with and without templated AuNPs (a) and Q and Qx in the presence and absence of 25 v/v % TFE and with and without templated AuNPs (b). K2D analysis of data from Figure 2.28 and Figure 2.31 ..... 151

Figure 2.33. Transmission electron micrographs for proteins in the presence of 25 v/v % TFE, C + TFE (a), Q + TFE (b), Cx + TFE (c), and Qx + TFE (d). Scale bars are 200 nm in (a), (c), and (d) and 2 µm in (b). ..... 152

Figure 2.34. Fourier transform self-deconvoluted spectra of proteins in the presence of 25 v/v % TFE in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials. ..... 153

Figure 2.35. Fourier transform self-deconvoluted spectra of proteins in the presence of 25 v/v % TFE and after gold templation in 50 mM PB 8 for C (a) and Q (b). Data was not obtained for Cx·AuNP and Qx·AuNP with TFE due to problems with sample measurement. Each spectrum represents the average of two trials. ..... 153

Figure 2.36. Transmission electron micrographs for cleaved proteins in the absence presence of 25 v/v % TFE after AuNP templation: Cx·AuNP without and with TFE (a) and Qx·AuNP without and with TFE (b). Scale bars are 50 nm in top micrographs in (a) and (b), 200 nm in lower (a) and 2 µm in lower (b). ..... 154

Figure 2.37. Nyquist plots obtained for 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (4, 8, 10, 15, 25, 35 µM) of proteins C (a) and Q (b). ..... 156

Figure 2.38. Nyquist plots obtained for 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (4, 8, 10, 12 µM) of proteins Cx (a) and Qx (b) in the absence (triangle markers) and presence (circle markers) of AuNPs. ..... 157

Figure 2.39. Nyquist plots for varying concentrations (4, 8, 10, 15, 25, 35 µM) of proteins C (a) and Q (b) in the presence of AuNPs. All electrochemical measurements were performed in the presence of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes. ..... 158

Figure 2.40. Cyclic voltammograms of PB (black), C (red), Q (blue), Cx (red dashed), and Qx (blue dashed) (a) and PB (black), PB_AuNP (black dashed), C (red), and C·AuNP (red dashed) (b) in the presence of 10 mM $K_4Fe(CN)_6 \cdot 3H_2O$ and 10 mM $K_3Fe(CN)_6$. Cathodic and –anodic peak currents of different concentrations of C in the absence of Au are plotted in (c). ..... 159

Figure 2.41. Cyclic voltammograms of 8 mM C, Cx, C_AuNP, and Cx_AuNP (a) and 8 mM Q, Qx, Q_AuNP, and Qx_AuNP (b). All electrochemical measurements were performed in the presence of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes. ..... 160

Figure 2.42. Protein nanoparticle film formation results from protein self-assembly, dictated by structure. Templation of gold nanoparticles on structured proteins leads to film formation via aggregation, while unstructured proteins solvate nanoparticles. In the presence of TFE, structure is induced in cleaved proteins allowing solvated nanoparticles to aggregate. Electrochemical properties indicate that these materials can be suitable for biosensor development. ..... 160

Figure 3.1: Ferrite unit cell structure (a), and an expanded view of the magnetite crystal structure, composed of $Fe^{2+}$, $Fe^{3+}$, and $O^{2-}$ atoms (b).[215] ..... 164

Figure 3.2. Magnetic response curves from a hypothetical situation where ferromagnetic particles ranging in size from nanometer to mesoscale are injected in a blood vessel (a). M-H curves are shown for diamagnetic (DM) (b), paramagnetic (PM) (c), ferromagnetic (FM) (d), and superparamagnetic (SPM) (e). In FM particles the response can either be multi-domain (dashed), single-domain (solid), or SPM, depending on the size of the particles.[219] ..... 165

Figure 3.3. Nanoparticles produced by synthetic pathways involving ferritin in constrained reaction environments.[220] ..... 167

Figure 3.4: Magnetic bacteria and magnetosome chain formation. *Magnetospirillum magneticum* magnetic particles arranged in a chain with an average size of 50-100 nm (a),[230] electron micrograph of the magnetic bacterial cell (b), and a schematic view looking down the interior axis of a *Magnetospirillum magneticum* where 1 - cytoskeletal filament, 2 - inner membrane, and 3 – magnetosome (c).[143] Electron hologram of two double chains of magnetite in a single magnetotactic

bacteria, where magnetic field lines are oriented such that the vector points toward the northwest in the image (d).[220] ..... 168

Figure 3.5: TEM images and magnetization curves of synthesized magnetic particles. Image (a) shows those synthesized by recombinant Mms6 while (b) contains particles extracted from *Magnetospirillum magneticum*. Particle dimensions were very similar in both cases.[218] Magnetization loops for Mms6 compared to ferritin, lipocalin, and no protein at 5 K (c) and 105 K (d).[233] ..... 169

Figure 3.6. PyMOL representation of half of the E2 protein cage with 30 iron-binding sites on the interior highlighted in red (a) and assembled protein-MNP cage structures with E2-LFtn (b) and E2-LE6 (c).[225] Scale bars in (b) and (c) represent 50 nm.[225] ..... 171

Figure 3.7. Examples of protein fibers organizing inorganic nanoparticles. TEM image of 5 nm functionalized streptavidin-gold nanoparticles on SAF-p1-biotin self-assembling fiber (a).[236] STEM image of FePt nanoparticles on engineered viral templates, where scale bar represents 350 nm (b).[228] TEM images of Zn and Ni nanoparticles templated on histidine tagged COMPcc protein fibers, where scale bars represent 200 nm in the main images and 1 μm in the insets (c).[88] Nanotube composites using the tobacco mosaic virus as the protein scaffold (d).[239] ..... 172

Figure 3.8. Charge/discharge curves for hematite samples of increasing surface area (S4>S3>S2>S1) at a current density of 0.2 mA/cm$^2$ (a).[5] Sensor sensitivity towards ethanol (b) and $H_2$ (c) for hematite nanotubes (filled markers) and nanoparticles (empty markers).[243] ..... 174

Figure 3.9. Schematic of biosensor detection of MC-LR through aggregation of magnetic nanoparticles (a).[245] SEM images of nanobiocomposite/ITO electrode composed of CH-$Fe_3O_4$ particles (b) and nanobiocomposite/ITO electrode composed of Ur-GLDH/CH-$Fe_3O_4$ (c).[246] ..... 175

Figure 3.10. Purification schemes for magnetic separation of histidine tagged proteins using functionalized MNPs.[250] ..... 177

Figure 3.11. Effect of internalization of SPM MNPs in cells on MRI relaxation times. Protons in cells with MNPs exhibit shorter relaxation times (a) than those without (b).[219] ..... 178

Figure 3.12. Chemical structure of azidohomoalanine (AHA) and sequences of C+AHA and Q+AHA. .. 180

Figure 3.13. Incorporation schematic of unnatural amino acid AHA by methionine auxotrophic M15MA cells. Uptake of AHA by the cells is accomplished by washing of native methionine with 0.9 % NaCl and subsequent introduction of AHA into the culture media. MetRS binds to AHA permitting tRNA to incorporate and translate the amino acid analog.[263] ..... 182

Figure 3.14. Examples of AHA-incorporated proteins used in click chemistry reactions. Western blot analysis of the products of the Staudinger ligation, where Lane 1 is mDHFR + AHA and anti-FLAG peptide, Lane 2 is mDHFR + AHA and no peptide, Lane 3 is mDHFR + met and anti-FLAG, and Lane 4 is mDHFR + met and no peptide (a).[258] Two-dye labeling strategy for protein identification and imaging, where purple squares and orange circles represent the functionalized amino acids (b).[264] Schematic of the approach for identification of newly synthesized proteins translationally regulated by YB-1 (c).[265] ..... 183

Figure 3.15. Schematic overview of functionalization of C+AHA and Q+AHA with an alkyne-functionalized prg-CMms6 via click chemistry. ..... 185

Figure 3.16. Azide resonance structures (a). General reactions for click chemistry between azides and alkynes in the absence (b, c) and presence (c, d) of a Cu(I) catalyst. Click chemistry mechanisms of 1,3-dipolar cycloaddition for C – N bond formation in the absence of any catalyst (c), and in the presence of Cu(I) as the catalyst (d), by way of copper acetylide. Addition of copper catalyst results in only 1,4-disubstituted-1,2,3-triazole, while use of heat only and no catalyst results in a mixture of 1,4- and 1,5- regioisomers.[271] ..... 186

Figure 3.17. Synthesis schematic of AHA from protected diaminobutyric acid.[254] ..... 190

Figure 3.18. Absorption and emission spectra of Chromeo494.[277] ..... 196

Figure 3.19. Schematic of setup for magnetic separation experiments. ..... 198

Figure 3.20. SDS-PAGE gels showing AHA incorporation with expression with 19 and 20 amino acids as controls. (a) 1. Ladder, 2. C pre-induction, 3. Q pre-induction, 4. C+AHA post-induction, 5. Q+AHA post-induction, 6. C 19 aa post-induction, 7. Q 19 aa post-induction. (b) 1. Ladder, 2. C pre-induction,

3. Q pre-induction, 4. C+AHA post-induction, 5. Q+AHA post-induction, 6. C post-induction, 7. Q post-induction. ..... 201

Figure 3.21. SDS-PAGE analysis of purification for C+AHA (a) and Q+AHA (b). Lanes represent: 1. supernatant, 2. flow through, 3. wash 1 (20 mM imidazole), 4. wash 2 (20 mM imidazole), 5. 100 mM imidazole elution, 6 and 7. 200 mM imidazole elutions, 8-10. 500 mM imidazole elutions, 11-13. 1 M imidazole elutions. ..... 201

Figure 3.22. Amino acid analysis for AHA. ..... 202

Figure 3.23. Amino acid analysis for C (a), C+AHA (b), Q (c), and Q+AHA (d). ..... 203

Figure 3.24. Chromeo494 bearing an alkyne moiety and reaction schematic for covalent attachment via azide group (a). SDS-PAGE showing results of click chemistry reaction with incubation times of 4, 8, 12, 24, 36, and 48 h for Chromeo494 and C+AHA, C, Q+AHA, and Q, imaged with fluorescence filter (b). Results of ImageQuant gel analysis quantifying the RFU over time for C+AHA and C (c) and Q+AHA and Q (d). ..... 204

Figure 3.25. SDS-PAGE showing results of click chemistry reaction with of Chromeo494 with C, C+AHA, Q, and Q+AHA with incubation time of 48 h stained with Coomassie blue (a) and unstained and imaged with fluorescence filter (b). ..... 206

Figure 3.26. SDS-PAGE showing varying concentrations of prg-CMms6. 1. Ladder, 2. 2.25 mM, 3. 1 mM, 4. 500 μM, 5. 250 μM, 6. 100 μM, 7. 50 μM. The molecular weight of prg-CMms6 is 4.44 kDa. ... 207

Figure 3.27. SDS-PAGE showing Coomassie blue stained gel (a); 1. Mark12 ladder, 2. C+AHA + Chromeo, 3. C+AHA + CMms6, 4. Q+AHA + Chromeo, and 5. Q+AHA + CMms6. Same gel viewed with the fluorescence filter (b), where bands seen in lanes 3 and 5 are sample from lanes 2 and 4 that spilled over when loading the gel. ..... 208

Figure 3.28. SDS-PAGE showing click chemistry reactions with varying concentrations of prg-CMms6 with no sodium ascorbate after incubation at 50 °C for 48 h. 1 and 16. Mark12 ladder, 2-5: C+AHA + varying concentrations of prg-CMms6, 6-9: Q+AHA + varying concentrations of prg-CMms6, 10-13: C + varying concentrations of prg-CMms6, 14, 15, 17, 18: Q+ varying concentrations of prg-CMms6, where lanes 2, 6, 10, and 14 contain 100 μM prg-CMms6; 3, 7, 11, and 15 contain 200 μM prg-CMms6; 4, 8, 12, 17 contain 400 μM prg-CMms6; and 5, 9, 13, and 18 contain 500 μM prg-CMms6. ..... 210

Figure 3.29. SDS-PAGE of click chemistry on pure protein with no sodium ascorbate. 1. and 14. Ladder, 2-5: 100 μM prg-CMms6; 6-9: 200 μM prg-CMms6; 10-13: 400 μM prg-CMms6; 15-19: 500 μM prg-CMms6, where lanes 2, 6, 10, and 15 contain C+AHA; 3, 7, 11, and 16 contain Q+AHA; 4, 8, 12, and 17 contain C; and 5, 9, 13, and 18 contain Q. ..... 212

Figure 3.30. SDS-PAGE of click chemistry as a function of time, with sodium ascorbate and varying concentrations of prg-CMms6. (a) C+AHA lysate. 1. Ladder, 2-5: No prg-CMms6, 6-9: 200 μM prg-CMms6, 10-13: 500 μM prg-CMms6, where lanes 2, 6, and 10 were taken at the beginning of the reaction; 3, 7, and 11 represent 20 h of incubation; 4, 8, and 12 represent 28 h of incubation; and 5, 9, and 13 represent 48 h of incubation. (b) Q+AHA lysate. 1. Ladder, 2-5: No prg-CMms6, 6-9: 200 μM prg-CMms6, 10-13: 500 μM prg-CMms6, where lanes 2, 6, and 10 were taken at the beginning of the reaction; 3, 7, and 11 represent 20 h of incubation; 4, 8, and 12 represent 28 h of incubation; and 5, 9, and 13 represent 48 h of incubation. ..... 213

Figure 3.31. Secondary structure of C, C+AHA, Q, and Q+AHA proteins at 50 mM PB pH 8. Circular dichroism variable wavelength scans of C (black, solid), C+AHA (black, dashed), Q (grey, solid), and Q+AHA (grey, dashed). Inset represents scans of C+AHA and Q+AHA at a different y-axis scale. Data is averaged from three replicates and was obtained with 10 μM protein concentration for C and Q and 14 μM protein concentration for C+AHA and Q+AHA. ..... 214

Figure 3.32. Transmission electron micrographs for 23 μM C+AHA 50 mM PB pH 8 (a-c) and 20 μM C at 10 mM PB pH 8 at room temperature (d, e). Predominant morphology seen in C+AHA was aggregates and protofibrils in C (e). Scale bars are 2 μm in (a), 0.5 μm in (b), 200 nm in (c) and (e), and 100 nm in (d). ..... 215

Figure 3.33. Transmission electron micrographs for 11 μM Q+AHA 50 mM PB pH 8 (a-c) and 20 μM Q at 10 mM PB pH 8 at room temperature (d, e). Q+AHA contained some aggregates and some fibers,

which appeared in aggregated networks, whereas Q appeared as sheets with some observable striations (d). (e). Scale bars are 0.5 μm in (a), 50 nm in (b) and (d), 100 nm in (c), and 1 μm in (e). Histogram of protofibril diameters ($n$ = 50) of Q (black) and Q+AHA (grey), 50 mM PB pH 8..... 217

Figure 3.34. Transmission electron micrographs for 200 μM prg-CMms6 in 50 mM PB, pH 8. Predominant morphology seen was aggregates. Scale bars are 200 nm (left and center) and 0.5 μm (right). ..... 218

Figure 3.35. Transmission electron micrographs for MNPs formed via coprecipitation in 50 mM PB, pH 8 (a, b). White lines in (b) are superimposed on lattice fringes observed at high resolution. Electron diffraction pattern of MNPs (c) displayed rings that were used to calculate $d$-spacing of the crystals. Scale bars are 200 nm (a), 20 nm (b), and 5 $nm^{-1}$ (c). ..... 218

Figure 3.36. Transmission electron micrographs for MNPs formed in the presence of 1 mM prg-CMms6 via coprecipitation in 50 mM PB, pH 8 (a, b). Electron diffraction pattern of MNPs formed in the presence of prg-CMms6 (c) displayed rings that were used to calculate $d$-spacing of the crystals. Scale bars are 100 nm (a), 20 nm (b), and 5 $nm^{-1}$ (c). ..... 219

Figure 3.37. TEM of 200 μM prg-CMms6 + MNP (a-c), 500 μM prg-CMms6 + MNP (d-f), and 1 mM prg-CMms6 + MNP (g-i). Scale bars represent 200 nm in (a) and (d), 100 nm in (b) and (e-g), 50 nm in (c) and (h), and 20 nm in (i). ..... 221

Figure 3.38. Histogram of MNP diameters ($n$ = 50) reduced in the presence of 50 mM PB only (black) and 200 μM (dark grey), 500 μM (light grey), and 1 mM prg-CMms6 (white) (a). Negative linear correlation between prg-CMms6 concentration and MNP size (b). ..... 222

Figure 3.39. TEM of unclicked C+AHA + MNP (a, b) and unclicked Q+AHA + MNP (c, d). Concentration of C+AHA was 23 μM and Q+AHA was 11 μM, with molar ratios of iron salts and reducing agent normalized accordingly. Presence of needle-like crystals is indicated by white arrows. Scale bars represent 100 nm in (b) and (c) and 50 nm in (a) and (d). Samples imaged without uranyl acetate stain. ..... 223

Figure 3.40. Phase contrast microscopy of synthesized MNPs in 50 mM PB pH 8 before and after the application of a magnetic field. Direction of the magnetic field is indicated by the white arrow. Scale bars represent 100 μm. ..... 224

Figure 3.41. Phase contrast microscopy of MNPs synthesized in the presence of 200 μM prg-CMms6 in 50 mM PB pH 8 before (a) and after (b, c, d) the application of a magnetic field. Direction of the magnetic field is indicated by the white arrows. Scale bar represents 100 μm in (a) and 75 μm in (b), (c), and (d). ..... 225

Figure 3.42. Phase contrast microscopy of MNPs synthesized in 50 mM PB pH 8 incubated with C+AHA lysate and C+AHA+prg-CMms6 lysate before (top) and after 2 washes (middle) and 4 washes (bottom) via magnetic separation. Scale bar represents 100 μm in top left, middle, and bottom left images and 250 μm in top right and bottom right images. ..... 226

Figure 3.43. Phase contrast microscopy of MNPs synthesized in 50 mM PB pH 8 incubated with Q+AHA lysate and Q+AHA+prg-CMms6 lysate before (top) and after 2 washes (middle) and 4 washes (bottom) via magnetic separation. Scale bar represents 100 μm in top left, middle left, and bottom left images and 75 μm in top right, middle right, and bottom right images. ..... 227

Figure 3.44. Magnetization measurements as a function of applied field for MNPs synthesized in the presence of 50 mM PB pH 8 (white circles) and CMms6 (black circles) (a) and in the presence of whole cell lysate containing expressed C+AHA (white circles) and Q+AHA (black circles), or clicked C+AHA + CMms6 (white triangles) and Q+AHA + CMms6 (black triangles) (b). ..... 229

# List of Tables

Table 1.1. Protein names and sequences. 9
Table 1.2. pH sensitivity of select coiled-coil proteins.[31] 18
Table 1.3. Peptides susceptible to temperature control. 20
Table 1.4. Peptides susceptible to control by additives. 23
Table 1.5. Protein sequence information for wt, C, Q, and L, from N-terminus to C-terminus, with the histidine tags in bold. 29
Table 1.6. Primer sequences used to generate fragments for C, Q, and L. 38
Table 1.7. DNA sequences for C, Q, and L. 39
Table 1.8. Contents of restriction reactions of PheRS** xL1 blue vector DNA. 40
Table 1.9. Ligation reagent concentrations for ligating designed DNA with PheRS** plasmid DNA and for φ ladder control reactions. 41
Table 1.10. Contents of reaction tubes used during restriction of PheRS**/insert DNA. 42
Table 1.11. Expected base pair lengths of DNA fragments. 51
Table 1.12. Sequence identity between designed DNA sequences and results from sequencing experimentally generated DNA for C, L, and Q. 57
Table 1.13. Mean residual ellipticities and secondary structure content from circular dichroism measurements at pH 4, 8, and 10 for wt, C, Q, and L. The data presented represents the average of at least two replicates. 64
Table 1.14. Secondary structure conformation based on ATR-FTIR data for 10 μM Q in 50 mM PB at pH 4, 8, and 10. Percent composition was determined from relative areas of peaks fit to spectra (Figure 1.40). 66
Table 1.15. Thermodynamic constants obtained from thermal melts of 10 μM protein measured by circular dichroism, with the intercept of 1/T vs. lnK not equal to zero. Data is averaged from two replicates. 75
Table 1.16. Thermodynamic constants obtained from thermal melts of 10 μM protein measured by circular dichroism, with the intercept of $1/T$ vs. ln$K$ equal to zero. Data is averaged from two replicates. 76
Table 1.17. Extinction coefficients, binding constants, and number of binding sites for wt, C, and Q determined by the Benesi-Hildebrand or Scatchard equations. *Modified Q data represents the adjustment of a singular data point (absorbance at 420 nm at 20 μM concentration of curcumin). 92
Table 1.18. Secondary structure conformation based on ATR-FTIR data for 10 μM Q in 50 mM PB at pH 4 in the presence of 50 μM curcumin. Percent composition was determined from relative areas of peaks fit to spectra (Figure 1.64). 94
Table 2.1. AuNP templation by select proteins. 115
Table 2.2. Sequence information for C, Q, Cx, and Qx, from N-terminus to C-terminus, with the histidine tags in bold. 129
Table 2.3. Average and standard deviation of AuNP diameters templated in the presence of various proteins ($n$ = 50). 144
Table 2.4. Secondary structure conformation based on ATR-FTIR data for 10 μM protein in 50 mM PB pH 8 before and after templation of AuNPs, and in the presence of 25 v/v % TFE. Percent composition was determined from relative areas of peaks fit to spectra. Data was not obtained for Cx·AuNP and Qx·AuNP with TFE due to problems with sample measurement. 155
Table 2.5. Electron transfer resistance ($R_{et}$) of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (0, 4, 8, 10, 15, 25, 35 μM) C, Q, Cx, and Qx as calculated from Nyquist plots. (Data was not obtained for cleaved proteins above 10 μM as higher protein concentrations were not achieved.) 157

Table 2.6. Electron transfer resistance ($R_{et}$) of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (0, 4, 8, 10, 15, 25, 35 μM) C, Q, Cx, and Qx in the presence of AuNP as calculated from Nyquist plots. (Data was not obtained for cleaved proteins above 10 μM as higher protein concentrations were not achieved.) ..... 158

Table 3.1. (*hkl*) planes and corresponding *d*-spacings for maghemite ($\gamma$-$Fe_2O_3$) and magnetite ($Fe_3O_4$). .. 164

Table 3.2. Reagents and suppliers of materials used in AHA incorporation. ..... 189

Table 3.3. Actual molecular weights and estimated molecular weight shift by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and Chromeo494. Estimated reaction yields were obtained by the molecular weight shift. ..... 206

Table 3.4. Actual molecular weights and observed molecular weight shift by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and prg-CMms6. Observed fraction of proteins with two clicked prg-CMms6 peptides were obtained by the molecular weight shift. ..... 209

Table 3.5. Estimated molecular weights by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and varying concentrations of prg-CMms6. Yield and estimated fraction of proteins with two clicked prg-CMms6 peptides were obtained by the estimated molecular weights, calculated via calibration with the ladder in ImageQuant. ..... 211

Table 3.6. MRE values corresponding to $\theta_{222}$, $\theta_{min}$, $\theta_{222}/\theta_{min}$, and the deviation of $\theta_{222}/\theta_{min}$ from unity for C, C+AHA, Q, and Q+AHA as determined from wavelength scans presented in Figure 3.31. ..... 215

Table 3.7. Average and standard deviations of protofibril diameters for Q at pH 4 ($n$ = 210) and pH 8 ($n$ = 50) and Q+AHA at pH 8 ($n$ = 50). ..... 217

Table 3.8. Calculation of *d*-spacing for the rings resulting from electron diffraction pattern of MNPs (Figure 3.35c) and corresponding crystallographic planes. ..... 219

Table 3.9. Calculation of *d*-spacing for the rings resulting from electron diffraction pattern of MNPs formed in the presence of 1 mM prg-CMms6 (Figure 3.36c) and corresponding crystallographic planes. ... 220

# 1 Self-assembling protein fibers

The text in this chapter is partially from the paper:[1]

Jasmin Hume, J. Sun, R. Jacquet, P. D. Renfrew, J. A. Martin, R. Bonneau, M. L. Gilchrist, J. K. Montclare. Engineered Coiled-Coil Protein Microfibers. *Biomacromolecules* **15,** 3503-3510 (2014).

The print version of this paper can be found in Appendix 5.1.1.

## 1.1 Abstract

The fabrication of *de novo* proteins able to self-assemble on the nano- to meso- length scales is critical in the development of protein-based biomaterials in nanotechnology and medicine. Here we report the design and characterization of a protein-engineered coiled-coil that not only assembles into microfibers but also can bind hydrophobic small molecules. Under ambient conditions the protein forms fibers with nanoscale structure possessing large aspect ratios formed by bundles of $\alpha$-helical homopentameric assemblies, which further assemble into mesoscale fibers in the presence of curcumin through aggregation. Surprisingly, these biosynthesized fibers are able to form in conditions of remarkably low concentrations. Unlike previously designed coiled-coil fibers, these engineered protein microfibers can bind the small molecule curcumin throughout the assembly, serving as a depot for encapsulation and delivery of other chemical agents within protein-based 3D microenvironments.

## 1.2 Introduction

### *1.2.1 Bio-inspired self-assembling proteins*

For thousands of years, humans have drawn insight for solutions to technical problems from the natural world. It is this type of insight that has led to the design of organic, as well as inorganic, materials that are able to interface seamlessly with organic compounds, greatly benefitting medical technology as well as technologies for electronics and renewable energy sources.[2–5] Protein-based materials, the building blocks that nature uses

for myriad biological applications, prove to be a powerful platform when it comes to engineering new materials for state of the art applications ranging from biomedicine to nanostructured electronic components.

Precise molecular organization across the nano- and meso- length scales is crucial for producing materials with defined physicochemical properties.[6–8] Nature provides elegant examples of proteins that self-assemble to form a variety of complex structures, ranging from protein vesicles and cages to a variety of different fibrous assemblies (Figure 1.1).[9] In fact, protein assemblies exist in every biological cell – from unicellular bacteria, archaea, and eukaryotes to complex multicellular organisms. Understanding the assembly of these structures enables engineers to design and create functional biomaterials through bottom-up fabrication for applications in nanobiotechnology.

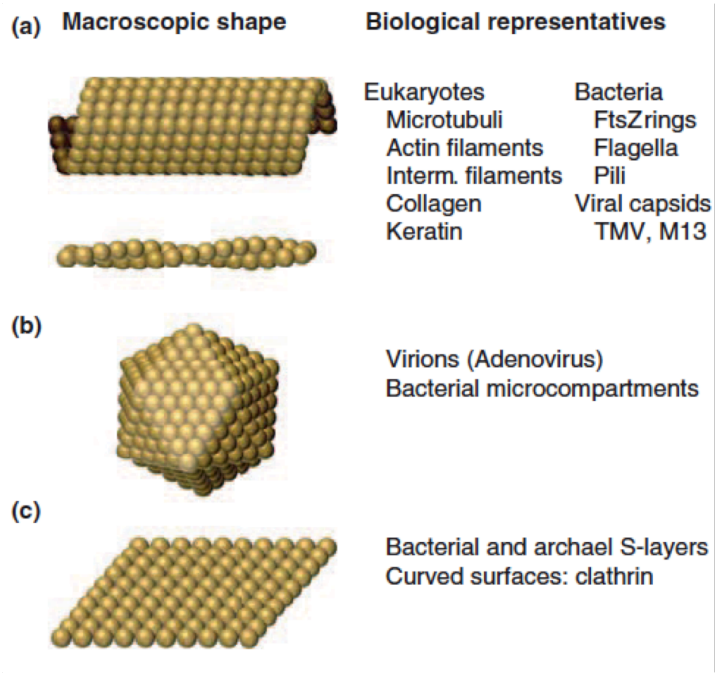



Figure 1.1. Protein assemblies and their biological representatives.[9]

For simplicity, we divide these structures between those that are non-fibrous and fibrous protein assemblies.

#### 1.2.1.1 Non-fibrous protein assemblies

Non-fibrous protein assemblies form shapes that can be categorized as cages, vesicles, and spheres (Figure 1.1b).[9] These types of protein materials frequently find function in

storing or encapsulating other materials. A typical example is the viral capsid that encapsulates DNA or RNA to protect the genetic material from the external environment, delivers it to a targeted location, and subsequently disassembles upon response from an external trigger.[10] Spherical protein assemblies also act as carriers of nanoscale biological cargo, such as enzymes or toxins. Walls of these spherical cargo carriers are often nanoporous, allowing for permeation of small organic substrates and products while withholding the larger enzyme within the sphere, leading to an improvement in enzymatic activity.[11]

Protein spheres can be composed of a single type of protein or have multiple protein components.[12–15] The C-terminal fragment of the heavy chain of clathrin, for example, is known to self-assemble into homotrimers forming truncated triseklia, which then further polymerize to form open cage-like structures as well as branched assemblies.[13] The shells of several bacterial microcompartments are in fact composed of two different proteins, typically the first assembling into hexagons creating icosahedral facets and the second forming pentamers at the vertices of said icosahedra. Collagen mimetic peptides (CMPs), commonly associated with fiber-forming structures, have also been used to create non-fibrous, dual component spherical micelles.[12] In this case, micelles are created by the combination of two amphiphilic fatty-acid CMP conjugates that self-assembled in aqueous solution.[15] As these systems are the focus of more and more studies, computational simulations are now being employed to design these dual-component spherical assemblies in a predictable and accurate manner.[14] Figure 1.2 demonstrates the high correlation between experimentally produced transmission electron micrographs of several of these designed dual-component assemblies with predictive models resulting from computational simulations.[14] As can be inferred from the natural function of these structures, synthetically produced versions of these non-fibrous protein structures hold great potential in applications in the area of drug delivery.

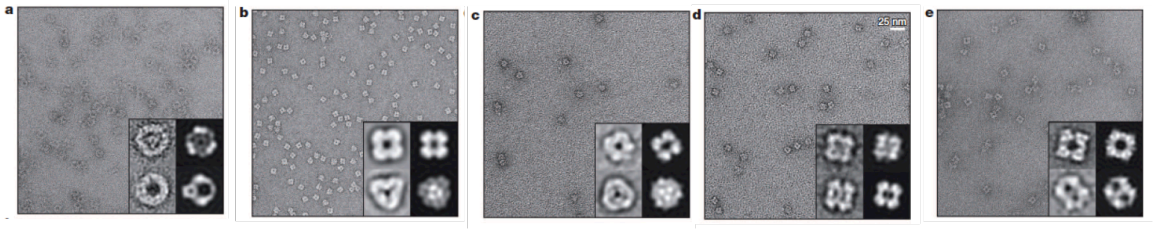



Figure 1.2. Transmission electron micrographs of designed two-component protein cages, protein components being T32-28 (a), T33-09 (b), T33-15 (c), T33-21 (d) and T33-28 (e). Insets represent the two class averages of the particles (top and bottom) for the actual particles (left) and computational models (right).[14]

#### 1.2.1.2 Fibrous protein assemblies

Nature provides elegant examples of proteins that self-assemble to form fibers both on the nanometer scale and on the mesoscale. On the nanoscale, examples include *β*-amyloids like those responsible for Alzheimer's disease[16,17] as well as optically active self-assembling reflectins.[18,19]

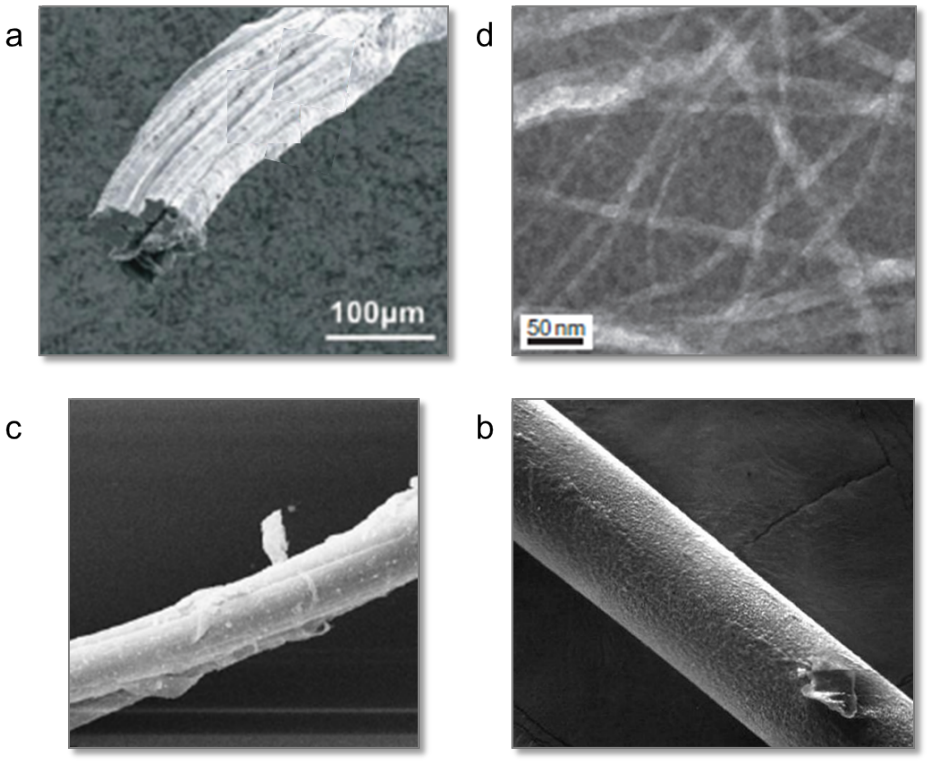



Figure 1.3. Mesoscale protein fibers. SEM of the byssal threads of *Tridacna maxima* (a),[20] TEM of collagen fibers (b),[21] SEM of natural spider silk (c),[22] and SEM of spun spider silk made recombinantly in mammalian cells (d).[23] Scale bar is 100 μm in (a), 50 nm in (b), and fibers are approximately 15 μm in width in (c) and 20 μm in (d).

Fibrous proteins on the mesoscale include the bundled $\alpha$-helical coiled-coil elastic protein of the Giant Clam, *Tridacna maxima* (Figure 1.3a),[20] $\alpha$-helical intermediates of amyloid fibers,[17,24] fibrinogen,[25] $\alpha$-keratin fibers found in hair, nails, quills, and horns,[26] actin filaments that make up muscle fibers,[27] collagen (Figure 1.3b),[8,12,21,28] bacterial pili, spider silk (Figure 1.3c, d), and the less well-known assembly of the sterile $\alpha$-motif.[29] The mesoscale fibers shown in micrographs in Figure 1.3, are, of course, comprised of organized nanostructures that bundle together to form the observed fibers. A well-studied example of this hierarchical assembly is found in bacterial pili, in which $\beta$-sheet proteins link together to make robust mesoscale fibers exhibiting impressive structural flexibility (Figure 1.1a).[9] The formation of $\beta$-sheet, amyloid-like fibers is controlled by inter-chain hydrogen bonding within the backbone of the protein. In Section 1.2.2 we discuss the self-assembly of helical proteins to form fibers, where the backbone hydrogen bonding is centered within the helices and not between chains. This reduced promiscuity of hydrogen bonding between protein subunits allows for greater control of the interactions between helices, allowing us to better manipulate fiber design, achieving fibers of specific dimensions or with interesting binding capabilities correlating with structure.

While nature has created many elaborate proteins capable of complex self-assembly and ligand binding, generating synthetic materials with the same level of structural and molecular specificity on various length scales remains a challenge. Attempts by various groups at addressing this undertaking, with respect to $\alpha$-helical protein fibers specifically, are presented in the following sections.

### *1.2.2 α-helical coiled-coil proteins and their fibers*

$\alpha$-helical proteins are typically right-handed spirals held together by hydrogen bonds occurring between the oxygen of the C=O on a residue on the top of a coil and the hydrogen of the N-H of a residue on the bottom of the coil, exactly 4 residues apart.[30] The $\alpha$-helix completes one turn every 3.6 residues, creating a pitch of 1.5 Å per residue. Each $\alpha$-helix is defined by heptad residue repeats, designated by positions *a-f* (Figure 1.4).

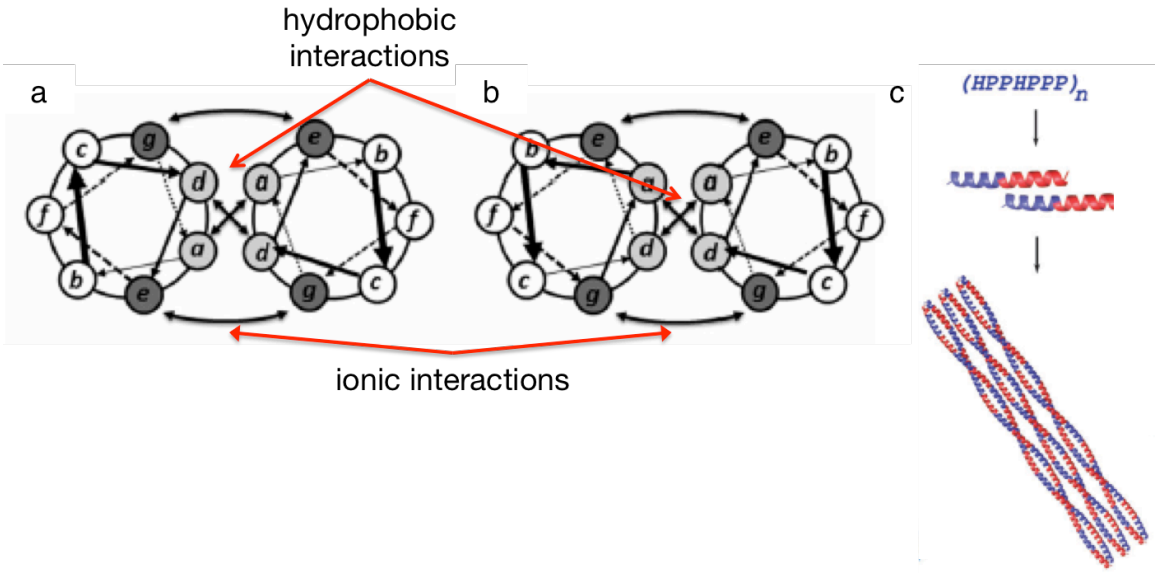



Figure 1.4. Dimeric coiled-coils. Helical wheels for parallel dimeric coiled-coil (a), and antiparallel dimeric coiled-coil (b),[31] where hydrophobic interactions occur between the *a* and *d* residues and ionic interactions between *g/e* (a) and *e/e or g/g* residues (b). Amphipathic assembly of a coiled-coil dimer (c).[32]

$\alpha$-helical proteins can assemble to form superstructures called coiled-coils, where the helices wind around one another with an overall left-handed twist.[33] In order for helices to assemble into a coiled-coil configuration, their heptad composition must contain hydrophobic residues (H) in positions *a* and *d* of the helix and polar residues (P) in positions *b*, *c*, *e*, *f*, and *g*. The resulting heptad template for coiled-coil proteins is thus $(HPPHPPP)_n$ (Figure 1.4c).

Oligomeric assemblies have been created using $\alpha$-helices as building blocks, where $\alpha$-helices assemble by taking advantage of hydrogen bonding and van der Waals' forces to gain stability.[34] Interactions between coiled-coils have been used to aid rational design of protein nanofibers. Coiled-coils have been designed using a wide range of combinations leading to interesting characteristics. They can possess various oligomerization states, such as dimeric, trimeric, pentameric, and so on.[30,35] It is the identity of the residues in the hydrophobic positions of the heptad that determine the oligomerization state of the coiled-coil, which was discovered by manipulation of these residues in the yeast transcription factor GCN4.[36] The main determining factor of whether the most stable form of a coiled-coil is obtained by two, three, or more chains is the shape of the hydrophobic side chains within the protein.[37] Crystallographic studies

proved that when *a* and *d* positions are occupied by (i) isoleucine and leucine, respectively, the result was dimers, (ii) isoleucine in both positions generated trimers, and (iii) leucine and isoleucine, respectively, resulted in tetramers.[36,37] Some examples of coiled-coil assemblies of various states of oligomerization are shown in Figure 1.5.

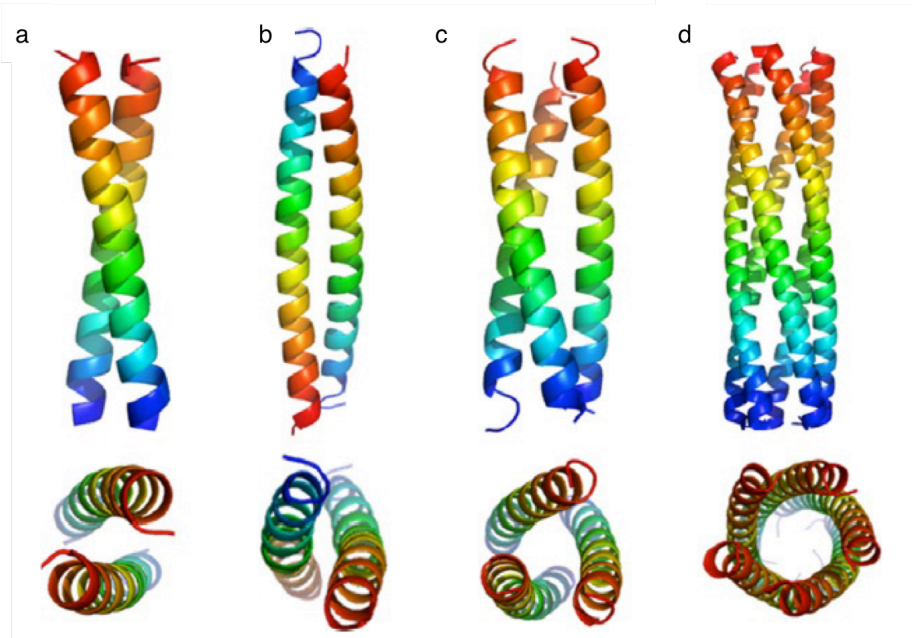



Figure 1.5. Orthogonal views of several coiled-coil structures: parallel dimer (PDB 2ZTA) (a), antiparallel dimer (PDB 1HF9) (b), parallel trimer (PDB 1BB1) (c), and parallel pentamer (PDB 2GUV) (d).[38]

Coiled-coil assemblies can be composed of a single type of $\alpha$-helix (homo) or different $\alpha$-helices (hetero).[31,37] Additionally, coiled-coil oligomers can be "blunt ended", meaning that $\alpha$-helices comprising the coiled-coil end and being in the same place (Figure 1.6a center), or "sticky ended", where the $\alpha$-helices are staggered with respect to one another, resulting in an overhanging region on each helix (Figure 1.6a right). Unlike coiled-coiled structures that are found in nature, which exhibit blunt-ended coils, several designed coiled-coils use sticky ends to program these helices to pack in slipped arrangements. Fibrillogenesis is triggered by these overhangs, or sticky-ends, which are energetically driven to further assemble longitudinally (Figure 1.6b).

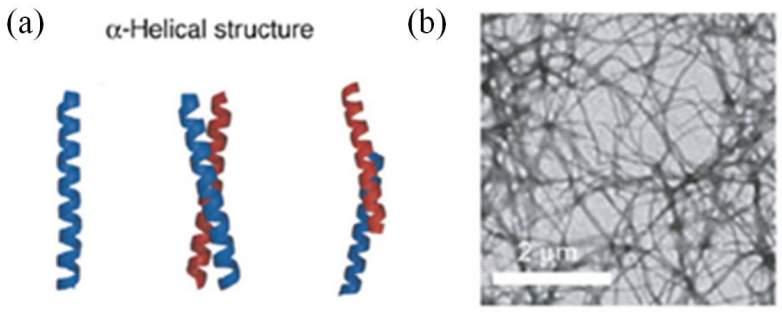



Figure 1.6: Protein fiber assembly by sticky-end coiled-coil dimers. (a) From left to right: single $\alpha$-helical protein, blunt-end protein assembly, and sticky-end protein assembly. (b) TEM micrograph of fiber formation due to sticky-end assembly.[34]

Coiled-coil assemblies contain structure-stabilizing interactions between hydrophobic residues, and ionic interactions between residues in positions *e* and *g* (see Figure 1.4).[31,33] They can be arranged in parallel or antiparallel fashion, as is shown schematically for dimers in Figure 1.4a, b and Figure 1.5a, b, respectively.

Table 1.1. Protein names and sequences.

| Name | Sequence | Ref. |
|---|---|---|
| SAF-p1 | K IAALKQK IASLKQE IDALEYE NDALEQ | 39–41 |
| SAF-p2 | K IRALKAK NAHLKQE IAALEQE IAALEQ | |
| SAF-p2a | K IRRLKQK NARLKQE IAALEYE IAALEQ | |
| Ac-SAF-p1-$NH_2$ | Ac-K IAALKQK IASLKQE IDALEYE NDALEQ-$NH_2$ | |
| Ac-SAF-p2a-$NH_2$ | Ac-K IRRLKQK NARLKQE IAALEYE IAALEQ-$NH_2$ | |
| SAF-p1-ext | K IAALKQK IASLKQE IDALEYE NDALEQK IAALEQ | |
| SAF-p2a-ext | K IRRLKQK NARLKQK IAALEQE IAALEYE IAALEQ | |
| Construct I | 2 x (KIAALKQ-EIAALEQ-KIAALKQ-EIAALEQ) | 42 |
| Construct II | 2 x (KIAALKQ-KIAALKQ-EIAALEQ-EIAALEQ) | |
| YZ1 | EIAQLEKEIQALEKENAQLEKKIQALRYKIAQLREKNQALRE | 43 |
| $\alpha_3$-peptide | $(LQTLAKA)_3$ | 44 |
| CMP | Based on the sequence of natural collagens, following 3 amino acid repeat of XYG, where X is predominantly occupied by proline and Y is most commonly 4R-hydroxyproline | 45 |
| CMP tethered to tris(2-aminoethyl) amine succinic acid | 3 amino acid repeat of XYG, where X is predominantly occupied by proline and Y is most commonly 4R-hydroxyproline and tethered to tris(2-aminoethyl) amine succinic acid derivatives | 46 |
| *α*FFP | QLAREL-$(QQLAREL)_4$ | 47 |
| MW1 | KIKALKYEIAALEQEIAALEQKIAALKQ | 48 |
| MW1-K3A | KIAALKYEIAALEQEIAALEQKIAALKQ | |
| MW1-Y7Q | KIKALKQEIAALEQEIAALEQKIAALKQ | |
| MW1-K3E | KIEALKYEIAALEQEIAALEQKIAALKQ | |
| MW1-K3nL | KInALKYEIAALEQEIAALEQKIAALKQ | |
| MW1-Y7➔21 | KIKALKQEIAALEQEIAALEYKIAALKQ | |
| MW1-K3➔17 | KIAALKYEIAALEQEIKALEQKIAALKQ | |
| DSAg | See Appendix 5.5.1 | 49 |
| DSD | | |

Woolfson *et al.* have produced a great body of work on self-assembling fibers (SAFs) containing either blunt ends and sticky ends (Figure 1.6a); the latter leads to longitudinal self-assembly resulting in linear fibers with diameters of about 20 nm (Figure 1.6b, Table 1.1) while the former is unable to form fibers altogether.[39] The fibers produced via sticky end assembly have been several hundred micrometers in length; however exact length was difficult to verify as fibers were intertwined.[39] A schematic of oligomeric assembly and TEM image of fiber formation is shown in Figure 1.6.

Zimenkov *et al.* designed an $\alpha$-helical oligopeptide called YZ1 programmed to self-assemble with a dimeric strandedness comprised of six heptad repeats (Table 1.1).[43] The authors followed the model of dimeric GCN4 coiled-coils, with isoleucine and leucine in the *a* and *d* positions, and also incorporated features to promote a staggered assembly promoting fibrillogenesis. The result was high aspect ratio fibers of uniform diameter and stable $\alpha$-helical signature as interpreted via circular dichroism.[43]

Some of the earliest work on *de novo* $\alpha$-helical protein fibers has been done by Kojima *et al.* with the homotetrameric $\alpha_3$-peptide, which assembled to form fibers 5-10 nm in diameter with visible protofibrils.[44] These short $\alpha_3$-peptides consist of three repeats of a seven-residue sequence, LQTLAKA (Table 1.1). Protofibrils 5-10 nm in width have been observed to bundle together to form larger fibers, confirmed by TEM. Collagen mimetic peptide (CMP) is a natural choice as a template for the formation of a variety of fibers, as collagen naturally forms robust tetrameric fibers.[12,21,45,46] CMPs have been used to form self-assembling homo and heterotrimers, as well as higher-order structures (Table 1.1).[45] Fibers formed by heterotrimers have also been produced by Li *et al.* who used collagen mimetic peptide as a basis for construction and covalently tethering them to a tris(2-aminoethyl) amine succinic acid derivative (Table 1.1).[46] The authors observe that the thermal stability and folding rates of these assemblies increased with the number of helical stabilizing amino acids (eg/ hydroxyproline) present in the sequence.[46]

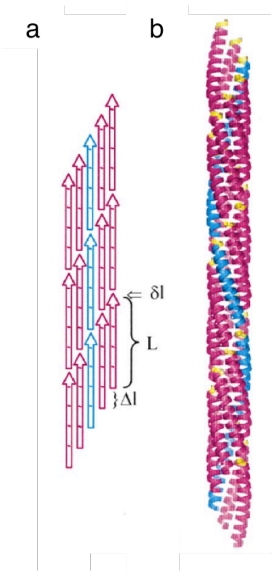



Figure 1.7. Schematic representation of pentameric fibril created by assembly of staggered $\alpha$-helices (a) and the structural model of the same (b).[47]

Potekhin *et al.* made a significant impact on the field of rationally designed coiled-coil assemblies with their work on an $\alpha$-helical peptide that self-assembled into a staggered homopentamer formation producing soluble fibers of defined diameter (Table 1.1).[47] They designed the peptides in such a manner as to achieve an offset of one heptad in length ($\Delta l$, as shown in Figure 1.7a), allowing each helix to be staggered with respect to its neighbor. This created an energetically favorable scenario for elongation of a pentameric coiled-coil fiber (Figure 1.7b).

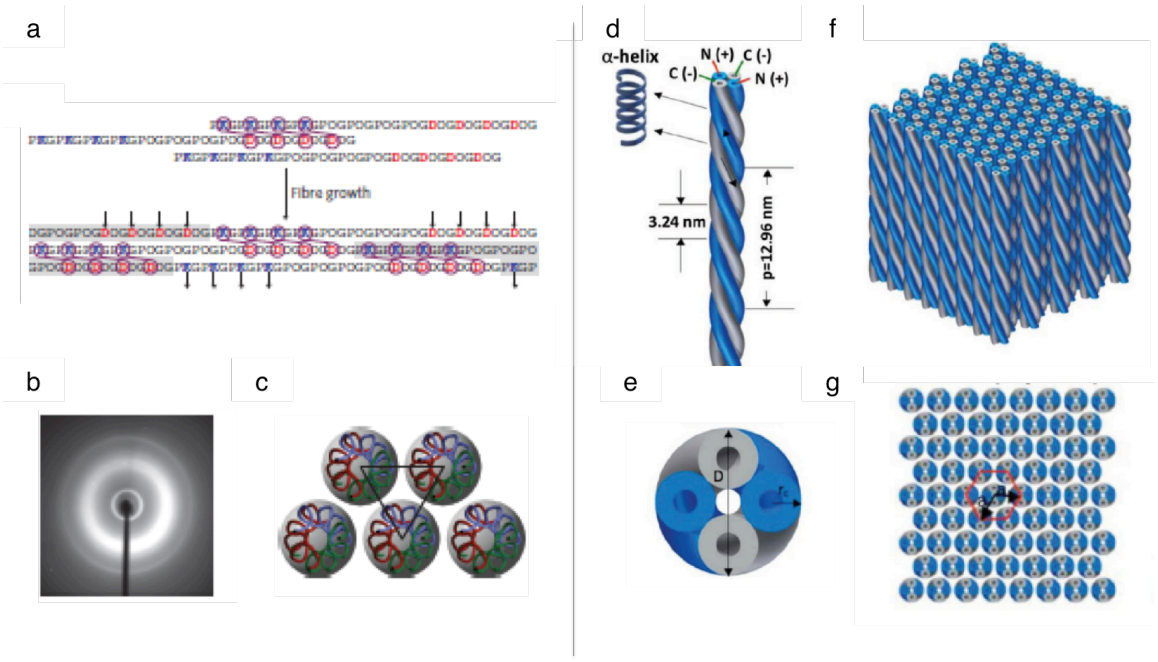



Figure 1.8. Helical coiled-coil assemblies bundle to form fibers. Sequence of collagen mimetic peptide which forms homotrimers (a),[21] whose fiber structure was solved using X-ray diffraction (b), quasihexagonal packing with 1.2 nm protofibrils, corresponding to the cross-section of the triple helix (c).[21] Structural model of the four-stranded nanofibrils of *Tridacna maxima* (d), nanofibril cross-section (e), and bundling of nanofibrils to form mesoscale fibers (f) through hexagonal packing (g).[20]

Even more remarkable than the ability of $\alpha$-helical peptides to self-assemble into coiled-coils is the bundling of coiled-coils into higher order structures (Figure 1.8). With so many different ways to design coiled-coil assemblies, we have combined two of the more successful techniques in forming robust, three-dimensional bundled fibers widely used in the literature in the design of our novel proteins (detailed design methodology presented in Section 1.2.5). These techniques involve the manipulation of electrostatic (Section 1.2.2.1) and 3D swapping (Section 1.2.2.2) of protein domains.

### 1.2.2.1 Electrostatic domains

Understanding of electrostatic interactions between specific residues or segments in $\alpha$-helical protein chains can be used to organize regions of the protein to obtain fibers. Secondary structure can be conserved while the tertiary structure is modified by maintaining the positions of residues in the heptad while moving whole heptads with respect to one another in the protein sequence. Electrostatic charge distribution in patches

has previously been used to facilitate and direct self-assembly of coiled-coil protein fibers, confirmed by cryo-TEM, X-ray crystallography/diffraction, and modeling.[40,50]

As mentioned previously, Woolfson *et al.* have produced SAFs[41,51,52] that form nanofibers from tens to hundreds of nanometers in diameter and micrometers in length (Table 1.1).[35,39] These fibers are made up of protofibrils corresponding to dimeric coiled-coils 2 nm in diameter that bundle to form fibers 40-80 nm in thickness have also been seen with modular $\alpha$-helical protein units.[42] These $\alpha$-helical peptide fibers assemble through electrostatic and hydrophobic interactions between complementary subunits. Dimensions of previously reported $\alpha$-helical peptide fibers range from 10-100 nm in diameter and micrometers in length.[39,41,43] To date, peptide or recombinant protein fibers with diameters on the micrometer scale have never been achieved.

Notable examples of this type of design include two different types of two-stranded $\alpha$-helical coiled-coil rods made from SAFs, where each heptad has been programmed with an overall positive or negative charge, as is determined by the solvent exposed residues in the heptad.[40,41,48] The first of these examples has proteins containing four heptads, two positively charged followed by two negatively charged, named SAF-p1 and SAF-p2a (Table 1.1). These proteins are designed in such a manner that the association of the positive and negative heptads generates two sticky ends, allowing the fibers to extend longitudinally through electrostatic interactions.[40] Not only does longitudinal extension occur along the axis of the fibers, but each fiber can also interact with neighboring fibers laterally, perpendicular to the fiber axis (Figure 1.9a). The second of these examples has been dubbed the "MagicWand" protein by Woolfson *et al.* as a result of its "+ - - +" charge pattern of the heptads (Table 1.1).[48] Similar to the previous example, the MagicWand also possesses sticky ends. The sequence and assembly of this coiled-coil dimer is demonstrated in Figure 1.9b and c, respectively.

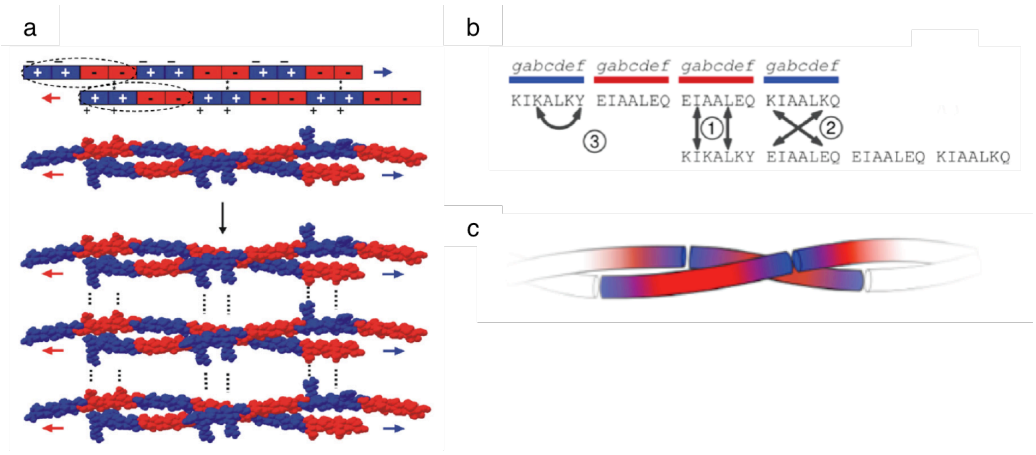



Figure 1.9. Electrostatic assembly of dimeric coiled-coils by SAFs with electrostatic charge patterns of "+ + - -" (a)[40] and "+ - - +" (b, c),[48] where both assemblies take advantage of electrostatic interactions and sticky ends. Blue represents positively charged heptads, red represents negatively charged heptads. Sequence of the MagicWand (b) leads to efficient self-assembly (c).[48]

Concentrations necessary to achieve these structures are typically greater than 50 μM,[41,43,51,53] with only a few examples of stable structures formed at lower concentrations.[41,54] It is our aim to be able to form fibers that self-assemble via electrostatic charge patches with diameters ranging from the nanometer to micrometer scale with concentrations less than 50 μM.

#### 1.2.2.2 3D domain swapping

Three dimensional (3D) domain swapping is yet another method for constructing fibrous protein materials that self-assemble in a defined manner. Specifically, it is a mechanism in which an identical structural element (a "domain") is exchanged within a protein or between proteins in order to generate oligomeric assemblies.[55] The most well-known cases of 3D domain swapping involve swapping regions present in the N and C-termini of proteins. In many of these cases, half the protein is swapped with respect to the other.[56,57]

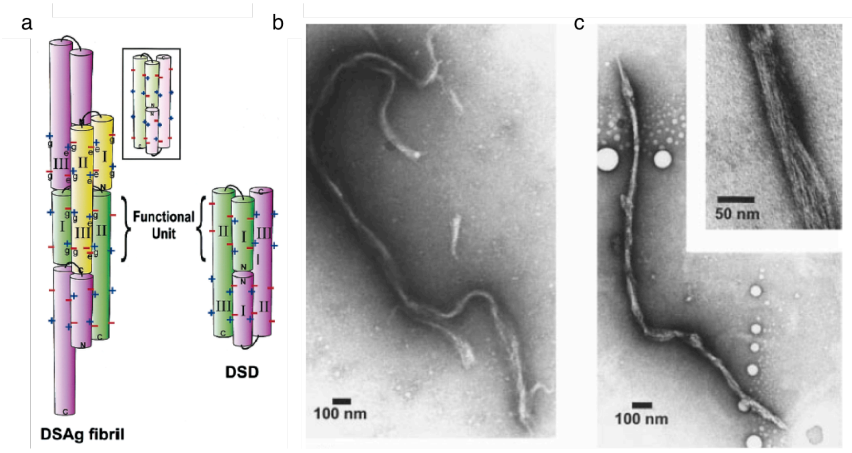



Figure 1.10. 3D domain swapped DSAg protein designed to form fibrils (a), demonstrated by TEM under neutral conditions (pH 6.5) (b) and acidic conditions (pH 2.4) (c). Inset of (c) shows clear protofibril formation.[49]

3D domain swapping has been used to generate three-dimensional fibers that are composed of bundles of protofibrils (Figure 1.10). Fibers with distinct protofibrils have been constructed from DSAg proteins through swapping portions of $\alpha$-helical subunits in the 3-helix bundle, using design principles of 3D domain swapping to create fibers 30-70 nm across with protofibrils of width corresponding to the trimeric coiled-coil (Table 1.1).[49] Ogihara *et al.* have created a domain swapped dimer (DSD) and demonstrated how the technique of 3D domain swapping can be extended to domain swapped oligomers. The fibrils seen in Figure 1.10b and c possess diameters between 30-80 nm, composed of bundles of protofibrils (inset Figure 1.10c) possessing widths comparable to the diameter of the 3-helix bundle. This observation confirms that the protofibrils are indeed consistent with the designed DSAg fibril.[49]

The ability to design fibrous proteins through electrostatic interactions and 3D domain swapping offers the possibility to program defined properties useful in many applications into self-assembling protein materials. Not only has much work been done on the design of these fibers, but external triggers may also be used to modify the structure and function of these materials, as is discussed in the following section.

### *1.2.3 Control of fiber assembly*

The ability to rationally design fibers that have known mechanisms of assembly is made multitudes more powerful by being able to control, or trigger, the assembly or disassembly of those structures. Proteins are naturally subject to structural, and therefore functional, changes by external triggers.[58–62] In living organisms these versatile biomolecules are known to fluctuate between native and denatured states, depending on ambient solvent conditions. The structural conformation of the protein governs its interactions with other protein subunits and their assembly, ligand binding capabilities, and its overall activity or function.

Switching states of proteins that self-assemble to form fibers can be triggered through a number of variable conditions. Herein we will explore several proteins that have been programmed to alternate states by way of modifying pH conditions, temperature, and through the ligand binding.

#### 1.2.3.1 pH control

Switching of proteins between different structural states can be designed according to the identity of amino acids in distinct positions of the heptad repeats within monomers of a coiled-coil, given by the pattern $(HPPHPPP)_n$. For example, the transition from $\beta$-sheet under acidic conditions (pH 4.0) to $\alpha$-helix at neutral pH (7.4) has been programmed by Pagel *et al.* by including L in positions *a* and *d*, Q and K in positions *e* and *g*, K in positions *b*, *f*, and *e*, and V in positions *b*, *c*, and *f* in the peptide sequence of VW19 (Table 1.2).[63] Self-assembly of $\alpha$-helical fibers formed from axial staggering of pentameric subunits (called αFPP) has been described by Potekhin *et al.*, where these fibers were shown to self-assemble into stable homopentamers producing fibers of controlled diameters under slightly acidic conditions (pH < 5.5) (Table 1.2).[47] In this case, the peptide sequence is defined by L in positions *a* and *d*, A and Q in positions *e* and *g*, respectively.[47] At neutral conditions, the peptide forms into spherical aggregates.

Programming the assembly of a coiled-coil protein based on pH conditions has also been accomplished through the design of modular cationic and anionic heptads. Woolfson *et al.* have created a library of homodimeric SAFs through this technique,

whose structure and assembly exhibited varying susceptibility to pH.[42] Sticky ended coiled-coils with charged patches comprised of the homodimeric Construct II peptides produce dense networks of fibers, but disassemble under acidic pH conditions (Table 1.1). Blunt ended coiled-coil assemblies were shown to be $\alpha$-helical at pH 7, but were unable to assemble into higher order structures. Upon exposure of these blunt ended proteins to elevated temperatures (up to 100 °C), these proteins underwent an irreversible $\alpha$ to $\beta$ transition[64] and the formation of amyloid-like fibers immune to a pH shift between 7 and 9 (Table 1.2).[42]

Table 1.2. pH sensitivity of select coiled-coil proteins.[31]

| Name | Sequence | Oligomerization state | pH sensitivity | Ref. |
|---|---|---|---|---|
| VW19 | LKVE-LKELKKE-LVVLKSE-LKELKKE-L | Homodimer | *α*-helical at pH > 4, positive charges on K in solvent exposed positions destabilizes *α*-helix at pH < 4 | 63 |
| αFFP | QLAREL-$(QQLAREL)_4$ | Homopentamer | High *α*-helical content at pH < 5.5, assembles into long fibrils 2.5 nm in diameter in acidic conditions | 47 |
| Construct I | 2 x (KIAALKQ-EIAALEQ-KIAALKQ-EIAALEQ) | Homodimer | *α*-helical at pH 7 | 42 |
| Construct II | 2 x (KIAALKQ-KIAALKQ-EIAALEQ-EIAALEQ) | Homodimer | *α*-helical at pH 7, disassembles at low pH | |
| Construct V | $C_{18}$-KIAALEQKIAALEY | Blunt-ended homodimers | *α*-helical at pH 7 but unable to assemble into higher order structures, formation of amyloid-like fibers immune to a pH shift between 7 - 9 | |
| Construct VI | KICCLKQKICCLKQEIAALEYEIAALEQ | Homodimer | Uniform fibrils at pH 7.4, undergoes *α*➔*β* transitions under mild alkaline conditions (pH 7.5-8). Fibrils remain intact up to pH 8.5. | |
| CPII | $(PRG)_4$-$(PHYPG)_4$-$(QHYPG)_4$ | Homotrimer | Stable triple helix at acidic conditions, forms nanofibers by adjusting pH to 7 via a high salt buffer | 65 |
| TZ1H | Ac-E-IAQHEKE-IQAIEKK-IAQHEYK-IQAIEEK-IAQHKEK-IQAIK-$NH_2$ | Homotrimer | *α*-helix at pH range 6.5–8; Random coil at pH > 5.5 | 66 |
| PA | See Appendix 5.5.2 | Homomultimer | Self-assemble to form insoluble materials at pH < 4, disassemble at neutral pH. Long-term stability at pH 8 achieved through oxidation. | 67 |
| CCSL | Ac-CAALESE-VSALESE-VASLESE-VAALGRGDMP-LAAVKSK-LSAVKSK-LASVKSK-LAAC-$NH_2$ | Intramolecular dimer | 80 % *α*-helical content in the pH range of 0.6-9.4 | 64 |
| TM-36 | Ac-$(KLEALEG)_5$-K-$NH_2$ | Homodimer | Coiled-coil is more stable at pH 2.5 than 7 | 68 |
| E/E35 | CGGDSE-$(VSSLESE)_3$-VSSLESKY | Heterodimer | Highest *α*-helix content of E/E35:K/K35 heterodimer is at pH = 5. At pH < 5, $(E/E35)_2$ homodimers are formed, while at pH > 11 $(K/K35)_2$ homodimers are formed | 69 |
| K/K35 | CGGDSK-$(VSSLKSK)_3$-VSSLKSKY | | | |
| IZ-2dE | YGG-IEKKIEA-EEKKIEA $(IEKKIEA)_2$ | Homotrimer | *α*-helix at pH range 4–5; Random coil at pH > 6 | 70 |
| IZ-2aE | YGG-IEKKIEA-IEKEIEA-$(IEKKIEA)_2$ | | | |
| EQ, CHAIN I | AC-Q-CGALEKQ-LGALEKQ-AGA-LEKQ-LGALEKQ-LGALEK-$NH_2$ | Homodimer | Coiled-coil more stable at pH 3 than pH 7 | 71 |
| EQ, CHAIN II | AC-Q-CGALEKQ-LGALEKQ-AGA-LEKQ-LGALEKQ-LGALEK-$NH_2$ | | | |

There are indeed many examples of coiled-coil rearrangement as a result of pH shifts, and a collection of these are presented in Table 1.2. A variant of CMP termed CPII, which forms a stable triple helix at acidic conditions, may be triggered to form nanofibers displaying a high level of organization through interdigitation of the triple helices by adjusting the pH to 7 via a high salt buffer (Table 1.2).[65] Another homotrimer, TZ1H, has been designed by Conticello *et al.* to explicitly assemble and disassemble in response to changes in pH (Table 1.2).[66] CD spectroscopy has been used to study pH-dependency of TZ1H from pH 4.0 to 8.0. A strong $\alpha$-helical signal between pH 6.5 and 8.0 is observed, whereas the pH range of 4.0 to 5.6 results in a loss of $\alpha$-helical signal and shows the onset of random coil formation, with the helix-coil transition defined at 203 nm in the CD signature.[66] Peptide-amphiphiles (PA) have also been designed to respond to changes in pH, where these molecules self-assemble at pHs below 4 and disassemble when the pH is increased to neutral (Table 1.2).[67] PA molecules (Appendix 5.5.2) self-assemble as a homomultimeric cylindrical micelle.[67] Interestingly, oxidation of PAs imparted a long-term stability at pH 8, where fibers formed did not disassemble over months, whereas non-oxidized versions disassembled within minutes. Reduction of the PAs by dithiothreitol allowed the nanofibers to regain their pH sensitivity and rapidly disassemble at pH 8.[67] The ability to control assembly of fibrous protein materials via pH can be exploited for use in controlled release of ligand molecules.

#### 1.2.3.2 Temperature control

Temperature control is a very intuitive method of provoking a fibrous protein assembly to dissociate, as all proteins demonstrate a melting point above which they denature and lose structure. Structural denaturation of proteins as a result of surpassing their melting point is often irreversible, as denatured proteins frequently aggregate at elevated temperatures. Aggregate protein solutions typically become turbid, facilitating structural characterization as a function of temperature by many techniques, including CD and UV-vis spectroscopy,[72] as well as dynamic light scattering (DLS).[73]

Table 1.3. Peptides susceptible to temperature control.

| Name | Sequence | Temperature response behavior | Ref. |
|---|---|---|---|
| CMA | Type-I collagen (3 polypeptide chains that oligomerize into an extended triple helix), where each chain is a GXY sequence, where X and Y represent any amino acid but primarily are Hyp and P. CMA is Type-I collagen with covalently bound methacrylate groups on the free amines of the lysine groups | Thermoreversible oscillating between a liquid macromer suspension and a semisolid fibrillar hydrogel between 10 and 37 °C | [73] |
| X22d | CGGE-VGALKAE-VGALKAQ-IGA**X**QKQ-IGALQKE-VGALKK<br>Central *d* position (**X**) was systematically substituted with L, M, I, Y, F, V, E, A, W, N, H, T, K, S, N, Q, R, and G. | Decreasing order of thermodynamic stability for *d* mutants of: L, M, I, Y, F, V, E, A, W, N, H, T, K, S, N, Q, R, and G. Overall stability range from lowest analog to highest was 7.4 kcal/mol. | [74] |
| X19a | CGGE-VGALKAQ-VGALQAQ-**X**GALQKE-VGALKKE-VGALKK<br>*a* position (**X**) was systematically substituted with I, L, N, A and E. | Decreasing order of thermodynamic stability for *a* mutants of: I, L, N, A and E. Overall stability range from lowest analog to highest was 2.2 kcal/mol. | [75] |

Nanda *et al.* have recently developed a derivative of Type I collagen with methacrylate groups covalently bound to free amines on lysine groups of the Type I collagen, creating so-called CMA (Table 1.3). Interestingly, CMA is able to thermoreversibly self-assemble into porous fiber networks, which is demonstrated on the molecular and supramolecular scale via CD, DLS, micro-rheology, TEM, and SEM.[73] As can be seen in Figure 1.11a, CMA demonstrates a fully recoverable storage modulus with temperature cycling between 4 and 37 °C and the ability to form organized, self-assembled fibers both before melting and upon cooling (Figure 1.11b-e).

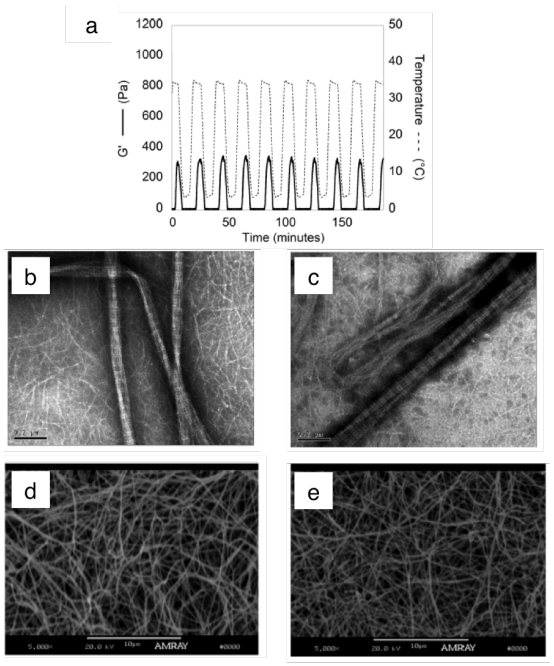



Figure 1.11. Self-assembly of CMA with respect to temperature cycling, demonstrating thermoreversibility of storage modulus *G'* between 4 and 37 °C (a), fiber formation before melting (b, d) and upon cooling (c, e).[73]

Hodges *et al.* investigated the contribution of different residues placed at heptad positions *a* or *d* in dimeric coiled-coils five heptads in length (Table 1.3).[74,75] These studies categorized a decreasing order of thermodynamic stability for *a* mutants of: I, L, N, A and E, and for *d* mutants of: L, M, I, Y, F, V, E, A, W, N, H, T, K, S, N, Q, R, and G.[31,74,75] Identity of residues in positions *a* and *d* could therefore be programmed to achieve increased thermal stability of coiled-coil assemblies.

### 1.2.3.3 Additives

Designs of self-assembling proteins based on ligand binding offer a natural mechanism of assembly control based on conformational changes proteins undergo upon binding. The number of ligands that can potentially influence structure of fibrous proteins is virtually limitless, but in this particular section we will limit our focus to the polyphenolic compound curcumin and divalent metal ions. In Sections 1.4 and 2.4.5 we present experiments studying the effect of an organic solvent called trifluoroethanol (TFE) on protein structure and assembly, but exclude TFE from this introductory section, as it does not quality as a ligand, per se.

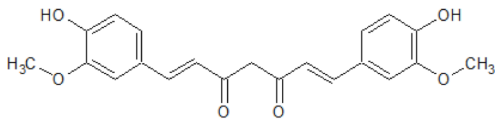



Figure 1.12. Molecular structure of curcumin.

The polyphenolic compound curcumin (Figure 1.12), has long been used for many therapeutic purposes due to its antiproliferative,[76] antibacterial,[77] and anti-inflammatory[78] properties, but exhibits limitations in delivery methods due to its low solubility in aqueous solutions.[79] We therefore choose to examine the binding of curcumin with protein fibers under acidic conditions to attempt to embed the polyphenolic molecule within the protein superstructures, as presented in Section 1.4.6. Curcumin has been shown to induce aggregation of protein fibers, such as collagen and the acidic $\alpha$-helical intermediate of PrP, a precursor to amyloid fibers.[79–81] Fathima *et al.* have characterized the curcumin-induced aggregation mechanism of collagen using CD, surface tension, and viscosity measurements.[79,80] At pH 4, fully protonated curcumin will have strong interactions with negatively charged amino acids on protein surfaces. Morphology and aggregation of $\alpha$-helical protein fibers results from an energetically favorable balance between the burial of the hydrophobic residues and the solvation of the surface-facing polar groups.[79] In literature, CD studies have confirmed that increasing molar ratios of curcumin affects only helical packing, not protein conformation.[80] Curcumin increases the surface activity of collagen, as a result of local restructuring of water leading to some exposure of non-polar groups.[80] The increased exposure of non-polar groups provides a driving force for aggregation, as the system strives to reduce overall surface energy, explaining the mechanism of curcumin-induced aggregation of $\alpha$-helical protein fibers.

Table 1.4. Peptides susceptible to control by additives.

| Name | Sequence | Response to additives | Ref. |
|---|---|---|---|
| H-byp | $(PXG)_4$ with a bipyridyl-modified lysine residue as X | $Fe^{2+}$ triggered formation of branched fibers. $Cu^{2+}$ was also seen to promote assembly | 82 |
| TZ1H | EIAQHEKEIQAIEKKIAQHEYKI QAIEEKIAQHKEKIQAIK | In the absence of $Ag^{+}$, no fibers seen; long aspect ratio fibers seen in the presence of equimolar $Ag^{+}$ | 83 |
| ZiCo | YIHALHRKAFAKIARLERHIRA LEHAA | Switches between a coiled-coil state to zinc finger type motif upon binding to $Zn^{2+}$ | 84 |
| B18 | LGLLLRHLRHHSNLLANI | Transforms from random to *α*-helical structure when bound to $Zn^{2+}$ and mimicked the function of the parent protein | 85 |
| IZ-3aH | YGGIEKKIEAIEKKIEAIEKKHE AIEKKIEA | Differentially binds divalent metals with affinity towards $Cu^{2+}$ and $Zn^{2+}$ and assembles into a trimeric coiled-coil but is incapable of binding $Ni^{2+}$ | 86 |
| Mastoparan X | INWKGIAAMAKKLL. Divalent metal binding site was incorporated by adding two histidine residues or a glutamic acid and a histidine residue were built into *i*, *i+4* positions | Addition of $Zn^{2+}$ does not affect the structure or lytic ability of one of the designed peptides, Mst-HAH, $Ni^{2+}$ binding promoted helix formation and hemolysis | 87 |
| $His_6$-C | MRGSHHHHHHGSGDLAPQML RELQETNAALQDVRELLRQQV KEITFLKNTVMESDASGKLN | $His_6$-C and $His_6$-T40A self-assemble into fibers in the absence of metals, in the presence of $Zn^{2+}$ fibers interact and aggregate, in the presence of $Ni^{2+}$ fibers assemble into micrometer sized aggregates disrupting $\alpha$-helices | 88 |
| $His_6$-T40A | MRGSHHHHHHGSGDLAPQML RELQE**A**NAALQDVRELLRQQV KEITFLKNTVMESDASGKLN | | |
| $His_6$-L44A | MRGSHHHHHHGSGDLAPQML RELQETNAA**A**QDVRELLRQQV KEITFLKNTVMESDASGKLN | Incapable of forming $\alpha$-helical structures. Particles seen in the presence and absence of $Cu^{2+}$, $Ni^{2+}$, $Zn^{2+}$ | |

Metal ion binding has been studied by many groups, with a wide variety of metal triggers.[82,88] The structures of many proteins are naturally triggered to change upon complexation with several divalent metal ions, such as $Ag^{2+}$, $Cu^{2+}$, $Ni^{2+}$, $Au^{2+}$, $Fe^{2+}$, and $Zn^{2+}$, among others. The collagen triple helix has been used in the design of metal-triggered radial growth of the helices, called H-byp, into fibrils (Table 1.4).[82] The interaction of H-byp with $Fe^{2+}$ was directed through the incorporation of a bipyridyl-

modified lysine residue in the place of the hydroxyproline residue in the collagen tripeptide sequence. This study demonstrated a triggered formation of fibers through the addition of $Fe^{2+}$ with the appearance of branched fibers occurring below the stoichiometry of ligand-metal binding.[82]

Dublin *et al.* described the manipulation of trimeric coiled-coil TZ1H peptide by the addition of soluble $Ag^{+}$ ions (Table 1.4).[83] This trimer possesses layers of three proximal histidine residues within the coiled-coil pore and therefore provide a potential metal ion binding site which the authors hypothesized could be used to modify self-assembly.[83] Under solution conditions of 10 mM PB pH 5.6 and in the absence of $Ag^{+}$ TZ1H did not form fibers, whereas the addition of an equimolar amount of $Ag^{+}$ promoted the formation of long aspect ratio fibers.[83]

Zinc is a well-studied additive capable of affecting protein structure.[84–88] It is omnipresent in biological systems and its presence often acts as a trigger for the procedure of necessary cellular functions.[85,89–91] Woolfson *et al.* have described the synthesis of a *de novo* peptide with structural duality called ZiCo (Table 1.4).[84] They designed a reversible switch manipulated through metal binding that changes the structure of ZiCo from one folded state to another upon introduction of $Zn^{2+}$.[84] ZiCo was shown to reversibly switch between a trimeric $\alpha$-helical coiled-coil to a zinc-bound folded monomer in the presence of $Zn^{2+}$.[84] The influence of $Zn^{2+}$ has also been studied with the B18 peptide, representing the minimal membrane-binding motif of a membrane-associated acrosomal protein called bindin from the sea urchin (Table 1.4).[85] Authors found that B18, similar to its parent protein, was able to bind $Zn^{2+}$ through a histidine rich motif that results in induction of $\alpha$-helical conformation of B18 and subsequent oligomerization of the peptide.[85] Kiyokawa *et al.* studied the interaction of de novo peptide IZ-3aH (Table 1.4) with $Zn^{2+}$ and $Cu^{2+}$, and $Ni^{2+}$.[86] The trimeric coiled-coil IZ-3aH was able to bind $Cu^{2+}$ and $Zn^{2+}$ with an increase in $\alpha$-helicity through the presence of a histidine residue in the sequence, but not able to bind $Ni^{2+}$.[86] DeGrado *et al.* engineered the $\alpha$-helical amphiphilic cell-lytic peptide Mastoparan X (Table 1.4) to bind with divalent cations including $Zn^{2+}$ and $Ni^{2+}$.[87] One of the variants produced, Mst-HAH, replaced the helix-breaking G in the fifth position with A in an effort to increase helical

conformation.[87] This mutant was found to bind $Ni^{2+}$ with enhanced helicity and hemolytic properties but not $Zn^{2+}$, which had no effect on helicity or enzymatic ability.[87]

The interaction of pentameric coiled-coil COMPcc (described further in Section 1.2.4) and its mutants has been examined by Gunasekar *et al.* of our group, specifically studying the effect of $Cu^{2+}$, $Ni^{2+}$, and $Zn^{2+}$ on fiber formation (Table 1.4).[88] $Zn^{2+}$ significantly increases $\alpha$-helical content of the protein systems, where $Ni^{2+}$ has the opposite effect. In accordance with the structural effects of the divalent metals, the binding ability of the pentameric proteins to curcumin is similarly affected.[88] The effect of gold and iron oxides on protein structure is discussed in greater detail in Sections 2 and 3.

### *1.2.4 COMPcc*

Our system is based on the coiled-coil domain of cartilage oligomeric matrix protein.[33,92] Cartilage oligomeric matrix protein (COMP) is found in cartilage, tendon, and ligament tissue.[32] $\alpha$-helical COMP assembles into a pentameric bouquet composed or five equal subunits which arrange to form a coiled-coil structure (Figure 1.13).[33] This protein is comprised of various domains, but its ability to assume a pentameric structure is attributed to its N-terminal coiled-coil region, denoted COMPcc.[34] The pentamer is stabilized by electrostatic interactions between aligned heptad units, creating a 73 Å long hydrophobic core 2-6 Å in width (Figure 1.13b) between subunit chains, as determined from the crystal structure (PDB file 1VDF). [34-35] The outer diameter of the COMPcc pentamer is approximately 30 Å.[93] Cysteine residues (positions 68 and 71) in COMPcc create inter-chain disulfide bridges between strands, as is depicted in Figure 1.13c.

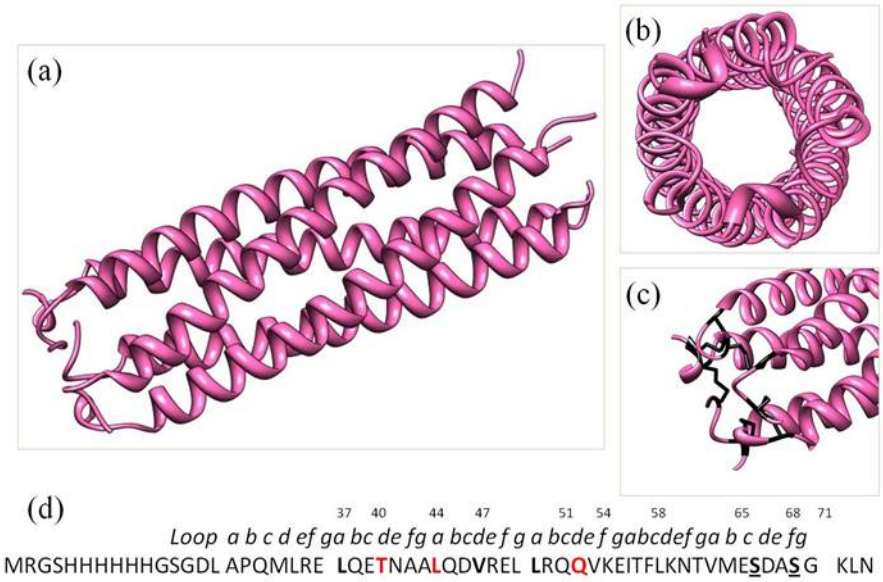



Figure 1.13. Crystal structure of COMPcc depicts a homopentamer 73 Å in length (a), a hydrophobic pore 2-6 Å in width (viewed from the C-terminal end of the protein) (b), and a C-terminus with serine residues shown in black creating inter-chain disulfide bonds (c). The primary amino acid sequence of COMPcc$^{s}$ with respective heptad positions as well as the corresponding residue numbers is shown in (d).

Previously, our group has engineered COMPcc mutating two cysteine residues (C68 and C71 in the original sequence), responsible for pentamer-stabilizing disulfide bonds, to serine to investigate the protein without complication of oxidation.[92] The COMPcc mutant with serine in positions 68 and 71 is referred to as wt (also referred to as COMPcc$^{s}$). The mutation of the cysteines to serines has been performed to slightly destabilize the pentamer and enable the study of the protein on a single strand basis. The primary sequence of wt is given in Figure 1.13d, where letters *a-f* denote a residue's position in the helical wheel.

As is expected with coiled-coil assemblies, the residues inhabiting the *a* and *d* positions of COMPcc are hydrophobic and non-polar. One of these residues, glutamine in *d* position 54, located centrally within the protein sequence, acts as a divider within the pore, separating it into two cavities. These cavities are where small molecules have been shown to bind within the hydrophobic core of wt, with volume capacities of 450 and 515 Å$^3$ for the N-terminal and C-terminal cavities, respectively.[94] The wt protein consists of a repeat of three leucines (L37, L44, L51) in the *a* site within the N-terminal pocket along

with a valine (V47) in the *d* site of the adjacent helix, known to be indispensable for the formation of stable pentamers (Figure 1.13a).[92]

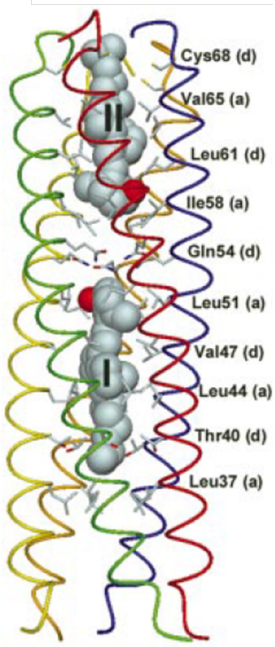



Figure 1.14. COMPcc complexed with two molecules of vitamin $D_3$, one in each binding pocket. The ion channel within COMPcc is divided in two cavities, with residue Q54 defining the separation.[94]

Wt has demonstrated binding ability to all-*trans*-retinol, retinoic acid, vitamin $D_3$, benzene, and other small hydrophobic molecules.[92,94,95] Recently, the crystal structures of COMPcc and several fatty acids has also been solved.[93] These fatty acids are myristic acid, palmitic acid, stearic acid, and oleic acid.[93] Importantly, only fatty acids possessing an elongated configuration are able to selectively bind with COMPcc due to the restriction created by the five Met33 residues at the entrance of the ion channel.[93] This study is also influential because McFarlane *et al.* demonstrated that palmitic acid bound not only within the hydrophobic channel of COMPcc, but also on the outside of the pentamer,[93] as is shown in Figure 1.15.

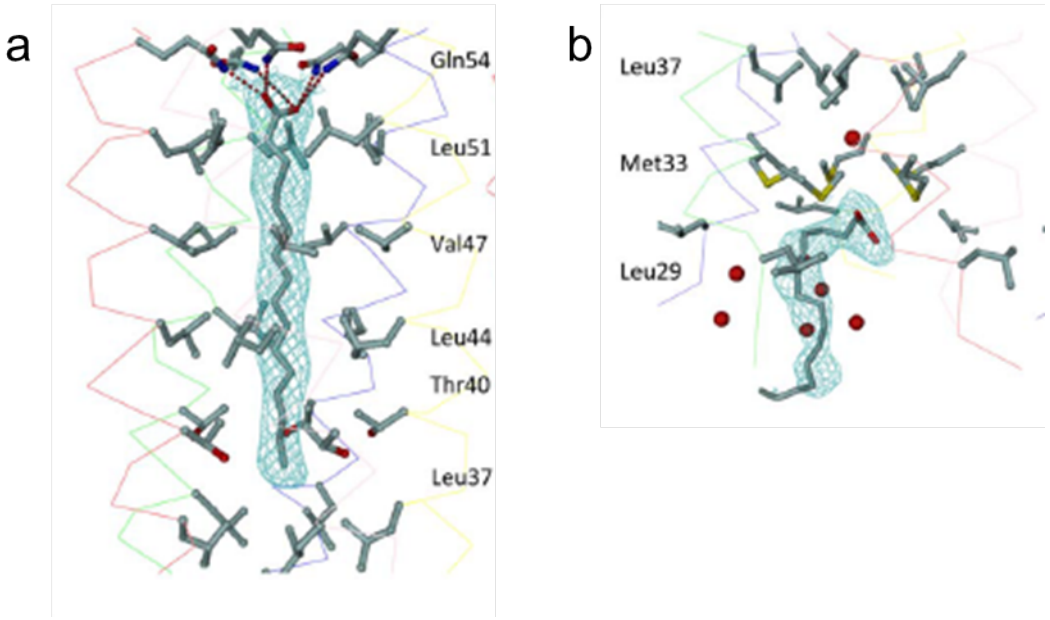



Figure 1.15. Palmitic acid binding inside (a) and outside (b) the hydrophobic channel of COMPcc.[93]

The binding ability of COMPcc to these molecules is attributed to the structure of the pentamer, specifically associated with several important aliphatic residues, namely L37, L44, V47, and L51.[92] Mutation of these residues in the *a* and *d* positions along the pore of the wt protein to alanine have been shown to be responsible for a loss in binding ability to small hydrophobic molecules, including curcumin.[92] While mutation of these aliphatic residues to alanine results in a loss of secondary structure of COMPcc and binding ability, mutation of two polar residues, T40 and Q54, into alanine results in increased α-helical stability and binding capability.[88,92] This knowledge has been implemented in the design of several novel proteins used in this work, and are described in detail in the following section.

### *1.2.5 Design methodology*

Recently, we discovered that wt assembles into fibers with diameters of 10-15 nm (Figure 1.51).[88] In addition to the nanofibers, nanoparticles were also formed, leading to a mixture. This was used as inspiration to study the effect of binding of curcumin to engineered proteins with swapped domains, leading to unexpected and influential results.

Novel proteins have been designed in a manner to homogeneously self-assemble based on electrostatic interactions between heptad repeats, the sequences for which are

given in Table 1.5. In all engineered COMPcc variants described herein an IEGR recognition sequence has been engineered into the primary sequence directly C-terminal to the His tag. This IEGR cleavage site allows for facile removal of the His tag via enzymatic cleavage by Factor Xa protease (discussed in Section 2.2.3.4).

Table 1.5. Protein sequence information for wt, C, Q, and L, from N-terminus to C-terminus, with the histidine tags in bold.

| | His tag | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg | |
|---|---|---|---|---|---|---|---|---|
| | *15* | *30* | *37* | *44* | *51* | *58* | *65* | *72* |
| wt | MRGS**HHHHHH**GSGDL | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | VMESDAS | GKLN |
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* | |
| C | MRGS**HHHHHH**GSIEGR | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL | |
| | *1* | *17* | *20* | *27* | *34* | *41* | *48* | |
| Q | MRGS**HHHHHH**GSIEGR | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL | |
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* | |
| L | MRGS**HHHHHH**GSIEGR | LQDVREL | LRQQVKE | ITFLKNT | APQMLRE | LQETNAA | SKL | |

The following subsections describe the novel proteins designed to self-assemble by truncating and swapping regions of wt, dubbed C and Q, and a negative control swap protein, called L.

#### 1.2.5.1 C

C was generated by truncating the original sequence of wt by one heptad repeat and adding the IEGR cleavage site (*vida supra*). The crystal structure of wt indicates that the C-terminus has higher number of residues in the loop region compared with the N-terminus. As such, the VMESDAS heptad in the loop region has been removed by truncating the wt sequence after the ITFLKNT heptad. The resulting sequence for C contains five complete heptads, as can be seen in Table 1.5. The helical wheel for C is shown in Figure 1.16.

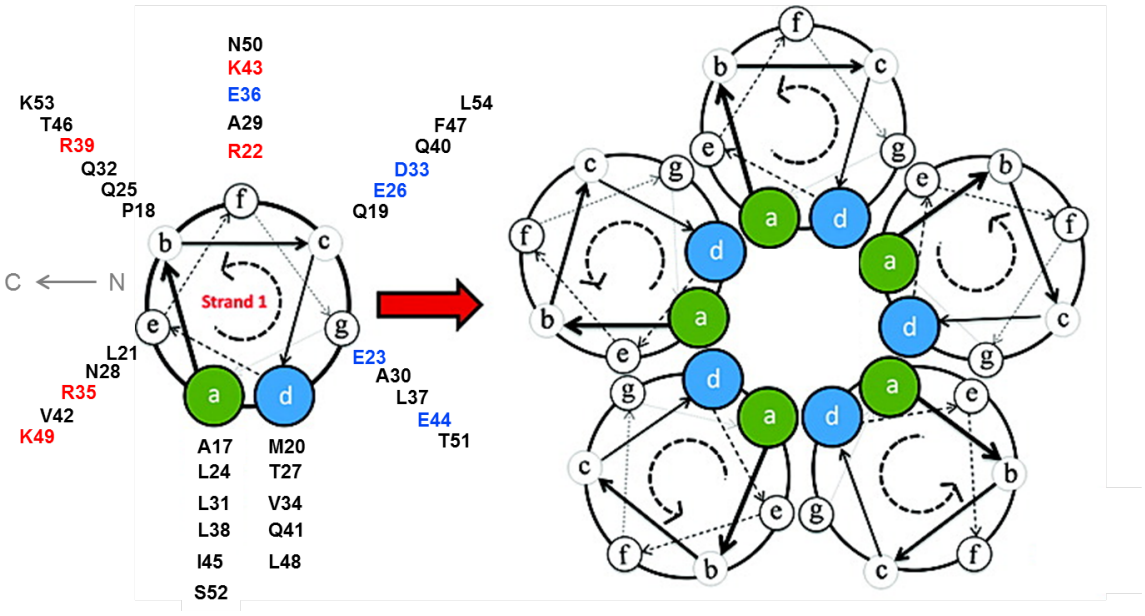



Figure 1.16. Helical wheel for C, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K43 and K49, in the *f* and *e* positions, respectively) and arginine residues (R22 and R39, in the *f* and *b* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E23, E26, and E36 in the *g*, *c*, and *f* positions, respectively) and asparagine residues (D33 in the *c* position). Individual $\alpha$-helices self-assemble to form homomeric pentamers.

#### 1.2.5.2 Q

Q was designed consistent with two of the same features as C; namely the insertion of the IEGR cleavage site and the removal of the extraneous last heptad of the wt protein. As discussed in Section 1.2.4, the glutamine residue at position 54 in wt divides the hydrophobic core into two cavities, or binding pockets. When Q was mutated to alanine in wt, mutants show enhanced helical stability and increased melting temperature.[92] The Q protein designed in this work swaps the order of the N-terminal and C-terminal segments of wt with respect to Q54, maintaining each of the binding pockets intact. In Q, the C-terminal portion after the glutamine residue at the *d* position was swapped to the N-terminus to better distribute the surface charge for lateral assembly (Figure 1.35a). The resulting sequence for Q is given in Table 1.5 and the helical wheel for Q is shown in Figure 1.17.

The aliphatic residues that were determined to be crucial for pentamer formation and thus thermodynamic stability in wt (L37, L44, L51 in the *a* site and V47 in the *d*

site)[92] are maintained in the same heptad positions within Q (corresponding residues in Q being L34, L41, L48 and V44). Once again, the residues responsible for maintaining core interactions in the peptide are not disturbed, maintaining small molecule binding ability.[92]

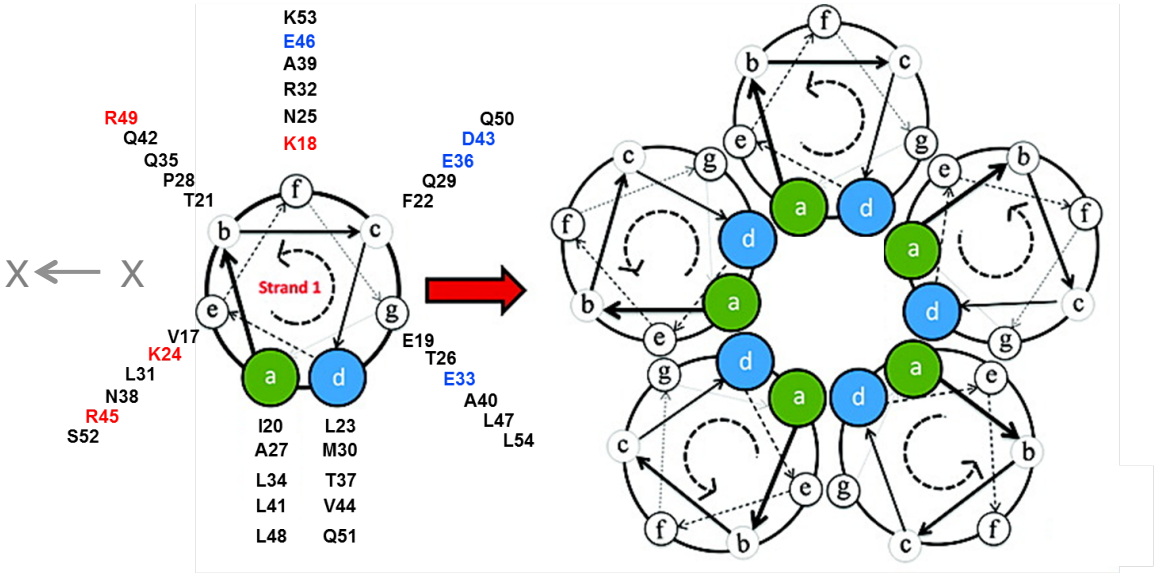



Figure 1.17. Helical wheel for Q, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K18 and K24, in the *f* and *e* positions, respectively) and arginine residues (R45 and R49, in the *e* and *b* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E33, E36, and E46 in the *g*, *c*, and *f* positions, respectively) and asparagine residues (D43 in the *c* position). Individual $\alpha$-helices self-assemble to form homomeric pentamers.

Q was originally designed in a manner to self-assemble with C based on electrostatic interactions between heptad repeats. The objective was that upon mixing, C and Q would self-assemble longitudinally to form fibers via sticky ends. The design proposed by combining C and Q is shown schematically with the amino acid sequences of the proteins in Figure 1.18.

...LQDVREL LRQQVKE ITFLKNT/APQMLRE LQETNAA LQDVREL LRQQVKE ITFLKNT ...
... LQDVREL LRQQ/VKE ITFLKNT APQMLRE LQETNAA LQDVREL LRQQ/VKE ITFLKNT ...

Figure 1.18. Sticky-ended fiber formed by longitudinal assembly of C (red) and Q (black). Slashes denote end/beginning of protein subunits.

There are of course several possible combinations of pentameric arrangements if C and Q were indeed staggered throughout the assembly. One possible combination would be one

C helix with four Q helices (Figure 1.19a, b) or one Q helix with four C helices (Figure 1.19c, d).

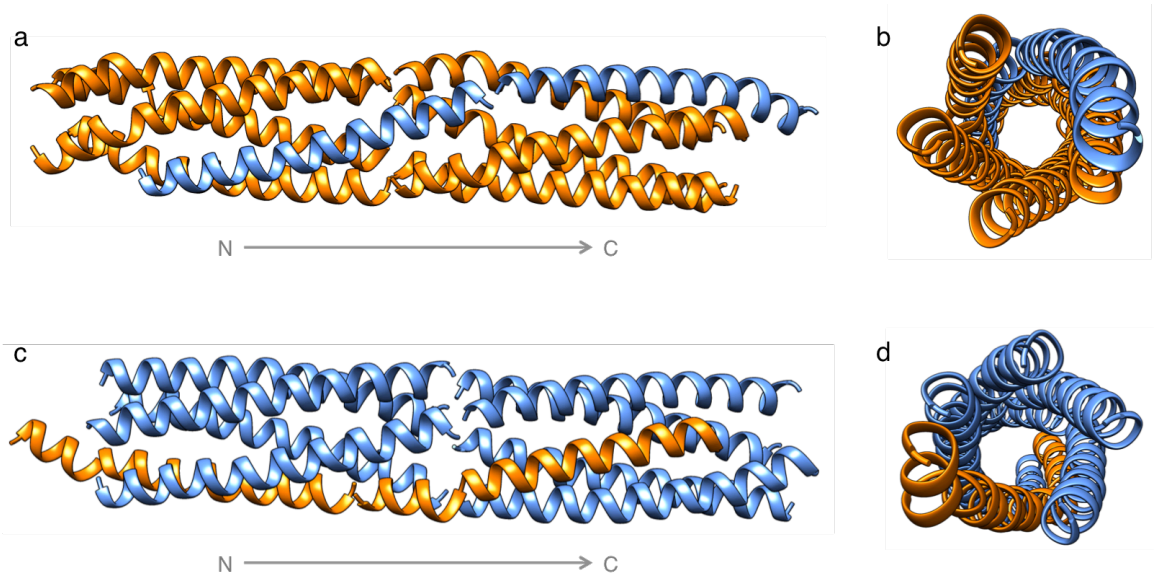



Figure 1.19. Longitudinally arranged pentameric assemblies comprised of 4 Q helices and 1 C helix (a, b) and 4 C helices and 1 Q helix (c, d), viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively.

The other possible heteropentameric arrangements of C and Q may adopt one of two possible configurations, which are referred to as A and B. A presents the case when C and Q are beside one another in the pentamers (Figure 1.20a, c) and B is the configuration in which they are interspersed (Figure 1.20b, d).

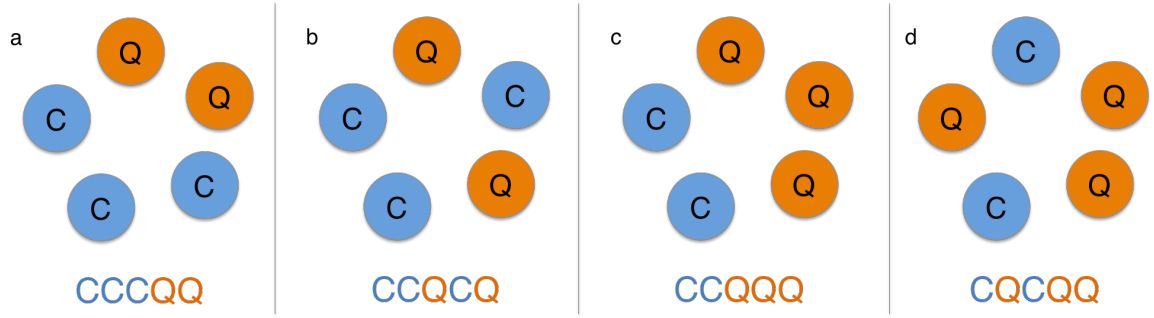



Figure 1.20. Cartoons of pentameric assemblies comprised of 2 Q helices and 3 C helices (a, b) and 3 Q helices and 2 C helices (c, d), viewed down the pentamer axis. In these configurations two distinct arrangements of the subunits are possible, which are referred to as A when C and Q are beside one another in the pentamers and B when they are interspersed. C and Q are represented in blue and orange, respectively.

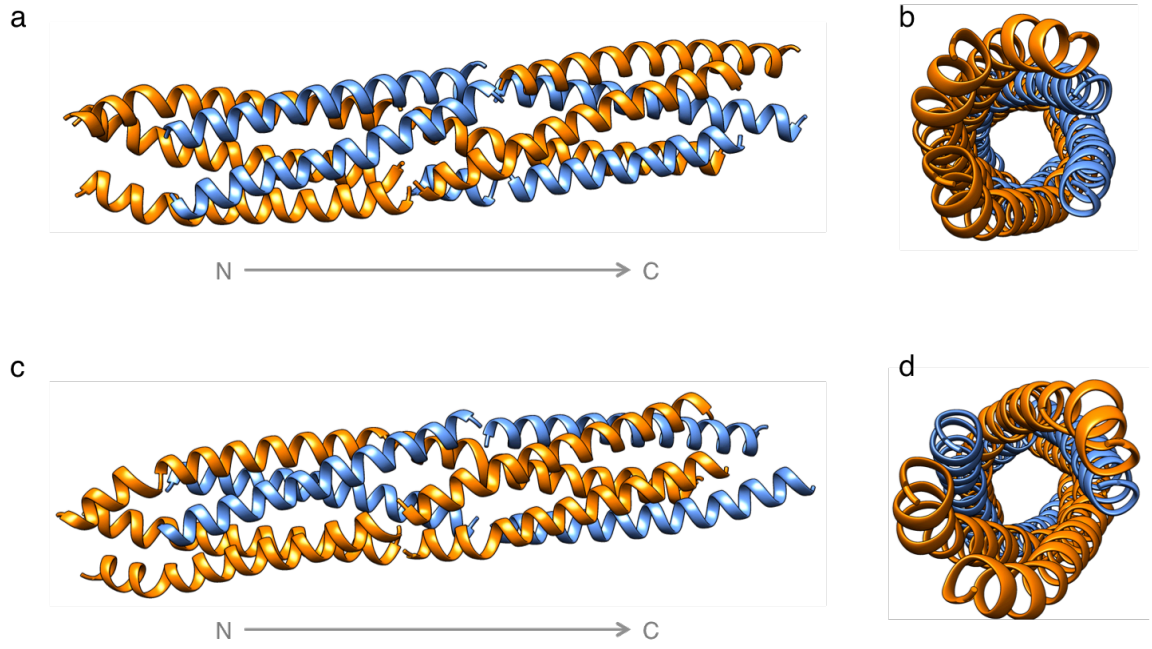



Figure 1.21. Longitudinally arranged pentameric assemblies comprised of 2 C helices and 3 Q helices in the A (a, b) and B (c, d) configurations, viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively.

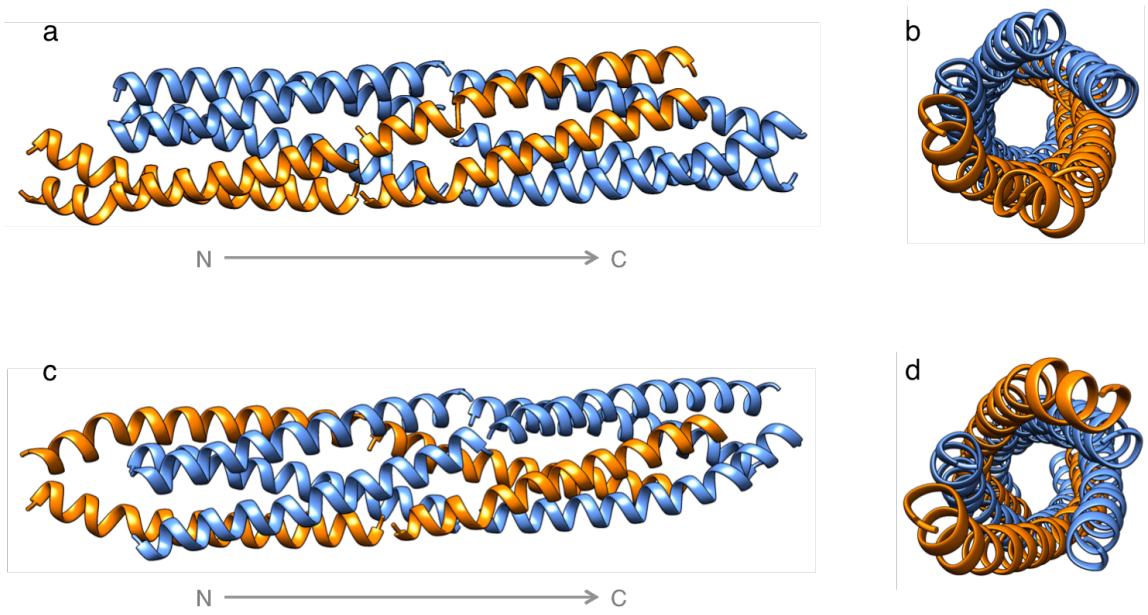



Figure 1.22. Longitudinally arranged pentameric assemblies comprised of 3 C helices and 2 Q helices in the A (a, b) and B (c, d) configurations, viewed along (a, c) and down (b, d) the pentamer axis. C and Q are represented in blue and orange, respectively.

Models of pentameric assemblies displaying these two types of configurations are shown in Figure 1.21 for the case of two C helices with three Q helices and in Figure 1.22 for the case of three C helices with two Q helices. As can be seen in Figure 1.19, Figure 1.21, and Figure 1.22 the interspersion of C and Q subunits within the pentamers results in overhanging portions of each protein, creating sticky ends by which longitudinal assembly can theoretically progress.[39,51,96]

### 1.2.5.3 L

In order to compare structure and assembly characteristics of engineered swap protein Q, we also designed a negative control, L, which was engineered to disrupt the N-terminal pocket crucial region for structure, stability, and pentamer formation by swapping at the leucine at position 44 (Figure 1.35a).[92] The L44 residue has been shown to be important for maintaining the oligomeric state of the protein, the mutation of which makes wt unable to form pentamers and decreases α-helical stability.[92] The resulting sequence of L is given in Table 1.5 and the helical wheel for L is shown in Figure 1.23.

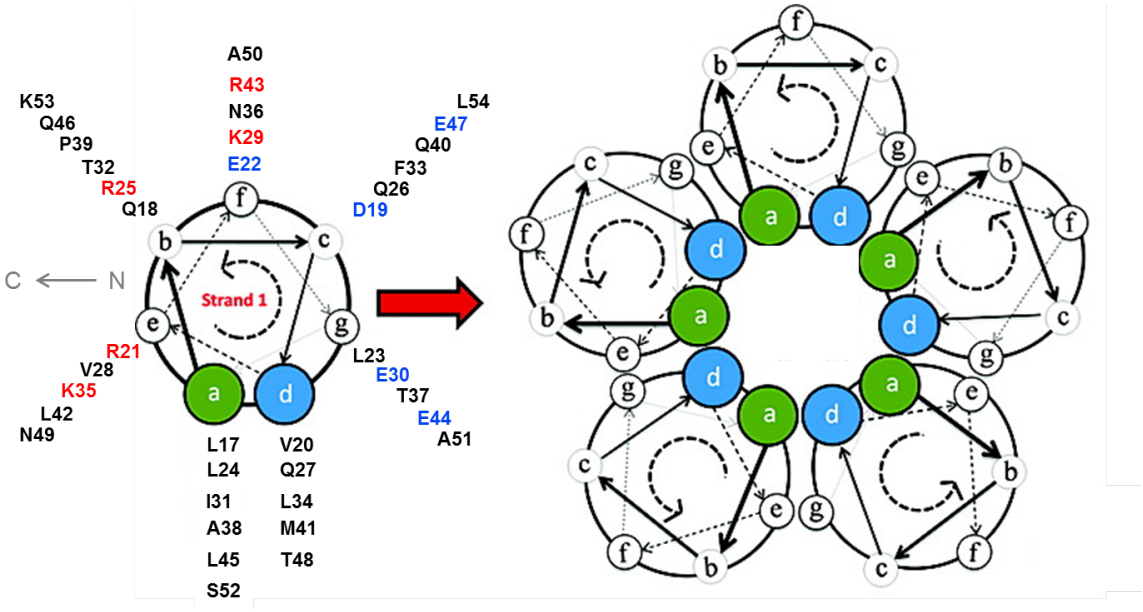



Figure 1.23. Helical wheel for L, showing the positioning of residues within heptad repeats. Residues highlighted in red contribute to positive surface charge, where residues highlighted in blue contribute to negative surface charge. Positive charge attributed to solvent-exposed lysine (K29 and K35, in the *f* and *e* positions, respectively) and arginine residues (R25 and R43, in the *b* and *f* positions, respectively). Negative charge attributed to solvent-exposed glutamine (E22, E30, and E47 in the *f*, *g*, and *c* positions, respectively) and asparagine residues (D19 in the *c* position). Individual *α*-helices self-assemble to form homomeric pentamers.

The inability of the negative control protein, L, to form fibers will prove the design methodology is sound. Characterization of secondary structure and macromolecular assembly of L will confirm its inability to self-assemble.

#### 1.2.5.4 A note on proline

Here we must address the proline residue in the *b* position of all three proteins. Proline is traditionally known as a helix breaking residue.[97] In wt and C, the proline occurs at the N-terminus of the protein, thus having a minimal effect on helical structure. Swapping of the N and C-termini of wt to generate Q and L results in proline being displaced to the center of the proteins, in positions 28 and 39 in Q and L, respectively. As the design of Q and L focuses on examining the effect of swapping the regions of wt, proline was not mutated out. This decision is made based on work that has shown that the presence of proline in the middle of an $\alpha$-helix can be accommodated by a local break in structure, where only the residue preceding the proline (A27 in Q and A38 in L) is not in $\alpha$-helical conformation.[98] The $\alpha$-helical structure resumes unbroken either side of the proline, however.[98]

### *1.2.6 Applications*

With all the work that has gone in to predictive structure-function relationships in the design of self-assembling protein-based materials, it is no surprise that they find many applications in diverse areas of bioengineering. As these naturally biocompatible materials can be programmed to self-assemble and disassemble by various triggers (Section 1.2.3), they are excellent candidates as scaffold materials in tissue engineering and as vehicles for drug delivery.

Fibers made of self-assembling proteins are ideal construction materials for scaffolds in tissue engineering. These materials are biocompatible, can be programmed to release therapeutic molecules and safely biodegrade with the formation of toxic degradation byproducts, and are able to form nanoporous, durable frameworks upon which cells can be conditioned to grow.[8,99,100] The remarkable feature of synthetic protein materials is their chemical and structural resemblance to natural structural proteins, and the facile introduction of these biomimetic materials into living systems.[6,8,101] Fibrinogen,

a natural structural protein involved in clotting, has been synthetically produced to form fibril scaffolds upon which the growth of hydroxyapatite crystals takes place.[25] Hydroxyapatite is an excellent bone implant material, making the use fibrinogen nanofibers of great interest in this area of tissue engineering. Fibrinogen nanofibrils have also been programmed to self-assemble on surfaces that can be used in tissue engineering and biomedicine.[102]

With the specific binding capability of many protein fibers, and more specifically, coiled-coils, the use of these systems as vehicles for drug delivery is being widely explored. Coiled-coils, when covalently bound to water soluble molecules, have been used to create hydrogels that undergo engineered volume changes, triggered by either temperature or pH.[103,104] COMPcc has been of particular interest in the area of drug delivery, as its hydrophobic channel and polar exterior has afforded it a "Trojan horse"-like quality that allows non-soluble drug molecules or genetic information to be safety transported to desired locations within its core for treatment of disease.[105,106] Elaborate networks of electrospun zein, a group of prolamines found in corn endosperm, have demonstrated the ability not only to form scaffolds, but also to be loaded with varying amounts of curcumin.[107] The subsequent enzymatic degradation of the zein provided a sustained release of curcumin over time, exemplifying the potential of these protein-based fibers to deliver drugs in a controlled fashion.[107] Self-assembling peptide materials are also of great interest in regenerative medicine, where cell signaling capabilities of proteins being are being exploited to complex structures with high therapeutic efficiency.[108]

Protein-based fibers and other self-assembling materials also find applications in other diverse fields. The use of these materials in chemical sensing, biomedical imaging, and bioelectronics applications through the templation of inorganic nanostructures are described in depth in Sections 2.2 and 3.2.

## 1.3 Materials and methods

### *1.3.1 Materials*

Sodium phosphate (monobasic and dibasic), trifluoroethanol, and nickel-nitrilotriacetic acid resins were purchased from Sigma-Aldrich. Ampicillin, isopropyl-*β*-D-thiogalactopyranoside (IPTG), tryptone, urea, tris-HCl, and sodium chloride were obtained from Fisher Scientific. Yeast extract, methanol, and curcumin were purchased from Acros Organics and BCA kit was obtained from Pierce. Imidazole was purchased from Alfa Aesar and copper grids for TEM were purchased from Ted Pella. Dialysis tubing (3.5 kDa molecular weight cut-off) was obtained from Thermo Scientific. See Table A 2 for a complete list of reagents and suppliers.

### *1.3.2 Genetic engineering*

#### 1.3.2.1 PCR amplification

DNA fragments for C, Q1, Q2, L1, and L2 have been generated via polymerase chain reaction amplification. Template DNA was COMPcc$^{s}$ (see Appendix 5.4.1), whose gene sequence is TCA ATT GTG AGC GGA TAA CAA TTT CAC ACA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG AGA GGA TCG CAT CAC CAT CAC CAT CAC GGA TCC GGT GAC CTG GCG CCG CAG ATG CTG CGT GAA CTG CAG GAA ACC AAC GCG GCG CTG CAG GAC GTT CGT GAA CTG CTG CGT CAG CAG GTT AAA GAA ATC ACC TTC CTG AAA AAC ACC GTT ATG GAA TCT GAC GCG TCT GGT AAG CTT AAT TAG CTG AGC TTG GAC TCC TGT TGA TAG ATC CAG TAA TGA CCT CAG AAC TCC ATC TGG ATT. Primer sequences are presented in Table 1.6.

Table 1.6. Primer sequences used to generate fragments for C, Q, and L.

| Primer | Primer sequence (5' → 3') |
|---|---|
| COMP FWD | GCATGGGATCCATCGAAGGTCGCGCGCCGCAGATGCTGCGT |
| COMP REV | GCATGAAGCTTTGAGGTGTTTTTCAGGAAGG |
| L44 Rev1 | CATCTGCGGCGCGGTGTTTTTCAGGAAGG |
| L44 Rev2 | GCATGAAGCTTTGACGCCGCGTTGGTTTCCTG |
| L44 Fwd1 | GCATGGGATCCATCGAAGGTCGCCTGCAGGACGTTCGTGAAC |
| L44 Fwd2 | CTGAAAAACACCGCGCCGCAGATGCTGCGT |
| Q54 Fwd1 | GCATGGGATCCATCGAAGGTCGCGTTAAAGAAATCACCTTC |
| Q54 Rev2 | GCATGAAGCTTTGACTGCTGACGCAGCAGTTC |

C was generated with the primers COMP FWD/COMP REV, Q1 with Q54 Fwd1/L44 Rev1, Q2 with Q54 Fwd2/L44 Rev2, L1 with L44 Fwd1/L44 Rev1, and L2 with L44 Fwd2/L44 Rev2. Phire Hot Start II DNA polymerase enzyme (Thermo Scientific, MA, USA) was used in the PCR reactions described herein. Concentrations of the reagents used for PCR amplification of DNA fragments were: 0.7 µL (200 ng) COMPcc$^{s}$ template DNA, 10 µL (1x final concentration) reaction buffer, 1 µL (0.2 mM) dNTPs, 1 µL dimethyl sulfoxide (DMSO), 1 µL (10 µM) respective forward primer, 1 µL (10 µM) respective reverse primer, 34.3 µL 2x filtered deionized water, and 1 µL Phire enzyme for a total reaction volume of 50 µL.

The same protocol was used to amplify all DNA fragments and was: 98 °C for 2 minutes, [98 °C for 5 seconds, 54 °C for 5 seconds, 72 °C for 20 seconds] repeated for 30 cycles, 72 °C for 1 minute, and 4 °C until reaction tube was removed from PCR apparatus. For electrophoresis of amplified DNA, 5 µL of dye was added to the 50 µL reaction tube and 15 µL of the sample was loaded into a 2 % (1 g agarose in 50 mL 1x TAE buffer) agarose gel. The gel was run for 30 minutes at 100 V in 1x TAE buffer. DNA was purified from the gel using a ZYMO DNA purification kit (Zymo Research Corporation, CA, USA) suspended in 2x filtered deionized water. Concentrations were measured using the NanoDrop 1000 (Thermo Scientific, MA, USA) at 260 nm. The DNA for Q541 and Q542 was subsequently frozen in the -80 °C freezer for 15 minutes and

lyophilized for 3 hours. This DNA was then resuspended in 2x filtered deionized water to reach a concentration of 33 ng/μL for future use in PCR assembly experiments.

1.3.2.2 PCR assembly

The DNA fragments for Q were generated through PCR assembly of Q1/Q2, and L was generated through assembly of L1/L2. Template DNA for generate Q was Q1 and Q2, where template to generate L was L1 and L2. For the generation of Q, primers Q Fwd1 and Q Rev2 were used, and primers L Fwd1 and L Rev2 were used to generate L (Table 1.6). Concentrations of the reagents used for the first step of PCR assembly of Q and L DNA were: 3 μL (100 ng/μL) Q1 or L1 template DNA, 3 μL (33 ng/μL) Q2 or L2 template DNA, 10 μL (1x final concentration) reaction buffer, 1 μL (0.2 mM) dNTPs, 1 μL dimethyl sulfoxide (DMSO), 29 μL 2x filtered deionized water, and 1 μL Phire enzyme for a total reaction volume of 50 μL. To anneal DNA fragments prior to PCR amplification the following protocol was used: 98 °C for 2 minutes, [98 °C for 2 seconds, 54 °C for 5 seconds, 72 °C for 20 seconds] repeated for 10 cycles, 72 °C for 1 second, and 4 °C until reaction tube was removed from PCR apparatus. Tubes were removed and 1 μL (10 μM) respective forward primer, 1 μL (10 μM) respective reverse primer were added to Q and L tubes. The same protocol was then used to amplify C, Q, and L DNA: 98 °C for 2 minutes, [98 °C for 5 seconds, 54 °C for 5 seconds, 72 °C for 20 seconds] repeated for 35 cycles, 72 °C for 1 minute, and 4 °C until reaction tube was removed from PCR apparatus. The DNA sequences of the genes generated are listed in Table 1.7.

Table 1.7. DNA sequences for C, Q, and L.

| Gene | DNA sequence (5' → 3') |
|---|---|
| L | GGATCCATCGAAGGTCGCCTGCAGGACGTTCGTGAACTGCTGCGTCAGCAGGT TAAAGAAATCACCTTCCTGAAAAACACCGCGCCGCAGATGCTGCGTGAACTGCA GGAAACCAACAATCAAAGCTT |
| Q | GGATCCATCGAAGGTCGCGTTAAAGAAATCACCTTCCTGAAAAACACCGCGCC GCAGATGCTGCGTGAACTGCAGGAAACCAACGCGGCGCTGCAGGACGTTCGT GAACTGCTGCGTCAGCAGTCAAAGCTT |
| C | GGATCCATCGAAGGTCGCGCGCCGCAGATGCTGCGTGAACTGCAGGAAACCA ACGCGGCGCTGCAGGACGTTCGTGAACTGCTGCGTCAGCAGGTTAAAGAAATC ACCTTCCTGAAAAACACCTCAAAGCTT |

For electrophoresis of amplified DNA, 5 µL of dye was added to the 50 µL reaction tube and 15 µL of the sample was loaded into a 2 % (1 g agarose in 50 mL 1x TAE buffer) agarose gel. The gel was run for 30 minutes at 100 V in 1x TAE buffer. DNA was purified from the gel using a ZYMO DNA purification kit suspended in 2x filtered deionized water. Concentrations were measured using the NanoDrop 1000 (Thermo Scientific, MA, USA) at 260 nm.

#### 1.3.2.3 Restriction I

Plasmid vector and insert (C, Q, and L) DNA were restricted prior to ligation. Restriction of plasmid DNA containing the mutation for PheRS** was performed in 3 samples of 50 µL each reaction tube. The contents of all restriction reaction tubes are given in Table 1.8.

Table 1.8. Contents of restriction reactions of PheRS** xL1 blue vector DNA.

| **Reagent** | **C (µL)** | **Q (µL)** | **L (µL)** | **PheRS** (µL)** |
|---|---|---|---|---|
| DNA | 11 | 11 | 11 | 2 |
| Buffer B | 5 | 5 | 5 | 5 |
| 2x $dH_2O$ | 30 | 30 | 30 | 41 |
| BamHI | 2 | 2 | 2 | 1 |
| HindIII | 2 | 2 | 2 | 1 |
| **TOTAL (µL)** | 50 | 50 | 50 | 50 |

The reaction tubes were incubated at 37 °C overnight to allow the reaction to go to completion. Restricted DNA was purified after electrophoresis and the concentration was measured using the NanoDrop 1000 (Thermo Scientific, MA, USA) at 260 nm.

#### 1.3.2.4 Ligation

Purified and restricted C, Q, and L DNA were ligated to purified and restricted PheRS** plasmid DNA. Positive (with insert) and negative (without insert) controls were prepared according to Table 1.9. In order to check whether ligation was successful, a φ ladder (New England BioLabs Inc., MA, USA) was used to visualize ligation results. The conditions of this ligation (+ and – controls) are also outlined in Table 1.9.

Table 1.9. Ligation reagent concentrations for ligating designed DNA with PheRS** plasmid DNA and for φ ladder control reactions.

| Reagent | -DNA control (µL) | +DNA control (µL) | - $\phi$ ladder (µL) | + $\phi$ ladder (µL) |
|---|---|---|---|---|
| Plasmid DNA | 2 | 2 | 2 | 2 |
| insert DNA | - | 5 | | |
| 2x $dH_2O$ | 15 | 10 | 16 | 15 |
| T4 buffer | 2 | 2 | 2 | 2 |
| T4 ligase | 1 | 1 | - | 1 |
| **TOTAL (µL)** | 20 | 20 | 20 | 20 |

Reaction tube was incubated for two days and nights at 16 °C. – and + φ ladder restriction reaction samples were run on a 2 % agarose gel to visualize ligation. Ligated plasmid was then used for transformation.

1.3.2.5 Transformation

Transformation was performed for insert DNA/PheRS** ligated plasmid vectors. Two reactions were prepared for each DNA type: (1) 5 µL DNA/PheRS** DNA was added to 100 µL Zymo XL1 blue cells (Zymo Research Corporation, CA, USA), and (2) 15 µL DNA/PheRS** DNA was added to 100 µL Zymo xL1 blue cells. Cells were kept on ice and thawed on ice for 30 minutes. After 30 minutes, 700 µL of Luria Bertani broth (at 37 °C) was added to each reaction, and was shaken at 350 rpm at 37 °C for 45 minutes on a shaking incubator (model 1570, VWR, PA, USA). TSA plates containing 0.2 mg/mL ampicillin were used to grow colonies. 5 plates were grown for each DNA type: (-) control sample and 4 (+) control samples: (1) 200 µL of reaction 1, (2) 600 µL of reaction 1, (3) 200 µL of reaction 2, and (4) 600 µL of reaction 2. Plates were left in the incubator for 20 minutes half covered to allow them to dry. Then were closed completely, inverted, and incubated at 37 °C overnight in a Yamato general purpose incubator IC400 (Yamato Scientific America, Inc., CA, USA).

1.3.2.6 Restriction II

Plasmid vector (containing the mutation for PheRS**) and C, Q, and L DNA were restricted, ligated, and transformed into chemically competent Zymo XL1 blue cells (Zymo Research Corporation, CA, USA) (see Appendix 5.6.1). Bacterial colonies grown

on TSA plates containing 0.2 mg/mL ampicillin were obtained and starter cultures were grown. The DNA from these cells was obtained through use of the MiniPrep kit (Zymo Research Corporation, CA, USA) and was purified. Restriction of the PheRS**/insert DNA was performed for 6 starter cultures for each DNA type in 50 µL volumes. The contents of the negative (no restriction enzymes) and positive (both restriction enzymes) control restriction reactions are given in Table 1.10.

Table 1.10. Contents of reaction tubes used during restriction of PheRS**/insert DNA.

| **Reagent** | **Uncut (µL)** | **Double cut (µL)** |
|---|---|---|
| PheRS**/insert DNA | 3 | 3 |
| Buffer B | 5 | 5 |
| 2x $dH_2O$ | 42 | 41 |
| BamHI | - | 0.5 |
| HindIII | - | 0.5 |
| **TOTAL (µL)** | 50 | 50 |

The reaction tubes were incubated at 37 °C overnight to allow the reaction to go to completion. The following day 10 µL of each sample was mixed with 3 µL dye and run in a 1 % agarose gel (0.5 g agarose in 50 mL 1x TAE buffer) alongside with a 1 kb DNA ladder in 1x TAE buffer. Images were taken at 30 and 45 minutes of electrophoresis at 100 V to visualize the inserts and plasmid vector DNA.

#### 1.3.2.7 DNA sequencing

Purified DNA (20 µL of 75 ng/µL) was sent to Eurofins MWG Operon (AL, USA) for sequencing.

### *1.3.3 Protein expression*

Colonies were initially grown on TSA plates containing 0.2 mg/mL ampicillin as a selection antibiotic at 37 °C for 16 hours. Ampicillin was used to select for genetically modified host cells containing the plasmid vector only.[109] Single colonies were selected from the plates and inoculated in 5 mL of LB with 0.2 mg/mL ampicillin and incubated at 37 °C, for 12 hours, at 350 rpm on a shaking incubator (model 1570, VWR, PA, USA).

Proteins were expressed in 200 mL volumes contained in 1 L baffled Erlenmeyer flasks. 1 mL of starter culture was added to 200 mL of LB and 0.2 mg/mL ampicillin. Glycerol stocks were prepared from starter cultures. 1 mL pre-expression samples were taken from each culture when optical density at 600 nm > 0.8, at which point cultures were induced with 0.2 mg/mL isopropyl $\beta$-D-1-thiogalactopyranoside. Induction was allowed to occur for 3 h, at which point post expression samples were taken. Cultures were shaken continuously at 250 rpm at 37 °C on an shaking incubator (model AJ125, Appropriate Technical Resources, MD, USA). Post induction, the cells were harvested by centrifugation on a Beckman Coulter Allegra$^{TM}$ centrifuge (IN, USA), Beckman Coulter TS-5.1-500 swinging bucket rotor (IN, USA), at 4 °C and 8000 rpm for 20 minutes. The supernatant was disposed of, and the cell pellets were stored in -80 °C until purification. A freeze/thaw cycle at -80 °C was carried out to facilitate lysis, followed by probe sonication (Sonicator 500, Qsonica, LLC, CT, USA) for 1 minute, with pulse 5 s on/5 s off at 35 % amplitude.

Pre/post expression samples for each culture were run on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels to evaluate protein expression. Purification under denaturing conditions was carried out using 50 mM tris-GCL, 0.5 M NaCl, 20 mM imidazole, 6 M urea pH 8 buffer. The soluble crude lysate was bound to Ni-NTA beads (Sigma Aldrich, MO, USA) and allowed to equilibrate for at least 3 h at 4 °C under continuous rotation in a Thermolyne Labquake shaker rotisserie (model 415110, Thermo Scientific, MA, USA). The proteins were eluted with increasing gradient of imidazole (20 mM – 1 M). Pure fractions were refolded via stepwise dialysis in pH adjusted phosphate buffer (10 or 50 mM), halving the urea concentration successively. For experiments involving mixtures of C and Q proteins (discussed in Section 1.4.4.1), bicinchoninic acid analysis (BCA) measurements were performed after purification and mixtures of varying molar ratios of C and Q were prepared in denatured conditions. The mixtures were then co-dialyzed according to the same conditions described above. The enhanced protocol of BCA (PIERCE, IL, USA) (see Appendix 5.6.6) was used to estimate protein concentration with bovine serum albumin (BSA) as a standard.

### *1.3.4 Conditions for curcumin binding*

A stock solution of 1 mM curcumin was prepared in non-denatured methanol and sonicated for 15 minutes to remove any aggregates. Appropriate volumes of the ccm/MeOH stock solution were added to the protein solution to obtain desired molar ratios of curcumin to protein, varying from 0 to 8.

### *1.3.5 Circular dichroism*

Circular dichroism (CD) measurements were conducted on a Jasco J-815 CD spectrometer (Jasco, Inc. MD, USA). Wavelength and temperature scans were conducted with 10 μM ($6.3x10^{-2}$ mg/mL) protein concentrations.

#### 1.3.5.1 Wavelength scans

The wavelength spectrum was measured over a range from 190 to 250 nm with a step size of 1 nm. Mean residue ellipticity (MRE) was calculated from raw data according to the procedure described in Gunasekar *et al.* 2009.[92] Secondary structure analysis of $\alpha$-helical, $\beta$-sheet, and random coil content was calculated with the K2D method using DichroWeb software.[110,111]

#### 1.3.5.2 Temperature scans

Temperature scans of each protein were performed over a range of 20 – 85 °C with a temperature step of 1 °C/min at 222 nm. Scans were also performed at of 2 °C/min[43] and 5 °C/min[112] to evaluate dependence of thermal melt signatures on scan speed. All measurements were made in duplicates of independently prepared proteins and data represents the average. Thermodynamic properties of wt, C, and Q were determined through analysis of thermal melts. A two-state model was used, where assumptions included monophasic behavior and reversible melting behavior.[112,113] These assumptions were confirmed experimentally by melting (from 20 to 85 °C) and cooling (from 85 to 20

°C) the proteins. Calculation of thermodynamic parameters including $T_m$, $\Delta G_m$, $\Delta H_m$, and $\Delta S_m$ were performed according to the method described by Greenfield.[112]

Fraction folded ($\alpha$) was determined through:

$$\alpha = \frac{\theta_t - \theta_u}{\theta_f - \theta_u}$$

Equation 1.1

where $\theta_t$ is the ellipticity at any given temperature, $\theta_u$ is the ellipticity of the unfolded form (at 85 °C), and $\theta_f$ is the ellipticity of the fully folded form (at 20 °C). The melting temperature ($T_m$) was defined as the temperature at which $\alpha$ = 0.5, Equation 1.2. $K$ was determined through Equation 1.3.

$$T_m = T|_{\alpha=0.5}$$

Equation 1.2

$$K = \frac{\alpha}{1 - \alpha}$$

Equation 1.3

The Gibbs free energy of folding ($\Delta G_m$) is given by:

$$\Delta G_m = -RTlnK$$

Equation 1.4

where $\Delta H_m$ and $\Delta S_m$ represent the standard enthalpy and entropy changes at the melting point, respectively. Reworking Equation 1.4 we obtain:

$$lnK = \frac{-\Delta G_m}{RT} = \frac{-\Delta G_m}{R}\left(\frac{1}{T}\right)$$

Equation 1.5

which allows us to obtain a linear relationship between ln$K$ and the inverse temperature. Melting curve relationships were linearized by plotting ln$K$ versus $T^{-1}$ ($K^{-1}$) and truncating the data to the linear region. Slope of these linear relationships allows for

calculation of $\Delta H_m$, where

$$\Delta H_m = -slope\ x\ R$$

Equation 1.6

At the transition midpoint, $\Delta G_m$ is equal to zero, which can be used to calculate $\Delta S_m$ as per Equation 1.7, where $T_m$ is given by Equation 1.2.

$$\Delta S_m = \frac{\Delta H_m}{T_m}$$

Equation 1.7

The Gibbs free energy of folding is then calculated at 25 °C, or 298 K by Equation 1.8.

$$\Delta G_m = \Delta H_m - T\Delta S_m$$

Equation 1.8

Two minima used to calculate helicity are located at 222 nm ($\theta_{222}$) and the wavelength between 195 and 210 nm with the minimum MRE value ($\theta_{min}$). The ratio of these two minima, $\theta_{222}/\theta_{min}$ results in the helicity.[114]

#### 1.3.5.3 Absorbance

Absorbance at 420 nm was obtained via CD.

### *1.3.6 Nuclear magnetic resonance*

1D $^1$H nuclear magnetic resonance (NMR) was performed on a Bruker Ultrashield 500 Plus instrument (MA, USA) and data was collected and analyzed using TopSpin 3.2 software (Bruker, MA, USA). Protein concentrations were kept constant at 20 μM, with buffer conditions of 50 mM PB pH 4 with 1 % (v/v) methanol and 1 % (v/v) $D_2O$. NMR was performed in the absence and presence of curcumin, in a 5:1 molar ratio of curcumin:protein.

### *1.3.7 ATR-FTIR*

Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy **(**ATR-FTIR) experiments were performed using Perkin Elmer System 2000 FT-IR (Perkin Elmer, MA,

USA) with DuraSamplIR II T diamond ATR accessory and equipped with a MCT-A detector (Smiths Detection Inc., CT, USA). Approximately 5 μL of peptide solution (10 μM in 50 mM PB, pH 4, 8, and 10) was added on the diamond ATR surface. The spectrum (128 scans) was measured at room temperature over a range of 4000–400 $cm^{-1}$ with 0.5 $cm^{-1}$ resolution. PeakFit software was used to process the data, which involved a 2nd derivative zero baseline correction of the amide I region between 1700–1600 $cm^{-1}$ and deconvolution of peaks with a Gaussian function.[115] All readings represent the average of two trials.

### *1.3.8 TEM*

Phillips CM-100 (CCNY) (Philips, Netherlands), JEOL JEM-1400 (BNL) (Jeol, MA, USA), and JEOL JEM-2100 HRTEM (Hunter College) (Jeol, MA, USA) transmission electron microscopes were used to view the supramolecular protein structure. Approximately 3 μL of 10 μM protein in 50 mM PB was spotted on copper grids. After 1 minute, the grids were blotted using filter paper and rinsed with 3-4 drops Milli Q water to remove excess salts from the buffer. After blotting with filter paper, the sample was negatively stained by adding 3 μL of 1 % filtered uranyl acetate, blotted using filter paper, and dried at room temperature for 10-15 minutes. ImageJ software (National Institutes of Health, MD, USA) was used to measure the fibers dimensions.[116]

### *1.3.9 Zeta potential*

Zeta potential measurements were performed on a Zetasizer Nano Series model Nano ZS90 (Malvern Instruments, Inc., MA, USA). 50 mM phosphate buffer at pH 4 was used to saturate the clear polycarbonate disposable zeta cell DTS1060C (Malvern Instruments, Inc., MA, USA) prior to injecting 750 μL protein samples at protein concentrations of 10 μM. The following settings were used for zeta potential measurements in the DTS (Nano) software: 90 °C instrument settings, Smoluvchoski model, material protein, dispersant PBS, temperature of 25 °C, viscosity 1.0200 cP, and dielectric constant of 1.34. Measurements were taken in triplicates, conducting 10 runs for each measurement, with a delay time of 2 seconds between each measurement.

### *1.3.10 Dynamic light scattering*

Dynamic light scattering (DLS) measurements were performed on a Zetasizer Nano Series model Nano ZS90 (Malvern Instruments, Inc., MA, USA). In measuring DLS, 50 mM phosphate buffer at pH 4 was used to wash the low volume disposable cuvette DTS0112 cell (Malvern Instruments, Inc., MA, USA). Approximately, 750 µL of 10 µM protein sample with varying amounts of curcumin was then added to the cell. To measure size, the following settings were applied: material protein (refractive index 1.450), absorption 0.001, dispersant PBS with a viscosity of 1.0200 cP and refractive index of 1.335. Measurements were taken in triplicates, conducting 10 runs for each measurement, with a delay time of 2 seconds between each measurement.

### *1.3.11 Scatchard binding analysis*

Equation 1.9 represents the interaction between protein and curcumin, binding to form a complex.

$$n\,Curcumin + protein \leftrightarrow Complex$$

Equation 1.9

The equilibrium constant for the above equation is given by:

$$K = \frac{[Complex]}{[protein][ccm]^n}$$

Equation 1.10

Assuming a 1:1 complex formation, the Benesi-Hildebrand equation can be applied to determine extinction coefficients at 420 nm:

$$\frac{1}{\Delta A} = \frac{1}{K\Delta\varepsilon_{420}[protein]}\left(\frac{1}{[ccm]}\right) + \frac{1}{\Delta\varepsilon_{420}[protein]}$$

Equation 1.11

where $\Delta A$ is the change in absorbance at 420 nm at different curcumin concentrations from 0 to 160 µM. $K$ is calculated from

Equation 1.11, where the slope of the linear fits represent $1/(K\Delta\varepsilon[\text{protein}])$. $\Delta\varepsilon_{420}$, the extinction coefficients for the protein at 420 nm and the protein in presence of curcumin at 420 nm, are calculated from linear plots of Equation 1.11. These values are then applied to the Scatchard equation (Equation 1.12), to calculate number of binding sites available for curcumin in the protein, *n*, and the number of moles of curcumin bound to the total number of moles of protein, *r*.

$$\frac{1}{r} = \frac{1}{nK}\left(\frac{1}{[ccm]}\right) + \frac{1}{n}$$

Equation 1.12

A linear relationship is produced by the Scatchard analysis, where *r* was calculated as per Equation 1.13:

$$r = \frac{Abs_{420}}{[protein]\varepsilon_{420}l}$$

Equation 1.13

The binding constant *K* is estimated using Equation 1.12 and the linear relationships generated from Equation 1.11.

### *1.3.12 SEM*

SEM images were taken on a Zeiss Gemini FE-SEM instrument (Carl Zeiss Microscopy, LLC, NY, USA) with an EHT of 2 kV. Samples were prepared by spotting approximately 3 μL of 10 μM protein in 50 mM PB on copper TEM grids. After 1 minute, the grids were blotted using filter paper and rinsed with 3-4 drops Milli Q water to remove excess salts from the buffer and dried at room temperature for 10-15 minutes.

### *1.3.13 AFM*

AFM images were taken on an Asylum MFD-3D-BIO Environmental Closed Loop ATM instrument (Asylum Research, CA, USA) operating in the tapping mode with standard silicon tips of 40 N/m force. Typically, the surface was scanned at 1 Hz with 256 lines/image resolution and 1.0–2.0 V setpoint. Samples were prepared by depositing protein in solution on freshly cleaved mica surfaces and dried under vacuum.

### *1.3.14 Fluorescence microscopy*

Upon binding protein to curcumin, images of protein in solution were obtained using a 40x-2000x Infinity Plan Epifluorescence microscope (AmScope, CA, USA). Approximately 7 μL of protein and curcumin solution was spotted on a glass microscopy slide (Clay Adams, CA, USA) and sealed with a coverslip (Fisherbrand, PA, USA). Images were taken with an Infinity HD camera. Unfortunately, the dimensions were not obtainable as there was no scale-bar inserted on each image taken due to software restrictions. Confocal microscopy was selected to overcome the 2D limitations of the fluorescence microscope.

### *1.3.15 Confocal microscopy*

Samples were imaged using a Leica TCS SP2 AOBS confocal microscope system (Leica, IL, USA) equipped with argon ion and HeNe lasers. A 63×/1.4 NA oil-immersion objective was used for all of the images. Lab-Tek II chambered #1.5 German coverglass system (Thermo Scientific, MA, USA) was used as the imaging slide. Curcumin was excited using the 458 nm line of the argon laser, and images were taken with the detection window set between 465 and 560 nm. The pinhole aperture was set at an Airy value of 1.0, which was equivalent to sampling an ~500 nm vertical z slice of the fiber, as estimated by the axial resolution, $r_{z,confocal} \approx 1.4\ \lambda_{em}\ n/NA^2$ (NA, numerical aperture; n, refractive index; $\lambda_{em}$, emission wavelength (525 nm)). Interference contrast images were obtained using the Leica tube optics HC 1X/B apparatus with a focusing Bertrand lens. The 3D reconstructions were constructed using ImageJ 64 1.43 in concert with Amira 5.43, employing the Volren 3D rendering routine.

## 1.4 Results and discussion

### *1.4.1 Genetic engineering*

Genetic engineering was employed to obtain gene constructs for the proteins of interest, namely C, Q, and L. C was obtained through introduction of a stop codon into the original wt gene sequence to truncate the C-terminal heptad repeat. Primers for C were generated to also introduce and IEGR cleavage site C-terminal to the histidine tag. Q and

L are constructed through a method called PCR assembly, in which overlapping DNA fragments are “assembled” to generate a full gene comprised of two gene fragments. The fragments used to construct Q are referred to as Q1 and Q2, while the fragments used in the construction of L are referred to as L1 and L2.

#### 1.4.1.1 PCR amplification

Amplification of DNA fragments C, L1, L2, Q1, and Q2 was carried out according to conditions outlined in Section 1.3.2.1 and led to successful gene products (Figure 1.24). The expected length of these fragments is given in Table 1.11.

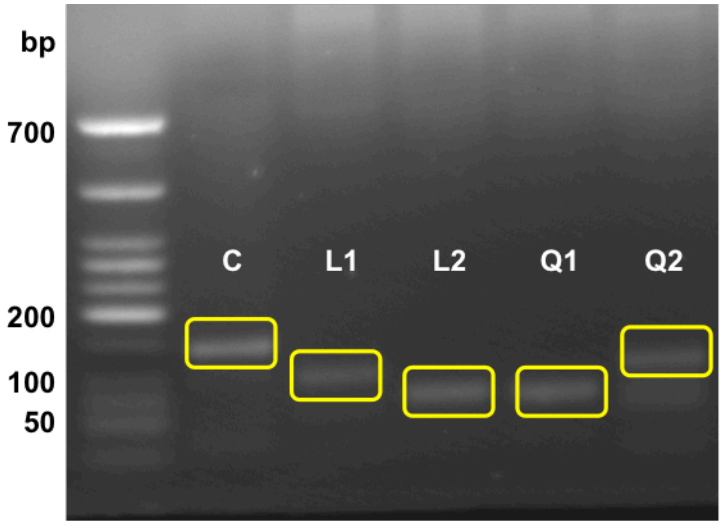



Figure 1.24. DNA gel of amplified fragments for C, L1, L2, Q1, and Q2 on a 2 % agarose gel. Each fragment is located at its expected base pair length.

Table 1.11. Expected base pair lengths of DNA fragments.

| Fragment | Length (bp) |
|---|---|
| L1 | 98 |
| L2 | 68 |
| L | 142 |
| Q1 | 65 |
| Q2 | 101 |
| Q | 142 |

As can be seen in Figure 1.24, each DNA fragment is located at its expected base pair length. DNA fragments of the expected length were identified, excised, and purified.

1.4.1.2 PCR assembly

The chemical synthesis of DNA provides a powerful tool to modify genes and obtain high expression levels of particular translation products of these genes. One particularly effective method of gene modification is through PCR assembly. PCR assembly is general a three step process where the steps consist of (1) the synthesis and amplification of individual fragments of interest with approximately 20 bp overlap in the design of a complete gene comprised of those fragments, (2) synthesis of the entire DNA sequence of interest via mixing of desired fragments in a PCR reaction with a high fidelity polymerase, and (3) final amplification of the desired synthesized DNA.[117]

The results for PCR assembly of L and Q are shown in Figure 1.25. The expected length of these fragments is 142 bp, which is in correspondence with the electrophoresis results, indicating that PCR assembly was successful.

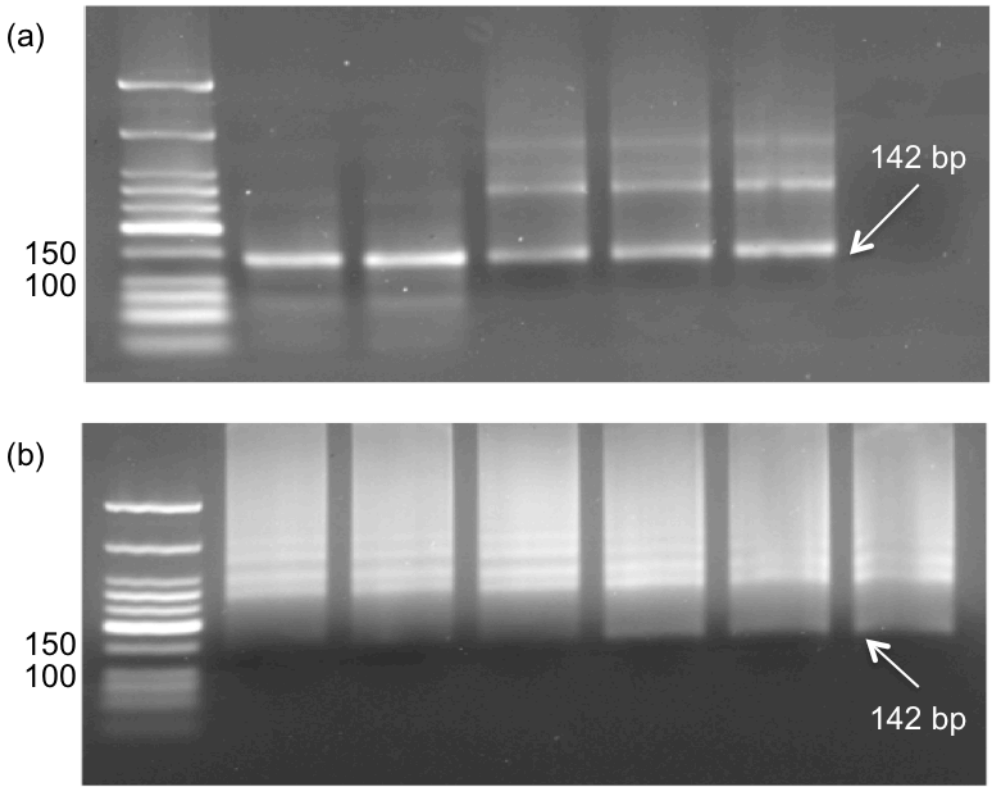



Figure 1.25. PCR assembled L (a) and Q (b) DNA on a 2 % agarose gel. The expected length of these fragments is 142 bp.

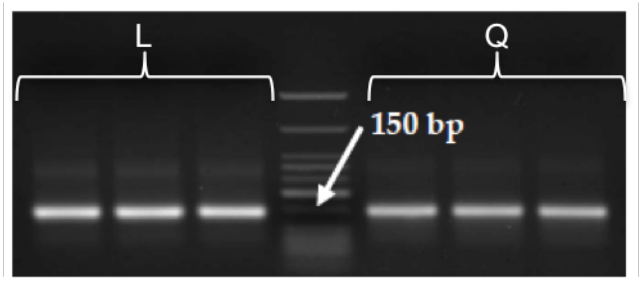



Figure 1.26. PCR amplified L and Q DNA on a 2 % agarose gel. The expected length of these fragments is 142 bp.

Upon completion of PCR assembly of these fragments, it was desired to generate more of this new template DNA via PCR amplification and the two end primers necessary for each type. The result of the PCR amplification of the assembled fragments is shown in Figure 1.26. For both L and Q the band located at approximately 142 bp was extracted and purified.

### 1.4.1.3 Restriction

Restriction was performed in order to determine that the desired PheRS** plasmid vector contained the correct insert DNA. In these experiments restriction enzymes BamHI and HindIII recognize a specific sequence of nucleotides and are able to cut the double stranded DNA. The products of the restriction reactions are run on agarose gels to determine the presence of target gene DNA (142 bp) and the PheRS** plasmid DNA (4512 bp). The first experiments were done on plasmic vectors containing C DNA, which was restricted with BamHI and HindIII. PheRS** plasmid DNA and C DNA were restricted and run on the same gel, as shown in Figure 1.27.

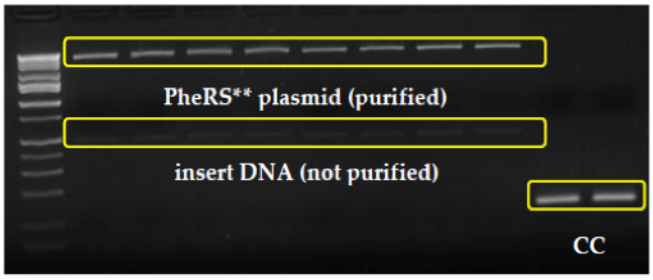



Figure 1.27. Restriction DNA gel for PheRS** plasmid and restricted C fragment on a 2 % agarose gel.

PheRS**/C DNA extracted from 20 bacterial colonies was restricted in order to determine whether the correct vector and insert DNA were obtained. The results of this restriction for 20 starter cultures are shown in Figure 1.28. The PheRS** vector DNA (expected at 4512 bp) is found in the correct location in samples 1, 3, 4, 6-9, 11-14, 16-18, and 20. Sample 11, however, is seen to have an insert of 500 bp, corresponding not to C DNA but a previous insert and is therefore removed from future experimentation. No bands were seen for the C insert (expected at 142 bp), as visualizing DNA of this length is difficult in a 1 % gel.

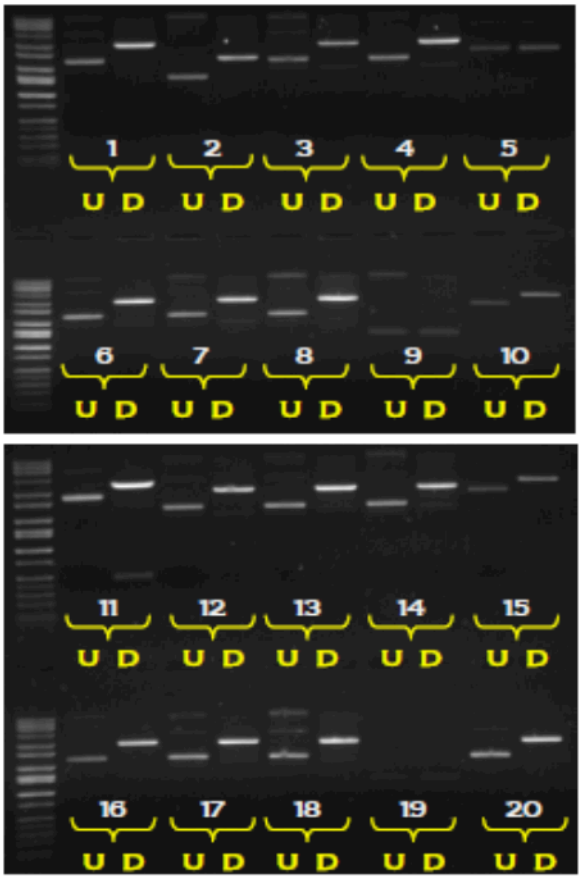



Figure 1.28. Restriction DNA gel for PheRS**/C DNA of uncut (U) and double cut (D) samples on a 1 % agarose gel.

Figure 1.29 shows the DNA gel of restricted C, L, and Q DNA, which genes are expected at a length of 142 bp. Each DNA fragment appeared at the proper location on the gel.

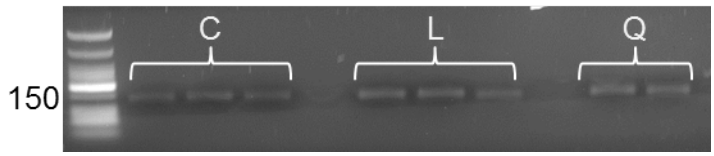



Figure 1.29. Restriction of C, L, and Q DNA. The expected length of these fragments is 142 bp on a 2 % agarose gel.

PheRS**/insert DNA extracted from six bacterial colonies per gene type were restricted in order to determine whether the correct vector and insert DNA were obtained. The results of this restriction for eighteen total starter cultures are shown in Figure 1.30, where the left lane in each sample pair represents the uncut DNA, with the right lane being the restricted DNA. The PheRS** vector DNA (4512 bp) is found in the correct location in all samples. Insert DNA bands (142 bp) were seen in samples C1-5, L5, L6, and Q2-6. The lower half of the first gel was difficult to visualize in general due to the image quality, and it is presumed that these samples also contained insert DNA.

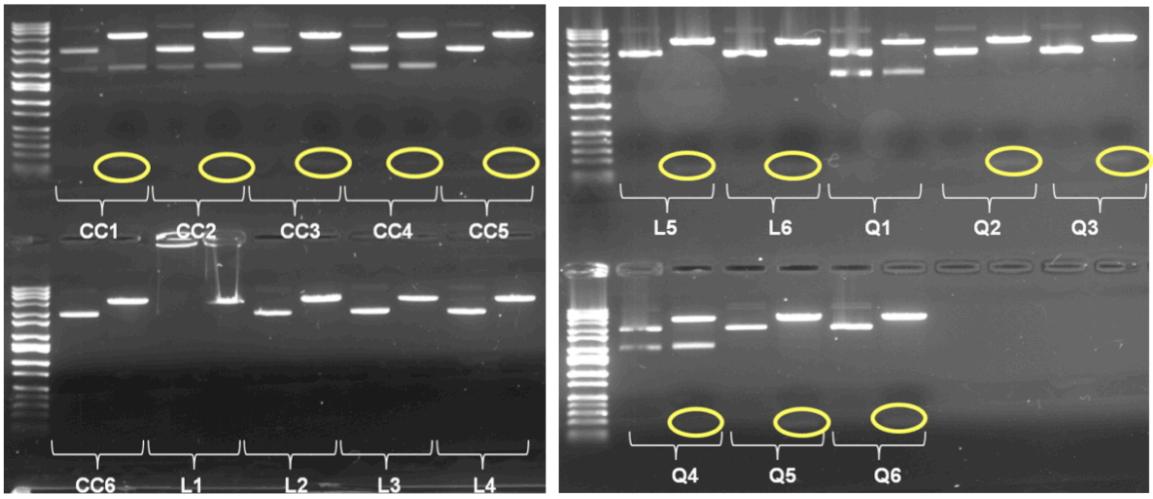



Figure 1.30. Restriction DNA gel for PheRS**/insert DNA for six samples of C, Q, and L. C1-5, L5, L6, and Q2-6 were ligated and transformed.

#### 1.4.1.4 Ligation

To assess whether ligation conditions were acceptable and successfully met a control φ ladder ligation was performed (Figure 1.31). The bright band in the left hand lane (corresponding to the + control) indicated that the individual fragments (seen in the – control lane) were successfully ligated together by the action of $T_4$ ligase. This

experiment was run as an indicator of successful ligation of the PheRS** plasmid vector and our C, L, and Q genes as the experiments were run under the same conditions.

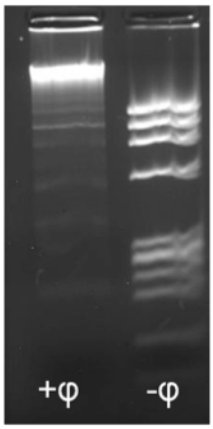



Figure 1.31. + and –φ ladder ligation samples. Solid band in the + control samples proves successful ligation was achieved.

### 1.4.1.5 Test expression and DNA sequencing

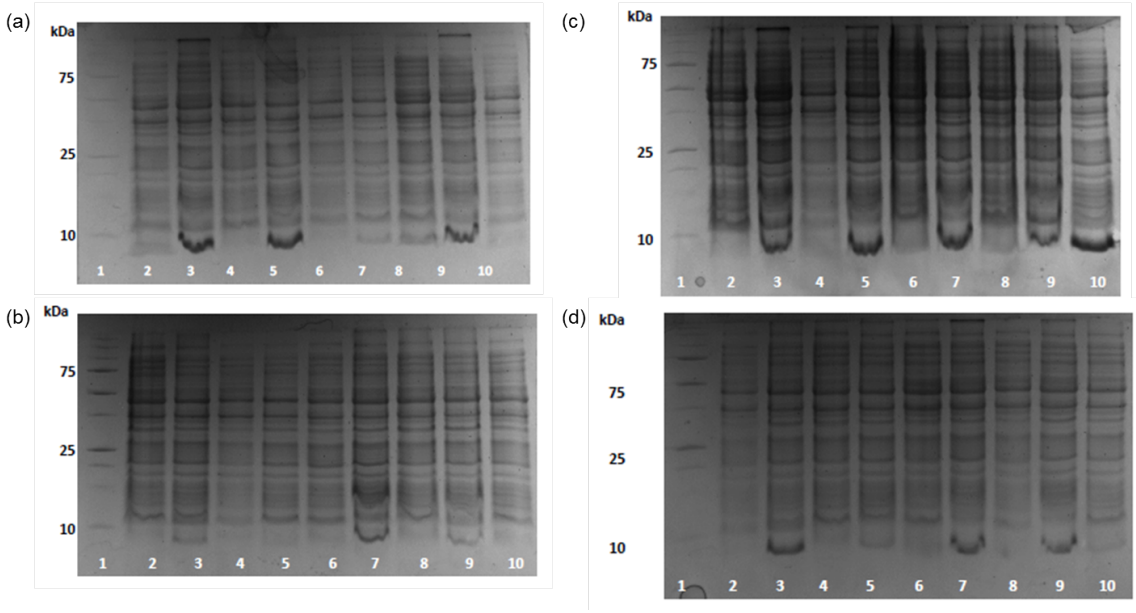



Figure 1.32. 12 % SDS-PAGE protein gels for pre/post expression of six samples of C, L, and Q. Lanes in (a): 1. Ladder, 2. C1 pre, 3. C1 post, 4. C2 pre, 5. C2 post, 6. C3 pre, 7. C3 post, 8. C4 pre, 9. C4 post, 10. C6 pre. (b): 1. Ladder, 2. L1 pre, 3. L1 post, 4. L2 pre, 5. L2 post, 6. L3 pre, 7. L3 post, 8. L4 pre, 9. L4 post, 10. L5 pre. (c): 1. Ladder, 2. Q1 pre, 3. Q1 post, 4. Q2 pre, 5. Q2 post, 6. Q3 pre, 7. Q3 post, 8. Q4 pre, 9. Q4 post, 10. C6 post. (d): 1. Ladder, 2. C5 pre, 3. C5 post, 4. L6 pre, 5. L6 post, 6. Q6 pre, 7. Q6 post, 8. Q5 pre, 9. Q5 post, 10. L5 post.

A test expression was run after transforming ligated PheRS** plasmids and our genes of interest into XL1blue cells according to the procedure described in Section 1.3.2.5. Six colonies were grown for each gene type and small scale expression was run to evaluate protein expression. The results are summarized for all eighteen test expressions in Figure 1.32. Given the results of the small scale expressions of C, L, and Q, three samples from each protein type were selected for DNA sequencing analysis. Miniprep was performed on these samples to obtain pure DNA and the genetic information was sent to Eurofins MWG Operon for sequencing. The samples that were selected displayed prominent bands for the protein of interest in the post expression samples, and were: C3-5, L2, L5, and L6, and Q4-6. The results for DNA sequencing from Eurofins MWG Operon were compared with the designed insert DNA sequences for each protein type. Sequence identity was calculated with LALIGN software (William Pearson's program, which implements the algorithm of Huang and Miller[118]) and are presented below in Table 1.12.

Table 1.12. Sequence identity between designed DNA sequences and results from sequencing experimentally generated DNA for C, L, and Q.

| **Protein** | **C** | | | **L** | | | **Q** | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample | 3 | 4 | 5 | 1 | 3 | 4 | 4 | 5 | 6 |
| Sequence identity (%) | 99.2 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 99.2 |

Samples C4, L3, and Q4 were selected for large scale protein expression (henceforth referred to simply as C, L ,and Q), as these samples showed 100 % sequence identity, strong post expression bands on protein gels, and highest OD at 600 nm 8 h after inoculation.

### *1.4.2 Protein expression and purification*

Amino acid sequences for wt, C, Q, and L that were encoded by the genetically engineered DNA described in Section 1.4.1 are listed in Table 1.5. Proteins were expressed according to the conditions described in Section 1.3.3. SDS-PAGE gel of pre and post expression samples of C, Q, and L (molecular weight 6.3 kDa) and wt are shown in Figure 1.33.

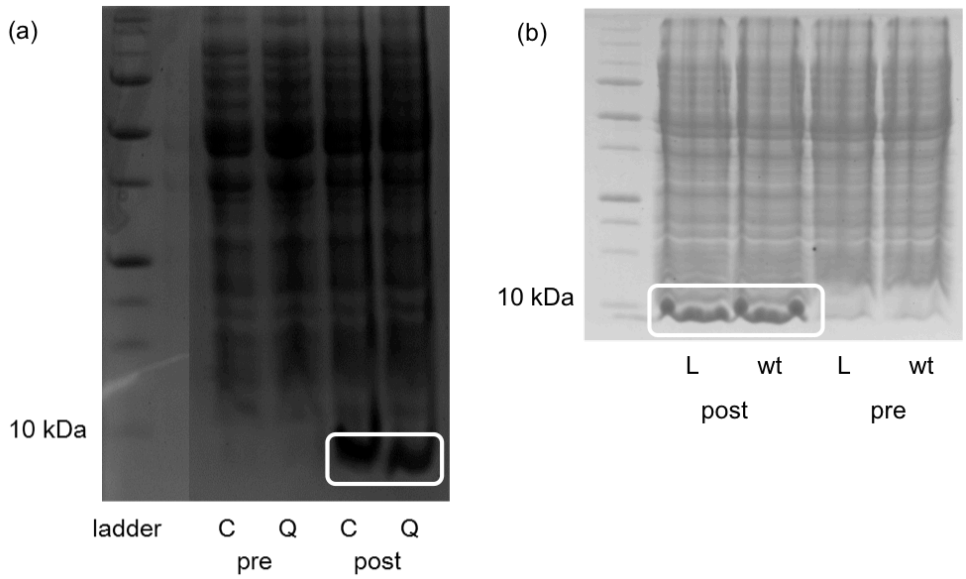



Figure 1.33. 12 % SDS-PAGE protein gels for pre/post expression of C and Q (a), and L and wt (b). Strong bands in the post expression lanes can be observed for all proteins, where the MW of C, Q, and L is 6.3 kDa and the MW of wt is 6.9 kDa.

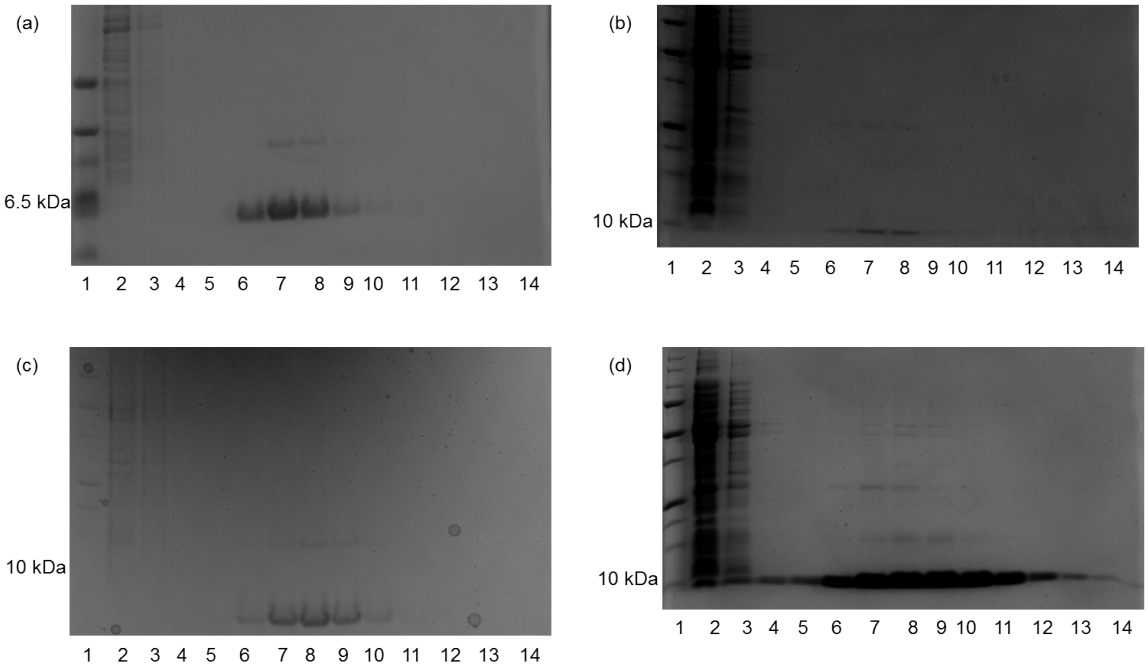



Figure 1.34. 12 % SDS-PAGE protein gels for purification of C (a), L (b), Q (c), and wt (d) by increasing imidazole gradient. Pure protein was typically eluted with imidazole concentrations of 100-500 mM (lanes 6-9). Lanes correspond to the following in gels (a) – (d): 1: ladder, 2: supernatant, 3: flowthrough, 4: wash 1, 5: wash 2, 6: elution 1, 7: elution 2, 8: elution 3, 9: elution 4, 10: elution 5, 11: elution 6, 12: elution 7, 13: elution 8, 14: elution 9.

Upon confirmation of protein expression, cell lysate containing the protein of interest was purified via affinity chromatography, according to the method described in Section 1.3.3 and in further detail in Appendix 5.6.4. Protein was typically eluted with imidazole concentrations in the range of 100 – 500 mM, as can be seen in Figure 1.34. Pure fractions containing sufficiently high concentrations of protein were combined and dialyzed against several liters of 50 mM PB and concentration was subsequently measured via the enhanced protocol of BCA (Appendices 5.6.5 and 5.6.6).

### *1.4.3 Modeling*

Helices and homopentameric assemblies generated from wt, C, and Q subunits were visualized using Chimera (Figure 1.35 - Figure 1.37). Individual helical subunits are presented in Figure 1.35, where the histidine tag has been omitted from the models (for wt the residue sequence MDL precedes the first heptad repeat, whereas C and Q models begin with their first heptad repeats). In Figure 1.35a the loop region of wt can be seen at the C terminal end, which has been removed from C and Q.

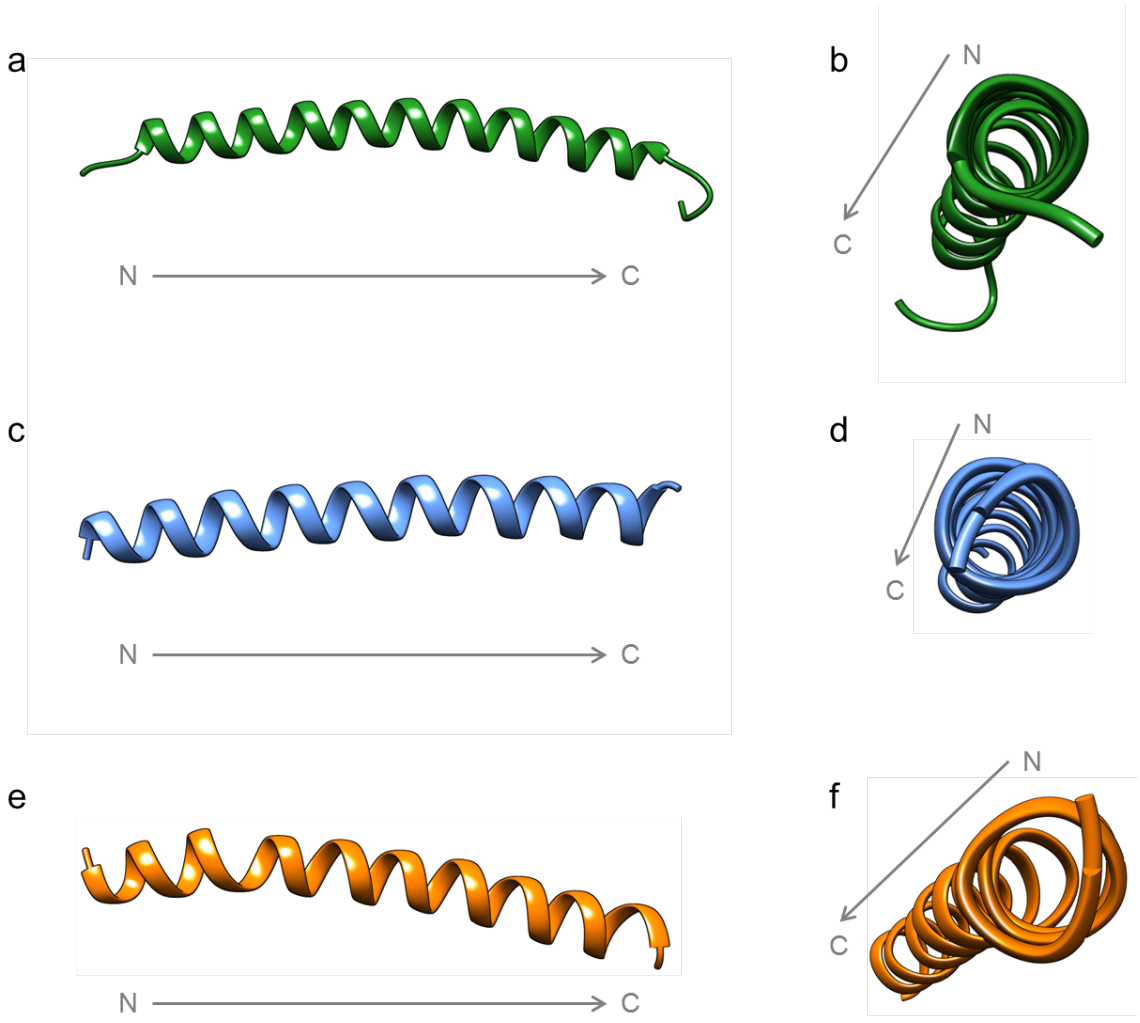



Figure 1.35. Helical subunits of wt (a, b), C (c, d), and Q (e, f) viewed along (a, c, e) and down (b, d, f) the helical axis.

In Section 1.2.5.4 we discussed proline, and its propensity to disrupt helical structure. As can be seen in Figure 1.35e, there is in fact a slight deformation in the helix in the region surrounding the proline residue, however the $\alpha$-helical structure resumes unbroken either side of the proline.[98] Another interesting point to note is the twist observed when looking at the helices along the helical axis, as in Figure 1.35b, d, and f. In wt, the individual helices are not perfectly aligned, contributing to the twist that is observed in the homopentamers when those helices assemble (Figure 1.36). This same twist is present in Q, but much less so in C, which is most likely a result of the repositioning of the proline residue within the sequence.

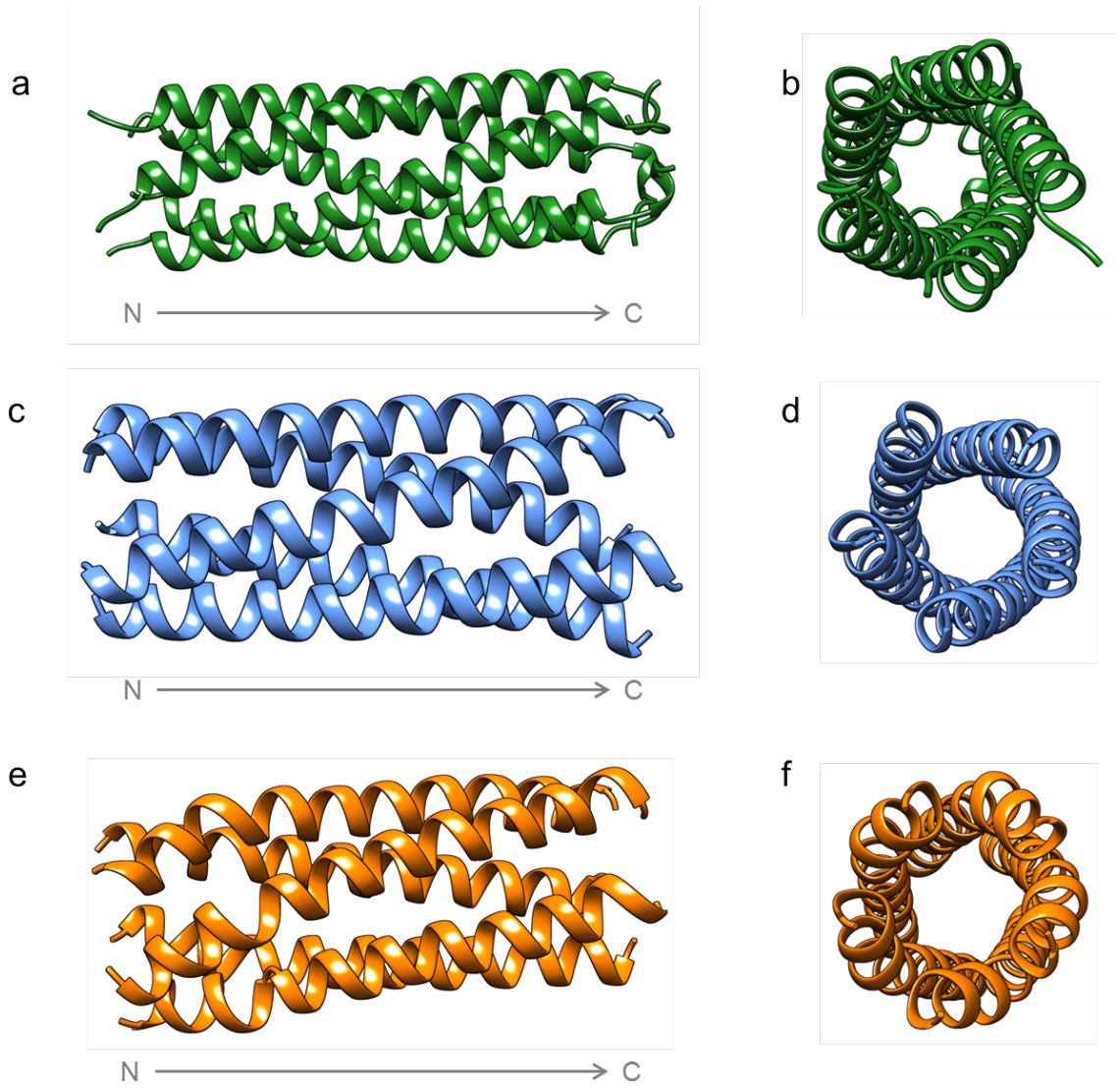



Figure 1.36. Pentameric assemblies of wt (a, b), C (c, d), and Q (e, f) viewed along (a, c, e) and down (b, d, f) the pentamer axis.

Homopentamers of wt, C, and Q are shown in Figure 1.36. The crystal structure for wt (PDB file 1VDF) indicates that the length of the pentamer is 73 Å and an average diameter of 30 Å.[119] The pore in the center of the pentamers in Figure 1.36b, d, and f is an ion trap formed by a ring of conserved glutamines, which has been shown to bind chlorine,[119] all-*trans*-retinol, vitamin $D_3$,[95] as well as several fatty acids.[93] Modeling of C and Q demonstrated that this pentameric bouquet formation, as well as the hydrophobic channel created by the protein subunits, is conserved.

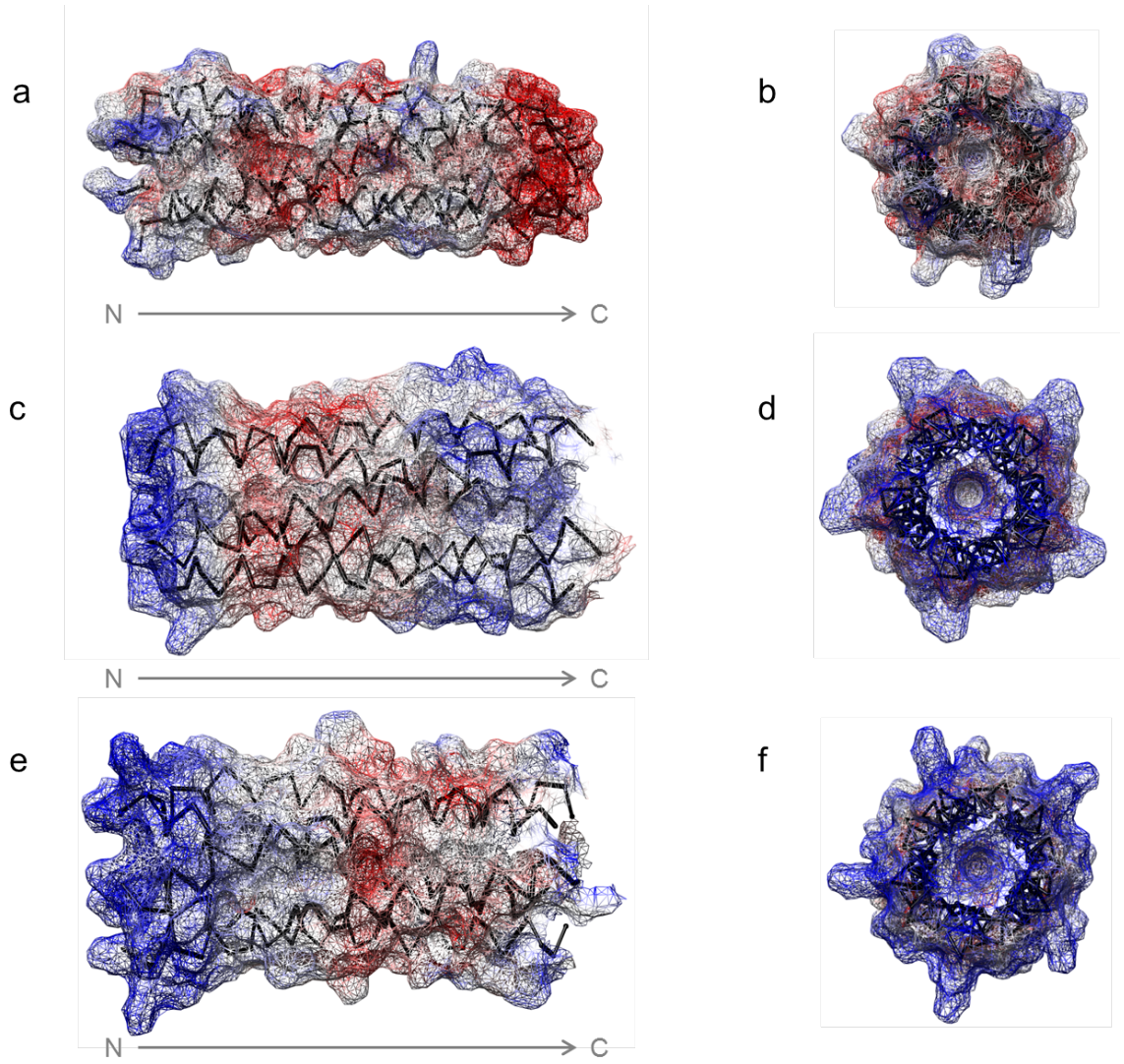



Figure 1.37. Surface charge representation of wt (a and b), C (c and d), and Q (e and f) pentamers under acidic conditions viewed along (a, c, e) and down (b, d, f) the pentamer axis. Positive blue patches are attributed to solvent-exposed lysine and arginine residues. Negative red patches are attributed to solvent-exposed glutamate and aspartate residues.

Surface charge representation was generated by fully protonating the acidic residues histidine, glutamic acid, aspartic acid, and lysine in Chimera to best represent the charge distribution at acidic pH conditions (Figure 1.37). Surface charge distribution of wt did not show charged patches, but rather exhibited an overwhelmingly negative surface charge along the entire pentamer (Figure 1.37). Blue regions on the termini of the Q coiled-coil were positively charged as a result of solvent-exposed arginine and lysine residues, while the center of the pentamer was negatively charged due to solvent-exposed

glutamine and asparagine residues. The overwhelmingly negative and positive "patches" that were produced along the length of the Q pentamer contribute to its ability to self-assemble.

### *1.4.4 Secondary structure and thermodynamic stability*

#### 1.4.4.1 Secondary structure

To assess the secondary structure of the proteins under acidic as well as neutral and basic conditions, CD and ATR-FTIR measurements were performed. CD was used to evaluate secondary structure in solution conditions, while secondary structure in solid state was studied using ATR-FTIR in the presence of 50 mM PB at all three pH conditions.

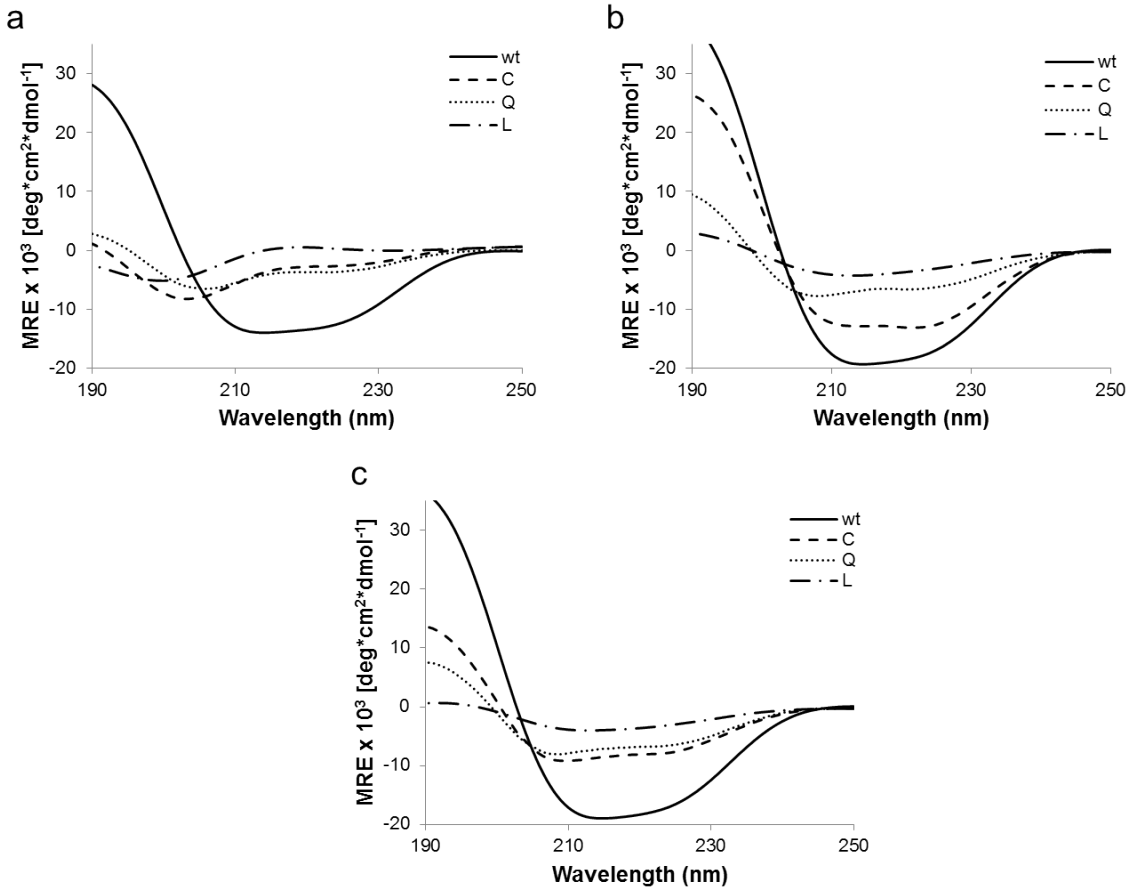



Figure 1.38. Secondary structure of proteins in 50 mM PB at 25 °C at pH 4 (a), 8 (b), and 10 (c). Circular dichroism variable wavelength scans of wt (solid line), C (dashed line), Q (dotted line), and L (dashed, dotted line) at pH 4 (a), 8 (b), and 10 (c). Data is averaged from at least two replicates and was obtained with 10 μM protein concentration.

Wavelength scans of wt, C, Q, and L are at a concentration of 10 μM are shown in Figure 1.38. $\alpha$-helical proteins are characterized by two minima occurring at 208 and 222 nm.[120] wt presents the most stable $\alpha$-helical structure across all pHs evaluated, displaying two deep minima in the wavelength curves. C displays a strong $\alpha$-helix at pH 8, but this structure appears at neutral pH conditions and slightly less so at basic conditions, with the structure appearing more random under acidic conditions. The loss in helical structure at pH 4 is demonstrated by the increase in the signal around 222 nm. This increase is also observed in L, which is completely unstructured at pH 4. Q maintains relatively similar CD signature across all pH values, however this is not the case for ATR-FTIR measurements, as will be discussed later.

Analysis of the secondary structure composition of each of the proteins was conducted using K2D method in the online software platform Dichroweb and is presented in Table 1.13 and Figure 1.39.[110,111] CD measurements showed that under acidic conditions, Q exhibited some helical structure with a double minima and a helical content of 9 % (Figure 1.38, Table 1.13). Maximum $\alpha$-helical content for wt and C was achieved at pH 8, whereas helicity of Q and L increased with pH.

Table 1.13. Mean residual ellipticities and secondary structure content from circular dichroism measurements at pH 4, 8, and 10 for wt, C, Q, and L. The data presented represents the average of at least two replicates.

| | pH 4 | | | | pH 8 | | | | pH 10 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | wt | C | L | Q | wt | C | L | Q | wt | C | L | Q |
| $-\theta_{222}$ x $10^3$ /deg cm$^2$ dmol$^{-1}$ | 13.18 | 2.73 | -0.41 | 3.71 | 18.28 | 13.15 | 3.54 | 6.63 | 17.84 | 8.03 | 3.40 | 6.76 |
| $-\theta_{min}$ x $10^3$ /deg cm$^2$ dmol$^{-1}$ | 13.02 | 8.27 | 5.18 | 6.54 | 17.61 | 12.33 | 4.14 | 7.79 | 17.13 | 9.20 | 3.89 | 8.07 |
| α helices /% | 50 | 8 | 9 | 9 | 64 | 46 | 13 | 22 | 63 | 28 | 16 | 25 |
| β sheet /% | 18 | 45 | 49 | 36 | 6 | 23 | 38 | 21 | 6 | 15 | 32 | 18 |
| random coil /% | 32 | 47 | 43 | 55 | 30 | 31 | 49 | 57 | 31 | 57 | 52 | 57 |

The values of $-\theta_{222}$ and $-\theta_{min}$ are indicative of the signal strength of a particular sample, and can also be impacted by the protein concentration. The removal of the peripheral heptad and reordering from wt led to a dampened signal in C across all pH values. Under acidic conditions Q exhibited a helical signature with a double minimum of -3.71 x $10^3$ deg cm$^2$ dmol$^{-1}$ and -6.54 x $10^3$ deg cm$^2$ dmol$^{-1}$at 222 nm and 206 nm, respectively (Table 1.13). The presence of residue P28 towards the center of Q may have influenced the secondary structure, showing decreased helical content when compared to wt and C at pH

4 and 10 (Figure 1.39). The negative control, L, was completely unstructured, devoid of helical content at pH 4.

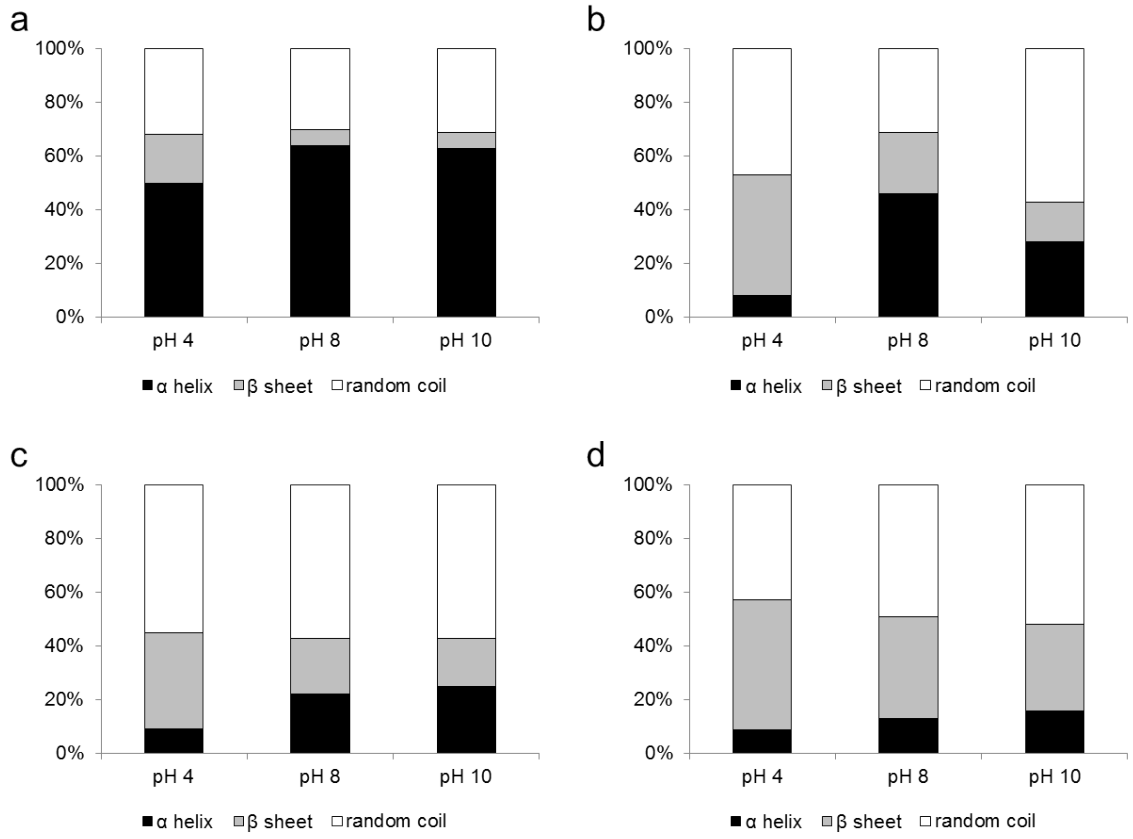



Figure 1.39. $\alpha$-helical (black), $\beta$-sheet (grey), and random coil (white) content for wt (a), C (b), Q (c), and L (d) at pH 4, 8, and 10 calculated from wavelength scans in Figure 1.38.

At neutral conditions, Q revealed more helical structure with a helical content of 22 % (Table 1.13, Figure 1.39). The negative control, L, did not illustrate $\alpha$-helical structure at pH 8; rather it presented a single minimum at 213 nm (Figure 1.38). $\beta$-sheet content of L was 49 % at pH 4, and decreased with increasing pH level to 38 % at pH 8 and 32 % at pH 10. At pH 10, Q maintained helical conformation with little change in the double minimum values, while L exhibited the same single minima (Figure 1.38, Table 1.13).

While residue P39 in L may have contributed to a small loss in structure (as was seen in Q), the complete loss in helical structure across all pH conditions can be attributed to the disruption of the N-terminal pocket due to domain swapping. As the pH increased, the helical content of soluble Q protein increased from 9 to 25 %. L exhibited a

small increase in helical content (9, 13, and 16 % at pH 4, 8, and 10, respectively) and a moderate increase in random coil content (42, 49, and 52 % at pH 4, 8, and 10, respectively).

While CD is able to provide information of the secondary structure content of the proteins that are soluble, it is unable to provide information on insoluble protein assemblies, such as fibers. Additionally, large discrepancies between structural information determined by CD versus X-ray data have been found, whereas ATR-FTIR data provides a much more accurate picture of a protein's secondary structure and is in much better agreement with crystallographic data.[121] As a result, ATR-FTIR experiments were conducted to evaluate secondary structure of insoluble Q fibers at pH 4, 8, and 10 (Figure 1.40 and Table 1.14).

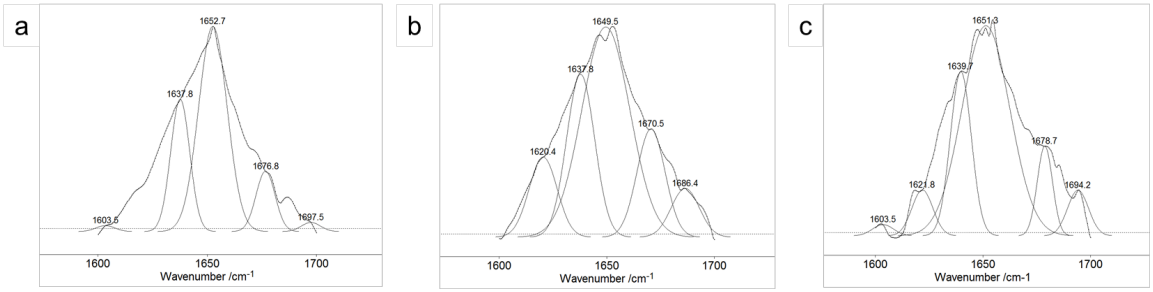



Figure 1.40. ATR-FTIR of 10 μM Q at pH 4 (a), pH 8 (b), and pH 10 (c) measured at room temperature.

Table 1.14. Secondary structure conformation based on ATR-FTIR data for 10 μM Q in 50 mM PB at pH 4, 8, and 10. Percent composition was determined from relative areas of peaks fit to spectra (Figure 1.40).

| | | % composition | | |
|---|---|---|---|---|
| Conformation | Wavelength [$cm^{-1}$][122] | pH 4 | pH 8 | pH 10 |
| $\beta$-sheet | 1625-1640, 1675-1695 | 38 | 37 | 32 |
| Random coil | 1640-1648 | - | - | - |
| $\alpha$ helix | 1648-1660 | 62 | 63 | 68 |

The frequency measured in the regions of the amide I and amide II absorptions of a protein correlate to the secondary structural motifs within the protein,[122] and were thereby used to assess conformation of the protein in its solid-state (Figure 1.40). Positions of amide I peaks in deconvoluted IR spectra of coiled-coil proteins have been shown

previously to differ compared to peak locations arising from purely $\alpha$-helical, monomeric proteins.[123] Deviations are related to pitch values of the $\alpha$-helices within the coiled-coils, with dimers showing the largest deviation (corresponding to helix deformations) and higher order oligomeric coiled-coils more closely resembling $\alpha$-helical proteins.[123] Our data correlate well with these observations, as the significant peak weights lie near the classical $\alpha$-helical band position of 1650-1653 $cm^{-1}$ (Figure 1.40). ATR-FTIR measurements of Q at pH 4 results in a helical content of 62 %, with helicity increasing to 63 and 68 % at pH 8 and 10, respectively (Table 1.14). This trend of increasing helical content with increasing pH is consistent with the CD data (Figure 1.39). As will be further explained in Section 1.4.5.1, solid-state ATR-FTIR data of Q secondary structure confirmed structured fibers visualized in TEM are indeed $\alpha$-helical.

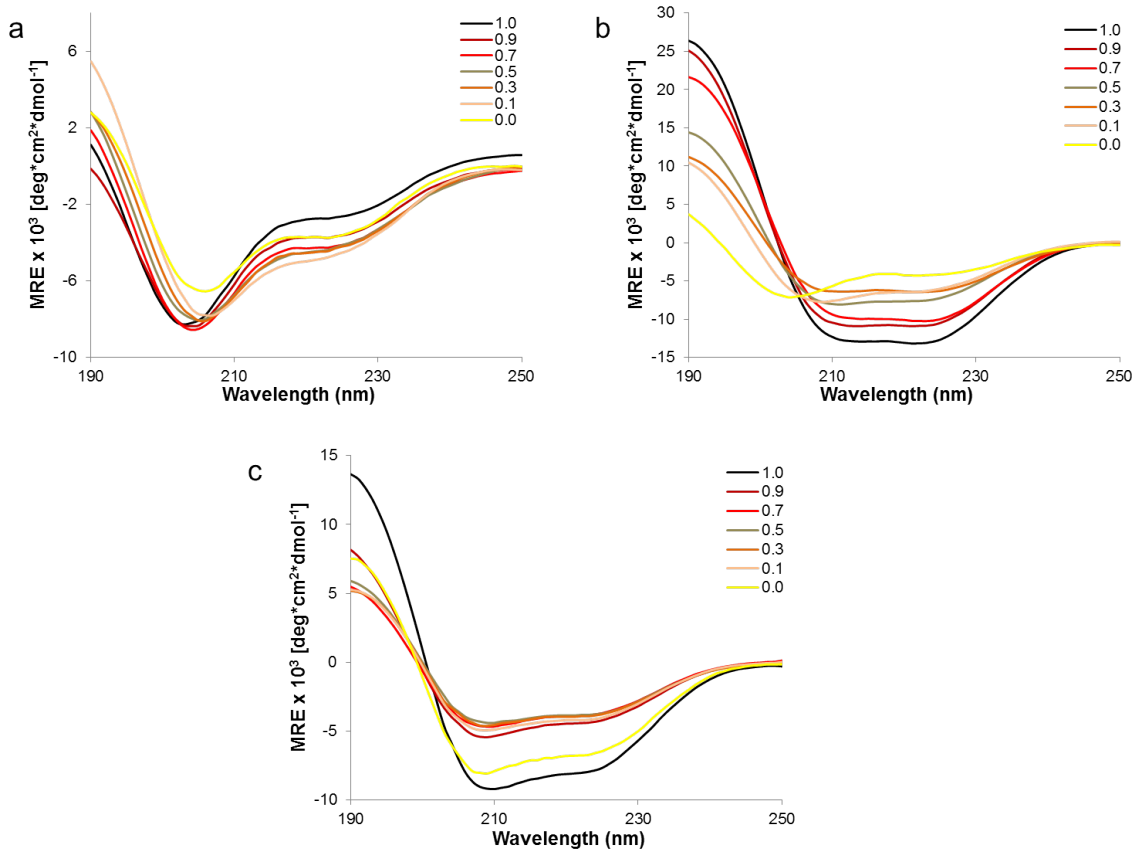



Figure 1.41. Secondary structure represented by circular dichroism variable wavelength scans of mixtures of C and Q of proteins at pH 4 (a), pH 8 (b), and pH 10 (c). Data is averaged from at least two replicates and was obtained with 50 mM PB and 10 µM protein concentration at 25 °C.

Circular dichroism measurements were also performed on mixtures of C and Q proteins, as it was initially hypothesized that these proteins, when mixed in equimolar ratios, would be able to form robust fibers. The conditions that were used to generate these mixtures involve co-dialysis and are described in Section 1.3.3. Molar ratios presented in Figure 1.41 and Figure 1.42 represent the concentration molar ratio of C protein with respect to the total (C and Q). As such, 1.0 corresponds to C alone and 0.0 corresponds to Q alone. The secondary structure of these mixtures was evaluated at pH 4, 8, and 10.

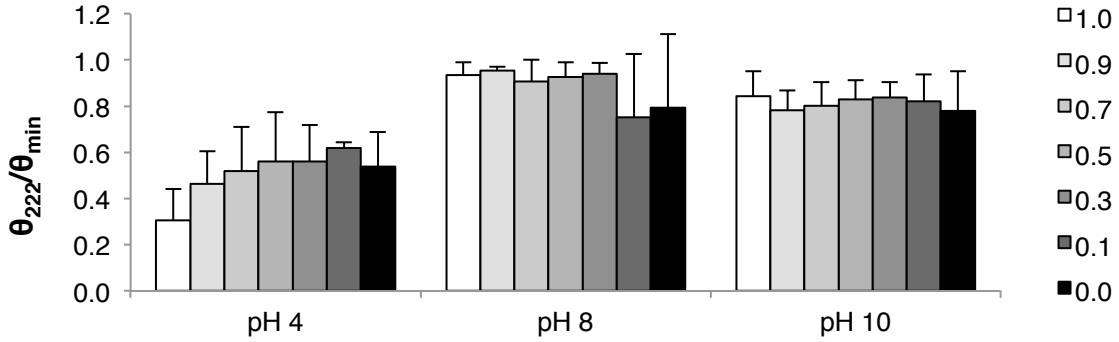



Figure 1.42. Ratio of $\theta_{222}$ to $\theta_{min}$ for mixtures of C and Q at various pH conditions.

The ratio of the ellipticity minima at 222 nm and that between 195 – 210 nm is a good indicator of helical content, where helical content of unity indicates a perfectly helical structure.[43] As can be seen in Figure 1.42, at pH 4 $\theta_{222}/\theta_{min}$ increases as concentration of C decreases within the mixture. Taking into consideration the standard deviations of the values at pH 8 and 10, the variation in the composition of C and Q plays a small role in altering the ratio of $\theta_{222}/\theta_{min}$. That being said, Figure 1.41 shows a very strong helical signature of C and pH 8 compared with Q alone. In fact, decreasing the concentration of C increases the value of $\theta_{222}$ successively. At pH 10 the strongest negative minima in the region of 208 and 222 nm correspond with the pure proteins C and Q, whereas mixtures of the two result in diminished signal intensity. The secondary structure of mixtures of C and Q were affected by pH, with an overall increase in helicity under neutral and basic conditions, consistent with other coiled-coil proteins that have been designed to self-assemble into fibers.[66]

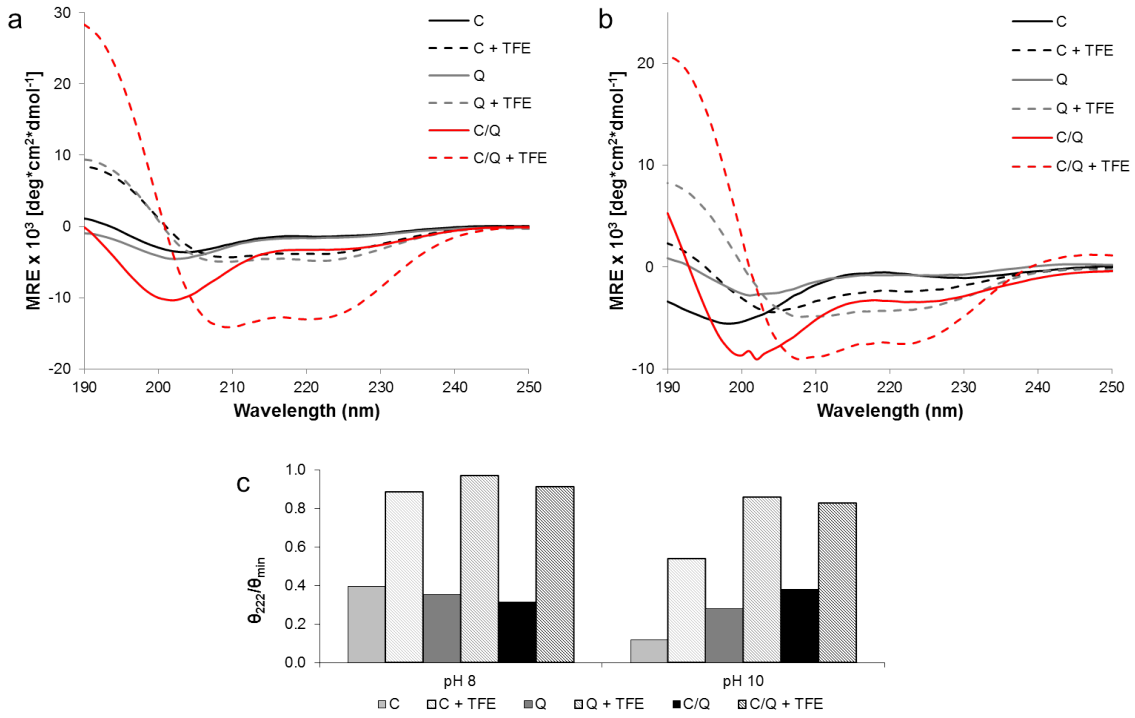



Figure 1.43. Secondary structure represented by circular dichroism variable wavelength scans of mixtures of C (black), Q (grey), and an equimolar mixture of the two (red) at pH 8 (a), and pH 10 (b) in the presence (dashed lines) and absence (solid lines) of 25 v/v % TFE at 25 °C. Proteins in the absence of TFE were in a 10 mM PB solution and at a concentration of 13.3 µM. Scans in the presence of TFE were in a 10 mM PB with 25 v/v % TFE at a concentration of 10 µM. Ratio of $\theta_{222}$ to $\theta_{min}$ for C, Q, and the equimolar mixture at pH 8 and 10 in the presence and absence of TFE (c).

The structure of C and Q were evaluated in the presence of trifluoroethanol (TFE), an alcohol frequently used to stabilize unstable protein formations.[124] Proteins were diluted in 25 v/v % TFE solutions, where the supporting buffer was 10 mM PB at pH 8 (Figure 1.43a) or 10 (Figure 1.43b). It should be noted that the CD signatures for C, Q, and an equimolar mixture of the two in 10 mM PB differs from the signature in 50 mM PB (Figure 1.41). As can be seen in the wavelength scans in Figure 1.43, TFE had a significant impact on the secondary structure of both C and Q individually, and even more so when they were in an equimolar mixture. At pH 8 the ratio of $\theta_{222}/\theta_{min}$ reached approximately unity for C and Q alone, as well as the equimolar mixture of the two. In the absence of TFE at pH 8 $\theta_{222}/\theta_{min}$ of the equimolar mixture was lower than either protein individually, whereas the opposite effect was observed at pH 10.

### 1.4.4.2 Thermodynamic stability

Thermodynamic properties were assessed via thermal melt of the proteins at 222 nm using circular dichroism measurements. Prior to calculating thermodynamic constants, assumptions of monophasic behavior and melt reversibility were confirmed experimentally for wt, Q, and C at all pHs by melting and cooling proteins. These curves are shown in Figure 1.44, where plots (a) - (c) demonstrate reversibility of wt, Q, and C at pH 8.

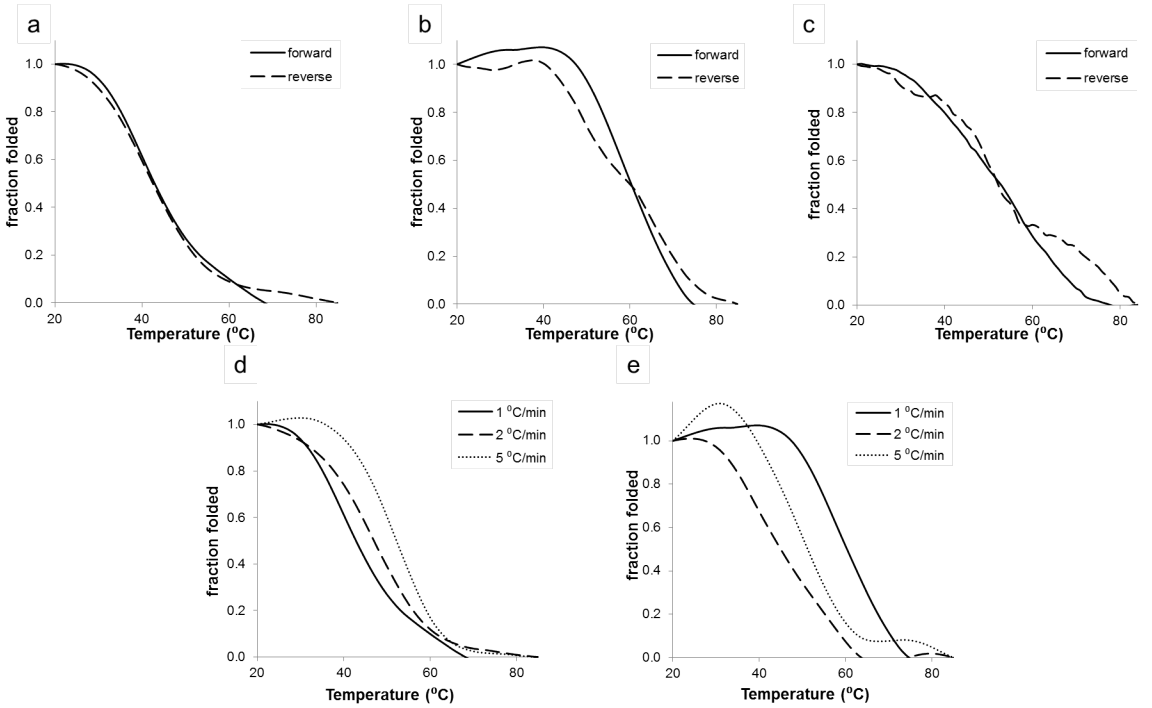



Figure 1.44. Thermal melt scans at 222 nm for wt (a), Q (b), and C (c) at pH 8. Melts were run in the forward, 20 to 85 °C (solid lines), and reverse, 85 to 20 °C (dashed lines), directions to confirm reversibility. Complete reversibility allowed for the van't Hoff analysis to be performed. Thermal melt scans at 222 nm for wt (a) and Q (b) at pH 8 at different scan speeds. Scan speeds of 1 (solid lines), 2 (dashed lines), and 5 (dotted lines) °C/min were run in the forward direction for both proteins.

Melts were also run at different ramping rates for wt and Q at pH 8 (Figure 1.44d and e). The scan rates selected were 1, 2, and 5 °C/min. As can be seen in Figure 1.44, the melt curves for wt and Q at pH 8 are dependent on the scan rate used. In light of this information, all thermodynamic properties have been calculated at the same scan rate, 1 °C/min. Thermodynamic properties are typically only reported at a single scan rate due to the melt's dependency on scan rate.[43,112]

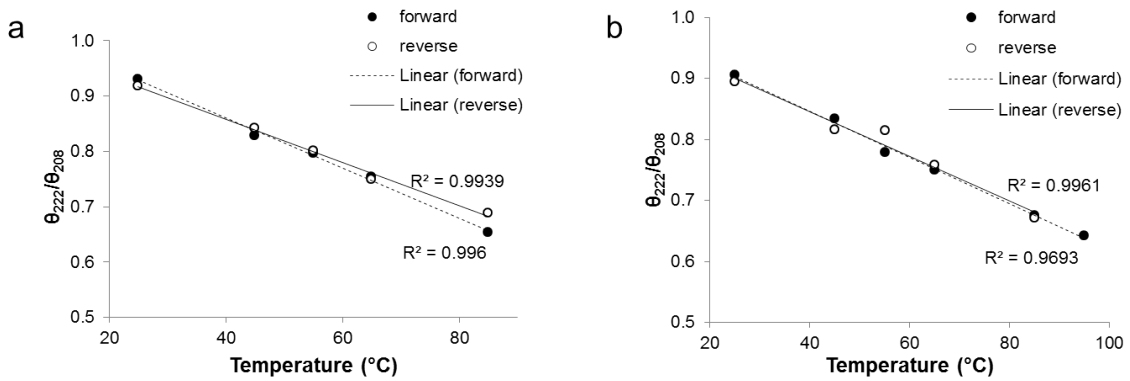



Figure 1.45. Melt reversibility for 20 µM C (a) and Q (b) at 50 mM PB pH 8.

Here it must be mentioned that L was excluded from analysis of thermodynamic properties. As L was completely unstructured, thermal melts of L did not yield a significant enough gradient in the ellipticity at 222 nm to calculate thermodynamic properties. Wavelength isotherms at 25, 45, 55, 65, and 85 °C of L are presented in Figure 1.46. As can be seen from the wavelength curves, the characteristic shape of the unstructured protein remains the same over this temperature range.

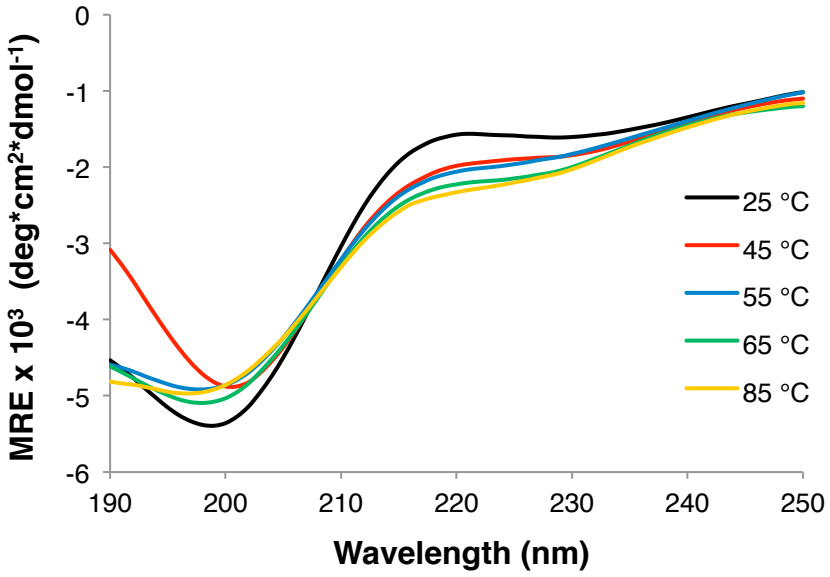



Figure 1.46. Thermal melt isotherms for 10 µM L at 50 mM PB pH 4, where the gradient in the MRE signal at 222 nm was not large enough over this temperature range for determination of thermodynamic properties. Wavelength scans were performed at 25 °C (black), 45 °C (red), 55 °C (blue), 65 °C (green), and 85 °C (yellow).

Thermal melts for wt, Q, and C at pH 4, 8, and 10 are plotted in Figure 1.47. Fraction folded was determined according to the method described in Section 1.3.5.2, where the protein is assumed to be completely folded at 25 °C and completely denatured at 85 °C. Melting temperatures of the proteins were determined to be at the temperature where the fraction folded was equivalent to 0.5. Wt displayed a very consistent melting point that changed little with fluctuations in pH, with $T_m$ equivalent to 48.2, 48.8, and 46.3 °C at pH 4, 8, and 10, respectively (Figure 1.47). Previously reported values of $T_m$ of wt COMPcc$^s$ from our group were on the order of 41 °C.[92] These values fall within the range as the values obtained herein, and the discrepancy can be explained by the use of two different methods for calculating $T_m$, as this value can also be obtained by plotting the first derivative of fraction folded versus the temperature. $T_m$s of Q and C showed more variability with pH, with higher $T_m$s obtained at neutral buffer conditions. $T_m$s of C were 43.8, 52.0, and 50.6 °C at pH 4, 8, and 10, respectively. $T_m$s of Q were 55.3, 61.0, and 46.4 °C at pH 4, 8, and 10, respectively.

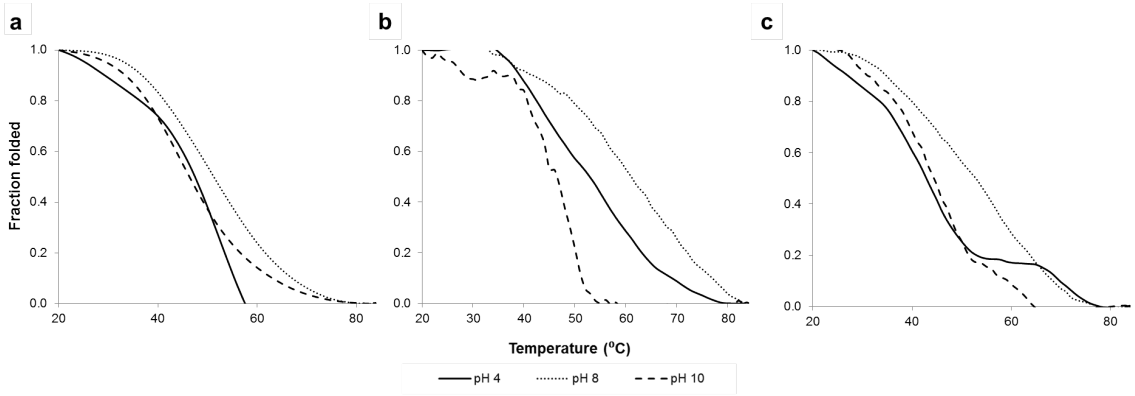



Figure 1.47. Thermal melt scans at 222 nm for wt (a), Q (b), and C (c) at pH 4 (solid lines), pH 8 (dotted lines), and pH 10 (dashed lines). All scans represent an average of two trials.

The van't Hoff equation was applied to thermal melts of the proteins in the range of 20-85 °C. The exact method of applying the van't Hoff equation is outlined in detail in Section 1.3.5.2. In addition to the pure proteins C, Q, and wt, thermodynamic constants were also calculated for mixtures of C and Q with varying molar ratios. Based on the calculations of the thermodynamic constants summarized in Table 1.15, the values for $\Delta H_m$ and $\Delta S_m$ are defined at $T_m$. In reality, $\Delta H$ is dependent on temperature and found to increase significantly with temperature as the heat capacity of the unfolded protein,

$C_p(U)$, is greater than that of the folded protein, $C_p(F)$. The change in heat capacity is therefore defined by Equation 1.14, and can be solved for by the Kirchoff equation (Equation 1.15).[113]

$$\Delta C_p = C_p(U) - C_p(F)$$

Equation 1.14

$$\Delta C_p = \frac{d[\Delta H]}{d[T]}$$

Equation 1.15

Solving for $\Delta C_p$ in this manner is essentially taking the second derivative of experimental data, which needs to be exceptionally good in order for accurate estimations to be made.[113] In order to characterize the relationship between $\Delta H$ and temperature as a function, the temperature dependent protein folding data needs to be collected over a wide temperature range. Differential scanning calorimetry is recommended in order to achieve high quality experimental results with which to perform this type of analysis.[113] The temperature dependence of protein folding depends on $\Delta C_p$ and $\Delta C_p$ depends on the ordering of water molecules around solvent-exposed non-polar groups on proteins. There have been six main sources identified that contribute to heat capacity and entropy changes in the unfolding of proteins, which are: hydrophobic effect, electrostatic charges, hydrogen bonds, conformational entropy, intramolecular vibrations, and changes in equilibria.[125] Amongst these forces it is the hydrophobic effect that has been identified as the major driving force in protein folding.[126] Liu *et al.* have analyzed thermodynamic parameters including free energy, enthalpy, entropy, and heat capacity changes for denaturation of various proteins, discovering that enthalpy changes were largely compensated by a corresponding change in entropy, resulting in a smaller free energy change in total.[127]

Overall, Q exhibited excellent stability across all pH values with melting temperatures ($T_m$) of 46.4-61.0 °C and Gibbs free energy ($\Delta G_m$) of -3.3, -2.7, and -3.1 kcal/mol at pH 4, 8, and 10, respectively (Table 1.15). Compared to the parent wt, Q

demonstrated a 7.1 °C and 12.2 °C increase in $T_m$ at pH 4 and 8, respectively, affirming that the modification made for the design of Q was indeed stabilizing (Table 1.15). At pH 10 the $T_m$ of Q and wt was essentially equal. Overall, Q was more stable at acidic and neutral pH conditions relative to wt, which can be attributed to the surface charge distribution along the pentamer subunits.

Table 1.15. Thermodynamic constants obtained from thermal melts of 10 μM protein measured by circular dichroism, with the intercept of 1/T vs. lnK not equal to zero. Data is averaged from two replicates.

| pH | Protein | $T_m$ (°C) | $\Delta H_m$ (kcal mol$^{-1}$) | $\Delta S_m$ (kcal mol$^{-1}$ K$^{-1}$)[a] | $\Delta G_m$ (kcal mol$^{-1}$)[b] |
|---|---|---|---|---|---|
| 4 | **1.0 (C)** | 43.8 | -32.5 | -102.8 | -1.8 |
| | **0.9** | 55.4 | -31.2 | -95.0 | -2.9 |
| | **0.7** | 49.7 | -30.7 | -95.1 | -2.4 |
| | **0.5** | 51.5 | -22.0 | -69.2 | -1.4 |
| | **0.3** | 56.3 | -32.7 | -98.6 | -3.3 |
| | **0.1** | 47.7 | -44.4 | -137.1 | -3.5 |
| | **0.0 (Q)** | 55.3 | -32.3 | -97.1 | -3.3 |
| | **wt** | 48.2 | -29.9 | -93.0 | -2.2 |
| 8 | **1.0 (C)** | 52.0 | -27.1 | -83.3 | -2.3 |
| | **0.9** | 55.0 | -30.4 | -92.4 | -2.9 |
| | **0.7** | 52.8 | -30.0 | -91.8 | -2.7 |
| | **0.5** | 51.5 | -30.3 | -93.3 | -2.5 |
| | **0.3** | 51.9 | -41.5 | -128.5 | -3.2 |
| | **0.1** | 60.6 | -27.4 | -82.0 | -2.9 |
| | **0.0 (Q)** | 61.0 | -25.0 | -75.0 | -2.7 |
| | **wt** | 48.8 | -31.0 | -96.6 | -2.3 |
| 10 | **1.0 (C)** | 50.6 | -42.1 | -133.4 | -2.4 |
| | **0.9** | 54.2 | -31.7 | -98.8 | -2.3 |
| | **0.7** | 47.0 | -18.3 | -57.2 | -1.3 |
| | **0.5** | 49.1 | -29.1 | -90.4 | -2.2 |
| | **0.3** | 46.5 | -24.4 | -75.5 | -1.9 |
| | **0.1** | 46.6 | -21.7 | -67.0 | -1.7 |
| | **0.0 (Q)** | 46.4 | -46.4 | -145.2 | -3.1 |
| | **wt** | 46.3 | -29.8 | -93.3 | -2.0 |

a) At the transition midpoint, $\Delta G_m = 0$. Hence, the change in entropy $\Delta S_m = \Delta H_m / T_m$. b) Free energy of folding at 25 °C calculated according to the expression $\Delta G_m = \Delta H_m - T\Delta S_m$.

Thermodynamic parameters $\Delta H_m$, $\Delta S_m$, and $\Delta G_m$ were also calculated by assuming that the linear regression of ln$K$ versus 1/$T$ has a y-intercept value of zero (Table 1.16). The values obtained are approximately two orders of magnitude smaller than the values obtained with a y-intercept unequal to zero (Table 1.15). Several values for $\Delta G_m$ presented in Table 1.16 are positive, mainly occurring at acidic and neutral pHs. There are no overarching trends observed from data provided by this estimation regarding comparison of the thermodynamic stability of C versus Q or wt.

Table 1.16. Thermodynamic constants obtained from thermal melts of 10 μM protein measured by circular dichroism, with the intercept of 1/$T$ vs. ln$K$ equal to zero. Data is averaged from two replicates.

| pH | Protein | $\Delta H_m$ (kcal mol$^{-1}$) | $\Delta S_m$ (kcal mol$^{-1}$ K$^{-1}$)[a] | $\Delta G_m$ (kcal mol$^{-1}$)[b] |
|---|---|---|---|---|
| 4 | **1.0 (C)** | -0.1 | -0.2 | 6.1E-03 |
| | **0.9** | 0.2 | 0.8 | -2.0E-02 |
| | **0.7** | 0.1 | 0.3 | -8.1E-03 |
| | **0.5** | -0.1 | -0.3 | 8.0E-03 |
| | **0.3** | -0.3 | -1.1 | 2.8E-02 |
| | **0.1** | -0.3 | -0.9 | 2.4E-02 |
| | **0.0 (Q)** | -0.3 | -1.1 | 2.8E-02 |
| | **wt** | -0.3 | -1.0 | 1.8E-02 |
| 8 | **1.0 (C)** | 0.0 | 0.1 | 1.6E-02 |
| | **0.9** | -0.3 | -1.0 | 3.1E-02 |
| | **0.7** | -0.1 | -0.4 | 2.0E-02 |
| | **0.5** | 0.1 | 0.2 | -3.0E-04 |
| | **0.3** | -0.2 | -0.7 | 3.6E-03 |
| | **0.1** | -0.3 | -1.0 | 2.5E-02 |
| | **0.0 (Q)** | -0.6 | -2.0 | 5.1E-02 |
| | **wt** | 0.1 | 0.3 | -2.1E-02 |
| 10 | **1.0 (C)** | 0.2 | 0.5 | 1.0E-02 |
| | **0.9** | -0.3 | -0.9 | -2.2E-02 |
| | **0.7** | -0.1 | -0.3 | -6.5E-03 |
| | **0.5** | 0.0 | 0.2 | 3.7E-03 |
| | **0.3** | -0.1 | -0.3 | -7.3E-03 |
| | **0.1** | 0.0 | -0.1 | -1.4E-02 |
| | **0.0 (Q)** | -0.6 | -2.0 | -4.2E-02 |
| | **wt** | 0.3 | 1.3 | -3.5E-02 |

a) At the transition midpoint, $\Delta G_m = 0$. Hence, the change in entropy $\Delta S_m = \Delta H_m / T_m$. b) Free energy of folding at 25 °C calculated according to the expression $\Delta G_m = \Delta H_m - T\Delta S_m$.

### *1.4.5 Fiber morphology*

Previous studies of wt and a mutant of wt, T40A, revealed fiber formation.[88] This led us to believe that fibers would also be formed by C and Q, and this fiber assembly and morphology was studied under different pH conditions and various states of mixtures for the proteins using SEM, AFM, and TEM. Mixed proteins were studied with all three techniques, while fibers formed by individual species were studied extensively in TEM under different pH conditions. We have also includ TEM data on wt in Section 1.4.5.3 as a basis for comparison of fiber morphology.

### 1.4.5.1 SEM

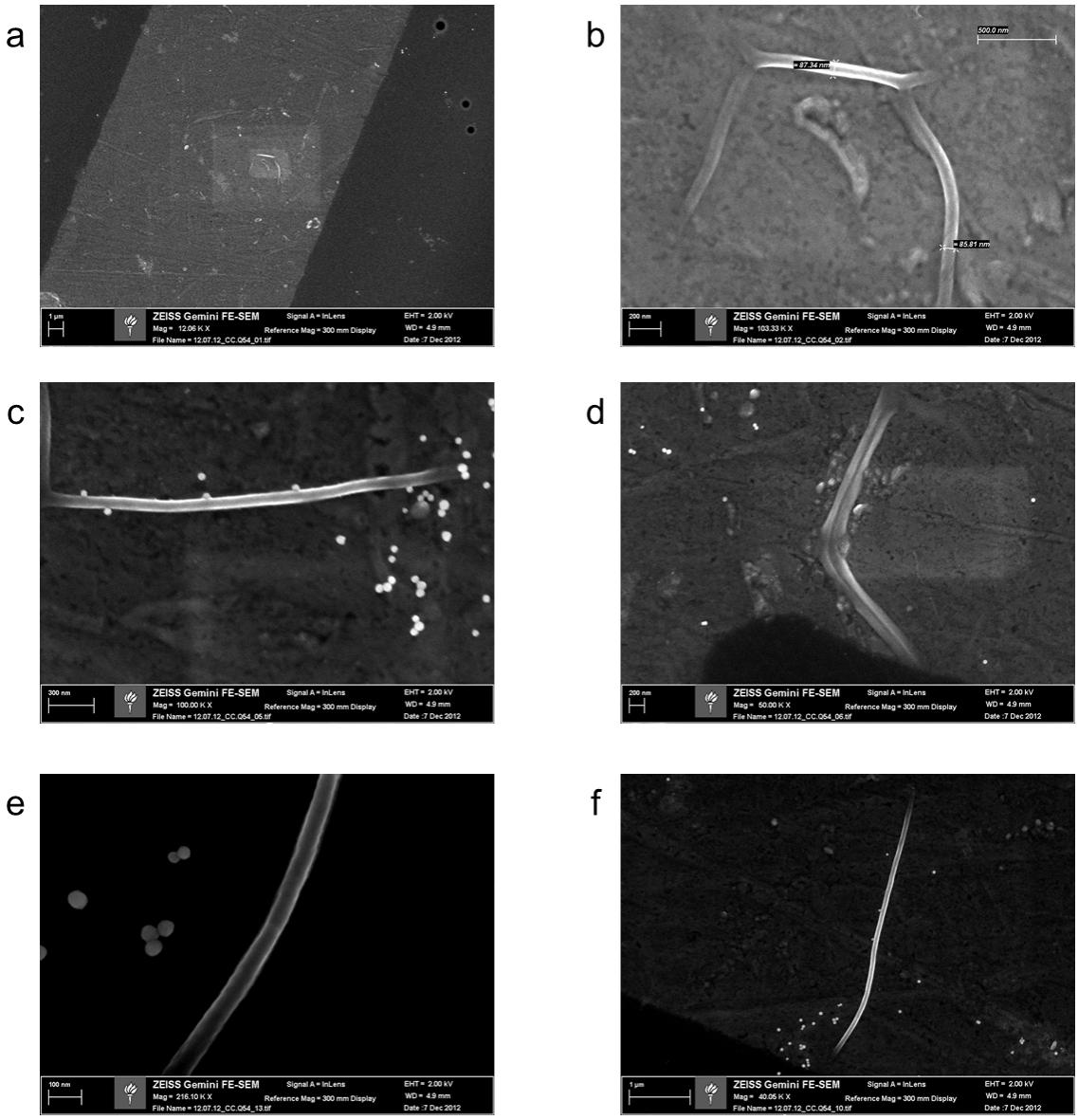



Figure 1.48. SEM images of an equimolar mixture of C and Q at in 50 mM PB pH 8. Scale bars in (c) and (e) represent 300 and 100 nm, respectively, in (a), (c), and (f) represent 1 μm, and in (b) and (d) represent 200 nm.

Scanning electron microscopy (SEM) was used to study shape and size of protein fibers. Preparation of the samples imaged via SEM is described in Section 1.3.12. The fibers that were seen in samples of equimolar mixtures of C and Q at pH 8 are shown in Figure 1.48. The fibers seen in these micrographs possess diameters on the range of 85 – 200 nm, with individual fibers measuring 96.7 ± 6.4 nm (Figure 1.48b), 92.4 ± 7.4 nm (Figure 1.48c),

268.1 ± 17.5 nm (Figure 1.48d), 70.1 ± 4.9 nm (Figure 1.48e), and 117.0 ± 7.9 nm (Figure 1.48f). Figure 1.48d appears to be two fibers bundled together, which would account for the larger diameter. The fact that these fibers were visualized without any metallic coating was very interesting; the contrast provided by the fibers themselves indicates a certain degree of charging within the protein itself.[128] The spherical structures observed in images Figure 1.48c-f are most likely dried salt aggregates.

### 1.4.5.2 AFM

Atomic force microscopy (AFM) was performed on samples of equimolar mixture of C and Q in solvent conditions of 10 mM PB at pH 8 in tapping mode, as is recommended to prevent damage in soft samples.[129] Figure 1.49 depicts an entire fiber of high aspect ratio that is on a mica surface. A topographical heat map was produced by scanning the surface and shows a fiber whose height is approximately 3 nm. In Figure 1.49b the surface scans have been used to generate a 3-dimensional rendering of the surface where the fiber was found.

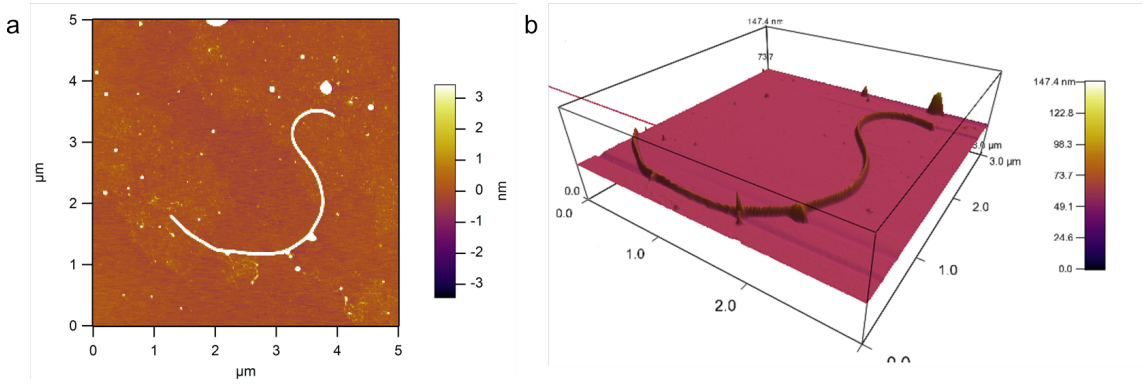



Figure 1.49. Equimolar mixture of C and Q at 10 mM PB pH 8 on a mica surface at room temperature. Heat map indicates topographical height of features on the mica surface, where a fiber on the surface is displayed in white in (a), corresponding to a height of about 3 nm. The 3 dimensional rendering of topographical scan of equimolar mixture of C and Q on mica surface is shown in (b).

A portion of another fiber was studied at a higher magnification, the micrographs for which are shown in Figure 1.50. In these scans the higher magnification allowed for visualization of smaller fibrils that appears to be coming off of the larger fiber, seen in Figure 1.50a, b. The line scan in Figure 1.50c shows that fibril height is approximately 2 nm while larger fiber has a height of approximately 12 nm, while the width of the fibril is

about 30 nm and the width of the adjacent larger fiber is approximately 60 nm. It can also be noted from Figure 1.49 and Figure 1.50 that there is very little material on the mica surface aside from the protein fibers in these scans. AFM measurements are consistent with results obtained in TEM studies (Section 1.4.5.3).

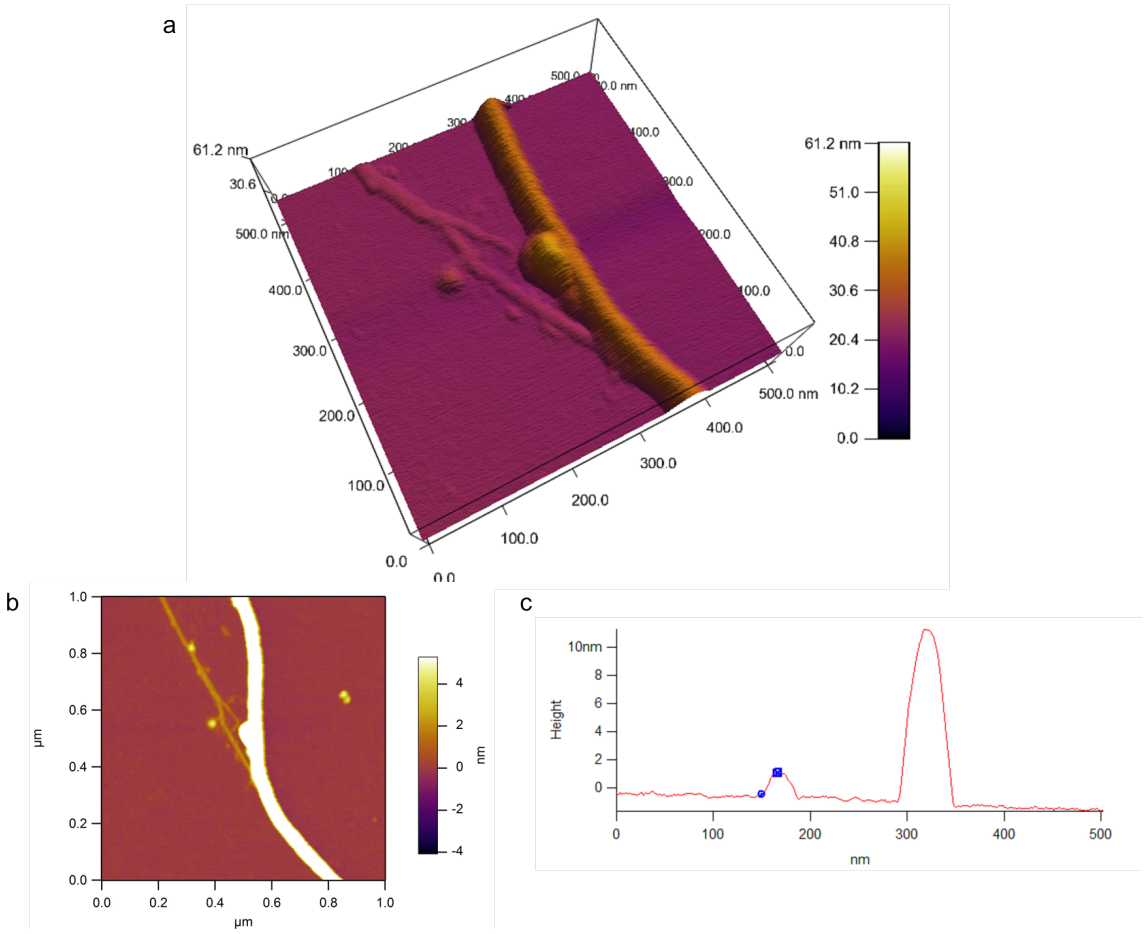



Figure 1.50. Equimolar mixture of C and Q at 10 mM PB pH 8. 3 dimensional rendering of topographical scan of equimolar mixture of C and Q on mica surface shown in (a). Fiber with adjacent protofibril can be seen in the scans in (a) and (b). Heat map indicates topographical height of features on the mica surface, where a fiber on the surface is displayed in white in (b). Line scan of the fiber is shown in (c).

### 1.4.5.3 TEM

Transmission electron microscopy (TEM) was used extensively to study the size and shape of the protein fibers that were generated. In the images presented and discussed in this section, the samples were stained using 1 % uranyl acetate solution, which provides an outline to the edge of the protein material and gives the protein enough contrast to be visualized without difficulty.

As mentioned previously, wt protein has been seen to form robust networks of fibers under TEM.[88] Early in our studies of the engineered variants of the wt protein, we conducted TEM measurements of wt at pH 8, 10 mM PB, with a protein concentration of 10 μM and observed fiber networks such as those depicted in Figure 1.51. At a lower magnification, many fibers can be seen scattered across the surface of the TEM grid (Figure 1.51a). Upon closer examination, one can see that dense networks of protein fibers are formed on the surface, where the fibers are so long that they appear to be intertwined with one another. This intertwining makes if very challenging to measure the length of singular fibers. The diameters of these fibers are between 10-15 nm.

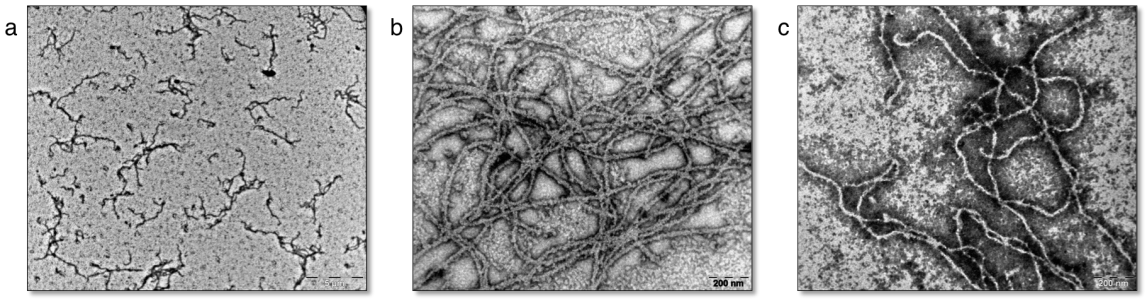



Figure 1.51. Transmission electron micrographs of wt fibers, 10 μM, pH 8, 10 mM PB at room temperature. Scale bar in (a) represents 5 μm and in (b) and (c) represents 200 nm.

The first experiments involving TEM imaging of C and Q proteins were performed under the same solvent conditions that had previously been used to generate fibers of wt. Initially, it was our focus to test our hypothesis that an equimolar mixture of C and Q would produce robust protein fibers. The CD data of secondary structure (Section 1.4.4.1) indicated that mixtures of C and Q in these proportions gave high $\theta_{222}/\theta_{min}$ ratios, and SEM and AFM experiments proved the presence of fibers (Sections 1.4.5.1 and 1.4.5.2). Surprisingly, TEM experiments under these conditions did not reveal distinct fibers, but large sheets of protein (Figure 1.52). Sheet formations of the kind seen in Figure 1.52 are distinct for the equimolar mixtures of C and Q proteins. Assemblies with this type of morphology may arise due to the complementarity between the C and Q subunits and lateral assembly of the pentamers formed by a mixture of both C and Q (Figure 1.19, Figure 1.21, and Figure 1.22).

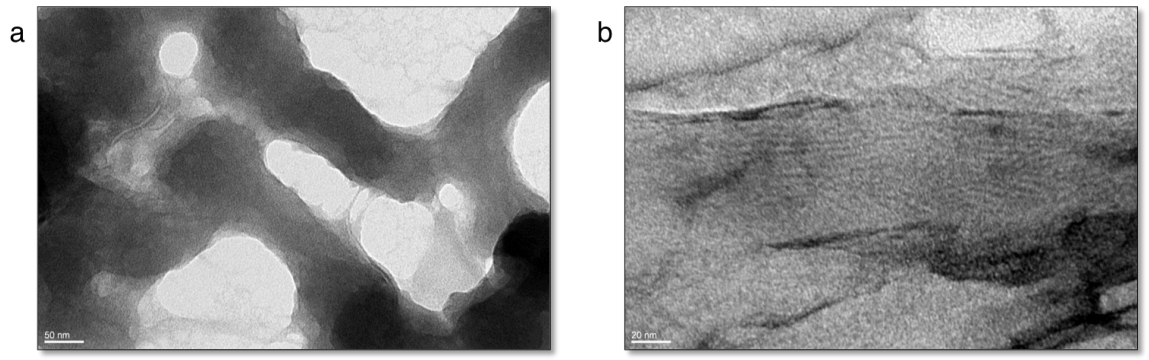



Figure 1.52. Transmission electron micrographs of sheets formed by equimolar mixture of C and Q, 10 μM, pH 8, 10 mM PB at room temperature. Scale bar in (a) represents 50 nm and in (b) represents 20 nm.

TEM was also performed on equimolar mixtures of C and Q in the presence of 25 % TFE (Figure 1.53), as this was seen to greatly enhance helical content at pH 8 (Figure 1.43a). Under these conditions distinct fibers were observed, with protofibrils visible within the fibers. In addition to fibers, there were several non-structured protein aggregates also present on the TEM grids. The micrographs in Figure 1.53 provide evidence that the protofibrils bundle together to form larger scale fibers.

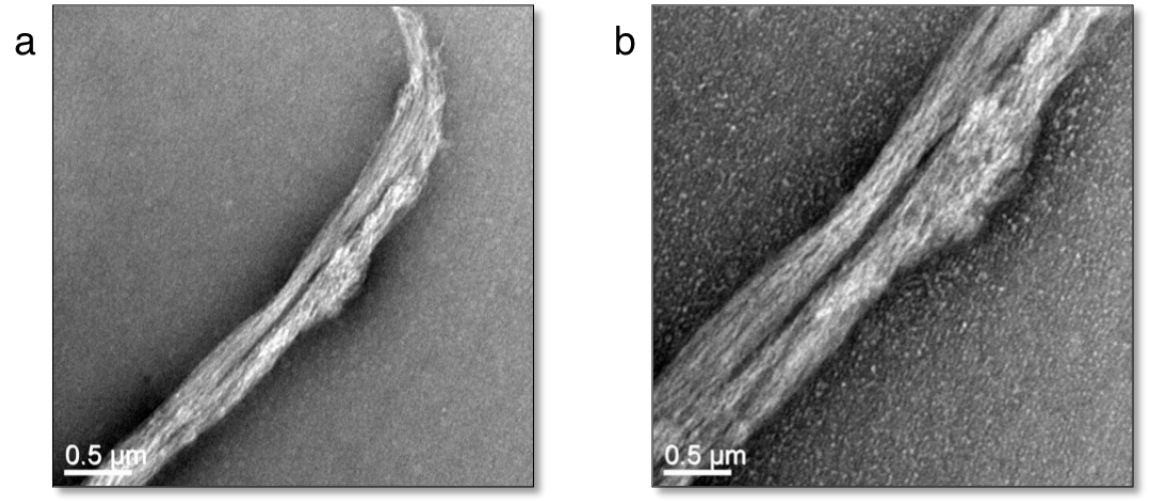



Figure 1.53. Transmission electron micrographs of equimolar mixtures of C and Q, 20 μM, 10 mM PB pH 8 in the presence of 25 v/v % TFE at room temperature. Distinct fibers could be seen, with protofibrils. Scale bars represent 0.5 μm in both (a) and (b).

It was also desired to study the morphology of the proteins assemblies under acidic and basic conditions, as was also done in studies concerning the secondary structure and thermodynamic properties of these materials (Section 1.4.4). Several

micrographs of the equimolar mixtures of C and Q at pH 4, 8, and 10 are shown in Figure 1.54. At each pH condition several morphologies were observed: fibers, sheets, and non-structured aggregates. That being said, a higher proportion of fibers was observed under acidic conditions than under neutral or basic conditions. Fibers were seen across all pH conditions, determining the composition of the fibers proved challenging (*vida infra*).

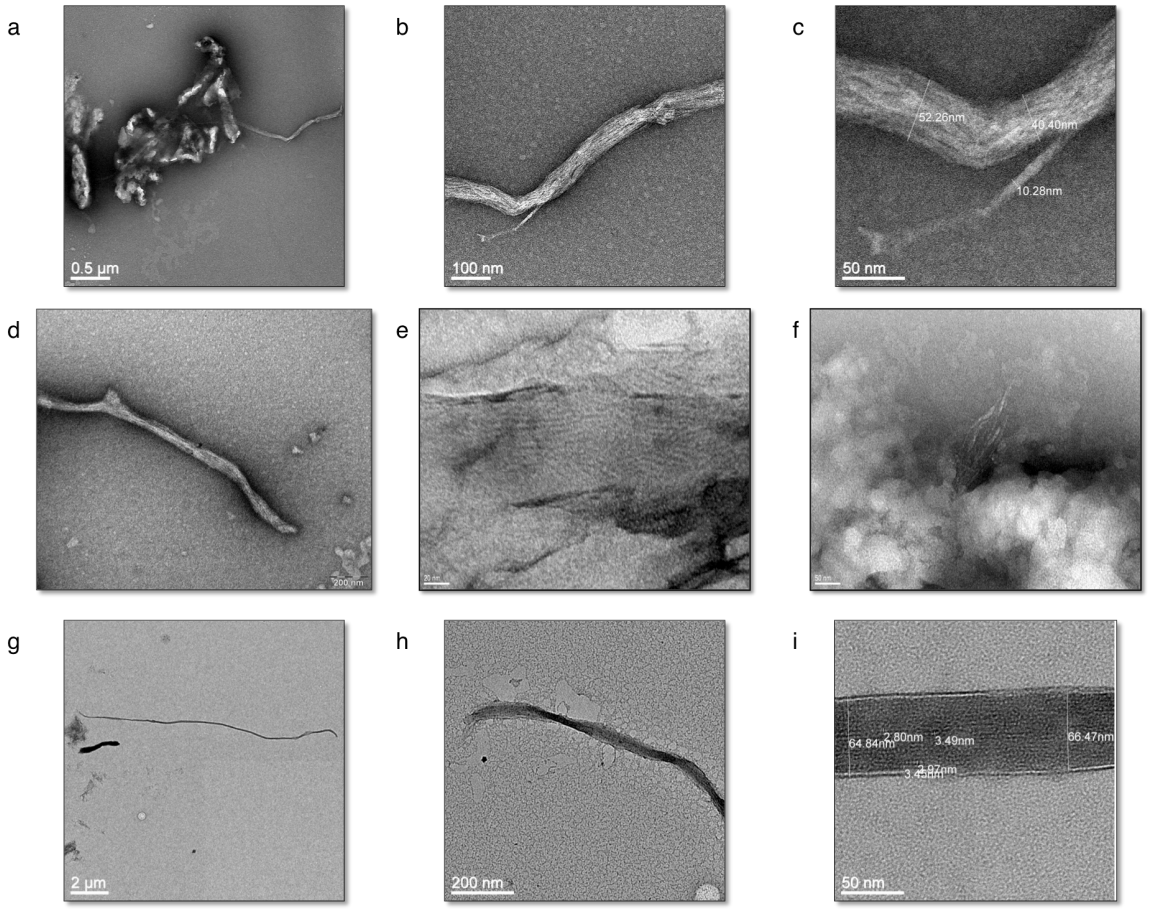



Figure 1.54. Transmission electron micrographs of equimolar mixtures of C and Q, 5 µM, 10 mM PB at various pH conditions: pH 4 (a)-(c), pH 8 (d)-(f), and pH 10 (g)-(i) at room temperature. Distinct fibers could be seen at pH 4. Distinct fibers were not seen at pH 8, although striations were observed that had a regularity of approximately 3 nm. Distinct fibers were seen at pH 10, and striations were observed that had a regularity of approximately 3 nm. Scale bars represent 0.5 µm in (a), 100 nm in (b), 50 nm in (c), (f), and (i), 200 nm in (d) and (h), 20 nm in (e), and 2 µm in (g).

Several studies have noted increased propensity for fiber formation with solvents of higher ion concentrations.[25,66,130,131] As such, we also dialyzed protein mixtures in 50 mM PB at pH 8. An increase in the proportion of fibers, with respect to other macromolecular structures, was observed in the case of equimolar mixtures of C and Q at

a protein concentration of 20 μM. These high aspect ratio fibers are shown in Figure 1.55. Protofibrils were also visible within these fibers.

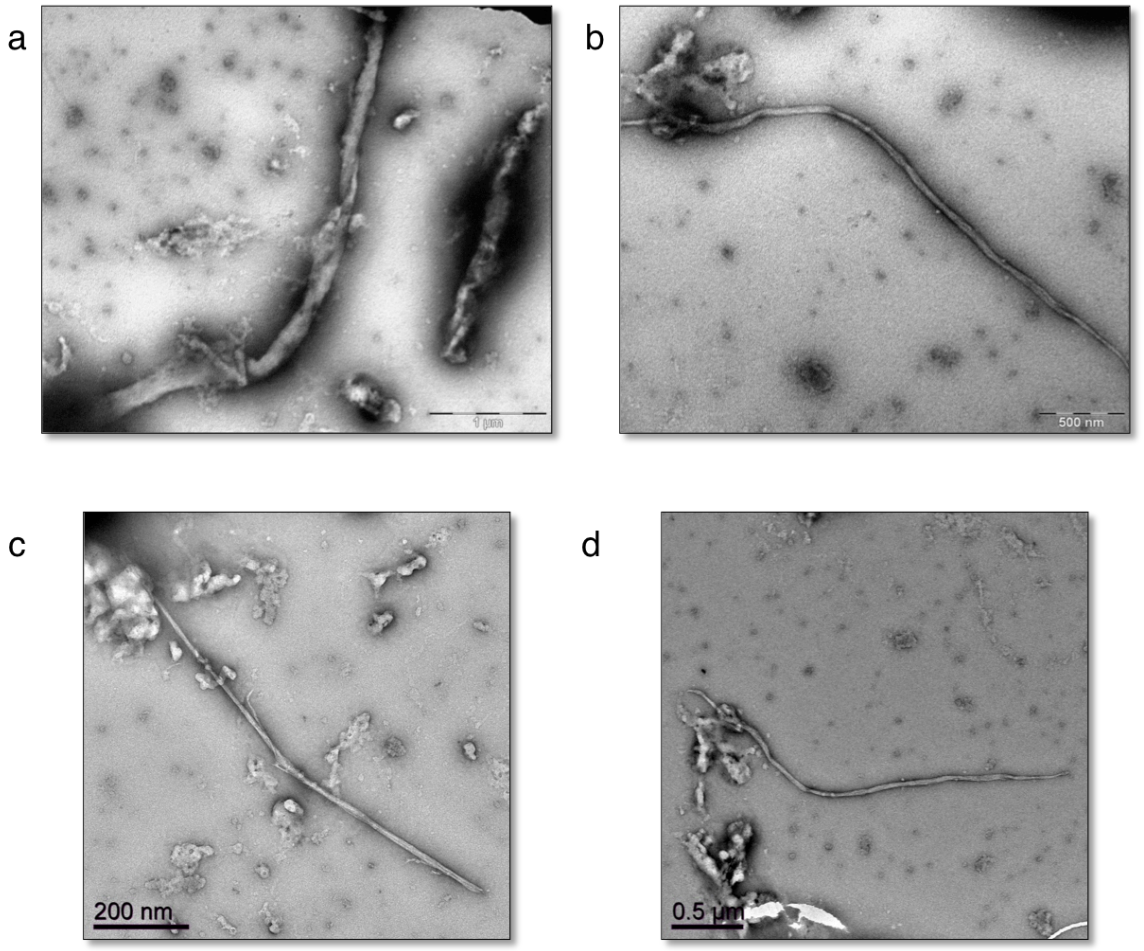



Figure 1.55. Transmission electron micrographs of equimolar mixture of C and Q, 20 μM, 50 mM PB pH 8 at room temperature. Distinct fibers could be seen, with visible protofibrils. Scale bars represent 1 μm in (a), 500 nm in (b), 200 nm in (c), and 0.5 μm in (d).

Despite the presence of protein sheets and fibers in mixtures of C and Q, the study of these mixtures by TEM proved inappropriate, as these micrographs provide no information on what composition the protein sheets and fibers in Figure 1.52 - Figure 1.55 possessed. Although it was clear that macromolecular assembly of the proteins was indeed taking place, the distribution of C and Q within those materials could not be determined via TEM. As there were several types of structures visualized – fibers, sheets, and aggregates – it was not possible to determine which protein comprised which type of assembly and in what proportion. Each of these different assemblies could have been

entirely C, entirely Q, or a combination of both proteins. Without the ability to probe the composition of the assemblies, we chose to evaluate protein fiber formation of homogeneous C or Q protein.

Homogeneous samples of C and Q were first evaluated at pH 8 in 10 mM PB at room temperature (Figure 1.56). C was observed in Figure 1.56a-d, where what resemble flattened fibers could be seen. Networks of these fibers were seen intertwined with one another, and sheet-like assemblies were also seen (Figure 1.56b). The uranyl acetate stain permits us to visualize protofibrils that are parallel with the axis of the fiber. The overall diameter of these fibers ranged from approximately 50 – 300 nm, with specific fibers measuring 67.7 ± 14.3 nm (Figure 1.56a), 250.1 ± 34.6 nm (Figure 1.56c), and 166.2 ± 17.5 nm (Figure 1.56d). Homogeneous assemblies of Q could be seen in Figure 1.56e-f. Contrary to C, under these conditions Q did not form visible fibers, rather forming large sheets of protein. Upon close inspection, areas of these sheets appeared striated (Figure 1.56e). These striations were reminiscent of the protofibrils seen in many of the protein fibers visualized, also possessing widths of 3-5 nm. These regions of the protein sheets could in fact contain areas of organized protein pentamers.

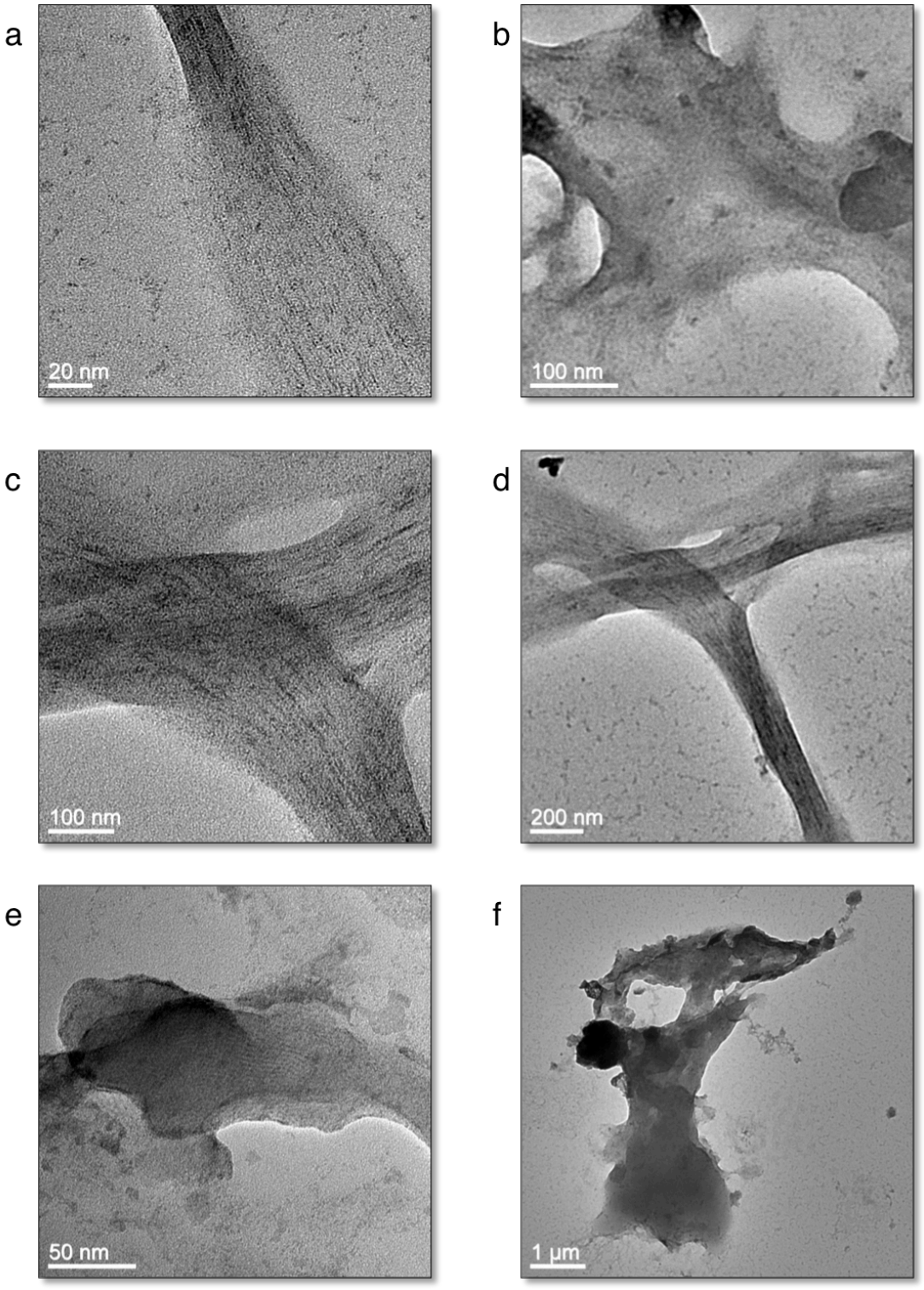



Figure 1.56. Transmission electron micrographs of C (a)-(d) and Q (e)-(f), 20 µM, 10 mM PB pH 8 at room temperature. Distinct fibers could be seen, with protofibrils for C whereas striations could be seen in large sheet-like aggregates for Q under these conditions. Scale bars represent 20 nm in (a), 100 nm in (b) and (c), 200 nm in (d), 50 nm in (e), and 1 µm in (f).

While limited fiber formation was observed for Q at neutral pH, an abundance of fibers was observed under acidic conditions as expected from our design (Figure 1.57a, Figure 1.59), and consistent with previous work on coiled-coil fibers.[47,58] Contrary to previous work on the wt protein, however, the mechanism of protofibril bundling via self-assembly to form large, bundled fibers is the electrostatic charge distribution throughout the Q pentamers. A model of this assembly is shown in Figure 1.57b, with pentamers arranged longitudinally head to tail and stacking laterally with a slip that permits negative and positive patches of the pentamer to be in close contact with one another.

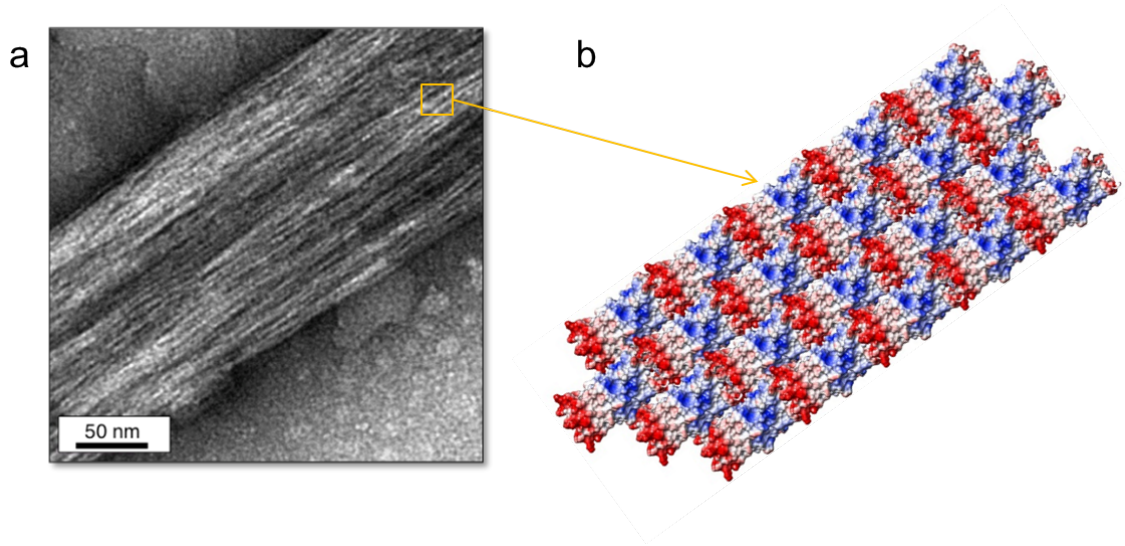



Figure 1.57. Microscopy and modeling of protein fibers. (a) Transmission electron micrograph of Q fiber, 10 μM, pH 4. (b) Schematic representation of Q fiber assembly with staggered positive (red) and negative (blue) regions of the pentamer.

Transmission electron micrographs of Q displayed the presence of bundled protofibrils forming high aspect ratio fibers (Figure 1.57a, Figure 1.59). These protofibrils possessed diameters of 3.5 ± 0.7 nm ($n$ = 210) (Figure 1.58a), which corresponded to the projected lateral dimension of pentameric helical bundles (Figure 1.57b, Figure 1.36e-f). The model shown in Figure 1.57b is an assumption based on the patched distribution of Coulombic surface charge in Q. TEM data confirmed the diameters of protein fibers varied from tens to hundreds of nanometers (20-560 nm, $n$ = 14, Figure 1.58b, Figure 1.59). At the upper limits, these fibers displayed tremendous

lateral assembly of hundreds of protofibrils and thus larger than other *de novo* designed coiled-coil protein fibers to date.[41,44,49]

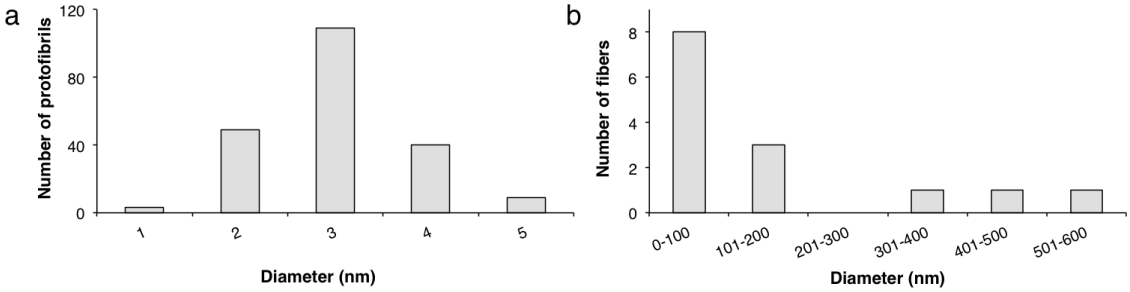



Figure 1.58. Histogram of Q pH 4, 50 mM PB protofibril diameters ($n$ = 210) (a) and fiber diameters ($n$ = 14) (b).

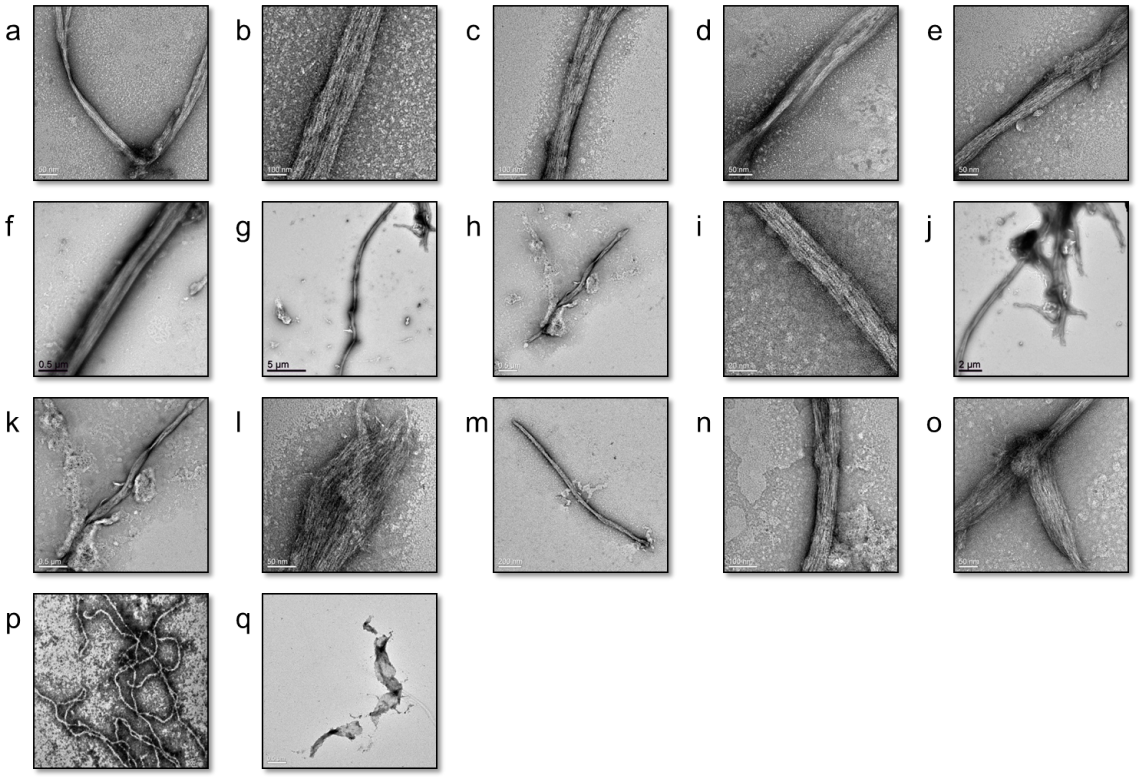



Figure 1.59. Transmission electron micrographs of protein fibers and sheets at room temperature. Transmission electron micrographs Q protein fibers at pH 4, 50 mM PB (a) – (o). A fibrous network is formed by wt (p) at pH 8, 10 mM PB. L (q) forms sheet-like assemblies and larger aggregates at pH 4, 50 mM PB conditions. Scale bars in (a), (d), (e), (l), (o) represent 50 nm, in (b), (c), (i), and (n) represent 100 nm, (m) and (p) represent 200 nm, (f), (h), (k), and (q) represent 0.5 μm, (j) represents 2 μm, and (g) represents 5 μm.

While wt demonstrated fiber formation (Figure 1.59p), the fibers possess diameters on the range of 10-15 nm[88] without any evidence for protofibrils. As expected, the L

negative control revealed non-fibrous aggregates affirming that $\alpha$-helical conformation was indeed important for fiber assembly (Figure 1.39). Results confirming the ability of Q to form high aspect ratio fibers under acidic conditions allowed us to explore its ability to bind with and encapsulate the small hydrophobic molecule curcumin, as is described in Section 1.4.6.

### *1.4.6 Curcumin binding*

Interactions of wt with curcumin have previously been investigated by our group.[92] In the design of Q, the hydrophobic pore was maintained to enable binding to small molecules. Thus, we decided to incubate Q with varying concentrations (0-80 μM) of curcumin at pH 4 in order to study its overall binding ability and effects on protein morphology, charge, and aggregation. The results of these studies are presented in the sections that follow.

#### 1.4.6.1 Binding kinetics

Curcumin concentration was varied from 0 to 160 μM, keeping the concentration of protein constant at 20 μM. Figure 1.60 displays the absorption spectral changes of wt, C, and Q in the presence of curcumin in the wavelength range of 350 – 650 nm at pH 8, 50 mM PB. A blueshift of approximately 15 nm was observed in the peak absorbance of wt and Q, with a negligible shift in absorbance in C.

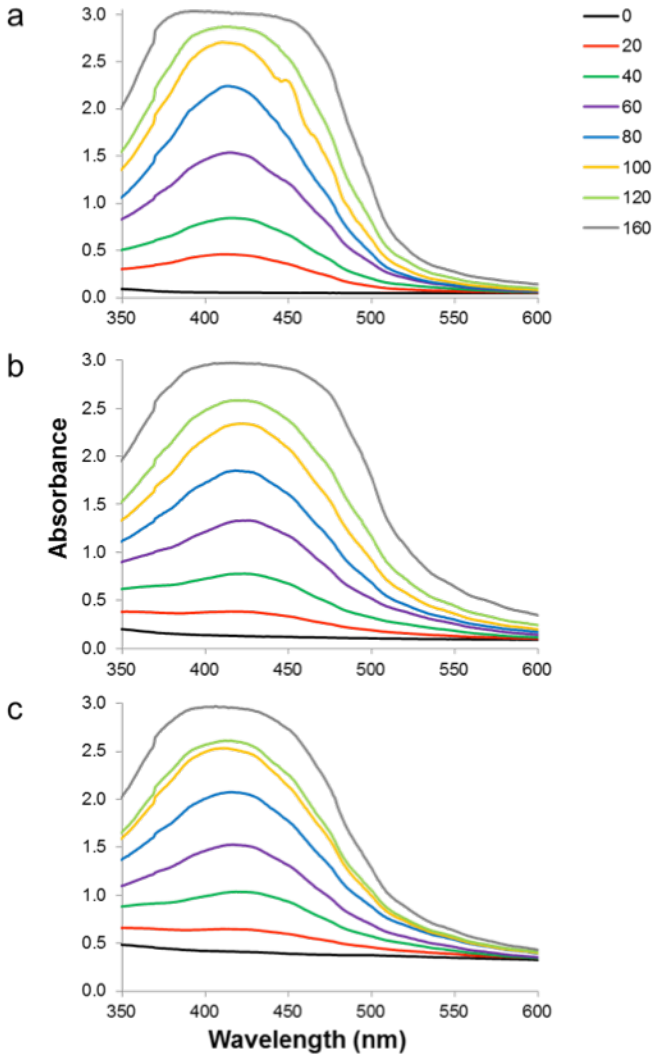



Figure 1.60. Absorption spectra of aqueous solutions containing 20 μM protein with varying concentrations of curcumin, from 0 to 160 μM at pH 8. Absorption spectra of wt and curcumin are presented in (a), C and curcumin in (b), and Q and curcumin in (c). Black lines are the protein in the absence of curcumin, red is 20 μM curcumin, green is 40 μM, purple is 60 μM, blue is 80 μM, yellow is 100 μM, green is 120 μM, and grey is 160 μM.

Section 1.3.11 describes in detail the mathematical manipulation that was employed to obtain extinction coefficients, binding constants, and number of binding sites for wt, C, and Q, as presented in Table 1.17. A linearized plot of the inverse in the change in absorbance over the concentration range versus the inverse in the curcumin concentration (Figure 1.61) was used to calculate the values presented in Table 1.17.

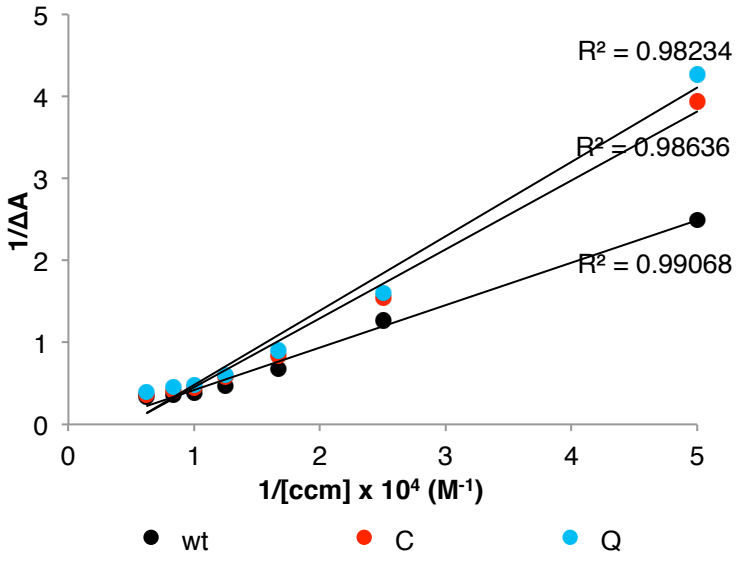



Figure 1.61. Linear plot for 1/ΔA vs. inverse curcumin (ccm) concentration.

The correlation between the inverse of the change in the absorbance and the inverse of the concentration of curcumin are highly linear in the case of all three proteins, as is seen in Figure 1.61. This is proven by correlation coefficients near unity of 0.99, 0.99, and 0.98 for wt, C, and Q, respectively. The linear correlation in this data permitted for estimations of extinction coefficients from the Benesi-Hildebrand equation and binding constants and number of binding sites via the Scatchard equation.

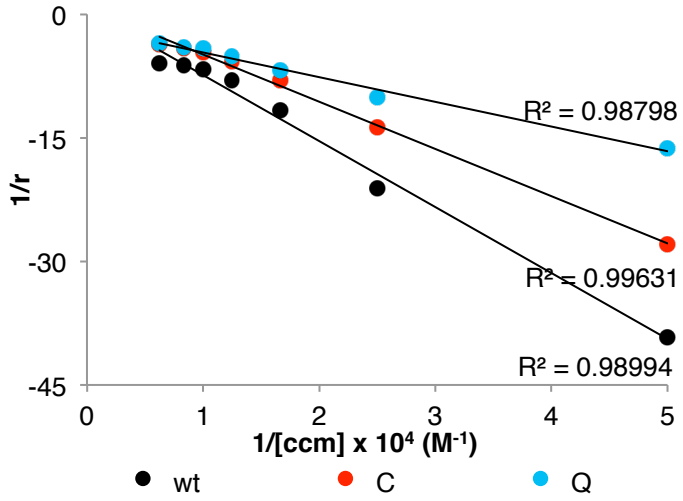



Figure 1.62. Linear plot for 1/r vs. inverse curcumin (ccm) concentration.

Correlation coefficients for wt, C, and Q in Figure 1.62 are all equal to 0.99. Binding constants and number of binding sites for wt, C, and Q are presented in Table 1.17. The data indicates that there is *at least* one binding site for curcumin on both wt and C. It is possible that the number of binding sites is slightly reduced in C compared to wt since it is truncated by one heptad repeat. Calculations for Q produced a negative number of binding sites (*n* equal to -0.6), resulting from a positive value of the binding constant, *K*. Looking into the data, the positive value of K was obtained due solely to a higher than expected absorbance at 420 nm in the presence of 20 μM curcumin (equal to 0.65 RFU), which can easily result from experimental error. Modeling this data point to account for this error we can assume an average value of $A_{420}$ for wt (equal to 0.45 RFU) and $A_{420}$ for C (0.38 RFU), equaling an estimated 0.42 RFU. By substituting this estimated $A_{420}$ value for Q, the Scatchard analysis reveals a *K* value of -0.15 and the number of binding sites (*n*) equals 1.7. These values are much more in line with those expected by this type of analysis, and experimental error can thus be blamed for the negative *n* value obtained from Q in Table 1.17.

Table 1.17. Extinction coefficients, binding constants, and number of binding sites for wt, C, and Q determined by the Benesi-Hildebrand or Scatchard equations. *Modified Q data represents the adjustment of a singular data point (absorbance at 420 nm at 20 μM concentration of curcumin).

| | | Protein | | | |
|---|---|---|---|---|---|
| | | wt | C | Q | |
| | | | | Original | Modified* |
| Benesi-Hildebrand | $\Delta\varepsilon_{420} \times 10^5_{protein}$ ($M^{-1}$ $cm^{-1}$) | -5.09 | -1.29 | -1.18 | -0.01 |
| | $\Delta\varepsilon_{420} \times 10^5_{protein + ccm}$ ($M^{-1}$ $cm^{-1}$) | -9.53 | -5.73 | -5.62 | -4.46 |
| Scatchard | κ ($M^{-1}$) | -0.08 | -0.15 | 0.55 | -0.15 |
| | n | 1.6 | 1.2 | -0.6 | 1.7 |

### 1.4.6.2 Secondary structure in the presence of curcumin

Secondary structure of wt and Q was evaluated in the presence of varying concentrations of curcumin via circular dichroism. ATR-FTIR measurements were performed for Q in the presence of curcumin in a 5:1 molar ratio of curcumin:protein. These measurements were done in acidic buffer conditions at pH 4.

Circular dichroism measurements reveal that the conformation of wt and Q is not disturbed upon interaction with curcumin, which is in agreement with published structural data on the binding of curcumin to collagen.[80] In fact, the absolute value of the MRE value at 222 nm, an indicator of helicity, of wt and Q increased linearly with increasing concentrations of curcumin (insets of Figure 1.63). This suggests a stabilization of coiled-coil assembly for both wt and Q in the presence of curcumin.

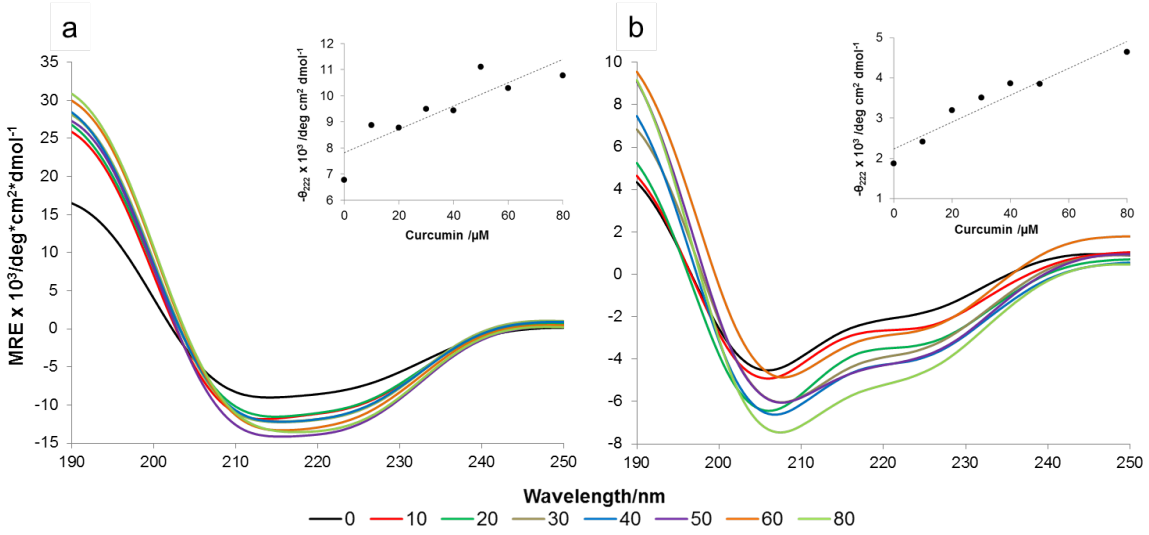



Figure 1.63. Helicity of wt and Q increase as a result of increasing curcumin concentration, pH 4 50 mM PB at room temperature. Circular dichroism wavelength scans of wt (a) and Q (b) with varying molar ratios of curcumin. Legend represents curcumin concentration in μM. Inserts show relationship between negative MRE at 222 nm of the mixtures with increasing curcumin/protein molar ratio, displaying near linear fits over the examined range.

ATR-FTIR measurements were performed for Q in the presence of a 5:1 molar ratio of curcumin:protein in pH 4 buffer conditions. The peak at approximately 1653 $cm^{-1}$ characterizes α-helical content. As can be seen in Figure 1.64, this peak dwarfs the peaks corresponding to other secondary structures, such as β-sheets (1625-1640 and 1675-1695 $cm^{-1}$) and random coils (1640-1648 $cm^{-1}$). The spectrum in Figure 1.64 can be compared to the spectrum in Figure 1.40a, corresponding to Q in the absence of curcumin. Comparison of these two spectra reveals a significant increase in the α-helical content of Q in the presence of curcumin under these conditions.

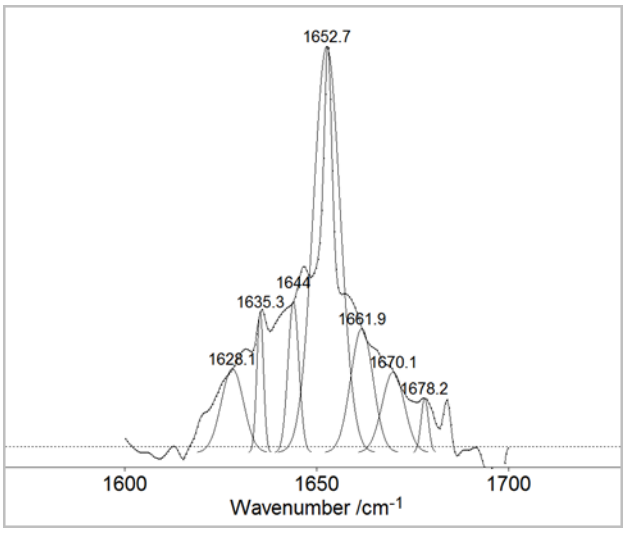



Figure 1.64. Fourier transform self-deconvoluted spectra of Q in 50 mM PB pH 4 in the presence of curcumin at a 5:1 molar ratio. Spectrum represents the average of two trials.

Deconvolution of these spectra allows for estimation of the percent composition for each secondary structure, which is presented for Q in the presence of curcumin in Table 1.18. α-helical content increased from 62 % in the absence of curcumin (Table 1.14) to 79 % in the presence of curcumin at a 5:1 molar ratio. This quantification of the increase in helicity supports the data obtained in CD studies (Figure 1.63b).

Table 1.18. Secondary structure conformation based on ATR-FTIR data for 10 μM Q in 50 mM PB at pH 4 in the presence of 50 μM curcumin. Percent composition was determined from relative areas of peaks fit to spectra (Figure 1.64).

| | | % composition |
|---|---|---|
| Conformation | Wavelength [$cm^{-1}$][122] | Curcumin |
| *β* sheet | 1625-1640, 1675-1695 | 10 |
| Random coil | 1640-1648 | 12 |
| *α* helix | 1648-1660 | 79 |

### 1.4.6.3 Macromolecular assembly

As curcumin bound to structured protein exhibits fluorescence (Figure 1.60),[92] fluorescence and confocal microscopy were performed (Figure 1.72). Surprisingly, fluorescence microscopy revealed many fluorescent protein fibers with diameters on the micrometer scale. These measurements were performed on an equimolar mixture of C and Q at pH 4, 8, and 10 (Figure 1.65) and also of homogeneous C and Q at pH 8 (Figure

1.66). These micrographs show exceptionally bright protein fibers, both within dense networks of fibers (insets of Figure 1.65a, b) and on their own (Figure 1.65, Figure 1.66).

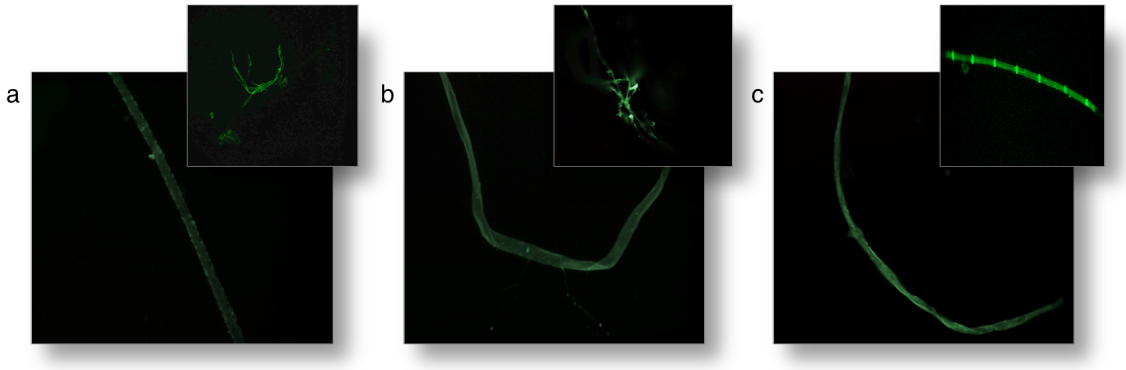



Figure 1.65. Fluorescence microscopy of an equimolar mixture of C and Q with embedded curcumin at 50 mM PB pH 4 (a), pH 8 (b), and pH 10 (c). Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Scale bars omitted due to software limitations. Insets in (a) and (b) show fiber networks which were observed, and inset of (c) shows a protein fiber at pH 10 displaying particularly bright spots along the fiber.

Interestingly, at pH 10 fibers were observed with particularly bright regions that appeared down the length of the protein fiber with some regularity (inset Figure 1.65c).

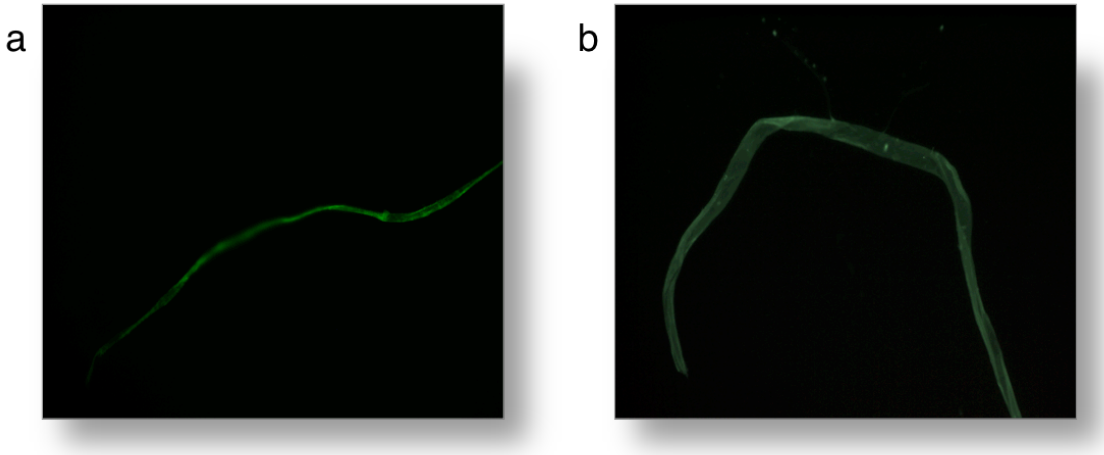



Figure 1.66. Fluorescence microscopy of curcumin-embedded C (a) and Q (b) protein fibers at pH 8 at a curcumin to protein molar ratio of 5:1. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Scale bars omitted due to software limitations.

Unfortunately, exact dimensions were not obtainable from fluorescence micrographs as the ability to inset a scale-bar on each image was not possible due to software limitations. Preliminary images of fluorescent protein fibers (Figure 1.65, Figure

1.66) suggested a three dimensional structure as there were areas of the protein fibers that were always out of focus. To overcome the limitations of fluorescence microscopy, confocal microscopy was performed.

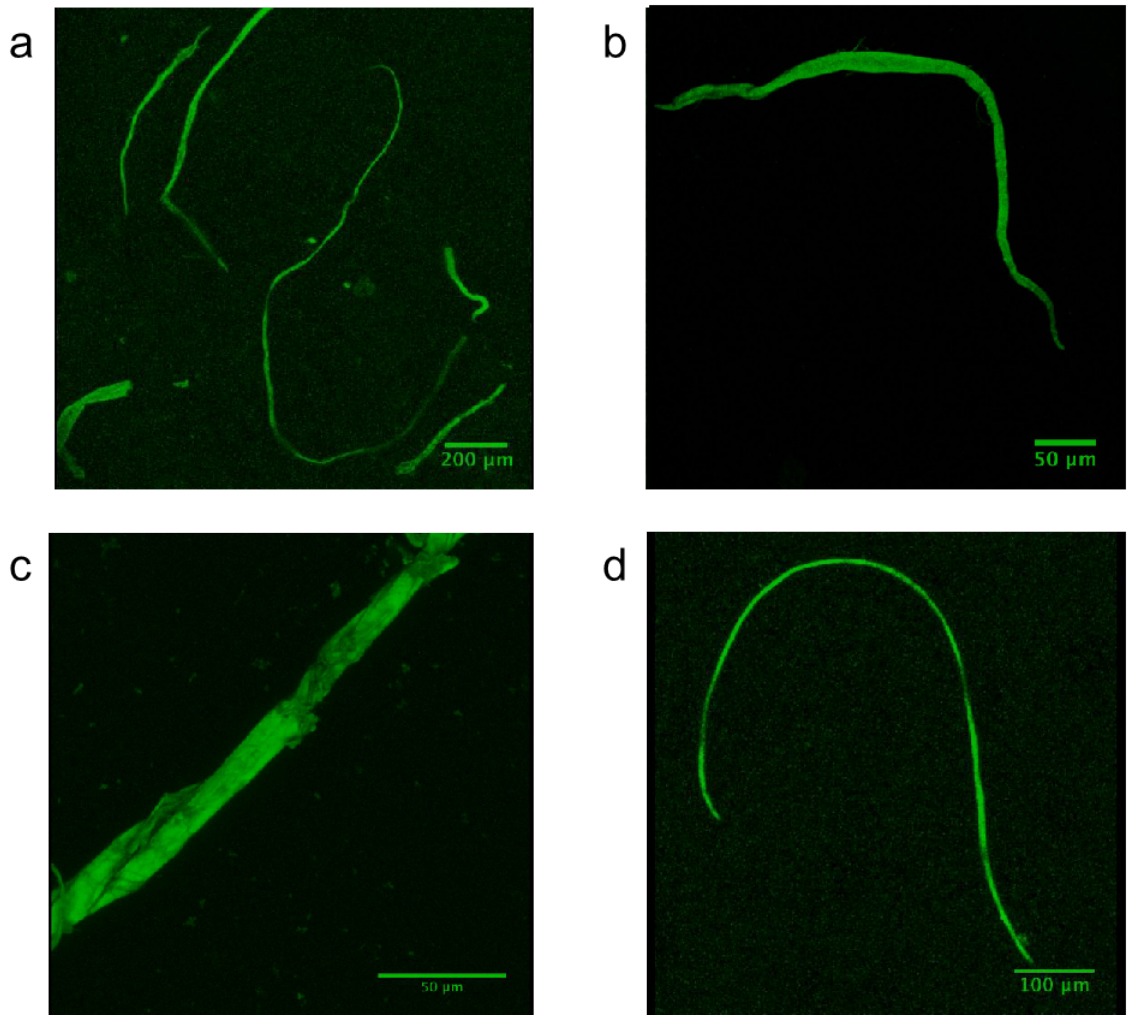



Figure 1.67. Confocal microscopy of an equimolar mixture of C and Q with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 µM was used. Fiber widths in (a) range from 18 - 21 µm, with the fiber in (b) possessing a width of 15.2 µm, (c) 12.8 µm, and (d) 7.4 µm.[1]Scale bars represent 200 µm (a), 50 µm (b-c), and 100 µm (d).

Confocal microscopy was used to examine 3D structure and morphology of the protein fibers in solution. High resolution confocal scans were able to provide not only micron scale structural information on the protein fibers, but also resolved finer features

[1] Image adapted from the undergraduate thesis of Jennifer Sun, May 2014.
[2] Image courtesy of undergraduate thesis of Jennifer Sun, May 2014.

within the fibers, such as striations and folds. Equimolar mixtures of C and Q (Figure 1.67), as well as homogeneous wt, C, and Q (Figure 1.69-Figure 1.71) were studied via confocal microscopy. The abundance of fibers observed under TEM at pH 4, as well as curcumin's stability under acidic conditions, prompted us to focus on analysis of confocal results at pH 4. Confocal microscopy allowed us to characterize fiber widths. Average width of fibers produced by an equimolar mixture of C and Q at pH 4 was 19.1 ± 7.4 μm, where the width ranged from approximately 7 to 21 μm (Figure 1.67, Figure 1.68).

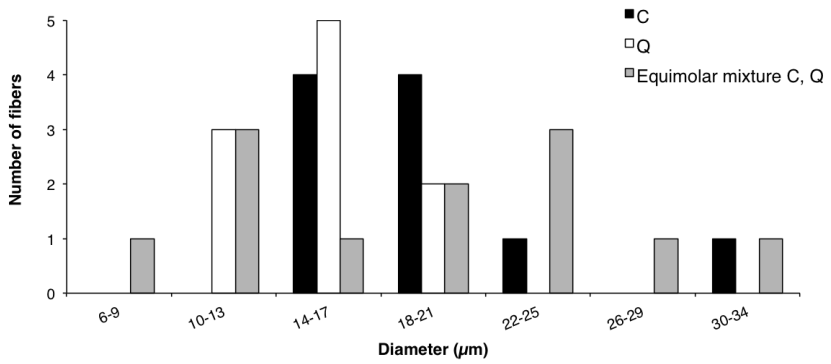



Figure 1.68. Histogram of fiber diameters in the presence of 100 μM curcumin as viewed in confocal microscopy for fibers of C (black), Q (white), and an equimolar mixture of C and Q (grey).

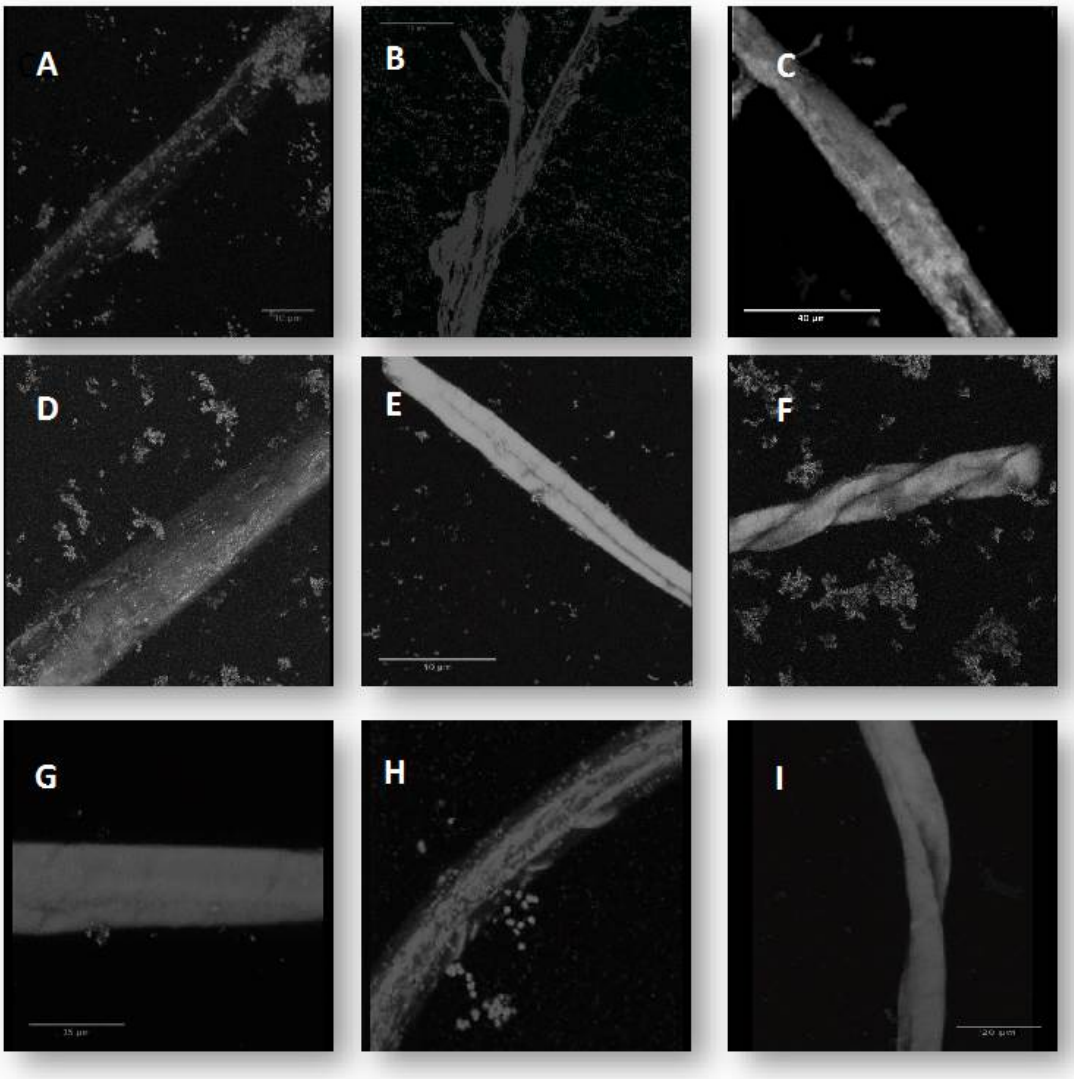



Figure 1.69. Confocal images of C (a-c), Q (d-f), and wt (g-i) with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used.[2]

Confocal micrographs of homogeneous C, Q, and wt show thick, uniform fibers (Figure 1.69). The average width of C was approximately 20.1 ± 5.2 μm ($n$ = 15) (Figure 1.69a-c), and Q was 16.0 ± 5.6 μm ($n$ = 22) (Figure 1.69d-f, Figure 1.72, Figure 1.71). Given the standard deviations, the difference in average width between the two proteins is insignificant. This dramatic increase in the fiber diameter from the absence of

[2] Image courtesy of undergraduate thesis of Jennifer Sun, May 2014.

curcumin to the presence of a 5:1 molar ratio of curcumin in solution indicates that the presence of curcumin further promotes fiber assembly.

3D reconstructions of C and Q fibers with embedded curcumin at pH 4 can be manipulated via software to view the protein fibers at all angles, by rotating the reconstruction 360°. Videos of these complete rotations of the fluorescent protein microfibers are available for viewing at http://bit.ly/1zDOSu5 (C) and http://bit.ly/1vv1Xb0 (Q).

An interesting feature of protein fiber analysis via confocal is the ability to examine different cross-sectional portions of the 3D reconstruction to be able to see "within" the protein fiber by creating orthogonal slices of the reconstructed structure. Orthogonal slices of wt fiber in 50 mM PB at pH 4 demonstrated the presence of a uniform and symmetric pore in the center of the fiber along its axis, as is shown in Figure 1.70a-c. Figure 1.70b, a circular cross section of the fiber, clearly shows the dark region in the middle, indicating either a hollow region in the center of the fiber *or* the absence of the fluorescence probe curcumin.

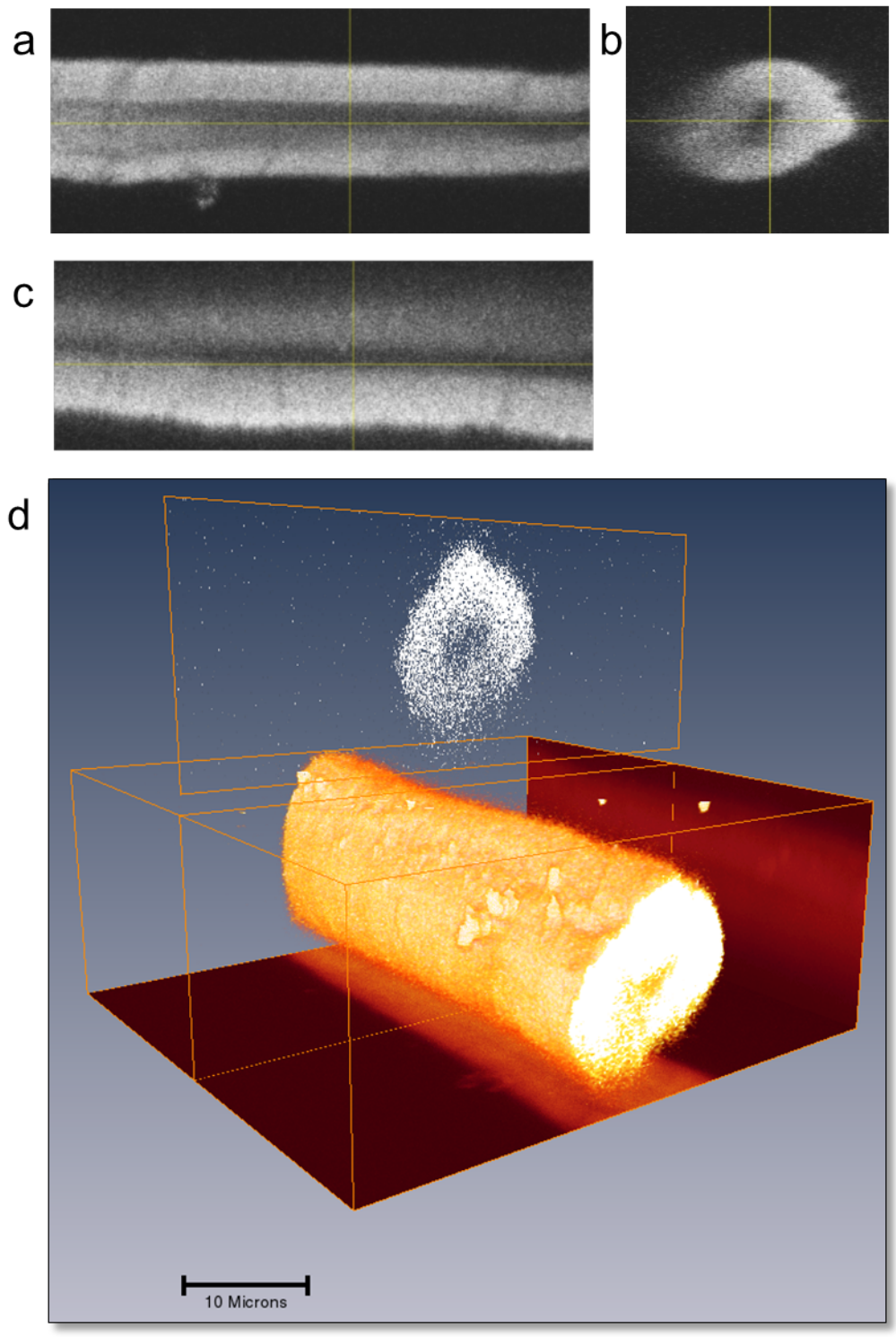



Figure 1.70. Orthogonal views of wt with embedded curcumin at 50 mM PB pH 4 at room temperature. Curcumin to protein molar ratio was 5:1 and protein concentration of 20 μM was used. Yellow crosshairs pinpoint the location of the orthogonal slices within the fiber. The presence of a hollow cavity can be seen in these orthogonal views.

Q was extensively studied in the presence of curcumin via confocal microscopy. A collection of several homogeneous Q fibers visualized by confocal are shown in Figure 1.71. As can be seen in Figure 1.71a-p, the fibers seen appeared twisted, rolled, bent, or intertwined with other fibers in various cases. In some images, fluorescent particulate matter can be seen beside the fibers (Figure 1.71b, g, j, l-o), which is most likely aggregated curcumin that was visualized in those cases because the samples were centrifuged before imaging, forcing a greater deal of the aggregates to collect at the bottom of the plate used for imaging.

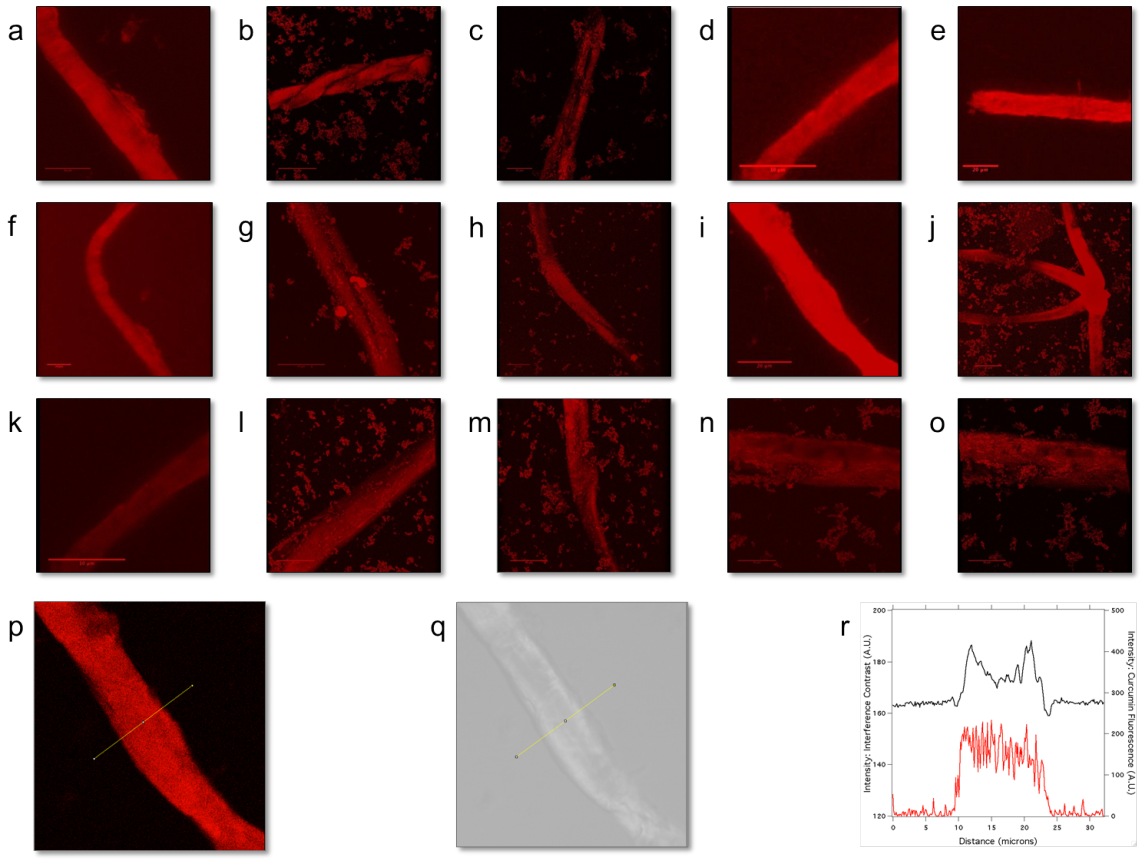



Figure 1.71. Confocal microscopy fluorescence images of several Q protein fibers in the presence of 50 µM curcumin at pH 4, 50 mM PB at room temperature. Scale bars in (d), (k), and (l) represent 10 µm, (a)-(c) and (e)-(j) represent 20 µm, and (b) and (l)-(o) represent 30 µm. Panels (p)-(r) originate from a single z-slice (119 of 227) within the XYZ data set from panel (i). The image line profiles of the confocal (red trace) and interference contrast (black trace) resulting from the yellow sampling line are shown in panel (r), indicating the boundaries of the fiber and the curcumin fluorescence are coincident.

Figure 1.72 displays an image from the 3D reconstruction of a Q fiber with embedded curcumin. Curcumin is found distributed homogeneously throughout the fiber

(Figure 1.72b). This is also demonstrated by examining Figure 1.71p and Figure 1.71q, which are the same image viewed in fluorescence and phase contrast modes, respectively. Figure 1.71r shows that interference contrast boundaries of the fibers coincide with the curcumin emission boundaries in the confocal images, where the red and black traces result from the yellow sampling line across the fibers in Figure 1.71p and Figure 1.71q, respectively. The width of that particular fiber is approximately 15 µm.

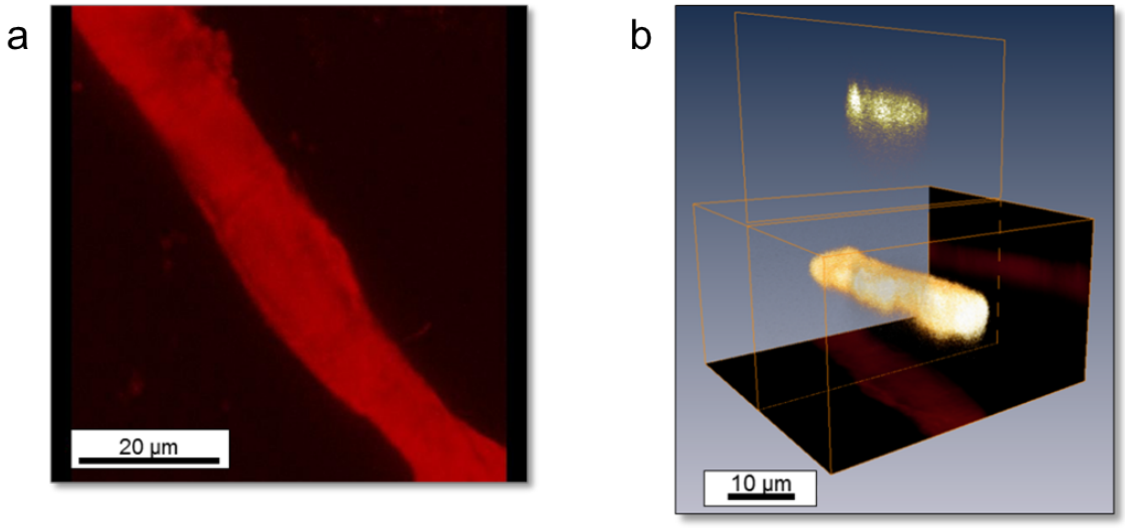



Figure 1.72. (a) Reconstruction of 3D confocal XYZ data of 10 µM Q protein in the presence of 50 µM curcumin (1:5 molar ratio of protein:curcumin) at room temperature. (b) 3D representation of the same Q fiber showing XZ and YZ orthogonal views and an oblique slice cross section above the 3D bounding box. Scale bars in (a) and (b) represent 20 and 10 µm, respectively.

TEM was also used to study aggregation of Q as a result of curcumin (Figure 1.73). Given the size of the large aggregates visualized in confocal, however, fibers of this dimension were not easily visualized via TEM as they did not easily adhere to the surface of the TEM grids and were easily wicked off when drying and preparing the samples. As a result, the fibers seen in Figure 1.73 are notably smaller (widths less than 500 nm) than the microfibers observed under confocal. The fibers that were seen in TEM were intertwined and exhibited protofibril features.

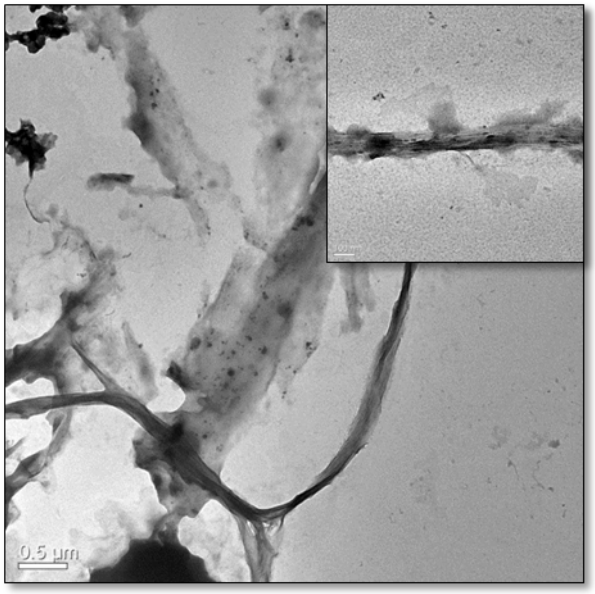



Figure 1.73. Transmission electron micrographs of 10 µM Q in the presence of 50 µM curcumin (molar ratio of 5:1 curcumin:protein) at pH 4, 50 mM PB at room temperature. Large aggregates were seen. Scale bars represent 0.5 µm in main image and 100 nm in inset.

NMR experiments confirmed peaks between 0 - 2 ppm (Figure 1.74), a region pertaining to aliphatic residues within the protein and one that rarely exhibits signals from organic molecules.[132]

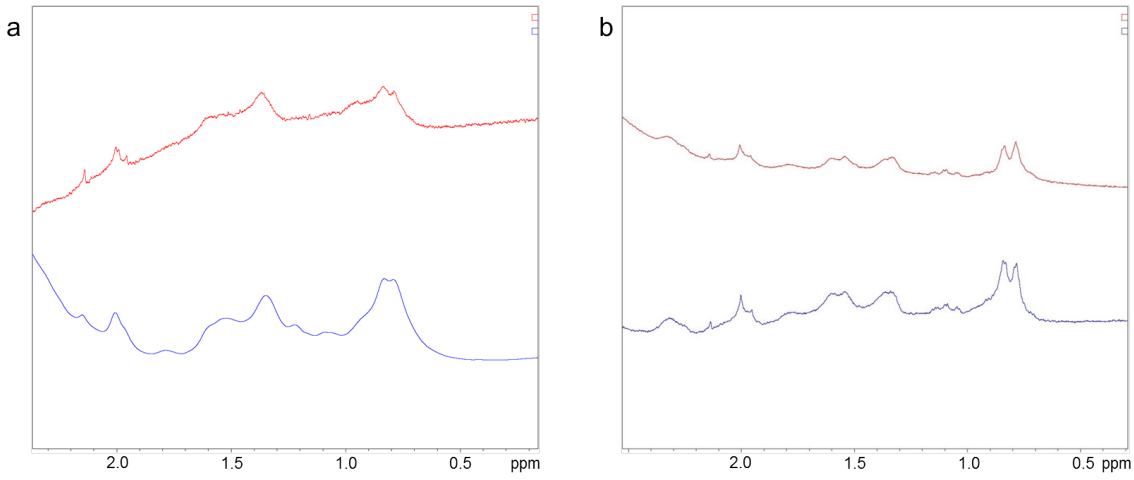



Figure 1.74. 1D $^{1}$H nuclear magnetic resonance scans for wt (a) and Q (b) in the absence (blue curves) and presence (red curves) of 100 µM curcumin (5:1 molar ratio of curcumin:protein). All protein concentrations were kept constant at 20 µM. Buffer conditions were 50 mM PB pH 4 with 1% (v/v) methanol and 1 % (v/v) $D_2O$. These protein fibers, both with and without curcumin, are not completely soluble but not crystalline and the methods commonly used to determine protein structure, e.g. nuclear magnetic resonance, provide a limited scope of data to interpret.

Peak broadening was demonstrated in the presence of curcumin in a 5:1 molar ratio of curcumin:protein, as is seen in Figure 1.74. These results suggest that curcumin is interacting with the nonpolar, hydrophobic residues located within the pore of wt and Q pentamers, however, do not preclude the possibility that curcumin may *also* be binding between the protofibrils. Fluorescence distributions from confocal measurements indicate that curcumin likely also binding between protofibrils, greatly contributing to observed aggregation effects. These protein fibers, both with and without curcumin, are not completely soluble but not crystalline and the methods commonly used to determine protein structure, e.g. nuclear magnetic resonance, provide a limited scope of data to interpret. NMR and confocal data suggest curcumin likely binds within the coiled-coil pore in addition to the surface of the pentamers, promoting supramolecular assembly (*vida infra*).

### 1.4.6.4 Aggregation studies

To further characterize the assembly and aggregation caused by the addition of curcumin, zeta potential, count rate, and absorbance at 420 nm of protein is studied (Figure 1.76, Figure 1.75).

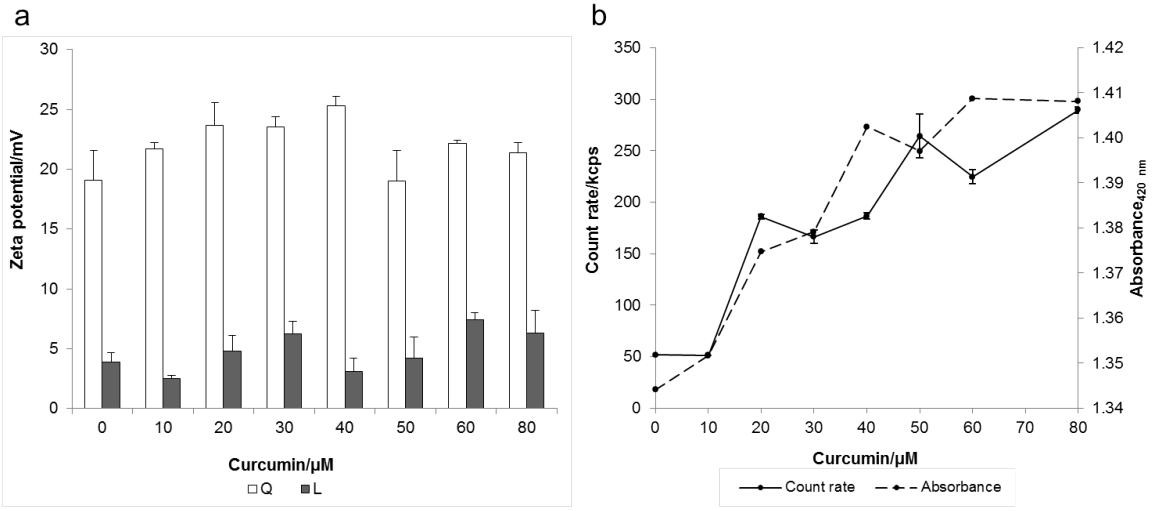



Figure 1.75. Aggregation of Q protein fibers due to increasing curcumin concentration. (a) Zeta potential as a function of curcumin concentration for Q (white bars) and L (grey bars). (b) Dynamic light scattering measurements of count rate for Q (solid line) as a function of curcumin concentration. Count rate was not obtained for L as signal at all curcumin concentrations was too low for detection. Absorbance measured at 420 nm as a function of curcumin concentration for Q (dashed line) at room temperature. Error bars in figures (a) and (b) represent an average of three trials.

With curcumin being fully protonated at pH 4, the interaction of curcumin with negative charges on the protein is likely.[79] Indeed, as curcumin concentration increases, the zeta potential exhibits a slight increase as well, indicating a neutralization of any negative charges through binding to curcumin (Figure 1.75a). For Q, zeta potential in the absence of curcumin is 19.1 ± 2.5 mV, with a steady increase up to a maximum value of 25.2 ± 0.8 mV at a curcumin concentration of 40 μM (Figure 1.75a). Above a molar concentration of 40 μM curcumin, however, zeta potential drops to values similar to Q in the absence of curcumin suggesting that neutralization is occurring by curcumin on the fiber surface. By contrast, L possesses a much lower zeta potential than Q both in the absence and presence of curcumin at all concentrations studied (Figure 1.75a).

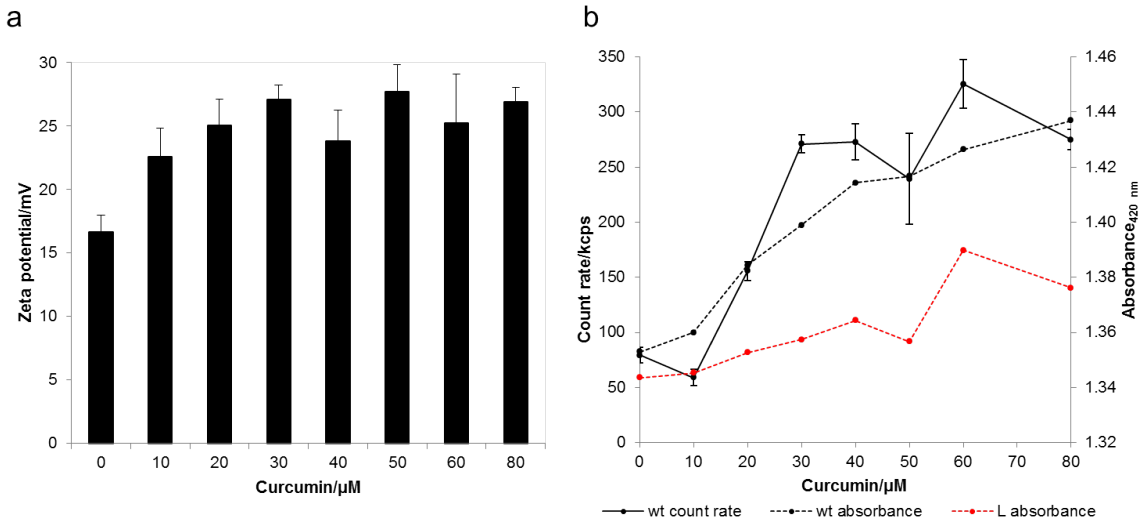



Figure 1.76. Aggregation of wt as a result of curcumin. (a) Zeta potential increases slightly as a function of increasing curcumin concentration for wt. (b) Count rate (solid line) and absorbance (dashed lines) at 420 nm of wt (black) and L (red) versus curcumin concentration. Absorbance measured at 420 nm as a function of curcumin concentration. As curcumin/protein molar ratio increases the absorbance increases correspondingly. Error bars in figures (a) and (b) represent an average of three trials.

Count rate from dynamic light scattering (DLS) is also studied to assess aggregation. For Q, a steady increase in response to higher curcumin concentrations (protein concentration held constant at 10 μM) is observed (Figure 1.75b), however count rate could not be obtained for L as the particles did not generate a high enough signal for detection, confirming the absence of aggregates. Absorbance of proteins at 420 nm in the presence of curcumin display a similar increasing trend (Figure 1.75b, Figure 1.76b).

While the zeta potential, count rate, and absorbance data values for Q are similar to those exhibited by wt on a macromolecular level (Figure 1.76), analysis of Coulombic surface charge shows that on a molecular level the two protein assemblies have *very different* charge distributions.

## 1.5 Conclusions

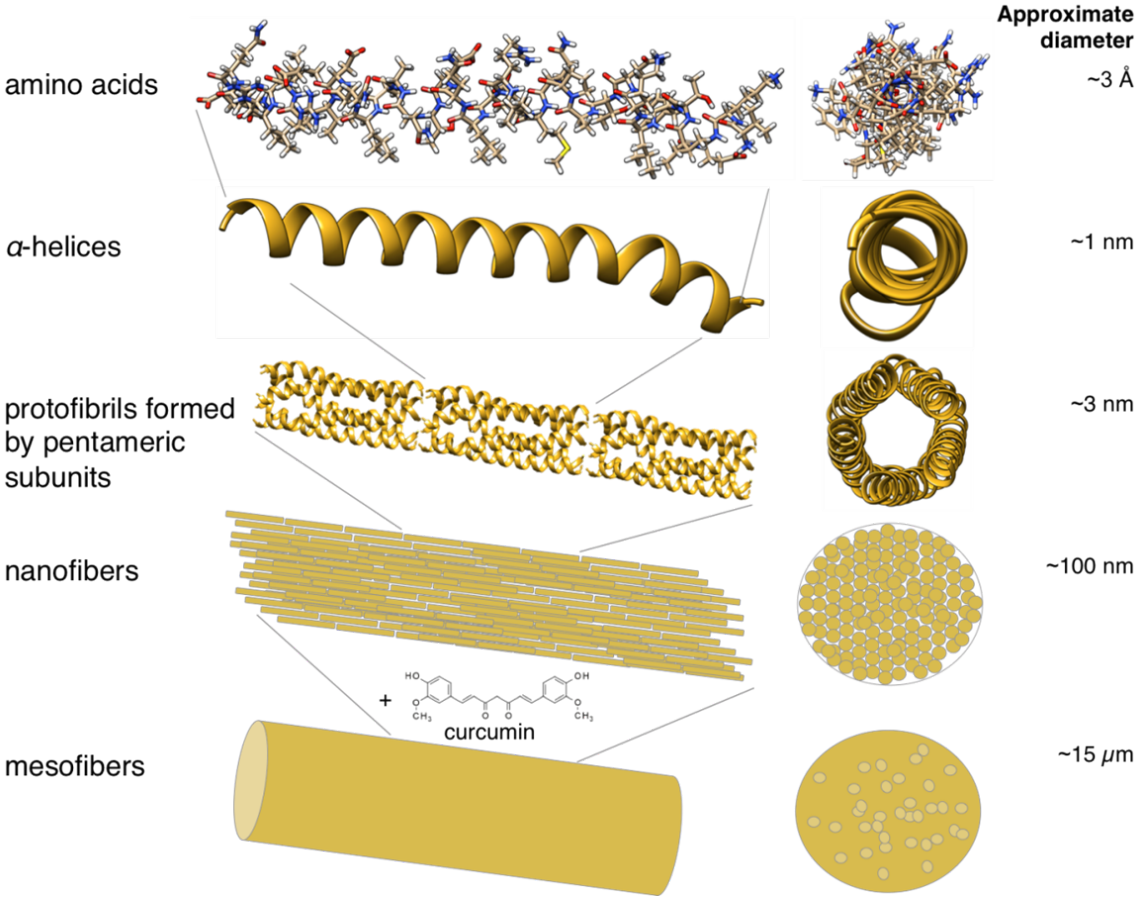



Figure 1.77. Schematic for fiber assembly. Schematic representation of fiber assembly of Q on various length scales. Amino acids form α-helices which assemble to form pentamers with approximate diameters of 3 nm. Protofibrils are formed by assembly of pentameric subunits and bundle together to create protein fibers ranging from tens to hundreds of nm in diameter, and upon addition of curcumin mesofibers with diameters on the micrometer scale are generated.

Aggregation of protein fibers has been thoroughly studied in the case of *β*-sheet amyloid fibers, as their aggregation is significant in the onset of Alzheimer's disease.[17] We have engineered proteins that can hierarchically assemble into mesofibers with the assistance of small molecules through encapsulation and aggregation (Figure 1.77). While studies

have been performed on controlling aggregation of coiled-coil protein fibers through macromolecular crowding[58] and adjusting solvent conditions such as pH[133] and ionic strength,[41] save a few exceptions,[45,79,80] comparatively little work has been done on inducing aggregation of $\alpha$-helical protein fibers through additives. In such examples, change on the order of nanometer to micrometer scale of *de novo* designed $\alpha$-helical protein fibers has not been previously reported. Additionally, the ability to generate fibers possessing nanoscale organization with concentrations on the micromolar range provides a tremendous technical advantage in future applications as biomaterials. Herein, we have demonstrated the synthesis of protein fibers of dimensions comparable to $\alpha$-keratin, collagen, and spider silk[130] with the added functionality of small molecule binding.[134,135] These fibers hold tremendous potential for tissue engineering and delivery of curcumin and other therapeutic small molecules.

# 2 Tunable Conformation-Dependent Protein·Gold Nanoparticle Nanocomposites

The text in this chapter is partially from the paper:

Jasmin Hume, R. Chen, R. Jacquet, M. Yang, J. K. Montclare. Tunable Conformation-Dependent Engineered Protein-Gold Nanoparticle Nanocomposites. *Biomacromolecules.* DOI: 10.1021/acs.biomac.5b00098. Publication Date (Web): April 13, 2015.

The proof of the print version of this paper can be found in Appendix 5.1.2.

## 2.1 Abstract

We demonstrate the fabrication of protein·gold nanoparticle (AuNP) nanocomposites *in situ*, leading to very distinct assemblies and properties that are dependent upon the protein secondary structure. In the presence of pentameric coiled-coil proteins C and Q, which contain a histidine tag and have helical contents of 54 and 65 %, templation of AuNP results in precipitation of the protein·AuNP composites with AuNPs approximately 6.5 nm in diameter, creating macromolecular assemblies on the micrometer scale. In the absence of the histidine tags the resulting Cx and Qx proteins, which exhibit lower helical contents of 37 and 45 %, respectively, stabilize AuNPs approximately 4.5 nm in diameter leading to soluble protein·AuNP composites for several days without aggregating. By manipulating protein structure via external triggers, such as TFE, we obtain control over the macromolecular conformation and overall physicochemical properties. These hybrid protein·AuNP assemblies can be readily deposited on electrodes, where they can serve as a tunable bio-nanocomposite kinetic barrier. Proteins that template and stabilize nanoparticles can serve as agents for biodetection and imaging while the formation of nanoparticle-embedded protein films have applications in catalysis and biosensing.

## 2.2 Introduction

### *2.2.1 Gold nanomaterials*

Noble metals exhibit highly desirable material properties such as plasmon absorption resonances, conductivity, and catalytic activity. The ability to tune these properties lies in the material dimensions and features; decrease in dimension increases the capacity for customization and adaptation for specific application.[136] Gold nanomaterials are of great interest due to their superior optical and electronic properties; both spherical and rod-like Au nanoparticles (AuNPs) exhibit strong surface plasmon resonance (SPR). As shown in Figure 2.1a, nanospheres demonstrate peak absorption at lower wavelengths, while Au nanorods absorb in the near infrared region.[136,137] Absorption is dependent on both NP shape and size. AuNPs can be easily synthesized through chemical reduction, where well-defined techniques allow facile methods for controlling NP size as well as shape. Change in buffer conditions can result in AuNPs of various size distributions that correspond to different optical properties, as determined by TEM in UV-vis spectroscopy in Figure 2.1b-d. Increasing diameter of AuNPs corresponds with an increase in the peak absorption wavelength.[138] Templation of AuNPs to proteins has been examined by Slocik *et al.*, where UV-vis spectroscopy results indicate that biomimetically synthesized gold nanoparticles demonstrate a peak in absorption centered around 545 nm[139,140]

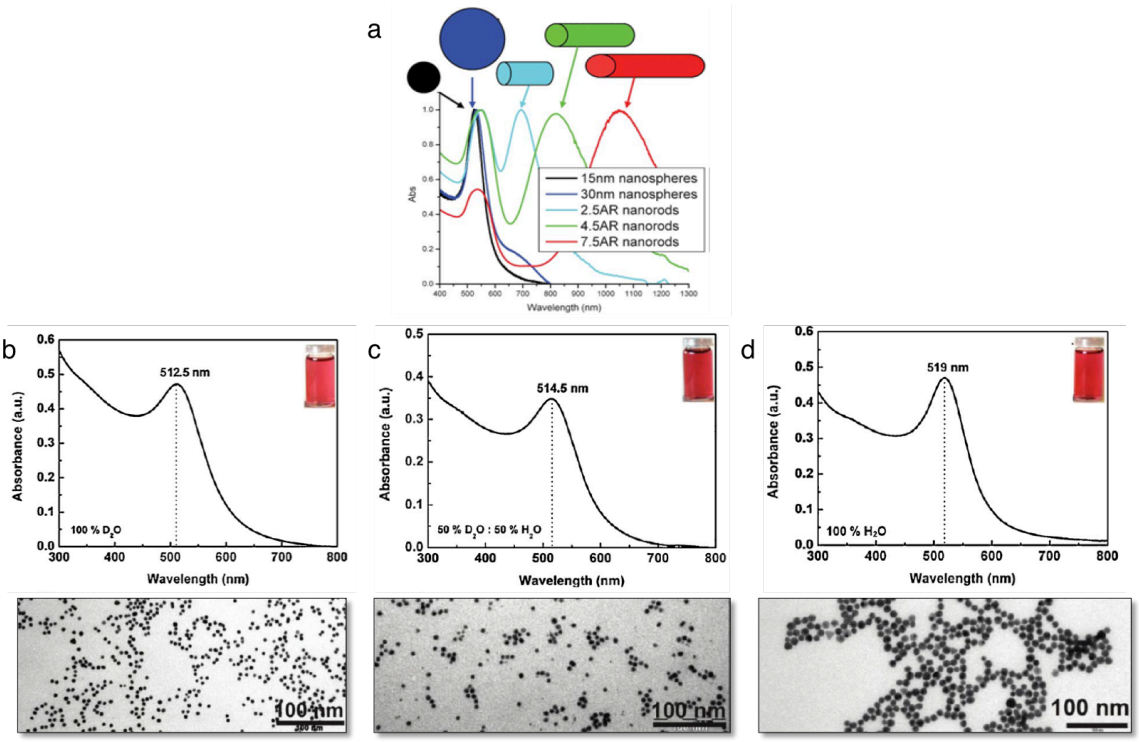



Figure 2.1. UV-vis absorption of gold nanoparticles of various sizes and shapes (a).[137] UV-vis and TEM characterization of AuNPs synthesized chemically in various solvents: 100 % $D_2O$ (b), 50 % $D_2O/H_2O$ (c), and 100 % $H_2O$ (d).[138]

Nanoparticles in general are of great interest as reaction catalysts (either in solution or immobilized), as their reduced dimensions offer maximized surface areas increasing catalytic rates when compared to the bulk material.[137] There are additional advantages associated with properties of noble metals. Noble metals, including gold, exhibit surface plasmon oscillation, due to collective oscillation of electrons in the conduction band of the material. SPR, and hence energy transfer, is dictated by nanocrystal arrangement.[141] Nanocrystals generate electromagnetic fields affecting their local environment, which can enhance fluorescence of the metal itself, the Raman signal of a molecule on the surface of the metal, and the scattering of light. These physical properties exhibited by gold and other noble metals have led to numerous applications in sensing and imaging techniques, as is subsequently discussed in Section 2.2.5. Of particular interest are applications involving optical properties of AuNPs permitting detection of molecular binding and fluorescence emission studies of AuNP energy transfer pairs used to measure distance and interactions between many types of

molecules.[137] Hybrid organic-inorganic nanocomposites are an emerging area with great potential to create advanced materials, many of their interactions designed based on natural examples of biomineralization.

### *2.2.2 Biomineralization*

#### 2.2.2.1 Natural systems

The number and variety of organisms that are able to template inorganic particles is substantial. Organisms incorporate inorganic materials from their surroundings into their anatomy through biomineralization; the process by which organisms form materials from inorganic solids and biomolecules.[142] Biomineralization is widespread in nature and is performed by a very diverse range of organisms. Some of the simplest of organisms, such as diatoms, produce intricate silica shells with nanoscale structures, as well as algae and plants that produce metal nanoparticles as a byproduct of detoxification functions.[143] Highly complex organisms, such as mammals, biomineralize compounds such as calcium in the bones and teeth, and magnetite in the tissues of the human brain.[62] A familiar example of biomineralization by humans, and one that can be detrimental to our health is the formation of kidney stones, which are microcrystals of calcium oxalate monohydrate.[144]

Many microorganisms are capable of metal templation, as these organisms are essential for metal cycling and mineral formation on the Earth's surface. Metallophilic bacteria thrive in environments rich in heavy minerals, and have evolved methods of detoxification and metal particle isolation within the cell to reduce toxic effects of oxidative stress caused by the metals. *Cupriavidus metallidurans*, a metallophilic bacteria, is capable of metabolizing Au(III) from its surroundings to generate metallic Au nanoparticles (see Figure 2.2).[145] Reduction of Au in *C. metallidurans* to is attributed to cellular processes governed by enzymes.[145] The ability of naturally occurring enzymes and peptides to reduce Au has inspired researchers to exploit the natural ability of these biological molecules to generate AuNPs for electronic applications.

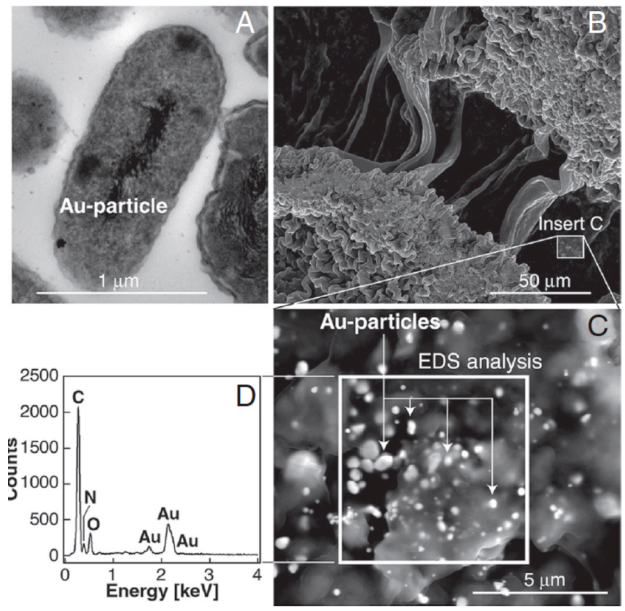



Figure 2.2. Biomineralization of Au in *C. metallidurans*. TEM micrograph of a *C. metallidurans* cell containing a AuNP in the periplasmic space (a), SEM micrographs (b, c) and corresponding EDAX analysis (d) of biofilm containing mineralized Au.[145]

In addition to biomineralization of metal atoms to impart electronic properties in natural systems, examples of organisms that possess near-metallic conductivities attributed to unique biomacromolecules are also of interest in the study of electronic biomaterials. The discovery of bacterial nanowires from the *Geobacter sulfurreducens* species that are able to transport electrons across great distances with respect to the length of the bacteria has recently been reported.[146] This bacteria naturally produces nanowires composed of pilin nanofilaments and has been used as a biofilm coating in an effort to generate novel materials that are capable of achieving the same level of conduction as synthetic organic metallic nanostructures.[146]

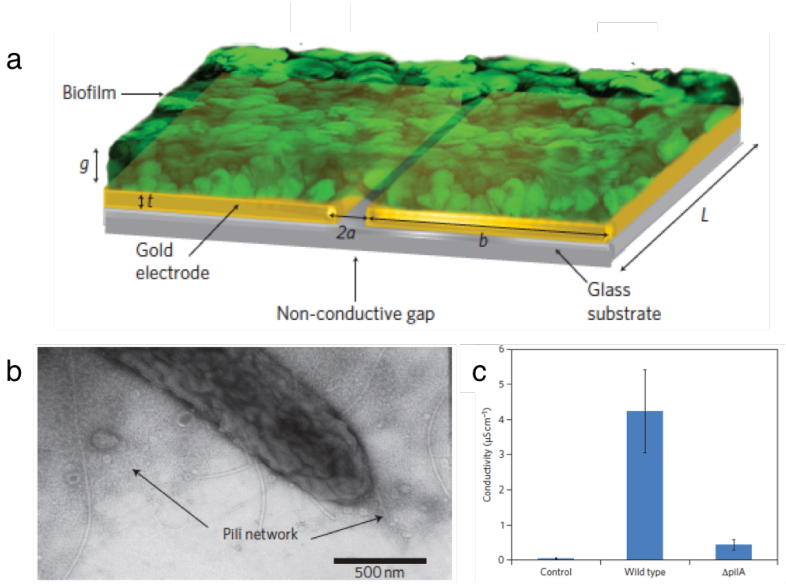



Figure 2.3. Setup for measuring *in situ* conductivity of a biofilm composed of *Geobacter sulfurreducens*, grown over two gold electrodes separated by a non-conductive gap (a). TEM micrograph of biofilm and pili nanofilaments grown on the electrode surface (b) and the conductivity of the wild type biofilm compared to a control of buffer and PilA-deficient mutant strain (c).[146]

The molecular organization of the pili filaments is not yet known, due to complications in the crystallization of pili filaments. The authors hypothesize that the metallic-like conductivity they observed is due to intermolecular electron delocalization within the pili filaments.[146] These biologically based systems are an alternative to expensive, difficult to produce conventional electronic materials.

#### 2.2.2.2 Synthetic systems

The ability of natural systems to template metal nanoparticles from ionic compounds is often used as inspiration for biomimicry in synthetic designed systems. For example, Au nanowires are currently being fabricated for use as energy transfer systems.[147–149] Notable advances have been made through routes of organic synthesis[147] and wet chemical strategies.[148] Ultrathin Au nanowires (1.6 nm diameter and ~ 4 μm long) have been produced by simply mixing $HAuCl_4$ and oleylamine.[147] Nanowires have also been created through deposition of AuNPs on polymer nanofibers for nanodevice fabrication, where electrically conductive poly(allylamine hydrochloride) is electrospun and gold is either directly spun with the polymer or deposited via UV photoirradiation.[150] Although these methods are indeed successful in their generation of nanowires, traditional wet chemical

synthesis is limited in the amount of flexibility the engineer has in the nanostructural design of these materials. These limitations have spurred research on templating AuNPs using bioconjugates.

Research on generating stable, multi-purpose AuNP assemblies using bioconjugates has generated much interest, especially in the area of creating biosensors of superior sensitivity and selectivity.[151–153] Some of the bioconjugate materials of interest in creating these compound materials include oligonucleotides, enzymes, viruses, lipids, and peptides.[154] Phospholipid tubules have been used as a template material for creating tubules decorated with nickel and subsequently magnetically aligned, as is seen in Figure 2.4a.[155] DNA is also an excellent choice as a scaffold material for the construction of inorganic-organic hybrid materials, as it possesses excellent recognition capabilities, excellent programmable specificity, mechanical rigidity, and high precision processing. Conductive DNA wires have been synthesized with silver nanoparticles (Figure 2.4b, c) with a resistivity of 86 Ω, however they are unstable at pH values above 10 and temperatures greater than 60 ºC.[156]

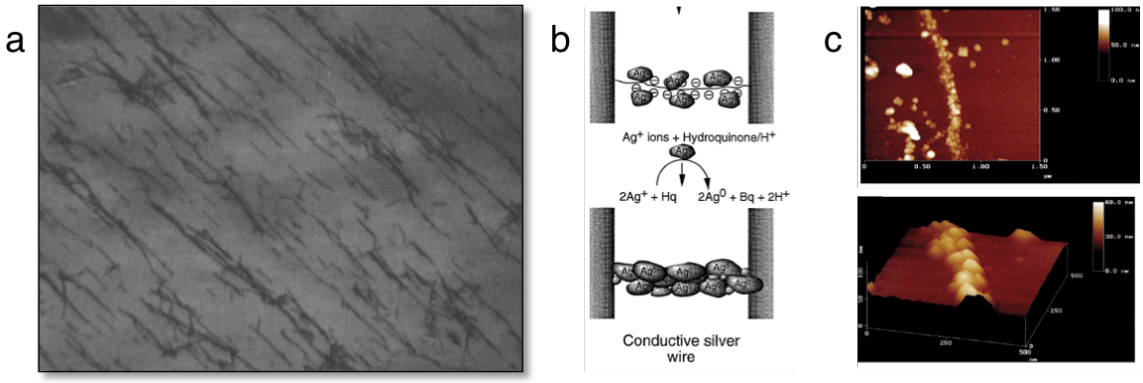



Figure 2.4. Phospholipid tubules plated with nickel and aligned with a bar magnet (magnification of 120x) (a).[155] Conductive DNA nanowires constructed with AgNPs shown schematically between two gold electrodes (b) and AFM images of the nanowires (c).[156]

Peptides provide a particularly interesting advantage in that they include self-assembling materials that can be tailored to generate a wide variety of nanoscale materials with high reproducibility.[157] The conjugation and stabilization of AuNPs with proteins takes advantage of the incredible physical properties of AuNPs while adding biological functionalization leading to protein·AuNP assemblies.[13,22,139,140,158,159]

Table 2.1. AuNP templation by select proteins.

| Protein | | | AuNP | | |
|---|---|---|---|---|---|
| **Name** | **Sequence** | **Secondary structure** | **Precursor materials** | **Size (nm)** | **Ref.** |
| $NM^{K184C}$ | Gln/Asn-rich N-terminal of Sup35 (1-253), with K184 mutated to C | *β*-sheet | Monomaleimido Nanogold® | 1.4 | 149 |
| Hub-$His_6$ | C-terminal (1074–1675) fragment of clathrin heavy chain with N-terminal hexahistidine tag | Structured, self-assembles into cage-like structures | $(CH_3)_3PAuCl$ in $H_2O$, reduction via HEPES buffer, pH 6.5 | ~12 nm, oblong with aspect ratio of 3.5 | 13 |
| Histidine rich peptide | AHHAHHAAD | NA | $ClAuPMe_3$ reduced with 2 M hydrazine hydrate | 6 | 160 |
| A3 | AYSSGAPPMPPF | NA | $HAuCl_4$, no additional reducing agent | 12.8 ± 2.9 | 161,162 |
| Flg-A3 | DYKDDDDKPAYSSGAPPMPPF | Unstructured | $HAuCl_4$, no additional reducing agent or premade AuNPs. Interacts via non-covalent interactions | 11.4 ± 1.2 ($HAuCl_4$), 10 (premade AuNPs) | |
| E5 | CGGEVSALEKEVSALEKEVSALE KEVSALEKE VSALEK | *α*-helical | Premade AuNPs. Interacts via N-terminal cysteine thiol linkage | 10 | |
| HPI and SAS S-layers | HPI is 6 identical monomers, 98 kDa each from *D. radiodurans*, SAS is dimers geometrically arranged into triads surrounding a central pore from *S. acidocaldarius* | Structured, p6 rotational symmetry (HPI) or p3 symmetry (SAS) | Citrate-stabilized, pre-fabricated AuNPs | 5.6 ± 0.3 | 163 |
| *α*-amylase | See Appendix 5.5.3 | Structured<br>Unstructured | Citrate stabilized AuNPs, pre-fabricated by reduction of $HAuCl_4$ with 0.5 M trisodium citrate 2-hydrate | ~10 nm<br>~20 nm, high size dispersity | 164 |
| *β*-sheet silk films | See Appendix 5.5.4 | Structured, *β*-sheet<br>Unstructured | 1 mM $HAuCl_4$ at room temperature in 0.1 M borate buffer, pH 10.5. No further reducing agent. | 6.9 ± 1.1<br>17 ± 5 | 165 |

(Table 2.1, continued).

| Protein | | | AuNP | | |
|---|---|---|---|---|---|
| **Name** | **Sequence** | **Secondary structure** | **Precursor materials** | **Size (nm)** | **Ref.** |
| $C_{12}$-$PEP_{Au}$ | Succinimide-activated dodecanoic acid to the N-terminus of AYSSGAPPMPPF | Structured, *β*-sheet | $HAuCl_4$ reduced with HEPES buffer | 8.2 ± 1.0 | 158,166 |
| BP124NAB | See Appendix 5.5.5 | Structured, *α*-helical | $HAuCl_4$ or $NaAuCl_4$, no additional reducing agent | 5 | 167 |
| Gelator 1 with 2a or 2b | See Appendix 5.5.6 | Structured, *α*-helical | Tetrahydrothiophene added to a solution of $HAuCl_4 \cdot 3H_2O$, UV photoreduction | Between 2 – 200 nm, depending on duration of UV irradiation | 168 |
| poly(*γ*-methyl L-glutamate-co-L-glutamic acid) (MG/GA) | 28 mol % glutamic acid residues, number-average MW of 73 kDa | Structured, *α*-helical in TFE | 0.2 mM $HAuCl_4$ mixed with alkanethiols at a molar ratio of 6, reduction with $NaBH_4$ | 2.61 | 169 |
| | | Unstructured in water | | 5.86 | |
| Cty c | See Appendix 5.5.7. apoCty c is obtained by removing the heme of Cty c | Structured, *α*-helical | Pre-fabricated Au-TAsp NPs featuring thiol ligands, a tetra(ethylene glycol) segment, and a dianionic aspartate | Increased from 10.2 – 15.1 with increasing ratio of protein:AuNP | 170 |
| apoCyt c | | Unstructured | | Ranges from 10.6 – 1010 | |
| Peptide 1 | Acidic leucine zipper-like peptide, glutamic acid occupying 75 % of heptad positions *e* and *g* | Structured, *α*-helical | Premade, thiol bond with a terminal cysteine residue present on the peptides | 8.5 | 171 |
| Peptide 2 | Basic leucine zipper-like peptide, lysine occupying 75 % of heptad positions *e* and *g* | Structured, *α*-helical | | 53 | |

These systems utilize the fact that certain amino acids are capable of binding Au as a result of hydrophobic interactions or hydrogen bonding, or even covalently through thiol bonds with cysteine residues. Conducting nanowires have been constructed with self-assembling amyloid protein fibers as a scaffold (Table 2.1).[149] In this work by Scheibel *et al.* functionalized monomaleimido Nanogold® AuNPs were covalently cross-

linked with protein fibers from a prion determinant of *Saccharomyces cerevisiae*. The N-terminal and middle region (referred to as "NM") of the Sup35p protein forms *β*-sheet-rich amyloid fibers with diameters of 9–11 nm where the lysine in position 184 was mutated to a cysteine residue (dubbed $NM^{K184C}$).[149] These fibers exhibited a remarkable stability that allowed them to maintain structure upon metallization, which included a gold toning procedure of silver enhancement and gold enhancement after the AuNPs had been bound (Figure 2.5a). The biotemplated wires (Figure 2.5b, c) demonstrated conductive properties on par with a solid metal wire, characterized by low resistance and ohmic behavior.[149]

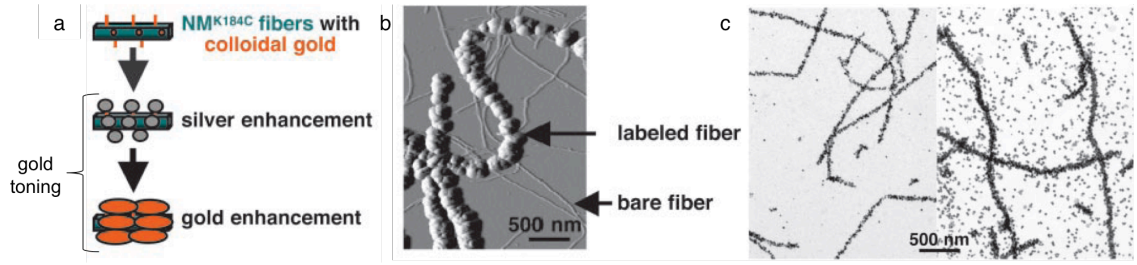



Figure 2.5. Procedure of gold toning (a) used to obtain conductive amyloid fibers deposited on $Si_3N_4$ viewed by AFM (b) and TEM (c), where diameters of coated fibers were enlarged from 50 nm to 100 nm by increasing the silver enhancement time from 3 minutes (c, left) to 5 minutes (c, right).[149]

Another residue that has shown the ability to template AuNPs is histidine, which has a high affinity for metal ions as a result of the nitrogen contained in its imidazole ring. Djalali *et al.* have successfully synthesized uniform size distribution AuNPs 6 nm in diameter on an immobilized histidine-rich peptide (sequence AHHAHHAAD, Figure 2.6a, Table 2.1) resulting in a nanotube of Au particles (Figure 2.6a, b).[160] Slocik *et al.* have used several peptides in the synthesis of 10 – 20 nm AuNPs, with sequences AYSSGAPPMPPF (termed A3) and DYKDDDDKPAYSSGAPPMPPF (termed Flg-A3), among others (Table 2.1). Both A3 and Flg-A3 are capable of reducing AuNPs upon the introduction of $HAuCl_4$ solution within 5 minutes without an additional reducing agent (Figure 2.6c, d). Slocik *et al.* evaluated the ability of a set of peptides to synthesize particles based on their amino acid sequences, underscoring the exploitation of peptides as an alternative to chemical methods.[161]

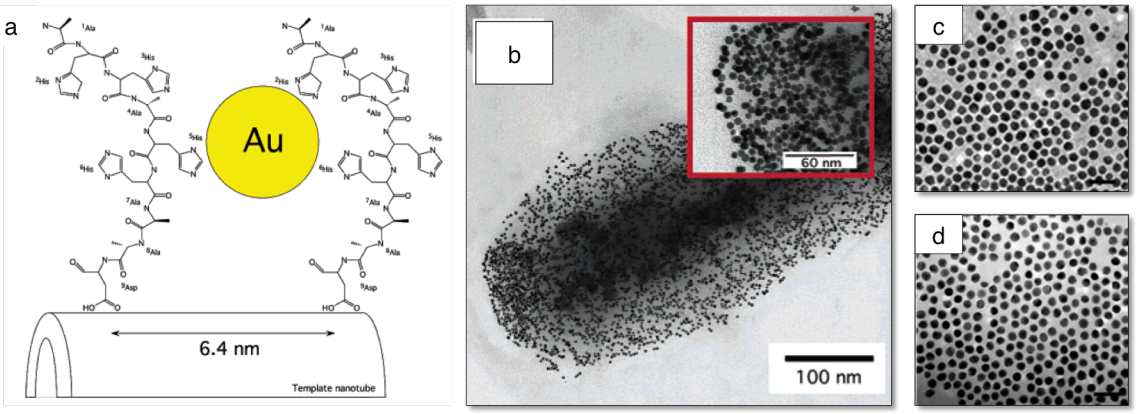



Figure 2.6. Schematic of the peptide·AuNP nanotube complex (a) and TEM images of AuNPs templated outside and inside the nanotube at pH 8.0 (b).[160] AuNPs templated with peptides A3 (c) and Flg-A3 (d), where scale bars indicate 50 nm.[161]

Not only can functional protein nanomaterials be created as wires or nanotube configuations, but two dimensional protein layers have also been used for templation of AuNPs.[172] S-layers, naturally occurring two dimensional crystalline protein sheets found in cell membranes, are able to self-assemble into arrays of oblique, square, or hexagonal symmetry. These protein arrays have been used as scaffolds for nanoparticle templation, where Mark *et al.* were able to demonstrate templation of citrate-stabilized AuNPs in a site-specific manner on S-layers via electrostatic, van der Waals, and/or non-covalent hydrogen bonding interactions (Table 2.1).[163] The bionanofabrication used to generate these materials is suitable for bottom-up approach for creating active nanoelectronic devices with protein scaffolds.

### *2.2.3 Impact of protein structure on assembly and templation*

Proteins, compared with other biomacromolecules that may be used to template metal nanoparticles, have evolved to succumb to several external triggers that modify the protein structure, and hence assembly. This extreme environmental sensitivity demonstrated by proteins is very attractive when designing composite materials with physical properties dictated by structure, and hence underlying chemistry. Protein secondary structure prescribes tertiary, as well as quaternary arrangement, and can therefore be provoked in controlled manners to govern self-assembly of functionalized protein-based materials.

Although previous work on protein·AuNP assemblies has focused predominantly on templation of AuNPs by structured protein,[158,162,164–168,171] there are also examples of templation by unstructured protein.[162,164,169,170] Distribution of hydrophobic and hydrophilic groups in native and unstructured proteins has been shown to have a significant impact on dimensions and spacing of templated NPs,[158,168] to the extent that protein secondary structure can be ascertained by probing of chiroptic properties of AuNP nanocomposites.[164,173] As secondary structure plays an important role in NP templation, it follows that macromolecular assembly of AuNP nanocomposites is often also dictated by peptide structure; templation by unstructured peptides frequently results in soluble AuNPs which do not aggregate whereas templation by structured protein allows for self-assembly, causing protein·AuNP nanocomposites to aggregate.[164,169] This dependence on structure is seen for *α*-helical peptide motifs[164,167–171] as well as *β*-sheets.[158,165,166]

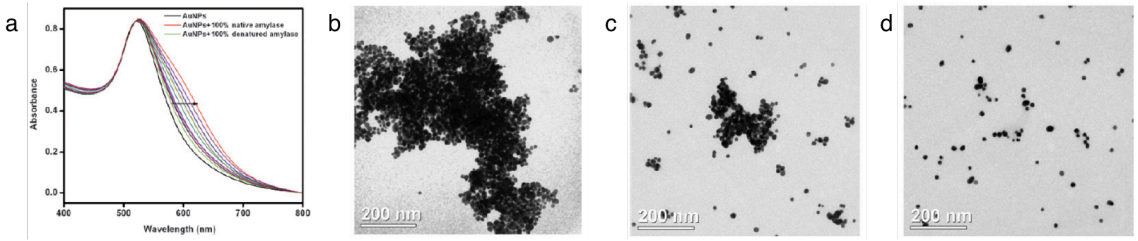



Figure 2.7. UV-vis spectra of citrate-stabilized AuNPs in the presence of varying volume fractions of native vs. denatured *α*-amylase (a). TEM of AuNPs with 100 % structured *α*-amylase (b), 50 % structured/denatured *α*-amylase (c), and 100 % denatured *α*-amylase (d).[164]

The optical properties of AuNPs have been exploited by Deka *et al.* to use these nanoparticles as a probe with which they measure the secondary structure of *α*-amylase, bovine serum albumin (BSA) (Appendix 5.5.8), and amyloglucosidase proteins (Table 2.1).[164] Mixtures of native and denatured proteins in the presence of citrate-stabilized AuNPs resulted in broadening of the absorption spectra (Figure 2.7a) corresponding to the fractional content of the conformations. Structured *α*-amylase was shown to cause AuNPs to aggregate (Figure 2.7b), whereas denatured *α*-amylase in the presence of AuNPs resulted in disperse NPs (Figure 2.7d).[164]

Protein secondary structure has also been seen to influence AuNP diameter. Formation of AuNPs from $HAuCl_4$ and subsequent chemical reduction onto *β*-sheet silk

films by Kharlampieva *et al.* demonstrated a change in AuNP diameter dependent upon secondary structure of the protein films; larger diameter AuNPs were formed by unstructured proteins, whereas smaller, more monodisperse AuNPs were obtained in the presence of structured *β*-sheet protein (Table 2.1).[165] This work suggested that that the formation of smaller, more disperse AuNPs was facilitated by ordered tyrosine residues in silk II and preventing the aggregation of AuNPs.[165] Smaller diameters of AuNPs result as crystallization of gold can be capped due to surrounding proteins that inhibit nanoparticle growth.[159,174]

#### 2.2.3.1 Protein influence on gold nanoparticle size

Synthesis of AuNPs by various means has been thoroughly characterized, to the point where nanoparticle engineers are able to chemically synthesize NPs to control both their shape and size.[137,154,175] Control on the nanometer scale is very important as plasmonic effects of AuNPs are highly defined by their size, where spherical AuNPs with particle diameters from five to a few tens of nanometers are red in color, and become increasingly purple/blue at larger AuNP sizes.[176] The aggregation of small AuNPs also shifts the color from red to blue.[176]

Some of the chemical techniques for producing monodisperse AuNPs include citrate reduction, microemulsions of surfactants, polyelectrolytes, and reverse micelles, or via the Brust-Schiffrin method of stabilization of the AuNPs via thiol bonds.[154] The Brust-Schiffrin method, first published in 1994, has been used to synthesize AuNPs that are thermally stable and air-stable of controlled size ranging between 1.5 nm and 5.2 nm in diameter.[154] In this type of synthesis the size of the AuNPs is controlled by the molar ratio of thiol to gold, where larger thiol/gold molar ratios lead to smaller core sizes of the AuNPs.[154] In addition to the AuNP size control offered by the Brust-Schiffrin method employing thiol groups, there has been a tremendous amount of research on AuNP templation utilizing proteins to impact NP size.[176–180]

a

| Size (diameter) (nm) | Surface area ($nm^2$) | $K_{sv} \times 10^{-8}$ ($dm^3\ mol^{-1}$) |
|---|---|---|
| 8 | 201 | 5.95 |
| 10 | 314 | 5.79 |
| 16 | 804 | 5.13 |
| 25 | 1963 | 4.37 |
| 34 | 3631 | 3.60 |
| 41 | 5280 | 2.90 |
| 47 | 6939 | 2.63 |
| 55 | 9501 | 2.39 |
| 70 | 15,394 | 2.09 |

b

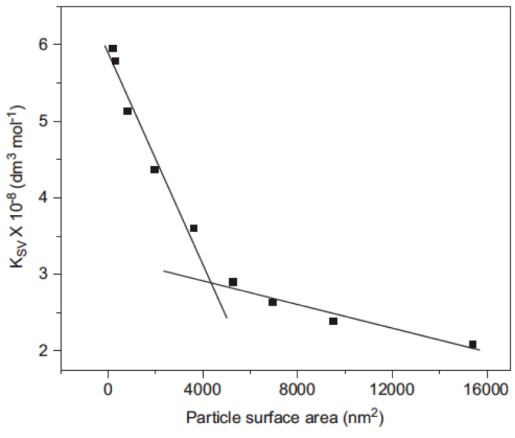



Figure 2.8. Stern-Volmer quenching constant ($K_{SV}$) of BSA by AuNPs of different sizes (a) and $K_{SV}$ as a function of AuNP surface area (b).[177]

Clusters of nanometer-sized gold particles show interesting size-dependent chemical properties that are conveniently monitored and characterized by fluorescence and absorption spectroscopy, allowing engineers to probe and modify biological systems using AuNPs templated with biomolecules.[137,177] Pramanik *et al.* have studied the interaction of BSA with AuNPs of different sizes, ranging from 8 – 70 nm in diameter.[177] When the surface area of the AuNPs is plotted against the Stern-Volmer quenching constant ($K_{SV}$) it was determined that smaller AuNPs quench BSA fluorescence to a much greater extent than larger AuNPs (Figure 2.8).[177] As AuNP diameters decrease the ratio of surface area to volume increases, increasing effects due to the surface energy and charge density on the AuNP surface.[177] Smaller particles that have a higher relative amount of surface area more efficiently quench molecular fluorescence of BSA.[177] The difference in quenching efficiency is defined by the two slopes seen in Figure 2.8b, with a transition in quenching efficiency observed at particles 30 nm in diameter.[177]

There is much evidence that certain amino acid residues have a significant effect on controlling AuNP size and shape, including cysteine, tyrosine, and histidine.[160,161,180,181] Viral capsids have also been used in AuNP templation, where active

tyrosine residues on the capsid surface reduce $AuCl_{4-}$, forming AuNPs on the surface.[180] Reduction of AuNPs in this system is enhanced by incorporating more tyrosine sites within the capsid cavity or on the surface.[180] Cowpea mosaic virus (CPMV) has been modified to include several cysteine residues on the surface of the viral capsid to control spacing of AuNPs templated on the surface.[176] Slocik *et al.* have also shown that varying the ratio of tyrosine to histidine can enhance AuNP templation.[161] Control of size and shape of AuNPs has also been achieved by adjusting the ratio of gold salt and cell-free filtrate of the fungus *Sclerotium rolfsii*.[180]

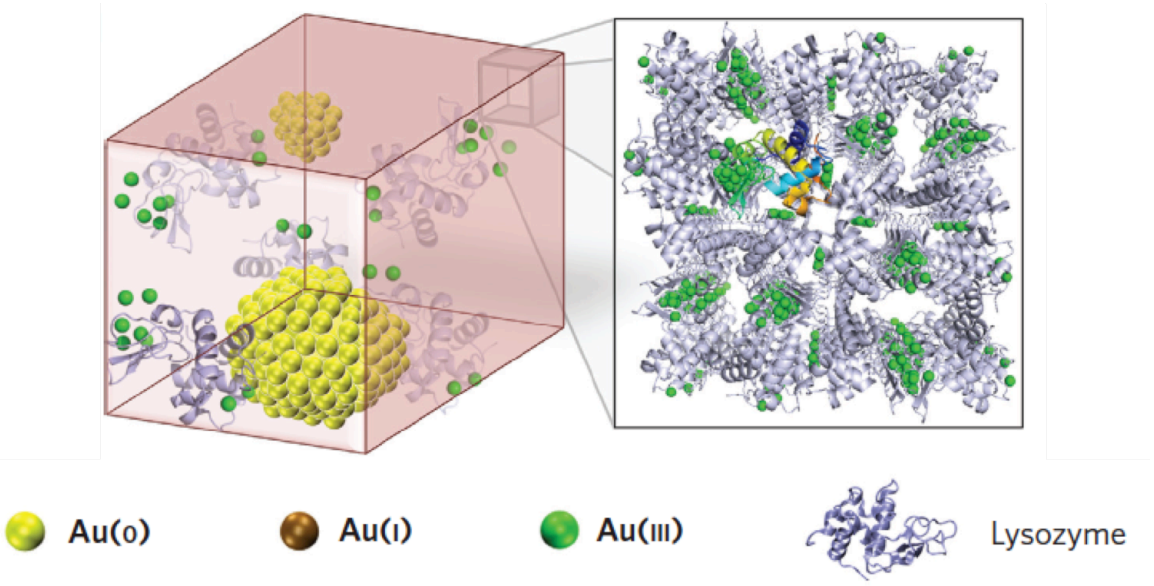



Figure 2.9. Experimentally determined crystallographic packing of 20 nm AuNPs within the lysozyme protein crystal.[178]

Wei *et al.* have used histidine residues within lysozyme to template AuNPs without disturbing the crystal structure of the enzyme.[178] Reaction kinetics, including the speed at which templation occurs has also been used to control AuNP size. Protein crystals of non-cross-linked lysozyme have been grown in the presence of $ClAuS(CH_2CH_2OH)_2$, resulting in *in situ* templation of AuNPs within the protein crystals.[178] The growth of the AuNPs in the presence of these proteins was much slower than in the absence of the proteins, leading to fine control over the AuNP diameter up to 20 nm (Figure 2.9).[178] In summary, proteins can control the process of AuNP

crystallization through residue-specific interactions as well as using the presence of other chemicals and ions to fine-tune the growth process.

#### 2.2.3.2 Templation by structured proteins

Templation of metal nanoparticles onto structured protein scaffolds is a clever and facile method of generating materials with predictable nanoparticle spacing and sizes.[158,162,164–168,171] Structured proteins that template AuNPs and form robust, sensitive nanocomposite membranes with nanoscale thickness and macroscopic lateral dimensions enable the design of biosensors capable of detecting a range of molecules.[22,153] More importantly, the preparation of such protein·AuNP assemblies embodies a simple way to generate influential and sophisticated materials.

Work by Rosi *et al*. provides an elegant example of *β*-sheet $C_{12}$-$PEP_{Au}$ amphiphile peptides that further assemble to form double helices that can be decorated with AuNPs (Figure 2.10, Table 2.1).[158,166] $C_{12}$-$PEP_{Au}$, a peptide with sequence AYSSGAPPMPPF that has a succinimide-activated dodecanoic acid N-terminus, was evolved to display a high affinity towards gold surfaces and isolated via phage-display techniques (Table 2.1).[161] Structural parameters of AuNP templation, such as identity of the reducing agent and adjusting reaction time and temperature, can affect the density of templation on $C_{12}$-$PEP_{Au}$ to achieve distinct nanoparticle superstructures.[158]

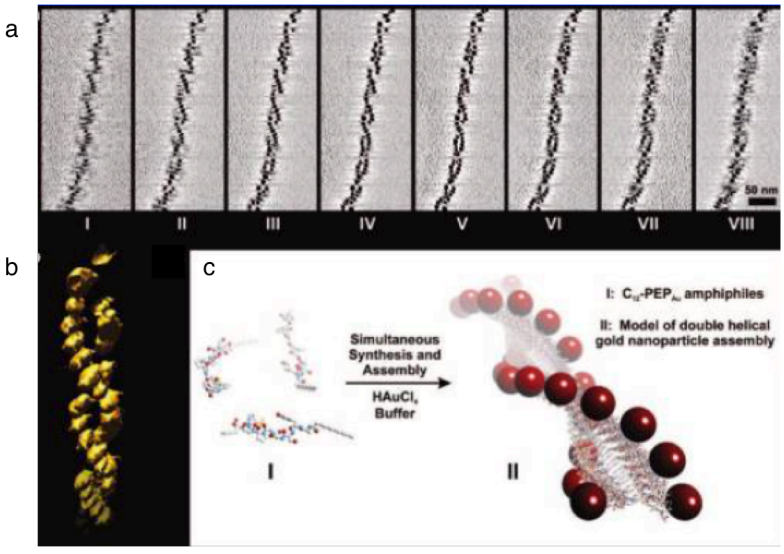



Figure 2.10. Electron tomography (a, b) of left-handed double helices of *β*-sheet $C_{12}$-$PEP_{Au}$ decorated with AuNPs and a schematic rendering of their assembly (c).[166]

There are also several examples of structured helical proteins that template AuNPs to generate interesting materials. Stability of AuNPs in aqueous solution over time can be obtained through complexation or encapsulation with self-assembling proteins, as has been shown by Nishikawa *et al.*[167] This group used a hydrophobic helical peptide combined with a hydrophilic poly(ethylene glycol) unit to create an amphiphilic block copolymer that self-assembled into a vesicular structure capable of reducing $Au^{3+}$ into AuNPs and encapsulating the NPs in solution without any additional reducing agent.[167] The protein copolymer·AuNP assemblies were stable for several days.

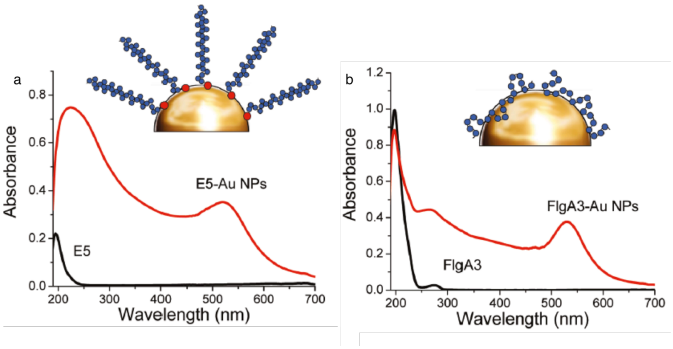



Figure 2.11. UV-vis spectra of E5 (a) and Flg-A3 (b) in the presence and absence of AuNPs.[162]

Recently Jung *et al.* obtained spatial control of AuNPs with a helical gelator, generating chiral materials for photonics and sensing applications.[168] Right-handed or left-handed helicity and formation of nanofibers was tuned by the addition of two different chiral compounds, and subsequent photoreduction of Au(I) at 254 nm resulted in nanoparticle superstructures with control over the surface plasmon absorption wavelengths.[168] Slocik *et al*. have also used proteins to create optically active chiral gold nanoparticle structures.[162] The unstructured Flg-A3 (Figure 2.6d, Section 2.2.2.2) and helical E5 peptides were used in templation of AuNPs 10 nm in diameter, where Flg-A3 interacts with AuNPs via non-covalent interactions while E5 forms a thiol linkage to AuNPs via an N-terminal cysteine residue present in E5 enabling the peptide to extend outwards from the NP surface, as is shown in Figure 2.11. Chiroptical properties introduced to nanosystems by way of structured proteins can have implications in the areas of design of nanostructured materials and developing novel recognition and separation processes for chiral molecules.

#### 2.2.3.3 Templation by unstructured proteins

Design of protein·NP assemblies with long-range order is arguably more easily obtained with the use of structured protein template materials. Be that as it may, there are also several examples of templation by unstructured protein.[162,164,169,170] As we have just discussed in Section 2.2.3.2, not only was helical E5 able to complex with AuNPs, but the unstructured FlgA3 peptide also interacted with the AuNPs through a flat, two dimensional conformation through the interaction of multiple residues with the gold surface (Figure 2.11b).[162] As was mentioned in Section 2.2.3, Deka *et al.* studied templation of unstructured protein with citrate-stabilized, soluble AuNPs.[164] AuNP complexation with denatured $\alpha$-amylase resulted in minimum agglomeration of the NPs (Figure 2.7d), the extent of which varies with the mole fraction of native (or denatured) protein corresponding to a given broadening of the UV-vis spectrum, thus providing a quantitative estimation of the conformation content of a protein structure in a mixture.[164]

Higuchi *et al*. described the preparation of peptide-coated AuNPs whose assembly is controlled by a conformational transition of the surface-bound $\alpha$-helical protein.[169] The dipole-dipole interactions between surface peptide chains lead to fibril formation by the $\alpha$-helical proteins. Conversely, when peptides on the surface of the AuNPs were in random coil configurations, peptide·AuNPs existed in a dispersed globular state. The secondary structure of surface-bound peptides was modified via an external trigger, as is discussed in Section 2.2.3.4. Bayraktar *et al*. investigated binding of functionalized AuNPs to folded Cyt c and unfolded apoCyt c. Formation of $\alpha$-helices was induced in soluble AuNPs at low concentrations, whereas at high concentrations they obtained micron-scale insoluble aggregates.[170] Both Cyt c and apoCyt c were able to direct the assembly of nanocomposite materials on the micron scale through complementary electrostatic interactions.[170] As we shall see, examples provided by Higuchi, Bayraktar *et al.* provide an excellent basis for comparison for the results obtained in this study, as is discussed in detail in Section 2.4.

#### 2.2.3.4 Switching by external triggers

Manipulation of secondary structure, and hence control over assembly, can be accomplished using a variety of external triggers, including modification of solvent pH

and polarity,[61,170,171,182] photothermal perturbations,[183,184] altering peptide concentration,[170,173] or through the use of additives, such as zinc.[88,185]

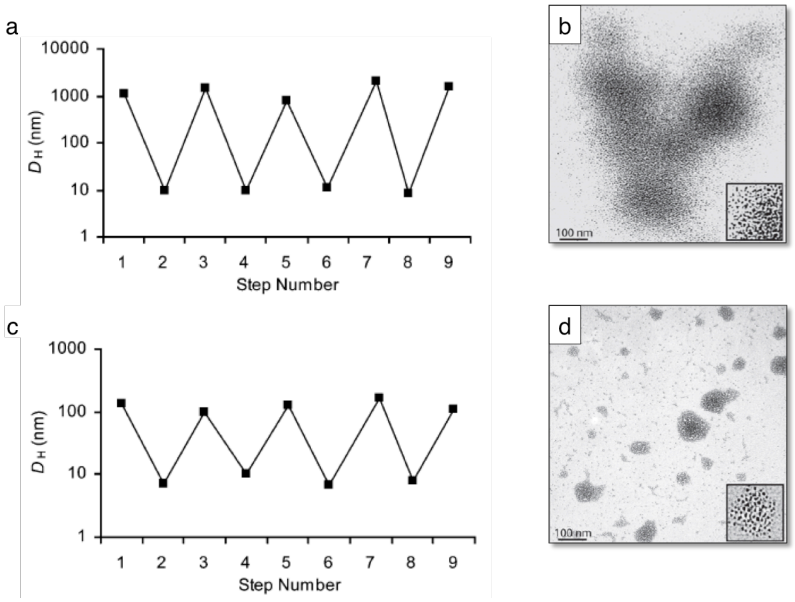



Figure 2.12. Dependence of hydrodynamic radius of apoCyt c·AuNP complexes at a 10:1 (a, b) and 20:1 (c, d) molar ratio between pH 7.4 and pH 10.8. TEM images of the complexes show smaller particles at a molar ratio of 20:1 (d) compared to large aggregates at 10:1.[170]

Specific examples include reversible assembly/disassembly of coiled-coil protein·AuNP aggregates upon pH switching between acidic and neutral or basic conditions.[61,171] In such cases, decrease in pH results in change of conformation from unpaired, unstructured $\alpha$-helices to coiled-coil dimers with pre-made AuNPs chemically conjugated to the proteins or peptides.[61,171] As discussed in Section 2.2.3.3, Bayraktar *et al*. studied nanocomposites constructed from random apoCyt c or helical Cyt c.[170] In this work the authors used pre-fabricated Au-TAsp NPs featuring thiol ligands containing an aliphatic segment for nanoparticle stability, a tetra(ethylene glycol) segment to minimize nonspecific interactions, and a dianionic to provide a charged surface (Table 2.1). In this work they also evaluated the dependence of the micron scale aggregates on pH, to the conclusion that assembly of the nanocomposites could be controlled by altering the pH from 7.4 to 10.8, at two ratios of apoCyt c to AuNPs. Higher ratios produced smaller nanocomposites while lower ratios resulted in larger assemblies, and hydrodynamic radius was reversibly switched between pH 7.4 and 10.8, as is seen in Figure 2.12.

Another example of pH tunable protein·AuNP assemblies comes from Tirrell *et al.*, who used two leucine zipper-like peptides (one acidic, one basic) that form coiled-coils to functionalized the surface of premade 8.5 nm AuNPs via a thiol bond with a terminal cysteine residue present on the peptides (Table 2.1).[171] The acidic leucine zipper peptide (termed peptide 1) revealed an increase in helicity as pH was decreased from 11.5 to 4.5, as determined by CD studies. At pH 11.5, AuNPs in the presence of peptide 1 were well dispersed, and aggregated as pH decreased to 4.5.[171] These studies highlight the flexibility afforded to the engineer in using proteins susceptible to conformational changes upon pH modification for medical and bioengineering applications.

Adjusting the concentration of a particular peptide or the molar ratio of peptide to AuNP is another strategy that has been used in controlling the assembly of nanocomposites. Bayraktar *et al.* (*vida supra*) observed that low concentrations of protein resulted in soluble AuNP nanocomposites, whereas high protein concentration generated insoluble assemblies.[170] This group also studied varying the molar ratio of apoCyt c to AuNPs, where smaller particles were obtained at a molar ratio of 20:1 compared to large aggregates at 10:1 (Figure 2.12, Table 2.1).[170]

The effect of additives on the secondary structure is also a useful tool in modifying protein·NP assembly, and has been introduced in Section 1.2.3.3. Gunsekar *et al.* from our group have used curcumin to modify the assembly of COMPcc-based proteins and several divalent metals, as is discussed in Section 1.2.3.3. Higuchi *et al*. used TFE to switch from unstructured poly($\gamma$-methyl L-glutamate-co-L-glutamic acid) (MG/GA) protein (number average MW of 73 kDa) assemblies to create fibrillar assemblies of AuNP and structured protein (Figure 2.13, Table 2.1).[169] In this study AuNPs were synthesized prior to binding with MG/GA by mixing 0.2 mM $HAuCl_4$ with alkanethiols at a molar ratio of 6 followed by reduction with $NaBH_4$. MG/GA peptides were bound on to the AuNP surface through ring-opening polymerization of N-carboxyanhydride of $\alpha$-amino acid with free amino groups fixed on the AuNP surface as an initiator.[169] This allowed Higuchi *et al.* to control the surface coverage of peptide chains on the AuNPs by adjusting the number of fixed amino groups on the surface of the NPs.

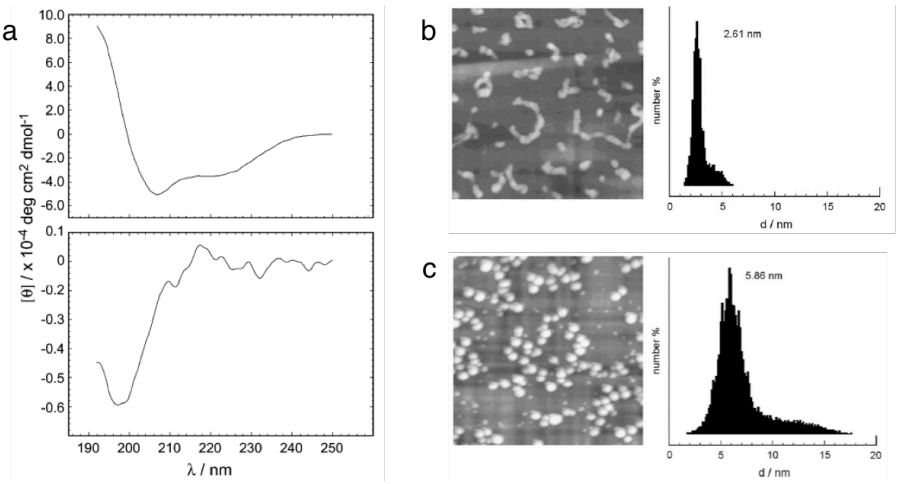



Figure 2.13. CD wavelength scans of MG/GA·AuNP nanocomposites in TFE (a, top) and $dH_2O$ (a, bottom). AFM scans (b, c left) and particle height analysis from cross-section evaluation of AFM scans (b, c right) of assemblies in TFE (b) and $dH_2O$ (c).[169]

When peptide·AuNP assemblies were dissolved in $dH_2O$, the peptides were random coil (Figure 2.13a, bottom) and NPs appeared as dispersed globular species 2.61 nm in height, on average (Figure 2.13b).[169] In TFE, peptides became $\alpha$-helical, with the $\alpha$-helical peptide chains on the AuNPs connecting the NPs to form a fibril assembly (Figure 2.13c), resulting from dipole-dipole interaction between the surface peptide chains.[169]

The combination of governable proteins with inorganic nanoparticles brings new possibilities to construct and deconstruct functional hybrid materials – often an unattainable goal for bulk metals. It is our aim to expand the control afforded by external triggers while also fabricating protein·AuNP assemblies under ambient conditions in one-pot. Herein we develop coiled-coil proteins capable of templating AuNPs *in situ*, resulting in tunable conformation-dependent protein·AuNP nanocomposites that are readily deposited on electrodes.

### *2.2.4 Design methodology*

Our group has recently developed engineered proteins based on the coiled-coil region of COMPcc which possess unique abilities to self-assemble into fibers.[1] One of these proteins, Q, was created by swapping the N and C-termini of the original wild type (wt) COMPcc sequence to obtain a charge distribution enabling self-assembly of protein fibers (Section 1.2.5.2). In addition to Q, another variant of wt COMPcc was generated

maintaining the original orientation of the N and C termini, but truncated by one heptad repeat originally in the C terminal region of wt proximal to the loop region and not necessary for structure. This protein has been termed C (see Section 1.2.5.1). In accordance with the conclusions of protein characterization studies,[1] self-assembling C and Q proteins have been selected as scaffolds for Au templation. Both C and Q contain a dual purpose 6-His tag near the N terminus (Table 2.2); the 6-His tag aids in purification and serves as a region capable of inducing templation of divalent metal nanoparticles, as determined with $Ni^{2+}$ and $Zn^{2+}$ in previous studies by our group.[88]

Table 2.2. Sequence information for C, Q, Cx, and Qx, from N-terminus to C-terminus, with the histidine tags in bold.

| | His tag | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg |
|---|---|---|---|---|---|---|---|
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* |
| C | MRGS**HHHHHH**GSIEGR | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | *1* | *17* | *20* | *27* | *34* | *41* | *48* |
| Q | MRGS**HHHHHH**GSIEGR | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL |
| | | *1* | *8* | *15* | *22* | *29* | *36* |
| Cx | | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | | *1* | *4* | *11* | *18* | *25* | *32* |
| Qx | | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL |

In some cases, identity of the secondary structure has been taken advantage of to manipulate macromolecular assembly,[133,169,171,185] whereas in others incorporation of certain residues or organic molecule appendages influences templation of metal NPs.[158,166,186,187] The 6-His tag on C and Q could be eliminated to produce Cx and Qx, respectively (Table 2.2).[13] Herein the cleavage of the 6-His tag allows us to take advantage of both of these strategies to modify macromolecular assembly of protein·AuNP structures.

### *2.2.5 Applications*

Gold is an inherently stable metal with a high oxidation resistance and excellent biocompatibility, which leads to its use in a wide range of applications. AuNP assemblies, particularly those that are stabilized by biomolecules, are of increasing

interest as reaction catalysts,[188] biomedical imaging,[189] biological labeling,[154] and sensing applications.[136,190]

Unique properties of AuNPs have made them extremely applicable in the areas of biomedical imaging and photothermal therapies. AuNPs have been used for biomedical imaging and photothermal therapy, owing to the strong near-infrared absorption (NIR) of gold nanorods, nanocages, and hollow gold nanospheres.[136] Radiative properties such as absorption, scattering, and SPR make these nanoparticle assemblies very useful for molecular cancer imaging. The choice of size of AuNPs for a particular biomedical application is very important, as SPR and NIR are dependent upon particle size and shape, where the ratio of scattering to absorption increases with particle size.[183] For imaging applications, larger particles are often selected because of their high scattering efficiency, whereas photothermal applications call for smaller NPs that efficiently absorb light and convert it to heat which is then used in the destruction of targeted tissues and cells.[189] Slocik *et al.* have demonstrated the use of a coiled-coil peptide motif in the assembly of two types of active NP composites that respond to NIR illumination via a thermally sensitive linker molecule.[183] Active nanostructures that exhibit light-inducing responses, such as the coiled-coil NP composite designed by Slocik *et al.*, allow for remote control of NP assembly by optical illumination, making them ideal materials for advanced *in vivo* photothermal therapies.[183]

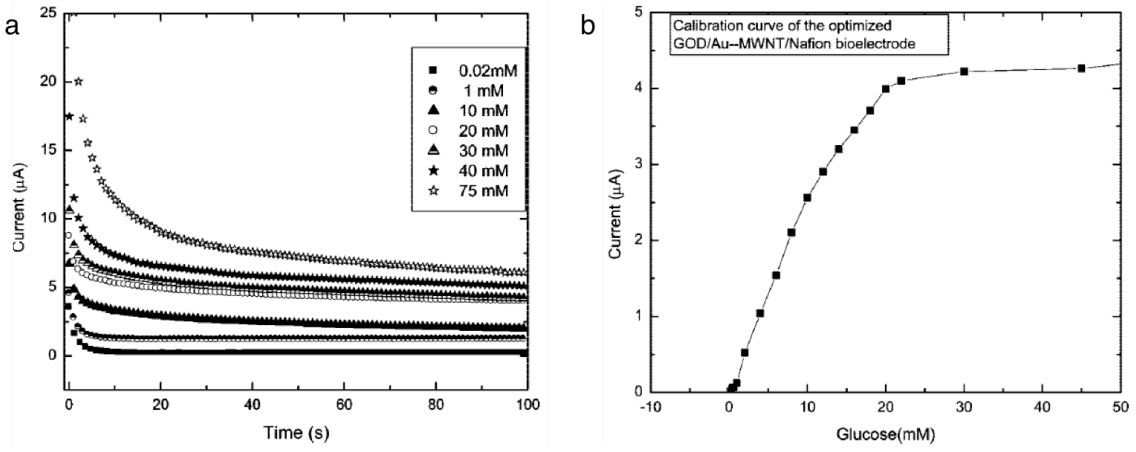



Figure 2.14. Amperometric measurements of a glucose biosensor with glucose oxidase-functionalized AuNPs on a multiwalled carbon nanotube (MWNT) electrode. *i-t* curves obtained at various concentrations of glucose (a) were translated into a calibration curve (b) for the biosensor.[191]

In medicine, the detection and quantification of enzyme concentrations are important tools in the screening of toxins and identification of pathologies. Nanoparticle-based assays have been used with specific proteins for enzymatically-controlled aggregation, which translates to detection of not only enzyme presence but also analytical quantification via spectroscopic and electrochemical methods.[191–193] Rakhi *et al*. developed a biosensor system with glucose oxidase functionalized AuNPs that was able to detect concentrations of glucose as low as 20 μM through electrochemical measurements (Figure 2.14). Notably, enzyme-NP systems in biodetection are able to reach detection levels comparable to fluorescence resonance energy transfer (FRET) or enzyme-linked immunosorbent (ELISA) assays.

Protein and gold nanocomposites are also being studied for applications in chemical and biological sensing and catalysis. The key to these types of applications is imparting electronic properties to the composite material via the introduction of metal nanoparticles, as proteins are generally considered electronic insulators (with few exceptions, including pilin filaments[146]). NP assemblies, with increased surface area compared to bulk metal materials and expression of phenomena such as SPR, make them materials prone to extreme sensitivity of environmental conditions. As SPR is dependent on NP size and shape, parameters such as these can be smartly controlled by protein matrices, as we have seen in Section 2.2.3. Electrodes with deposited AuNP films, compared to electrodes of bulk gold, exhibit very high sensitivity for electrocatalytic activity of CO oxidation, $O_2$ reduction, and hydrogenation of aldehydes and alcohols.[152] Proper coupling of abiotic and biotic materials offers numerous possibilities for overcoming barriers in nanobioelectronics.

## 2.3 Materials and methods

### *2.3.1 Materials*

Sodium phosphate (monobasic and dibasic), nickel-nitrilotriacetic acid resins, $HAuCl_3 \cdot 3H_2O$ salt, sodium borohydride, potassium hexacyanoferrate(II) trihydrate, potassium hexacyanoferrate(III), and trifluoroethanol (TFE) were purchased from Sigma-Aldrich. Ampicillin, isopropyl-*β*-D-thiogalactopyranoside (IPTG), tryptone, urea, tris-

HCl, and sodium chloride were obtained from Fisher Scientific. Yeast extract and methanol were purchased from Acros Organics and BCA kit was obtained from Pierce. Imidazole was purchased from Alfa Aesar and copper grids for TEM were purchased from Ted Pella. Dialysis tubing (3.5 kDa molecular weight cut-off) was obtained from Thermo Scientific. Factor Xa cleavage kit was purchased from Novagen. Glassy carbon electrodes were obtained from CH Instruments, Inc. See Table A 2 for a complete list of reagents and suppliers.

### *2.3.2 Expression, purification, and cleavage*

Gene sequences for C and Q were generated via polymerase chain reaction (PCR) amplification (see Section 1.4.1 for details).[1] Transformation was performed for insert DNA/PheRS** ligated plasmid vectors. Two reactions were prepared for each DNA type: (1) 5 µL DNA/PheRS** DNA was added to 100 µL Zymo xL1 blue cells, and (2) 15 µL DNA/PheRS** DNA was added to 100 µL Zymo xL1 blue cells. Cells were kept on ice and thawed on ice for 30 minutes. After 30 minutes, 700 µL of Luria Bertani broth (at 37 °C) was added to each reaction, and was shaken at 350 rpm at 37 °C for 45 minutes. Amp/TSA plates were used to grow colonies. 5 plates were grown for each DNA type: (-) control sample and 4 (+) control samples: (1) 200 µL of reaction 1, (2) 600 µL of reaction 1, (3) 200 µL of reaction 2, and (4) 600 µL of reaction 2. Plates were left in the incubator for 20 minutes half covered to allow them to dry. Then were closed completely, inverted, and incubated at 37 °C overnight.

Colonies were initially grown on TSA plates containing 0.2 mg/mL AMP as a selection antibiotic at 37 °C for 16 hours. AMP was used to select for genetically modified host cells containing the plasmid vector only.[109] Single colonies were selected from the plates and inoculated in 5 mL of LB with 0.2 mg/mL AMP and incubated at 37 °C, for 12 hours, at 350 rpm.

Proteins were expressed in 200 mL volumes contained in 1 L baffled Erlenmeyer flasks. 1 mL of starter culture was added to 200 mL of LB and 0.2 mg/mL ampicillin. Glycerol stocks were prepared from starter cultures. 1 mL pre expression samples were taken from each culture at $OD_{600} > 0.8$, at which point cultures were induced with 0.2

mg/mL IPTG. Induction was allowed to occur for 3 h, at which point post expression samples were taken. Cultures were shaken continuously at 250 rpm at 37 °C. Post induction, the cells were harvested by centrifugation on a Beckman Coulter Allegra™ centrifuge, Beckman Coulter TS-5.1-500 swinging bucket rotor, at 4 °C and 8000 rpm for 20 minutes. The supernatant was disposed of, and the cell pellets were stored in -80 °C until purification. A freeze/thaw cycle at -80 °C was carried out to facilitate lysis, followed by probe sonication (Sonicator 500, Qsonica, LLC, CT, USA) for 1 minute, with pulse 5 s on/5 s off at 35 % amplitude.

Pre/post expression samples for each culture were run on SDS-PAGE gels to evaluate protein expression. Purification under denaturing conditions was carried out using 50 mM tris-GCL, 0.5 M NaCl, 20 mM imidazole, 6 M urea pH 8 buffer. The soluble crude lysate was bound to Ni-NTA beads and allowed to equilibrate for at least 3 h at 4 °C under continuous rotation. The proteins were eluted with increasing gradient of imidazole (20 mM – 1 M). Pure fractions were refolded via stepwise dialysis in pH adjusted phosphate buffer (50 mM, pH 8), halving the urea concentration successively. The enhanced protocol of BCA (PIERCE) (see Appendix 5.6.6) was used to estimate protein concentration with bovine serum albumin (BSA) as a standard.

For experiments in TFE, protein was diluted in PB containing 25 v/v % TFE after dialysis and BCA. After confirming purity and protein concentration of > 10 μM, cleavage was initiated. Cleavage of the histidine tags was performed by using Factor Xa cleavage kit (Novagen). Cleavage was performed according to the manufacturer's protocol, with the following changes: protein to 10x cleavage buffer volume ratio was 9:1 and protein to enzyme molar ratio was maintained at 50:1. The reaction mixture (204 μL) was incubated at room temperature for 4 h. Factor Xa and histidine tag capture were performed according to the manufacturer's protocol, with the exception that cleaved protein was isolated from Ni-NTA beads using purification. After histidine tag capture, cleaved protein was dialyzed into 50 mM phosphate buffer pH 8 and concentration was determined via BCA (Appendix 5.6.6).

### *2.3.3 Gold binding*

Molar concentration of protein to Au to sodium borohydride was kept constant at a ratio of 1:40:100.[194] Au salt solution ($HAuCl_3 \cdot 3H_2O$) was added so that the final concentration was 400 μM ($dH_2O$) and the final concentration of the protein was 10 μM (in 50 mM PB, pH 8). For templation in the presence of TFE, 25 v/v % TFE was added to protein solution for a final protein concentration of 10 μM, followed by Au incubation for 15 minutes and reduced. To reduce the metal, a final concentration of sodium borohydride (1 mM) was added.

### *2.3.4 Absorbance*

Absorbance measurements were carried out using 10 μM of protein (50 mM PB, pH 8) in the absence and presence of 400 μM Au at room temperature. Absorbance was measured using a SpectraMax Plus M2 instrument (Molecular Devices, CA, USA). Spectra were measured over a range of 400–700 nm.

### *2.3.5 Circular dichroism*

CD measurements were conducted on a Jasco J-815 CD spectrometer (Jasco Inc. MD, USA), as described elsewhere (Section 1.3.5).[1] Wavelength scans were conducted with 10 μM protein concentrations in 50 mM PB pH 8 at room temperature. CD was also performed in the presence of 25 v/v % TFE and 50 mM PB pH 8 with 10 μM protein in the presence and absence of AuNPs. Mean residue ellipticity (MRE) was calculated from raw data according to the procedure described previously.[92]

### *2.3.6 ATR-FTIR*

Attenuated total reflectance Fourier transform infrared spectroscopy **(**ATR-FTIR) experiments were performed using Perkin Elmer System 2000 FT-IR (Perkin Elmer, MA, USA) with DuraSamplIR II T diamond ATR accessory and equipped with a MCT-A detector (Smiths Detection Inc., CT, USA), as described previously (Section 1.3.7).[1] Approximately 5 μL of protein solution (10 μM in 50 mM PB pH 8) with or without AuNPs was added on the diamond ATR surface. ATR-FTIR was also performed in the presence of 25 v/v % TFE and 50 mM PB pH 8 with 10 μM protein in the presence and

absence of AuNPs. PeakFit software was used to process the data, which involved a 2nd derivative zero baseline correction of the amide I region between 1700–1600 $cm^{-1}$ and deconvolution of peaks with a Gaussian function.[115] All readings represent the average of two trials.

### *2.3.7 TEM*

Approximately 3 μL of 10 μM protein in 50 mM PB pH 8 with or without AuNPs was spotted on copper grids, as done before (Section 1.3.8).[1] After 1 minute, the grids were blotted using filter paper and rinsed with 3-4 drops Milli Q water to remove excess salts from the buffer. After blotting with filter paper, protein samples in the absence of Au were negatively stained by adding 3 μL of 1 % filtered uranyl acetate and blotted using filter paper. Samples with AuNPs were not stained. Samples were then dried at room temperature for 10-15 minutes. TEM was also performed in the presence of 25 v/v % TFE and 50 mM PB pH 8 at 10 μM protein in the presence and absence of AuNPs. Prepared grids were viewed on a JEOL JEM-1400 TEM (Jeol, MA, USA) at an accelerating voltage of 120 kV. The same grids were further analyzed with an EDAX energy dispersive X-ray spectrometer to obtain elemental maps. ImageJ software was used to measure nanoparticle dimensions.[116]

TEM was performed according to the procedure outlined in Section 1.3.8, with the exception that uranyl acetate stain was not used in imaging of AuNP-containing samples.

### *2.3.8 Cyclic voltammetry and impedance*

Electrochemical measurements were performed with a three-electrode system comprising a platinum wire counter electrode (MW-1032, BASi, IN, USA), an Ag/AgCl SDR2 Super Dri-Ref reference electrode (World Precision Instruments, FL, USA), and a glassy carbon disk electrode (d = 3.0 mm) as the working electrode (MF-2012, BASi, IN, USA). Measurements were collected on a CHI660d electrochemical workstation (CH Instruments, Inc., TX, USA). The polished disk electrode was exposed to 8 μM protein and protein·AuNP assemblies. CV is used to characterize surface coverage of a working electrode in relation to peak current, where a decrease in peak current corresponds to

adsorption onto the electrode surface blocking detection of the redox reaction of $Fe(CN)_6^{3-/4-}$.[195] 8 μM protein concentration was chosen for CV measurements as this was seen to provide peak resistance values in impedance measurements. Faradaic impedance spectra (or Nyquist plots) are generated for all protein·AuNP assemblies. Nyquist plots consist of a semi-circular portion, corresponding to the electron-transfer-limited process and a linear portion that corresponds to diffusion-limited electron-transfer.[196] The diameter of the semi-circular portion corresponds to the electron-transfer resistance, $R_{et}$, and is helpful for characterizing electrochemical properties of surface-modified electrodes. Cyclic voltammograms and Nyquist plots were recorded in 50 mM phosphate buffer pH 8 in the presence of 10 mM $K_4Fe(CN)_6 \cdot 3H_2O$ and 10 mM $K_3Fe(CN)_6$. CV measurements were carried out at room temperature using a scan rate of 0.1 mV/s. Impedance measurements were recorded at room temperature at a bias potential of 210 mV within the frequency range of $10^{-1}$–$10^5$ Hz.

## 2.4 Results and discussion

Herein we present results relating to the templation of gold nanoparticles on proteins C and Q bearing 6-His tags and their counterparts, Cx and Qx, where the 6-His tags have been enzymatically cleaved. Upon templation of AuNPs, these proteins displayed varied macromolecular assemblies. Remarkably, uncleaved C and Q were able to bind and template gold forming large assemblies that precipitated out of solution, while the Cx and Qx proteins were able to solvate AuNPs. Figure 2.15 illustrates schematically the two different types of nanocomposite assemblies that were observed.

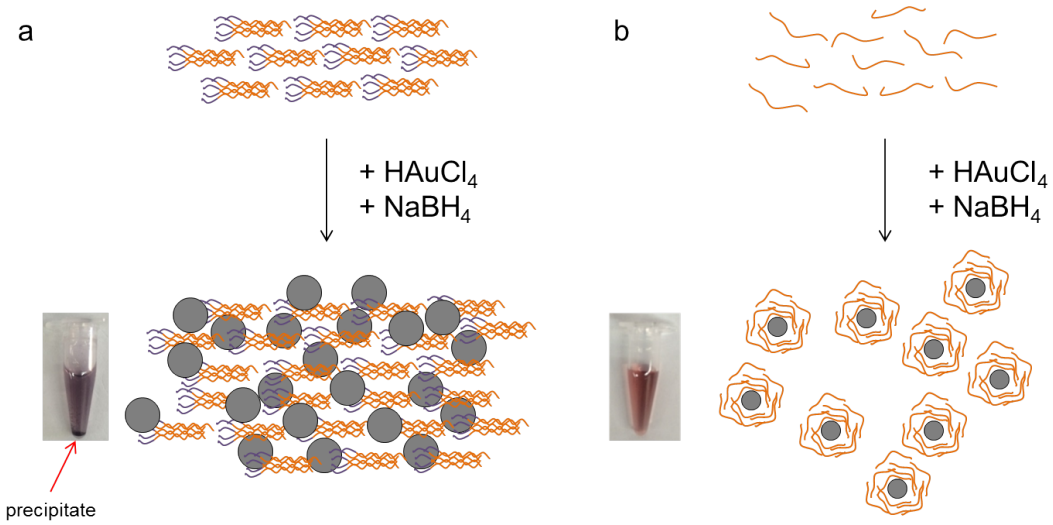



Figure 2.15. Schematic illustration of protein and nanoparticle assembly upon gold templation. Protein units are illustrated in orange, his tags in purple, and AuNPs as grey spheres. (a) $\alpha$-helical proteins C and Q with his tags form ordered pentamer assemblies, and upon gold templation exhibit a decrease in helical content and template AuNPs forming large aggregates of disordered AuNP-embedded protein. Photo demonstrates the formation of precipitates in C-AuNP 24 h post templation. (b) Cleaved C and Q with his tags removed are unstructured, and upon gold templation exhibit an increase in helical content, surrounding AuNPs and stabilizing them in solution. Photo demonstrates characteristic purple color of protein-stabilized Cx-AuNP 24 h post templation.

These findings are concluded from the results of several measurement techniques, including UV-vis absorption spectroscopy (Section 2.4.2), TEM and EDAX (Section 2.4.3), circular dichroism and ATR-FTIR (Section 2.4.4), and electrochemical techniques such as cyclic voltammetry and electrochemical impedance spectroscopy (Section 2.4.6).

### *2.4.1 Histidine tag cleavage*

Proteins C and Q were constructed to bear a cleavage site via recombinant approaches and purified (see Sections 1.4.1 and 1.4.2). Successful cleavage of the 6-His tag from C and Q by Factor Xa was confirmed by SDS-PAGE gel, with a MW of 6-His proteins corresponding to 6.31 kDa for C and Q versus cleaved MW of 4.45 kDa for Cx and Qx (Figure 2.16).

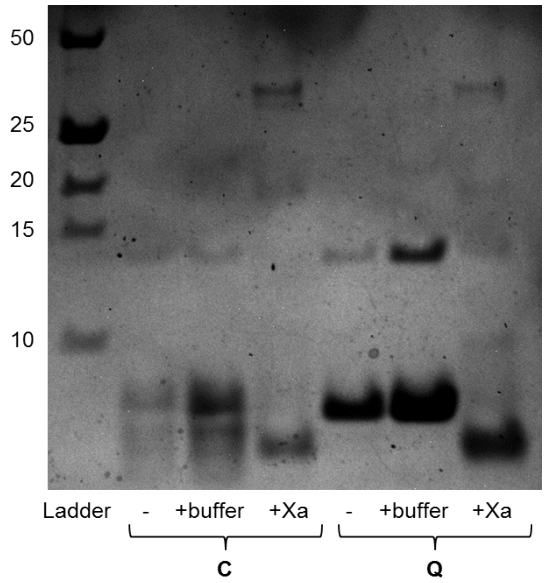



Figure 2.16. SDS-PAGE gel showing cleavage of histidine tags for both C and Q. Lanes (from left to right) are ladder, C protein (6.31 kDa), C with cleavage buffer (6.31 kDa), C after incubation with Factor Xa (4.45 kDa), Q protein (6.31 kDa), Q with cleavage buffer (6.31 kDa), and Q after incubation with Factor Xa (4.45 kDa).

### *2.4.2 Absorption profiles*

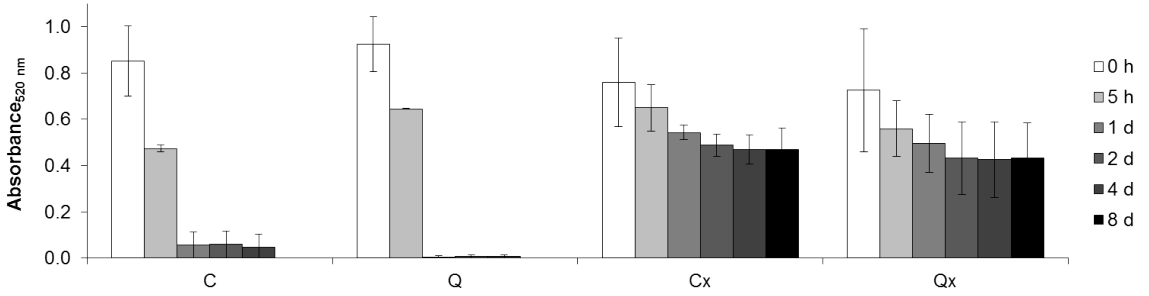



Figure 2.17. Absorbance of protein-AuNP complexes at 520 nm, where precipitation of C and Q after 1 d results in a dramatic decrease in absorbance and Cx and Qx samples have absorbances that stabilize after 2 days and cease to decrease in value.

Since spherical AuNPs exhibit peak absorbance corresponding to their surface plasmon resonance,[137] the absorption at 520 nm of C and Q and their cleaved counterparts immediately after reduction of $HAuCl_4$ with $NaBH_4$, as well as 5 h and 1 d, 2 d, 4 d, and 8 d thereafter were evaluated (Figure 2.17). The absorbance of all four proteins was high immediately after AuNP reduction, with C and Q proteins illustrating slightly higher

value compared to cleaved proteins at that time; this was indicated visually by a grey/purple color in C and Q compared to a pinkish hue for Cx and Qx (Figure 2.18).

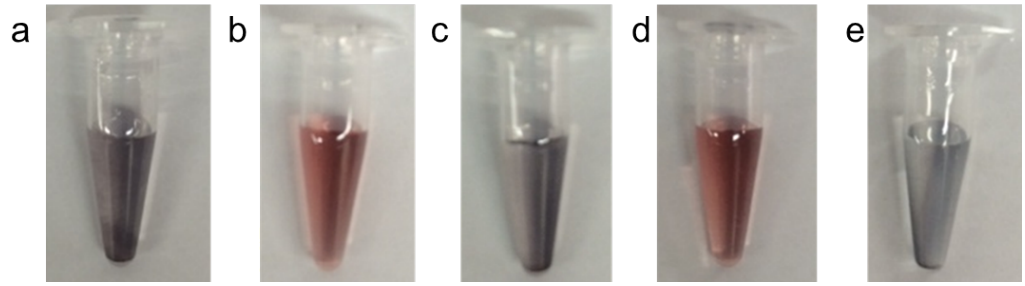



Figure 2.18. Photos of protein templated with AuNPs, taken immediately after templation. C (a) and Q (c) have a dark purple hue, similar to that seen in the absence of any protein (e). Cleaved proteins Cx (b) and Qx (d), on the other hand, have a pinkish hue that is maintained over a period of 8 days.

A dramatic decline in the absorption of C and Q was seen after just 5 h, followed by a complete drop after 1 d, after which the value remained close to zero for the remainder of the period studied (Figure 2.17). This drop in absorbance could be attributed to the formation of large protein·AuNP precipitates, sequestering all AuNPs present. Cleaved proteins Cx and Qx, on the other hand, revealed a drop in absorption that stabilized after 1 day at approximately 70 % its original value (Figure 2.17). These quantitative measurements confirmed what was observed visually; C and Q formed large precipitates loaded with AuNPs, causing the bulk solution to lose color after a few hours. By contrast, no precipitation was observed when AuNPs were reduced in the presence of cleaved proteins, whose solutions maintained their pink hue over an extended period of time. AuNPs reduced in the absence of protein (PB alone) fell out of solution over time; however they remained discrete particles rather than aggregates (Figure 2.18, Figure 2.23).

### *2.4.3 Nanocomposite morphology and elemental characterization*

#### 2.4.3.1 TEM

All proteins were visualized via TEM in conditions of 50 mM PB at pH 8. Morphology of C in the absence of gold was mainly aggregates, Q formed high aspect ratio fibers (Section 1.4.5.3),[1] and both Cx and Qx displayed sheet-like morphologies (Figure 2.19).

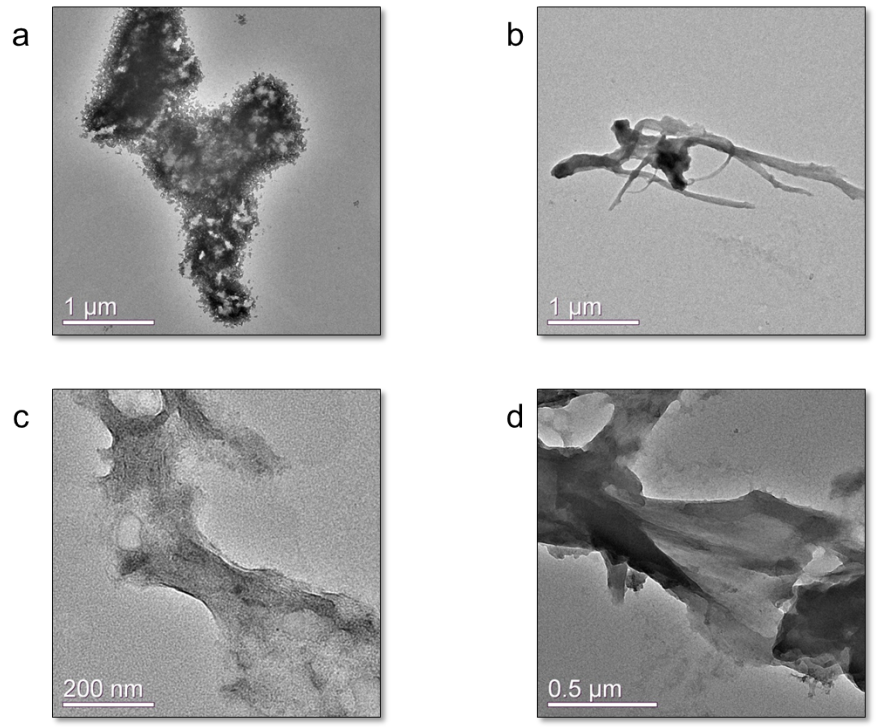



Figure 2.19. Transmission electron micrographs aggregates formed by C (a), fibers of Q (b), and sheet-like structures seen from cleaved proteins Cx (c) and Qx (d). Scale bars are 1 µm in (a) and (b), 200 nm in (c) and 0.5 µm in (d).

TEM measurements were very insightful in the study the structure of protein·AuNP assemblies on the nanometer scale. Before we introduce TEM results of AuNPs templated on C, Q, Cx, and Qx, we must discuss the templation of AuNP on wt protein as a basis for comparison, as well as templation on BSA, a protein that lacks a histidine tag and is not known to assemble into higher order protein structures. AuNPs were templated on BSA* and wt according to the procedure outlined in Section 2.3.3. As can be seen from the micrographs in Figure 2.20, AuNPs templated on BSA did not generate structures that aggregated or assembled in any organized fashion. NPs were indeed reduced, but it seems that they are present within the protein as well as in solution surrounding the protein. This is demonstrated by the presence of high contrast NPs co-localized with grey protein matter and on the lighter TEM grid itself in Figure 2.20. Despite the presence of a histidine tag, several other amino acids present in BSA are known to template AuNPs, including those with thiol (serine), amine (threonine), and

* Sequence information for BSA is provided in Appendix 5.5.8.

hydroxyl (tyrosine). BSA, in contrast to the COMPcc-based proteins used in this work, contains several serine residues (sequence information provided in Appendix 5.5.1). Thiol groups have the greatest ability to bind with AuNPs via surface interactions on the gold nanoparticles, followed by amine groups and then hydroxyl groups.[197]

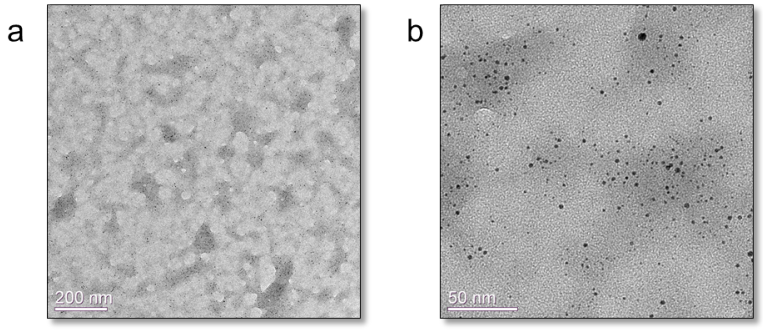



Figure 2.20. Transmission electron micrographs for BSA·AuNP in 50 mM PB pH 8 at various magnifications. Scale bars are 200 nm in (a) and 50 nm in (b).

Fibrous assemblies of AuNPs templated on wt protein are shown in Figure 2.21. Micrographs shown in Figure 2.21 demonstrate that wt protein fibers are capable of templating AuNPs while maintaining their fibrous network structure. Based on the micrographs in Figure 2.21, wt proteins display a very high coverage of AuNPs, although AuNPs also exist in solution beside the fibers (Figure 2.21c). As these samples were unstained, the contrast of the fiber networks observed in Figure 2.21a and b is solely due to the presence of templated AuNPs within the fibers. Due to the magnification of these images individual AuNPs are not visible at these length scales and are only perceived as black fibrous networks.

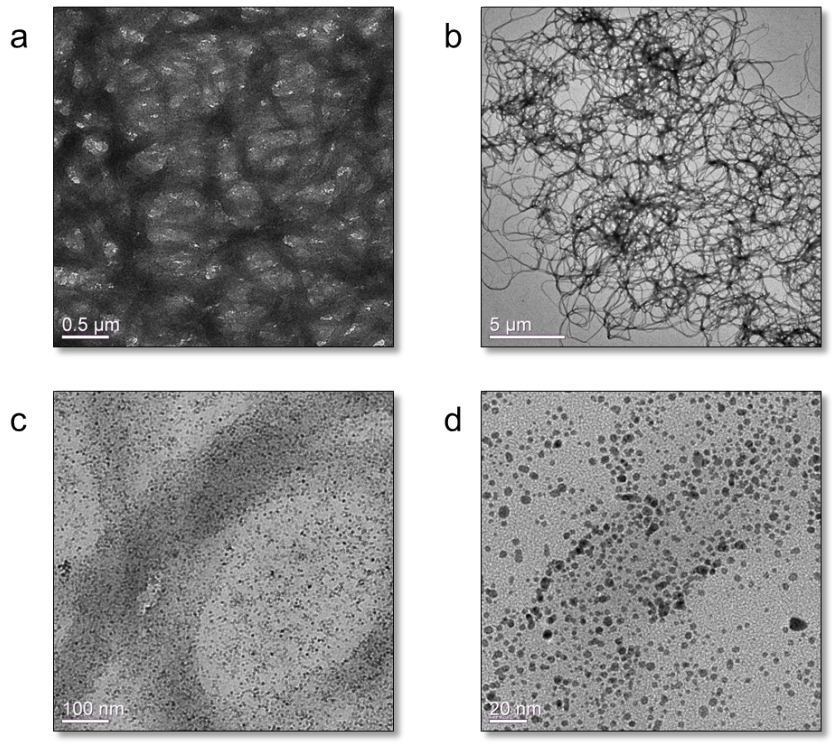



Figure 2.21. Transmission electron micrographs for wt·AuNP in 50 mM PB pH 8 at various magnifications. Scale bars are 0.5 µm in (a), 5 µm in (b), 100 nm in (c), and 20 nm in (d).

As the organization of wt pentamers within each fiber is unknown, the arrangement of the 6-His tags is also unknown. As seen by the TEM images of wt·AuNP, the arrangement of the AuNPs cannot be described as having any distinct regularity within the wt fibers.

Measurements of the diameters of the AuNPs templated on BSA and wt were performed and are compared in the histogram in Figure 2.22. AuNPs templated in the presence of BSA possessed an average diameter of 3.1 ± 0.8 nm, whereas AuNPs templated on wt protein were 4.0 ± 0.8 nm in diameter.

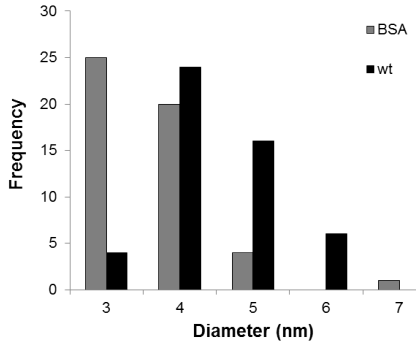



Figure 2.22. Histogram of AuNP diameters templated on BSA (grey) and wt (black).

Assemblies of AuNPs templated on C, Q, Cx, and Qx were also studied via TEM. These measurements confirmed that AuNPs were completely sequestered by C and Q, leaving few AuNPs outside of the protein matrix (Figure 2.23a and b). There was no observable difference in the AuNP density on C and Q assemblies, indicating that the presence of the 6-His tag on both proteins enables them to sequester gold from the solution and that the coiled-coil structures further assemble to form macromolecular films of AuNP-embedded proteins. Cleaved proteins, Cx and Qx, showed completely different organization of nanoparticles with AuNPs scattered across the grid in seemingly random dispersion (Figure 2.23c and d). Upon careful inspection of the Cx·AuNP and Qx·AuNP micrographs (Figure 2.23c and d), cleaved protein surrounded individual AuNPs as represented by a dark shadow around the nanoparticles.

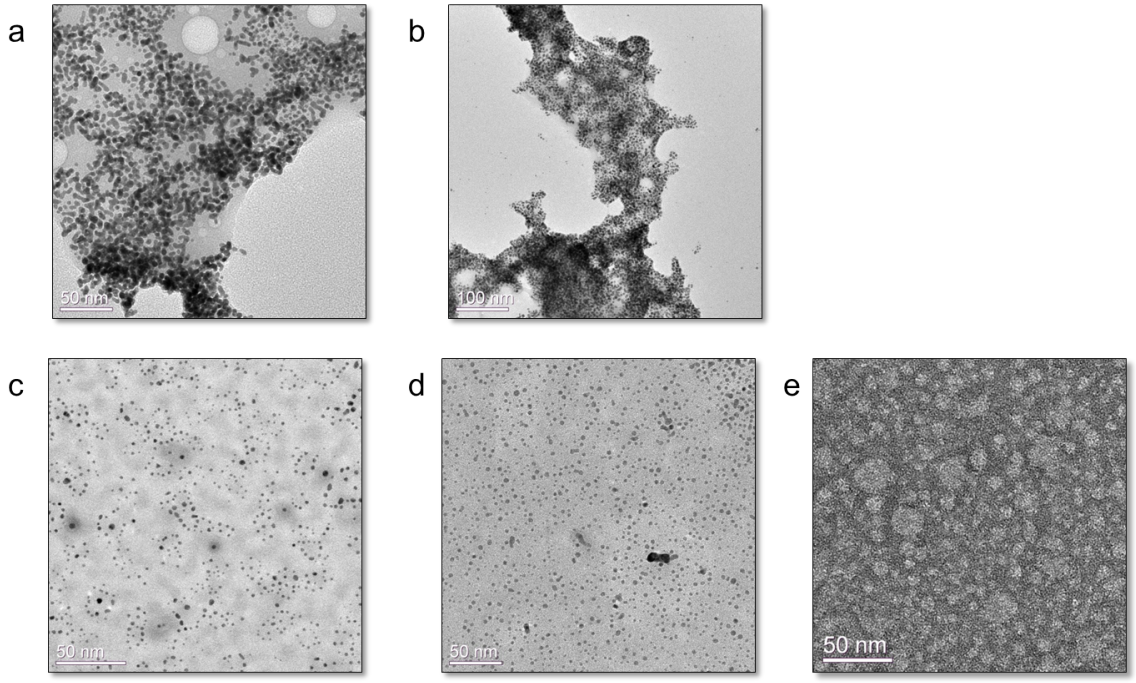



Figure 2.23. Transmission electron micrographs for C·AuNP (a), Q·AuNP (b), Cx·AuNP (c), Qx·AuNP (d), and AuNP in the absence of protein (in phosphate buffer) (e). Scale bars are 50 nm in (a), (c), (d), and (e) and 100 nm in (b).

AuNPs templated on C and Q were larger in diameter than those solvated by Cx and Qx (Figure 2.24). For C and Q assemblies, the AuNPs measured 6.7 ± 1.2 nm and 6.2 ± 0.9 nm, respectively, compared to 4.5 ± 0.9 nm for Cx and 4.0 ± 0.7 nm for Qx (Figure 2.24).

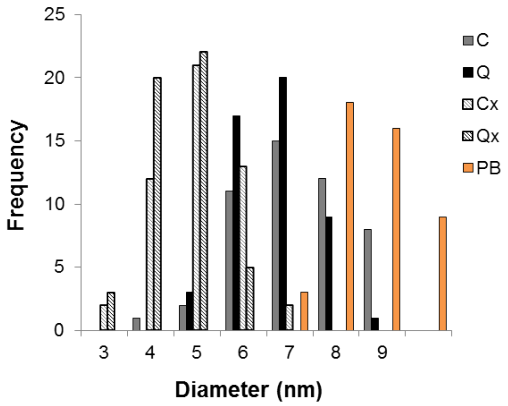



Figure 2.24. Histogram of nanoparticle sizes shows that cleaved proteins template smaller AuNPs than 6-His tagged proteins.

AuNPs reduced in the absence of protein possessed diameters of 7.3 ± 1.0 nm (Figure 2.24). Table 2.3 summarizes the AuNP sizes in the presence of each protein type. The smaller diameters observed in the presence of cleaved proteins resulted as crystallization of gold was capped due to the proteins surrounding the nanoparticles preventing them from growing further.[159,174] Dependence of AuNP diameter on secondary structure of protein for AuNPs synthesized from $HAuCl_4$ *in situ* was also observed by Kharlampieva *et al.*, who proposed that the formation of smaller, individual, AuNPs was facilitated by ordered tyrosine residues in silk II and prevented the aggregation of these AuNPs.[165] AuNPs templated with C and Q formed larger diameters that kinetically favored aggregation as increased AuNP surface area promoted interaction with neighboring 6-His tags, leading to aggregation of the protein·AuNP complex.[198]

Table 2.3. Average and standard deviation of AuNP diameters templated in the presence of various proteins ($n$ = 50).

| Protein | AuNP diameter (nm) |
|---|---|
| BSA | 3.1 ± 0.8 |
| Wt | 4.0 ± 0.8 |
| C | 6.7 ± 1.2 |
| Q | 6.2 ± 0.9 |
| Cx | 4.5 ± 0.9 |
| Qx | 4.0 ± 0.7 |
| - | 7.3 ± 1.0 |

### 2.4.3.2 EDAX

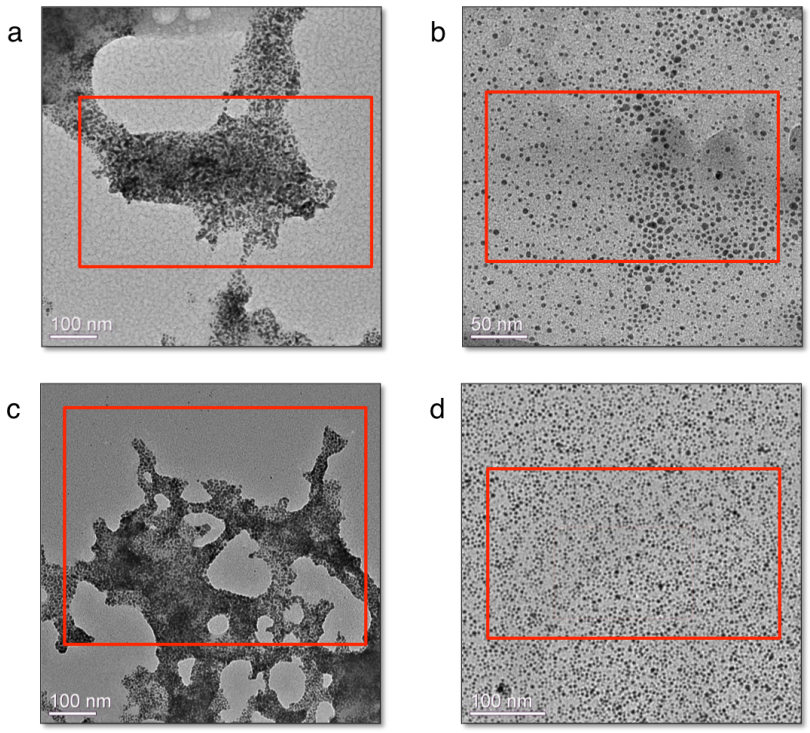



Figure 2.25. Transmission electron micrographs and elemental maps for C (a), Cx (b), Q (c), and Qx (d) with templated AuNPs. Area outlined in red was selected for data acquisition for elemental maps in Figure 2.26. Scale bars in TEM micrographs are 50 nm in (b), 100 nm in (a), (c), and (d).

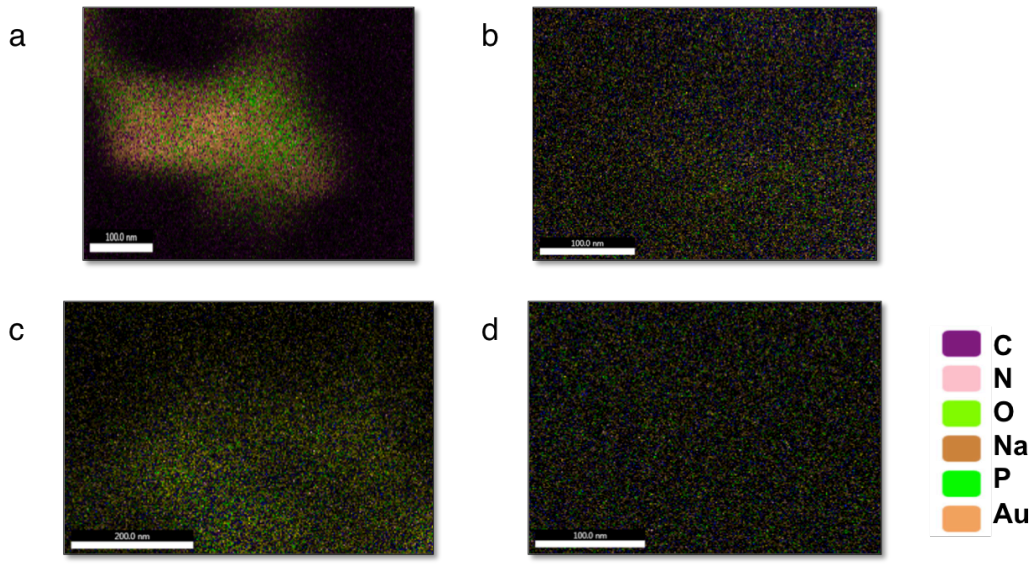



Figure 2.26. Elemental maps of C (a), Cx (b), Q (c), and Qx (d) with templated AuNPs. Maps resulted from the area outlined in red in corresponding TEM micrographs in Figure 2.25. Maps display location of signals resulting from C (purple), N (pink), oxygen (lime green), Na (rust), P (neon green), and Au (orange). Scale bars in EDAX maps are 100 nm in (a), (b), and (d) and 200 nm in (c).

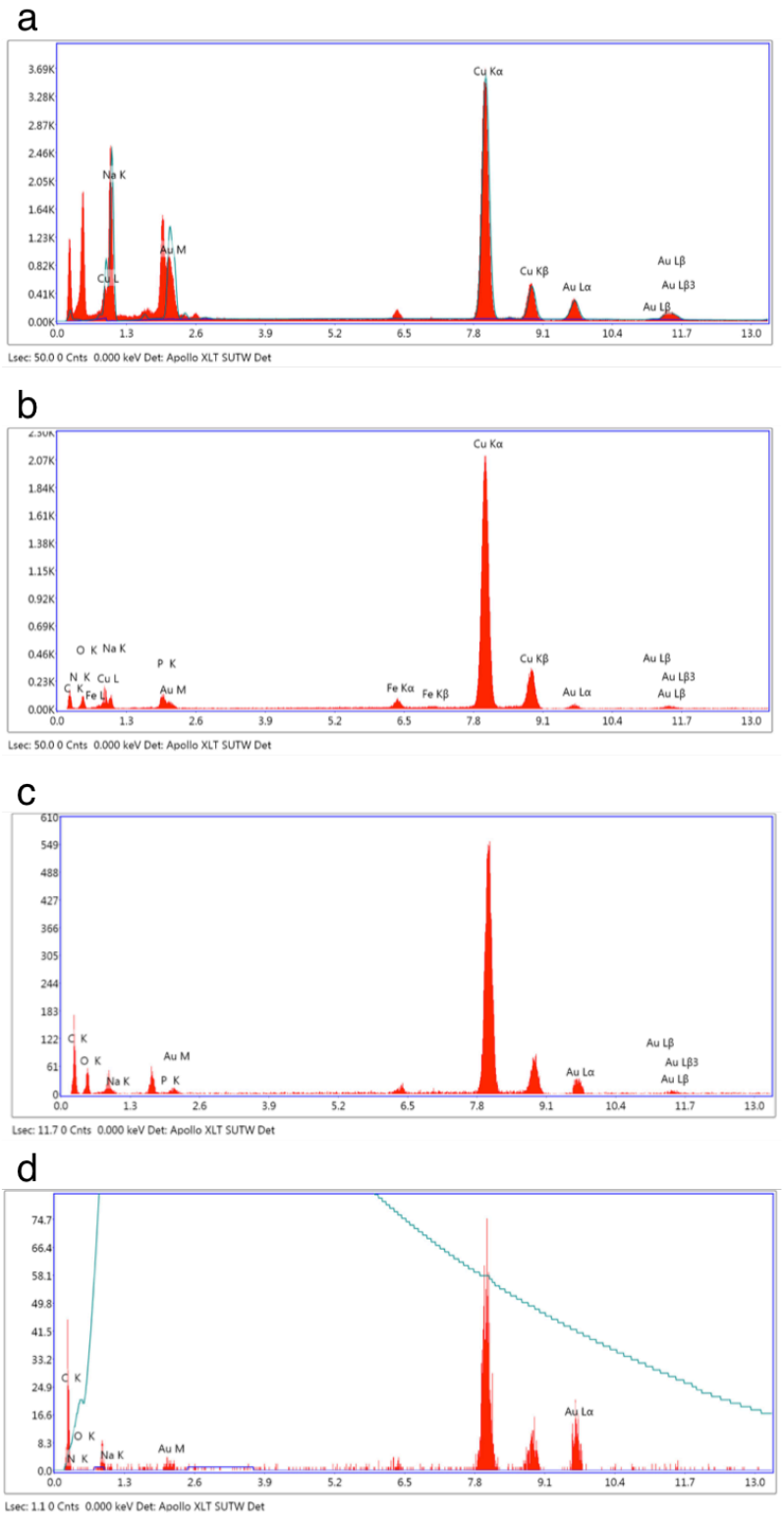



Figure 2.27. EDS spectra from C·AuNP (a), Q·AuNP (b), Cx·AuNP (c), and Qx·AuNP (d).

Elemental mapping via energy dispersive X-ray spectroscopy (EDAX) revealed elements corresponding to protein (C, N, O, P) co-localized with the AuNP in the case of the 6-His proteins (Figure 2.25, Figure 2.26), confirming that C and Q sequestered AuNPs completely from the solution. X-ray signals of elements corresponding to protein were observed throughout the elemental maps of cleaved proteins, Cx and Qx, indicating that protein material surrounded the AuNPs (Figure 2.25, Figure 2.26). Cx and Qx were able to solvate the AuNPs, enabling them to maintain their solubility over time (Figure 2.23c and d, Figure 2.25b and d). Higher count rates of gold X-ray signals from C and Q samples were observed compared with Cx and Qx, resulting from the high density of AuNPs in the aggregates formed by the uncleaved proteins (Figure 2.27).

### *2.4.4 Secondary structure*

Secondary structure analysis of C, Q, Cx, and Qx in the absence and presence of AuNP was performed via CD and ATR-FTIR to better understand the structural changes the uncleaved and cleaved proteins undergo upon AuNP templation (Figure 2.28 - Figure 2.30, Table 2.4).

CD was used to measure soluble protein structure while ATR-FTIR provided information on structural composition of insoluble protein assemblies.[122] CD data of both C and Q also displayed double minima at 208 and 222 nm, indicative of helical structure (Figure 2.28).

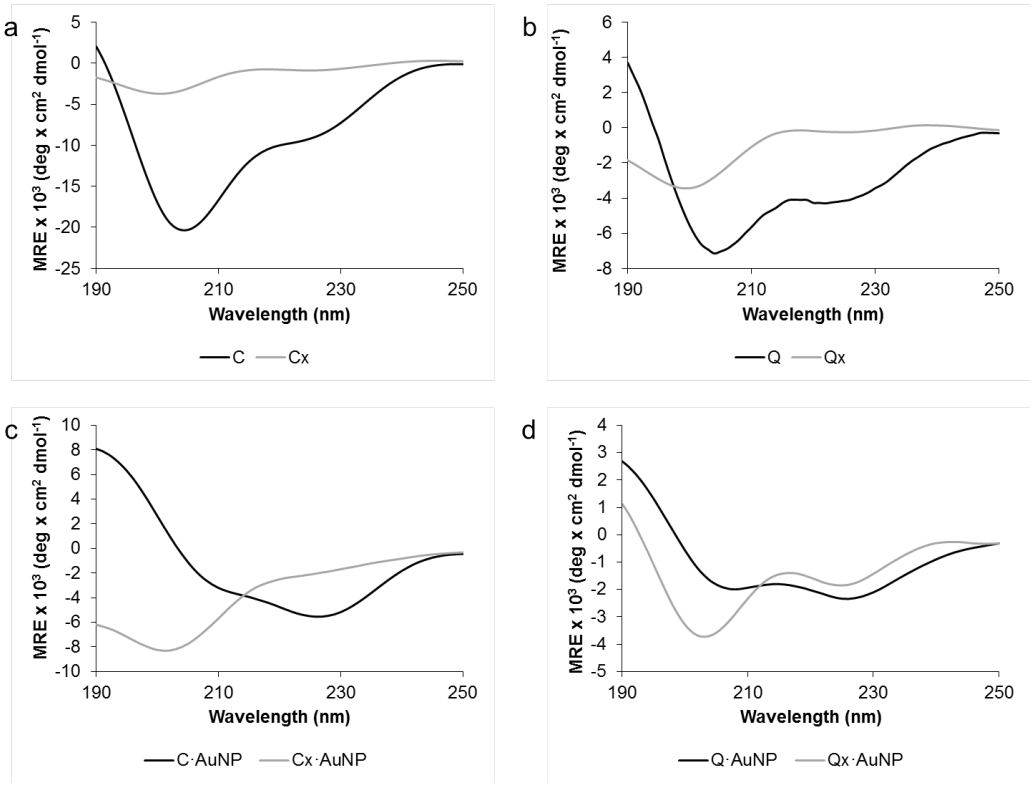



Figure 2.28. Secondary structure via circular dichroism, 50 mM PB pH 8 10 μM protein concentration. Wavelength scans of 6x-His and cleaved proteins with C, Cx in (a) and Q, Qx in (b). Wavelength scans of 6x-His and cleaved proteins with templated AuNPs with C·AuNP, Cx·AuNP in (c) and Q·AuNP, Qx·AuNP in (d).

Due to the recent debate as to the accuracy of CD in estimating protein secondary structures,[121] ATR-FTIR data was more heavily relied upon (Figure 2.29, Figure 2.30). Helical content, determined through peak deconvolution and fitting of ATR-FTIR data of C and Q (Figure 2.29, Figure 2.30, Table 2.4) prior to templation was 54 and 65 %, respectively. Upon AuNP templation, values of helical content according to FTIR analysis dropped to 30 % for both proteins (Figure 2.30, Table 2.4). For cleaved proteins Cx and Qx, helical content prior to templation was 37 and 45 %, respectively, and in the presence of AuNPs increased to 54 and 48 %, respectively (Figure 2.29, Figure 2.30, Table 2.4).

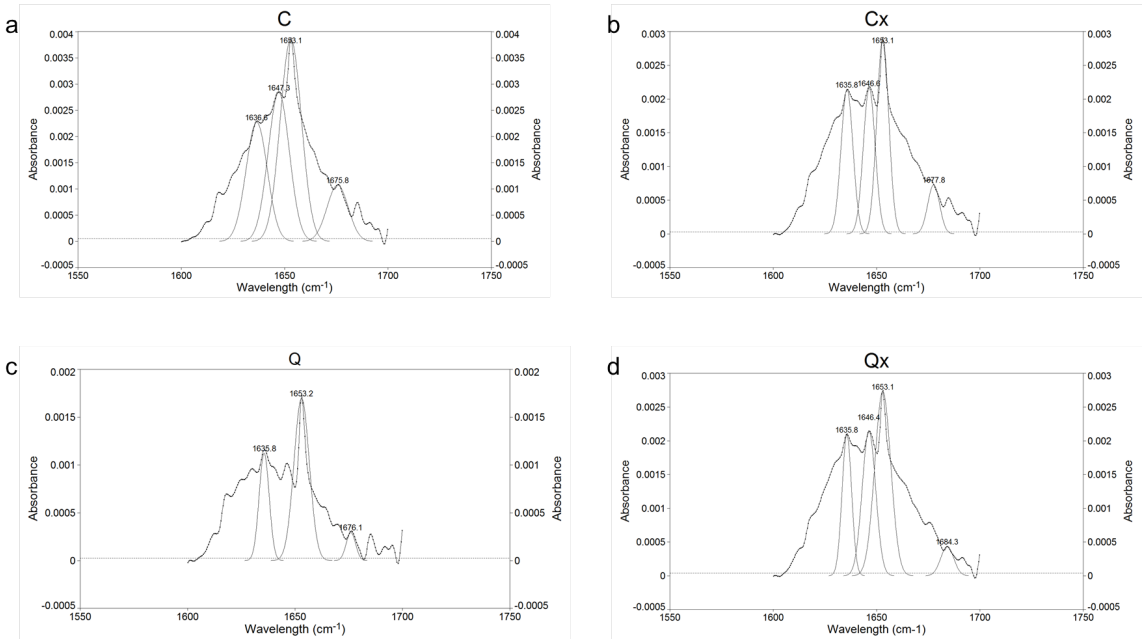



Figure 2.29. Fourier transform self-deconvoluted spectra of proteins in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials.

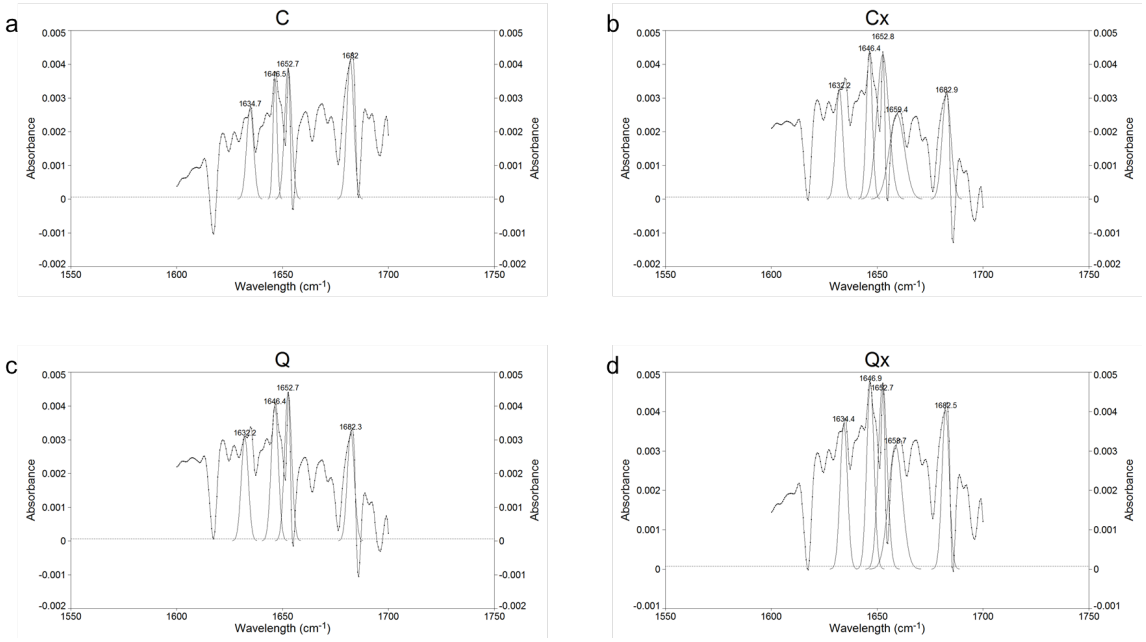



Figure 2.30. Fourier transform self-deconvoluted spectra of proteins after gold templation in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials.

ATR-FTIR and CD data both confirmed that helical content of Cx and Qx increased upon templation, whereas the helical content of uncleaved proteins C and Q decreased upon

AuNP templation (Figure 2.28 - Figure 2.30, Table 2.4). Changes in the secondary structure of the proteins that result from templation of AuNPs depended on the presence of the 6-His tag. This was consistent with findings of Naik *et al.* in which both $\alpha$-helical as well as unstructured proteins revealed conformational changes in the presence of AuNPs.[162]

### *2.4.5 Triggered assembly*

As C and Q exhibited a loss in helical content upon cleavage, we investigated AuNP templation on cleaved proteins after the addition of 25 v/v % trifluoroethanol (TFE). TFE is known to stabilize $\alpha$-helical structure in proteins, the propensity for which is not explicitly correlated to the helical content of the native protein.[124] CD data of C, Cx, Q, Qx in the presence and absence of Au with 25 v/v % TFE is presented in Figure 2.31.

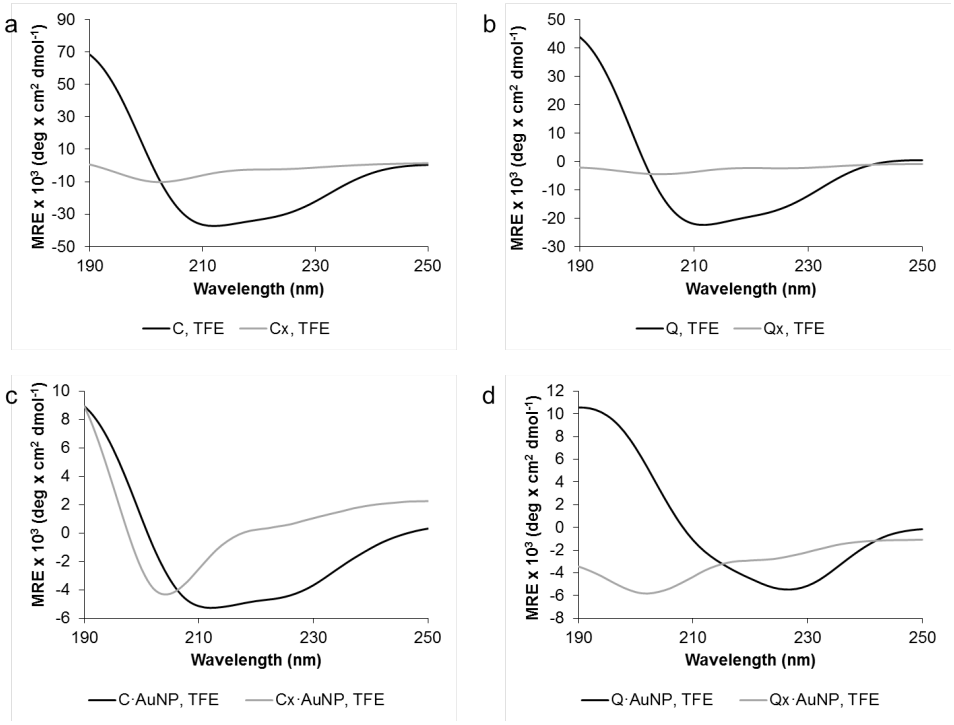



Figure 2.31. Secondary structure via circular dichroism, pH 8 10 µM protein concentration. Wavelength scans of 6x-His and cleaved proteins in the presence of 25 v/v % TFE, with C, Cx in (a) and Q, Qx in (b). Wavelength scans of 6x-His and cleaved proteins with templated AuNPs in the presence of 25 v/v % TFE, with C·AuNP, Cx·AuNP in (c) and Q·AuNP, Qx·AuNP in (d).

The wavelength scans in Figure 2.31 demonstrate a significant increase in helical character in certain proteins. The K2D method of secondary structure analysis allowed

for the quantification of $\alpha$-helical, $\beta$-sheet, and random coil content (described in Section 1.3.5.1).

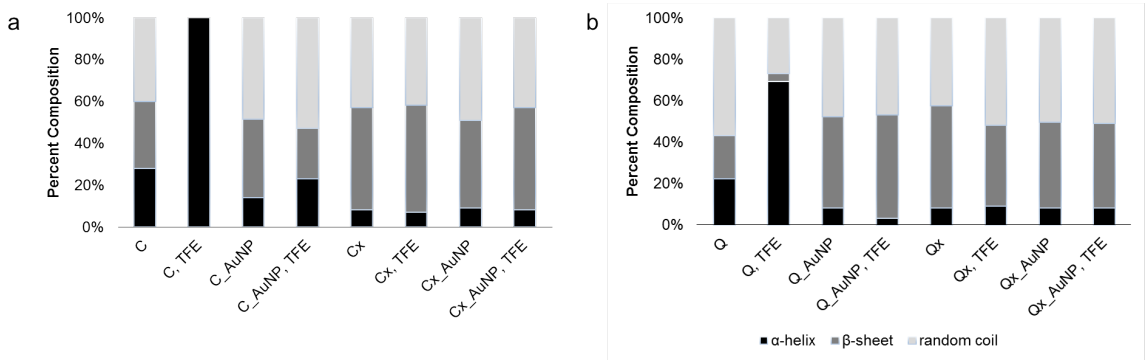



Figure 2.32. $\alpha$-helical (black), $\beta$-sheet (grey), and random coil (light grey) content for C and Cx in the presence and absence of 25 v/v % TFE and with and without templated AuNPs (a) and Q and Qx in the presence and absence of 25 v/v % TFE and with and without templated AuNPs (b). K2D analysis of data from Figure 2.28 and Figure 2.31.

A significant increase in $\alpha$-helical content was observed in C and Q upon the introduction of TFE (Figure 2.32). In the case of the cleaved proteins the CD data did not demonstrate a significant increase in the helical content in the presence of TFE, however observation of these assemblies in TEM revealed significant presence of insoluble composites, whose secondary structure characteristics are better defined by ATR-FTIR studies (Figure 2.34, Figure 2.35).

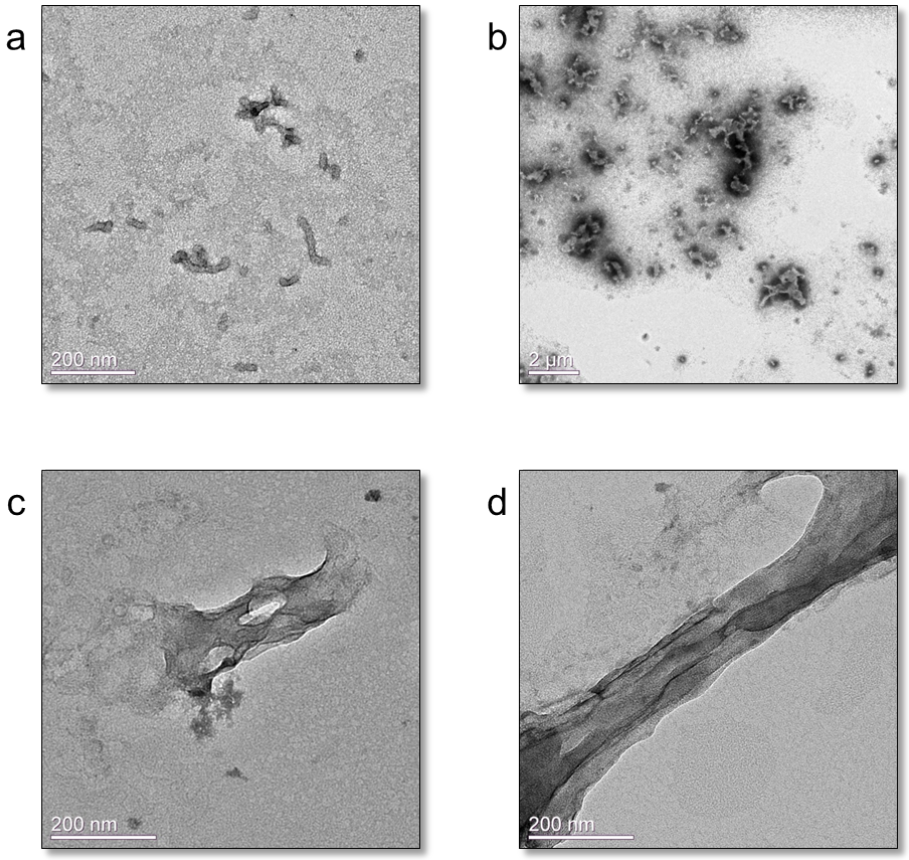



Figure 2.33. Transmission electron micrographs for proteins in the presence of 25 v/v % TFE, C + TFE (a), Q + TFE (b), Cx + TFE (c), and Qx + TFE (d). Scale bars are 200 nm in (a), (c), and (d) and 2 µm in (b).

Interestingly, C and Q in the presence of TFE manifested itself as aggregate formation when these assemblies were viewed in TEM (Figure 2.33a and b). Cx and Qx, on the other hand, formed sheets in the presence of TFE (Figure 2.33c and d), which are not dissimilar to the structures observed in the absence of TFE (Figure 2.19c and d). The results obtained by ATR-FTIR analysis of C, Q, Cx, and Qx in the presence of TFE revealed that the deconvoluted peaks in the 1648-1660 $cm^{-1}$ region corresponded to $\alpha$-helical character, and the smaller peaks in the 1675-1695 $cm^{-1}$ region corresponded to $\beta$-sheet character. ATR-FTIR results showed an increase in helical content for all proteins studied in the presence of TFE (Table 2.4, Figure 2.34, Figure 2.35). Upon addition of TFE, helical content increased to 68 and 71 % for C and Q, respectively, and 50 and 51 % for Cx and Qx, respectively.

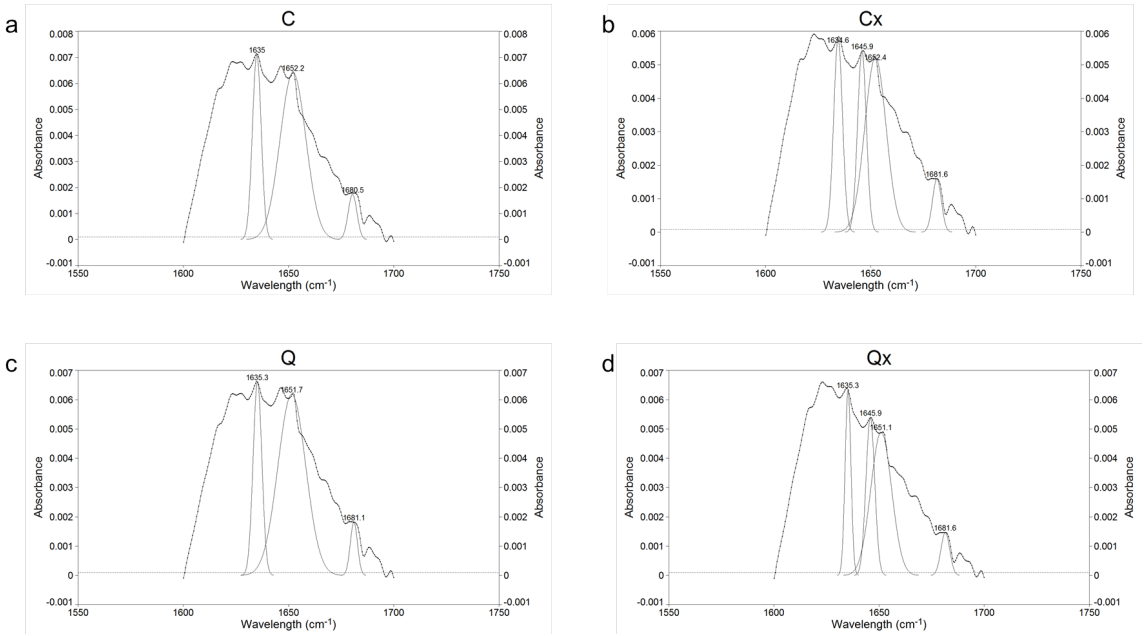



Figure 2.34. Fourier transform self-deconvoluted spectra of proteins in the presence of 25 v/v % TFE in 50 mM PB 8: C (a), Cx (b), Q (c), Qx (d). Each spectrum represents the average of two trials.

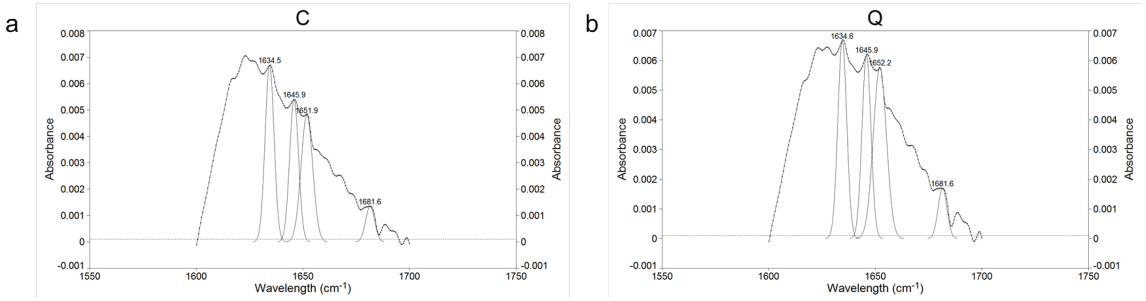



Figure 2.35. Fourier transform self-deconvoluted spectra of proteins in the presence of 25 v/v % TFE and after gold templation in 50 mM PB 8 for C (a) and Q (b). Data was not obtained for Cx·AuNP and Qx·AuNP with TFE due to problems with sample measurement. Each spectrum represents the average of two trials.

Proteins in the presence of TFE were then used for templation of AuNPs. Cleaved proteins Cx and Qx exhibited a very interesting change in morphology upon AuNP templation, which is illustrated by TEM micrographs in Figure 2.36. TFE induced an increase in helical content of the proteins, which in the presence of AuNPs, allowed for the formation of nanocomposite precipitates, reminiscent of those formed by 6-His proteins C and Q in phosphate buffer (Figure 2.23a and b). The precipitates were not as cohesive as those formed by 6-His proteins in phosphate buffer, as can be seen by the

presence of several fragmented precipitates in the lower micrograph of Figure 2.36b. Nonetheless, the ability of TFE to induce structural change on the secondary structure level that impacts the morphology of nanoparticle composites formed with said proteins is quite remarkable.

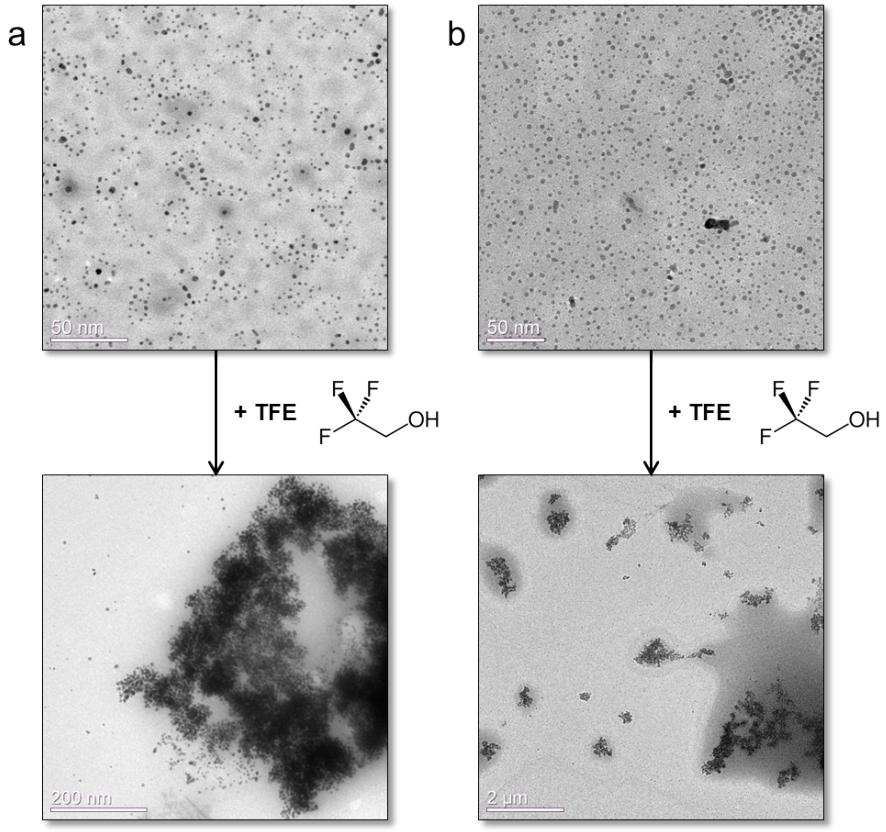



Figure 2.36. Transmission electron micrographs for cleaved proteins in the absence presence of 25 v/v % TFE after AuNP templation: Cx·AuNP without and with TFE (a) and Qx·AuNP without and with TFE (b). Scale bars are 50 nm in top micrographs in (a) and (b), 200 nm in lower (a) and 2 µm in lower (b).

The increase in structure of the cleaved proteins Cx and Qx, as demonstrated by ATR-FTIR results, enabled further assembly and the formation of aggregates as opposed to the soluble AuNPs generated in the absence of TFE (Table 2.4, Figure 2.36). This supported work by Higuchi *et al.*, who demonstrated that the peptide·AuNP assemblies could be controlled by modifying peptide secondary structure.[169] Thus, an external trigger such as TFE led to conformational changes that impacted the macromolecular assembly of protein·AuNP complexes.

Table 2.4. Secondary structure conformation based on ATR-FTIR data for 10 μM protein in 50 mM PB pH 8 before and after templation of AuNPs, and in the presence of 25 v/v % TFE. Percent composition was determined from relative areas of peaks fit to spectra. Data was not obtained for Cx·AuNP and Qx·AuNP with TFE due to problems with sample measurement.

| | Conformation | | |
|---|---|---|---|
| **Protein** | **α helix** | **β sheet** | **random coil** |
| C | 54 | 46 | 0 |
| C·AuNP | 30 | 54 | 16 |
| C + TFE | 68 | 32 | 0 |
| C + TFE, AuNP | 32 | 41 | 27 |
| Q | 65 | 35 | 0 |
| Q·AuNP | 30 | 43 | 27 |
| Q + TFE | 71 | 29 | 0 |
| Q + TFE, AuNP | 37 | 36 | 27 |
| Cx | 37 | 36 | 27 |
| Cx·AuNP | 54 | 30 | 16 |
| Cx + TFE | 50 | 29 | 21 |
| Qx | 45 | 26 | 29 |
| Qx·AuNP | 48 | 32 | 20 |
| Qx + TFE | 51 | 25 | 24 |

### *2.4.6 Electrochemical Characterization of Protein·AuNP Assemblies*

In order to evaluate the resistivity, and hence conductivity of protein·AuNP assemblies, electrochemical impedance and cyclic voltammetry (CV) measurements have been performed in the presence of the $Fe(CN)_6^{3-/4-}$ redox pair on a glassy carbon electrode.[195] Uncleaved C·AuNP and Q·AuNP films spontaneously adsorb onto the electrode surface, whereas AuNPs templated with cleaved protein remain soluble.

#### 2.4.6.1 Electrochemical impedance spectroscopy

Faradaic impedance spectra (or Nyquist plots) are generated for all protein·AuNP assemblies (Figure 2.38, Figure 2.39) as well as proteins in the absence of AuNPs (Figure 2.37, Figure 2.38). Nyquist plots consist of a semi-circular portion, corresponding to the electron-transfer-limited process and a linear portion that corresponds to diffusion-limited electron-transfer.[196] The diameter of the semi-circular portion corresponds to the

electron-transfer resistance, $R_{et}$, and is helpful for characterizing electrochemical properties of surface-modified electrodes.

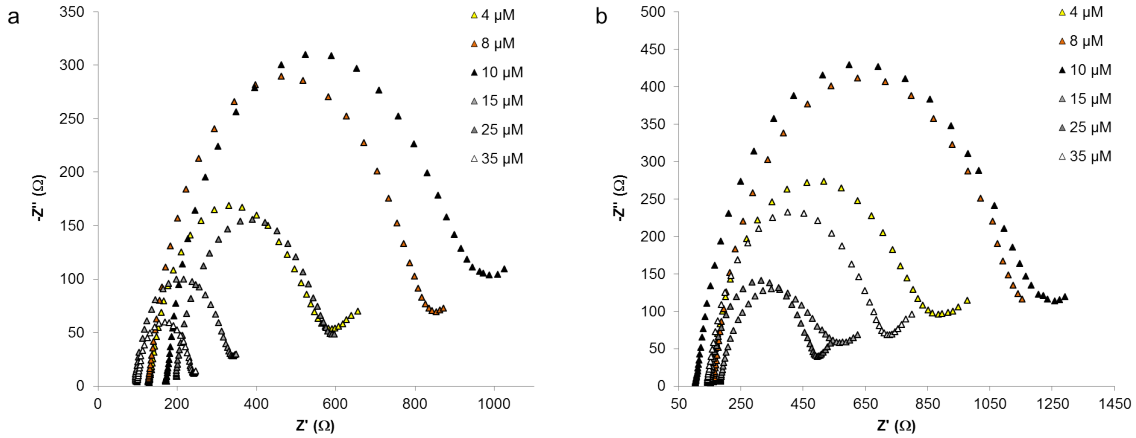



Figure 2.37. Nyquist plots obtained for 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (4, 8, 10, 15, 25, 35 µM) of proteins C (a) and Q (b).

Electrochemical impedance spectroscopy measurements for C and Q in the absence of AuNPs are presented in Figure 2.37. Maximum $R_{et}$ values were obtained at a protein concentration of 10 µM for both C and Q, equaling 409.4 and 603.7 Ω, respectively. The electrochemical resistance values for C and Q in the absence of AuNP at concentrations of 4, 8, and 10 µM were higher than $R_{et}$ values at the same concentrations for their cleaved counterparts, Cx and Qx (Table 2.5). It should be noted that the $R_{et}$ values from the Nyquist plots of the cleaved proteins at 10 µM with and without AuNPs do not change significantly; indicating that surface coverage of the electrode is not heavily affected by the presence of AuNPs (Figure 2.38). This was supported by absorbance and TEM data demonstrating AuNPs solubilized by Cx and Qx.

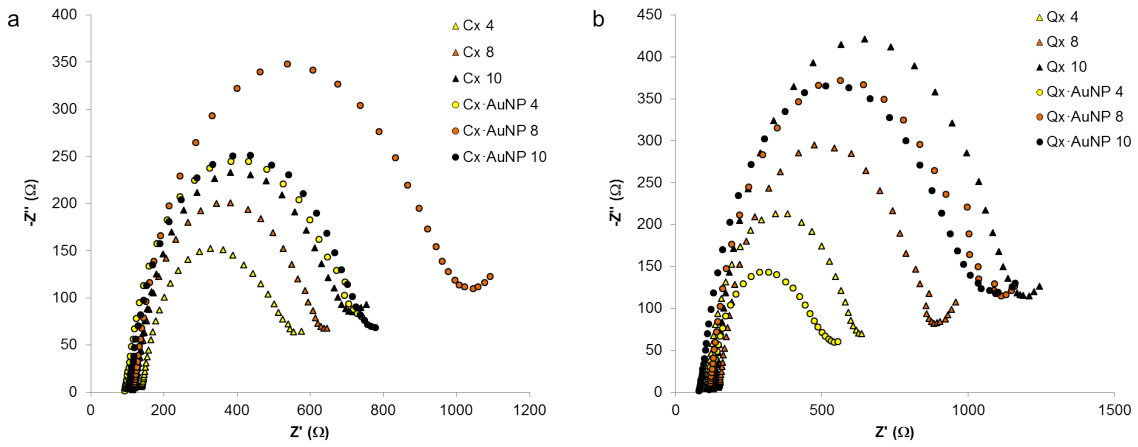



Figure 2.38. Nyquist plots obtained for 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (4, 8, 10, 12 µM) of proteins Cx (a) and Qx (b) in the absence (triangle markers) and presence (circle markers) of AuNPs.

Table 2.5. Electron transfer resistance ($R_{et}$) of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (0, 4, 8, 10, 15, 25, 35 µM) C, Q, Cx, and Qx as calculated from Nyquist plots. (Data was not obtained for cleaved proteins above 10 µM as higher protein concentrations were not achieved.)

| | $R_{et}$ (Ω) | | | |
|---|---|---|---|---|
| **Concentration (µM)** | **C** | **Q** | **Cx** | **Qx** |
| 0 | 28.4 | 28.4 | 28.4 | 28.4 |
| 4 | 266.7 | 395.5 | 259.4 | 281.0 |
| 8 | 367.0 | 558.8 | 300.3 | 403.2 |
| 10 | 409.4 | 603.7 | 374.4 | 541.3 |
| 15 | 138.1 | 193.3 | | |
| 25 | 234.8 | 203.7 | | |
| 35 | 114.5 | 310.1 | | |

Nyquist plots for 6-His proteins C and Q with templated AuNPs are presented in Figure 2.39. $R_{et}$ values increase until reaching a peak at 8 µM, with the largest resistance values attributed to C·AuNP and Q·AuNP at 812.0 and 902.1 Ω, respectively (Figure 2.39, Table 2.6). These values are nearly double the $R_{et}$ values observed at the same concentration in the absence of AuNPs (367.0 and 558.8 Ω for C and Q, respectively). The formation of micrometer sized precipitates that cover the electrode surface in the case of C·AuNP and Q·AuNP accounts for this significant increase in $R_{et}$.

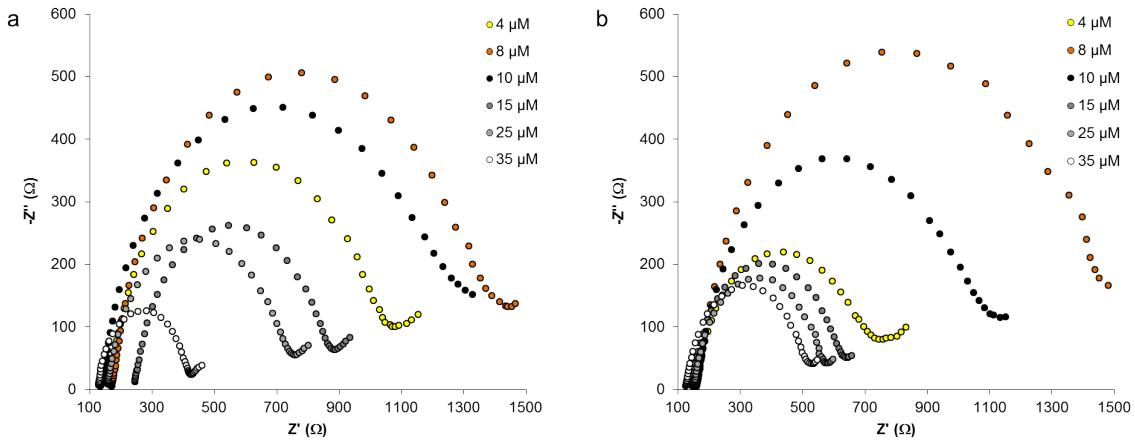



Figure 2.39. Nyquist plots for varying concentrations (4, 8, 10, 15, 25, 35 μM) of proteins C (a) and Q (b) in the presence of AuNPs. All electrochemical measurements were performed in the presence of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes.

As colloidal gold multilayers undergo an insulator-conductor transition at critical particle density,[199] the decrease in $R_{et}$ above 8 μM indicates that surface coverage on the electrode is less than one monolayer in thickness below this concentration. Above 10 μM, the uncleaved protein·AuNP films display $R_{et}$ values that stabilize in the 200-400 Ω range (Table 2.6). Similar $R_{et}$ values have been reported for the functionalization of electrode surfaces with DNA stabilized AuNPs.[152] The same trend of increasing and decreasing $R_{et}$ as concentration increases has been observed with chitosan and citrate-stabilized nanoparticles as well.[200]

Table 2.6. Electron transfer resistance ($R_{et}$) of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes with varying concentrations (0, 4, 8, 10, 15, 25, 35 μM) C, Q, Cx, and Qx in the presence of AuNP as calculated from Nyquist plots. (Data was not obtained for cleaved proteins above 10 μM as higher protein concentrations were not achieved.)

| | $R_{et}$ (Ω) | | | |
|---|---|---|---|---|
| **Concentration (μM)** | **C·AuNP** | **Q·AuNP** | **Cx·AuNP** | **Qx·AuNP** |
| 0 | 114.5 | 114.5 | 114.5 | 114.5 |
| 4 | 602.2 | 417.0 | 364.2 | 233.3 |
| 8 | 812.0 | 902.1 | 506.1 | 523.8 |
| 10 | 726.6 | 643.4 | 375.7 | 509.5 |
| 15 | 434.5 | 352.3 | | |
| 25 | 391.0 | 291.1 | | |
| 35 | 169.3 | 223.0 | | |

### 2.4.6.2 Cyclic voltammetry

CV is used to characterize surface coverage of a working electrode in relation to peak current, where a decrease in peak current corresponds to adsorption on to the electrode surface blocking detection of the redox reaction of $Fe(CN)_6^{3-/4-}$.[195] Hysteresis curves demonstrate intense anodic and cathodic peaks in the absence of protein (Figure 2.40a and b) diminish in intensity upon protein deposition (Figure 2.40c, Figure 2.41).

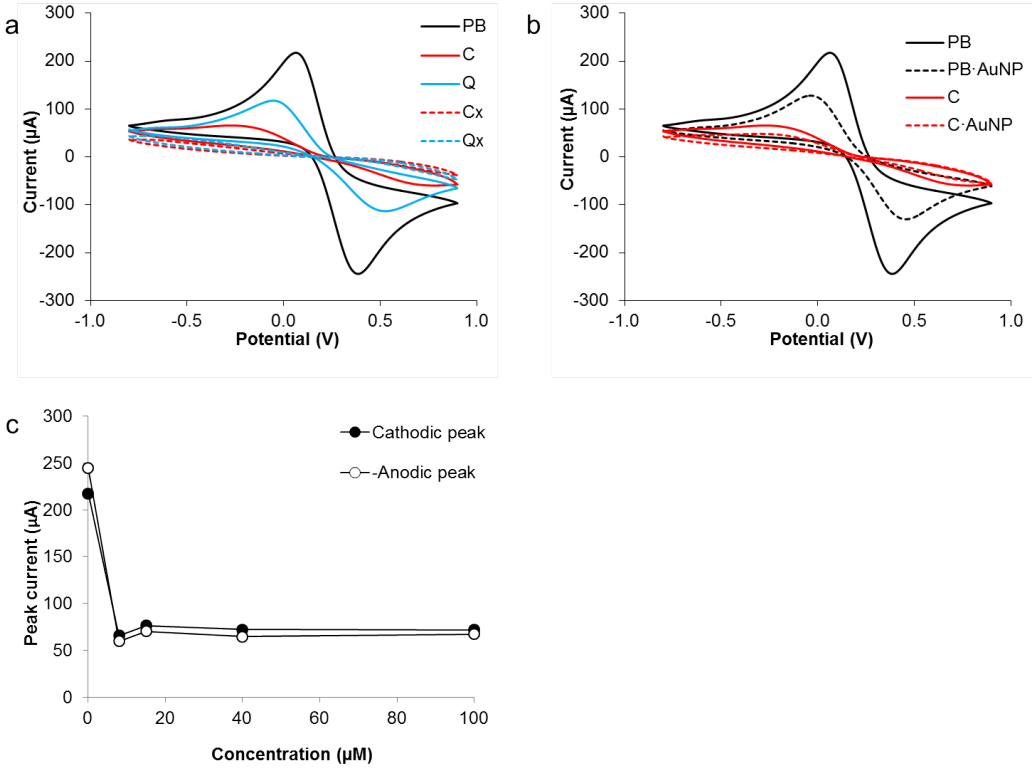



Figure 2.40. Cyclic voltammograms of PB (black), C (red), Q (blue), Cx (red dashed), and Qx (blue dashed) (a) and PB (black), PB_AuNP (black dashed), C (red), and C·AuNP (red dashed) (b) in the presence of 10 mM $K_4Fe(CN)_6 \cdot 3H_2O$ and 10 mM $K_3Fe(CN)_6$. Cathodic and –anodic peak currents of different concentrations of C in the absence of Au are plotted in (c).

In the absence of AuNPs, proteins exhibit a distinct level of adsorption onto the electrode surface, with cleaved proteins having a significantly lower peak current than uncleaved counterparts (Figure 2.40a, Figure 2.41). This can be explained as the absorption of a particular material onto the electrode surface is a complex function of hydrophobicity, electric charges of both the electrode and the protein, and protein conformation.[195]

Interestingly, uncleaved proteins exhibit a decrease in peak current upon the templation of AuNPs, whereas cleaved proteins possess an increase in peak current with

templated AuNPs. Proteins C and Q show a decrease in peak current equivalent to 17.4 μA and 9.0 μA, respectively (Figure 2.41). This data confirms that proteins bearing 6-His tags template AuNPs, generating dense films that adsorb onto the electrode surface, decreasing the electron transfer in the redox reaction of $Fe(CN)_6^{3-/4-}$ ions at the electrode surface.

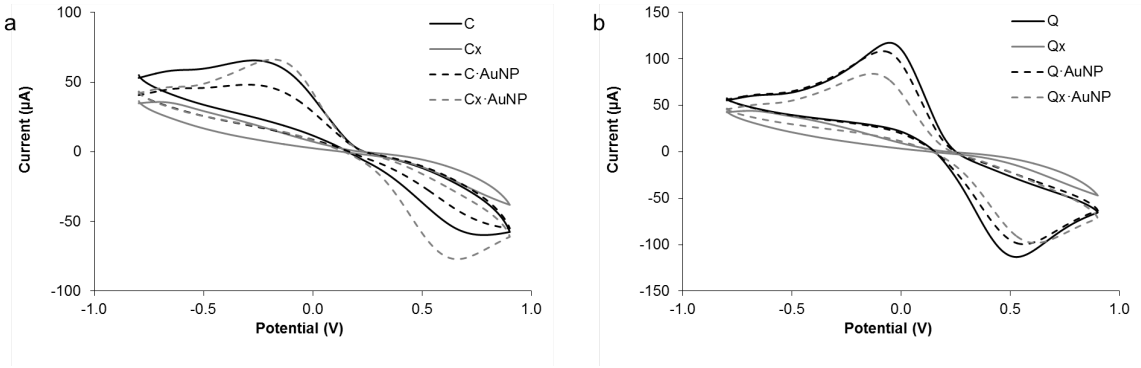



Figure 2.41. Cyclic voltammograms of 8 mM C, Cx, C_AuNP, and Cx_AuNP (a) and 8 mM Q, Qx, Q_AuNP, and Qx_AuNP (b). All electrochemical measurements were performed in the presence of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB pH 8 on glassy carbon electrodes.

Cleaved proteins Cx and Qx exhibit an increase in peak current, equivalent to 30.3 μA and 39.9 μA, respectively (Figure 2.41). In the case of cleaved proteins, protein-solvated AuNPs in the bulk volume increase the overall electron transfer of the iron ions at the electrode surface. These CV and electrical impedance measurements indicate that protein assemblies of structured proteins can be used to form a tunable kinetic barrier on the electrode surface.

## 2.5 Conclusions

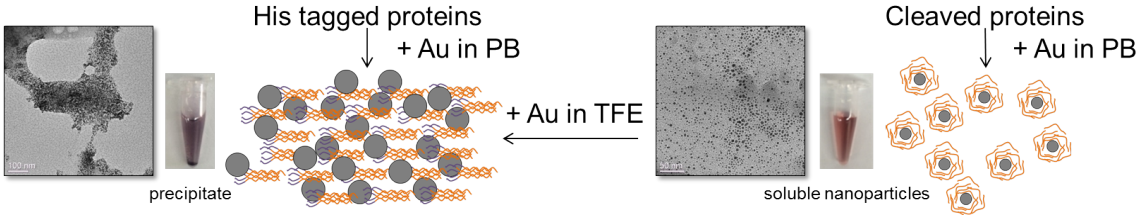



Figure 2.42. Protein nanoparticle film formation results from protein self-assembly, dictated by structure. Templation of gold nanoparticles on structured proteins leads to film formation via aggregation, while unstructured proteins solvate nanoparticles. In the presence of TFE, structure is induced in cleaved proteins allowing solvated nanoparticles to aggregate. Electrochemical properties indicate that these materials can be suitable for biosensor development.

Previous examples of protein·AuNP nanocomposites with the ability to switch assembly conformation have been constructed by the conjugation of protein to prefabricated AuNPs.[61,170,171,173,182,183] We have shown the ability to fabricate protein·AuNP nanocomposites *in situ*, leading to very distinct assembly morphology, dependent upon the protein secondary structure. Templation by unstructured peptides frequently results in soluble AuNPs which do not aggregate, whereas templation by structured protein allows for self-assembly, causing protein·AuNP nanocomposites to aggregate.[164,169] The ability of C and Q proteins to sequester NPs from the solution is strong, but the exact mechanism behind this AuNP sequestration has not been elucidated. In addition to the histidine tags, other functional groups such as primary amines, carboxylates, sulfur, and aromatic rings also dictate interactions with gold.[201] By manipulating protein structure via external triggers, such as TFE, we obtain control over the macromolecular conformation of these nanocomposites (see schematic in Figure 2.42). Such external triggers can lead to switching from unstructured to structured assemblies. These hybrid materials can be readily deposited on electrodes, where they can serve as a tunable bio-nanocomposite kinetic barrier.[202] Proteins that stabilize nanoparticles may prove important as agents for biodetection and imaging as biocompatible stabilizing agents for nanoparticles in aqueous solutions while formation of nanoparticle-embedded protein films may be of interest in catalysis and biosensing applications.[161,203]

# 3 Magnetite templation on functionalized protein biomaterials

## 3.1 Abstract

Coiled-coil proteins based on COMPcc have been used as a building block in the creation of functional materials capable of templating magnetic nanoparticles. Unnatural amino acid incorporation of an azide-bearing residue, L-azidohomoalanine (AHA), into proteins C and Q, whose sequences are based on COMPcc, offers a chemical handle upon which an orthogonal magnetite templating peptide, CMms6, is attached. Alkyne-functionalized propargylglycine CMms6 (prg-CMms6) is attached to the AHA-bearing proteins through a copper catalyzed click chemistry reaction and monitored molecular weight shifts in SDS-PAGE gel electrophoresis. Magnetic nanoparticles (MNPs) were synthesized through coprecipitation of $FeCl_3$ and $FeSO_4$ and reduced with NaOH to produce octahedral nanoparticles approximately 8.6 nm in diameter in the absence of prg-CMms6 and diameters ranging from 16.3 to 12.2 nm in the presence of prg-CMms6 concentrations ranging from 200 μM to 1 mM, respectively. Transmission electron microscopy confirmed that regular, octahedral MNPs were synthesized in the presence of prg-CMms6, whereas MNPs crystallized with only AHA protein resulted in MNPs of irregular morphology, including needle-shaped crystals. Whole cell lysate containing C+AHA and Q+AHA conjugated clicked with prg-CMms6 is used for templation of MNPs and studied with TEM. Crystallization and stabilization of MNPs by proteins is important in biomedical fields where magnetic materials can be made biocompatible or ligand-targeted with proteins in areas such as magnetic resonance imaging as well as targeted drug delivery through use of magnetic fields.

## 3.2 Introduction

### *3.2.1 Ferric oxide nanomaterials*

Synthesis of magnetic materials for applications in sensors, nanoelectronics, and biomedical imaging techniques is widespread, particularly those structured on the nanometer scale.[5] Function of all materials is highly correlated to structure, and the

ability to finely tune function is enabled through modification of chemical and physical properties of materials. In general, reduced-scale materials permit finer tuning. Alongside the surge in applications of magnetic materials comes the necessity to adjust the function of these materials, and hence, the structure. While magnetic materials, such as iron oxides, are abundant in nature,[204,205] at the current time industrial design and synthesis is necessary to achieve materials with nanoscale-defined features. Specific magnetic properties, sizes, shapes, and surface characteristics can only be obtained through defined methods of preparation. Industrial production of iron oxides employs gas phase, liquid phase, two-phase, sol-gel, mixed oxide sintering processes, and high pressure hydrothermal methods for generating magnetic materials.[206] While these methods produce suitable materials, they require the use of elaborate equipment, high temperatures (> 900 °C in gas phase methods and even 1500 °C in the mixed oxide process), elevated pressures, harsh precursors and are thus often quite costly.[207,208]

In an effort to create such magnetic materials while avoiding harsh processing methods, more natural approaches of fabricating highly structured magnetic nanomaterials are being explored.[209–211] Iron oxides are ubiquitous in nature, the most common representations being magnetite ($Fe_3O_4$), maghemite ($\gamma$-$Fe_2O_3$), and hematite ($\alpha$-$Fe_2O_3$).[209,212–214] Most biological magnetic systems contain magnetite as their source of magnetism, which is known to crystallize as an inverse spinel that is face-centered cubic in orientation (shown in Figure 3.1) with structural formula $Fe^{3+}(Fe^{2+}Fe^{3+})O_4$. Both magnetite and maghemite are superparamagnetic at room temperature, and have *hkl* planes that are indistinguishable when examined with X-ray or electron diffraction techniques (Table 3.1), as the *d*-spacings for each crystal structure are within fractions of an Ångström from one another.

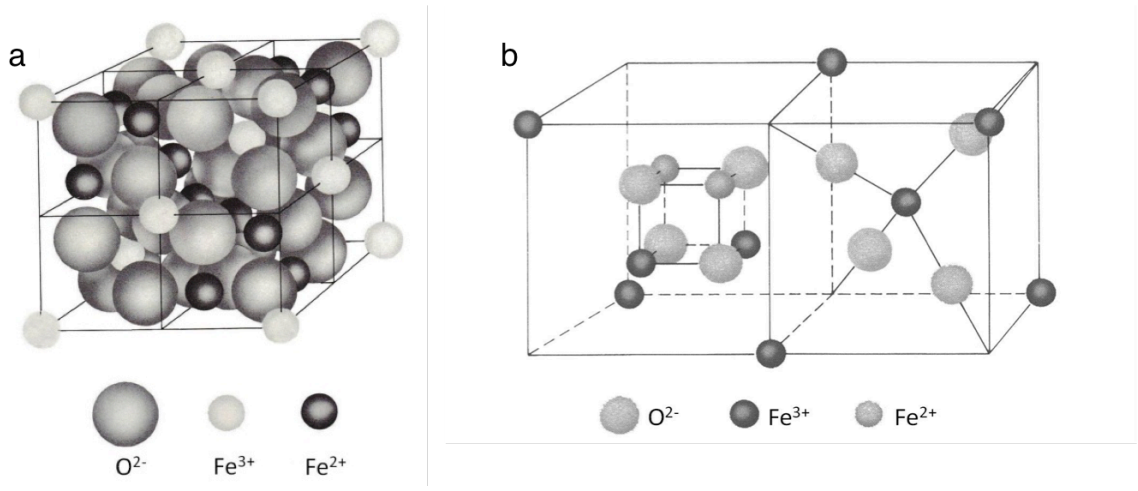



Figure 3.1: Ferrite unit cell structure (a), and an expanded view of the magnetite crystal structure, composed of $Fe^{2+}$, $Fe^{3+}$, and $O^{2-}$ atoms (b).[215]

Table 3.1. (*hkl*) planes and corresponding *d*-spacings for maghemite ($\gamma$-$Fe_2O_3$) and magnetite ($Fe_3O_4$).

| ***γ*-$Fe_2O_3$ reference** | **$Fe_3O_4$ reference** | **(*hkl*) planes** |
|---|---|---|
| 2.950 | 2.970 | (220) |
| 2.520 | 2.530 | (311) |
| 2.230 | - | (321) |
| 2.080 | 2.097 | (400) |
| 1.700 | 1.714 | (422) |
| 1.610 | 1.615 | (511) |
| 1.480 | 1.484 | (440) |

At room temperature, magnetite is classified as ferromagnetic,[208] meaning that the magnetic moments of iron atoms on different sublattices are opposed and unequal, due to the presence of two types of iron atoms ($Fe^{2+}$ and $Fe^{3+}$). These unequal and opposing magnetic moments cause a net moment to result in the material. The magnetic properties of magnetite nanocrystals are highly dependent on the crystal morphology as well as the

particle size (Figure 3.2).[208] Figure 3.2 presents the magnetization curves (or *M-H* curves, where $M = m/v$ represents the magnetization or magnetic moment $m$ per unit volume $v$, and $H$ is the strength of the magnetic field) for a hypothetical situation of ferromagnetic particles of various sizes in a blood vessel. As is shown in Figure 3.2b-e, the magnetic responses of particles is highly dependent on their size, and therefore the applications a particular ferromagnetic particle is suited for is wholly determined by its physical properties.[216,217]

### *3.2.2 Biologically generated magnetic particles*

Organisms capable of templation of magnetic particles have evolved the ability to specifically generate particles of a given size or crystal structure, dependent on the purpose, and the control of this outcome is determined in most cases by proteins. Biological systems that are able to control mineralization of magnetic particles at the genetic level are of great interest due to their high levels of specificity and the ability to tailor material properties through genetic engineering.[218]

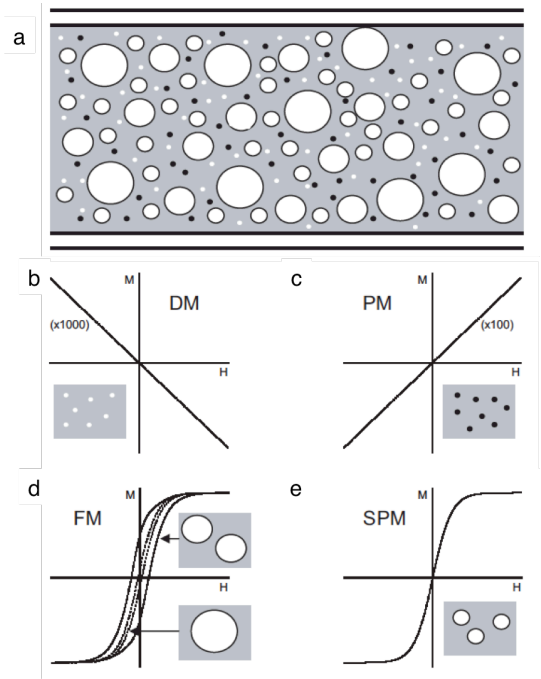



Figure 3.2. Magnetic response curves from a hypothetical situation where ferromagnetic particles ranging in size from nanometer to mesoscale are injected in a blood vessel (a). M-H curves are shown for diamagnetic (DM) (b), paramagnetic (PM) (c), ferromagnetic (FM) (d), and superparamagnetic (SPM) (e). In FM particles the response can either be multi-domain (dashed), single-domain (solid), or SPM, depending on the size of the particles.[219]

Many organisms have evolved the ability to template either diamagnetic or ferromagnetic materials within their structures.[204,212] The range of materials that various biological systems are capable of templating demonstrates an impressive level of control over composition, orientation, and morphology. Of particular interest to materials scientists is the ability of these biological systems to form highly regular nanocomposites typically between 10-30 nm[210] under mild reaction conditions and from natural precursor materials.[220]

#### 3.2.2.1 Iron oxide mineralization by ferritin

Ferritins represent a well-known class of proteins that are capable of mineralizing hydrated ferric oxide within the core of the protein assembly.[221,222] Ferritins are found in prokaryotic and eukaryotic organisms alike, with a shared structural similarity of 24 protein subunits that arrange themselves into a cage-like structure composed of heavy chains and light chains.[220,221] The ratio of heavy to light chains varies within organisms. Mineralization of ferrihydrite by ferritin has been thoroughly studied, and can be broken down in to a multi-step process consisting of Fe(II) oxidation, hydrolysis, nucleation, and crystal growth.[213,220,223] It has been shown that the reaction stoichiometry of Fe oxidation by $O_2$ depends on the ratio of Fe to protein. Low levels of Fe result in the stoichiometry given by Equation 3.1, where high levels of Fe result in the reaction represented in Equation 3.2.[224] Ferrihydrite are produced as a result of both reactions, but low levels of Fe also produce $H_2O_2$ as a byproduct.

$$2Fe^{2+} + O_2 + 4H_2O \rightarrow 2FeOOH + H_2O_2 + 4H^+$$

Equation 3.1

$$2Fe^{2+} + O_2 + 6H_2O \rightarrow 4FeOOH + 8H^+$$

Equation 3.2

Ferritin is also capable of synthesizing many different types of magnetic nanoparticles, dependent on the reaction conditions.[220,225] Magnetite ($Fe_3O_4$) can be synthesized by ferritin anaerobically under $N_2$ at elevated temperatures (60-65 °C) and pH 8.5.[213,220,226] Ferritin from the obligate anaerobe and hyperthermophilic archaeon *Pyrococcus furiosus* has been shown to be incredibly thermostable, with a half-life of 48 h at 100 °C and no

detectable metling temperature observed up to 120 °C.[226] Some of the other nanoparticles that may be synthesized by ferritin are shown schematically in Figure 3.3.

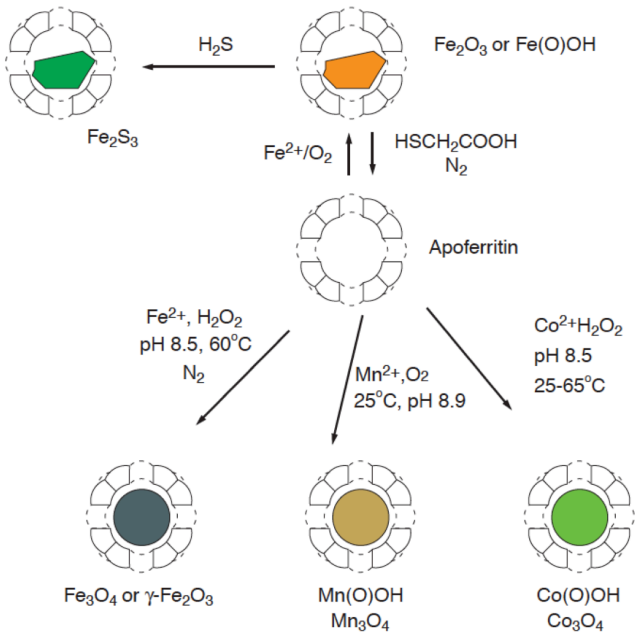



Figure 3.3. Nanoparticles produced by synthetic pathways involving ferritin in constrained reaction environments.[220]

### 3.2.2.2 Iron oxide mineralization by magnetotactic bacteria

One particularly interesting type of biomineralization takes advantage of the earth's magnetic fields to allow certain species to navigate along these lines or simply align themselves accordingly.[224,227,228] Nano-scale biogenic magnetite particles are formed as a basis for geomagnetic field perception in migrating birds, as well as in bacteria (Figure 3.4a). Magnetotaxis performed by bacteria is associated with aerotactic sensory mechanisms; bacteria can locate and migrate to low oxygen environments required for metabolism from the orientation of intracellular magnetic particles.[143] Magnetic bacteria, including a species named *Magnetospirillum magneticum* (Figure 3.4b), synthesize magnetic nanoparticles 20-30 nm in diameter composed of either magnetite ($Fe_3O_4$) or greigite ($Fe_3S_4$) that are contained in inclusion bodies called magnetosomes.[142,229] These magnetosomes are arranged in a chain (Figure 3.4c) that stretches from one end of the bacterial cell to the other, making it possible for the cell to align itself when a magnetic force is present.[143] Each particle, in fact, possesses a magnetic moment, and magnetic

interactions between these particles in the same chain allows the cell to have a large collective magnetic moment enabling them to sense and migrate along geomagnetic field lines (Figure 3.4d).

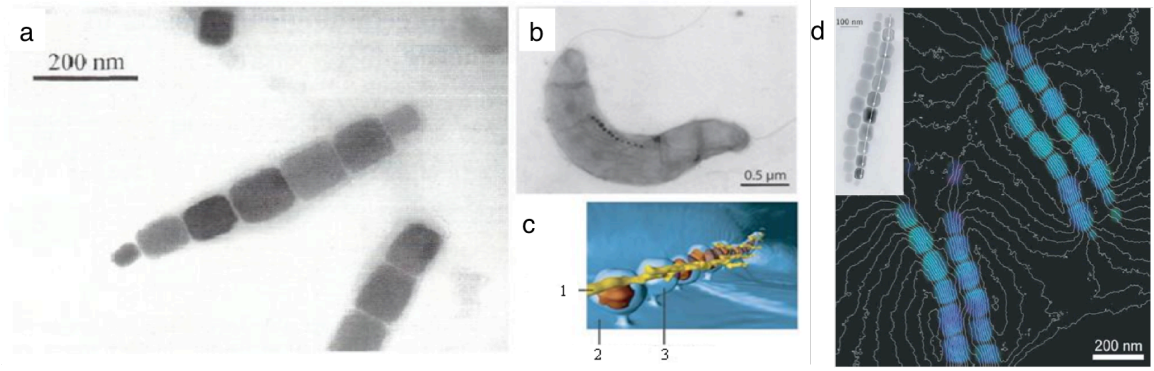



Figure 3.4: Magnetic bacteria and magnetosome chain formation. *Magnetospirillum magneticum* magnetic particles arranged in a chain with an average size of 50-100 nm (a),[230] electron micrograph of the magnetic bacterial cell (b), and a schematic view looking down the interior axis of a *Magnetospirillum magneticum* where 1 - cytoskeletal filament, 2 - inner membrane, and 3 – magnetosome (c).[143] Electron hologram of two double chains of magnetite in a single magnetotactic bacteria, where magnetic field lines are oriented such that the vector points toward the northwest in the image (d).[220]

The ability of magnetic bacteria to take advantage of magnetic fields in order to optimize their oxygen uptake has spurred research into the mechanisms of how bacteria are able to synthesize magnetic particles. Proteins found tightly bound to bacterial magnetic particles isolated from *Magnetospirillum magneticum* strain (Mms) AMB-1 included several low molecular weight proteins, named Mms and numbered accordingly.[229] These proteins expressed on the membrane of bacterial magnetic particles included Mms5, Mms6, Mms7, and Mms13.[229] Mms6, a particularly tightly-bound protein, was seen to direct the crystallization of magnetite.[218,229,231,232] The Mms6 amino acid sequence:

GGTIWTGK**GLGLGLGLGLGAWGPIILGVVGAGA**VYAYMKSRDIESAQSDEEVELRDALA

contains a hydrophobic domain (bold) and a hydrophilic domain (highlighted in grey) capable of capturing metal ions and interacting with the mineral phase, directing the synthesis of magnetite precipitates of cuboidal morphology 20 to 30 nm in diameter.[229] This magnetosome-associated protein possesses a hydrophobic N-terminus and a

hydrophilic C-terminus, and has been shown to bind strongly with $Fe^{3+}$, with a $K_d$ of Mms6 for $Fe^{3+}$ determined to be $10^{-16}$ M.[205] Recombinant, purified Mms6 has been shown to replicate the magnetite crystals (average size of 20 ± 4 nm) extracted from magnetic bacteria, maintaining cuboidal morphology and particle dimensions (as seen in Figure 3.5) through partial oxidation of ferrous hydroxide.[218] Control experiments studying the crystallization of MNPs by Mms6, mammalian iron-storage protein ferritin, and two proteins not known to template MNPs lipocalin and BSA revealed that Mms6 achieved the formation of uniform isomorphic MNPs while the other proteins each bound iron in a different manner.[233] A higher magnetic moment per particle was observed in the case of MNPs crystallized by Mms6, as indicated by the *M-H* curves in the presence of Mms6 reaching saturation levels in smaller fields (Figure 3.5c, d).

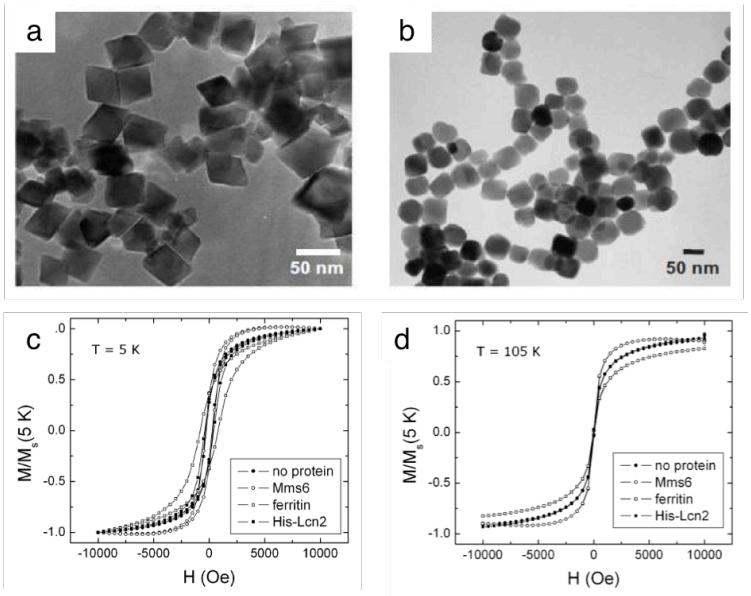



Figure 3.5: TEM images and magnetization curves of synthesized magnetic particles. Image (a) shows those synthesized by recombinant Mms6 while (b) contains particles extracted from *Magnetospirillum magneticum*. Particle dimensions were very similar in both cases.[218] Magnetization loops for Mms6 compared to ferritin, lipocalin, and no protein at 5 K (c) and 105 K (d).[233]

The C-terminal domain of Mms6 (YAYMKSRDIESAQSDEEVELRDALA), termed CMms6, is the portion of the protein in which almost all amino acid residues capable of chelating $Fe^{2+}$ or $Fe^{3+}$.[205] As a result, the C-terminal domain alone has been shown to promote the growth of superparamagnetic nanoparticles and is responsible for high affinity stoichiometric iron binding.[205] The C-terminal domain of Mms6 contributes to its

quaternary structure, as was discovered by Mallapragada *et al.* by increasing concentrations of CMms6 with respect to Mms6 and observing protein precipitation.[205] Their results suggested that CMms6 interacts on the surface of the micelle, allowing the authors to conclude that it is the N-terminal domain of Mms6 that anchors the protein in the magnetosome membrane *in vivo*, while the C-terminal domain interacts with iron to form magnetite nanoparticles.[205]

#### 3.2.2.3 Iron oxide templation by protein assemblies

As nature has shown us, magnetically active nanoparticles can be synthesized by several organisms, such as magnetotactic bacteria.[218,227] The biological machinery that is used by magnetotactic bacteria to direct the synthesis of these MNPs is largely due to protein assemblies, specifically the protein Mms6, as has been described in Section 3.2.2.2. Notably, even recombinantly expressed Mms6 is successfully in templating MNPs, binding tightly to MNPs and regulating their size to produce crystals of narrow size distribution.[218] As we have also seen in Section 3.2.2.1, ferritin assemblies are also responsible for templating iron oxide nanoparticles in a very regular fashion. In addition to these two examples provided by natural systems, work is currently being done to generate synthetic, protein-based systems for templation of iron oxide NPs.

One of these systems uses the tobacco mosaic virus (TMV) to template iron oxide nanoparticles. TMV is made up of 2130 identical protein building blocks that form a hollow tube 300 nm long by 18 nm wide with a 4 nm wide hollow channel on the inside.[234] Charged amino acid residues present on the polar surface of TMV has been used to direct the templation of iron oxyhydroxides, although the magnetic properties of these hybrid materials have not yet been characterized.[220] Microemulsions of human serum albumin-coated $Fe_3O_4$ MNPs radiolabeled with $^{188}Re$ for *in vivo* regional target therapy have been generated, resulting in protein-coated MNPs 200 nm in diameter.[235]

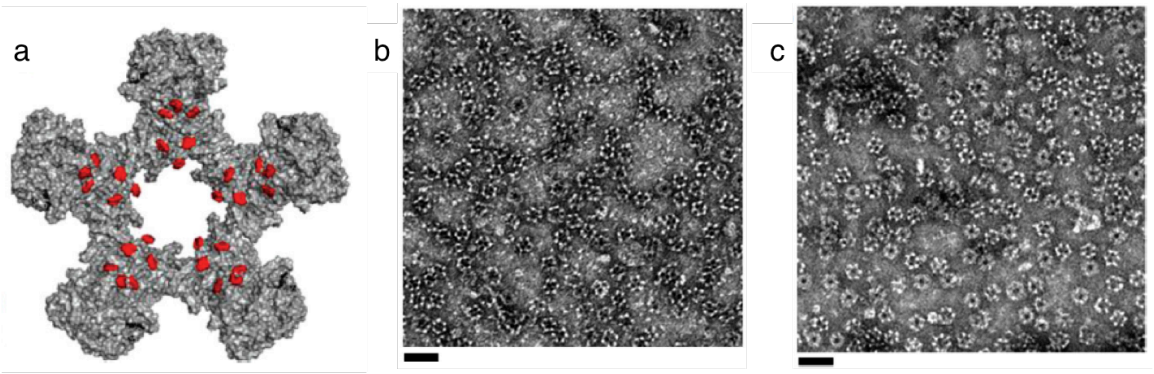



Figure 3.6. PyMOL representation of half of the E2 protein cage with 30 iron-binding sites on the interior highlighted in red (a) and assembled protein-MNP cage structures with E2-LFtn (b) and E2-LE6 (c).[225] Scale bars in (b) and (c) represent 50 nm.[225]

Another interesting example of engineered protein assemblies for MNP templation is the work of Peng *et al.*, who designed a protein cage from the E2 core protein capable of templating MNPs through incorporation of ferritin-mimicking peptides (sequence QEHEDE, referred to as E2-LFtn) or negatively charged glutamine residues (EEEEEE, referred to as E2-LE6).[225] E2 is part of a multi-enzyme complex from *Geobacillus stearothermophilus* with 60 identical subunits that self-assembles to form a hollow cage structure.[225] Iron oxide templation was directed through incorporation of either a ferritin-mimicking peptide or a series of glutamine residues within a loop region of the E2 protein.[225] Nanocomposite cages from either E2-LFtn or E2-LE6 were generated through coprecipitation of ferrous ions in the presence of protein and measured approximately 25 nm in diameter (Figure 3.6b and c).[225]

#### 3.2.2.4 Organization of inorganic nanoparticles by protein fibers

Protein fibers able to self-assemble to form organized structures under benign and ambient conditions are highly desirable scaffold materials for tissue engineering as well as in coatings for diagnostic devices.[236–238] There are many examples of protein fibers $\alpha$-helical in nature[236] or with $\beta$-sheet structure[22,158,166] assembling inorganic nanoparticles along the length of the protein fibers.

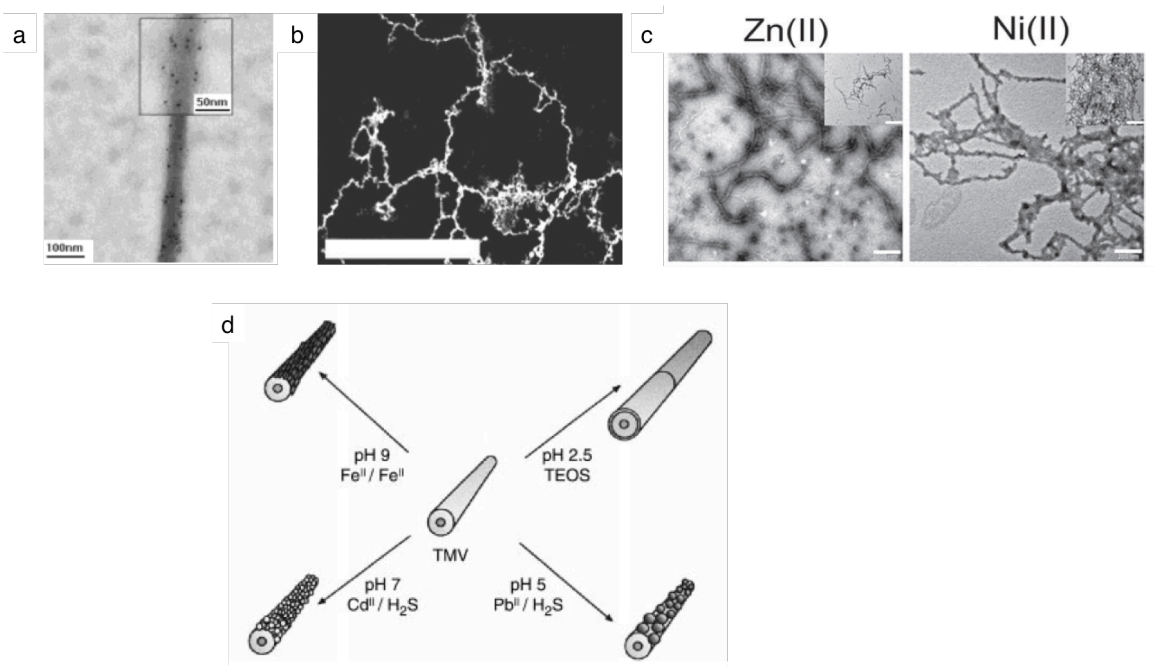



Figure 3.7. Examples of protein fibers organizing inorganic nanoparticles. TEM image of 5 nm functionalized streptavidin-gold nanoparticles on SAF-p1-biotin self-assembling fiber (a).[236] STEM image of FePt nanoparticles on engineered viral templates, where scale bar represents 350 nm (b).[228] TEM images of Zn and Ni nanoparticles templated on histidine tagged COMPcc protein fibers, where scale bars represent 200 nm in the main images and 1 μm in the insets (c).[88] Nanotube composites using the tobacco mosaic virus as the protein scaffold (d).[239]

There are several methods by which inorganic nanoparticles may be recruited and organized along the length of protein fibers to create protein nanowires. One of these methods is through affinity interactions between antibody-antigen compounds that can be used to functionalize the protein fibers and nanoparticles.[236] Woolfson *et al.* have used their SAF systems in order to compose protein fibers that are capable of templating gold nanoparticles through exploitation of the high affinity between streptavidin and biotin (Figure 3.7a).[236] These SAFs were unique in that they were conjugated with small molecules, such as biotin, which were present on the surface of the peptide fibers and used to recruit gold nanoparticles functionalized with streptavidin.[236] Another technique for directing the templation of inorganic nanoparticles on protein fibers includes the incorporation of specific amino acid sequences that are known to organize NPs.[218,228,233,240] Phage display was used by Reiss *et al.* to identify an amino acid sequence HNKHLPSTQPLA that selectively binds to FePt nanoparticles and fused this sequence to the gP3 protein of the M13 bacteriophage virus to direct templation of FePt

nanoparticles approximately 4 nm in diameter (Figure 3.7b).[228] Another example of using specific amino acids to direct nanoparticle templation on protein fibers includes the use of the TMV (Figure 3.7d).[234] Surface exposed charged amino acids direct biomineralization of iron, lead, and cadmium NPs on the surface of the tube under specific buffer conditions.[234]

One residue in particular has been used in many different applications for the templation of inorganic nanoparticles on protein fibers. Histidine tags consisting of several histidine residues are commonly used to template divalent metal nanoparticles.[88,160,181,241] Gunasekar *et al.* have used a 6-histidine tag on the N-terminal end of COMPcc to non-covalently template Zn and Ni nanoparticles (Figure 3.7c),[88] and Rosi and Matsui's groups have used histidine to template Au nanoparticles,[160,166,181,242] as has been described in Section 2.2.3.

### *3.2.3 Applications of magnetic nanomaterials*

The ability of naturally occurring systems to template magnetically active nanoparticles with defined morphological and physical properties has piqued the interest of materials scientists striving to design the next generation of functional nanoscale materials under benign reaction conditions, for which proteins hold the key. The ability of recombinant Mms6 to synthesize inorganic magnetic particles allows researchers to apply techniques in biological engineering to create highly functional nanostructured materials through the use of engineered proteins. By taking advantage of existing proteins that are well characterized the development of novel nanostructured magnetic materials is virtually limitless.

Iron oxides are of scientific and technological importance because of their widespread usage in magnetic materials, gas and environmental sensors, and lithium-ion batteries.[5] There are several examples of nanorods composed of $\alpha$-$Fe_2O_3$ being used in each of these applications. Wu *et al.* have discovered the dependence of hematite nanorod dimensions on charge/discharge responses for $Li^+$ batteries.[5] The dependence on the NP dimension is related to the number of holes on the hematite surface that allow for the introduction of fortein atoms or ions, such as $Li^+$.[5] NPs with higher surface areas can

effectively shorten the diffusion length of $Li^+$ ions, increasing capacity of the batteries with respect to potential (Figure 3.8a).[5] Hematite nanostructures, which are n-type semiconductors with electrical conductivity that is highly sensitive to gaseous environments, have also been employed in gas sensors for detection of formaldehyde, $H_2$, and ethanol (Figure 3.8b, c).[5,243]

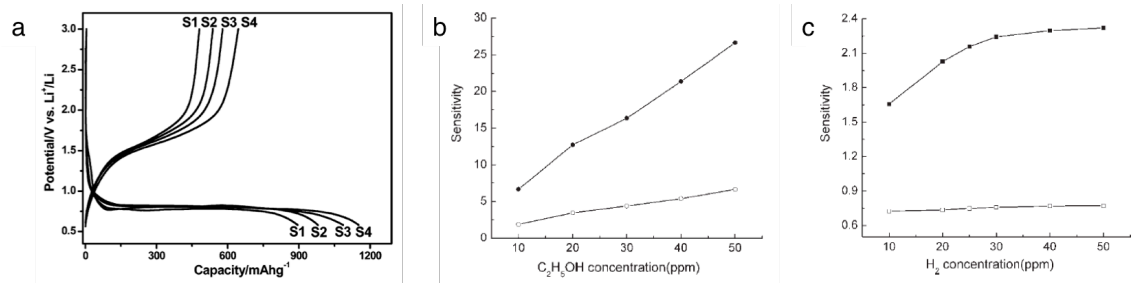



Figure 3.8. Charge/discharge curves for hematite samples of increasing surface area (S4>S3>S2>S1) at a current density of 0.2 mA/cm$^2$ (a).[5] Sensor sensitivity towards ethanol (b) and $H_2$ (c) for hematite nanotubes (filled markers) and nanoparticles (empty markers).[243]

As the current methods for producing synthetic magnetic materials is costly and requires harsh processing conditions,[208,244] the engineers of novel nanomaterials for electronics and sensors are turning towards biological systems. The use of biological molecules for the development of biosensors is a natural choice, due to their inherent biocompatibility and finely tuned properties via biological engineering.

Biological sensors use magnetic NPs to detect analytes in different matrices (ie. toxicity studies) without disturbance from scattering or absorption from surrounding tissues, where biocompatibility of the NPs is paramount (Figure 3.9).[245] Much attention is being focused on the use of nanocomposites for high sensitivity biosensor applications.[245,246] For example, spin-spin relaxation times of water $T_2$ signals can be detected by biocompatible magnetic nanoparticle assemblies by switching from dispersed to aggregated states (Figure 3.9a).[245] In this biosensor antigen was conjugated to MC-LR on MNPs and the presence of antibodies induced aggregation of the assemblies.[245] These types of biosensors are referred to as magnetic-relaxation switches (MRS).

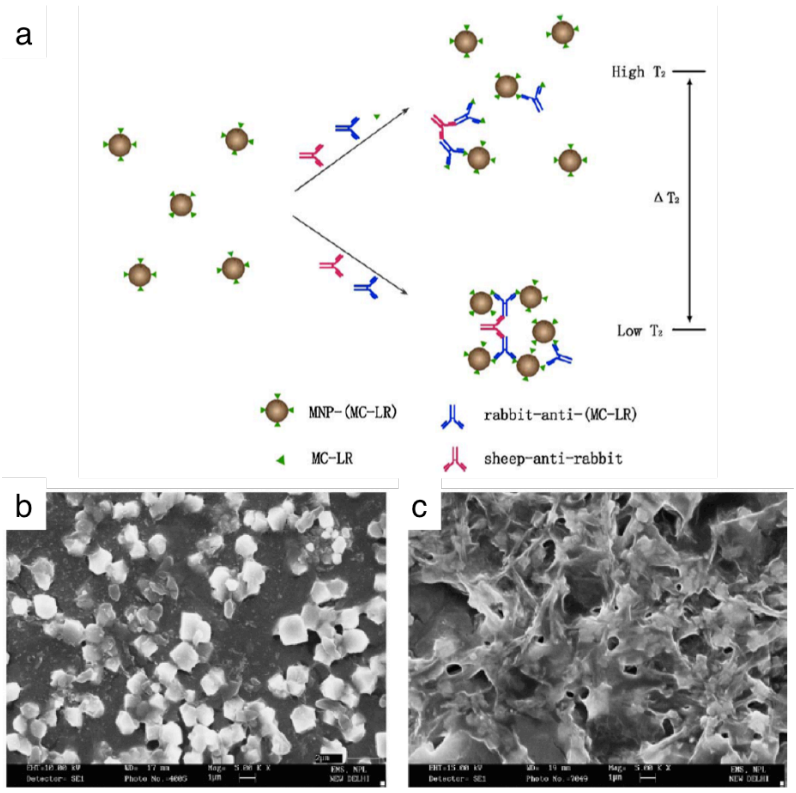



Figure 3.9. Schematic of biosensor detection of MC-LR through aggregation of magnetic nanoparticles (a).[245] SEM images of nanobiocomposite/ITO electrode composed of CH-$Fe_3O_4$ particles (b) and nanobiocomposite/ITO electrode composed of Ur-GLDH/CH-$Fe_3O_4$ (c).[246]

An iron oxide-chitosan biosensor has been developed by Kaushik *et al.* to detect urea using urease (Ur) and glutamate dehydrogenase (GLDH) coimmobilized onto a film composed of chitosan (CH) and MNPs.[246] CH-$Fe_3O_4$ nanocomposite demonstrated a globular and porous morphology (Figure 3.9b), while Ur-GLDH on CH-$Fe_3O_4$ had a more regular form (Figure 3.9c), leading the authors to conclude that high loading of Ur-GLDH was facilitated by the CH-$Fe_3O_4$ film.[246] The biosensor formed from these materials was shown to have high sensitivity towards urea and have negligible influence of interferents.[246]

There exists a wide range of applications for magnetic nanocomposite materials that can be synthesized via biomimetic routes, including materials for use in spintronics, magnetic inks, high-density magnetic memory devices, as magnetic seals in motors, as well as many interesting applications in biomedical fields.[233] Magnetic nanoparticle materials are relevant in biomedicine, including magnetic separation, drug delivery, and

as a contrast agent in magnetic resonance imaging (MRI).[219] MNPs are particularly well suited for these applications as they are small enough to be localized in desired regions within the body physiology and superparamagnetic properties allow them to travel easily within the body by application of local magnetic field gradients. Optimal sizes of these nanoparticles are between approximately 10-30 nm.[219,247] The proper stabilization of these particles is of upmost importance, as loss of magnetization properties or aggregation of the particles can be very problematic in *in vivo* scenarios.[248] Aggregation or precipitation of MNPs can result in cytotoxicity and rapid clearance of these particles from the body by the reticuloendothelial system, limiting their efficacy.[211,223]

Magnetic separation is used in biological systems to isolate particular biological molecules from their native solutions so that they can be analyzed or concentrated for other purposes.[223] The targeted isolation of a particular biomolecule from a mixture with MNPs is usually achieved through coating or functionalization of the MNP with biocompatible molecules such as dextran and phospholipids such that the functionalized MNPs can recognize and bind with the biomolecule of interest and sequester it upon application of a magnetic field. The two-step process that takes place in magnetic separation is therefore (1) tagging the molecule of interest with the magnetic particle and (2) separating the tagged molecules through a fluid-based magnetic separation device.[219] These types of MNPs have been used for magnetic separation in binding immunospecific agents to red blood cells, lung cancer and urological cancer cells, bacteria, and Golgi vescicles.[219] The size of MNPs can be controlled to range from a few nanometers to tens of nanometers. This means that they can therefore achieve the same dimension range of many relevant biological entities including cells (10-100 μm), viruses (20-450 nm), proteins (5-50 nm), and even genes (2 nm in width and 10-100 nm in length).[219] Biological MNP nanocomposites can therefore be designed to coat MNPs with certain biomolecules or have the MNPs and biological entity of the same size such that they can interact side by side.[219]

Coated MNPs have been used to immobilize a variety of enzymes, antibodies, oligonucleotides, as well as histidine-tagged proteins (Figure 3.10).[223,249] Research efforts in this area are currently focused on increasing MNP-ligand binding specificities and obtaining mechanisms for reversible and controllable attraction of MNPs within fluid

biological systems.[249] Hung *et al.* have used the technique of magnetic separation to perform purification of histidine-tagged proteins by immobilizing copper ions (represented as small diamonds in Figure 3.10) onto the surface of MNPs that show a high affinity towards histidine residues.[250] Only the target protein displaying the histidine tag will interact with the copper-functionalized MNPs, enabling impurities within the cell lysate to be washed away.[250] The target protein was then released through addition of imidazole.[250]

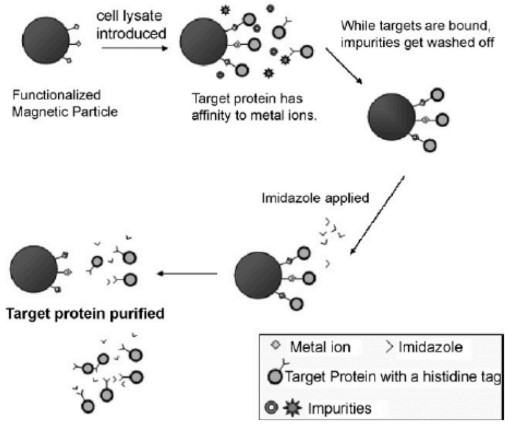



Figure 3.10. Purification schemes for magnetic separation of histidine tagged proteins using functionalized MNPs.[250]

One of the major concerns with administering chemotherapy treatment is its nonspecificity. Targeted drug delivery has the potential to selectively deliver therapeutic agents to the site of a diseased tissue without affecting surrounding healthy tissues.[251,252] The principle behind using MNPs for targeted drug delivery is similar to magnetic separation, where functionalized nanoparticles can be directed to a particular site with the aid of a magnetic field gradient. Ferrofluids containing MNPs used in drug delivery are typically administered intravenously, and properties determining the efficacy of these treatments include material-specific properties of the MNPs such as hydrodynamic radius, biocompatibility, and circulation time.[253] Ferrofluids are colloidal suspensions of magnetic particles that are susceptible to manipulation by magnetic fields while maintaining fluidity.[253] Coatings for the MNPs can be loaded with cytotoxic drugs that can then be released upon a given trigger that modifies the chemistry of the particular coating material.[211]

Magnetic resonance imaging relies on the difference between the very small magnetic moment of a proton and the incredibly large number of protons found in tissue, which in the presence of a very strong magnetic field will have a measureable effect.[219] Despite this naturally occurring magnetic moment observed in human tissue, relaxation times necessary to observe these effects can be somewhat lengthy. In order to reduce these relaxation times, superparamagnetic (SPM) nanoparticles are frequently used.[247] Magnetic nanoparticles are multidomain materials defined by their response to an externally applied magnetic field.[209] Size reductions of magnetic materials result in superparamagnetic materials, which are single domain materials due to size restictions and are inherently nonmagnetic but become magnetized in the presence of a magnetic field.[209] The effect of using SPM particles on the relaxation time is demonstrated in Figure 3.11.

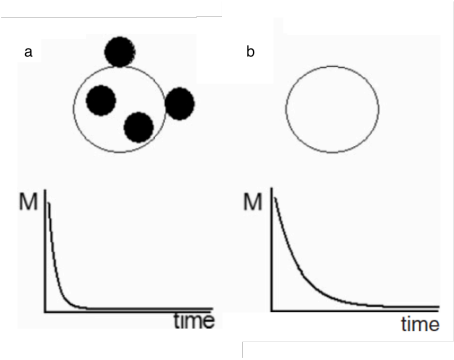



Figure 3.11. Effect of internalization of SPM MNPs in cells on MRI relaxation times. Protons in cells with MNPs exhibit shorter relaxation times (a) than those without (b).[219]

MRI contrast relies upon the differential uptake of MNPs by different tissue types. The use of magnetic nanoparticles as contrast agents in MRI imaging largely relies on the ability to generate MNPs of uniform size and dispersity; MNPs with diameters > ~30 nm are rapidly taken up by the liver and spleen, while MNPs < ~10 nm are not easily recognized and travel easily throughout the vasculature.[219,247] The best size range of MNPs for imaging purposes is therefore between 10-30 nm.[219,247] Currently much work is being done to develop biologically functionalized MNPs that may be used as contrast agents in MRI that are dually capable of targeting specific tumors or drug delivery as a result of biological specificity.[251]

### *3.2.4 Design methodology*

Herein we describe the design methodology employed in the attempt to generate self-assembling, functional proteins that bear an orthogonal magnetite-binding group to template and order inorganic nanoparticles. Proteins C and Q (described and characterized in detail in Sections 1 and 2) have been selected as proteins to act as templated for MNP assembly in this work as they can self-assemble to form organized fibers[1] and are also amenable to functionalizion of orthogonal peptides through incorporation of unnatural amino acids. The goal of this work is to develop protein assemblies that can order and crystallize MNPs under ambient conditions.

#### 3.2.4.1 Incorporation of azidohomoalanine

In order to attach the magnetite binding peptide CMms6 to the C and Q proteins, an unnatural amino acid analog L-azidohomoalanine (AHA) is incorporated into the proteins via residue-specific incorporation to generate C+AHA and Q+AHA, respectively (Figure 3.12). While not present in nature, the azide group has recently emerged as a highly useful tool in chemical biology and bioconjugation as a result of its high reaction specificity with alkyne moieties, in particular.[254] This highly useful synthetic amino acid residue has been synthesized via direct azidification (Section 3.3.2) of protected 2-aminobutanoic acid (Figure 3.17).

As is shown in Figure 3.12, there are two methionine residues in the sequences of C and Q. Methionine residues will be replaced by AHA residues in C+AHA and Q+AHA. The total distance of each wild type protein is approximately 73 Å.[95] Based on this approximation, the distance between methionine residues in C and Q would be 27 and 41 Å, respectively, when assembled.

| | His tag | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg |
|---|---|---|---|---|---|---|---|
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* |
| C | MRGS**HHHHHH**GSIEGR | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* |
| C+AHA | **$A_{HA}$**RGS**HHHHHH**GSIEGR | APQ**$A_{HA}$**LRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | *1* | *17* | *20* | *27* | *34* | *41* | *48* |
| Q | MRGS**HHHHHH**GSIEGR | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL |
| | *1* | *17* | *20* | *27* | *34* | *41* | *48* |
| Q+AHA | **$A_{HA}$**RGS**HHHHHH**GSIEGR | VKE | ITFLKNT | APQ**$A_{HA}$**LRE | LQETNAA | LQDVREL | LRQQSKL |

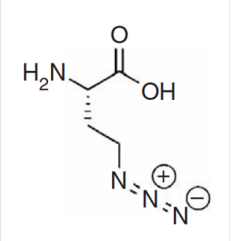
AHA 

Figure 3.12. Chemical structure of azidohomoalanine (AHA) and sequences of C+AHA and Q+AHA.

#### 3.2.4.1.1 Sensitivity of AHA to photolytic decomposition

It has been recently reported that alkyl azides are prone to photodecomposition under conditions as mild as exposure to laboratory light at room temperature.[255] As imagined, photosensitivity of AHA has ramifications on chemical reactivity and biological properties of azides. Degradation of azides by light exposure results in aldehyde and monoacyl aminal as the major products.[255] Both these degradation products have highly reactive functional groups capable of contaminating azide transformations in chemical and biological applications, which combined with the reduced amount of active azide left to react with, can have a significant impact on the yield of a particular reaction with an azide group as a reactant.[255] In light of this photosensitivity, extreme precautions must be taken when working with AHA to prevent its exposure to light and ensure its chemical reactivity remains uncompromised.

#### 3.2.4.1.2 Methionine auxotroph M15MA

Methionine auxotroph strain M15MA can be used for protein expression in the presence of AHA via residue-specific incorporation.[256] Residue-specific incorporation involves an exchange of a natural amino acid with a specific unnatural amino acid supplemented in

the growth medium, using an organism's translation system to replace a natural amino acid with a desired analog (Figure 3.13).[257] Auxotrophic *E. coli* strains, such as M15MA, allow engineers to expand the structural diversity of protein-based materials with the addition of functional unnatural amino acids capable of engaging in a large number of synthetic reactions.[257] M15MA is incapable of heterotrophic synthesis of methionine and thus any methionine or methionine analogs must be introduced systematically for complete protein expression in M15MA cells. In the incorporation of synthetic amino acids, the aminoacyl-tRNA synthetases (MetRS in the case of methionine and its analogs) mediates the acceptance or rejection of specific amino acids to be coupled with a given tRNA.[257] Incorporation efficiency of AHA in *E. Coli* M15MA cells has been reported to be between 70-90 %, and in the case of incorporation into the murine dihydrofolate reductase (mDHFR) enzyme reaching effiencies approaching 100 %.[256,258] Residue specific incorporation takes advantage of the fact that these aaRSs ligate to a specific amino acid with high fidelity based on the identity of the tRNA.[259] MetRS is able to readily incorporate AHA into proteins in place of methionine when culture media is depleted of the natural amino acid analog as AHA is very similar in size to methionine, thereby not putting any strain on the binding pocket of MetRS.[260,261] AHA is activated by wild type MetRS, but not all methionine analogs are as easily accepted by MetRS as the endogenous enzyme may not be able to incorporate the bulkier moieties.[262] Engineered MetRS mutants may be introduced into expression hosts in order to accommodate for bulkier residues, which is critical for efficient expression of recombinant proteins with a wide range of unnatural amino acids.[259]

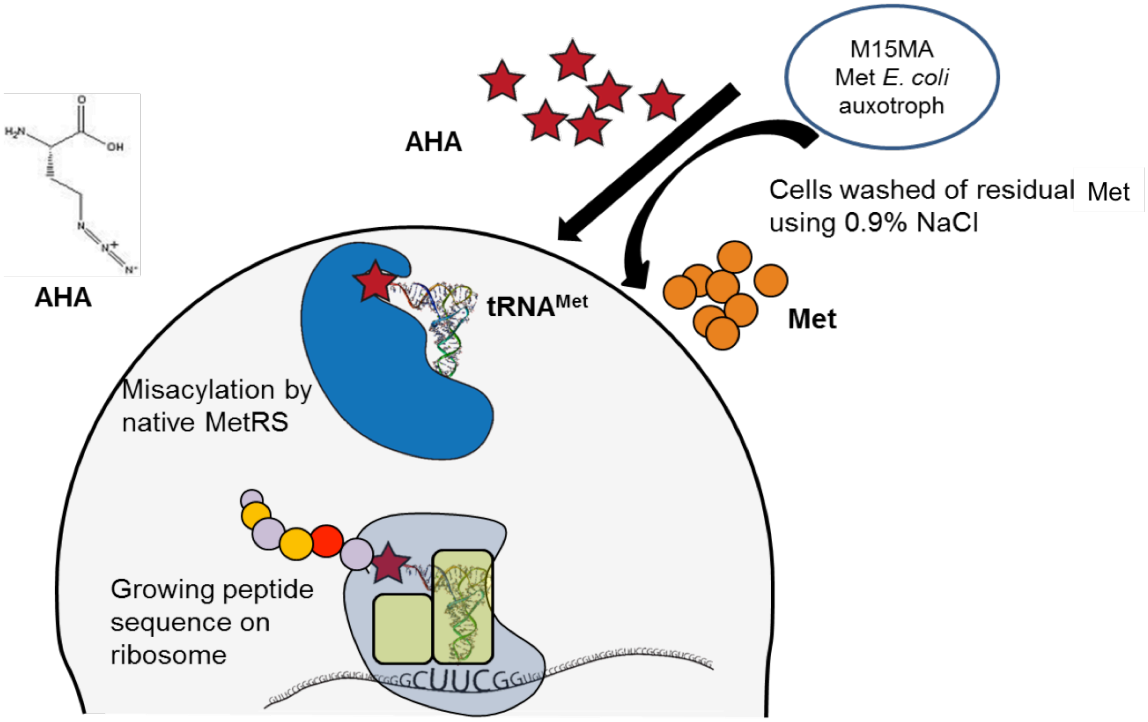



Figure 3.13. Incorporation schematic of unnatural amino acid AHA by methionine auxotrophic M15MA cells. Uptake of AHA by the cells is accomplished by washing of native methionine with 0.9 % NaCl and subsequent introduction of AHA into the culture media. MetRS binds to AHA permitting tRNA to incorporate and translate the amino acid analog.[263]

### 3.2.4.2 Click chemistry

Many groups have used click chemistry to selectively bind ligands to peptides and proteins functionalized with AHA.[258,264–266] Methionine auxotrophic *E. Coli* cells have been used to incorporate azide-functionalized amino acids into cell-surface glycoproteins and subsequently used in Staudinger ligation to attach molecular probes to the cell's surface.[267] This was remarkable as it was observed that the azide group was able to survive the cell's metabolism and selectively react with alkyne ligands under mild conditions.[267] AHA has been identified by Kiick *et al.* as an excellent methionine surrogate, being efficiently incorporated into the mDHFR enzyme then used in the Staudinger ligation with a phosphine bearing an antigenic FLAG peptide (Figure 3.14a).[258] Beatty *et al.* have incorporated AHA and homopropargylglycine into newly synthesized protein populations to then label these populations with reactive rhodamine, dimethylaminocoumarin, and bodipy dyes, according to the scheme shown in Figure 3.14b.[264]

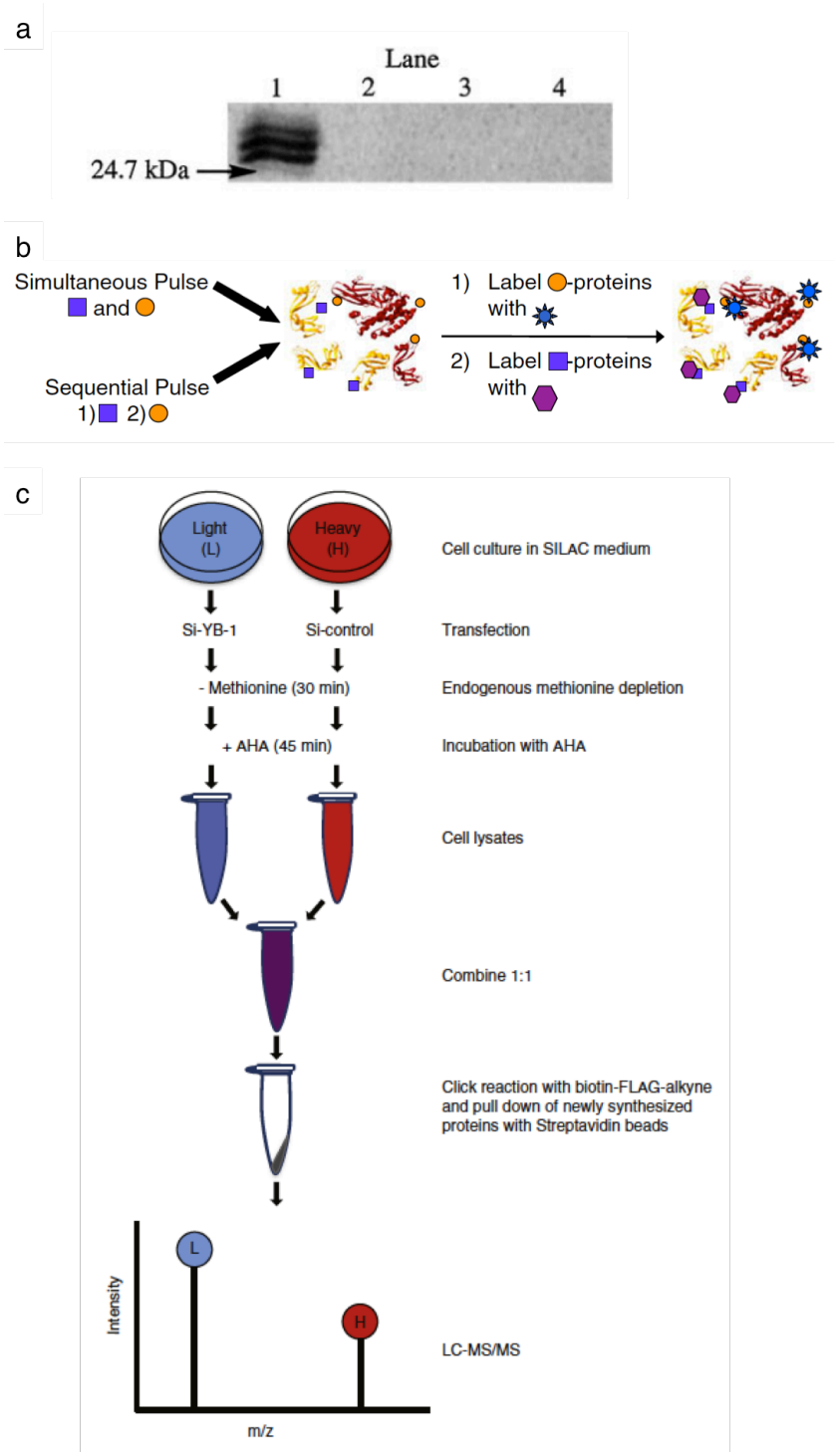



Figure 3.14. Examples of AHA-incorporated proteins used in click chemistry reactions. Western blot analysis of the products of the Staudinger ligation, where Lane 1 is mDHFR + AHA and anti-FLAG peptide, Lane 2 is mDHFR + AHA and no peptide, Lane 3 is mDHFR + met and anti-FLAG, and Lane 4 is mDHFR + met and no peptide (a).[258] Two-dye labeling strategy for protein identification and imaging, where purple squares and orange circles represent the functionalized amino acids (b).[264] Schematic of the approach for identification of newly synthesized proteins translationally regulated by YB-1 (c).[265]

Y-box binding protein-1 (YB-1), a protein that is involved in cancer progression, has been modified with AHA and click chemistry in order to identify proteins that are translationally regulated by YB-1.[265] Cells expressing YB-1 were labeled with either light or heavy isotopologs and were then cotranslationally pulsed with AHA (Figure 3.14c).[265] By lysing the cells and subjecting their contents to a click chemistry reaction with an alkyne biotin tag followed by affinity purification, Somasekharan *et al.* were able to quantify newly synthesized proteins that are regulated by YB-1 in only 45 minutes.[265] In general, for proteomic applications click chemistry proves to be highly selective and efficient.[266]

Inspired by the organisms that are naturally able to template iron oxide nanoparticles (Section 3.2.2), click chemistry will be used to add the C-terminal (magnetite binding) end of the Mms6 peptide bearing a propargylglycine (prg) group, termed prg-CMms6 (sequence prg-YAKMKSRDIESAQSDEEVELRDALA), to the constructed AHA functionalized proteins (C+AHA and Q+AHA). In this click chemistry reaction, AHA acts as the linker molecule covalently attaching the N-terminus of the prg-CMms6 peptide to the AHA residue(s) within C+AHA and Q+AHA (Figure 3.15). The hypothetical amount of prg-CMms6-bearing proteins will depend on the incorporation levels of AHA and the success of the click chemistry reaction. Alkyne and azide groups have been employed extensively as useful and highly reactive yet specific bio-orthogonal moieties in biomacromolecules such as proteins.[268] As azido and alkynyl groups are entirely orthogonal to endogenous chemistries of naturally occurring functional groups in proteins,[269] they are highly useful in establishing controlled bioconjugation.

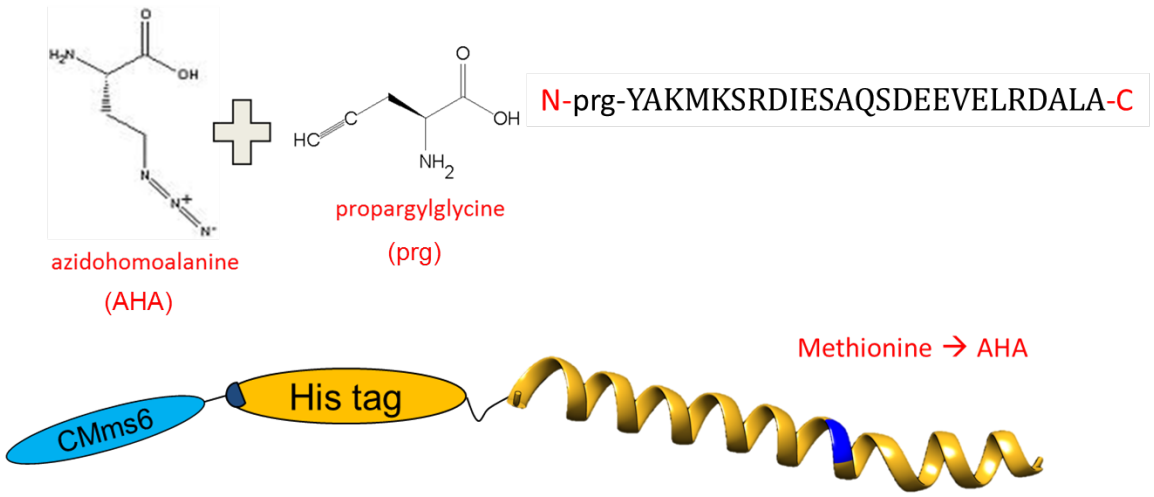



Figure 3.15. Schematic overview of functionalization of C+AHA and Q+AHA with an alkyne-functionalized prg-CMms6 via click chemistry.

Click chemistry provides numerous advantages compared to other reactions designed to covalently attach peptides, including high chemo- and regioselectivity, stability at room and physiological temperatures, relatively simple reaction conditions and facile product isolation.[270] These reactions are typically simple to perform with high yields and relatively few byproducts.[271] 1,3-dipolar cycloaddition of azides and alkynes is thermodynamically favorable (-50 to -65 kcal/mol) and also highly selective, as azide and alkyne units are inert to most chemical functionalities and stable in a wide range of solvents, temperatures, and pHs.[271]

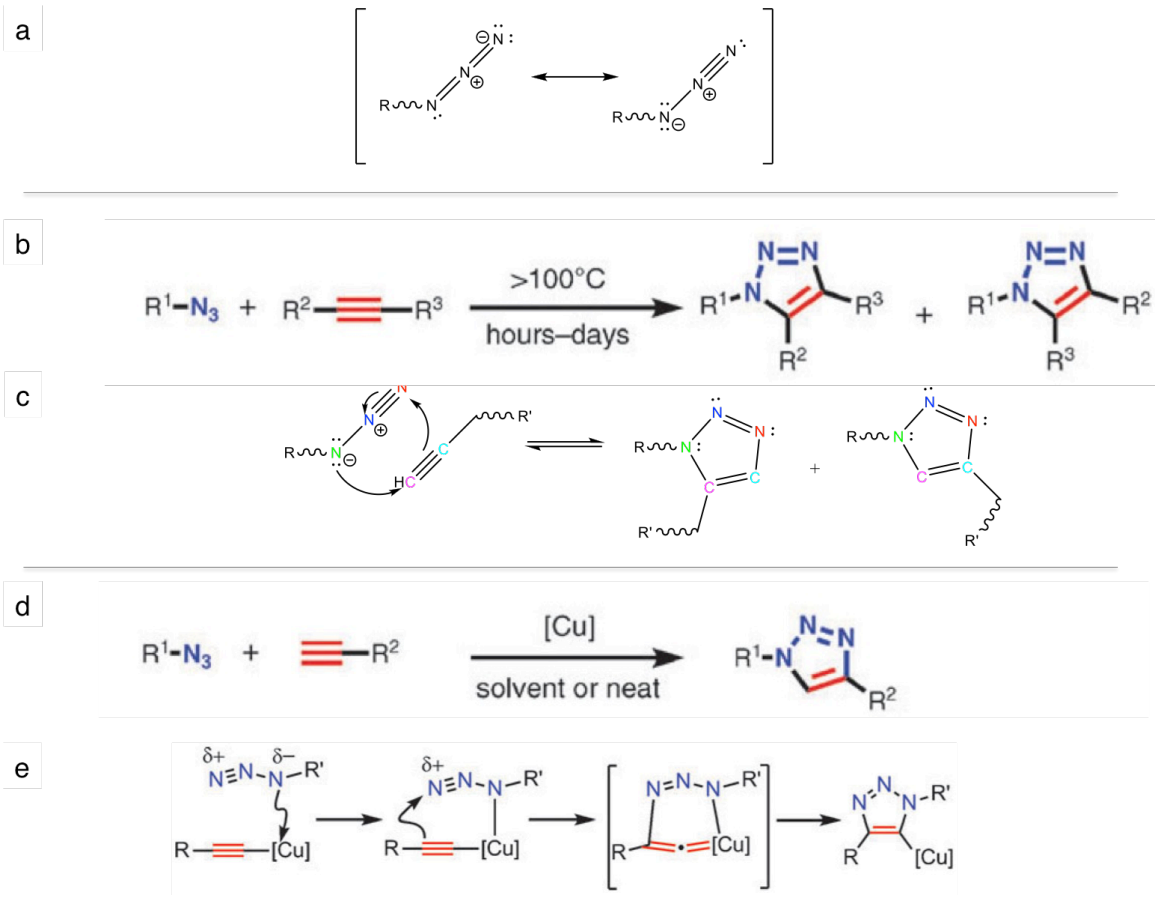



Figure 3.16. Azide resonance structures (a). General reactions for click chemistry between azides and alkynes in the absence (b, c) and presence (c, d) of a Cu(I) catalyst. Click chemistry mechanisms of 1,3-dipolar cycloaddition for C – N bond formation in the absence of any catalyst (c), and in the presence of Cu(I) as the catalyst (d), by way of copper acetylide. Addition of copper catalyst results in only 1,4-disubstituted-1,2,3-triazole, while use of heat only and no catalyst results in a mixture of 1,4- and 1,5-regioisomers.[271]

Click chemistry reaction between azides and alkynes (discovered in 1893) have inherently low reaction rates when uncatalyzed (Figure 3.16b, c), with a high activation barrier of ~25 kcal/mol.[271] It has been shown that the use of copper (I) is capable of increasing the reaction rate by a factor of $10^7$ - $10^8$.[271] The discovery of a copper (I) catalyst is significant as it enables the reaction to proceed at much lower temperatures as a result of an effective reduction in the activation energy necessary, and also provides regioselectivity where only one isomer (a 1,4-disubstituted-1,2,3-triazole) is obtained instead of two (Figure 3.16d, e).[268,271,272] Copper-catalyzed azide-alkyne cycloadditions are carried out by binding terminal alkynes to the azide groups in the presence of catalytic amounts of copper (I).[268]

In order to maintain high turnover rates of the copper catalyst, a constant supply of Cu(I) must be made present within the solution. The use of a reducing agent is therefore necessary to maintain copper in the Cu(I) form as opposed to the Cu(II) form. Sodium ascorbate ($C_6H_7NaO_6$) is a powerful reducing agent, capable of providing a constant source of active Cu(I) catalyst ions in the presence of atmospheric oxygen.[273] Hong *et al.* has determined that a sodium ascorbate concentration of 2.5 mM in solution is sufficient to maintain active copper in the +1 oxidation state.[273] Rostovtsev *et al.* have proved the robustness of the copper catalyzed click chemistry reaction with sodium ascorbate as the reducing agent as only a low concentration of sodium ascorbate is necessary, with the reaction proceeding to completion between 6 and 36 h at room temperature in a variety of solvents, with high reaction yields achieved at an impressively wide pH range of 4 – 12.[274]

#### 3.2.4.3 Templation of ferric oxide nanoparticles

Coprecipitation is one of the simplest and most efficient methods of obtaining magnetic particles.[223,275] Iron oxides can be obtained by stoichiometric mixtures of ferrous and ferric salts in an aqueous medium, where magnetite and maghemite structures are obtained. A stoichiometric ratio of 2:1 $Fe^{3+}/Fe^{2+}$ in an aqueous environment between pH 8 and 12 will result in complete precipitation of $Fe_3O_4$, as given by Equation 3.3.

$$Fe^{2+} + 2Fe^{3+} + 8OH^- \rightarrow Fe_3O_4 + 4H_2O$$

Equation 3.3

Magnetite is not very stable, however, and in the presence of oxygen it is easily converted to maghemite, as is described by Equation 3.4.

$$Fe_3O_4 + 2H^+ \rightarrow \gamma Fe_2O_3 + 2Fe^{2+} + H_2O$$

Equation 3.4

The structure of maghemite is explained through the oxidation of ferrous ions causing the migration of cations through the crystal lattice, resulting in cationic vacancies that maintain the overall charge balance in the unit cell.[224] Maghemite differs from magnetite by the presence of these cationic vacancies within the octahedral site.[224]

Templation of iron oxide nanoparticles will occur through coprecipitation and subsequent reduction of $FeCl_3$ and $FeSO_4 \cdot H_2O$ precursors by NaOH.[218] This procedure is standard and successful at room temperature, with superparamagnetic nanoparticles crystallized with narrow size distributions (10-30 nm diameter particles) within seconds of adding the reducing agent.[205,218,231] This procedure may be performed in the absence or presence of any peptide. In the presence of clicked peptides, purification may be performed by applying a magnetic field and isolating proteins that are successfully bound to magnetic nanoparticles.

## 3.3 Materials and methods

### *3.3.1 Materials*

Sodium phosphate (monobasic and dibasic), nickel-nitrilotriacetic acid resins, were purchased from Sigma-Aldrich. Ampicillin, isopropyl-$\beta$-D-thiogalactopyranoside (IPTG), tryptone, urea, tris-HCl, and sodium chloride were obtained from Fisher Scientific. Yeast extract and methanol were purchased from Acros Organics and BCA kit was obtained from Pierce. Imidazole was purchased from Alfa Aesar and copper grids for TEM were purchased from Ted Pella. Uranyl acetate for staining TEM grids was purchased from Electron Microscopy Science. Glass microscopy slides used in optical microscopy were per-cleaned Gold Seal beveled micro slides (model 3061) obtained from Clay Adams. Dialysis tubing (3.5 kDa molecular weight cut-off) was obtained from Thermo Scientific. ZipTip pipette tips were obtained from EMD Millipore.

Materials used in expression of C+AHA and Q+AHA include those presented in Table 3.2.

Table 3.2. Reagents and suppliers of materials used in AHA incorporation.

| **Reagent** | **Supplier** |
|---|---|
| Calcium chloride | EM Science |
| D-glucose monohydrate | Alfa Aesar |
| Glycine | Amresco |
| Kanamycin sulfate | Amresco |
| L-alanine | Fisher Scientific |
| L-arginine | Fisher Scientific |
| L-asparagine monohydrate | Fisher Scientific |
| L-aspartic acid | Fisher Scientific |
| L-azidohomoalanine | Click Chemistry Tools |
| L-cysteine hydrochloride monohydrate | Fisher Scientific |
| L-glutamic acid | Fisher Scientific |
| L-glutamine | Sigma Aldrich |
| L-histidine | Amresco |
| L-isoleucine | Fisher Scientific |
| L-leucine | Fisher Scientific |
| L-lysine hydrochloride | Fisher Scientific |
| L-methionine | Amresco |
| L-phenylalanine | Fisher Scientific |
| L-proline | Sigma Aldrich |
| L-serine | Fisher Scientific |
| L-threonine | Fisher Scientific |
| L-tryptophan | Fisher Scientific |
| L-tyrosine | Fisher Scientific |
| L-valine | Fisher Scientific |
| Magnesium sulfate | Sigma Aldrich |
| Vitamin B (Thiamine hydrochloride) | Sigma Aldrich |

The following materials were used for click chemistry. Iron (II) sulfate heptahydrate ($FeSO_4 \cdot 7H_2O$) and iron (III) chloride ($FeCl_3$) were obtained from Sigma Aldrich. Copper wire was obtained from Arcor Electronics and copper sulfate ($CuSO_4$)

was obtained from Amresco. NaOH was procured from BDH Chemicals and sodium ascorbate was obtained from Acros Organics. Alkyne-functionalized Chromeo494 was purchased from Active Motif. L-azidohomoalanine (AHA) was purchased from Click Chemistry Tools. CMms6 (sequence prg-YAKMKSRDIESAQSDEEVELRDALA, functionalized with an N-terminal L-propargylglycine, prg) was synthesized and purchased from LifeTein. See Table A 2 for a complete list of reagents and suppliers.

### *3.3.2 Synthesis of L-azidohomoalanine*

L-azidohomoalanine was chemically synthesized as described by Link *et al.* (Figure 3.17).[254] This reaction was scaled up by 1.802 (= 1.77 g/3.19 g) to adjust for an increased amount of Boc-dab starting material. All glassware used was flame dried.

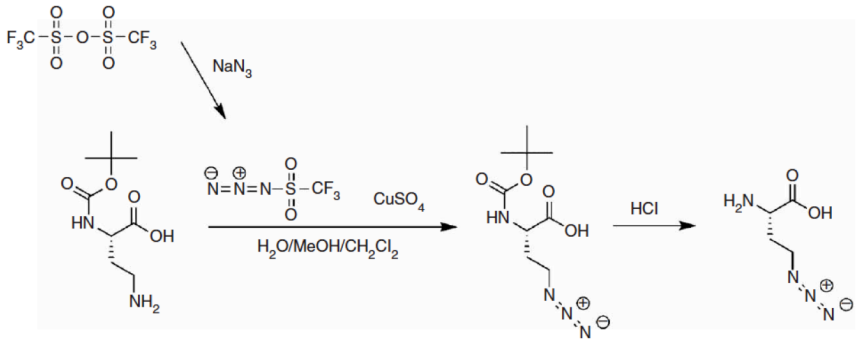



Figure 3.17. Synthesis schematic of AHA from protected diaminobutyric acid.[254]

Distillation of triflic anhydride was the first step in the synthesis. 970 mg of $P_2O_5$ was added to to 15.44 g $Tf_2O$ *slowly* in the fume hood. If $Tf_2O$ contains water, this will result in a highly exothermic reaction. An oil bath was heated at 95 °C until all liquid was distilled, with the condenser reflux water at 10 °C. 9.5 g sodium azide was added to to 23.4 mL $dH_2O$ in a round bottom flask, which was covered and put in an ice bath. 4.87 mL triflic anhydride was then added to the flask, which was then removed from the ice bath and the mixture was stirred at room temperature for 2 h. In another round bottom flask, 46.9 mL $dH_2O$ and 93.7 mL methanol were combined. 3.2 g of Boc-dab, 3.03 g $K_2CO_3$, and 56.36 mg $CuSO_4 \cdot 5H_2O$ were then added. The reaction mixture was transferred to a separatory funnel and extracted with 14.4 mL methylene chloride. The separatory funnel was shaken by inverting and swirling, occasionally venting to release

gas buildup. The organic bottom layer was collected and set aside. The aqueous layer was extracted with another 14.4 mL methylene chloride, mixed, and separated.

Neutralization was performed with $Na_2CO_3$ to neutralize any remaining triflic acid. Methylene chloride layers were combined in a separatory funnel with 45 mL saturated $Na_2CO_3$. The methylene chloride layer (bottom), which contained triflic azide, was then collected. The methylene chloride/triflic azide solution was added dropwise to the flask and stirred overnight at room temperature.

The following day methanol and methylene chloride were removed by rotary evaporation at 40 °C leaving behind about 45 mL of an aqueous solution. This solution was bright blue and pH ~ 9.3. 6 N HCl was added slowly to the solution until the pH reached 6. Upon reaching pH ~ 8.5 the solution turned green. Phosphate buffer containing 0.25 M $KH_2PO_4$ and 0.25 M $K_2HPO_4 \cdot 3H_2O$ was pH adjusted to 6.2. 180 mL of the phosphate buffer was added to the green solution and the pH was readjusted to 3 with 6 N HCl. The aqueous layer was extracted twice with 2 x 180 mL ethyl acetate.

Ethyl acetate layers were combined and washed in a separatory funnel with 90 mL saturated NaCl. Ethyl acetate layers were collected and 9 g drying agent ($MgSO_4$) was added to remove any water from the organic layer. The solution was allowed to dry for 30 min. $MgSO_4$ was removed by gravity filtration through filter paper. Ethyl acetate was removed by rotary evaporation at 40 °C. The result was 5 g of oil containing protected AHA. The oil was removed from the rotary evaporator and 18 mL of concentrated HCl was added. The solution was stirred at room temp for 1 h and then diluted with 162 mL $dH_2O$. The resulting solution was then stored in the 4 °C fridge overnight.

45 g of Dowex resin was added to a large glass chromatography column and washed with:

3 x 180 mL 1 N $NH_4OH$
1 x 7 L $dH_2O$
1 x 180 mL 1 N HCl
1 x 7 L $dH_2O$

After washing the eluent was slightly acidic (pH 6-7). The oil solution from the previous step was added to the column and the flow through was collected and reintroduced into the column once more.

The column was then washed with 10.8 L $dH_2O$, or until the pH of the eluent remained constant (pH 6-7). The pH was frequently tested with Litmus paper. When the pH reached a constant value, the amino acid was eluted from the column by adding 720 mL 1 N $NH_4OH$. The aqueous $NH_4OH$ was removed from the product using a rotary evaporator at 65 °C. The product obtained in the previous step was dissolved in a minimal amount of $dH_2O$ in a 100 mL round bottom flask. Acetone was added drop wise until the solution started to cloud. The mixture was swirled in a 75 °C water bath until the mixture became clear again. Once the solution was clear crystalline AHA was collected by vacuum filtration on filter paper.

### *3.3.3 Expression with unnatural amino acid incorporation*

pQE30 plasmid vectors containing C and Q gene sequences were transformed into methionine auxotrophic M15MA cells.[260] Colonies were initially grown on TSA plates containing 0.2 mg/mL ampicillin and 0.035 mg/mL kanamycin as selection antibiotics at 37 °C for approximately 16 hours. Ampicillin and kanamycin were used to select for genetically modified host cells containing the target pQE30 plasmid vector and methionine auxotropic strain.[109] Single colonies were selected from the plates and inoculated in 5 mL of M9 media containing all 20 natural amino acids (recipe for 20 aa M9 media provided in Appendix 5.6.3) with 0.2 mg/mL ampicillin, 0.035 mg/mL kanamycin, 0.033 mg/mL vitamin B, 0.085 mg/mL D-glucose, 0.11 mg/mL magnesium sulfate, and 0.01mg/mL calcium chloride and incubated at 37 °C, at 350 rpm, for 12 h or until $OD_{600} > 1$.

Proteins were then expressed in 200 mL media volumes contained in 1 L baffled Erlenmeyer flasks. 4 mL of starter culture was added to 200 mL of 1X M9 and 0.2 mg/mL ampicillin, 0.035 mg/mL kanamycin, 0.033 mg/mL vitamin B, 0.085 mg/mL D-glucose, 0.11 mg/mL magnesium sulfate, and 0.01mg/mL calcium chloride. Pre-induction was allowed to incubate for 6 h at 350 rpm at 37 °C. Cells were then pelleted

by centrifugation on a Beckman Coulter Allegra™ centrifuge, Beckman Coulter TS-5.1-500 swinging bucket rotor, at 4 °C and 4000 rpm for 10 minutes. 1 mL of the pre-induction culture was sampled and stored for later analysis at -20 °C. In order to remove residual methionine before the introduction of AHA, three washes of the culture with ice cold 0.9 % NaCl solution was performed. Pre-induction culture was transferred to centrifuge bottles and spun at 4000 rpm for 10 min at 4 °C using a Beckman Coulter Allegra™ centrifuge (IN, USA), Beckman Coulter TS-5.1-500 swinging bucket rotor (IN, USA). Supernatant was discarded and resuspended in 150 mL 0.9 % ice cold NaCl solution. Centrifugation and rinsing with NaCl solution was repeated a total of three times.

200 mL of 1X M9 media bearing 19 amino acids (no methionine), 0.2 mg/mL ampicillin, 0.035 mg/mL kanamycin, 0.033 mg/mL vitamin B, 0.085 mg/mL D-glucose, 0.11 mg/mL magnesium sulfate, and 0.01mg/mL calcium chloride was added to the rinsed cell pellet and resuspended via shaking and vortexed using a Fisher Scientific™ 50/60 Hz 150 W Vortex Mixer (MA, USA) Resuspended pellets were transferred back to 1 L Erlenmeyer flasks that were then sterilized over a flame. Cells were starved of methionine by incubating at 37 °C, 350 rpm for 15 min. To induce protein expression, 200 μL of 200 mg/mL IPTG (final concentration of 0.2 μg/mL) and 57.6 μL of 100 mg/mL AHA (final concentration 0.2 mM AHA) was added to the flask. Induction was allowed to proceed at 300 rpm, 37 °C for no more than 3 h. Aliquots of post expression samples were stored for later analysis at - 20 °C. Post induction, $OD_{600}$* was measured and cells were harvested by centrifugation at 4 °C and 4000 rpm for 15 minutes using a Beckman Coulter Allegra™ centrifuge (IN, USA), Beckman Coulter TS-5.1-500 swinging bucket rotor (IN, USA). Cell pellets were stored in -80 °C until purification. Pre/post expression samples for each culture were run on 12 % on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to evaluate protein expression according to the procedure outlined in Appendix 5.6.7.1.

Purification under denaturing conditions was carried out using 50 mM tris-HCl, 0.5 M NaCl, 20 mM imidazole, 6 M urea pH 8 buffer. The soluble crude lysate was

* Ideally the post-induction OD is double that of pre-induction.

bound to Ni-NTA beads (Sigma Aldrich, MO, USA) and allowed to equilibrate for at least 3 h at 4 °C under continuous rotation in a Thermolyne Labquake shaker rotisserie (model 415110, Thermo Scientific, MA, USA). The proteins were eluted with increasing gradient of imidazole (20 mM – 1 M). Pure fractions were refolded via stepwise dialysis using dialysis tubing with a 3.5 kDa molecular weight cut-off in pH 8 50 mM phosphate buffer, halving the urea concentration successively. For experiments involving mixtures of C+AHA and Q+AHA proteins, bicinchoninic acid analysis (BCA) measurements were performed after purification and mixtures of varying molar ratios of C and Q were prepared in denatured conditions. The mixtures were then co-dialyzed according to the same conditions described above. The enhanced protocol of BCA (PIERCE, IL, USA) (see Appendix 5.6.6) was used to estimate protein concentration with bovine serum albumin (BSA) as a standard.

#### 3.3.3.1 Amino acid analysis

Amino acid analysis of purified C+AHA and Q+AHA (1.5 mL of 10 μM protein in 50 mM PB pH 8) as well as pure AHA (1.5 mL of 100 mg/mL in 50 mM PB pH 8) was performed by the Molecular Structure Facility at the University of California, Davis. The percentage of AHA incorporation in C+AHA and Q+AHA was determined as follows.

$$n_a = \frac{2\ x\ I}{n_{mol}}$$

Equation 3.5

Peak areas from amino acid analysis corresponding to each amino acid were obtained ($n_{mol}$) and adjusted for fluctuations in the internal standard, norleucine ($I$) (Equation 3.5). The adjusted number of moles ($n_a$) was determined for each amino acid in the spectra and totaled ($T_{na}$) (Equation 3.6).

$$T_{na} = \sum n_a$$

Equation 3.6

Mole fraction ($f_{mol}$) was calculated as the fraction of $n_a$ over $T_{na}$ (Equation 3.7).

$$f_{mol} = \frac{n_a}{T_{na}}$$

Equation 3.7

As AHA demonstrates the same retention time as lysine (46 min), the amount of lysine in C+AHA and Q+AHA was taken as the sum of the theoretical lysine (3) and AHA (2) residues in the sequence. The experimentally determined number of AHA residues ($A_e$) was calculated as the observed mole fraction corresponding to the lysine peak minus contribution from 3 lysine residues multiplied by the total number of residues in the sequence ($T$ = 54) (Equation 3.8). Percent incorporation ($PI$) was determined as the ratio between the experimentally observed number of AHA residues and the expected number of AHA residues ($A$ = 2) (Equation 3.10).

$$A_e = f_{mol}\, x\, T$$

Equation 3.8

$$PI = 100\, x\, \frac{A_e}{A}$$

Equation 3.9

### *3.3.4 ImageQuant analysis*

ImageQuant analysis software (GE Healthcare Life Sciences) was used to perform 1D electrophoresis gel analysis.[276] A molecular weight ladder was used to obtain a molecular weight calibaration curve that is automatically calculated by the software. Molecular size calibration selects standards from a library within the software. Bands of interest are detected within the gel and their molecular weight is estimated by comparison with the molecular weight standard curve.

### *3.3.5 Click chemistry*

Click chemistry was performed using the C and Q incorporated with AHA (C+AHA and Q+AHA, respectively) with an alkyne-bearing magnetite binding peptide, propargylglycine CMms6 (prg-CMms6). Before performing click chemistry with prg-CMms6, conjugation with a fluorescent probe, Chromeo494 that bears an alkyne was

performed. Stock reagents used in the click chemistry reaction included copper wire, alkyne ligand either 10 mM Chromeo 494 (dissolved in MeOH/$dH_2O$) or 2.25 mM prg-CMms6 (dissolved in $dH_2O$), 400 mM sodium ascorbate, 100 mM $CuSO_4$ (dissolved in $dH_2O$), and pure 50 µM AHA protein or whole cell lysate ($OD_{600}$ = 20) containing AHA protein, all dissolved in 50 mM PB, pH 8. Click chemistry was performed with either purified AHA protein or lysate containing expressed AHA protein. Conditions for these reactions are described in this section.

#### 3.3.5.1 Pure protein

Pure AHA protein was obtained after dialysis using dialysis tubing with a 3.5 kDa molecular weight cut-off in pH 8 50 mM phosphate buffer. Alkyne ligand (prg-CMms6 or Chromeo494), $CuSO_4$, sodium ascorbate, and Cu wire were added to 50 µM pure protein and diluted with 50 mM PB pH 8.[258] Final concentrations of reaction conditions were:

50 µM azide-functionalized protein
1 mM alkyne ligand (prg-CMms6 or Chromeo494)
25 mM sodium ascorbate
500 mM $CuSO_4$
1 mg Cu wire

Reaction tubes were gently mixed by inverting several times. Samples were then incubated at 50 or 37 °C and shaken continuously at 300 rpm for 24-48 h using a Brinkmann Instruments™ Thermomixer R (HH, DE). 5 µL SDS dye (composition provided in Appendix 5.7) was added to each sample and they were heated at 95 °C for 20 minutes prior to before subjecting them to 12 % SDS-PAGE. Gels were visualized with fluorescence filter 630BP30 (detection range 550-650 nm) with Chromeo494 alkyne, which has a peak emission at 628 nm (Figure 3.18).

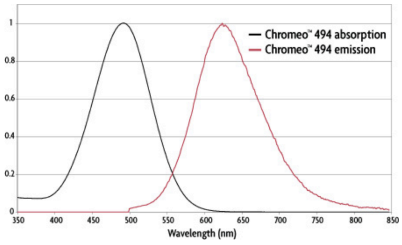



Figure 3.18. Absorption and emission spectra of Chromeo494.[277]

#### 3.3.5.2 Cell lysate

Frozen cell pellets with C+AHA or Q+AHA protein were thawed and resuspended in an appropriate volume (approximately 20 mL) of 50 mM PB pH 8 to obtain an $OD_{600}$ = 20. To facilitate lysis, lysate was sonicated by probe sonication (Sonicator 500, Qsonica, LLC, CT, USA) for 1 minute, with a pulse 5 s on/5 s off at 35 % amplitude then heated at 95 °C for 20 minutes using a Brinkmann Instruments™ Thermomixer R (HH, DE). Alkyne ligand (prg-CMms6 or Chromeo494), $CuSO_4$, sodium ascorbate, and Cu wire were added to the cell lysate and diluted with 50 mM PB pH 8.[258] Final concentrations of reaction conditions were:

$OD_{600}$ = 15 for lysate (containing azide-functionalized protein)
1 mM alkyne ligand (CMms6 or Chromeo494)
25 mM sodium ascorbate
1 mM $CuSO_4$
1 mg Cu wire

Reaction tubes were gently mixed by inverting several times. Samples were then incubated at 50 or 37 °C and shaken continuously at 300 rpm for 24-48 h using a Brinkmann Instruments™ Thermomixer R (HH, DE). 5 µL SDS dye (composition provided in Appendix 5.7) was added to each sample and they were heated at 95 °C for 20 minutes prior to before subjecting them to 12 % SDS-PAGE. Gels were visualized with fluorescence filter 630BP30 (detection range 550-650 nm) with Chromeo494 alkyne, which has a peak emission at 628 nm (Figure 3.18).

### *3.3.6 Ferric oxide nanoparticle templation*

Ferric oxide nanoparticle templation was performed via co-reduction of two iron oxide salts in the presence and absence of prg-CMms6-containing proteins.[205,247,278] Stock solutions for the reactions were 600 mM $FeCl_3$, 300 mM $FeSO_4$, and 3 M NaOH (all dissolved in $dH_2O$). A molar ratio of ferric ($Fe^{3+}$) to ferrous ($Fe^{2+}$) ions was kept at 2:1 for co-precipitation of iron oxide nanoparticles.[229,279] $FeCl_3$ and $FeSO_4$ were added to protein solutions to obtain final concentrations of 60 mM $FeCl_3$ and 30 mM $FeSO_4$ and samples were degassed for 15 minutes in a desiccator at room temperature. After 15 minutes appropriate volumes of NaOH to reach a final NaOH concentration of 300 mM were added to samples reducing iron salts to precipitate nanoparticles. Successful formation of

ferric oxide nanoparticles was confirmed visually through the appearance of a black precipitate and magnetic ability was confirmed by manipulation with a bar magnet.

#### 3.3.6.1 Magnetic separation

Magnetic separation experiments were performed after binding of lysate samples with MNPs generated through coprecipitation of ferric and ferrous salts in the absence of protein in 50 mM PB pH 8 (Section 3.3.6). 250 µL of MNP solution was added to 500 µL of lysate containing C+AHA, Q+AHA, C+AHA+CMms6, or Q+AHA+CMms6. Samples were incubated at room temperature for 30 minutes. Separation was performed by isolation of MNPs with a neodymium magnet (model 07212, Magnet Source) normal to gravity (Figure 3.19).

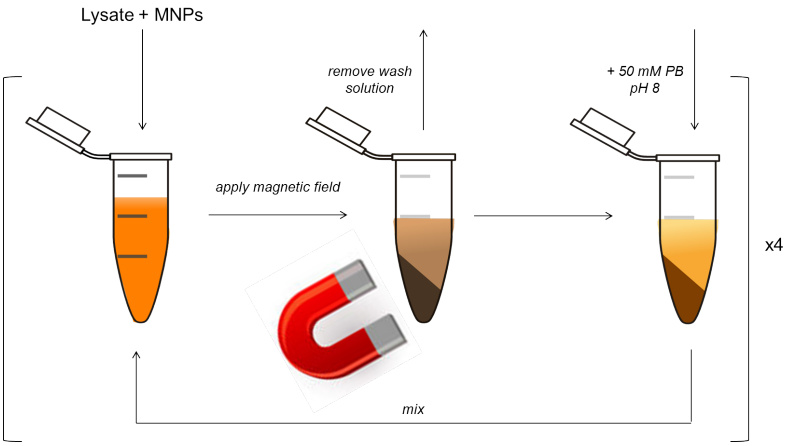



Figure 3.19. Schematic of setup for magnetic separation experiments.

500 µL of soluble fraction was pipetted out, replaced with 500 µL of 50 mM PB pH 8, and mixed thoroughly. 50 µL samples were taken after each wash for imaging. Magnetic isolation was performed again as described above and this process was repeated for a total of four buffer washes.

### *3.3.7 Optical microscopy*

Micrographs and time-lapse videos of magnetic nanoparticles were imaged with a Leica DMI4000 B microscope (IL, USA). Approximately 20 µL of solution was spotted on a glass microscopy slide (Clay Adams). Micrographs were obtained without drying the samples before imaging, and without the use of a coverslip, allowing MNPs to remain

soluble and travel within the liquid sample upon application of a magnetic field. Images were taken with a Leica DFC310 FX camera (IL, USA). Time-lapse videos were made constructed from images taken at regular intervals of 0.5 s over a 20-30 s time period. Images were compressed to 10 frames per second to generate videos.

### *3.3.8 Circular dichroism*

Circular dichroism (CD) measurements were conducted on a Jasco J-815 CD spectrometer (Jasco, Inc. MD, USA). Wavelength scans were conducted with 10 μM protein concentrations in 50 mM PB pH 8 at room temperature. The wavelength spectrum was measured over a range from 190 to 250 nm with a step size of 1 nm. Mean residue ellipticity (MRE) was calculated from raw data according to the procedure described in Gunasekar *et al.* 2009.[92]

### *3.3.9 TEM and electron diffraction*

Approximately 3 μL of 10 μM protein in 50 mM PB pH 8 with or without MNPs was spotted on copper grids, as done before (Section 1.3.8).[1] After 1 minute, the grids were blotted using filter paper and rinsed with 3-4 drops Milli Q water to remove excess salts from the buffer. After blotting with filter paper, samples in the absence of MNPs were negatively stained by adding 3 μL of 1 % filtered uranyl acetate and blotted using filter paper. Samples with MNPs were not stained. Samples were then dried at room temperature for 10-15 minutes. Prepared grids were viewed on a JEOL JEM-1400 TEM (Jeol, MA, USA) at an accelerating voltage of 120 kV. ImageJ software was used to measure nanoparticle dimensions.[116]

Selected-area electron diffraction for single crystals was also obtained on the JEOL JEM-1400 TEM (Jeol, MA, USA). Lattice *d*-spacing was calculated from measurements of diffraction rings of single crystals according to Equation 3.10 and Equation 3.11, where *D* equals the ring diameter, *r* equals ring radius, and *d*-spacing represents the crystal lattice spacing.[280]

$$r = D/2$$

Equation 3.10

$$d - spacing = 1/r$$

Equation 3.11

### *3.3.10 Magnetometry*

A Quantum Design MPMS-3 Superconducting Quantum Interference Device (SQUID) magnetometer (CA, USA) was used to characterize the magnetic properties of MNPs synthesized in the presence or absence of protein. Hysteresis curves enable the study of anisotropy of ferric oxide nanoparticles and are a necessary means of characterization for these materials.[281,282] Sample magnetization was measured as a function of applied magnetic field in order to ascertain if the nanoparticles showed hysteresis or superparamagnetic behavior, and to estimate magnetic properties such as saturation magnetization, remanence and coercivity (if any), and the magnetic diameter (obtained from a fit to the Langevin function weighted by the lognormal size distribution). These measurements were made at a temperature of 300 K and in a field range of up to 7 T.

## 3.4 Results and discussion

### *3.4.1 Protein expression via unnatural amino acid incorporation*

In order to produce the C+AHA and Q+AHA proteins, M15MA methionine auxotrophic *E. coli* was used as an expression host.[260] Expression of C and Q with AHA was evaluated by SDS-PAGE analysis. Before the introduction of IPTG, no overexpression was observed for both C and Q, as expected (Figure 3.20a, b). In order to determine whether target proteins were expressed with AHA, control samples were run in post-induction conditions of either 19 amino acids only (missing methionine or an analogue of methionine) or the standard set of 20 amino acids (Figure 3.20a, b). As can be seen in Figure 3.20, over expression bands for the target proteins can be seen in the case of 19 amino acids + AHA and for the standard set of 20 amino acids, but protein expression was not obtained in the absence of methionine or its analogue (Figure 3.20a, b). These pre-induction and post-induction protein gels revealed over-expression of C and Q in the presence of AHA, indicative of incorporation (Figure 3.20). Incorporation of AHA shown

in SDS-PAGE in Figure 3.14 was performed with AHA synthesized by us in the Link Lab at Princeton University (synthesis procedure described in Section 3.3.2).

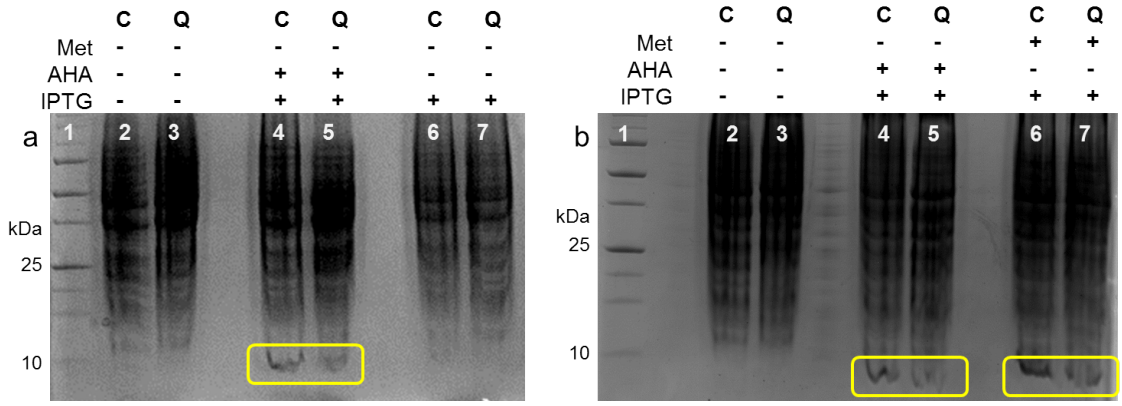



Figure 3.20. SDS-PAGE gels showing AHA incorporation with expression with 19 and 20 amino acids as controls. (a) 1. Ladder, 2. C pre-induction, 3. Q pre-induction, 4. C+AHA post-induction, 5. Q+AHA post-induction, 6. C 19 aa post-induction, 7. Q 19 aa post-induction. (b) 1. Ladder, 2. C pre-induction, 3. Q pre-induction, 4. C+AHA post-induction, 5. Q+AHA post-induction, 6. C post-induction, 7. Q post-induction.

Both C+AHA and Q+AHA were purified via affinity binding to Ni-NTA beads under denaturing conditions and eluted via increasing imidazole gradient (Figure 3.21).

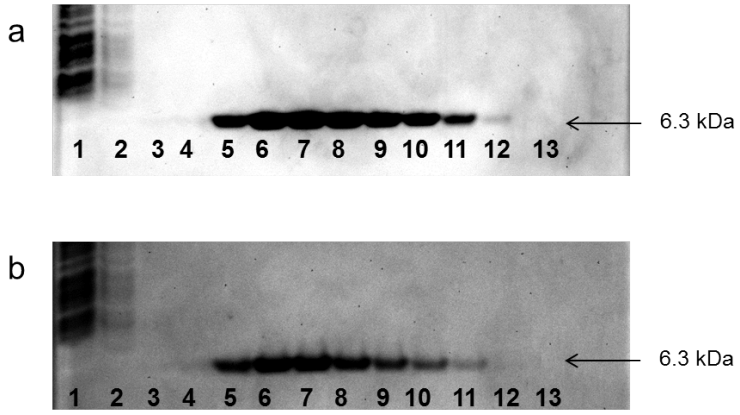



Figure 3.21. SDS-PAGE analysis of purification for C+AHA (a) and Q+AHA (b). Lanes represent: 1. supernatant, 2. flow through, 3. wash 1 (20 mM imidazole), 4. wash 2 (20 mM imidazole), 5. 100 mM imidazole elution, 6 and 7. 200 mM imidazole elutions, 8-10. 500 mM imidazole elutions, 11-13. 1 M imidazole elutions.

As can be seen in both gels in Figure 3.21, elution of the pure proteins began at a concentration of 100 mM imidazole and finished at 500 mM imidazole. A significant amount of pure protein was obtained and these AHA-containing proteins were subsequently assessed for percent incorporation via amino acid analysis (AAA).

The AAA signal for AHA overlaps with the region where lysine is detected, as was discovered upon running amino acid analysis on the AHA residue alone (Figure 3.22).

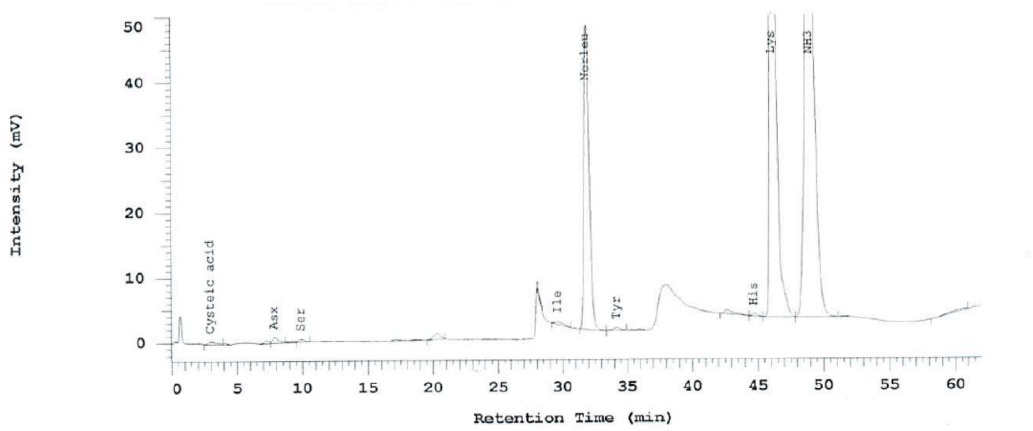



Figure 3.22. Amino acid analysis for AHA.

This precludes the possibility of observing the peak for AHA in the spectra, however the incorporation of AHA can be determined by taking into account the total area of the lysine peak in the spectra, as described in Section 3.3.3.1. Incorporation of AHA is also qualitatively confirmed by the absence of the methionine peak in the spectra for C+AHA and Q+AHA compared to C and Q (Figure 3.23). As can be seen in Figure 3.23a and c, C and Q both demonstrate a clear peak for methionine at a retention time of approximately 28 minutes. From these spectra pecent incorporation of AHA in C+AHA was determined to be 78.3 % and in Q+AHA was 94.3 %. These values are in agreement with previsouly reported incorporation levels of AHA in *E. Coli* M15MA cell of between 70-90 %.[256]

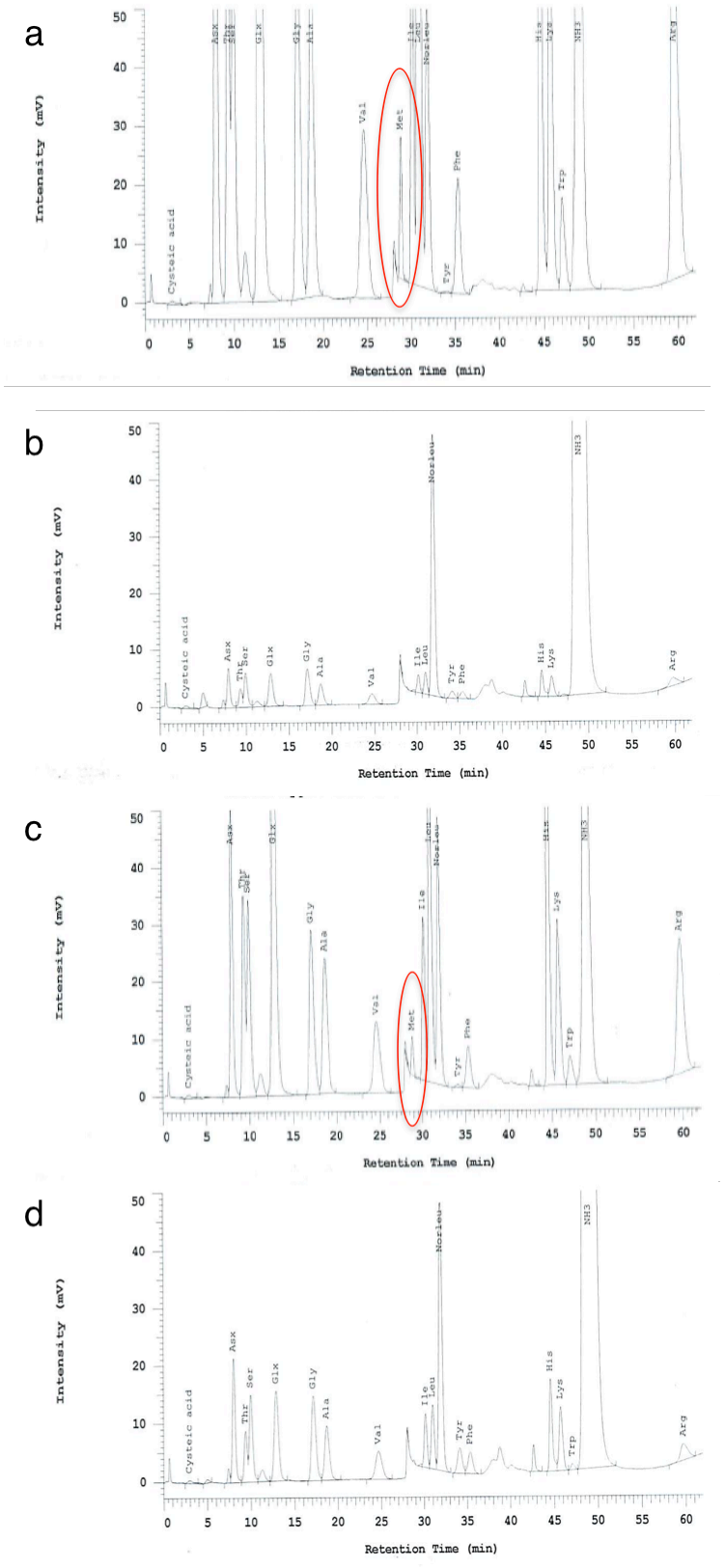



Figure 3.23. Amino acid analysis for C (a), C+AHA (b), Q (c), and Q+AHA (d).

### *3.4.2 Click chemistry*

#### 3.4.2.1 Chromeo494

To assess whether the C+AHA and Q+AHA proteins are able to undergo conjugation and determine appropriate conditions for click chemistry, a fluorescent probe called Chromeo494, which was functionalized with an alkyne group was used to engage in the copper-catalyzed azide-alkyne reaction. The use of Chromeo494, which demonstrates an excitation peak at 494 nm and an emission peak at 628 nm (Figure 3.18), imparts fluorescent properties to the AHA-containing protein upon successful conjugation. Reaction between Chromeo494 and C+AHA, C, Q+AHA, and Q was evaluated over a period of 48 h, where samples were taken at intervals of 4, 8, 12, 24, 36, and 48 h of incubation (Figure 3.24).

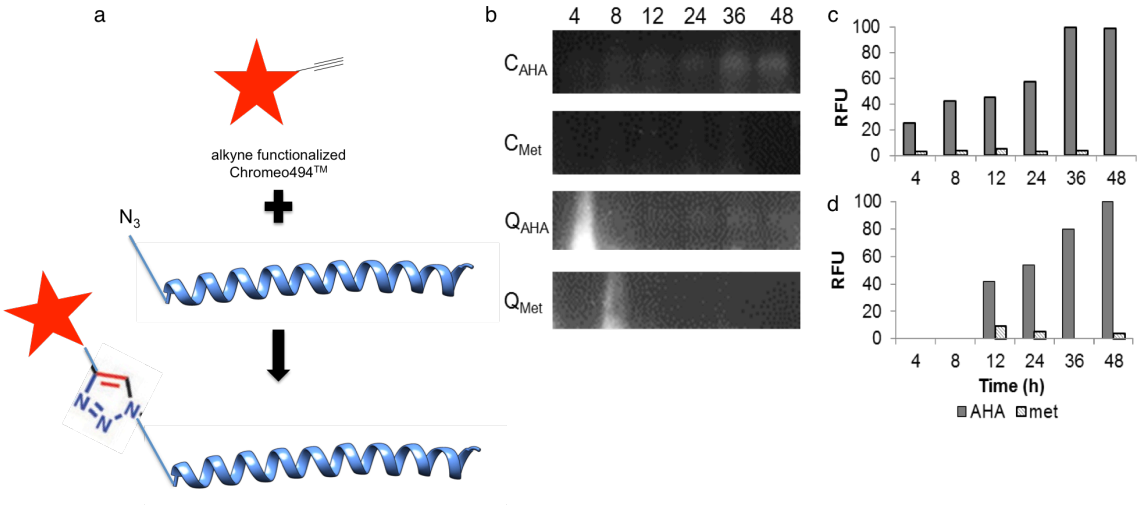



Figure 3.24. Chromeo494 bearing an alkyne moiety and reaction schematic for covalent attachment via azide group (a).* SDS-PAGE showing results of click chemistry reaction with incubation times of 4, 8, 12, 24, 36, and 48 h for Chromeo494 and C+AHA, C, Q+AHA, and Q, imaged with fluorescence filter (b). Results of ImageQuant gel analysis quantifying the RFU over time for C+AHA and C (c) and Q+AHA and Q (d).

Figure 3.24b has been cropped to show only the molecular weight bands corresponding to conjugated protein and Chromeo494 (6.85 kDa for AHA proteins and Chromeo, and a hypothetical 7.49 kDa for met proteins and Chromeo494). Bright fluorescent bands can

* Chemical structure of Chromeo494™ is unavailable as it is a registered trademarked product (Active Motif).

be seen in the C+AHA and Q+AHA gels that increase in intensity with increasing reaction times (Figure 3.24c, d). The exceedingly bright smudges that can be seen in the Q+AHA and Q gels at lanes corresponding to approximately 4 and 8 h, respectively, represent unbound Chromeo494 which ran fast through the gel to lower MWs but displayed residual fluorescence when imaging. These blurs were discounted when performing the quantitative analysis of these gels via ImageQuant.

ImageQuant analysis software (Section 3.3.4) was used to quantify the relative fluorescence units. RFUs were normalized to a maximal value (100 RFU) achieved after 48 h incubation. Both C+AHA and Q+AHA demonstrate increasing RFUs over time, reaching a plateau level after approximately 36 h in the case of C+AHA and 48 h for Q+AHA (Figure 3.24b, c). Notably, C and Q do not demonstrate high RFUs at any incubation point, and the RFUs observed do not increase over time. These results indicate that click chemistry of Chromeo494 was only successful with proteins bearing the azide group. Based on these experiments, we discovered that a 48 h incubation period was an appropriate time for the click chemistry reaction to proceed to produce to obtain the desired products.

In addition to viewing the Chromeo494 reaction products with the fluorescence filter, gels were also stained with Coomassie blue and imaged. In this case, gels were first viewed in the absence of Coomassie blue with the fluorescence filter (Figure 3.25b) and subsequently stained and viewed without the filter (Figure 3.25a) after incubation for 48 h. Fluorescent bands are seen in the case of Chromeo494 and the AHA-containing proteins, but absent in the lanes corresponding to the methionine-containing proteins (Figure 3.25). These bright fluorescent bands appear at the molecular weight distances that correspond to approximately 6.85 kDa, as was expected.

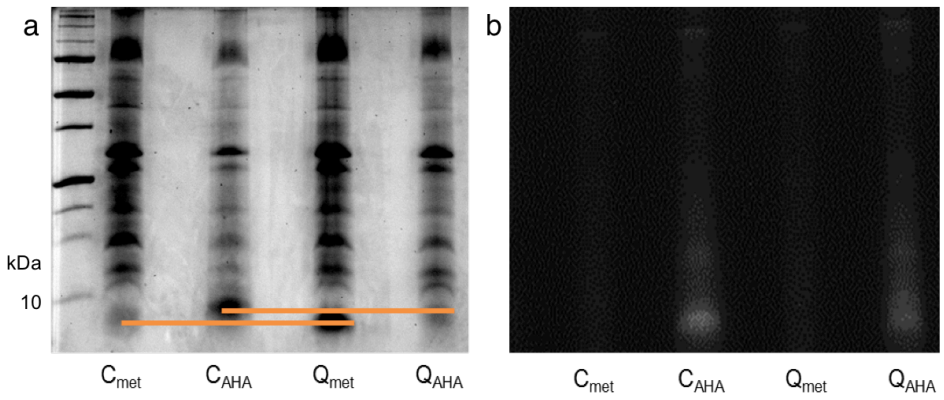



Figure 3.25. SDS-PAGE showing results of click chemistry reaction with of Chromeo494 with C, C+AHA, Q, and Q+AHA with incubation time of 48 h stained with Coomassie blue (a) and unstained and imaged with fluorescence filter (b).

Difference in molecular weights of the bands corresponding to C, Q, and C+AHA + Chromeo494 and Q+AHA + Chromeo494 (indictated by the orange lines in Figure 3.25a), as determined by ImageQuant software, were used to make a rough calculation of the yield of the click chemistry reaction (Table 3.3).

Table 3.3. Actual molecular weights and estimated molecular weight shift by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and Chromeo494. Estimated reaction yields were obtained by the molecular weight shift.

| | MW (kDa) | | (%) | | |
|---|---|---|---|---|---|
| | Actual | Shift | Gel shift yield | Incorporation | Overall yield |
| C, Q | 6.31 | | | | |
| C+AHA, Q+AHA | 6.30 | | | 78.3, 94.3 | |
| Chromeo494 | 0.55 | | | | |
| 2 x Chromeo494 | 1.11 | | | | |
| C+AHA + 2 x Chromeo494 | 7.40 | 0.91 | 83 | | 65 |
| Q+AHA + 2 x Chromeo494 | 7.40 | 0.77 | 70 | | 66 |

The shift in molecular weight observed between C and C+AHA + Chromeo494 was 0.91 kDa, and the shift between Q and Q+AHA + Chromeo494 was 0.77 kDa (Table 3.3). Given the actual molecular weights of these compounds, a reaction that conjugated two Chromeo494 ligands to the two AHA sites in the C+AHA and Q+AHA proteins, would

result in a molecular weight shift of 1.11 kDa, resulting in a yield of 83 % in the case of C+AHA and a yield of 70 % for Q+AHA (Table 3.3). The overall reaction yields are obtained by multiplying the yield of the gel shift by the percent incorporation of AHA in each protein (78.3 % for C+AHA and 94.3 % for Q+AHA) to obtain yields of 65 and 66 % for C+AHA+Chromeo494 and C+AHA+Chromeo494, respectively. These reaction yields were sufficiently high to begin experiments clicking on the alkyne-functionalized CMms6 peptide.

### 3.4.2.2 CMms6

With the successful outcomes of the click chemistry reaction of AHA-containing proteins and Chromeo494, experiments were run to evaluate the conjugation of magnetite binding peptide prg-CMms6 to C+AHA and Q+AHA in a similar fashion. Initially, varying concentrations of prg-CMms6 were studied to observe the effect of ligand concentration on click chemistry reaction yield. In order to get a sense of how the prg-CMms6 ran on the SDS-PAGE gels when unbound, concentrations ranging from 2.25 mM down to 50 μM were run alongside one another in a protein gel (Figure 3.26). As can be seen, higher concentrations of prg-CMms6 appear blurred on the gel (Figure 3.26). Concentrations used for click chemistry reactions were 500 μM and below.

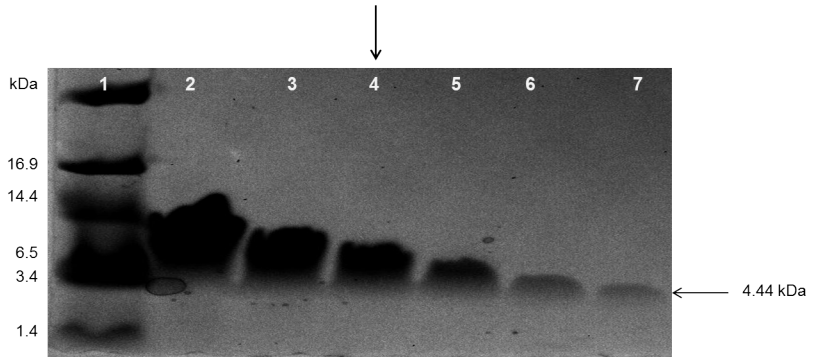



Figure 3.26. SDS-PAGE showing varying concentrations of prg-CMms6. 1. Ladder, 2. 2.25 mM, 3. 1 mM, 4. 500 μM, 5. 250 μM, 6. 100 μM, 7. 50 μM. The molecular weight of prg-CMms6 is 4.44 kDa.

Initial click chemistry reactions involving prg-CMms6 were run using the same reaction conditions that proved successful with Chromeo494. In order to compare, reactions were set up with C+AHA, Q+AHA, Chromeo494, prg-CMms6 and run alongside one another on an SDS-PAGE for analysis (Figure 3.27). This gel was viewed

with a fluorescence filter before it was stained (Figure 3.27b), and without the filter with the help of Coomassie blue staining after (Figure 3.27a). While the ladder in lane 1 is not fluorescent, bright bands are visible in lanes 2 and 3 corresponding to C+AHA + Chromeo494 and Q+AHA + Chromeo494, respectively (Figure 3.27b). The fluorescent bands at the bottom of the gel represent unreacted Chromeo494, which was difficult to remove from the gel before imaging (Figure 3.27b).

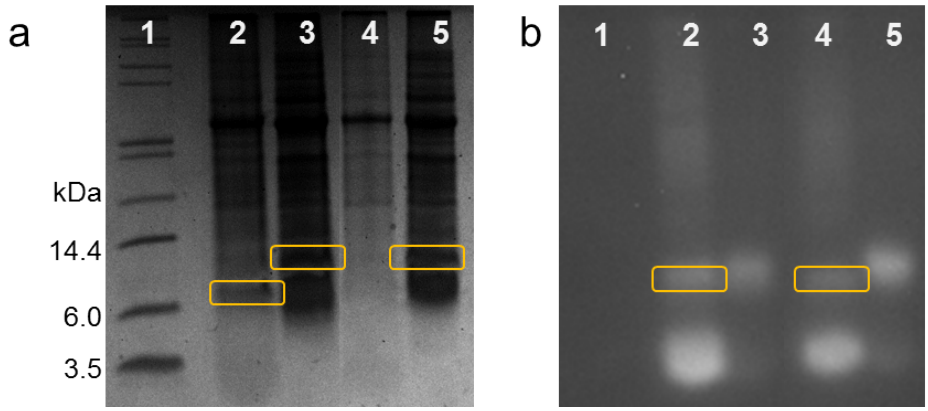



Figure 3.27. SDS-PAGE showing Coomassie blue stained gel (a); 1. Mark12 ladder, 2. C+AHA + Chromeo, 3. C+AHA + CMms6, 4. Q+AHA + Chromeo, and 5. Q+AHA + CMms6. Same gel viewed with the fluorescence filter (b), where bands seen in lanes 3 and 5 are sample from lanes 2 and 4 that spilled over when loading the gel.

In the Coomassie blue stained gel, bands for C+AHA + 2 x Chromeo494, C+AHA + prg-CMms6, and Q+AHA + prg-CMms6 can be seen clearly and are highlighted in orange (Figure 3.27a). Interestingly, the band for Q+AHA + Chromeo494 was not observed when stained with Coomassie blue but was visible under fluorescence visualization.

Table 3.4. Actual molecular weights and observed molecular weight shift by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and prg-CMms6. Observed fraction of proteins with two clicked prg-CMms6 peptides were obtained by the molecular weight shift.

| | MW (kDa) | | |
|---|---|---|---|
| | **Actual** | **Observed** | **Disubstituted fraction (%)** |
| C+AHA, Q+AHA | 6.30 | | |
| prg-CMms6 | 4.44 | | |
| 2 x prg-CMms6 | 8.88 | | |
| C+AHA + prg-CMms6 | 10.74 | 11.32 | 10 |
| C+AHA + 2 x prg-CMms6 | 15.18 | | |
| Q+AHA + prg-CMms6 | 10.74 | 11.46 | 15 |
| Q+AHA + 2 x prg-CMms6 | 15.18 | | |

Table 3.4 describes the actual and observed molecular weights from ImageQuant analysis (Section 3.3.4) of the clicked prg-CMms6 peptides in Figure 3.27a. The estimated MW of the bands corresponding to C+AHA + pgr-CMms6 and Q+AHA + pgr-CMms6 are 11.32 and 11.46 kDa (as calculated via ImageQuant), respectively. If one pgr-CMms6 peptide were successfully clicked on all AHA proteins, the expected MW would be 10.74 kDa, and if two pgr-CMms6 peptides were clicked on each AHA in these proteins the expected MW would be 15.18 kDa (Table 3.4). Accordingly, the estimated fraction of AHA proteins that clicked on two pgr-CMms6 peptides was approximately 10 % for C+AHA and 15 % for Q+AHA, taking into account the percent incorporation of AHA was well. The majority of these proteins were able to click on one pgr-CMms6 peptide, as the AHA residue at the N-terminus of the protein would be solvent exposed and easily accessed for binding.

Varying concentrations of pgr-CMms6 were used in the click chemistry reactions in order to study the effect of ligand concentration and obtain optimal conjugation. Concentrations examined in this study were 100, 200, 400, and 500 μM (Figure 3.28). An increase in MW was observed when C+AHA and Q+AHA were reacted with increasing concentrations of pgr-CMms6 (Figure 3.28).

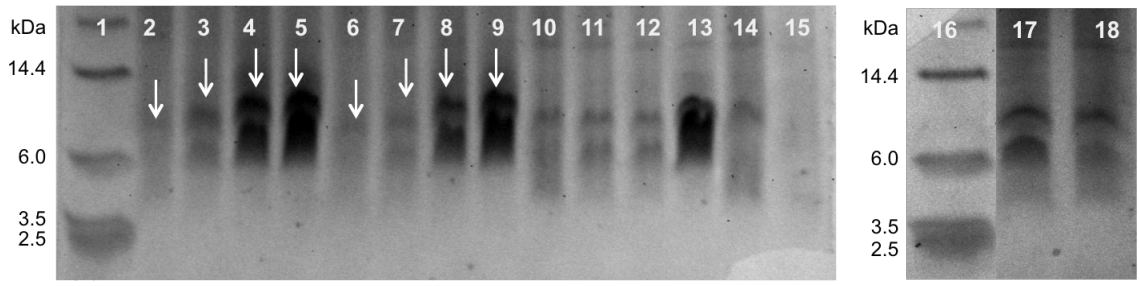



Figure 3.28. SDS-PAGE showing click chemistry reactions with varying concentrations of prg-CMms6 with no sodium ascorbate after incubation at 50 °C for 48 h. 1 and 16. Mark12 ladder, 2-5: C+AHA + varying concentrations of prg-CMms6, 6-9: Q+AHA + varying concentrations of prg-CMms6, 10-13: C + varying concentrations of prg-CMms6, 14, 15, 17, 18: Q+ varying concentrations of prg-CMms6, where lanes 2, 6, 10, and 14 contain 100 μM prg-CMms6; 3, 7, 11, and 15 contain 200 μM prg-CMms6; 4, 8, 12, 17 contain 400 μM prg-CMms6; and 5, 9, 13, and 18 contain 500 μM prg-CMms6.

A similar shift in the MW of bands in the lanes corresponding to C and Q containing was not observed (Figure 3.28). The MWs of the bands indicated with white arrows (Figure 3.28 were estimated via ImageQuant analysis (Section 3.3.4) and presented in Table 3.5. The yield of the click chemistry reaction with prg-CMms6 was determined as the MW of the bands in the gel compared to the expected MW of an AHA protein with a single bound prg-CMms6 peptide, 10.74 kDa. Although there is a discernable shift in the upper bands in lanes 2-9 in Figure 3.28, the band corresponding to unclicked protein (around 6 kDa) remains constant as concentrations of prg-CMms6 are increased. This band is of equal intensity as the band that increases in MW, and thus the overall yield is ultimately divided by a factor of two to take into account unclicked protein. The overall reaction yields are obtained by multiplying the yield of the gel shift by the percent incorporation of AHA in each protein (78.3 % for C+AHA and 94.3 % for Q+AHA) and then dividing by a factor of two. The highest concentration of prg-CMms6 that was used *without* obtaining disubstitution was 200 μM (Table 3.5). This concentration was selected for many of the subsequent click chemistry reactions as a result.

Table 3.5. Estimated molecular weights by ImageQuant analysis of SDS-PAGE click chemistry gels of protein and varying concentrations of prg-CMms6. Yield and estimated fraction of proteins with two clicked prg-CMms6 peptides were obtained by the estimated molecular weights, calculated via calibration with the ladder in ImageQuant.

| | | MW (kDa) | | % | | |
|---|---|---|---|---|---|---|
| | | Actual | Observed | Gel shift yield | Disubstituted fraction | Overall yield* |
| C+AHA/Q+AHA + prg-CMms6 | | 10.74 | | | | |
| C+AHA/Q+AHA + 2 x prg-CMms6 | | 15.18 | | | | |
| C+AHA | 100 μM prg-CMms6 | | 9.76 | 78 | | 31 |
| | 200 μM prg-CMms6 | | 10.4 | 92 | | 36 |
| | 400 μM prg-CMms6 | | 11.52 | 100 | 18 | 39 |
| | 500 μM prg-CMms6 | | 11.92 | 100 | 27 | 39 |
| Q+AHA | 100 μM prg-CMms6 | | 9.12 | 64 | | 30 |
| | 200 μM prg-CMms6 | | 9.68 | 76 | | 36 |
| | 400 μM prg-CMms6 | | 11.36 | 100 | 14 | 47 |
| | 500 μM prg-CMms6 | | 12.24 | 100 | 34 | 47 |

Click chemistry reactions on whole cell lysate that contained C+AHA and Q+AHA were successful, as indicated by the results presented (Figure 3.24, Figure 3.25, Figure 3.27, and Figure 3.28). The reaction was also run on pure proteins that were expressed in the presence of either AHA or methionine (Figure 3.29). The results from this experiment showed that bands corresponding to MWs of potential clicked products were not seen. While the intensity of the bands seen on the gel increase with increasing concentration of prg-CMms6 (Figure 3.29), the increase could be attributed to increase in

* The overall reaction yield is obtained by multiplying the yield of the gel shift by the percent incorporation of AHA in each protein (78.3 % for C+AHA and 94.3 % for Q+AHA) and then dividing by a factor of two to account for unclicked protein, which has a band of equal intensity in all lanes.

prg-CMms6 peptide alone (Figure 3.26). As indicated by the ladder in lane 1 of Figure 3.29, MWs between 3.4 and 6.5 kDa were not easily discernible. As prg-CMms6 and C, Q, C+AHA, and Q+AHA fall within this range, we were unable to discriminate between these proteins from this gel. The MWs of the bands seen in Figure 3.29 did not increase indicating that click chemistry of pure C, Q, C+AHA, and Q+AHA with prg-CMms6 in these reaction conditions was unsuccessful (Figure 3.23). The reason for this is that AHA in purified C+AHA and Q+AHA may have degraded upon exposure to light.[255] As purification was done under ambient laboratory light and purified protein was not stored in complete darkness this exposure may have been sufficient to degrade the azide functionalized amino acid. Contrary to pure C+AHA and Q+AHA, whole cell lysate samples containing the over expressed proteins were stored in darkness and not subject to ambient light as they were not purified. The likely degradation of AHA in the purified C+AHA and Q+AHA is the cause for the unsuccessful click chemistry reactions (Figure 3.23).

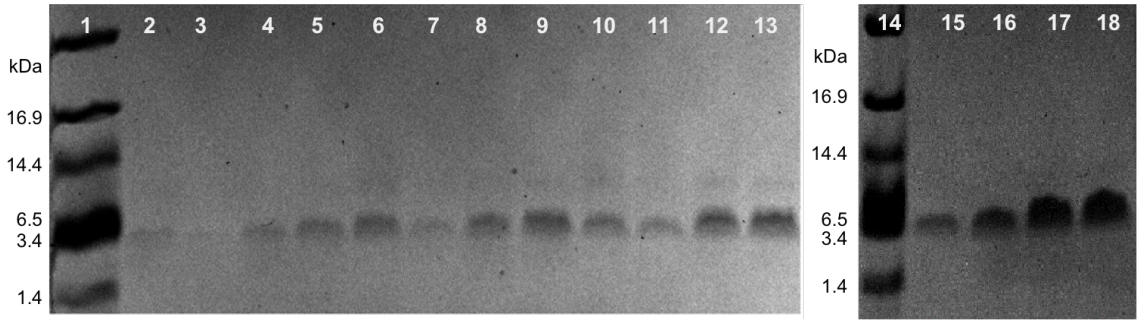



Figure 3.29. SDS-PAGE of click chemistry on pure protein with no sodium ascorbate. 1. and 14. Ladder, 2-5: 100 μM prg-CMms6; 6-9: 200 μM prg-CMms6; 10-13: 400 μM prg-CMms6; 15-19: 500 μM prg-CMms6, where lanes 2, 6, 10, and 15 contain C+AHA; 3, 7, 11, and 16 contain Q+AHA; 4, 8, 12, and 17 contain C; and 5, 9, 13, and 18 contain Q.

In addition to studying the effect of prg-CMms6 concentration on the reaction, we also studied the incubation time in the presence of sodium ascorbate.[273] As mentioned previously, sodium ascorbate is a powerful reducing agent, providing a constant source of Cu(I) ions in solution that assist in improving regioselectivity of the click chemistry reaction.[273] Click chemistry was performed in the presence of sodium ascorbate, in the absence of CMms6 and in the presence of 200 and 500 μM prg-CMms6 over time (Figure 3.24). These reactions were run with whole cell lysate containing C+AHA and

Q+AHA. In the presence of sodium ascorbate and 200 μM of prg-CMms6, a shift in MW was observed after approximately 28 h of incubation (Figure 3.24). This shift was not observed in the absence of prg-CMms6 (lanes 2-5 in Figure 3.30a and b).

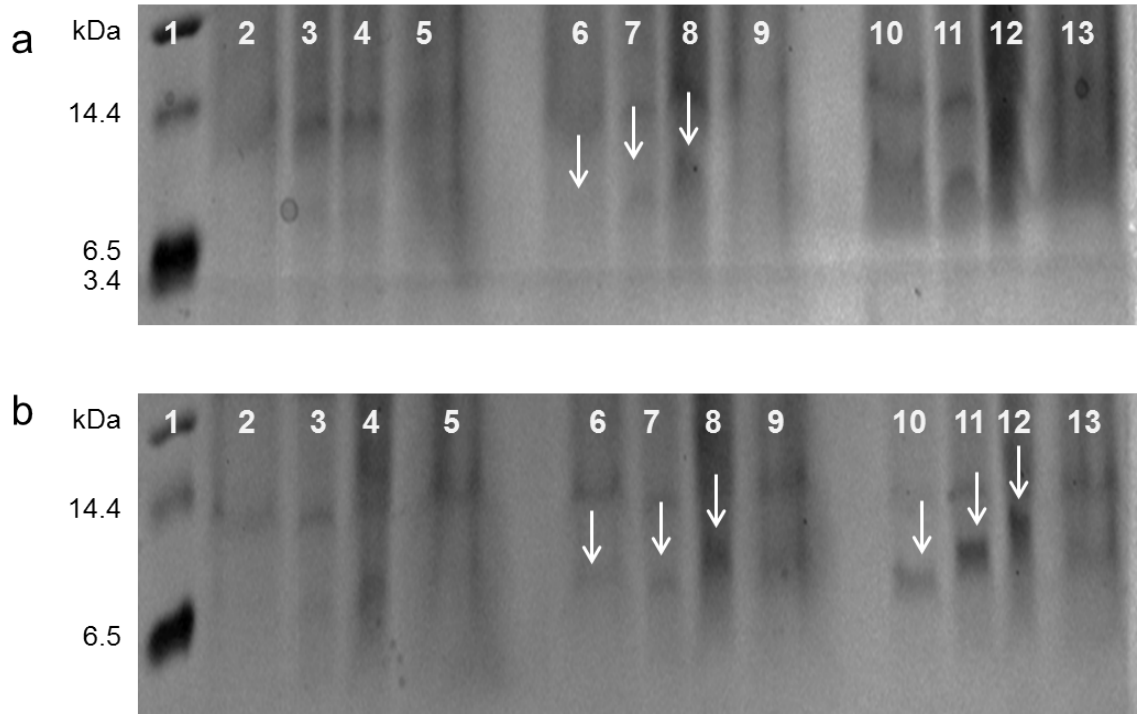



Figure 3.30. SDS-PAGE of click chemistry as a function of time, with sodium ascorbate and varying concentrations of prg-CMms6. (a) C+AHA lysate. 1. Ladder, 2-5: No prg-CMms6, 6-9: 200 μM prg-CMms6, 10-13: 500 μM prg-CMms6, where lanes 2, 6, and 10 were taken at the beginning of the reaction; 3, 7, and 11 represent 20 h of incubation; 4, 8, and 12 represent 28 h of incubation; and 5, 9, and 13 represent 48 h of incubation. (b) Q+AHA lysate. 1. Ladder, 2-5: No prg-CMms6, 6-9: 200 μM prg-CMms6, 10-13: 500 μM prg-CMms6, where lanes 2, 6, and 10 were taken at the beginning of the reaction; 3, 7, and 11 represent 20 h of incubation; 4, 8, and 12 represent 28 h of incubation; and 5, 9, and 13 represent 48 h of incubation.

### *3.4.3 Circular dichroism*

To assess the secondary structure of the AHA containing proteins, CD wavelength scans were performed in conditions of 50 mM PB pH 8. CD was used to evaluate secondary structure in solution conditions for these two proteins and the results of triplicate measurements for each protein (Figure 3.31).

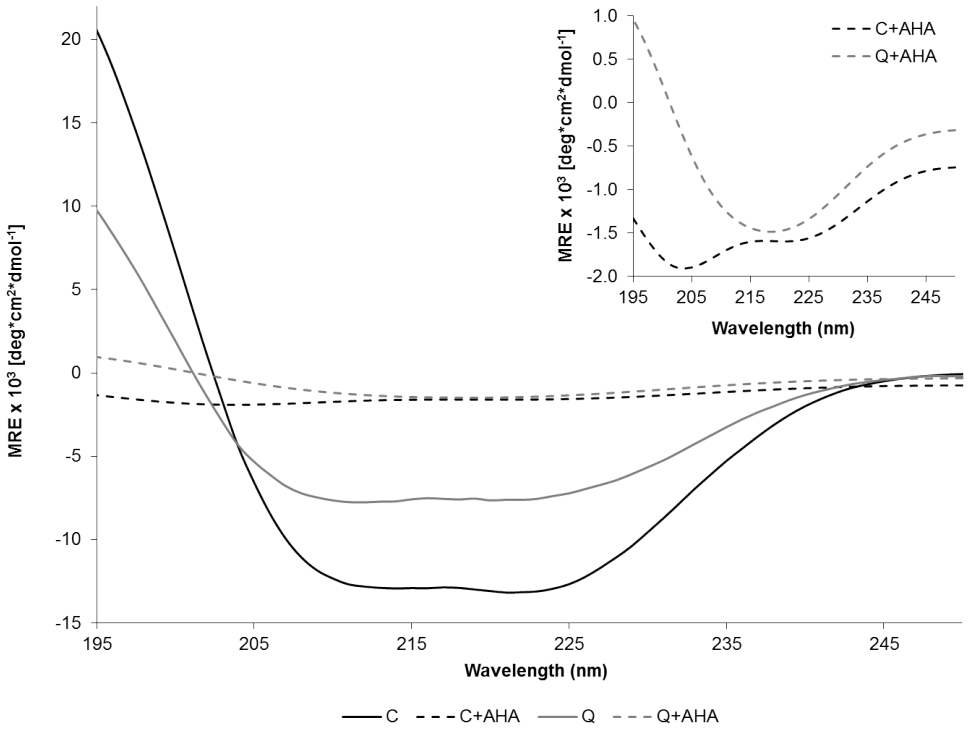



Figure 3.31. Secondary structure of C, C+AHA, Q, and Q+AHA proteins at 50 mM PB pH 8. Circular dichroism variable wavelength scans of C (black, solid), C+AHA (black, dashed), Q (grey, solid), and Q+AHA (grey, dashed). Inset represents scans of C+AHA and Q+AHA at a different y-axis scale. Data is averaged from three replicates and was obtained with 10 µM protein concentration for C and Q and 14 µM protein concentration for C+AHA and Q+AHA.

When compared to C and Q, C+AHA and Q+AHA demonstrate relatively weak CD signatures (Figure 3.31). Azides, as well as products of photolytic decomposition of azides including monoacyl aminals and aldehydes, absorb strongly in the UV-vis region.[255] This may explain the dimished signal of these proteins in Figure 3.31. Despite the weak signal strength of C+AHA and Q+AHA, structural comparisons can nonetheless be made by comparing MRE values at 222 nm and the minimum MRE value between 195-210 nm.[283] A $\theta_{222}/\theta_{min}$ ratio of unity indicates complete $\alpha$-helical character of a protein, and deviations from unity can be calculated as $1-|1-\theta_{222}/\theta_{min}|$.[283] C and Q exhibit theta ratios very close to, or equal to, one, respectively (Table 3.6). C+AHA and Q+AHA possess ratios of 0.84 and 0.78, respectively. Judging from this data, C+AHA and Q+AHA are both slightly less $\alpha$-helical than C and Q at the same conditions.

Table 3.6. MRE values corresponding to $\theta_{222}$, $\theta_{min}$, $\theta_{222}/\theta_{min}$, and the deviation of $\theta_{222}/\theta_{min}$ from unity for C, C+AHA, Q, and Q+AHA as determined from wavelength scans presented in Figure 3.31.

| | C | C+AHA | Q | Q+AHA |
|---|---|---|---|---|
| $\theta_{222}$ | -1.43 | -1.59 | -1.61 | -1.44 |
| $\theta_{min}$ | -3.63 | -1.91 | -4.56 | -1.19 |
| $\theta_{222}/\theta_{min}$ | 0.39 | 0.84 | 0.35 | 1.22 |
| $1-\lvert 1-\theta_{222}/\theta_{min}\rvert$ | 0.39 | 0.84 | 0.35 | 0.78 |

### *3.4.4 TEM*

Transmission electron microscopy was used to study morphological features of AHA containing proteins. Samples were prepared as described elsewhere, and a 1 % uranyl acetate stain was used to help improve contrast when visualizing protein materials. Micrographs of pure C+AHA appeared as predominantly large aggregates on the micrometer scale.

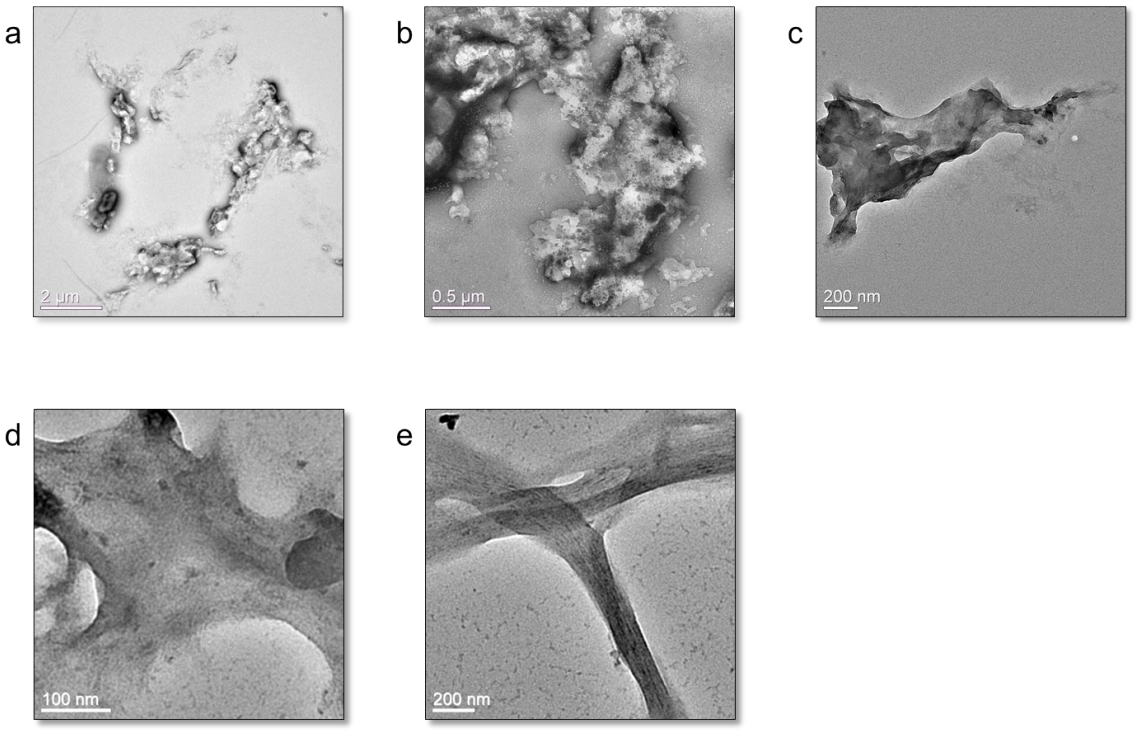



Figure 3.32. Transmission electron micrographs for 23 µM C+AHA 50 mM PB pH 8 (a-c) and 20 µM C at 10 mM PB pH 8 at room temperature (d, e). Predominant morphology seen in C+AHA was aggregates and protofibrils in C (e). Scale bars are 2 µm in (a), 0.5 µm in (b), 200 nm in (c) and (e), and 100 nm in (d).

The aggregates did not have any apparent regularity in their structures (Figure 3.32a-c). Figure 3.32c appears to be more sheet-like in comparison with other microstructures formed by C+AHA that were more globular (Figure 3.32a, b). C under the same buffer conditions presents itself in sheet-like formations or as fibers (Figure 3.32d, e) with fewer aggregates than were seen in C+AHA.

In contrast, Q+AHA presented fibrous protein structures (Figure 3.33b, c) in addition to aggregates (Figure 3.33a). The fibers formed by Q+AHA in these conditions (50 mM PB, pH 8) were reminiscent of fibers formed by Q under acidic conditions (pH 4), with diameters on the order of 100 nm and showing clear striations within the fibers indicating the bundling of pentameric protofibrils (Figure 3.33b). Protofibrils formed by Q+AHA at pH 8 averaged 5.1 ± 0.9 nm, which was larger than the protofibrils of Q under similar conditions, which were 2.6 ± 0.4 nm (Table 3.7). The larger protofibril diameter may result from the increase bulkiness of the incorporated AHA residue, causing pentamers to obtain a slightly larger outer diameter. Q+AHA fibers at pH 8 were also seen to form bundles or larger aggregates of fibers (Figure 3.33c), which were not commonly seen by Q at pH 4. The fact that Q+AHA fibers were formed under neutral solvent conditions and Q fibers were scarcer under these conditions points to increased stability due to AHA incorporation. Aggregation of these AHA protein fibers is also likely due to the incorporation of the azide functionalized unnatural amino acid.

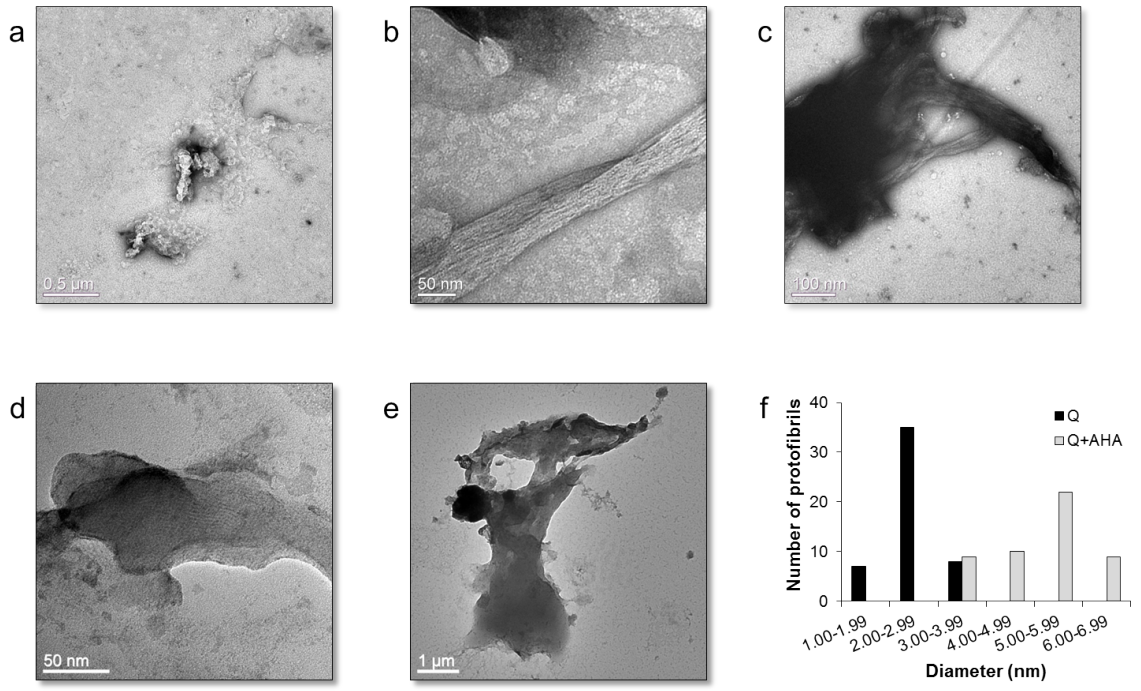



Figure 3.33. Transmission electron micrographs for 11 µM Q+AHA 50 mM PB pH 8 (a-c) and 20 µM Q at 10 mM PB pH 8 at room temperature (d, e). Q+AHA contained some aggregates and some fibers, which appeared in aggregated networks, whereas Q appeared as sheets with some observable striations (d). (e). Scale bars are 0.5 µm in (a), 50 nm in (b) and (d), 100 nm in (c), and 1 µm in (e). Histogram of protofibril diameters ($n$ = 50) of Q (black) and Q+AHA (grey), 50 mM PB pH 8.

Table 3.7. Average and standard deviations of protofibril diameters for Q at pH 4 ($n$ = 210) and pH 8 ($n$ = 50) and Q+AHA at pH 8 ($n$ = 50).

| | Q | | Q+AHA |
|---|---|---|---|
| **Protofibril diameter** | **pH 4** | **pH 8** | **pH 8** |
| **Average (nm)** | 3.5 | 2.6 | 5.1 |
| **Standard deviation (nm)** | 0.7 | 0.4 | 0.9 |

prg-CMms6 was visualized at 200 µM peptide concentration via TEM (Figure 3.34). prg-CMms6 formed large aggregates that were on the micrometer scale. The high concentration of prg-CMms6 used in TEM studies resulted in aggregates covering the majority of the carbon grid (Figure 3.29).

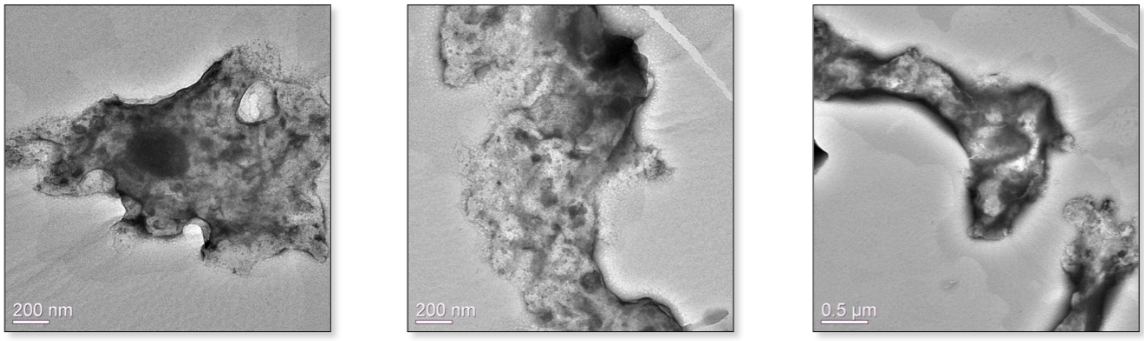



Figure 3.34. Transmission electron micrographs for 200 µM prg-CMms6 in 50 mM PB, pH 8. Predominant morphology seen was aggregates. Scale bars are 200 nm (left and center) and 0.5 µm (right).

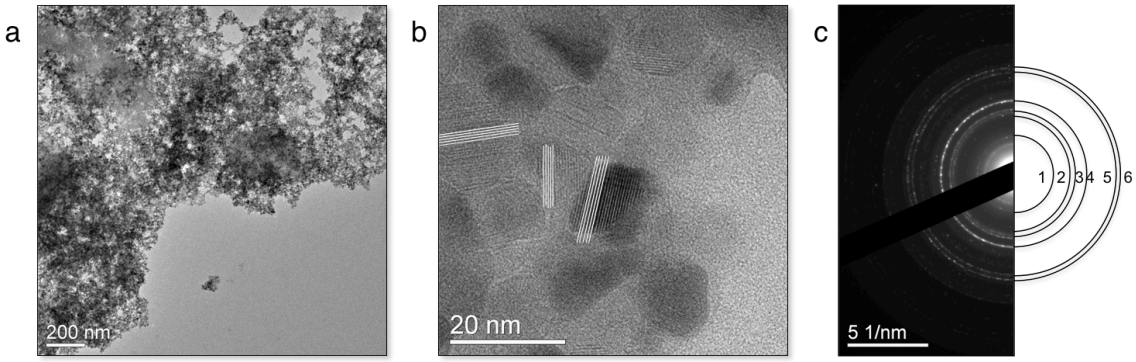



Figure 3.35. Transmission electron micrographs for MNPs formed via coprecipitation in 50 mM PB, pH 8 (a, b). White lines in (b) are superimposed on lattice fringes observed at high resolution. Electron diffraction pattern of MNPs (c) displayed rings that were used to calculate *d*-spacing of the crystals. Scale bars are 200 nm (a), 20 nm (b), and 5 $nm^{-1}$ (c).

MNPs reduced in the absence of protein were visualized in TEM (Figure 3.35). The nanoparticles that were observed were in very large clusters (Figure 3.35a), which was most likely an artifact of drying the soluble MNPs on the carbon grid during TEM sample preparation. MNPs were of regular size and very narrow size distribution (Figure 3.38), with an average diameter of 8.6 ± 1.2 nm. Shape of the MNPs was octahedral or almost spherical. This MNP dimension and shape was in agreement with the original synthesis of these particles by Massart *et al.*, who observed particles that were roughly spherical and 8 nm in diameter.[284] High resolution micrographs were obtained that showed lattice fringes within the magnetic crystals, outlined with white lines (Figure 3.35b). The lattice spacing determined from measuring the fringes in Figure 3.35b was determined to be 5.32 ± 0.21 Å. This is in excellent agreement with the *d*-spacing

calculated from ring 1 of the electron diffraction rings (Figure 3.35c), which was 5.36 ± 0.06 Å (Table 3.8).

Table 3.8. Calculation of *d*-spacing for the rings resulting from electron diffraction pattern of MNPs (Figure 3.35c) and corresponding crystallographic planes.

| Ring | *d*-spacing (Å) | Crystallographic plane |
|---|---|---|
| 1 | 5.36 ± 0.06 | |
| 2 | 2.69 ± 0.02 | |
| 3 | 2.48 ± 0.02 | (311) |
| 4 | 2.17 ± 0.02 | (321) |
| 5 | 1.58 ± 0.01 | (511) |
| 6 | 1.50 ± 0.01 | (440) |

The estimation of the main lattice spacing for the MNPs is more accurately determined through measurement of the diffraction rings, as is indicated by a smaller standard deviation of these values. The intensity of the rings is indicative of the predominance of a certain crystallographic plane (Figure 3.35c); the planes represented by the more intense rings (ring 4, for example) occur more frequently within the crystal structure (Figure 3.35c). The diffraction patterns shown by MNPs crystallized in the presence of prg-CMms6 were identical to those shown in Figure 3.35c (Figure 3.36c), indicating the same crystal structure of the MNPs synthesized in the presence or absence of protein. Lattice fringes on high resolution TEM images were more difficult to see on MNPs formed in the presence of prg-CMms6, however (Figure 3.36b), more likely due to coating of the MNPs with protein.

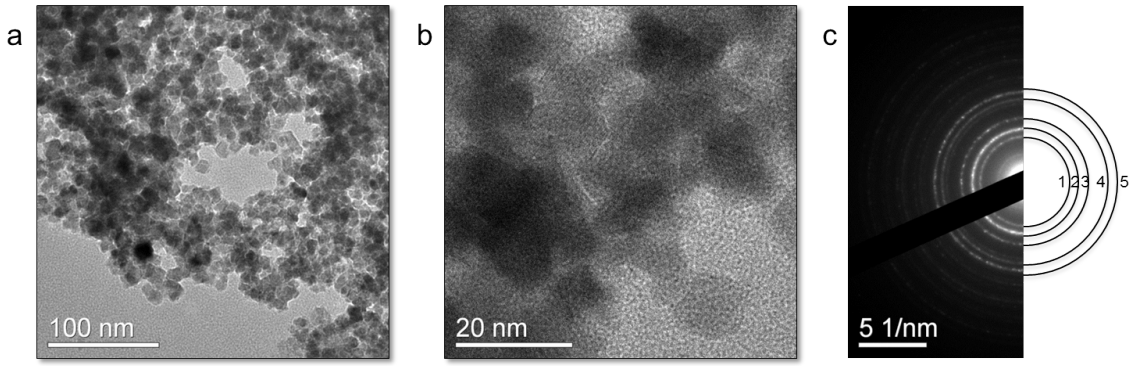



Figure 3.36. Transmission electron micrographs for MNPs formed in the presence of 1 mM prg-CMms6 via coprecipitation in 50 mM PB, pH 8 (a, b). Electron diffraction pattern of MNPs formed in the presence of

prg-CMms6 (c) displayed rings that were used to calculate *d*-spacing of the crystals. Scale bars are 100 nm (a), 20 nm (b), and 5 $nm^{-1}$ (c).

Table 3.9. Calculation of *d*-spacing for the rings resulting from electron diffraction pattern of MNPs formed in the presence of 1 mM prg-CMms6 (Figure 3.36c) and corresponding crystallographic planes.

| Ring | *d*-spacing (Å) | Crystallographic plane |
|---|---|---|
| 1 | 3.00 ± 0.04 | |
| 2 | 2.56 ± 0.03 | (331) |
| 3 | 2.10 ± 0.03 | (321) |
| 4 | 1.62 ± 0.02 | (511) |
| 5 | 1.48 ± 0.02 | (440) |

Based on the diffraction data (Figure 3.35, Table 3.8, Table 3.9), it is not possible to distinguish between magnetite ($Fe_3O_4$) and maghemite ($\gamma$-$Fe_2O_3$) (Table 3.1). As explained in Section 3.2.1, magnetite and maghemite are indistinguishable when examined with X-ray or electron diffraction techniques, as the *d*-spacings for each crystal structure are within fractions of an Ångström from one another. Magnetite is not very stable and in the presence of oxygen is easily converted to maghemite, therefore it is within reason that a combination of both magnetite and maghemite are present in these samples.

MNPs synthesized in the presence of the magnetite binding peptide prg-CMms6 resulted in spherical, uniform crystals (Figure 3.37). The protein matrix could be seen in the micrographs with as a film in which the MNPs were embedded with a low level of contrast, as no stain was used in imaging these samples. This film was notably absent in Figure 3.35a. Several concentrations of prg-CMms6 were used in these studies, ranging from 200 μM – 1 mM. Similar to the MNPs in buffer, the nanoparticles appeared aggregated in TEM images, which was due to drying of the samples while preparing the TEM grids. Measurement of the MNPs with prg-CMms6 using ImageJ software ($n$ = 50) revealed that the average diameter of the MNPs decreased as prg-CMms6 concentration increased.With 200 μM prg-CMms6 the MNPs measured 16.3 ± 2.2 nm, 500 μM prg-CMms6 resulted in MNPs 15.1 ± 3.0 nm in diameter, and 1 mM prg-CMms6 templated MNPs 12.2 ± 1.6 nm in diameter (Figure 3.37). A linear regression of the concentration versus MNP diameter returned an $R^2$ value of 0.993, indicating a very strong linear

relationship, defined by Equation 3.12, where $D$ equals the MNP diameter in nm and $C$ is CMms6 concentration in μM.

$$D = -0.0053C + 17.501$$

Equation 3.12

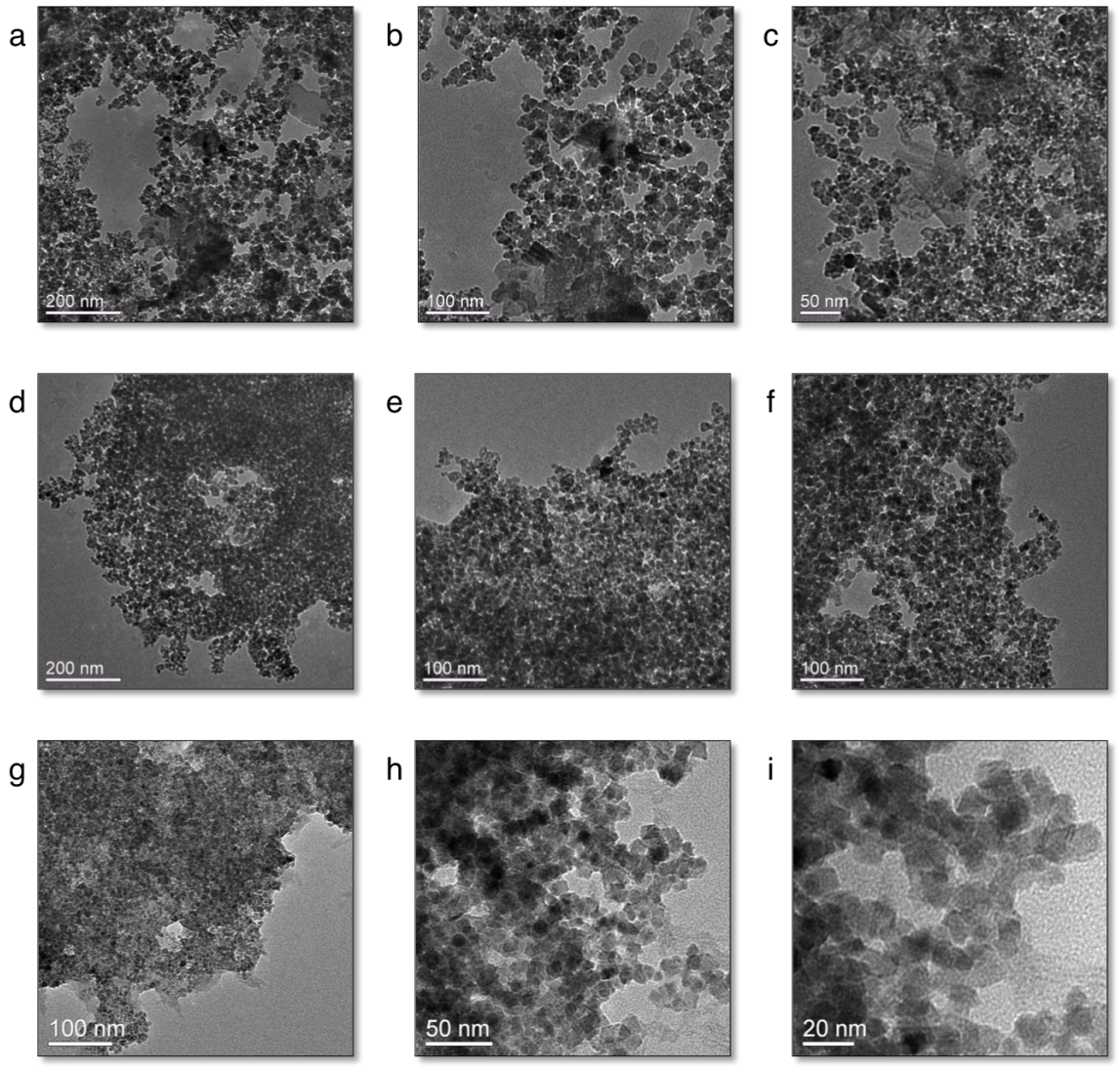



Figure 3.37. TEM of 200 μM prg-CMms6 + MNP (a-c), 500 μM prg-CMms6 + MNP (d-f), and 1 mM prg-CMms6 + MNP (g-i). Scale bars represent 200 nm in (a) and (d), 100 nm in (b) and (e-g), 50 nm in (c) and (h), and 20 nm in (i).

The size distribution of MNPs that were templated with prg-CMms6 was fairly narrow across all concentrations studied, but more broad than MNPs crystallized in only 50 mM PB (Figure 3.38a). MNPs formed through coprecipitation in the presence of 5.6

μg/mL Mms6 (the magnetite binding peptide in its entirety) were 21.2 ± 8.3 nm, with a narrow size distribution.[218] The uniformity in shape could be explained by hydroxyl group rich C-terminal section of one prg-CMms6 that bound tightly to MNPs and template for magnetite crystal formation, directing the shape of magnetite crystals formed, as previously observed by Arakaki *et al*.[229]

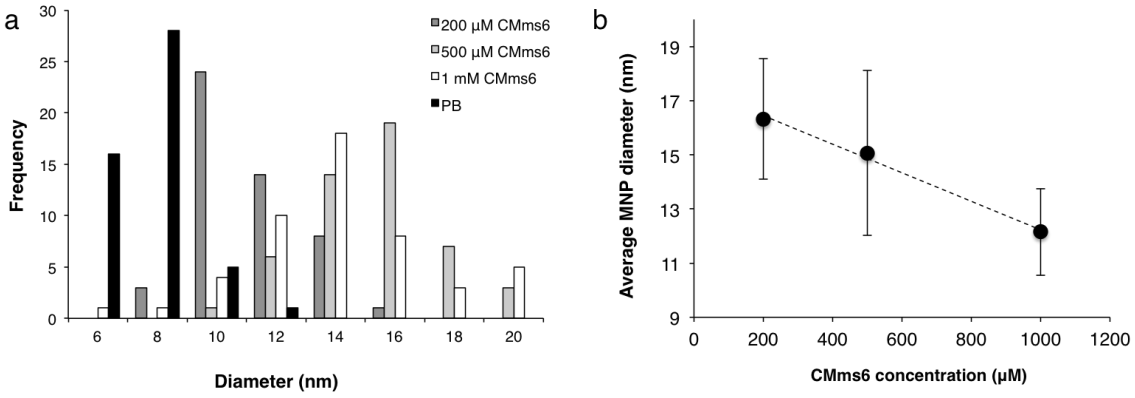



Figure 3.38. Histogram of MNP diameters ($n$ = 50) reduced in the presence of 50 mM PB only (black) and 200 μM (dark grey), 500 μM (light grey), and 1 mM prg-CMms6 (white) (a). Negative linear correlation between prg-CMms6 concentration and MNP size (b).

MNPs were also reduced in the presence of only pure C+AHA (23 μM) or Q+AHA (11 μM) in 50 mM PB pH 8 conditions (Figure 3.39). These nanoparticles were not uniform in morphology, with a high concentration of needle-like crystalline iron oxides (indicated by white arrows in Figure 3.39). As can bee seen in Figure 3.39, there is a mixed morphology of the iron oxide particles, with roughly 20 % of the crystals assuming the needle-like conformation. Arakaki *et al.* also observed similar amounts of needle shaped crystals in the absence of CMms6 produced by coprecipitation, which were similar to $\alpha$-FeOOH.[229]

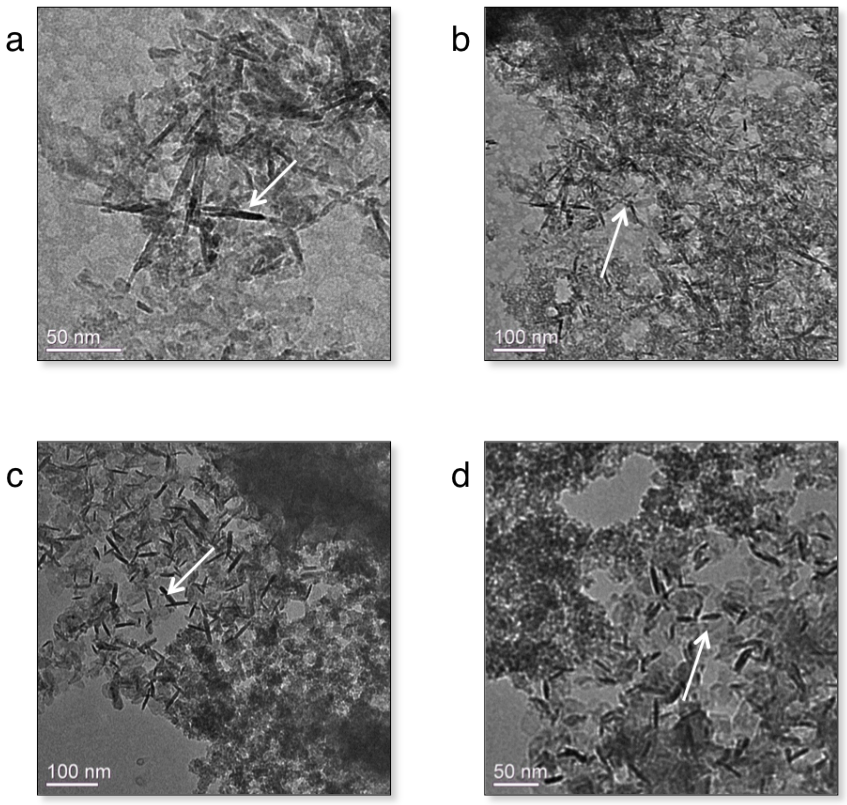



Figure 3.39. TEM of unclicked C+AHA + MNP (a, b) and unclicked Q+AHA + MNP (c, d). Concentration of C+AHA was 23 µM and Q+AHA was 11 µM, with molar ratios of iron salts and reducing agent normalized accordingly. Presence of needle-like crystals is indicated by white arrows. Scale bars represent 100 nm in (b) and (c) and 50 nm in (a) and (d). Samples imaged without uranyl acetate stain.

### *3.4.5 Optical microscopy*

Phase contrast microscopy was used to observe manipulation of MNP and protein composites synthesized via coprecipitation under magnetic influence at various stages of magnetic separation.

#### 3.4.5.1 MNPs in PB

MNPs synthesized in the absence of protein appear spherical and are highly susceptible to an applied magnetic field, as was observed by rapid movement of the particles through the buffer solution (Figure 3.40, see time-lapse videos at http://bit.ly/1BWjhCe). Particles appear to be aggregated and travel in flocculates (http://bit.ly/1BWjhCe).

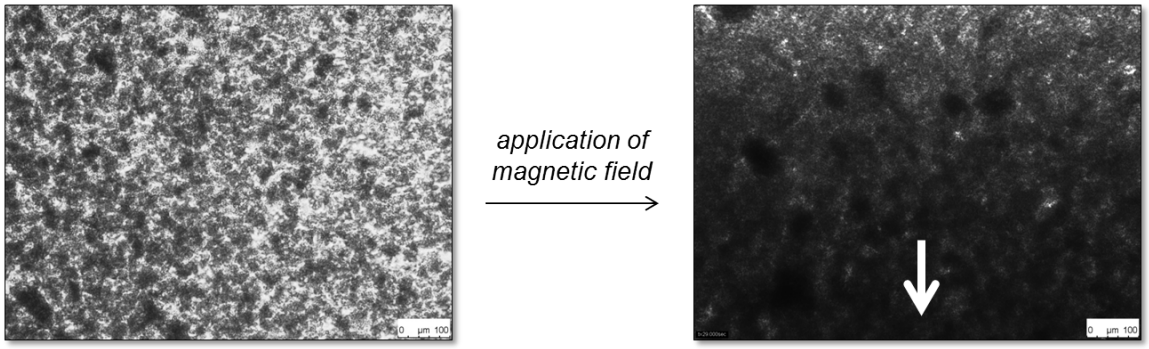



Figure 3.40. Phase contrast microscopy of synthesized MNPs in 50 mM PB pH 8 before and after the application of a magnetic field. Direction of the magnetic field is indicated by the white arrow. Scale bars represent 100 μm.

### 3.4.5.2 MNPs in CMms6

MNPs synthesized in the presence of 200 μM prg-CMms6 protein appear spherical and are highly susceptible to an applied magnetic field, as was observed by rapid movement of the particles through the buffer solution (Figure 3.41, see time-lapse videos at http://bit.ly/1DgKd4d). Prg-CMms6 was shown to assemble MNPs in chains, which align with a magnetic field when applied. These chains are very similar to the magnetosome chains formed by magnetotactic bacteria, whose purpose is to align the bacterium with magnetic fields.[205,218,233] Interestingly, time-lapse videos show that these prg-CMms6+MNP assemblies do not move significantly when a magnetic field is applied, but instead remain relatively stationary while being aligned (http://bit.ly/1DgKd4d). It is, in fact, the biological function of Mms6 to align chains of MNPs but not to travel towards a particular pole.[205,218,233]

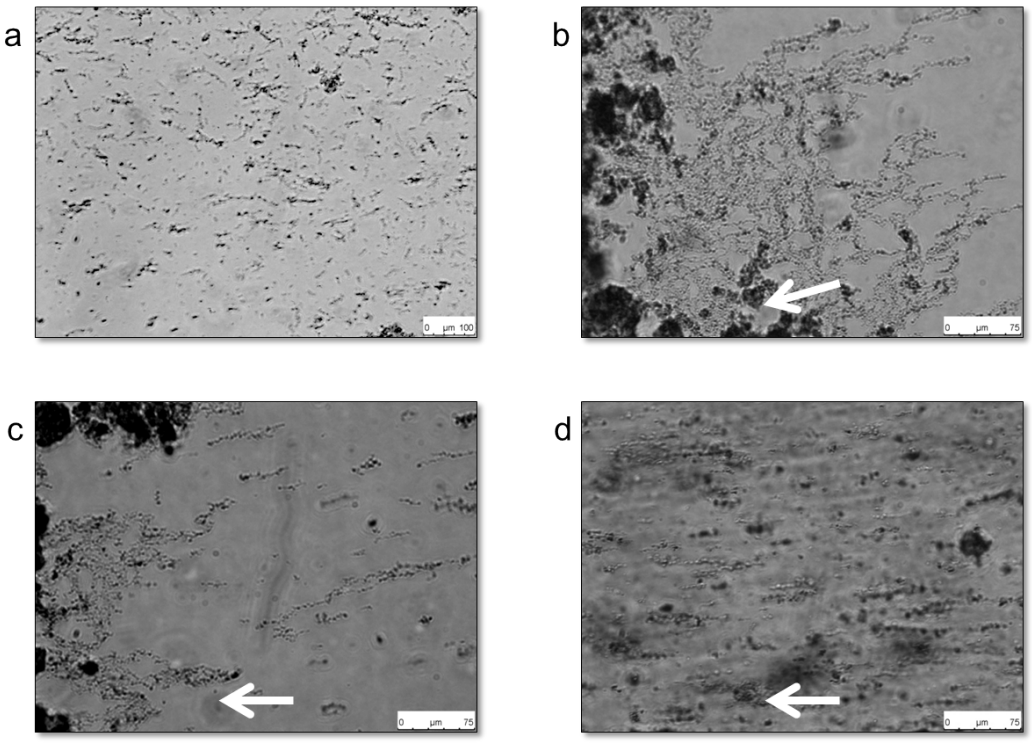



Figure 3.41. Phase contrast microscopy of MNPs synthesized in the presence of 200 μM prg-CMms6 in 50 mM PB pH 8 before (a) and after (b, c, d) the application of a magnetic field. Direction of the magnetic field is indicated by the white arrows. Scale bar represents 100 μm in (a) and 75 μm in (b), (c), and (d).

#### 3.4.5.3 Cell lysate washing

MNPs synthesized in buffer were incubated with lysate that contained over expressed C+AHA, Q+AHA, or clicked C+AHA+prg-CMms6 or Q+AHA+prg-CMms6. After incubation, solutions were subjected to four rounds of magnetic separation as described in Section 3.3.6.1. Samples of the MNP and cell lysate containing only C+AHA generated particles were very heterogenous in size (Figure 3.42). Smaller, more discrete, particles were seen in the presence of cell lysate containing C+AHA+prg-CMms6 (Figure 3.42). The manipulation of these solutions with a neodymium magnet demonstrated that C+AHA+prg-CMms6 samples were retarded in traveling through the solution compared to MNPs in C+AHA samples (time-lapse videos of C+AHA+MNP: http://bit.ly/1ucyVff, time-lapse videos of C+AHA+prg-CMms6+MNP: http://bit.ly/1C2tRu5). Particle retardation was also observed in prg-CMms6+MNPs (http://bit.ly/1DgKd4d).

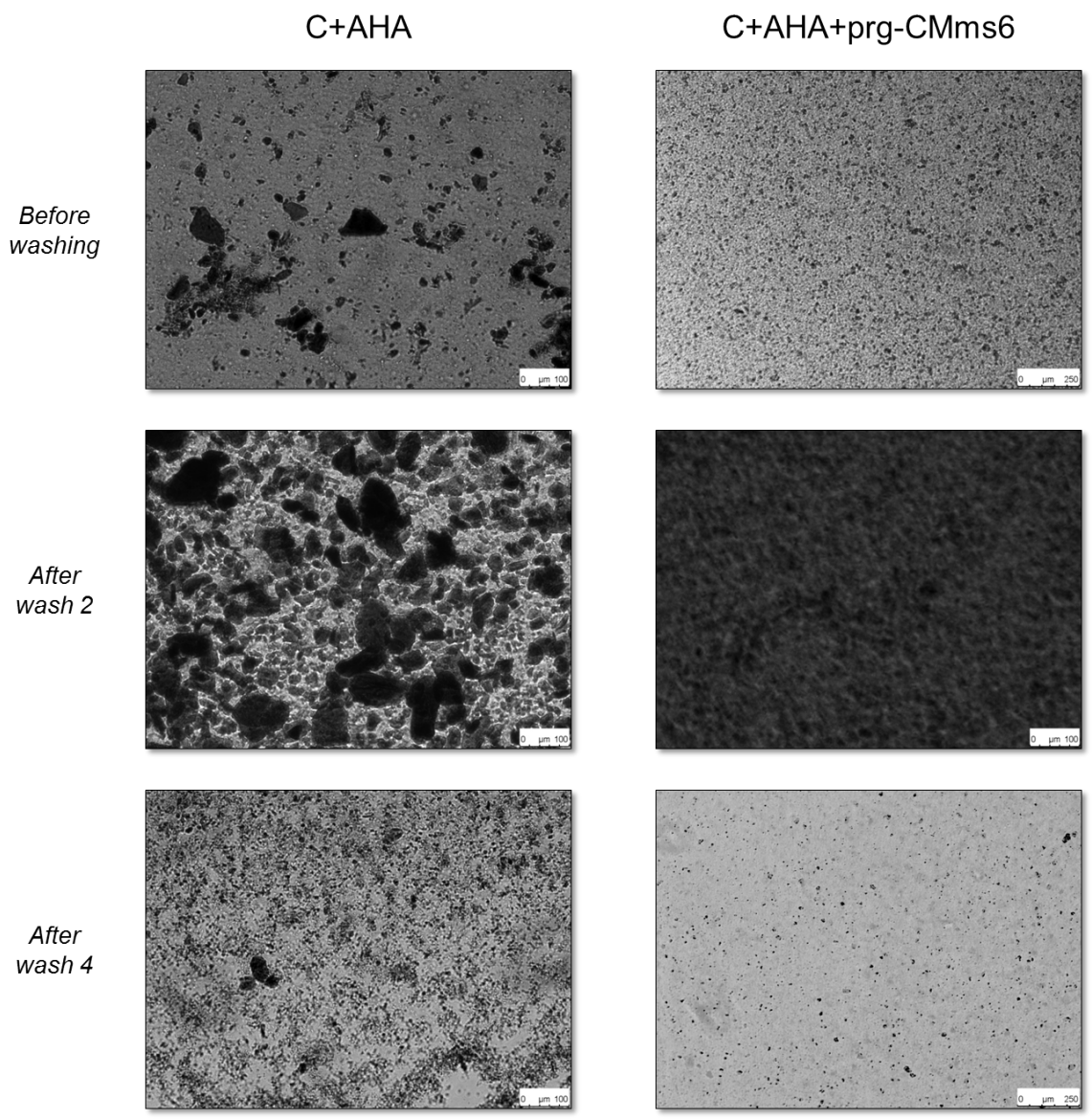



Figure 3.42. Phase contrast microscopy of MNPs synthesized in 50 mM PB pH 8 incubated with C+AHA lysate and C+AHA+prg-CMms6 lysate before (top) and after 2 washes (middle) and 4 washes (bottom) via magnetic separation. Scale bar represents 100 μm in top left, middle, and bottom left images and 250 μm in top right and bottom right images.

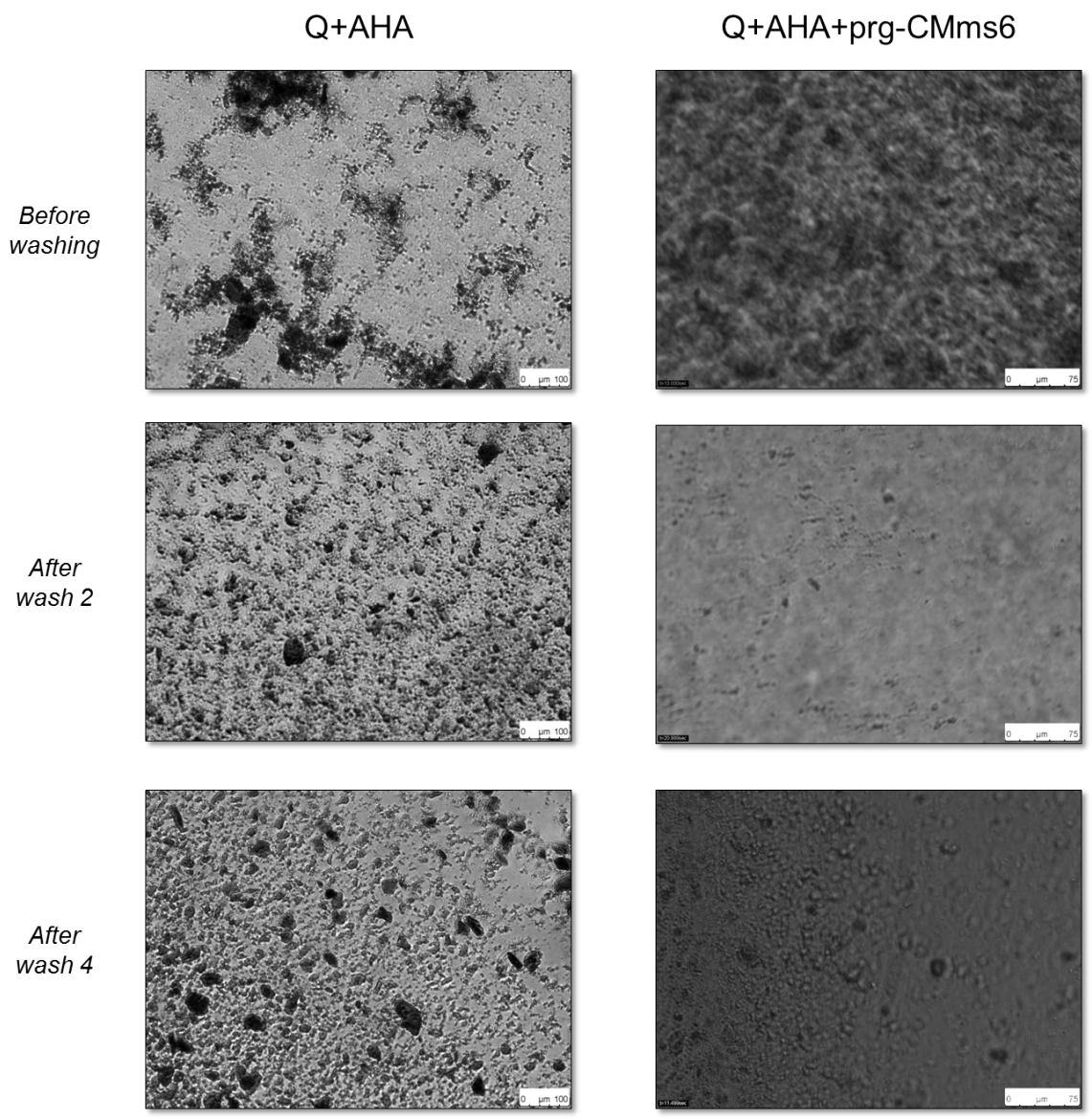



Figure 3.43. Phase contrast microscopy of MNPs synthesized in 50 mM PB pH 8 incubated with Q+AHA lysate and Q+AHA+prg-CMms6 lysate before (top) and after 2 washes (middle) and 4 washes (bottom) via magnetic separation. Scale bar represents 100 μm in top left, middle left, and bottom left images and 75 μm in top right, middle right, and bottom right images.

Heterogeneous, aggregated particles were seen in MNP and Q+AHA cell lysate samples (Figure 3.43). Similarly to C+AHA+prg-CMms6, smaller, more discrete, particles were seen in the presence of cell lysate containing Q+AHA+prg-CMms6 (Figure 3.43). The manipulation of these solutions with a neodymium magnet demonstrated that Q+AHA+prg-CMms6 samples were retarded in traveling through the solution compared to MNPs in Q+AHA samples (time-lapse videos of Q+AHA+MNP:

http://bit.ly/1AuKUae, time-lapse videos of Q+AHA+prg-CMms6+MNP: http://bit.ly/1zbceDT). Cohesive collections of MNPs that were traveling together in assemblies were viewed in Q+AHA+prg-CMms6 samples, particularly after the second wash (Figure 3.43, http://bit.ly/1zbceDT). These assemblies also traveled very slowly through the solution, which may be a result of the presence of prg-CMms6 (http://bit.ly/1zbceDT).

### *3.4.6 Magnetometry*

Magnetometry was performed using a SQUID magnetometer* to obtain hysteresis curves for MNPs reduced in the presence of only 50 mM PB pH 8 or pure prg-CMms6 (Figure 3.44a) or in whole cell lysate (Figure 3.44b) containing AHA proteins. There is a measureable difference of MNPs in the presence of prg-CMms6 versus in its absence (Figure 3.44a), supporting the conclusion that prg-CMms6 is capable of templating and organizing MNPs in solution and that they are responsive to applied magnetic fields. The preliminary data presented in Figure 3.44 indicates a diminished magnetization of MNPs templated in the presence of whole cell lysate. To address this, hysteresis curves should be obtained with MNPs templated in the presence of purified protein constructs.

* Experiments performed in collaboration with the University of Florida.

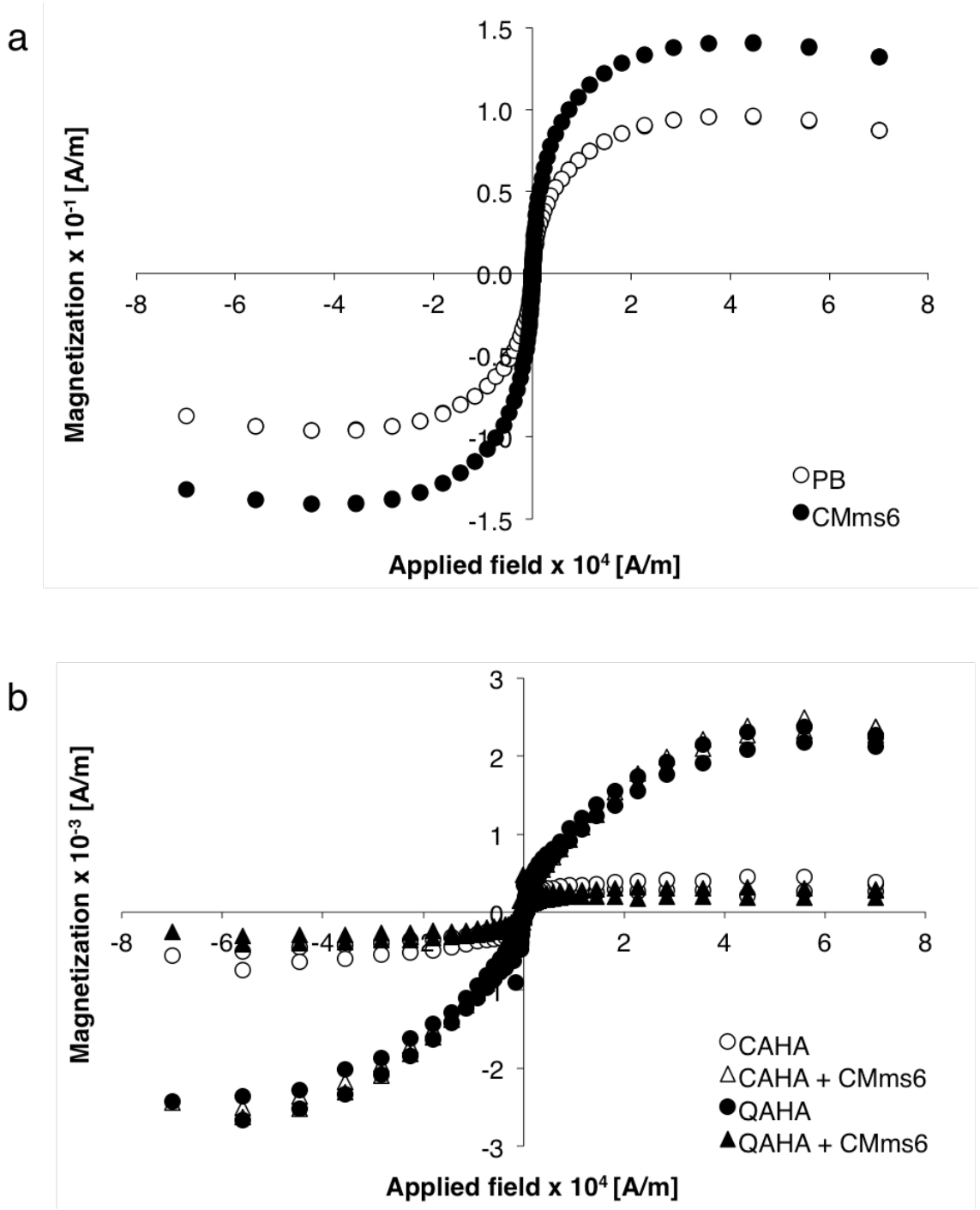



Figure 3.44. Magnetization measurements as a function of applied field for MNPs synthesized in the presence of 50 mM PB pH 8 (white circles) and CMms6 (black circles) (a) and in the presence of whole cell lysate containing expressed C+AHA (white circles) and Q+AHA (black circles), or clicked C+AHA + CMms6 (white triangles) and Q+AHA + CMms6 (black triangles) (b).

## 3.5 Conclusions and future work

Techniques of unnatural amino acid incorporation have been used to introduce AHA residues in discrete locations within the C and Q proteins. C+AHA and Q+AHA were used in click chemistry reactions that functionalized the proteins first with a fluorescent probe with an alkyne moiety, Chromeo494, which enabled us to determine appropriate reaction conditions to achieve high yields of desired reaction products. Upon determination of those conditions, the C-terminal portion of the magnetite binding peptide prg-CMms6, another alkyne-functionalized ligand, was used in the copper catalyzed click chemistry reaction. SDS-PAGE analysis of click chemistry with varying concentrations of prg-CMms6 and reaction times showed that effective conjugation with AHA proteins was achieved with 200 μM prg-CMms6 with a single prg-CMms6 peptide reacting to each protein subunit on average. Click chemistry was effective when overexpressed AHA proteins were contained in whole cell lysate, but purified AHA proteins did not show successful reaction with prg-CMms6 under similar reaction conditions. This can be attributed to the photolytic degradation of AHA while exposed to ambient light during purification. Unfortunately clicked reaction products obtained from the whole cell lysate samples were not successfully purified, likely because these samples were in native conditions and complexation with Ni-NTA beads used for purification may have been incomplete.

Transmission electron microscopy showed that we were able to synthesize magnetically active iron oxide nanoparticles via coprecipitation of $FeSO_4$ and $FeCl_3$ upon reduction with NaOH. In the absence of any proteins these nanoparticles were 8.6 ± 1.2 nm in diameter, and in the presence of prg-CMms6 it was observed that increasing concentrations of the magnetite binding peptide resulted in a decreasing MNP diameter that followed a linear relationship between 200 μM and 1 mM prg-CMms6. Due to the reduction of MNPs with NaOH, solution conditions go from pH 8 to significantly more basic after reduction of MNPs. As a result, proteins may be destabilized compared to neutral or acidic solution conditions (see thermodynamic properties in Table 1.15). Future work should include titrating the solution pH back to neutral or even acidic pH levels to ensure protein stability, as MNPs will be stable after crystallization. Magnetic

separation of these assemblies demonstrated that MNPs incubated with C+AHA+prg-CMms6 and Q+AHA+prg-CMms6 resulted in assemblies that retarded the movement of magnetic material through solution, which may be due to the contribution of conjugated prg-CMms6. Future work should include characterizing MNP templation on clicked, pure protein and magnetization studies of these nanocomposites, as protein-stabilized MNPs have tremendous potential for use in biomedical imaging, targeted drug delivery, and magnetic separation.

# 4 References


1. Hume, J. *et al.* Engineered Coiled-Coil Protein Microfibers. *Biomacromolecules* **15,** 3503–3510 (2014).

2. Carmeli, I. *et al.* Broad Band Enhancement of Light Absorption in Photosystem I by Metal Nanoparticle Antennas. *Nano Lett.* **10,** 2069–2074 (2010).

3. Fan, T.-X., Chow, S.-K. & Zhang, D. Biomorphic mineralization: From biology to materials. *Prog. Mater. Sci.* **54,** 542–659 (2009).

4. Slocik, J. M. & Wright, D. W. Biomimetic mineralization of noble metal nanoclusters. *Biomacromolecules* **4,** 1135–41 (2003).

5. Wu, C., Yin, P., Zhu, X., OuYang, C. & Xie, Y. Synthesis of hematite (alpha-Fe2O3) nanorods: diameter-size and shape effects on their applications in magnetism, lithium ion battery, and gas sensors. *J. Phys. Chem. B* **110,** 17806–12 (2006).

6. Mann, S. Molecular tectonics in biomineralization and biomimetic materials chemistry. *Nature* **365,** 499–505 (1993).

7. Lehn, J.-M. Perspectives in Supramolecular Chemistry—From Molecular Recognition towards Molecular Information Processing and Self-Organization. *Angew. Chemie Int. Ed. English* **29,** 1304–1319 (1990).

8. Sanchez, C., Arribart, H. & Guille, M. M. G. Biomimetism and bioinspiration as tools for the design of innovative materials and systems. *Nat. Mater.* **4,** 277–288 (2005).

9. Howorka, S. Rationally engineering natural protein assemblies in nanobiotechnology. *Curr. Opin. Biotechnol.* **22,** 485–91 (2011).

10. Goldsmith, L. E., Pupols, M., Kickhoefer, V. a, Rome, L. H. & Monbouquette, H. G. Utilization of a protein ‘shuttle’ to load vault nanocapsules with gold probes and proteins. *ACS Nano* **3,** 3175–83 (2009).

11. Chan, D. I. & Vogel, H. J. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430,** 1–19 (2010).

12. Yu, S. M., Li, Y. & Kim, D. Collagen mimetic peptides: progress towards functional applications. *Soft Matter* **7,** 7927 (2011).

13. Hom, N. *et al.* Anisotropic nanocrystal arrays organized on protein lattices formed by recombinant clathrin fragments. *J. Mater. Chem.* **22,** 23335–23339 (2012).

14. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510,** 103–108 (2014).

15. Gore, T., Dori, Y., Talmon, Y., Tirrell, M. & Bianco-peled, H. Self-Assembly of Model Collagen Peptide Amphiphiles. *Langmuir* **17,** 5352–5360 (2001).

16. Soto, C., Brahes, M. C., Alvarez, J. & Inestrosa, N. C. Structural Determinants of the Alzheimer's Amyloid b-Peptide. *J. Neurochem.* **63,** 1191–1198 (1994).

17. Rousseau, F., Schymkowitz, J. & Serrano, L. Protein aggregation and amyloidosis: confusion of the kinds? *Curr. Opin. Struct. Biol.* **16,** 118–126 (2006).

18. Kramer, R. M., Crookes-Goodson, W. J. & Naik, R. R. The self-organizing properties of squid reflectin protein. *Nat. Mater.* **6,** 533–538 (2007).

19. Crookes, W. J. *et al.* Reflectins: the unusual proteins of squid reflective tissues. *Science (80-. ).* **303,** 235–238 (2004).

20. Miserez, A., Li, Y., Cagnon, J., Weaver, J. C. & Waite, J. H. Four-stranded coiled-coil elastic protein in the byssus of the giant clam, Tridacna maxima. *Biomacromolecules* **13,** 332–341 (2012).

21. O'Leary, L. E. R., Fallas, J. A., Bakota, E. L., Kang, M. K. & Hartgerink, J. D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nat. Chem.* **3,** 821–8 (2011).

22. Singh, A., Hede, S. & Sastry, M. Spider silk as an active scaffold in the assembly of gold nanoparticles and application of the gold-silk bioconjugate in vapor sensing. *Small* **3,** 466–473 (2007).

23. Lazaris, A. *et al.* Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science (80-. ).* **295,** 472–6 (2002).

24. Anderson, V. L., Ramlall, T. F., Rospigliosi, C. C., Webb, W. W. & Eliezer, D. Identification of a helical intermediate in trifluoroethanol-induced alpha-synuclein aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **107,** (2010).

25. Wei, G., Reichert, J., Bossert, J. & Jandt, K. D. Novel biopolymeric template for the nucleation and growth of hydroxyapatite crystals based on self-assembled fibrinogen fibrils. *Biomacromolecules* **9,** 3258–67 (2008).

26. MacArthur, I. Structure of alpha keratin. *Nature* **152,** 38–41 (1943).

27. Dos Remedios, C. G. & Thomas, D. D. in *Results Probl. Cell Differ.* **32,** 1–7 (2001).

28. Buehler, M. J. Nanomechanics of collagen fibrils under varying cross-link densities: atomistic and continuum studies. *J. Mech. Behav. Biomed. Mater.* **1,** 59–67 (2008).

29. Kim, C. A. *et al.* Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* **20,** 4173–82 (2001).

30. Cohen, C. & Parry, D. a. Alpha-helical coiled coils and bundles: how to design an alpha-helical protein. *Proteins Struct. Funct. Genet.* **7,** 1–15 (1990).

31. Apostolovic, B., Danial, M. & Klok, H.-A. Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. *Chem. Soc. Rev.* **39,** 3541–75 (2010).

32. Woolfson, D. N. & Mahmoud, Z. N. More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials. *Chem. Soc. Rev.* **39,** 3464–79 (2010).

33. Gunasekar, S. K., Haghpanah, J. S. & Montclare, J. K. Assembly of bioinspired helical protein fibers. *Polym. Adv. Technol.* **19,** 454–468 (2008).

34. Woolfson, D. N. & Ryadnov, M. G. Peptide-based fibrous biomaterials: Some things old, new and borrowed. *Curr. Opin. Chem. Biol.* **10,** 559–67 (2006).

35. Ryadnov, M. G. & Woolfson, D. N. Engineering the morphology of a self-assembling protein fibre. *Nat. Mater.* **2,** 329–32 (2003).

36. Parry, D. a D., Fraser, R. D. B. & Squire, J. M. Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. *J. Struct. Biol.* **163,** 258–69 (2008).

37. Mason, J. M. & Arndt, K. M. Coiled coil domains: stability, specificity, and biological implications. *Chembiochem* **5,** 170–6 (2004).

38. Bromley, E. H. C. & Channon, K. J. *Alpha-helical peptide assemblies giving new function to designed structures. Prog. Mol. Biol. Transl. Sci.* **103,** (Elsevier Inc., 2011).

39. Pandya, M. J. *et al.* Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry* **39,** 8728–34 (2000).

40. Papapostolou, D. *et al.* Engineering nanoscale order into a designed protein fiber. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 10853–8 (2007).

41. Papapostolou, D., Bromley, E. H. C., Bano, C. & Woolfson, D. N. Electrostatic control of thickness and stiffness in a designed protein fiber. *J. Am. Chem. Soc.* **130,** 5124–30 (2008).

42. Ryadnov, M. G., Bella, A., Timson, S. & Woolfson, D. N. Modular design of peptide fibrillar nano- to microstructures. *J. Am. Chem. Soc.* **131,** 13240–1 (2009).

43. 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* **60,** 7237–7246 (2004).

44. Kojima, S., Kuriki, Y., Yoshida, T., Yazaki, K. & Miura, K. Fibril Formation by an Amphipathic alpha-Helix Forming Polypeptide Produced by Gene Engineering. *Proc. Japan Acad. Ser. B* **73,** 7–11 (1997).

45. Fallas, J. A., O'Leary, L. E. R. & Hartgerink, J. D. Synthetic collagen mimics: self-assembly of homotrimers, heterotrimers and higher order structures. *Chem. Soc. Rev.* **39,** 3510–27 (2010).

46. Li, Y., Mo, X., Kim, D. & Yu, S. M. Template-tethered collagen mimetic peptides for studying heterotrimeric triple-helical interactions. *Biopolymers* **95,** 94–104 (2011).

47. Potekhin, S. A. a *et al.* De novo design of fibrils made of short alpha-helical coiled coil peptides. *Chem. Biol.* **8,** 1025–32 (2001).

48. Gribbon, C. *et al.* MagicWand: A Single, Designed Peptide That Assembles to Stable, Ordered a-helical fibers. *Biochemistry* **47,** 10365–10371 (2008).

49. Ogihara, N. L. *et al.* Design of three-dimensional domain-swapped dimers and fibrous oligomers. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1404–9 (2001).

50. Sharp, T. H. *et al.* Cryo-transmission electron microscopy structure of a gigadalton peptide fiber of de novo design. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 13266–13271 (2012).

51. Bromley, E. H. C. *et al.* Assembly pathway of a designed alpha-helical protein fiber. *Biophys. J.* **98,** 1668–76 (2010).

52. Smith, A. M. *et al.* Polar assembly in a designed protein fiber. *Angew. Chemie* **44,** 325–8 (2004).

53. Wagner, D. E. *et al.* Toward the development of peptide nanofilaments and nanoropes as smart materials. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 12656–61 (2005).

54. Fletcher, J. M. *et al.* Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science (80-. ).* **340,** 595–599 (2013).

55. Eisenberg, D. & Liu, Y. 3D domain swapping: As domains continue to swap. *Protein Sci.* **11,** 1285–1299 (2002).

56. Bax, B. *et al.* X-ray analysis of bB2-crystallin and evolution of oligomeric lens proteins. *Nature* **347,** 776–780 (1990).

57. Åkansson, M. H., Svensson, A., Fast, J. & Linse, S. An extended hydrophobic core induces EF-hand swapping. *Protein Sci.* **10,** 927–933 (2001).

58. Munishkina, L. A., Ahmad, A., Fink, A. L. & Uversky, V. N. Guiding Protein Aggregation with Macromolecular Crowding. *Biochemistry* **47,** 8993–9006 (2008).

59. Matsumura, S., Uemura, S. & Mihara, H. Metal-triggered Nanofiber Formation of His-containing β-Sheet Peptide. *Supramol. Chem.* **18,** 397–403 (2006).

60. Wagner, S. C. On the interaction between proteins and nanoparticles : Coiled coil peptides organize nanoparticles and vice versa. (2010).

61. Minelli, C., Liew, J. X., Muthu, M. & Andresen, H. Coiled coil peptide-functionalized surfaces for reversible molecular binding. *Soft Matter* **9,** 5119–5124 (2013).

62. Tao, A. R. *et al.* The role of protein assembly in dynamically tunable bio-optical tissues. *Biomaterials* **31,** 793–801 (2010).

63. Pagel, K., Wagner, S. C., Samedov, K., Berlepsch, H. Von & Bo, C. Random Coils , b-Sheet Ribbons, and a-Helical Fibers: One Peptide Adopting Three Different Secondary Structures at Will. *J. Am. Chem. Soc.* **128,** 2196–2197 (2006).

64. Myszka, D. G. & Chaiken, I. M. Design and Characterization of an Intramolecular Antiparallel Coiled Coil Peptide. *Biochemistry* **33,** 2363–2372 (1994).

65. Rele, S. *et al.* D-periodic collagen-mimetic microfibers. *J. Am. Chem. Soc.* **129,** 14780–7 (2007).

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

67. Hartgerink, J. D., Beniash, E. & Stupp, S. I. Self-assembly and mineralization of peptide-amphiphile nanofibers. *Science (80-. ).* **294,** 1684–8 (2001).

68. Lau, S. Y. M., Taneja, A. K. & Hodges, S. Synthesis of a Model Protein of Defined Secondary and Quaternary Structure. *J. Biol. Chem.* **259,** 13252–13261 (1984).

69. Graddis, T. J., Myszka, D. G. & Chaiken, I. M. Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry* **32,** 12664–12671 (1993).

70. Suzuki, K., Yamada, T. & Tanaka, T. Role of the Buried Glutamate in the a-Helical Coiled Coil Domain of the Macrophage Scavenger Receptor. *Biochemistry* **38,** 1751–1756 (1999).

71. Zhou, N. E., Kay, C. M. & Hodges, R. S. The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Eng.* **7,** 1365–1372 (1994).

72. Gursky, O., Benjwal, S., Verma, S. & Röhm, K.-H. Monitoring protein aggregation during thermal unfolding in circular dichroism experiments. *Protein Sci.* **15,** 635–639 (2006).

73. Drzewiecki, K. *et al.* Methacrylation Induces Rapid, Temperature Dependent, Reversible Self-Assembly of Type-I Collagen. *Langmuir* **30,** 11204–11211 (2014).

74. Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T. & Hodges, R. S. Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position ‘d’. *J. Mol. Biol.* **300,** 377–402 (2000).

75. Wagschal, K., Tripet, B. & Hodges, R. S. De novo design of a model peptide sequence to examine the effects of single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils. *J. Mol. Biol.* **285,** 785–803 (1999).

76. Yallapu, M. M., Jaggi, M. & Chauhan, S. C. Curcumin nanoformulations: a future nanomedicine for cancer. *Drug Discov. Today* **17,** 71–80 (2012).

77. Rai, D., Singh, J. K., Roy, N. & Panda, D. Curcumin inhibits FtsZ assembly: an attractive mechanism for its antibacterial activity. *Biochem. J.* **410,** 147–55 (2008).

78. Zsila, F., Bikádi, Z. & Simonyi, M. Induced circular dichroism spectra reveal binding of the antiinflammatory curcumin to human alpha1-acid glycoprotein. *Bioorg. Med. Chem.* **12,** 3239–45 (2004).

79. Fathima, N. N., Dhathathreyan, A. & Ramasami, T. Directed 2-dimensional organisation of collagen: Role of cross-linking and denaturing agents. *J. Chem. Sci.* **122,** 881–889 (2010).

80. Nishad Fathima, N., Saranya Devi, R., Rekha, K. B. & Dhathathreyan, A. Collagen-curcumin interaction — A physico-chemical study. *J. Chem. Sci.* **121,** 509–514 (2009).

81. Hafner-Bratkovic, I., Gaspersic, J., Smid, L. M., Bresjanac, M. & Jerala, R. Curcumin binds to the alpha-helical intermediate and to the amyloid form of prion protein - a new mechanism for the inhibition of PrP(Sc) accumulation. *J. Neurochem.* **104,** 1553–64 (2008).

82. Przybyla, D. E. & Chmielewski, J. Metal-Triggered Radial Self-Assembly of Collagen Peptide Fibers. *J. Am. Chem. Soc.* **130,** 12610–12611 (2008).

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

84. Cerasoli, E., Sharpe, B. K. & Woolfson, D. N. ZiCo: A peptide designed to switch folded state upon binding zinc. *J. Am. Chem. Soc.* **127,** 15008–15009 (2005).

85. Ulrich, A. S., Otter, M., Glabe, C. G. & Hoekstra, D. Membrane fusion is induced by a distinct peptide sequence of the sea urchin fertilization protein bindin. *J. Biol. Chem.* **273,** 16748–16755 (1998).

86. Kiyokawa, T. *et al.* Binding of Cu(II) or Zn(II) in a de novo designed triple-stranded alpha-helical coiled-coil toward a prototype for a metalloenzyme. *J. Pept. Res.* **63,** 347–353 (2004).

87. Signarvic, R. S. & Degrado, W. F. Metal-binding dependent disruption of membranes by designed helices. *J. Am. Chem. Soc.* **131,** 3377–3384 (2009).

88. Gunasekar, S. K., Anija, L., Matsui, H., Montclare, J. K. & Anjia, L. Effects of Divalent Metals on Nanoscopic Fiber Formation and Small Molecule Recognition of Helical Proteins. *Adv. Funct. Mater.* **22,** 2154–2159 (2012).

89. Bent, P., Hospital, B., Kagi, J. H. R. & Vallee, B. L. Metallothionein: a Cadmium and Zinc-containing Protein from Equine Renal Cortex. *J. Biol. Chem.* **236,** (1961).

90. Nam, Y. S. *et al.* Virus-templated assembly of porphyrins into light-harvesting nanoantennae. *J. Am. Chem. Soc.* **132,** 1462–3 (2010).

91. Laganowsky, A. *et al.* Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci.* **19,** 1031–43 (2010).

92. Gunasekar, S. K. *et al.* N-terminal aliphatic residues dictate the structure, stability, assembly, and small molecule binding of the coiled-coil region of cartilage oligomeric matrix protein. *Biochemistry* **48,** 8559–8567 (2009).

93. McFarlane, A. *et al.* The pentameric channel of COMPcc in complex with different fatty acids. *PLoS One* **7,** e48130 (2012).

94. Özbek, S., Engel, J. & Stetefeld, J. Storage function of cartilage oligomeric matrix protein: the crystal structure of the coiled-coil domain in complex with vitamin D(3). *EMBO J.* **21,** 5960–8 (2002).

95. Guo, Y. *et al.* All-trans retinol, vitamin D and other hydrophobic compounds bind in the axial pore of the five-stranded coiled-coil domain of cartilage oligomeric matrix protein. *EMBO J.* **17,** 5265–72 (1998).

96. Dong, H., Paramonov, S. E. & Hartgerink, J. D. Self-assembly of alpha-helical coiled coil nanofibers. *J. Am. Chem. Soc.* **130,** 13691–5 (2008).

97. Chou, P. Y. & Fasman, G. D. Conformational Parameters for Amino Acids in Helical, b-Sheet, and Random Coild Regions Calculated from Proteins. *Biochemistry* **13,** 211–222 (1974).

98. Piela, L., Nemethy, G. & Harold, A. Proline-Induced Constraints in a-Helices. *Biopolymers* **26,** 1587–1600 (1987).

99. Altunbas, A., Lee, S. J., Rajasekaran, S. a, Schneider, J. P. & Pochan, D. J. Encapsulation of curcumin in self-assembling peptide hydrogels as injectable drug delivery vehicles. *Biomaterials* **32,** 5906–14 (2011).

100. Xu, C. *et al.* Rational Design of Helical Nanotubes from Self-Assembly of Coiled-Coil Lock Washers. *J. Am. Chem. Soc.* **135,** (2013).

101. Sellinger, A. *et al.* Continuous self-assembly of organic–inorganic nanocomposite coatings that mimic nacre. *Nature* **394,** (1998).

102. Reichert, J., Wei, G. & Jandt, K. D. Formation and Topotactical Orientation of Fibrinogen Nanofibrils on Graphite Nanostructures. *Adv. Eng. Mater.* **11,** B177–B181 (2009).

103. Wang, C., Stewart, R. J. & Kopec, J. Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* **397,** 417–420 (1999).

104. Dai, M. *et al.* Artificial protein block polymer libraries bearing two SADs: effects of elastin domain repeats. *Biomacromolecules* **12,** 4240–6 (2011).

105. McFarlane, A. A., Orriss, G. L. & Stetefeld, J. The use of coiled-coil proteins in drug delivery systems. *Eur. J. Pharmacol.* **625,** 101–7 (2009).

106. More, H. T., Frezzo, J. a, Dai, J., Yamano, S. & Montclare, J. K. Gene delivery from supercharged coiled-coil protein and cationic lipid hybrid complex. *Biomaterials* **35,** 7188–93 (2014).

107. Brahatheeswaran, D. *et al.* Hybrid fluorescent curcumin loaded zein electrospun nanofibrous scaffold for biomedical applications. *Biomed. Mater.* **7,** 045001 (2012).

108. Matson, J. B. & Stupp, S. I. Self-assembling peptide scaffolds for regenerative medicine. *Chem. Commun.* **48,** 26–33 (2012).

109. Berg, J. M., Tymoczko, J. L. & Stryer, L. *Stryer Biochemie*. *Biochem. Textb.* (2007).

110. Whitmore, L. & Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89,** 392–400 (2008).

111. Whitmore, L. & Wallace, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32,** W668–73 (2004).

112. Greenfield, N. J. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protoc.* **1,** 2527–35 (2006).

113. Pace, C. N. Measuring and increasing protein stability. *Trends Biotechnol.* **8,** 93–8 (1990).

114. Greenfield, N. & Fasman, G. D. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **8,** 4108–16 (1969).

115. Hu, X., Kaplan, D. & Cebe, P. Determining Beta-Sheet Crystallinity in Fibrous Proteins by Thermal Analysis and Infrared Spectroscopy. *Macromolecules* **39,** 6161–6170 (2006).

116. Abràmoff, M. D. & Magalhães, P. J. Image Processing with ImageJ. *Biophotonics Int.* **11,** 36–42 (2004).

117. Xiong, A.-S. *et al.* A simple, rapid, high-fidelity and cost-effective PCR-based two-step DNA synthesis method for long gene sequences. *Nucleic Acids Res.* **32,** e98 (2004).

118. Huang, X. & Miller, W. A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* **12,** 337–357 (1991).

119. Vladimir, N. *et al.* The Crystal structure of a five-stranded coiled coil in COMP: A prototype ion channel? *Science (80-. ).* **274,** 761–5 (1996).

120. Kelly, S. M. & Price, N. C. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* **1,** 349–84 (2000).

121. Lin, K., Yang, H., Gao, Z., Li, F. & Yu, S. Overestimated accuracy of circular dichroism in determining protein secondary structure. *Eur. Biophys. J.* **42,** 455–461 (2013).

122. Jackson, M. & Mantsch, H. H. The use and misuse of FTIR spectroscopy in the determination of protein structure. *Crit. Rev. Biochem. Mol. Biol.* **30,** 95–120 (1995).

123. Heimburg, T., Schünemann, J., Weber, K. & Geisler, N. FTIR-Spectroscopy of multistranded coiled coil proteins. *Biochemistry* **38,** 12727–34 (1999).

124. 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.* **245,** 180–194 (1995).

125. Sturtevant, J. M. Heat Capacity and Entropy Changes in Processes Involving Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 2236–2240 (1977).

126. Yang, A. S., Sharp, K. . & Honig, B. Analysis of the heat capacity dependence of protein folding. *J. Mol. Biol.* **227,** 889–900 (1992).

127. Liu, L., Yang, C. & Guo, Q. X. A study on the enthalpy-entropy compensation in protein unfolding. *Biophys. Chem.* **84,** 239–251 (2000).

128. Sharp, T. H. *Biomolecular Imaging at High Spatial and Temporal Resolution In Vitro and In Vivo*. (Springer, 2014).

129. Santos, N. C. & Castanho, M. a R. B. An overview of the biophysical applications of atomic force microscopy. *Biophys. Chem.* **107,** 133–49 (2004).

130. Hagn, F. *et al.* A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465,** 239–42 (2010).

131. Zhang, S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21,** 1171–1178 (2003).

132. Stebbins, J. L. *et al.* Nuclear magnetic resonance fragment-based identification of novel FKBP12 inhibitors. *J. Med. Chem.* **50,** 6607–17 (2007).

133. Anderson, V. L., Webb, W. W. & Eliezer, D. Interplay between desolvation and secondary structure in mediating cosolvent and temperature induced alpha-synuclein aggregation. *Phys. Biol.* **9,** 056005 (2012).

134. Gopinath, D. *et al.* Dermal wound healing processes with curcumin incorporated collagen films. *Biomaterials* **25,** 1911–1917 (2004).

135. Vogel, K. G. & Trotter, J. A. The Effect of Proteoglycans on the Morphology of Collagen Fibrils Formed In Vitro. *Coll. Relat. Res.* **7,** 105–114 (1987).

136. Guo, S. & Wang, E. Noble metal nanomaterials: Controllable synthesis and application in fuel cells and analytical sensors. *Nano Today* **6,** 240–264 (2011).

137. Eustis, S. & El-Sayed, M. A. Why gold nanoparticles are more precious than pretty gold: noble metal surface plasmon resonance and its enhancement of the radiative and nonradiative properties of nanocrystals of different shapes. *Chem. Soc. Rev.* **35,** 209–217 (2006).

138. Zhao, P., Li, N. & Astruc, D. State of the art in gold nanoparticle synthesis. *Coord. Chem. Rev.* **257,** 638–665 (2013).

139. Slocik, J. M., Kim, S. N., Whitehead, T. A., Clark, D. S. & Naik, R. R. Biotemplated metal nanowires using hyperthermophilic protein filaments. *Small* **5,** 2038–42 (2009).

140. Slocik, J. M., Moore, J. T. & Wright, D. W. Monoclonal Antibody Recognition of Histidine-Rich Peptide Encapsulated Nanoclusters. *Nano Lett.* **2,** 169–173 (2002).

141. Barnes, W. L., Dereux, A. & Ebbesen, T. W. Surface plasmon subwavelength optics. *Nature* **424,** 824–30 (2003).

142. Veintemillas-Verdaguer, S. *et al.* Colloidal dispersions of maghemite nanoparticles produced by laser pyrolysis with application as NMR contrast agents. *J. Phys. D. Appl. Phys.* **37,** 2054–2059 (2004).

143. Crookes-Goodson, W. J., Slocik, J. M. & Naik, R. R. Bio-directed synthesis and assembly of nanomaterials. *Chem. Soc. Rev.* **37,** 2403–12 (2008).

144. Wesson, J. A. & Ward, M. D. Medical Mineralogy and Geochemistry: Pathological Biomineralization of Kidney Stones. *Elements* **3,** 415–421 (2007).

145. Reith, F. *et al.* Mechanisms of gold biomineralization in the bacterium Cupriavidus metallidurans. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 17757–62 (2009).

146. Malvankar, N. S. *et al.* Tunable metallic-like conductivity in microbial nanowire networks. *Nat. Nanotechnol.* (2011). doi:10.1038/nnano.2011.119

147. Huo, Z., Tsung, C., Huang, W., Zhang, X. & Yang, P. Sub-Two Nanometer Single Crystal Au Nanowires. *Nano Lett.* **8,** 2041–2044 (2008).

148. Feng, H. *et al.* Simple and rapid synthesis of ultrathin gold nanowires, their self-assembly and application in surface-enhanced Raman scattering. *Chem. Commun.* 1984–1986 (2009). doi:10.1039/b822507a

149. Scheibel, T. *et al.* Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4527–4532 (2003).

150. Kundu, S., Gill, R. S. & Saraf, R. F. Electrospinning of PAH Nanofiber and Deposition of Au NPs for Nanodevice Fabrication. *J. Phys. Chem. C* **115,** 15845–15852 (2011).

151. Saha, K., Agasti, S. S., Kim, C., Li, X. & Rotello, V. M. Gold nanoparticles in chemical and biological sensing. *Chem. Rev.* **112,** 2739–79 (2012).

152. Lu, L.-P., Wang, S.-Q. & Lin, X.-Q. Fabrication of layer-by-layer deposited multilayer films containing DNA and gold nanoparticle for norepinephrine biosensor. *Anal. Chim. Acta* **519,** 161–166 (2004).

153. Jiang, C., Markutsya, S., Pikus, Y. & Tsukruk, V. V. Freely suspended nanocomposite membranes as highly sensitive sensors. *Nat. Mater.* **3,** 721–8 (2004).

154. Daniel, M. & Astruc, D. Gold Nanoparticles: Assembly, Supramolecular Chemistry, Quantum-Size-Related Properties, and Applications toward Biology, Catalysis, and Nanotechnology. *Chem. Rev.* **104,** 293–346 (2004).

155. Schnur, J. M., Price, R., Schoen, P. & Yager, P. Lipid-based tubule microstructures. *Thin Solid Films* **152,** 181–206 (1987).

156. Braun, E., Eichen, Y., Sivan, U. & Ben-Yoseph, G. DNA-templated assembly and electrode attachment of a conducting silver wire. *Nature* **391,** 775–8 (1998).

157. Channon, K. J., Devlin, G. L. & MacPhee, C. E. Efficient energy transfer within self-assembling peptide fibers: a route to light-harvesting nanomaterials. *J. Am. Chem. Soc.* **131,** 12520–1 (2009).

158. Chen, C. L. & Rosi, N. L. Preparation of unique 1-D nanoparticle superstructures and tailoring their structural features. *J. Am. Chem. Soc.* **132,** 6902–3 (2010).

159. Higashi, N., Kawahara, J. & Niwa, M. Preparation of helical peptide monolayer-coated gold nanoparticles. *J. Colloid Interface Sci.* **288,** 83–87 (2005).

160. Djalali, R., Chen, Y. & Matsui, H. Au nanocrystal growth on nanotubes controlled by conformations and charges of sequenced peptide templates. *J. Am. Chem. Soc.* **125,** 5873–9 (2003).

161. Slocik, J. M., Stone, M. O. & Naik, R. R. Synthesis of gold nanoparticles using multifunctional peptides. *Small* **1,** 1048–52 (2005).

162. Slocik, J. M., Govorov, A. O. & Naik, R. R. Plasmonic Circular Dichroism of Peptide-Functionalized Gold Nanoparticles. *Nano Lett.* **11,** 701–705 (2011).

163. Mark, S. S. *et al.* Bionanofabrication of Metallic and Semiconductor Nanoparticle Arrays Using S-Layer Protein Lattices with Different Lateral Spacings and Geometries. *Langmuir* 3763–3774 (2006).

164. Deka, J., Paul, A. & Chattopadhyay, A. Estimating conformation content of a protein using citrate-stabilized Au nanoparticles. *Nanoscale* **2,** 1405–1412 (2010).

165. Kharlampieva, E. *et al.* Redox-Active Ultrathin Template of Silk Fibroin: Effect of Secondary Structure on Gold Nanoparticle Reduction. *Chem. Mater.* **21,** 2696–2704 (2009).

166. Chen, C.-L., Zhang, P. & Rosi, N. L. A new peptide-based method for the design and synthesis of nanoparticle superstructures: construction of highly ordered gold nanoparticle double helices. *J. Am. Chem. Soc.* **130,** 13555–13557 (2008).

167. Nishikawa, H., Morita, T., Sugiyama, J. & Kimura, S. Formation of gold nanoparticles in microreactor composed of helical peptide assembly in water. *J. Colloid Interface Sci.* **280,** 506–510 (2004).

168. Jung, S. H., Jeon, J., Kim, H., Jaworski, J. & Jung, J. H. Chiral Arrangement of Achiral Au Nanoparticles by Supramolecular Assembly of Helical Nanofiber Templates. *J. Am. Chem. Soc.* **136,** 6446–6452 (2014).

169. Higuchi, M., Ushiba, K. & Kawaguchi, M. Structural control of peptide-coated gold nanoparticle assemblies by the conformational transition of surface peptides. *J. Colloid Interface Sci.* **308,** 356–363 (2007).

170. Bayraktar, H., Srivastava, S., You, C.-C., Rotello, V. M. & Knapp, M. J. Controlled nanoparticle assembly through protein conformational changes. *Soft Matter* **4,** 751–756 (2008).

171. Stevens, M. M., Flynn, N. T., Wang, C., Tirrell, D. A. & Langer, R. Coiled-Coil Peptide-Based Assembly of Gold Nanoparticles. *Adv. Mater.* **16,** 915–918 (2004).

172. Behrens, S. S. Synthesis of inorganic nanomaterials mediated by protein assemblies. *J. Mater. Chem.* **18,** 3788 (2008).

173. Tullman, J. A., Finney, W. F., Lin, Y.-J. & Bishnoi, S. W. Tunable Assembly of Peptide-coated Gold Nanoparticles. *Plasmonics* **2,** 119–127 (2007).

174. Shimmin, R. G., Schoch, A. B. & Braun, P. V. Polymer Size and Concentration Effects on the Size of Gold Nanoparticles Capped by Polymeric Thiols. *Langmuir* **20,** 5613–5620 (2004).

175. Li, N. *et al.* State of the art in gold nanoparticle synthesis. *Coord. Chem. Rev.* **257,** 638–665 (2013).

176. Heddle, J. Gold Nanoparticle-Biological Molecule Interactions and Catalysis. *Catalysts* **3,** 683–708 (2013).

177. Pramanik, S., Banerjee, P., Sarkar, A. & Bhattacharya, S. C. Size-dependent interaction of gold nanoparticles with transport protein: A spectroscopic study. *J. Lumin.* **128,** 1969–1974 (2008).

178. Wei, H. *et al.* Time-dependent, protein-directed growth of gold nanoparticles within a single crystal of lysozyme. *Nat. Nanotechnol.* **6,** 93–97 (2011).

179. Daniel, M. C. *et al.* Role of surface charge density in nanoparticle-templated assembly of bromovirus protein cages. *ACS Nano* **4,** 3853–3860 (2010).

180. Rai, M. & Posten, C. *Green Biosynthesis of Nanoparticles: Mechanisms and Applications*. (CABI, 2013).

181. Kaur, P. *et al.* Three-dimensional directed self-assembly of Peptide nanowires into micrometer-sized crystalline cubes with nanoparticle joints. *Angew. Chem. Int. Ed. Engl.* **49,** 8375–8 (2010).

182. Guo, Y., Ma, Y., Xu, L., Li, J. & Yang, W. Conformational Change Induced Reversible Assembly/Disassembly of Poly-L-Lysine-Functionalized Gold Nanoparticles. *J. Phys. Chem. C* **111,** 9172–9176 (2007).

183. Slocik, J. M., Tam, F., Halas, N. J. & Naik, R. R. Peptide-assembled optically responsive nanoparticle complexes. *Nano Lett.* **7,** 1054–1058 (2007).

184. Link, S., Burda, C., Nikoobakht, B. & El-Sayed, M. A. Laser-Induced Shape Changes of Colloidal Gold Nanorods Using Femtosecond and Nanosecond Laser Pulses. *J. Phys. Chem. B* **104,** 6152–6163 (2000).

185. Aili, D. *et al.* Folding induced assembly of polypeptide decorated gold nanoparticles. *J. Am. Chem. Soc.* **130,** 5780–5788 (2008).

186. Kogot, J. M., England, H. J., Strouse, G. F. & Logan, T. M. Single peptide assembly onto a 1.5 nm Au surface via a histidine tag. *J. Am. Chem. Soc.* **130,** 16156–7 (2008).

187. Feyer, V. *et al.* Adsorption of histidine and histidine-containing peptides on Au(111). *Langmuir* **26,** 8606–13 (2010).

188. Chen, M. S. & Goodman, D. W. The structure of catalytically active gold on titania. *Science* **306,** 252–255 (2004).

189. Huang, X. & El-Sayed, M. A. Gold nanoparticles: Optical properties and implementations in cancer diagnosis and photothermal therapy. *J. Adv. Res.* **1,** 13–28 (2010).

190. Krasteva, N. *et al.* Self-Assembled Gold Nanoparticle/Dendrimer Composite Films for Vapor Sensing Applications. *Nano Lett.* **2,** 551–555 (2002).

191. Rakhi, R. B., Sethupathi, K. & Ramaprabhu, S. A glucose biosensor based on deposition of glucose oxidase onto crystalline gold nanoparticle modified carbon nanotube electrode. *J. Phys. Chem. B* **113,** 3190–4 (2009).

192. Feng, J.-J., Zhao, G., Xu, J.-J. & Chen, H.-Y. Direct electrochemistry and electrocatalysis of heme proteins immobilized on gold nanoparticles stabilized by chitosan. *Anal. Biochem.* **342,** 280–6 (2005).

193. Zheng, L., Brody, J. P. & Burke, P. J. Electronic manipulation of DNA, proteins, and nanoparticles for potential circuit assembly. *Biosens. Bioelectron.* **20,** 606–19 (2004).

194. Polavarapu, L. & Xu, Q.-H. A single-step synthesis of gold nanochains using an amino acid as a capping agent and characterization of their optical properties. *Nanotechnology* **19,** 075601 (2008).

195. Guo, B., Anzai, J. & Osa, T. Adsorption Behavior of Serum Albumin on Electrode Surfaces and the Effects of Electrode Potential. *Chem. Pharm. Bull.* **44,** 800–803 (1996).

196. Ding, X., Yang, M., Hu, J., Li, Q. & McDougall, A. Study of the adsorption of cytochrome c on a gold nanoparticle – modified gold electrode by using cyclic voltammetry, electrochemical impedance spectroscopy and chronopotentiometry. *Microchim. Acta* **158,** 65–71 (2006).

197. Zhu, X., Yang, Q., Huang, J., Suzuki, I. & Li, G. Colorimetric Study of the Interaction Between Gold Nanoparticles and a Series of Amino Acids. *J. Nanosci. Nanotechnol.* **8,** 353–357 (2008).

198. Ojea-Jiménez, I. & Puntes, V. Instability of cationic gold nanoparticle bioconjugates: the role of citrate ions. *J. Am. Chem. Soc.* **131,** 13320–13327 (2009).

199. Shipway, A. N., Katz, E. & Willner, I. Nanoparticle Arrays on Surfaces for Electronic, Optical, and Sensor Applications. *ChemPhysChem* **1,** 18–52 (2000).

200. Huang, H. & Yang, X. Chitosan mediated assembly of gold nanoparticles multilayer. *Colloids Surfaces A Physicochem. Eng. Asp.* **226,** 77–86 (2003).

201. Zhou, J. C. *et al.* Characterization of gold nanoparticle binding to microtubule filaments. *Mater. Sci. Eng. C* **30,** 20–26 (2010).

202. Huang, H., Liu, Z. & Yang, X. Application of electrochemical impedance spectroscopy for monitoring allergen-antibody reactions using gold nanoparticle-based biomolecular immobilization method. *Anal. Biochem.* **356,** 208–14 (2006).

203. Zhang, L., Jiang, X., Wang, E. & Dong, S. Attachment of gold nanoparticles to glassy carbon electrode and its application for the direct electrochemistry and electrocatalytic behavior of hemoglobin. *Biosens. Bioelectron.* **21,** 337–45 (2005).

204. Sone, E. D. & Stupp, S. I. Bioinspired Magnetite Mineralization of Peptide-Amphiphile Nanofibers. *Chem. Mater.* **23,** 2005–2007 (2011).

205. Mallapragada, S. *et al.* Bacterial Protein That Promotes the Formation of Superparamagnetic Magnetite Nanoparticles of Uniform Size and Shape. *Biomacromolecules* **13,** 98–105 (2012).

206. Teja, A. S. & Koh, P.-Y. Synthesis, properties, and applications of magnetic iron oxide nanoparticles. *Prog. Cryst. Growth Charact. Mater.* **55,** 22–45 (2009).

207. Burda, C., Chen, X., Narayanan, R. & El-Sayed, M. a. Chemistry and properties of nanocrystals of different shapes. *Chem. Rev.* **105,** 1025–1102 (2005).

208. Teja, A. S. & Koh, P.-Y. Synthesis, properties, and applications of magnetic iron oxide nanoparticles. *Prog. Cryst. Growth Charact. Mater.* **55,** 22–45 (2009).

209. Faraji, M., Yamini, Y. & Rezaee, M. Magnetic Nanoparticles: Synthesis, Stabilization, Functionalization, Characterization, and Applications. *J. Iran. Chem. Soc.* **7,** 1–37 (2010).

210. Lu, A.-H., Salabas, E. L. & Schüth, F. Magnetic nanoparticles: synthesis, protection, functionalization, and application. *Angew. Chemie* **46,** 1222–44 (2007).

211. Kumar, C. S. S. R. & Mohammad, F. Magnetic nanomaterials for hyperthermia-based therapy and controlled drug delivery. *Adv. Drug Deliv. Rev.* **63,** 789–808 (2011).

212. Kirschvink, J. L., Kobayashi-Kirschvink, A. & Woodford, B. J. Magnetite biomineralization in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 7683–7 (1992).

213. Faivre, D. & Schüler, D. Magnetotactic bacteria and magnetosomes. *Chem. Rev.* **108,** 4875–98 (2008).

214. Sun, X. *et al.* Cyclodextrin-Assisted Synthesis of Superparamagnetic Magnetite Nanoparticles from a Single Fe (III) Precursor. *J. Phys. Chem. C* 17148–17155 (2008).

215. Nussbaum, A. *Electronic and magnetic behavior of materials.* (Prentice-Hall, Inc., 1967).

216. Pascal, C., Pascal, J. L. & Favier, F. Electrochemical Synthesis for the Control of γ-Fe 2 O 3 Nanoparticle Size. Morphology, Microstructure, and Magnetic Behavior. *Small* 141–147 (1999).

217. Lee, Y. *et al.* Large-Scale Synthesis of Uniform and Crystalline Magnetite Nanoparticles Using Reverse Micelles as Nanoreactors under Reflux Conditions. *Adv. Funct. Mater.* **15,** 503–509 (2005).

218. Amemiya, Y., Arakaki, A., Staniland, S. S., Tanaka, T. & Matsunaga, T. Controlled formation of magnetite crystal by partial oxidation of ferrous hydroxide in the presence of recombinant magnetotactic bacterial protein Mms6. *Biomaterials* **28,** 5381–9 (2007).

219. Pankhurst, Q. A., Connolly, J., Jones, S. K. & Dobson, J. Applications of magnetic nanoparticles in biomedicine. *J. Phys. D. Appl. Phys.* **36,** 167–181 (2003).

220. Klem, M. T., Young, M. & Douglas, T. Biomimetic magnetic nanoparticles. *Mater. Today* 28–37 (2005).

221. Dickerson, M. B., Sandhage, K. H. & Naik, R. R. Protein- and peptide-directed syntheses of inorganic materials. *Chem. Rev.* **108,** 4935–78 (2008).

222. Kim, S. U. *et al.* Nanoscale film formation of ferritin and its application to biomemory device. *Ultramicroscopy* **109,** 974–979 (2009).

223. Tartaj, P., Morales, M. P., González-Carreño, T., Veintemillas-Verdaguer, S. & Serna, C. J. Advances in magnetic nanoparticles for biotechnology applications. *J. Magn. Magn. Mater.* **290-291,** 28–34 (2005).

224. Laurent, S. *et al.* Magnetic iron oxide nanoparticles: synthesis, stabilization, vectorization, physicochemical characterizations, and biological applications. *Chem. Rev.* **108,** 2064–110 (2008).

225. Peng, T., Paramelle, D., Sana, B., Lee, C. F. & Lim, S. Designing Non-Native Iron-Binding Site on a Protein Cage for Biological Synthesis of Nanoparticles. *Small* 1–8 (2014). doi:10.1002/smll.201303516

226. Tatur, J., Hagedoorn, P. L., Overeijnder, M. L. & Hagen, W. R. A highly thermostable ferritin from the hyperthermophilic archaeal anaerobe Pyrococcus furiosus. *Extremophiles* **10,** 139–148 (2006).

227. Marshall, T. R., Parmenter, C. S. & Blakemore, R. Magnetotactic Bacteria. *Science (80-. ).* **190,** 1–3 (1975).

228. Reiss, B. D. *et al.* Biological Routes to Metal Alloy Ferromagnetic Nanostructures. *Nano Lett.* **4,** 1127–1132 (2004).

229. Arakaki, A., Webb, J. & Matsunaga, T. A novel protein tightly bound to bacterial magnetic particles in Magnetospirillum magneticum strain AMB-1. *J. Biol. Chem.* **278,** 8745–8750 (2003).

230. in (Baeuerlain, E.) (Wiley-VCH).

231. Arakaki, A., Masuda, F., Amemiya, Y., Tanaka, T. & Matsunaga, T. Control of the morphology and size of magnetite particles with peptides mimicking the Mms6 protein from magnetotactic bacteria. *J. Colloid Interface Sci.* **343,** 65–70 (2010).

232. Galloway, J. M. *et al.* Biotemplated Magnetic Nanoparticle Arrays. *Small* 204–208 (2011). doi:10.1002/smll.201101627

233. Prozorov, T. *et al.* Protein-Mediated Synthesis of Uniform Superparamagnetic Magnetite Nanocrystals. *Adv. Funct. Mater.* **17,** 951–957 (2007).

234. Niemeyer, C. M. Nanoparticles, Proteins, and Nucleic Acids: Biotechnology Meets Materials Science. *Angew. Chem. Int. Ed.* **40,** 4128 ± 4158 (2001).

235. Wu, W., He, Q. & Jiang, C. Magnetic iron oxide nanoparticles: Synthesis and surface functionalization strategies. *Nanoscale Res. Lett.* **3,** 397–415 (2008).

236. Ryadnov, M. G. & Woolfson, D. N. Fiber recruiting peptides: Noncovalent decoration of an engineered protein scaffold. *J. Am. Chem. Soc.* **126,** 7454–7455 (2004).

237. Gazit, E. Use of biomolecular templates for the fabrication of metal nanowires. *FEBS J.* **274,** 317–322 (2007).

238. Scheibel, T. Protein fibers as performance proteins: New technologies and applications. *Curr. Opin. Biotechnol.* **16,** 427–433 (2005).

239. Niemeyer, C. M. Nanoparticles, Proteins, and Nucleic Acids: Biotechnology Meets Materials Science. *Angew. Chemie* **40,** 4128–4158 (2001).

240. Prozorov, T. *et al.* Cobalt ferrite nanocrystals: Out-performing magnetotactic bacteria. *ACS Nano* **1,** 228–233 (2007).

241. Banerjee, I. a, Yu, L. & Matsui, H. Cu nanocrystal growth on peptide nanotubes by biomineralization: size control of Cu nanocrystals by tuning peptide conformation. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 14678–14682 (2003).

242. Chen, C.-L. & Rosi, N. L. Peptide-based methods for the preparation of nanostructured inorganic materials. *Angew. Chemie* **49,** 1924–42 (2010).

243. Chen, J., Xu, L., Li, W. & Gou, X. a-Fe2O3 nanotubes in gas sensor and lithium-ion battery applications. *Adv. Mater.* **17,** 582–586 (2005).

244. Burda, C., Chen, X., Narayanan, R. & El-Sayed, M. a. Chemistry and properties of nanocrystals of different shapes. *Chem. Rev.* **105,** 1025–1102 (2005).

245. Wang, L. *et al.* Nanoparticle-based environmental sensors. *Mater. Sci. Eng. R Reports* **70,** 265–274 (2010).

246. Kaushik, A. *et al.* Chemical Iron oxide-chitosan nanobiocomposite for urea sensor. *Sensors Actuators B Chem.* **138,** 572–580 (2009).

247. Buzug, T. M. & Borgert, J. *Magnetic Particle Imaging*. (2012).

248. Zhang, Y., Kohler, N. & Zhang, M. Surface modification of superparamagnetic magnetite nanoparticles and their intracellular uptake. *Biomaterials* **23,** 1553–61 (2002).

249. Lee, K. S. & Lee, I. S. Decoration of superparamagnetic iron oxide nanoparticles with Ni2+: agent to bind and separate histidine-tagged proteins. *Chem. Commun.* 709 (2008). doi:10.1039/b715796g

250. Hung, C.-W., Holoman, T. R. P., Kofinas, P. & Bentley, W. E. Towards oriented assembly of proteins onto magnetic nanoparticles. *Biochem. Eng. J.* **38,** 164–170 (2008).

251. Yallapu, M. M. *et al.* Multi-functional magnetic nanoparticles for magnetic resonance imaging and cancer therapy. *Biomaterials* **32,** 1890–905 (2011).

252. Ito, A., Shinkai, M., Honda, H. & Kobayashi, T. Medical application of functionalized magnetic nanoparticles. *J. Biosci. Bioeng.* **100,** 1–11 (2005).

253. Gupta, A. K. & Gupta, M. Synthesis and surface engineering of iron oxide nanoparticles for biomedical applications. *Biomaterials* **26,** 3995–4021 (2005).

254. Link, A. J., Vink, M. K. S. & Tirrell, D. A. Preparation of the functionalizable methionine surrogate azidohomoalanine via copper-catalyzed diazo transfer. *Nat. Protoc.* **2,** 1879–83 (2007).

255. Abbenante, G., Le, G. T. & Fairlie, D. P. Unexpected photolytic decomposition of alkyl azides under mild conditions. *Chem. Commun. (Camb).* **1,** 4501–3 (2007).

256. Strable, E. *et al.* Unnatural amino acid incorporation into virus-like particles. *Bioconjug. Chem.* **19,** 866–75 (2008).

257. Voloshchuk, N. & Montclare, J. K. Incorporation of unnatural amino acids for synthetic biology. *Mol. Biosyst.* **6,** 65–80 (2010).

258. Kiick, K. L., Saxon, E., Tirrell, D. a & Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 19–24 (2002).

259. Link, a J. *et al.* Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 10180–5 (2006).

260. Link, A. J. *et al.* Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 10180–5 (2006).

261. Hong, V., Udit, A. K., Evans, R. a & Finn, M. G. Electrochemically protected copper(I)-catalyzed azide-alkyne cycloaddition. *Chembiochem* **9,** 1481–6 (2008).

262. Tanrikulu, I. C., Schmitt, E., Mechulam, Y., Goddard, W. a & Tirrell, D. a. Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 15285–90 (2009).

263. Mittl, P. R. *et al.* The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 2562–2566 (2000).

264. Beatty, K. E. & Tirrell, D. a. Two-color labeling of temporally defined protein populations in mammalian cells. *Bioorg. Med. Chem. Lett.* **18,** 5995–9 (2008).

265. Somasekharan, S. P. *et al.* Identification and quantification of newly synthesized proteins translationally regulated by YB-1 using a novel Click-SILAC approach. *J. Proteomics* **77,** e1–10 (2012).

266. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A. & Bertozzi, C. R. A Comparative Study of Bioorthogonal Reactions with Azides. *ACS Chem. Biol.* **1,** 644–648 (2006).

267. Link, a J. & Tirrell, D. a. Cell surface labeling of Escherichia coli via copper(I)-catalyzed [3+2] cycloaddition. *J. Am. Chem. Soc.* **125,** 11164–5 (2003).

268. Lutz, J.-F. & Zarafshani, Z. Efficient construction of therapeutics, bioconjugates, biomaterials and bioactive surfaces using azide-alkyne 'click' chemistry. *Adv. Drug Deliv. Rev.* **60,** 958–70 (2008).

269. Ester, M., Kayser, B., Altman, J. & Beck, W. Alkyne Bridged et- Amino Acids by Palladium Mediated Coupling of Alkynes with N-t-Boc-4-iodo- phenylalanine Methyl Ester. *Tetrahedron* **53,** 2475–2484 (1997).

270. Le Droumaguet, B. & Velonia, K. Click Chemistry: A Powerful Tool to Create Polymer‐Based Macromolecular Chimeras. *Macromol. Rapid Commun.* **29,** 1073–1089 (2008).

271. Hein, J. E. & Fokin, V. V. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem. Soc. Rev.* **39,** 1302–1315 (2010).

272. Lewis, W. G., Magallon, F. G., Fokin, V. V & Finn, M. G. Discovery and Characterization of Catalysts for Azide - Alkyne Cycloaddition by Fluorescence Quenching. *J. Am. Chem. Soc.* **126,** 9152–9153 (2004).

273. Hong, V., Presolski, S. I., Ma, C. & Finn, M. G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed. Engl.* **48,** 9879–83 (2009).

274. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chemie - Int. Ed.* **41,** 2596–2599 (2002).

275. Feltin, N. & Pileni, M. P. New Technique for Synthesizing Iron Ferrite Magnetic Nanosized Particles. *Langmuir* **13,** 3927–3933 (1997).

276. *ImageQuant TL, Image analysis software, data file 28-9379-13 AA*.

277. *Chromeo 494 Alkyne product data sheet, Catalog number 15292*.

278. Laurenti, M. *et al.* Synthesis of Thermosensitive Microgels with a Tunable Magnetic Core. *Langmuir* (2011). doi:10.1021/la201723a

279. Cabrera, L., Gutierrez, S., Menendez, N., Morales, M. P. & Herrasti, P. Magnetite nanoparticles: Electrochemical synthesis and characterization. *Electrochim. Acta* **53,** 3436–3441 (2008).

280. Sun, Y. K., Ma, M., Zhang, Y. & Gu, N. Synthesis of nanometer-size maghemite particles from magnetite. *Colloids Surfaces A Physicochem. Eng. Asp.* **245,** 15–19 (2004).

281. Dodrill, B. C., Lindemuth, J. R. & Krause, J. K. *Magnetic Anisotropy : Measurements with a Vector Vibrating*.

282. Park, S.-J. *et al.* Synthesis and Magnetic Studies of Uniform Iron Nanorods and Nanospheres. *J. Am. Chem. Soc.* **122,** 8581–8582 (2000).

283. Andersen, N. H., Liu, Z. & Prickett, K. S. Efforts toward deriving the CD spectrum of a 3(10) helix in aqueous medium. *FEBS Lett.* **399,** 47–52 (1996).

284. Massart, R. & Cabuil, V. J. Massart, R.; Cabuil, V. J. Chim. Phys. 1987, 84,7. *J. Chim. Phys.* **84,** (1987).

# 5 Appendices

## 5.1 Publications

### 5.1.1 *Hume et al. Biomacromolecules 2014*

BioMACROMOLECULES

Article

pubs.acs.org/Biomac

## Engineered Coiled-Coil Protein Microfibers

Jasmin Hume,[†] Jennifer Sun,[†] Rudy Jacquet,[†] P. Douglas Renfrew,[‡] Jesse A. Martin,[§] Richard Bonneau,[‡] M. Lane Gilchrist,[§] and Jin Kim Montclare*[,†]

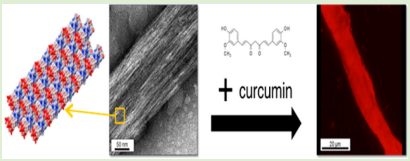
[†]Department of Chemical and Biomolecular Engineering, NYU Polytechnic School of Engineering, Brooklyn, New York 11201, United States

[‡]Center for Genomics and Systems Biology, Department of Biology, New York University New York, New York 10003, United States

[§]Departments of Chemical Engineering and Biomedical Engineering, The City College of the City University of New York, New York, New York 10031, United States

Supporting Information

**ABSTRACT:** The fabrication of de novo proteins able to self-assemble on the nano- to meso-length scales is critical in the development of protein-based biomaterials in nanotechnology and medicine. Here we report the design and characterization of a protein engineered coiled-coil that not only assembles into microfibers, but also can bind hydrophobic small molecules. Under ambient conditions, the protein forms fibers with nanoscale structure possessing large aspect ratios formed by bundles of $\alpha$-helical homopentameric assemblies, which further assemble into mesoscale fibers in the presence of curcumin through aggregation. Surprisingly, these biosynthesized fibers are able to form in conditions of remarkably low concentrations. Unlike previously designed coiled-coil fibers, these engineered protein microfibers can bind the small molecule curcumin throughout the assembly, serving as a depot for encapsulation and delivery of other chemical agents within protein-based 3D microenvironments.




## INTRODUCTION

Precise molecular organization across the nano- and meso-length scales is crucial for producing materials with defined physicochemical properties.[1−3] Nature provides elegant examples of proteins that self-assemble to form fibers both on the nanometer scale, including $\beta$-amyloids like those responsible for Alzheimer's disease,[4,5] as well as optically active self-assembling reflectins,[6,7] and on the mesoscale like the bundled $\alpha$-helical coiled-coil elastic protein of the Giant Clam, *Tridacna maxima*.[8] While nature has created many elaborate proteins capable of complex self-assembly and ligand binding, fabricating materials with the same level of structural and molecular specificity on various length scales remains a challenge. A range of self-assembling coiled-coil peptides have been designed to form nanofibers. Some of the earliest work on de novo $\alpha$-helical protein fibers has been done by Kojima et al. with the homotetrameric $\alpha_3$-peptide, which assembled to form fibers 5−10 nm in diameter with visible protofibrils.[9] The Woolfson group has produced self-assembling fibers, or SAFs,[10−12] that form nanofibers from tens to hundreds of nanometers in diameter and micrometers in length.[13,14] Fibers with visible protofibrils have also been constructed from DSAg proteins through swapping portions of $\alpha$-helical subunits in the 3-helix bundle, using design principles of 3D domain swapping to create fibers 30−70 nm across with protofibrils of a width corresponding to the trimeric coiled-coil.[15] Protofibrils corresponding to dimeric coiled-coils 2 nm in diameter that bundle to form fibers 40−80 nm in thickness have also been seen with modular $\alpha$-helical protein units.[16] While such $\alpha$-helical peptide fibers assemble through electrostatic and hydrophobic interactions between complementary subunits, the concentrations necessary to achieve these structures are typically greater than 50 $\mu$M,[10,11,17,18] with only a few examples of stable structures formed at lower concentrations.[11,19] Moreover, the dimensions of previously reported $\alpha$-helical peptide fibers range from 10 to 100 nm in diameter and micrometers in length.[11,13,18] To date, peptide or recombinant protein fibers with diameters on the micrometer scale have never been achieved.

We sought to design proteins capable of assembly in nanometer to micrometer scales akin to those found in nature: well-known examples including collagen,[3,20−22] $\alpha$-helical intermediates of amyloid fibers,[5,23] fibrinogen,[24] $\alpha$-keratin fibers,[25] actin,[26] and the less well-known assembly of the sterile $\alpha$-motif.[27] Rather than coiled-coil dimers or trimers,[15,18,28−30] we chose to focus on pentamer assemblies. Self-assembling $\alpha$-helical fibers formed from axial staggering of pentameric subunits have been described by Potekhin et al., producing fibers of controlled diameters.[31] Our system is based on the coiled-coil domain of cartilage oligomeric matrix protein

Received: April 3, 2014
Revised: June 12, 2014
Published: June 18, 2014

 © 2014 American Chemical Society 

(COMPcc).[32,33] Previously, our group engineered COMPcc mutating two cysteine residues (C68 and C71 in the original sequence), responsible for pentamer-stabilizing disulfide bonds, to serine to investigate the protein without the complication of oxidation.[33] The COMPcc mutant with serine in positions 68 and 71 is referred to as wt. Recently, we discovered that wt assembles into fibers with diameters of 10−15 nm;[34] in addition to the nanofibers, nanoparticles were formed, leading to a mixture. Additionally, mutation of several residues in the *a* and *d* positions along the pore of the wt protein to alanine were shown to be responsible for a loss in binding ability to small hydrophobic molecules, including curcumin.[35] This was used as inspiration to study the effect of binding of curcumin to engineered proteins with swapped domains, leading to unexpected and influential results. In this paper we describe the self-assembly of a novel protein designed by swapping regions of wt, dubbed Q, and compare it to wt as well as a negative control swap protein, called L. Q assembled into robust nanofibers with unprecedented diameters up to 560 nm under pH 4. In the presence of the small molecule curcumin, the Q fibers further assembled into microfibers with diameter of 16 $\mu$m, akin to natural keratin[25,35] and spider silk[36] fibers measuring tens of micrometers in diameter, providing the first example of an engineered protein microfiber.

## EXPERIMENTAL SECTION

**Materials.** Sodium phosphate (monobasic and dibasic) and nickel-nitrilotriacetic acid resins were purchased from Sigma-Aldrich. Ampicillin, isopropyl-$\beta$-D-thiogalactopyranoside (IPTG), tryptone, urea, tris-HCl, and sodium chloride were obtained from Fisher Scientific. Yeast extract, methanol, and curcumin were purchased from Acros Organics and BCA kit was obtained from Pierce. Imidazole was purchased from Alfa Aesar and copper grids for TEM were purchased from Ted Pella.

**Methods.** *Genetic Engineering.* Gene sequences for Q and L were generated via polymerase chain reaction (PCR) amplification and PCR assembly of DNA fragments of the wt gene (described in detail in Supporting Information).

*Protein Expression.* Approximately 1 mL of starter culture was added to 1 L of Luria Broth (LB) containing 0.2 mg/mL ampicillin and incubated at 37 °C, 250 rpm. After 9 h, the cultures were induced with 0.2 mg/mL IPTG and incubated for 3 h at 37 °C, 250 rpm. Purification under denaturing conditions was carried out using 50 mM tris-HCl, 0.5 M NaCl, 20 mM imidazole, 6 M urea, pH 8 buffer. The soluble crude lysate was bound to Ni-NTA beads and allowed to equilibrate for 3 h at 4 °C. The proteins were eluted with increasing gradient of imidazole (20 mM − 1 M). Pure fractions were refolded via stepwise dialysis in pH adjusted phosphate buffer (50 mM), halving the urea concentration successively. A BCA kit was used to estimate protein concentration with bovine serum albumin as a standard.

*Circular Dichroism.* Circular dichroism (CD) measurements were conducted on a Jasco J-815 CD spectrometer. Wavelength and temperature scans were conducted with 10 $\mu$M ($6.3 \times 10^{-2}$ mg/mL) protein concentrations. The wavelength spectrum was measured over a range from 190 to 250 nm with a step size of 1 nm. Mean residue ellipticity (MRE) was calculated from raw data according to the procedure described in Gunasekar et al. 2009.[33] Secondary structure analysis of $\alpha$-helical, $\beta$-sheet, and random coil content was calculated with the K2D method using DichroWeb software.[37,38] Temperature scans of each protein were performed over a range of 20−85 °C with a temperature step of 1 °C/min at 222 nm. Scans were also performed at of $2^{18}$ and 5 °C/min[39] to evaluate dependence of thermal melt signatures on scan speed. All measurements were made in duplicates of independently prepared proteins and data represents the average. Thermodynamic properties of wt and Q were determined through analysis of thermal melts. A two-state model was used, where assumptions included monophasic behavior and reversible melting behavior.[39,40] These assumptions were confirmed experimentally by melting (from 20 to 85 °C) and cooling (from 85 to 20 °C) the proteins. Calculation of thermodynamic parameters including $T_m$, $\Delta G°$, $\Delta H°$, and $\Delta S°$ were performed according to the method described by Greenfield.[39]

*Transmission Electron Microscopy.* A JEOL JEM-1400 transmission electron microscope (TEM) was used to study the supramolecular protein structure. Approximately 3 $\mu$L of 10 $\mu$M protein in 50 mM PB was spotted on copper grids. After 1 min, the grids were blotted using filter paper and rinsed with 3−4 drops Milli Q water to remove excess salts from the buffer. After blotting with filter paper, the sample was negatively stained by adding 3 $\mu$L of 1% filtered uranyl acetate, blotted using filter paper, and dried at room temperature for 10−15 min. ImageJ software was used to measure the fibers dimensions.[41]

*Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy.* Attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) experiments were performed using PerkinElmer System 2000 FT-IR with DuraSamplIR II T diamond ATR accessory and equipped with a MCT-A detector. Approximately 5 $\mu$L of peptide solution (10 $\mu$M in 50 mM PB, pH 4, 8, and 10) was added on the diamond ATR surface. The spectrum (128 scans) was measured at room temperature over a range of 4000−400 $cm^{-1}$ with 0.5 $cm^{-1}$ resolution. PeakFit software was used to process the data, which involved a second derivative zero baseline correction of the amide I region between 1700−1600 $cm^{-1}$ and deconvolution of peaks with a Gaussian function.[42] All readings represent the average of two trials.

*Zeta Potential.* Zeta potential measurements were performed on a Zetasizer Nano Series model Nano ZS90. 50 mM phosphate buffer at pH 4 was used to saturate the clear polycarbonate disposable zeta cell DTS1060C prior to injecting 750 $\mu$L of protein samples at protein concentrations of 10 $\mu$M. The following settings were used for zeta potential measurements in the DTS (Nano) software: 90 °C instrument settings, Smoluvchoski model, material protein, dispersant PBS, temperature of 25 °C, viscosity 1.0200 cP, and dielectric constant of 1.34. Measurements were taken in triplicates, conducting 10 runs for each measurement, with a delay time of 2 s between each measurement.

*Dynamic Light Scattering.* Dynamic light scattering (DLS) measurements were performed on a Zetasizer Nano Series model Nano ZS90. In measuring DLS, 50 mM phosphate buffer at pH 4 was used to wash the low volume disposable cuvette DTS0112 cell. Approximately, 750 $\mu$L of 10 $\mu$M protein sample with varying amounts of curcumin was then added to the cell. To measure size, the following settings were applied: material protein (refractive index 1.450), absorption 0.001, dispersant PBS with a viscosity of 1.0200 cP and refractive index of 1.335. Measurements were taken in triplicates, conducting 10 runs for each measurement, with a delay time of 2 s between each measurement.

*Confocal Microscopy.* Samples were imaged using a Leica TCS SP2 AOBS confocal microscope system equipped with argon ion and HeNe lasers. A 63×/1.4 NA oil-immersion objective was used for all of the images. Lab-Tek II chambered #1.5 German coverglass system was used as the imaging slide. Curcumin was excited using the 458 nm line of the argon laser, and images were taken with the detection window set between 465 and 560 nm. The pinhole aperture was set at an Airy value of 1.0, which was equivalent to sampling an ∼500 nm vertical *z* slice of the fiber, as estimated by the axial resolution, $r_{z,\mathrm{confocal}} \approx 1.4\lambda_{em}n/\mathrm{NA}^2$ (NA, numerical aperture; *n*, refractive index; $\lambda_{em}$, emission wavelength (525 nm)). Interference contrast images were obtained using the Leica tube optics HC 1X/B apparatus with a focusing Bertrand lens. The 3D reconstructions were constructed using ImageJ 64 1.43 in concert with Amira 5.43, employing the Volren 3D rendering routine.

*Nuclear Magnetic Resonance.* 1D $^1$H NMR was performed on a Bruker Ultrashield 500 Plus instrument and data was collected and analyzed using TopSpin 3.2 software. Protein concentrations were kept constant at 20 $\mu$M, with buffer conditions of 50 mM PB pH 4 with 1% (v/v) methanol and 1% (v/v) $D_2O$. NMR was performed in

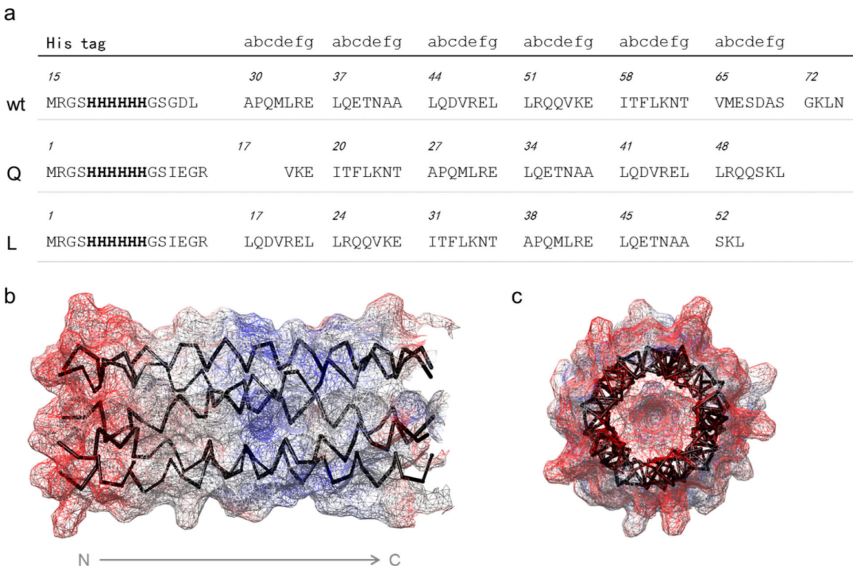



**Figure 1.** Protein sequences and surface charge representation for Q. (a) Sequence information for wt, Q, and L, from N-terminus to C-terminus, with the histidine tags in bold. Surface charge representation of Q pentamer under acidic conditions viewed along (b) and down (c) the pentamer axis. Positive red patches in (b) and (c) are attributed to solvent-exposed lysine and arginine residues. Negative blue patches in (b) and (c) are attributed to solvent-exposed glutamate and aspartate residues.

the absence and presence of curcumin, in a 5:1 molar ratio of curcumin/protein.

## RESULTS AND DISCUSSION

**Design Methodology.** Here, we designed Q, in which residues not essential for pentamer assembly in wt were eliminated and surface residues were distributed according to charge for optimal lateral assembly under acidic conditions[43,44] (Figure 1). As has been taken advantage of previously,[45] the thermodynamic driving force for self-assembly of protein-based fiber structures consisted of optimal distribution of surface charges on the solvent-exposed outside of the homopentamer combined with shielding of the aliphatic residues within the pore of the protein oligomer.

The wt consisted of a repeat of three leucines (L37, L44, L51) in the *a* site within the N-terminal pocket along with a valine (V47) in the *d* site of the adjacent helix, known to be indispensable for the formation of stable pentamers (Figure 1a).[33] The corresponding residues in Q became L34, L41, L48, and V44, all maintaining the same positions within the heptad. The extraneous last heptad of the wt protein was eliminated and the C-terminal portion after the glutamine residue at the *d* position was swapped to the N-terminus to better distribute the surface charge for lateral assembly, producing Q (Figure 1a). In order to compare structural and assembly characteristics of this engineered protein, we also designed a negative control, L, which was engineered to disrupt the N-terminal pocket crucial region for structure, stability, and pentamer formation by swapping at the leucine at position 44[33] (Figure 1a). Both the Q and L sequence was constructed via standard recombinant DNA methods and subsequently biosynthesized via bacterial expression followed by purification. As the design of Q and L focused on examining the effects of swapping the regions of wt, the proline residue that was present in the *b* position of wt was also in the engineered constructs. While proline is traditionally known as a helix breaking residue,[46] its presence had a minimal effect on helical structure. Swapping of the N- and C-terminal regions to generate Q and L result in proline being displaced to the center of the proteins, in positions 28 and 39 in Q and L, respectively. This decision was made based on work that has shown that the presence of proline in the middle of an $\alpha$-helix can be accommodated by a local break in the structure, where only the residue preceding the proline (A27 in Q and A38 in L) is not in $\alpha$-helical conformation.[47] The $\alpha$-helical structure resumes unbroken either side of the proline, however.[47]

The homopentameric assembly generated from Q subunits was visualized using Chimera (Figure 1b,c). Electrostatic charge distribution in patches has previously been used to facilitate and direct self-assembly of coiled-coil protein fibers, confirmed by cryo-TEM, X-ray crystallography/diffraction, and modeling.[48,49] The surface charge representation was generated by fully protonating the acidic residues in the software to best represent the charge distribution at acidic pH conditions (Figures 1b,S1, and S2). Red regions on the termini of the coiled-coil were positively charge as a result of solvent-exposed arginine and lysine residues, while the center of the pentamer was negatively charged due to solvent-exposed glutamine and asparagine residues (Figures 1b and S1). The overwhelmingly negative and positive "patches" that were produced along the length of the pentamer contribute to its ability to self-assemble. In contrast,

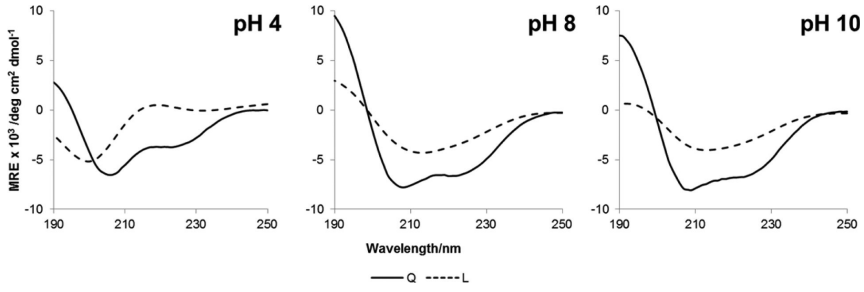



**Figure 2.** Secondary structure of proteins at different pH conditions. Circular dichroism variable wavelength scans of Q (solid line) and L (dashed line) at pH 4, 8, and 10. Data is averaged from at least two replicates and was obtained with 10 $\mu$M protein concentration.

surface charge distribution of wt did not show charged patches but rather exhibited an overwhelmingly negative surface charge along the entire pentamer (Figure S2).

**Secondary Structure in Solution.** To assess the secondary structure of the proteins under acidic as well as neutral and basic conditions in solution, CD measurements were performed.

While the removal of the peripheral heptad and reordering from wt led to a dampened signal, Q exhibited a helical signature with a double minimum of $-3.71 \times 10^3$ deg cm$^2$ dmol$^{-1}$ and $-6.54 \times 10^3$ deg cm$^2$ dmol$^{-1}$ at 222 and 206 nm at pH 4, respectively (Figure 2, Table S5). The presence of residue P28 toward the center of Q may have influenced the secondary structure, showing decreased helical content when compared to wt (Figures S3 and S4). The negative control, L, was completely unstructured, devoid of helical content at pH 4 (Figure 2). At neutral pH 8 conditions, Q revealed more helical structure (Figure 2, Table S5, and Figure S4). The negative control, L, did not illustrate $\alpha$-helical structure at pH 8; rather it presented a single minimum at 213 nm (Figure 2). At pH 10, Q maintained helical conformation with little change in the double minimum values, while L exhibited the same single minima (Figure 2 and Table S5). While residue P39 in L may have contributed to a small loss in structure (as was seen in Q), the complete loss in helical structure across all pH conditions can be attributed to the disruption of the N-terminal pocket due to domain swapping. As the pH increased, the helical content of soluble Q protein increased. L exhibited a small increase in helical content from pH 4 to pH 10 and a moderate increase in random coil content.

**Thermodynamic Properties.** Thermodynamic properties were assessed via thermal melt of the proteins at 222 nm. Prior to calculating thermodynamic constants, assumptions of monophasic behavior and melt reversibility were confirmed experimentally for wt and Q at all pHs by melting and cooling proteins (Figure S5). The van't Hoff equation was applied to thermal melts of the proteins in the range of 20−85 °C. Overall, Q exhibited excellent stability across all pH values with melting temperatures ($T_m$) of 46.4−63.5 °C and Gibbs free energy ($\Delta G^\circ$) of −3.3, −3.8, and −3.1 kcal/mol at pH 4, 8, and 10, respectively (Table 1).

As expected, thermal melts of L did not yield a significant enough gradient in the ellipticity at 222 nm to calculate thermodynamic properties. Compared to the parent wt, Q demonstrated a 7.1 and 14.7 °C increase in $T_m$ at pH 4 and 8, respectively, affirming that the modification made for the design of Q was indeed stabilizing (Table S6). At pH 10, the $T_m$ of Q and wt was essentially equal. Overall, Q was more stable at acidic and neutral pH conditions relative to wt, which could be attributed to the surface charge distribution along the pentamer subunits.

**Table 1. Thermodynamic Constants Obtained from Thermal Melts of 10 $\mu$M Q, Measured by Circular Dichroism**

| pH | $T_m$ (°C) | $\Delta H^\circ$ (kcal mol$^{-1}$)[a] | $\Delta S^\circ$ (kcal mol$^{-1}$ K$^{-1}$)[b] | $\Delta G^\circ$ (kcal mol$^{-1}$)[c] |
|---|---|---|---|---|
| 4 | 55.3 | −32.3 | −97.1 | −3.3 |
| 8 | 63.5 | −32.6 | −96.8 | −3.8 |
| 10 | 46.4 | −46.4 | −145.2 | −3.1 |

[a]van't Hoff enthalpy calculated as described in Supporting Information. [b]At equilibrium, $\Delta G^\circ = 0$. Hence, the change in entropy $\Delta S^\circ = \Delta H^\circ / T_m$. [c]Free energy of folding at 25 °C calculated according to the expression $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$.

**Protofibrils Bundle to Form Fibers.** To determine whether Q could self-assemble into fibers, TEM analysis was performed. While limited fiber formation was observed at neutral pH, an abundance of fibers was observed under acidic conditions, as expected from our design (Figures 3a and S6), consistent with previous work on coiled-coil fibers.[31,50] Contrary to previous work on the wt protein, however, the mechanism of protofibril bundling via self-assembly to form large, bundled fibers was the electrostatic charge distribution throughout the Q pentamers.

Transmission electron micrographs of Q displayed the presence of bundled protofibrils forming high aspect ratio fibers (Figure 3a). These protofibrils possessed diameters of 3.5 ± 0.5 nm ($n$ = 210), which corresponded to the projected lateral dimension of pentameric helical bundles (Figure 3b). Note that the model shown in Figure 3b was an assumption based on the patched distribution of Coulombic surface charge in Q, with X-ray diffraction and cryo-TEM experiments underway. TEM data confirmed the diameters of protein fibers varied from tens to hundreds of nanometers (20−560 nm, $n$ = 14, Figure S6). At the upper limits, these fibers displayed tremendous lateral assembly of hundreds of protofibrils and thus larger than other de novo designed coiled-coil protein fibers to date.[9,11,15] While wt demonstrated fiber formation (Figure S6), the fibers possess diameters on the range of 10−15 nm[34] without any evidence for protofibrils. As expected, the L negative control revealed nonfibrous aggregates, affirming that

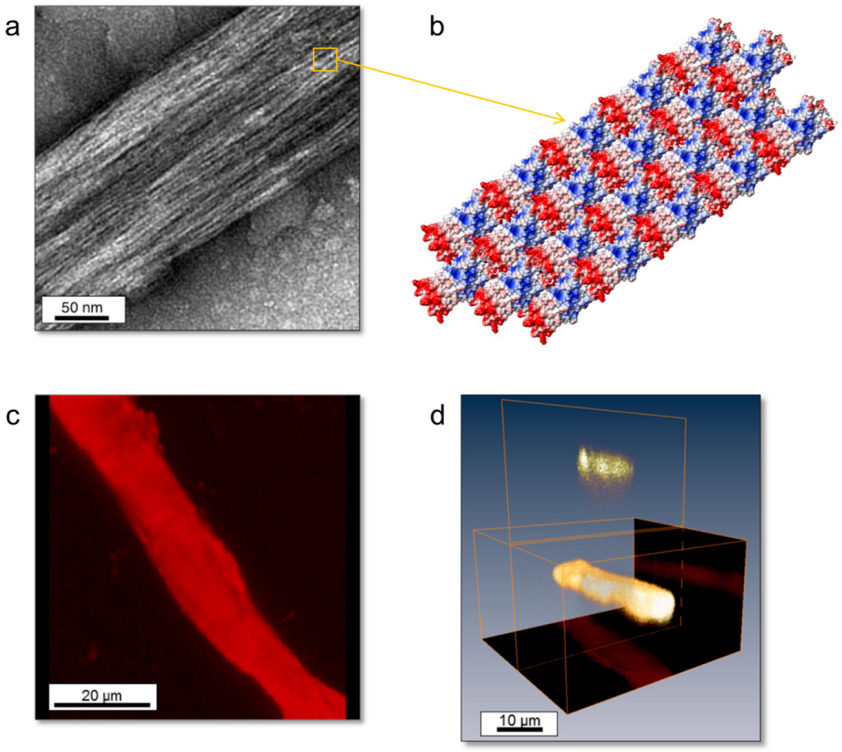



**Figure 3.** Microscopy and modeling of protein fibers. (a) Transmission electron micrograph of Q fiber, 10 $\mu$M, pH 4. (b) Schematic representation of Q fiber assembly with staggered positive (red) and negative (blue) regions of the pentamer. (c) Reconstruction of 3D confocal *XYZ* data of 10 $\mu$M Q protein in the presence of 50 $\mu$M curcumin (5:1 molar ratio of curcumin:protein). (d) 3D representation of the same Q fiber showing *XZ* and *YZ* orthogonal views and an oblique slice cross section above the 3D bounding box. Scale bars in (a), (c), and (d) represent 50 nm, 20 $\mu$m, and 10 $\mu$m, respectively.

**Table 2. Secondary Structure Conformation Based on ATR-FTIR Data for 10 $\mu$M Q in 50 mM PB at pH 4, 8, and 10, and in the Presence of Curcumin at a 5:1 Molar Ratio with Q at pH 4[a]**

| | | % composition | | | |
|---|---|---|---|---|---|
| conformation | wavelength[51] ($cm^{-1}$) | pH 4 | pH 8 | pH 10 | curcumin |
| $\beta$-sheet | 1625–1640, 1675–1695 | 38 | 37 | 32 | 10 |
| random coil | 1640–1648 | | | | 12 |
| $\alpha$-helix | 1648–1660 | 62 | 63 | 68 | 79 |

[a]Percent composition was determined from relative areas of peaks fit to spectra (Figure S7).

$\alpha$-helical conformation was indeed important for fiber assembly (Figure S4).

With the observation of defined fibers formed by Q, we were interested in quantifying the secondary structure of the protein in its solid-state. As a result, ATR-FTIR experiments were conducted to evaluate a secondary structure of insoluble Q fibers at pH 4, 8, and 10 (Table 2). The frequency measured in the regions of the amide I and amide II absorptions of a protein correlate to the secondary structural motifs within the protein[51] and were thereby used to assess conformation of the protein in its solid state (Figure S7).

Positions of amide I peaks in deconvoluted IR spectra of coiled-coil proteins have been shown previously to differ compared to peak locations arising from purely $\alpha$-helical, monomeric proteins.[52] Deviations are related to pitch values of the $\alpha$-helices within the coiled-coils, with dimers showing the

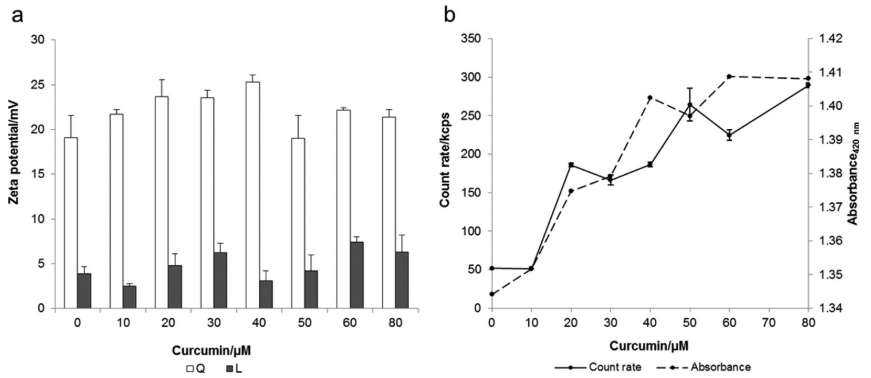



**Figure 4.** Aggregation of Q protein fibers due to increasing curcumin concentration. Protein concentration was held constant at 10 $\mu$M for zeta potential and DLS measurements. (a) Zeta potential as a function of curcumin concentration for Q (white bars) and L (gray bars). (b) Dynamic light scattering measurements of count rate for Q (solid line) as a function of curcumin concentration. Count rate was not obtained for L as signal at all curcumin concentrations was too low for detection. Absorbance measured at 420 nm as a function of curcumin concentration for Q (dashed line). Error bars in (a) and (b) represent an average of three trials.

largest deviation (corresponding to helix deformations) and higher order oligomeric coiled-coils more closely resembling $\alpha$-helical proteins.[52] Our data correlate well with these observations, as the significant peak weights lie near the classical $\alpha$-helical band position of 1650−1653 cm$^{-1}$ (Figure S7). ATR-FTIR measurements of Q at pH 4 results in a helical content of 62%, with helicity increasing to 63 and 68% at pH 8 and 10, respectively (Table 2). This trend of increasing helical content with increasing pH is consistent with the CD data (Figure S4). Solid-state ATR-FTIR data of Q secondary structure confirmed structured fibers visualized in TEM are indeed $\alpha$-helical.

**Curcumin Binding.** In the design of Q, the hydrophobic pore was maintained to enable binding to small molecules. The polyphenolic compound curcumin has long been used for many therapeutic purposes due to its antiproliferative,[53] antibacterial,[54] and anti-inflammatory[55] properties, but exhibits limitations in delivery methods due to its low solubility in aqueous solutions.[56] In addition, curcumin induces aggregation of protein fibers, such as collagen and the acidic $\alpha$-helical intermediate of PrP, a precursor to amyloid fibers.[56−58] Thus, we incubate Q with varying concentrations (0−80 $\mu$M) of curcumin at pH 4 in order to study its overall binding ability and effects on protein morphology, charge, and aggregation.

*Secondary Structure in the Presence of Curcumin.* Circular dichroism measurements reveal that the conformation of Q is not disturbed upon interaction with curcumin in accordance with structural data on the binding of curcumin to collagen.[57] In fact, the absolute value of the MRE value at 222 nm, an indicator of helicity, of Q increased linearly with increasing concentrations of curcumin (Figure S8). In addition, ATR-FTIR measurements were performed in the presence of a 5:1 molar ratio of curcumin to protein at pH 4. As can be seen in Table 2, $\alpha$-helical composition of insoluble protein fibers increases dramatically upon the addition of curcumin, going from 62 to 79% at pH 4. The ATR-FTIR spectra are particularly telling in this case (Figure S7), where spectra collected in the presence of curcumin demonstrate a very large, sharp peak of high intensity at 1653 cm$^{-1}$. These results suggest a stabilization of coiled-coil assembly in the presence of curcumin.

*Macromolecular Assembly.* As curcumin bound to structured protein exhibits fluorescence,[33] confocal microscopy was performed on the Q-curcumin complex (Figure 3c,d). Surprisingly, microfibers were observed with an average diameter of 16.0 ± 5.6 $\mu$m ($n$ = 22) (Figures 3c,d and S9), indicating that the presence of curcumin further promotes fiber assembly. TEM was also used to study aggregation of Q in the presence of curcumin (Figure S10). Given the size of the large aggregates visualized in confocal, however, fibers of this dimension were not easily visualized via TEM, as they did not adhere to the surface of the TEM grids and were easily wicked off when drying and preparing the samples. More importantly, curcumin was found distributed homogeneously throughout the fiber (Figure 3d), and furthermore, the interference contrast boundaries of the fibers coincided with the curcumin emission boundaries in the confocal images (Figure S9). NMR experiments confirmed peaks between 0 and 2 ppm, a region pertaining to aliphatic residues within the protein and one that rarely exhibits signals from organic molecules.[59] Peak broadening was demonstrated in the presence of curcumin in a 5:1 molar ratio of curcumin/protein (Figure S11). These results suggested that curcumin was interacting with the nonpolar, hydrophobic residues located within the pore of wt and Q pentamers, however, did not preclude the possibility that curcumin could also be binding between the protofibrils. Fluorescence distributions from confocal measurements indicated that curcumin could be binding between protofibrils, contributing to observed aggregation effects. NMR and confocal data revealed that curcumin likely bound within the coiled-coil pore in addition to the surface of the pentamers, promoting supramolecular assembly (vida infra).

*Evidence of Aggregation.* To further characterize the assembly and aggregation caused by the addition of curcumin, zeta potential, count rate, and absorbance at 420 nm of protein

is studied (Figure 4). With curcumin being fully protonated at pH 4, the interaction of curcumin with negative charges on the protein is likely.[56] Indeed, as curcumin concentration increases, the zeta potential exhibits a slight increase as well, indicating a neutralization of any negative charges through binding to curcumin (Figure 4a). For Q, zeta potential in the absence of curcumin is 19.1 ± 2.5 mV, with a steady increase up to a maximum value of 25.2 ± 0.8 mV at a curcumin concentration of 40 $\mu$M (Figure 4a). Above a molar concentration of 40 $\mu$M curcumin, however, zeta potential drops to values similar to Q in the absence of curcumin, suggesting that neutralization is occurring by curcumin on the fiber surface. By contrast, L possesses a much lower zeta potential than Q both in the absence and presence of curcumin at all concentrations studied (Figure 4a). Count rate from dynamic light scattering (DLS) is also studied to assess aggregation. For Q (in which the concentration was held constant at 10 $\mu$M), a steady increase in response to higher curcumin concentrations is observed (Figure 4b); however, count rate could not be obtained for L as the particles did not generate a high enough signal for detection, confirming the absence of aggregates. Absorbance of proteins at 420 nm in the presence of curcumin display a similar increasing trend (Figures 4b and S12). While the zeta potential, count rate, and absorbance data values for Q are similar to those exhibited by wt on a macromolecular level (Figure S12), analysis of Coulombic surface charge shows that on a molecular level the two protein assemblies have very different charge distributions.

The curcumin-induced aggregation mechanism of collagen has been characterized by Fathima et al. using CD, surface tension, and viscosity measurements.[56,57] At pH 4, fully protonated curcumin will have strong interactions with negatively charged amino acids on protein surfaces. Morphology and aggregation of $\alpha$-helical protein fibers results from an energetically favorable balance between the burial of the hydrophobic residues and the solvation of the surface-facing polar groups.[56] Our CD studies have confirmed that increasing molar ratios of curcumin affects only helical packing, not protein conformation. Curcumin increases the surface activity of collagen as a result of local restructuring of water, leading to some exposure of nonpolar groups.[57] The increased exposure of nonpolar groups provides a driving force for aggregation, as the system strives to reduce overall surface energy, explaining the mechanism of curcumin-induced aggregation of these $\alpha$-helical protein fibers.

## CONCLUSION

Aggregation of protein fibers has been thoroughly studied in the case of $\beta$-sheet amyloid fibers, as their aggregation is significant in the onset of Alzheimer's disease.[5] We have engineered proteins that can hierarchically assemble into mesofibers with the assistance of small molecules through encapsulation and aggregation (Figure S13). While studies have been performed on controlling aggregation of coiled-coil protein fibers through macromolecular crowding[50] and adjusting solvent conditions such as pH[60] and ionic strength,[11] save a few exceptions,[56,57,61] comparatively little work has been done on inducing aggregation of $\alpha$-helical protein fibers through additives. In such examples, change on the order of the nanometer to micrometer scale of de novo designed $\alpha$-helical protein fibers has not been previously reported. Additionally, the ability to generate fibers possessing nanoscale organization with concentrations on the micromolar range provides a tremendous technical advantage in future applications as biomaterials. Herein, we have demonstrated the synthesis of protein fibers of dimensions comparable to $\alpha$-keratin, collagen, and spider silk[62] with the added functionality of small molecule binding.[63,64] These fibers hold potential for tissue engineering and delivery of curcumin and other therapeutic small molecules. Future experiments to probe the fiber microstructure in the presence and absence of curcumin and other small molecules, such as X-ray diffraction, are underway.

## ASSOCIATED CONTENT

### Supporting Information

Details on genetic engineering employed to obtain Q and L constructs, CD, ATR-FTIR, NMR, galleries of TEM and confocal data collected, and experimental data for wt. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: montclare@nyu.edu.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank Joseph Frezzo, Chin Lin, and Kim Kisslinger for technical assistance and critical insight into the experiments and manuscript. TEM measurements were performed at the Center for Functional Nanomaterials at Brookhaven National Laboratory, supported by the Office of Basic Energy Sciences, U.S. DOE. Confocal microscopy was performed at CCNY. This work was supported by the ARO (W911NF-11-1-0449), a GK-12 Fellows Grant DGE-0741714, the NSF DMR BMAT program (NSF 1207480; M.L.G., J.A.M.), and in part by the NSF MRSEC Program under Award Number DMR-0820341.

## REFERENCES

(1) Mann, S. *Nature* **1993**, *365*, 499−505.
(2) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 1304−1319.
(3) Sanchez, C.; Arribart, H.; Guille, M. M. G. *Nat. Mater.* **2005**, *4*, 277−288.
(4) Soto, C.; Brahes, M. C.; Alvarez, J.; Inestrosa, N. C. *J. Neurochem.* **1994**, *63*, 1191−1198.
(5) Rousseau, F.; Schymkowitz, J.; Serrano, L. *Curr. Opin. Struct. Biol.* **2006**, *16*, 118−126.
(6) Kramer, R. M.; Crookes-Goodson, W. J.; Naik, R. R. *Nat. Mater.* **2007**, *6*, 533−538.
(7) Crookes, W. J.; Ding, L.-L.; Huang, Q. L.; Kimbell, J. R.; Horwitz, J.; McFall-Ngai, M. J. *Science* **2004**, *303*, 235−238.
(8) Miserez, A.; Li, Y.; Cagnon, J.; Weaver, J. C.; Waite, J. H. *Biomacromolecules* **2012**, *13*, 332−341.
(9) Kojima, S.; Kuriki, Y.; Yoshida, T.; Yazaki, K.; Miura, K. *Proc. Jpn. Acad., Ser. B* **1997**, *73*, 7−11.
(10) Bromley, E. H. C.; Channon, K. J.; King, P. J. S.; Mahmoud, Z. N.; Banwell, E. F.; Butler, M. F.; Crump, M. P.; Dafforn, T. R.; Hicks, M. R.; Hirst, J. D.; Rodger, A.; Woolfson, D. N. *Biophys. J.* **2010**, *98*, 1668−1676.
(11) Papapostolou, D.; Bromley, E. H. C.; Bano, C.; Woolfson, D. N. *J. Am. Chem. Soc.* **2008**, *130*, 5124−5130.
(12) Smith, A. M.; Acquah, S. F. a; Bone, N.; Kroto, H. W.; Ryadnov, M. G.; Stevens, M. S. P.; Walton, D. R. M.; Woolfson, D. N. *Angew. Chem.* **2004**, *44*, 325−328.
(13) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, a; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728−8734.
(14) Ryadnov, M. G.; Woolfson, D. N. *Nat. Mater.* **2003**, *2*, 329−332.

(15) Ogihara, N. L.; Ghirlanda, G.; Bryson, J. W.; Gingery, M.; DeGrado, W. F.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 1404−1409.
(16) Ryadnov, M. G.; Bella, A.; Timson, S.; Woolfson, D. N. *J. Am. Chem. Soc.* **2009**, *131*, 13240−13241.
(17) Wagner, D. E.; Phillips, C. L.; Ali, W. M.; Nybakken, G. E.; Crawford, E. D.; Schwab, A. D.; Smith, W. F.; Fairman, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12656−12661.
(18) Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P. *Tetrahedron* **2004**, *60*, 7237−7246.
(19) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. *Science* **2013**, *340*, 595−599.
(20) Yu, S. M.; Li, Y.; Kim, D. *Soft Matter* **2011**, *7*, 7927.
(21) Buehler, M. J. *J. Mech. Behav. Biomed. Mater.* **2008**, *1*, 59−67.
(22) O'Leary, L. E. R.; Fallas, J. a; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. *Nat. Chem.* **2011**, *3*, 821−828.
(23) Anderson, V. L.; Ramlall, T. F.; Rospigliosi, C. C.; Webb, W. W.; Eliezer, D. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18850−18855.
(24) Wei, G.; Reichert, J.; Bossert, J.; Jandt, K. D. *Biomacromolecules* **2008**, *9*, 3258−3267.
(25) MacArthur, I. *Nature* **1943**, *152*, 38−41.
(26) dos Remedios, C. G.; Thomas, D. D. *Results and Problems in Cell Differentiation*; Springer: New York, **2001**; Vol. *32*, pp 1−7.
(27) Kim, C. a; Phillips, M. L.; Kim, W.; Gingery, M.; Tran, H. H.; Robinson, M. a; Faham, S.; Bowie, J. U. *EMBO J.* **2001**, *20*, 4173−4182.
(28) Eisenberg, D.; Liu, Y. *Protein Sci.* **2002**, *11*, 1285−1299.
(29) Padilla, J. E.; Colovos, C.; Yeates, T. O. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2217−2221.
(30) Li, Y.; Mo, X.; Kim, D.; Yu, S. M. *Biopolymers* **2011**, *95*, 94−104.
(31) Potekhin, S. A. a; Melnik, T. N. N.; Popov, V.; Lanina, N. F. F.; Vazina, A. A. a; Rigler, P.; Verdini, A. S. S.; Corradin, G.; Kajava, A. V. V *Chem. Biol.* **2001**, *8*, 1025−1032.
(32) Gunasekar, S. K.; Haghpanah, J. S.; Montclare, J. K. *Polym. Adv. Technol.* **2008**, *19*, 454−468.
(33) Gunasekar, S. K.; Asnani, M.; Limbad, C.; Haghpanah, J. S.; Hom, W.; Barra, H.; Nanda, S.; Lu, M.; Montclare, J. K. *Biochemistry* **2009**, *48*, 8559−8567.
(34) Gunasekar, S. K.; Anjia, L.; Matsui, H.; Montclare, J. K. *Adv. Funct. Mater.* **2012**, *22*, 2154−2159.
(35) Briki, F.; Doucet, J.; Etchebest, C. *Biophys. J.* **2002**, *83*, 1774−1783.
(36) Lazaris, A.; Arcidiacono, S.; Huang, Y.; Zhou, J.-F.; Duguay, F.; Chretien, N.; Welsh, E. A.; Soares, J. W.; Karatzas, C. N. *Science* **2002**, *295*, 472−476.
(37) Whitmore, L.; Wallace, B. A. *Biopolymers* **2008**, *89*, 392−400.
(38) Whitmore, L.; Wallace, B. A. *Nucleic Acids Res.* **2004**, *32*, W668−73.
(39) Greenfield, N. J. *Nat. Protoc.* **2006**, *1*, 2527−2535.
(40) Pace, C. N. *Trends Biotechnol.* **1990**, *8*, 93−98.
(41) Abràmoff, M. D.; Magalhães, P. J. *Biophotonics Int.* **2004**, *11*, 36−42.
(42) Hu, X.; Kaplan, D.; Cebe, P. *Macromolecules* **2006**, *39*, 6161−6170.
(43) Guo, Y.; Bozic, D.; Malashkevich, V. N.; Kammerer, R. a; Schulthess, T.; Engel, J. *EMBO J.* **1998**, *17*, 5265−5272.
(44) McFarlane, A. a; Orriss, G. L.; Stetefeld, J. *Eur. J. Pharmacol.* **2009**, *625*, 101−107.
(45) Xu, C.; Liu, R.; Mehta, A. K.; Guerrero-ferreira, R. C.; Wright, E. R.; Dunin-horkawicz, S.; Morris, K.; Serpell, L. C.; Zuo, X.; Wall, J. S.; Conticello, V. P. *J. Am. Chem. Soc.* **2013**, *135*, 15565−15578.
(46) Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211−222.
(47) Piela, L.; Nemethy, G.; Harold, A. *Biopolymers* **1987**, *26*, 1587−1600.
(48) Papapostolou, D.; Smith, A. M.; Atkins, E. D. T.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10853−10858.
(49) Sharp, T. H.; Bruning, M.; Mantell, J.; Sessions, R. B.; Thomson, A. R.; Zaccai, N. R.; Brady, R. L.; Verkade, P.; Woolfson, D. N. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 13266−13271.
(50) Munishkina, L. A.; Ahmad, A.; Fink, A. L.; Uversky, V. N. *Biochemistry* **2008**, *47*, 8993−9006.
(51) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95−120.
(52) Heimburg, T.; Schünemann, J.; Weber, K.; Geisler, N. *Biochemistry* **1999**, *38*, 12727−12734.
(53) Yallapu, M. M.; Jaggi, M.; Chauhan, S. C. *Drug Discovery Today* **2012**, *17*, 71−80.
(54) Rai, D.; Singh, J. K.; Roy, N.; Panda, D. *Biochem. J.* **2008**, *410*, 147−155.
(55) Zsila, F.; Bikádi, Z.; Simonyi, M. *Bioorg. Med. Chem.* **2004**, *12*, 3239−3245.
(56) Fathima, N. N.; Dhathathreyan, A.; Ramasami, T. *J. Chem. Sci.* **2010**, *122*, 881−889.
(57) Nishad Fathima, N.; Saranya Devi, R.; Rekha, K. B.; Dhathathreyan, A. *J. Chem. Sci.* **2009**, *121*, 509−514.
(58) Hafner-Bratkovic, I.; Gaspersic, J.; Smid, L. M.; Bresjanac, M.; Jerala, R. *J. Neurochem.* **2008**, *104*, 1553−1564.
(59) Stebbins, J. L.; Zhang, Z.; Chen, J.; Wu, B.; Emdadi, A.; Williams, M. E.; Cashman, J.; Pellecchia, M. *J. Med. Chem.* **2007**, *50*, 6607−6617.
(60) Anderson, V. L.; Webb, W. W.; Eliezer, D. *Phys. Biol.* **2012**, *9*, 1−9.
(61) Fallas, J. a; O'Leary, L. E. R.; Hartgerink, J. D. *Chem. Soc. Rev.* **2010**, *39*, 3510−3527.
(62) Hagn, F.; Eisoldt, L.; Hardy, J. G.; Vendrely, C.; Coles, M.; Scheibel, T.; Kessler, H. *Nature* **2010**, *465*, 239−242.
(63) Gopinath, D.; Ahmed, M. R.; Gomathi, K.; Chitra, K.; Sehgal, P. .; Jayakumar, R. *Biomaterials* **2004**, *25*, 1911−1917.
(64) Vogel, K. G.; Trotter, J. a. *Collagen Relat. Res.* **1987**, *7*, 105−114.

nmt00 | ACSJCA | JCA10.0.1465/W Unicode | research.3f (R3.6.i8:4237 | 2.0 alpha 39) 2015/04/10 13:33:00 | PROD-JCA1 | rq_3516147 | 4/29/2015 09:35:12 | 8 | JCA-DEFAULT

Biomacromolecules

Article

pubs.acs.org/Biomac

# Tunable Conformation-Dependent Engineered Protein·Gold Nanoparticle Nanocomposites

Jasmin Hume,[†] Raymond Chen,[†] Rudy Jacquet,[‡] Michael Yang,[†] and Jin Kim Montclare*[,†,§,∥]

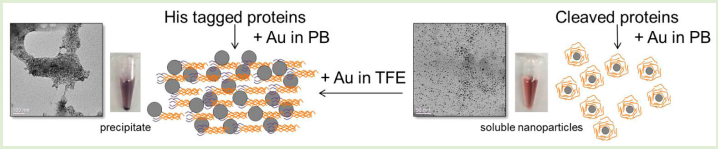
[†]Department of Chemical and Biomolecular Engineering, New York University Polytechnic School of Engineering, Brooklyn, New York 11201, United States

[‡]College of Physicians and Surgeons, Columbia University, New York, New York 10032, United States

[§]Department of Biochemistry, State University of New York Downstate Medical Center, Brooklyn, New York 11203, United States

[∥]Department of Chemistry, New York University, New York, New York 10003, United States

Supporting Information




**ABSTRACT:** We demonstrate the fabrication of protein·gold nanoparticle (AuNP) nanocomposites in situ, leading to distinct assemblies dependent upon protein secondary structure. In the presence of pentameric coiled-coil proteins C and Q, which contain histidine tags and have helicities of 54 and 65%, respectively, templation of AuNP results in precipitation of the protein·AuNP composites with AuNPs 6.5 nm in diameter, creating macromolecular assemblies on the micrometer scale. In the absence of the histidine tags, the resulting Cx and Qx proteins, which exhibit lower helicities of 37 and 45%, respectively, stabilize soluble protein·AuNP composites with AuNPs 4.5 nm in diameter for several days without aggregating. By manipulating protein structure via external triggers, such as TFE, we obtain control over the macromolecular conformation and overall physicochemical properties. These hybrid protein·AuNP assemblies can be readily deposited on electrodes, where they can serve as a tunable bionanocomposite kinetic barrier.

## INTRODUCTION

Nanoparticle assemblies, particularly those that are stabilized by biomolecules, are of increasing interest for reaction catalysis,[1] biomedical imaging[2], biological labeling,[3] and sensing applications.[4,5] Research on generating stable, multipurpose gold nanoparticle (AuNP) assemblies using bioconjugates has generated much interest, especially in the area of creating biosensors of superior sensitivity and selectivity.[6−8] The stabilization of AuNPs with proteins takes advantage of the incredible physical properties of AuNPs while adding biological functionalization, leading to protein·AuNP assemblies.[9−14]

Although previous work on protein·AuNP assemblies has focused predominantly on templation of AuNPs by structured proteins,[11,15−21] there are examples of templation by unstructured protein.[15,20,22,23] Distribution of hydrophobic and hydrophilic groups in structured and unstructured proteins has been shown to have a significant impact on dimensions and spacing of templated AuNPs,[11,16] to the extent that protein secondary structure can be ascertained by probing the chiroptic properties of protein·AuNP nanocomposites.[15,24] Because secondary structure plays an important role in AuNP templation, it follows that macromolecular assembly of the nanocomposites is often also dictated by peptide structure.[15,22] This dependence on structure is seen for $\alpha$-helical peptide motifs[15,16,19,21−23] as well as for $\beta$ sheets.[11,17,18]

Manipulation of secondary structure, and hence control over assembly, can be accomplished using a variety of external triggers, including modification of solvent pH and polarity,[21,23,25,26] photothermal perturbations,[27,28] alteration of peptide concentration,[23,24] or the use of additives such as zinc.[29,30] Specific examples include reversible assembly/disassembly of coiled-coil protein·AuNP aggregates upon pH switching between acidic and neutral or basic conditions.[21,25] In such cases, the decrease in pH results in the change of conformation from unpaired, unstructured $\alpha$ helices to coiled-coil dimers with premade AuNPs chemically conjugated to the proteins or peptides.[21,25] The combination of governable proteins with inorganic nanoparticles brings new possibilities to construct and deconstruct functional hybrid materials, often an unattainable goal for bulk metals. It is our aim to expand the control afforded by external triggers while also fabricating protein·AuNP assemblies under ambient conditions in a one-pot procedure. Herein, we develop coiled-coil proteins capable

Received: January 23, 2015
Revised: April 9, 2015

ACS Publications © XXXX American Chemical Society



DOI: 10.1021/acs.biomac.5b00098
Biomacromolecules XXXX, XXX, XXX−XXX

**Table 1. Sequence Information for C, Q, Cx, and Qx, from N Terminus to C Terminus, with the Histidine Tags in Bold**

| | His tag | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg | abcdefg |
|---|---|---|---|---|---|---|---|
| | *1* | *17* | *24* | *31* | *38* | *45* | *52* |
| C | MRGS**HHHHHH**GSIEGR | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | *1* | *17* | *20* | *27* | *34* | *41* | *48* |
| Q | MRGS**HHHHHH**GSIEGR | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL |
| | | *1* | *8* | *15* | *22* | *29* | *36* |
| Cx | | APQMLRE | LQETNAA | LQDVREL | LRQQVKE | ITFLKNT | SKL |
| | | *1* | *4* | *11* | *18* | *25* | *32* |
| Qx | | VKE | ITFLKNT | APQMLRE | LQETNAA | LQDVREL | LRQQSKL |

of templating AuNPs in situ, resulting in tunable conformation-dependent protein·AuNP nanocomposites that are readily deposited on electrodes.

## EXPERIMENTAL SECTION

**Materials.** Sodium phosphate (monobasic and dibasic), nickel–nitrilotriacetic acid (Ni-NTA) resins, $HAuCl_3{\cdot}3H_2O$ salt, sodium borohydride, potassium hexacyanoferrate(II) trihydrate, potassium hexacyanoferrate(III), and trifluoroethanol (TFE) were purchased from Sigma-Aldrich. Ampicillin, isopropyl-$\beta$-D-thiogalactopyranoside (IPTG), tryptone, urea, tris-HCl, and sodium chloride were obtained from Fisher Scientific. Yeast extract and methanol were purchased from Acros Organics, and the BCA kit was obtained from Pierce. Imidazole was purchased from Alfa Aesar, and copper grids for transmission electron microscopy (TEM) were purchased from Ted Pella. Dialysis tubing (3.5 kDa molecular weight cutoff) was obtained from Thermo Scientific. Factor Xa cleavage kit was purchased from Novagen. Glassy-carbon electrodes were obtained from CH Instruments, Inc.

**Methods.** *Genetic Engineering.* Gene sequences for C and Q were generated via polymerase chain reaction (PCR) amplification. Construction of the Q gene has been described previously,[31] and a detailed description of creating the C gene is in the Supporting Information.

*Protein Expression.* Conditions for expression and purification of C and Q remained the same as those described previously.[31] Protein was dialyzed into 50 mM phosphate buffer (PB) at pH 8. For experiments in TFE, protein was diluted in PB containing 25 v/v % TFE. After confirming purity and protein concentration of >10 $\mu$M, cleavage was initiated. Cleavage of the histidine tags was carried out by using Factor Xa cleavage kit (Novagen). Cleavage was carried out according to the manufacturer's protocol, with the following changes: protein to 10× cleavage buffer volume ratio was 9:1, and protein:enzyme molar ratio was maintained at 50:1. The reaction mixture (204 $\mu$L) was incubated at room temperature for 4 h. Factor Xa and histidine tag capture were carried out according to the manufacturer's protocol, with the exception that cleaved protein was isolated from Ni-NTA beads using purification. After histidine tag capture, cleaved protein was dialyzed into 50 mM PB, pH 8, and concentration was determined via BCA.

*Gold Binding.* Molar concentration of protein:Au:sodium borohydride was kept constant at a ratio of 1:40:100. Au salt solution ($HAuCl_3{\cdot}3H_2O$) was added so that the final concentration was 400 $\mu$M ($dH_2O$) and the final concentration of the protein was 10 $\mu$M (in 50 mM PB, pH 8). For templation in the presence of TFE, 25 v/v % TFE was added to the protein solution for a final protein concentration of 10 $\mu$M, followed by Au incubation for 15 min and reduction. To reduce the metal, a final concentration of sodium borohydride (1 mM) was added.

*Absorbance.* Absorbance measurements were carried out using 10 $\mu$M protein (50 mM PB, pH 8) in the absence and presence of 400 $\mu$M Au at room temperature. Absorbance was measured using a SpectraMax Plus M2 instrument. Spectra were measured over a range of 400–700 nm.

*Circular Dichroism.* Circular dichroism (CD) measurements were conducted on a Jasco J-815 CD spectrometer, as described elsewhere.[31] Wavelength scans were conducted with 10 $\mu$M protein concentrations in 50 mM PB, pH 8, at room temperature. CD was also carried out in the presence of 25 v/v % TFE and 50 mM PB, pH 8, with 10 $\mu$M protein in the presence and absence of AuNPs. Mean residue ellipticity (MRE) was calculated from raw data according to the procedure described previously.[32]

*Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy.* Attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) experiments were carried out using a PerkinElmer System 2000 FT-IR with DuraSamplIR II T diamond ATR accessory and equipped with a MCT-A detector, as described previously.[31] Approximately 5 $\mu$L of protein solution (10 $\mu$M in 50 mM PB, pH 8) with or without AuNPs was added on the diamond ATR surface. ATR-FTIR was also carried out in the presence of 25 v/v % TFE and 50 mM PB, pH 8, with 10 $\mu$M protein in the presence and absence of AuNPs. PeakFit software was used to process the data, which involved a second-derivative zero-baseline correction of the amide I region between 1700–1600 $cm^{-1}$ and deconvolution of peaks with a Gaussian function.[33] All readings represent the average of two trials.

*Transmission Electron Microscopy.* A JEOL JEM-1400 TEM instrument was used to study the supramolecular protein structure. Approximately 3 $\mu$L of 10 $\mu$M protein in 50 mM PB, pH 8, with or without AuNPs was spotted on copper grids, as done before.[31] After 1 min, the grids were blotted using filter paper and rinsed with 3–4 drops Milli Q water to remove excess salts from the buffer. After blotting with filter paper, protein samples in the absence of Au were negatively stained by adding 3 $\mu$L of 1% filtered uranyl acetate and blotted using filter paper. Samples with AuNPs were not stained. Samples were then dried at room temperature for 10–15 min. TEM was also carried out in the presence of 25 v/v % TFE and 50 mM PB, pH 8, with 10 $\mu$M protein in the presence and absence of AuNPs. Prepared grids were viewed at an accelerating voltage of 120 kV. The same grids were further analyzed with an EDAX energy dispersive X-ray spectrometer to obtain elemental maps. ImageJ software was used to measure nanoparticle dimensions.[34]

*Cyclic Voltammetry and Impedance.* Electrochemical measurements were carried out with a three-electrode system comprising a platinum wire counter electrode, a Ag/AgCl reference electrode, and a glassy-carbon disk electrode ($d$ = 3.0 mm) as the working electrode. The polished disk electrode was exposed to 8 $\mu$M protein and protein·AuNP assemblies. Cyclic voltammetry (CV) is used to characterize surface coverage of a working electrode in relation to peak current, where a decrease in peak current corresponds to adsorption onto the electrode surface, blocking detection of the redox reaction of $Fe(CN)_6^{3-/4}$.[35] An 8 $\mu$M protein concentration was chosen for CV measurements because this was seen to provide peak resistance values in impedance measurements. Faradaic impedance spectra (or Nyquist plots) are generated for all protein·AuNP assemblies. Nyquist plots consist of a semicircular portion, corresponding to the electron-transfer-limited process, and a linear portion, corresponding to diffusion-limited electron transfer.[36] The diameter of the semicircular portion corresponds to the electron-transfer resistance, $R_{et}$, and is



DOI: 10.1021/acs.biomac.5b00098
Biomacromolecules XXXX, XXX, XXX–XXX

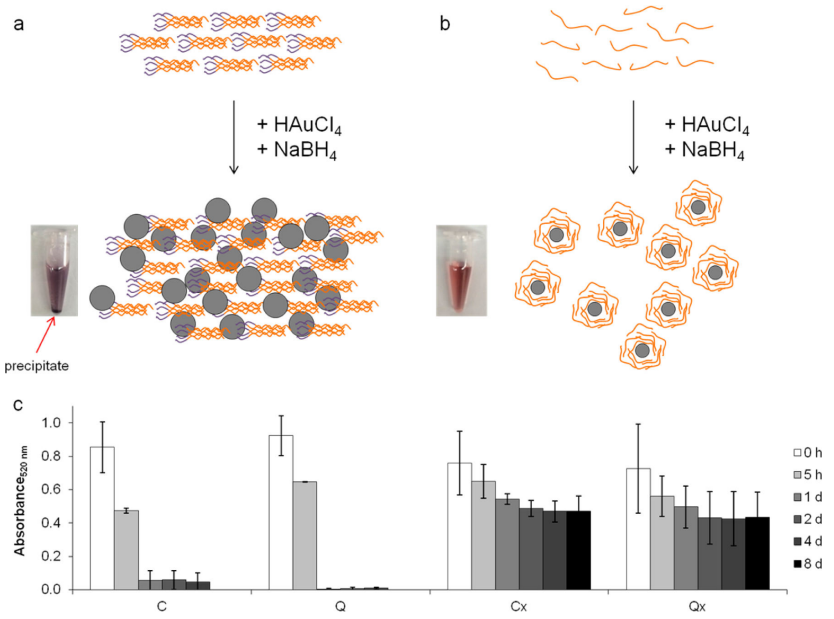



**Figure 1.** Schematic illustration of protein and nanoparticle assembly upon gold templation. Protein units are illustrated in orange, His tags in purple, and AuNPs as gray spheres. (a) $\alpha$-helical proteins C and Q with His tags form ordered pentamer assemblies, and upon gold templation exhibit a decrease in helical content and template AuNPs forming large aggregates of disordered AuNP·protein nanocomposite. Photo demonstrates the formation of precipitates in C·AuNP 24 h post templation. (b) Cleaved C and Q with His tags removed are unstructured and upon gold templation exhibit an increase in helical content, surrounding AuNPs and stabilizing them in solution. Photo demonstrates characteristic purple color of protein stabilized Cx·AuNP 24 h post-templation. (c) Normalized absorbance of protein·AuNP complexes at 520 nm. Error bars represent standard deviation of triplicate measurements.

helpful for characterizing electrochemical properties of surface-modified electrodes. CV data and Nyquist plots were recorded in 50 mM PB, pH 8, in the presence of 10 mM $K_4Fe(CN)_6{\cdot}3H_2O$ and 10 mM $K_3Fe(CN)_6$ immediately after protein composites were introduced to the electrode system. CV measurements were carried out at room temperature using a scan rate of 0.1 mV/s. Impedance measurements were recorded at room temperature at a bias potential of 210 mV within the frequency range of $10^{-1}$–$10^5$ Hz.

## RESULTS AND DISCUSSION

**Protein Design.** Two proteins, C and Q, were designed on the basis of the cartilage oligomeric matrix protein coiled-coil (COMPcc).[31,37] C protein possessed the same N-terminal sequence as COMPcc but truncated by one heptad repeat, whereas Q was domain-swapped and demonstrated fiber formation under acidic pH conditions (Table 1).[31] Both C and Q contained an N-terminal hexa-histidine tag (6-His) that could be eliminated to produce Cx and Qx, respectively (Table 1).[14] Successful cleavage of the 6-His tag from C and Q by Factor Xa was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE, Figure S1), with a molecular weight of the 6-His proteins corresponding to 6.31 kDa for C and Q versus 4.45 kDa for the corresponding cleaved proteins Cx and Qx (Table 1).

**Templation and Synthesis of the Protein·AuNP Nanocomposites.** All proteins were subjected to AuNP templation via addition of $HAuCl_4$ followed by reduction. Notably, all proteins were able to template gold in situ, and depending on the presence or absence of the 6-His tag, they displayed varied macromolecular assemblies. Remarkably, uncleaved C and Q were able to bind and template gold, forming large assemblies that precipitated out of solution, whereas the Cx and Qx proteins were able to solvate AuNPs (Figure 1a,b). Because spherical AuNPs exhibit peak absorbance corresponding to their surface plasmon resonance at 520 nm,[38] absorption of AuNPs templated in the presence of C and Q and their cleaved counterparts immediately after reduction with $NaBH_4$ as well as 5 h and 1, 2, 4, and 8 d thereafter were evaluated (Figure 1c). The absorbance of all four proteins was high immediately after AuNP reduction, with C and Q proteins illustrating slightly higher values compared to those of cleaved proteins at that time; this was indicated visually by a gray/purple color in the C and Q samples compared to a pinkish hue for Cx and Qx samples (Figure S2). A dramatic decline in the absorption of C and Q was seen after just 5 h, followed by a complete drop after 1 d, after which the value remained close to zero for the remainder of the period studied (Figure 1c). Cleaved proteins Cx and Qx, however, revealed a drop in absorption that stabilized after 1 day at approximately 70% its original value (Figure 1c). The formation of protein·AuNP precipitates in the case of C and Q resulted in a sharp



DOI: 10.1021/acs.biomac.5b00098
Biomacromolecules XXXX, XXX, XXX−XXX

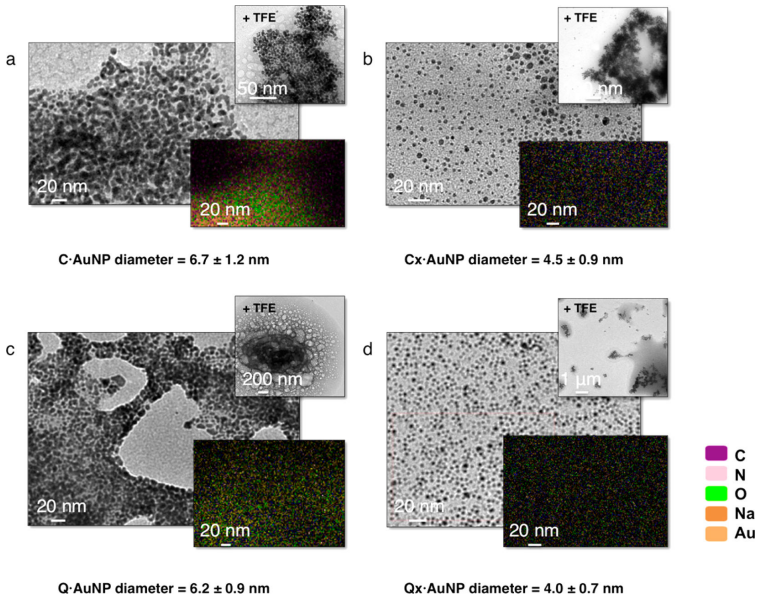



**Figure 2.** Transmission electron micrographs and elemental maps for C (a), Cx (b), Q (c), and Qx (d) obtained 8 days after reduction. Proteins in the presence of AuNP in buffer containing 25 v/v % TFE shown in top insets. Area analyzed in elemental maps and corresponding TEM image overlap. Maps display location of signals resulting from C (purple), N (pink), oxygen (lime green), Na (light brown), and Au (orange). Scale bars in TEM micrographs are 20 nm in (a), (b), (c), (d). TFE insets are 50 nm in (a), 100 nm in (b), 200 nm in (c) and 1 $\mu$m in (d). Scale bars on EDAX maps are 20 nm in (a), (b), (c), (d).

drop in absorbance at 520 nm within a few hours. This drop in absorbance could be attributed to the formation of large protein·AuNP precipitates, aggregating all AuNPs present. These quantitative measurements confirmed what was observed visually: C and Q formed large precipitates loaded with AuNPs, causing the bulk solution to lose color after a few hours (Figure S3). By contrast, no precipitation was observed when AuNPs were reduced in the presence of cleaved proteins, whose solutions maintained their pink hue over an extended period of time (Figure S3). Aggregates of AuNPs consisted of clustered, micrometer-sized assemblies of protein and AuNPs, like those seen in the presence of C and Q (Figure S4a,b).

AuNPs reduced in the absence of protein (PB alone) fell out of solution over time; however, they remained discrete particles rather than aggregates (Figures S2 and S4). Gold salt reduced with $NaBH_4$ in the absence of protein formed nanoparticles, as was observed via TEM (Figure S4e) and confirmed visually via a color change in solution (Figure S2e). Precipitation of AuNPs was observed in the absence of protein (Figure S2e), but these particles did not cluster together to form larger-scale assemblies, remaining as discrete particles that eventually precipitated to the bottom of sample containers as a result of gravity.

**Characterization of the Protein·AuNP Nanocomposites.** TEM was carried out to study the protein·AuNP assemblies on the nanometer scale. All proteins were visualized via TEM in conditions of 50 mM PB, pH 8, where morphology of C was mainly aggregates, Q formed high aspect ratio fibers,[31] and both Cx and Qx displayed sheetlike structures (Figure S5). In accordance with observations of absorbance studies, TEM confirmed that AuNPs were completely aggregated by C and Q, leaving few AuNPs outside of the protein matrix (Figures 2 and S3). There was no observable difference between the AuNP density on C and Q assemblies, indicating that the presence of the 6-His tag on both proteins enables them to aggregate gold from the solution and that the coiled-coil structures further assemble to form macromolecular films of AuNP-embedded proteins. Cleaved proteins Cx and Qx showed completely different organization of nanoparticles, with AuNPs scattered across the grid in a seemingly random dispersion. Upon careful inspection of the Cx·AuNP and Qx·AuNP micrographs (Figure S4), cleaved protein surrounded individual AuNPs, as represented by a dark shadow around the nanoparticles. AuNPs templated on C and Q were larger in diameter than those solvated by Cx and Qx (Figure S4). For C and Q assemblies, the AuNPs measured 6.7 ± 1.2 and 6.2 ± 0.9 nm, respectively, compared to 4.5 ± 0.9 nm for Cx and 4.0 ± 0.7 nm for Qx (Figure 2). AuNPs reduced in the absence of protein possessed diameters of 7.3 ± 1.0 nm. The smaller diameters observed in the presence of cleaved proteins resulting from crystallization of gold was capped because of the proteins surrounding the nanoparticles preventing them from further growth.[13,39] AuNPs templated with C and Q formed larger diameters that kinetically favored aggregation because increased AuNP surface area promoted interaction with neighboring 6-His tags, leading to aggregation of the protein·AuNP



DOI: 10.1021/acs.biomac.5b00098
*Biomacromolecules* XXXX, XXX, XXX−XXX

complex.[40] TEM results clearly demonstrated the distinct assembly morphologies of AuNPs templated in the presence of C and Q versus in the presence of Cx and Qx. C and Q templated large, micrometer-sized AuNP precipitates, whereas Cx- and Qx-templated AuNPs remained distinct and soluble.

Elemental mapping via energy dispersive X-ray spectroscopy (EDAX) revealed elements corresponding to protein (C, N, and O) colocalized with the AuNP in the case of the 6-His proteins (Figures 2 and S6), confirming that C and Q aggregated AuNPs completely from the solution. As a control, EDAX maps of protein in phosphate buffer solution in the absence of Au demonstrated high concentrations of carbon and oxygen corresponding to the proteins themselves and a lack of Au signal (Figure S7). X-ray signals of elements corresponding to protein were observed throughout the elemental maps of cleaved proteins, Cx and Qx, indicating that protein material surrounded the AuNPs (Figures 2 and S6). Cx and Qx were able to solvate the AuNPs, enabling them to maintain their solubility over time (Figure 1c). Higher count rates of gold X-ray signals from C and Q samples were observed compared with those from Cx and Qx, resulting from the high density of AuNPs in the aggregates formed by the uncleaved proteins (Figure S6). EDAX measurements confirmed that protein and gold were colocalized, as opposed to AuNPs aggregating or remaining as distinct particles independently and protein matter assembled apart from the AuNPs. This data supports the conclusion that AuNPs were embedded within the protein materials in the case of the 6-His proteins and that protein was surrounding the particles in the case of cleaved proteins.

Secondary structure analysis of C, Q, Cx, and Qx in the absence and presence of AuNP was carried out via CD and ATR-FTIR to better understand the structural changes that the uncleaved and cleaved proteins undergo upon AuNP templation (Figures S8–S10, Tables 2 and S5). Helical content, determined through peak deconvolution and fitting of ATR-FTIR data of C and Q (Table 2, Figures S9 and S10) prior to templation, was 54 and 65%, respectively. CD data of both C and Q also displayed double minima at 208 and 222 nm, indicative of helical structure (Figure S8). Upon AuNP templation, values of helical content according to FTIR analysis dropped to 30% for both proteins (Table 2, Figure S10). For cleaved proteins Cx and Qx, helical content prior to templation was 37 and 45%, respectively, and in the presence of AuNPs increased to 54 and 48%, respectively (Table 2, Figures S9 and S10). FTIR analysis of C, Q, Cx, and Qx demonstrated a loss in helical content upon cleavage of the 6-His tag. Cleavage of the 6-His tag indirectly affects the templation and aggregation of AuNPs through alteration of the secondary structure.

**Table 2. Secondary Structure Conformation Based on ATR-FTIR Data for 10 $\mu$M Protein in 50 mM PB before and after Templation of AuNPs and in the Presence of 25 v/v % TFE[a]**

| | conformation | | |
|---|---|---|---|
| protein | $\alpha$ helix | $\beta$ sheet | random coil |
| C | 53.5 ± 6.5 | 46.5 ± 6.5 | |
| C·AuNP | 30.5 ± 0.8 | 53.5 ± 0.6 | 16.0 ± 1.3 |
| C + TFE | 68.1 ± 1.3 | 31.9 ± 1.3 | |
| C·AuNP + TFE | 31.7 ± 0.9 | 40.9 ± 2.1 | 27.4 ± 1.7 |
| Q | 65.0 ± 1.4 | 35.0 ± 1.4 | |
| Q·AuNP | 29.9 ± 0.2 | 42.9 ± 0.1 | 27.2 ± 0.1 |
| Q + TFE | 70.7 ± 2.7 | 29.3 ± 2.7 | |
| Q·AuNP + TFE | 36.8 ± 0.8 | 36.3 ± 4.1 | 26.9 ± 1.9 |
| Cx | 36.5 ± 1.4 | 36.2 ± 1.2 | 27.3 ± 2.6 |
| Cx·AuNP | 54.0 ± 8.9 | 30.3 ± 4.5 | 15.7 ± 4.4 |
| Cx + TFE | 49.7 ± 0.5 | 29.0 ± 3.8 | 21.3 ± 3.3 |
| Qx | 44.6 ± 1.8 | 26.0 ± 0.5 | 29.4 ± 1.3 |
| Qx·AuNP | 48.2 ± 1.3 | 32.1 ± 0.9 | 19.6 ± 0.4 |
| Qx + TFE | 51.2 ± 0.3 | 25.0 ± 2.3 | 23.8 ± 2.1 |

[a]Percent composition was determined from relative areas of peaks fit to spectra (Figure S9–S13). Values represent the average and standard error from duplicate measurements. Data was not obtained for Cx·AuNP and Qx·AuNP with TFE because of problems with sample measurement.

Because C and Q exhibited a loss in helical content upon cleavage, we investigated AuNP templation on cleaved proteins after the addition of 25 v/v % trifluoroethanol (TFE). TFE is known to stabilize $\alpha$-helical structure in proteins, the propensity for which is not explicitly correlated to the helical content of the native protein.[41] TEM studies on proteins in the presence and absence of TFE served as a control to understand the protein morphology in the presence and absence of TFE (Figures S2 and S11). Protein structural changes are not easily observed in TEM in the absence and presence of TFE; however, data on secondary structure revealed that an increase in helical content was observed for all proteins studied in the presence of TFE (Table 2, Figures S12 and S13). Upon addition of TFE, helical content increased to 68 and 71% for C and Q, respectively, and 50 and 51% for Cx and Qx, respectively. Remarkably, this increase in structure of the cleaved proteins Cx and Qx enabled further assembly from aggregates as opposed to the soluble AuNPs generated in the absence of TFE (Table 2, Figure S11). This supports work by Higuchi et al., who demonstrated that peptide·AuNP assemblies could be controlled by modifying peptide secondary structure.[22] Thus, an external trigger such as TFE led to conformational changes that impacted the macromolecular assembly of protein·AuNP complexes. Our studies demonstrate that TFE-induced conformational changes of protein can influence the macromolecular assembly of protein·AuNP complexes, reinforcing the importance of secondary structure.

The presence of the 6-His tags in C and Q directly affected assembly because the AuNPs templated with C and Q were larger in diameter; moreover, the assemblies are demonstrated thorough aggregation. Aggregation of the protein·AuNP complexes with 6-His tags was kinetically favored because larger AuNP surface areas promoted interaction between 6-His tags.[40] An unexpected consequence of the 6-His tag cleavage was loss of secondary structure, which also indirectly affected aggregation because higher-order assembly of these proteins was precluded, preventing aggregation of the protein·AuNP complexes. The mechanism of protein aggregation could be explained by adsorption of proteins on the surface of AuNPs.[20] Surface protein exhibiting ordered secondary structure, in this case strong $\alpha$-helicity of C and Q with 6-His tags present, induced aggregation, as was directed by further assembly of the structured proteins. Adsorbed protein possessing random structure, such as Cx and Qx in the absence of the 6-His tag, precluded aggregation of AuNPs.

**Electrochemical Characterization.** To evaluate the resistivity, and hence conductivity, of protein·AuNP assemblies, electrochemical impedance spectroscopy (EIS) and CV measurements were been carried out in the presence of the $Fe(CN)_6^{3-/4-}$ redox pair on a glassy-carbon electrode.[35] Uncleaved C·AuNP and Q·AuNP films spontaneously adsorb



DOI: 10.1021/acs.biomac.5b00098
Biomacromolecules XXXX, XXX, XXX−XXX

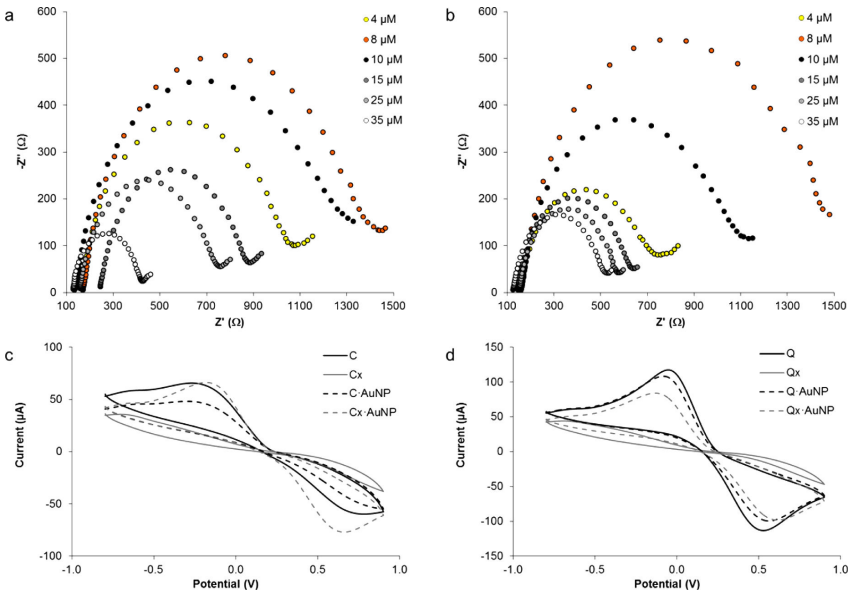



**Figure 3.** Nyquist plots for varying concentrations (4, 8, 10, 15, 25, and 35 $\mu$M) of proteins (a) C and (b) Q in the presence of AuNPs. Cyclic voltammograms of (c) 8 $\mu$M C, Cx, C·AuNP, and Cx·AuNP and (d) 8 $\mu$M Q, Qx, Q·AuNP, and Qx·AuNP. All electrochemical measurements were carried out in the presence of 10 mM $[Fe(CN)_6]^{3-}$ and 10 mM $[Fe(CN)_6]^{4-}$ in 50 mM PB, pH 8, on glassy-carbon electrodes.

onto the electrode surface, whereas AuNPs templated with cleaved protein remain soluble. The electron-transfer resistance ($R_{et}$) values increase until reaching a peak at 8 $\mu$M, with the largest resistance values attributed to C·AuNP and Q·AuNP at 812.0 and 902.1 Ω, respectively (Figure 3, Table S6). CV demonstrates intense anodic and cathodic peaks in the absence of protein (Figure S14), which diminish in intensity upon protein deposition (Figure 3). In the absence of AuNPs, proteins exhibit a distinct level of adsorption onto the electrode surface, with cleaved proteins having a significantly lower peak current than those of the uncleaved counterparts (Figure 3). Proteins C and Q show a decrease in peak current equivalent to 17.4 and 9.0 $\mu$A, respectively (Figure 3). Cleaved proteins Cx and Qx exhibit an increase in peak current, equivalent to 30.3and 39.9 $\mu$A, respectively (Figure 3). Overall, uncleaved proteins exhibit a decrease in peak current upon the templation of AuNPs, whereas cleaved proteins possess an increase in peak current with templated AuNPs.

Changes in protein conformation on a macromolecular level will result in fluctuations of interfacial charge, resistance, capacitance, and protein layer thickness on the electrode surface.[44] Variations in these properties affect the electron-transfer resistance at the protein−electrode interface, which can be measured via CV and EIS. Because colloidal gold multilayers undergo an insulator−conductor transition at critical particle density,[45] the decrease in $R_{et}$ above 8 $\mu$M indicates that surface coverage on the electrode is less than one monolayer in thickness below this concentration. Above 10 $\mu$M, the uncleaved protein·AuNP films display $R_{et}$ values that stabilize in the 200−400 Ω range (Table S6). Similar $R_{et}$ values have been reported for the functionalization of electrode surfaces with DNA-stabilized AuNPs.[7] The same trend of increasing and decreasing $R_{et}$ as concentration increases has been observed with chitosan- and citrate-stabilized nanoparticles as well.[46] Cleaved and uncleaved proteins adsorbed on the electrode surface exhibit difference peak currents in CV measurements. Cleaved proteins display a much lower peak current than do uncleaved proteins (Figure 3). This can be explained by the fact that the absorption of a particular material onto the electrode surface is a complex function of hydrophobicity, electric charges of both the electrode and the protein, and protein conformation.[35] As can be seen in Figure 3, a decrease in peak current is observed upon the templation of AuNPs in the case of uncleaved proteins (decrease of 17.4 $\mu$A for C and 9.0 $\mu$A for Q). Cleaved proteins, on the other hand, demonstrate an increase in peak current upon templation of AuNPs. This data confirms that proteins bearing 6-His tags template AuNPs, generating dense films that adsorb onto the electrode surface and decreasing the electron transfer in the redox reaction of $Fe(CN)_6^{3-/4-}$ ions at the electrode surface. In the case of cleaved proteins, protein-solvated AuNPs in the bulk volume increase the overall electron transfer of the iron ions at the electrode surface. These measurements indicate that protein assemblies of structured proteins can be used to form a tunable kinetic barrier on the electrode surface.

## CONCLUSIONS

Previous examples of protein·AuNP nanocomposites with the ability to switch assembly conformation have been constructed by the conjugation of protein to prefabricated AuNPs.[21,23−27]



DOI: 10.1021/acs.biomac.5b00098
Biomacromolecules XXXX, XXX, XXX−XXX

We demonstrate that protein·AuNP nanocomposites can be fabricated in situ, leading to very distinct assembly morphology that is dependent upon the protein secondary structure. Templation by unstructured peptides frequently results in soluble AuNPs that do not aggregate, whereas templation by structured protein allows for self-assembly, causing protein·AuNP nanocomposites to aggregate.[15,22] By manipulating protein structure via external triggers, such as TFE, we obtain control over the macromolecular conformation of these nanocomposites. Such external triggers can lead to switching from unstructured to structured assemblies. These hybrid materials can be readily deposited on electrodes, where they can serve as a tunable bionanocomposite kinetic barrier.[47] Proteins that stabilize nanoparticles may prove important as agents for biodetection and imaging and as biocompatible stabilizing agents for nanoparticles in aqueous solutions, whereas formation of nanoparticle-embedded protein films may be of interest in catalysis and biosensing applications.[48,49]

## ■ ASSOCIATED CONTENT

### Ⓢ Supporting Information

Details on genetic engineering employed to obtain the C construct, galleries of TEM and CD data collected, ATR-FTIR spectra, and additional Nyquist plots and cyclic voltammograms. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: montclare@nyu.edu.


### Author Contributions

The manuscript was written through contributions of all authors. J.H. designed and carried out experiments, analyzed data, and wrote the paper; R.C., R.J., and M.Y. carried out experiments and analyzed data; and J.K.M. supervised the project and wrote the paper. All authors have given approval to the final version of the manuscript.


### Funding

This work was supported by the ARO (W911NF-11−1−0449), GK-12 Fellows grant DGE-0741714, and in part by the NSF MRSEC Program under award number DMR-0820341.


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS


The authors thank Kim Kisslinger and Qi Zhang for technical assistance and critical insight into the experiments and manuscript. Research carried out in part at the Center for Functional Nanomaterials, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under contract no. DE-AC02-98CH10886.


## ■ ABBREVIATIONS

CD, circular dichroism; TEM, transmission electron microscopy; NP, nanoparticle; AuNP, gold nanoparticle; TFE, trifluoroethanol; CV, cyclic voltammetry; ATR-FTIR, Attenuated total reflectance Fourier transform infrared spectroscopy; IPTG, isopropyl-$\beta$-D-thiogalactopyranoside; BCA, bicinchoninic acid; COMPcc, cartilage oligomeric matrix protein coiled-coil; PCR, polymerase chain reaction

## ■ REFERENCES


(1) Chen, M. S.; Goodman, D. W. *Science* **2004**, *306*, 252−255.
(2) Huang, X.; El-Sayed, M. A. *J. Adv. Res.* **2010**, *1*, 13−28.
(3) Daniel, M.; Astruc, D. *Chem. Rev.* **2004**, *104*, 293−346.
(4) Krasteva, N.; Besnard, I.; Guse, B.; Bauer, R. E.; Mu, K.; Yasuda, A.; Vossmeyer, T. *Nano Lett.* **2002**, *2*, 551−555.
(5) Guo, S.; Wang, E. *Nano Today* **2011**, *6*, 240−264.
(6) Saha, K.; Agasti, S. S.; Kim, C.; Li, X.; Rotello, V. M. *Chem. Rev.* **2012**, *112*, 2739−2779.
(7) Lu, L.-P.; Wang, S.-Q.; Lin, X.-Q. *Anal. Chim. Acta* **2004**, *519*, 161−166.
(8) Jiang, C.; Markutsya, S.; Pikus, Y.; Tsukruk, V. V. *Nat. Mater.* **2004**, *3*, 721−728.
(9) Slocik, J. M.; Kim, S. N.; Whitehead, T. A.; Clark, D. S.; Naik, R. R. *Small* **2009**, *5*, 2038−2042.
(10) Slocik, J. M.; Moore, J. T.; Wright, D. W. *Nano Lett.* **2002**, *2*, 169−173.
(11) Chen, C. L.; Rosi, N. L. *J. Am. Chem. Soc.* **2010**, *132*, 6902−6903.
(12) Singh, A.; Hede, S.; Sastry, M. *Small* **2007**, *3*, 466−473.
(13) Higashi, N.; Kawahara, J.; Niwa, M. *J. Colloid Interface Sci.* **2005**, *288*, 83−87.
(14) Hom, N.; Mehta, K. R.; Chou, T.; Foraker, A. B.; Brodsky, F. M.; Kirshenbaum, K.; Montclare, J. K. *J. Mater. Chem.* **2012**, *22*, 23335−23339.
(15) Deka, J.; Paul, A.; Chattopadhyay, A. *Nanoscale* **2010**, *2*, 1405−1412.
(16) Jung, S. H.; Jeon, J.; Kim, H.; Jaworski, J.; Jung, J. H. *J. Am. Chem. Soc.* **2014**, *136*, 6446−6452.
(17) Kharlampieva, E.; Zimnitsky, D.; Gupta, M.; Bergman, K. N.; Kaplan, D. L.; Naik, R. R.; Tsukruk, V. V. *Chem. Mater.* **2009**, *21*, 2696−2704.
(18) Chen, C.-L.; Zhang, P.; Rosi, N. L. *J. Am. Chem. Soc.* **2008**, *130*, 13555−13557.
(19) Nishikawa, H.; Morita, T.; Sugiyama, J.; Kimura, S. *J. Colloid Interface Sci.* **2004**, *280*, 506−510.
(20) Slocik, J. M.; Govorov, A. O.; Naik, R. R. *Nano Lett.* **2011**, *11*, 701−705.
(21) Stevens, M. M.; Flynn, N. T.; Wang, C.; Tirrell, D. A.; Langer, R. *Adv. Mater.* **2004**, *16*, 915−918.
(22) Higuchi, M.; Ushiba, K.; Kawaguchi, M. *J. Colloid Interface Sci.* **2007**, *308*, 356−363.
(23) Bayraktar, H.; Srivastava, S.; You, C.-C.; Rotello, V. M.; Knapp, M. J. *Soft Matter* **2008**, *4*, 751−756.
(24) Tullman, J. A.; Finney, W. F.; Lin, Y.-J.; Bishnoi, S. W. *Plasmonics* **2007**, *2*, 119−127.
(25) Minelli, C.; Liew, J. X.; Muthu, M.; Andresen, H. *Soft Matter* **2013**, *9*, 5119−5124.
(26) Guo, Y.; Ma, Y.; Xu, L.; Li, J.; Yang, W. *J. Phys. Chem. C* **2007**, *111*, 9172−9176.
(27) Slocik, J. M.; Tam, F.; Halas, N. J.; Naik, R. R. *Nano Lett.* **2007**, *7*, 1054−1058.
(28) Link, S.; Burda, C.; Nikoobakht, B.; El-Sayed, M. A. *J. Phys. Chem. B* **2000**, *104*, 6152−6163.
(29) Gunasekar, S. K.; Anija, L.; Matsui, H.; Montclare, J. K.; Anjia, L. *Adv. Funct. Mater.* **2012**, *22*, 2154−2159.
(30) Aili, D.; Enander, K.; Rydberg, J.; Nesterenko, I.; Björefors, F.; Baltzer, L.; Liedberg, B. *J. Am. Chem. Soc.* **2008**, *130*, 5780−5788.
(31) Hume, J.; Sun, J.; Jacquet, R.; Renfrew, P. D.; Martin, J. A.; Bonneau, R.; Gilchrist, M. L.; Montclare, J. K. *Biomacromolecules* **2014**, *15*, 3503−3510.
(32) Gunasekar, S. K.; Asnani, M.; Limbad, C.; Haghpanah, J. S.; Hom, W.; Barra, H.; Nanda, S.; Lu, M.; Montclare, J. K. *Biochemistry* **2009**, *48*, 8559−8567.
(33) Hu, X.; Kaplan, D.; Cebe, P. *Macromolecules* **2006**, *39*, 6161−6170.
(34) Abràmoff, M. D.; Magalhães, P. J. *Biophotonics Int.* **2004**, *11*, 36−42.




DOI: 10.1021/acs.biomac.5b00098
*Biomacromolecules* XXXX, XXX, XXX−XXX

559 (35) Guo, B.; Anzai, J.; Osa, T. *Chem. Pharm. Bull.* **1996**, *44*, 800−
560 803.
561 (36) Ding, X.; Yang, M.; Hu, J.; Li, Q.; McDougall, A. *Microchim.*
562 *Acta* **2006**, *158*, 65−71.
563 (37) Gunasekar, S. K.; Haghpanah, J. S.; Montclare, J. K. *Polym. Adv.*
564 *Technol.* **2008**, *19*, 454−468.
565 (38) Eustis, S.; El-Sayed, M. A. *Chem. Soc. Rev.* **2006**, *35*, 209−217.
566 (39) Shimmin, R. G.; Schoch, A. B.; Braun, P. V. *Langmuir* **2004**, *20*,
567 5613−5620.
568 (40) Ojea-Jiménez, I.; Puntes, V. *J. Am. Chem. Soc.* **2009**, *131*,
569 13320−13327.
570 (41) Shiraki, K.; Nishikawa, K.; Goto, Y. *J. Mol. Biol.* **1995**, *245*, 180−
571 194.
572 (42) Vedantham, G.; Sparks, H. G.; Sane, S. U.; Tzannis, S.;
573 Przybycien, T. M. *Anal. Biochem.* **2000**, *285*, 33−49.
574 (43) Byler, D. M.; Susi, H. *Biopolymers* **1986**, *25*, 469−487.
575 (44) Liu, X.; Duckworth, P. A.; Wong, D. K. Y. *Biosens. Bioelectron.*
576 **2010**, *25*, 1467−1473.
577 (45) Shipway, A. N.; Katz, E.; Willner, I. *ChemPhysChem* **2000**, *1*,
578 18−52.
579 (46) Huang, H.; Yang, X. *Colloids Surf., A* **2003**, *226*, 77−86.
580 (47) Huang, H.; Liu, Z.; Yang, X. *Anal. Biochem.* **2006**, *356*, 208−
581 214.
582 (48) Zhang, L.; Jiang, X.; Wang, E.; Dong, S. *Biosens. Bioelectron.*
583 **2005**, *21*, 337−345.
584 (49) Slocik, J. M.; Stone, M. O.; Naik, R. R. *Small* **2005**, *1*, 1048−
585 1052.



DOI: 10.1021/acs.biomac.5b00098
*Biomacromolecules* XXXX, XXX, XXX−XXX

## 5.2 Abbreviations

Table A 1. Abbreviations and their explanations.

| **Acronym** | **Explanation** |
|---|---|
| AFM | Atomic force microscopy |
| AHA | azidohomoalanine |
| AuNP | Gold nanoparticle |
| BSA | Bovine serum albumin |
| CD | Circular dichroism |
| CMP | Collagen mimetic peptide |
| COMP | Cartilage oligomeric matrix protein |
| CV | Cyclic voltammetry |
| DLS | Dynamic light scattering |
| EDAX | Energy dispersive X-ray analysis |
| MNP | Magnetic nanoparticle |
| MQ | MilliQ water, or $dH_2O$ |
| MRE | Mean residual ellipticity |
| NP | Nanoparticle |
| PA | Peptide-amphiphiles |
| SEM | Scanning electron microscopy |
| TEM | Transmission electron microscopy |
| TFE | Trifluoroethanol |
| $T_m$ | Melting temperature |
| UV | Ultra violet |
| Wt | Wild type |

## 5.3 Source of materials

Table A 2. Reagents and suppliers.

| **Reagent** | **Supplier** |
|---|---|
| Ampicillin | Fisher Scientific |
| BCA kit | Pierce |
| Calcium chloride | EM Science |
| Chromeo494 | Active Motif |
| CMms6 (N-terminal L-propargylglycine) | LifeTein |
| Copper grids | Ted Pella |
| Copper sulfate | Amresco |
| Copper wire | Arcor Electronics |
| Curcumin | Acros Organics |
| D-glucose monohydrate | Alfa Aesar |
| Dialysis tubing (3.5 kDa MWCO) | Thermo Scientific |
| Factor Xa cleavage kit | Novagen |
| Glassy carbon electrodes | CH Instruments, Inc. |
| Glycine | Amresco |
| $HAuCl_3 \cdot 3H_2O$ | Sigma Aldrich |
| Imidazole | Alfa Aesar |
| IPTG | Fisher Scientific |
| Iron (II) sulfate heptahydrate | Sigma Aldrich |
| Iron (III) chloride | Sigma Aldrich |
| $K_2HPO_4$ (Dipotassium phosphate dibasic) | Fisher Scientific |
| Kanamycin sulfate | Amresco |
| L-alanine | Fisher Scientific |
| L-arginine | Fisher Scientific |
| L-asparagine monohydrate | Fisher Scientific |
| L-aspartic acid | Fisher Scientific |
| L-azidohomoalanine | Click Chemistry Tools |
| L-cysteine hydrochloride monohydrate | Fisher Scientific |
| L-glutamic acid | Fisher Scientific |

| | |
|---|---|
| L-glutamine | Sigma Aldrich |
| L-histidine | Amresco |
| L-isoleucine | Fisher Scientific |
| L-leucine | Fisher Scientific |
| L-lysine hydrochloride | Fisher Scientific |
| L-methionine | Amresco |
| L-phenylalanine | Fisher Scientific |
| L-proline | Sigma Aldrich |
| L-serine | Fisher Scientific |
| L-threonine | Fisher Scientific |
| L-tryptophan | Fisher Scientific |
| L-tyrosine | Fisher Scientific |
| L-valine | Fisher Scientific |
| Magnesium sulfate | Sigma Aldrich |
| Methanol | Acros Organics |
| Microscopy slides | Clay Adams |
| $Na_2HPO_4$ (sodium phosphate dibasic anhydrous) | Amresco |
| Sodium hydroxide | BDH Chemicals |
| $NH_4Cl$ | Fisher Scientific |
| Ni-NTA beads | Sigma Aldrich |
| Potassium hexacyanoferrate (II) trihydrate | Sigma Aldrich |
| Potassium hexacyanoferrate (III) | Sigma Aldrich |
| Sodium ascorbate | Acros Organics |
| Sodium borohydride | Sigma Aldrich |
| Sodium chloride | Fisher Scientific |
| Sodium phosphate dibasic | Sigma Aldrich |
| Sodium phosphate monobasic | Sigma Aldrich |
| Trifluoroethanol | Sigma Aldrich |
| Tris-HCl | Fisher Scientific |
| Tryptone | Fisher Scientific |
| Uranyl acetate | Electron Microscopy Science |
| Urea | Fisher Scientific |

| | |
|---|---|
| Vitamin B (Thiamine hydrochloride) | Sigma Aldrich |
| Yeast extract | Acros Organics |

## 5.4 Plasmid maps and DNA sequences

### *5.4.1 COMPcc$^{s}$*

Plasmid Name: pQE9-COMPss*

Vector size: 3.5 kBP

Cloning Site: BAMHI and HINDIII

Construction of pQE9-COMPss vector: The gene coding for COMPss was cloned into the *Bam*HI and *Hind*III sites to generate pQE9/COMPss.

Source available:

50 % Glycerol culture. Liquid Nitrogen GSK box. Culture conditions: Medium: LB/Amp Temp. 37 °C

50 % Glycerol culture. -80 °C GSK box. Culture conditions: Medium: LB/Amp. Temp: 37 °C

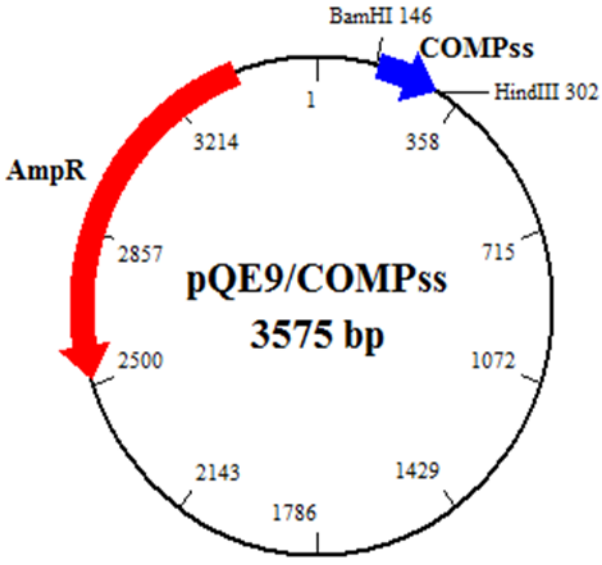



Figure A 1. Plasmid map for COMPcc$^{s}$.

* http://www1.qiagen.com/literature/pqesequences/pqe3x.pdf

## Positions of elements in bases

| | |
|---|---|
| Vector size | 3575 |
| T5 promoter/lac operator | 7-87 |
| T5 transcription start | 61 |
| 6X His-tag coding sequence | 127-144 |
| COMPss gene | 151-300 |
| β-lactamase (AmpR) coding sequence | 2513-3370 |

Full sequence

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGAT
TCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACT
ATGAGAGGATCGCATCACCATCACCATCACGGATCCGGTGACCTGGCGCCGCAGATG
CTGCGTGAACTGCAGGAAACCAACGCGGCGCTGCAGGACGTTCGTGAACTGCTGCGT
CAGCAGGTTAAAGAAATCACCTTCCTGAAAAACACCGTTATGGAATCTGACGCGTCT
GGTAAGCTTAATTAGAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAG
TAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTT
TTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTA
AAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTA
AAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTC
AGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATC
CGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGC
AATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTT
CCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCG
GCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTA
TTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGT
TTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCA
CCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGG
TTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACA
GTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCC
TTAAACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGCTC
AGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAG
TAGGACAAATCCGCCCTCTAGAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAAC
CTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGG
AGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGC
CATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCA
GAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTA
AGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGC
TCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTA
TCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA
AGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGA
CCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGC
TGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTC
TTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA

GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA
ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA
CCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG
GTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCA
ATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCC
AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATA
AACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCC
ATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT
TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA
TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGC
CGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA
TCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT
GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC
AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACA
TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAA
AAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATA
GGCGTATCACGAGGCCCTTTCGTCTTCAC

*5.4.2 C*

Plasmid Name: p**-C

Vector size: 3000 bp

Cloning Site: BamHI and HindIII

Construction of pQE30-p**-C vector: The original construct was obtained from Natalya Voloshchuk and given to Ching Yao Yang. This vector is used for incorporation of unnatural amino acid by the mutation of PheRS. Ching Yao passed the vector containing the p** mutation to me which I then restricted to release his gene and ligated mine to.

Source available:

50 % glycerol culture: - 80 $^{o}$C JH box

Culture conditions: Medium: LB/Amp Temp. 37 $^{o}$C

Plasmid DNA : JH - 20 $^{o}$C DNA box

Plasmid DNA : HTM - 20 $^{o}$C DNA box

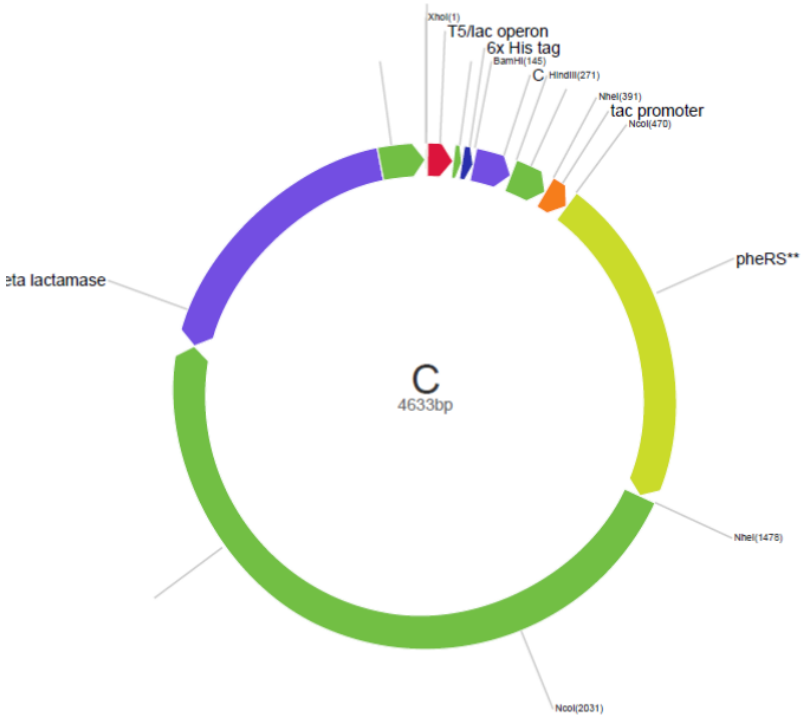



Figure A 2. Plasmid map for C.

Table A 3. Position of restriction sites in pQE30-p**C plasmid.

| Restriction site | Sequence | Position |
|---|---|---|
| XhoI | CTCGAG | 145 |
| NheI | GCTAGC | 271 |
| NcoI | CCATGG | 1 |
| BamHI | GGATCC | 291, 1478 |
| HindIII | AAGCTT | 470, 2031 |

Table A 4. Positions of DNA motifs in pQE30-p**C plasmid.

| Motif | Position |
|---|---|
| Start of numbering at XhoI (CTCGAG) | 1 |
| T5/lac operon | 7 |
| Gene (C) | 151 |
| tac promoter | 392 |

| | |
|---|---|
| pheRS** (alpha, T251G/A294G) | 472 |
| Beta lactamase gene | 3633 |

**Full Sequence**

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAAT
AGATTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAG
AAATTAACTATGAGAGGATCGCATCACCATCACCATCACGGATCCATCGAAG
GTCGCGCGCCGCAGATGCTGCGTGAACTGCAGGAAACCAACGCGGCGCTGCA
GGACGTTCGTGAACTGCTGCGTCAGCAGGTTAAAGAAATCACCTTCCTGAAA
AACACCTCAAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTA
ATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGC
GTTTTTTATTGGTGAGAATCCAAGCTAGCAGTTGACAATTAATCATCGGCTCG
TATAATGGATCGAATTGTGAGCGGAATCGATTTTCACACAGGAAACAGACCA
TGGCACATCTCGCAGAACTGGTTGCCAGTGCGAAGGCGGCCATTAGCCAGGC
GTCAGATGTTGCCGCGTTAGATAATGTGCGCGTCGAATATTTGGGTAAAAAA
GGGCACTTAACCCTTCAGATGACGACCCTGCGTGAGCTGCCGCCAGAAGAGC
GTCCGGCAGCTGGTGCGGTTATCAACGAAGCGAAAGAGCAGGTTCAGCAGGC
GCTGAATGCGCGTAAAGCGGAACTGGAAAGCGCTGCACTGAATGCGCGTCTG
GCGGCGGAAACGATTGATGTCTCTCTGCCAGGTCGTCGCATTGAAAACGGCG
GTCTGCATCCGGTTACCCGTACCATCGACCGTATCGAAAGTTTCTTCGGTGAG
CTTGGCTTTACCGTGGCAACCGGGCCGGAAATCGAAGACGATTATCATAACT
TCGATGCTCTGAACATTCCTGGTCACCACCCGGCGCGCGCTGACCACGACAC
TTTCTGGTTTGACACTACCCGCCTGCTGCGTACCCAGACCTCTGGCGTACAGA
TCCGCACCATGAAAGCCCAGCAGCCACCGATTCGTATCATCGCGCCTGGCCG
TGTTTATCGTAACGACTACGACCAGACTCACACGCCGATGTTCCATCAGATGG
AAGGTCTGATTGTTGATACCAACATCAGCTTTACCAACCTGAAAGGCACGCT
GCACGACTTCCTGCGTAACTTCTTTGAGGAAGATTTGCAGATTCGCTTCCGTC
CTTCCTACTTCCCGTTTGCCGAACCTTCTGCAGAAGTGGACGTCATGGGTAAA
AACGGTAAATGGCTGGAAGTGCTGGGCTGCGGGATGGTGCATCCGAACGTGT
TGCGTAACGTTGGCATCGACCCGGAAGTTTACTCTGGTTTCGGCTTCGGGATG
GGGATGGAGCGTCTGACTATGTTGCGTTACGGCGTCACCGACCTGCGTTCATT
CTTCGAAAACGATCTGCGTTTCCTCAAACAGTTTAAATAAGGCAGGAATAGA
TTATAATCTTGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATG
GAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTA
AAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACC
GTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACA
AGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGG
AATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCA
CCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGA
GTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTG
GCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATAT
GTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGT
GGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATA

CGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT
CTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCG
ATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAA
ACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGCGTTTCGGTGAT
GACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGT
GTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACC
ATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGC
GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG
CGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TCATCCATAGCTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGG
GCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC
GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGG
AAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATT
GCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTC
CGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAA
GCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGT
GTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCAT
CCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAA
TAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA
CCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTC
GGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA
CCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCT
GGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGC
GACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCA

TTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAA
AATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACG
TCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCAC
GAGGCCCTTTCGTCTTCAC

### *5.4.3 Q*

Plasmid Name: p**-Q

Vector size: 3000 bp

Cloning Site: BamHI and HindIII

Construction of pQE30-p**-Q vector: The original construct was obtained from Natalya Voloshchuk and given to Ching Yao Yang. This vector is used for incorporation of unnatural amino acid by the mutation of PheRS. Ching Yao passed the vector containing the p** mutation to me which I then restricted to release his gene and ligated mine to.

Source available:

50 % glycerol culture: -80 °C JH box

Culture conditions: Medium: LB/Amp Temp. 37 $^{o}$C

Plasmid DNA : JH - 20 °C DNA box

Plasmid DNA : HTM - 20 °C DNA box




Figure A 3. Plasmid map for Q.

Table A 5. Position of restriction sites in pQE30-p**Q plasmid.

| **Restriction site** | **Sequence** | **Position** |
|---|---|---|
| XhoI | CTCGAG | 145 |
| NheI | GCTAGC | 271 |
| NcoI | CCATGG | 1 |
| BamHI | GGATCC | 291, 1478 |
| HindIII | AAGCTT | 470, 2031 |

Table A 6. Positions of DNA motifs in pQE30-p**Q plasmid.

| **Motif** | **Position** |
|---|---|
| Start of numbering at XhoI (CTCGAG) | 1 |
| T5/lac operon | 7 |
| Gene (Q) | 151 |
| tac promoter | 392 |
| pheRS** (alpha, T251G/A294G) | 472 |
| Beta lactamase gene | 3633 |

## Full Sequence

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAAT
AGATTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAG
AAATTAACTATGAGAGGATCGCATCACCATCACCATCACGGATCCATCGAAG
GTCGCGTTAAAGAAATCACCTTCCTGAAAAACACCGCGCCGCAGATGCTGCG
TGAACTGCAGGAAACCAACGCGGCGCTGCAGGACGTTCGTGAACTGCTGCGT
CAGCAGTCAAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTA
ATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGC
GTTTTTTATTGGTGAGAATCCAAGCTAGCAGTTGACAATTAATCATCGGCTCG
TATAATGGATCGAATTGTGAGCGGAATCGATTTTCACACAGGAAACAGACCA
TGGCACATCTCGCAGAACTGGTTGCCAGTGCGAAGGCGGCCATTAGCCAGGC
GTCAGATGTTGCCGCGTTAGATAATGTGCGCGTCGAATATTTGGGTAAAAAA
GGGCACTTAACCCTTCAGATGACGACCCTGCGTGAGCTGCCGCCAGAAGAGC
GTCCGGCAGCTGGTGCGGTTATCAACGAAGCGAAAGAGCAGGTTCAGCAGGC
GCTGAATGCGCGTAAAGCGGAACTGGAAAGCGCTGCACTGAATGCGCGTCTG
GCGGCGGAAACGATTGATGTCTCTCTGCCAGGTCGTCGCATTGAAAACGGCG
GTCTGCATCCGGTTACCCGTACCATCGACCGTATCGAAAGTTTCTTCGGTGAG
CTTGGCTTTACCGTGGCAACCGGGCCGGAAATCGAAGACGATTATCATAACT
TCGATGCTCTGAACATTCCTGGTCACCACCCGGCGCGCGCTGACCACGACAC
TTTCTGGTTTGACACTACCCGCCTGCTGCGTACCCAGACCTCTGGCGTACAGA
TCCGCACCATGAAAGCCCAGCAGCCACCGATTCGTATCATCGCGCCTGGCCG
TGTTTATCGTAACGACTACGACCAGACTCACACGCCGATGTTCCATCAGATGG
AAGGTCTGATTGTTGATACCAACATCAGCTTTACCAACCTGAAAGGCACGCT
GCACGACTTCCTGCGTAACTTCTTTGAGGAAGATTTGCAGATTCGCTTCCGTC
CTTCCTACTTCCCGTTTGCCGAACCTTCTGCAGAAGTGGACGTCATGGGTAAA
AACGGTAAATGGCTGGAAGTGCTGGGCTGCGGGATGGTGCATCCGAACGTGT
TGCGTAACGTTGGCATCGACCCGGAAGTTTACTCTGGTTTCGGCTTCGGGATG

GGGATGGAGCGTCTGACTATGTTGCGTTACGGCGTCACCGACCTGCGTTCATT
CTTCGAAAACGATCTGCGTTTCCTCAAACAGTTTAAATAAGGCAGGAATAGA
TTATAATCTTGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATG
GAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTA
AAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACC
GTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACA
AGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGG
AATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCA
CCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGA
GTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTG
GCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATAT
GTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGT
GGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATA
CGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT
CTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCG
ATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAA
ACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGCGTTTCGGTGAT
GACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGT
GTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACC
ATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGC
GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG
CGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TCATCCATAGCTGTATGTATCCGCTCATGAGACAATAACCCTCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAA

CCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGT
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCT
CGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTT
ACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC
TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
TTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACC
GCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG
CATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCTTTGAATGTATTTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAA
GAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGC
CCTTTCGTCTTCAC

*5.4.4 L*

Plasmid Name: p**-L

Vector size: 3000 bp

Cloning Site: BamHI and HindIII

Construction of pQE30-p**-L vector: The original construct was obtained from Natalya Voloshchuk and given to Ching Yao Yang. This vector is used for incorporation of unnatural amino acid by the mutation of PheRS. Ching Yao passed the vector containing the p** mutation to me which I then restricted to release his gene and ligated mine to.

Source available:

50 % glycerol culture: - 80 °C JH box

Culture conditions: Medium: LB/Amp Temp. 37 °C

Plasmid DNA : JH - 20 °C DNA box

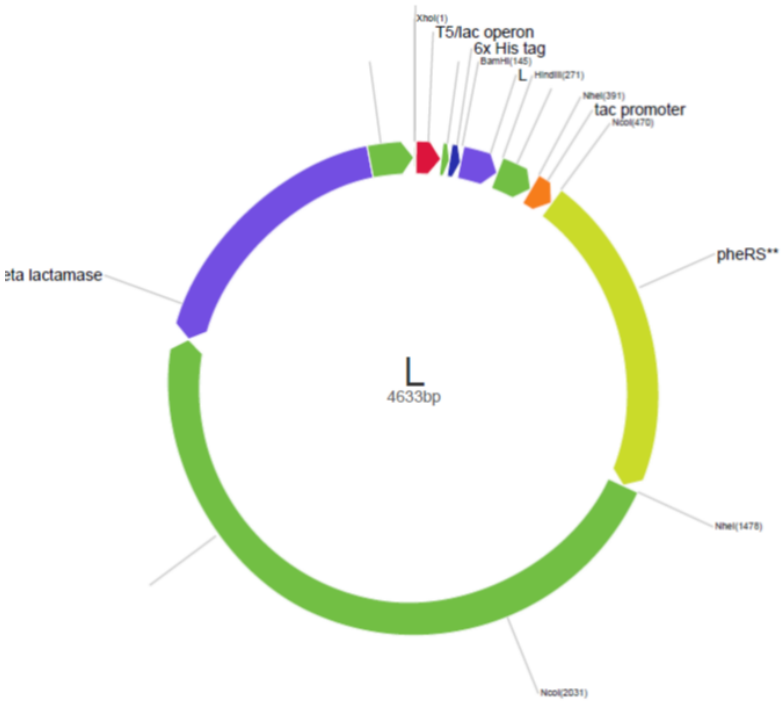



Figure A 4. Plasmid map for L.

Table A 7. Position of restriction sites in pQE30-p**L plasmid.

| **Restriction site** | **Sequence** | **Position** |
|---|---|---|
| XhoI | CTCGAG | 145 |
| NheI | GCTAGC | 271 |
| NcoI | CCATGG | 1 |
| BamHI | GGATCC | 291, 1478 |
| HindIII | AAGCTT | 470, 2031 |

Table A 8. Positions of DNA motifs in pQE30-p**L plasmid.

| **Motif** | **Position** |
|---|---|
| Start of numbering at XhoI (CTCGAG) | 1 |
| T5/lac operon | 7 |
| Gene (L) | 151 |
| tac promoter | 392 |
| pheRS** (alpha, T251G/A294G) | 472 |
| Beta lactamase gene | 3633 |

## Full Sequence

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAAT
AGATTCAATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAG
AAATTAACTATGAGAGGATCGCATCACCATCACCATCACGGATCCATCGAAG
GTCGCCTGCAGGACGTTCGTGAACTGCTGCGTCAGCAGGTTAAAGAAATCAC
CTTCCTGAAAAACACCGCGCCGCAGATGCTGCGTGAACTGCAGGAAACCAAC
GCGGCGTCAAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTA
ATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGC
GTTTTTTATTGGTGAGAATCCAAGCTAGCAGTTGACAATTAATCATCGGCTCG
TATAATGGATCGAATTGTGAGCGGAATCGATTTTCACACAGGAAACAGACCA
TGGCACATCTCGCAGAACTGGTTGCCAGTGCGAAGGCGGCCATTAGCCAGGC
GTCAGATGTTGCCGCGTTAGATAATGTGCGCGTCGAATATTTGGGTAAAAAA
GGGCACTTAACCCTTCAGATGACGACCCTGCGTGAGCTGCCGCCAGAAGAGC
GTCCGGCAGCTGGTGCGGTTATCAACGAAGCGAAAGAGCAGGTTCAGCAGGC
GCTGAATGCGCGTAAAGCGGAACTGGAAAGCGCTGCACTGAATGCGCGTCTG
GCGGCGGAAACGATTGATGTCTCTCTGCCAGGTCGTCGCATTGAAAACGGCG
GTCTGCATCCGGTTACCCGTACCATCGACCGTATCGAAAGTTTCTTCGGTGAG
CTTGGCTTTACCGTGGCAACCGGGCCGGAAATCGAAGACGATTATCATAACT
TCGATGCTCTGAACATTCCTGGTCACCACCCGGCGCGCGCTGACCACGACAC
TTTCTGGTTTGACACTACCCGCCTGCTGCGTACCCAGACCTCTGGCGTACAGA
TCCGCACCATGAAAGCCCAGCAGCCACCGATTCGTATCATCGCGCCTGGCCG
TGTTTATCGTAACGACTACGACCAGACTCACACGCCGATGTTCCATCAGATGG
AAGGTCTGATTGTTGATACCAACATCAGCTTTACCAACCTGAAAGGCACGCT
GCACGACTTCCTGCGTAACTTCTTTGAGGAAGATTTGCAGATTCGCTTCCGTC
CTTCCTACTTCCCGTTTGCCGAACCTTCTGCAGAAGTGGACGTCATGGGTAAA
AACGGTAAATGGCTGGAAGTGCTGGGCTGCGGGATGGTGCATCCGAACGTGT
TGCGTAACGTTGGCATCGACCCGGAAGTTTACTCTGGTTTCGGCTTCGGGATG

GGGATGGAGCGTCTGACTATGTTGCGTTACGGCGTCACCGACCTGCGTTCATT
CTTCGAAAACGATCTGCGTTTCCTCAAACAGTTTAAATAAGGCAGGAATAGA
TTATAATCTTGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATG
GAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTA
AAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTATAACCAGACC
GTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACA
AGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGG
AATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCA
CCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGA
GTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTG
GCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATAT
GTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGT
GGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGCAAATATTATA
CGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGT
CTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCG
ATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATTGGTGCCCTTAA
ACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAAGGCT
CAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCT
CCTGAGTAGGACAAATCCGCCGCTCTAGAGCTGCCTCGCGCGTTTCGGTGAT
GACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGT
GTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACC
ATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAG
GCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGTCGGCTGC
GGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG
GACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCT
GTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG
CGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGC
GGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGA
CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGA
TTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA
AAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TCATCCATAGCTGTATGTATCCGCTCATGAGACAATAACCCTCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTG
CAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAA

CCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGT
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCT
CGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTT
ACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC
TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTG
AGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTC
TTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAA
GTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACC
GCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG
CATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA
TGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCTTTGAATGTATTTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAA
GAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGC
CCTTTCGTCTTCAC

## 5.5 Additional information on proteins of interest

### *5.5.1 DSAg and DSD*

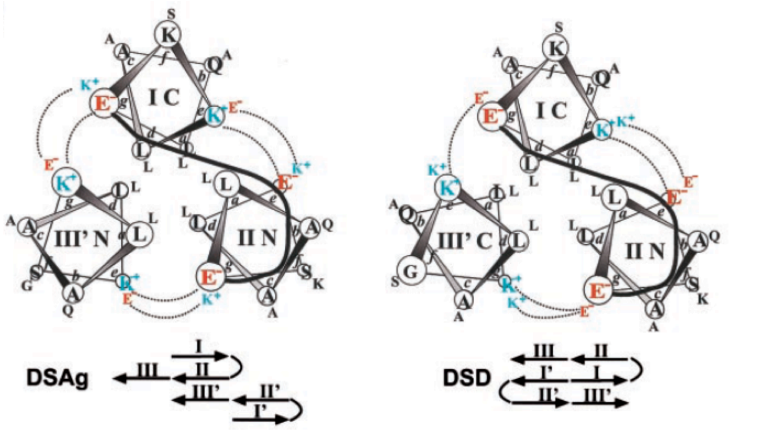



Figure A 5. Helical wheel diagrams and amino acid sequences of DSAg and DSD.[49]

### *5.5.2 PA*

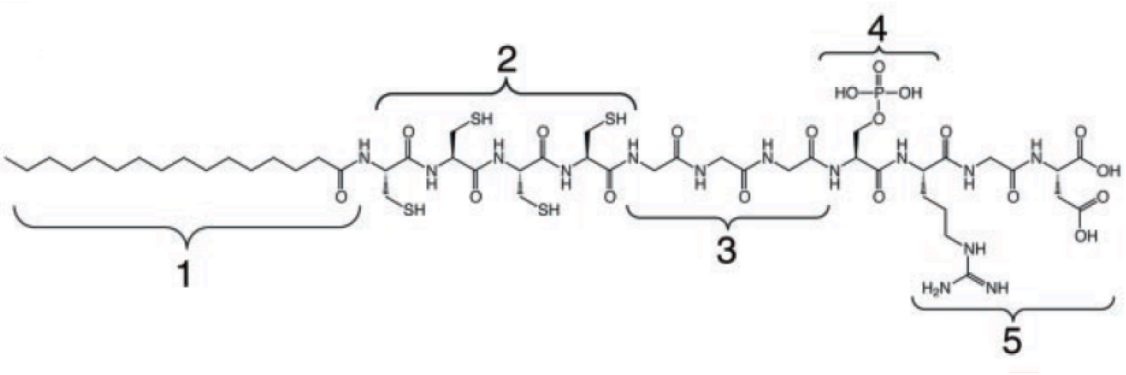



Figure A 6. Chemical structure of the peptide amphiphile (PA), where region 1 is a long alkyl tail that conveys hydrophobic character, region 2 is composed of four consecutive cysteine residues, region 3 is a flexible linker region of three glycine residues, region 4 is a single phosphorylated serine residue, and region 5 displays the cell adhesion ligand RGD.[67]

### *5.5.3 α-amylase*

Theoretical pI = 6.33

Molecular weight = 58.5 kDa

Amino acid sequence:

MKQQKRLYARLLTLLFALIFLLPHSAAAAANLNGTLMQYFEWYMPNDGQHWKRLQNDS
AYLAEHGITAVWIPPAYKGTSQADVGYGAYDLYDLGEFHQKGTVRTKYGTKGELQSAIKSL
HSRDINVYGDVVINHKGGADATEDVTAVEVDPADRNRVISGEHRIKAWTHFHFPGRGSTY
SDFKWHWYHFDGTDWDESRKLNRIYKFQGKAWDWEVSNENGNYDYLMYADIDYDHP
DVAAEIKRWGTWYANELQLDGFRLDAVKHIKFSFLRDWVNHVREKTGKEMFTVAEYW
QNDLGALENYLNKTNFNHSVFDVPLHYQFHAASTQGGGYDMRKLLNSTVVSKHPLKAVT
FVDNHDTQPGQSLESTVQTWFKPLAYAFILTRESGYPQVFYGDMYGTKGDSQREIPALKH
KIEPILKARKQYAYGAQHDYFDHHDIVGWTREGDSSVANSGLAALITDGPGGAKRMYVGR
QNAGETWHD ITGNRSEPVVINSEGWGEFHVNGGSVSIYVQR

### *5.5.4 Sericin silk protein*

Theoretical pI = 5.92

Molecular weight = 119.5 kDa

Amino acid sequence:

MRFVLCCTLIALAALSVKAFGHHPGNRDTVEVKNRKYNAASSESSYLNKDNDSISAGAHR
AKSVEQSQDKSKYTSGPEGVSYSGRSQNYKDSKQAYADYHSDPNGGSASAGQSRDSSLRE
RNVHYVSDGEAVAASSDARDENRSAQQNAQANWNADGSYGVSADRSGSASSRRRQANY
YSDKDITAASKDDSRADSSRRSNAYYNRDSDGSESAGLSDRSASSSKNDNVFVYRTKDSIG
GQAKSSRSSHSQESDAYYNSSPDGSYNAGTRDSSISNKKKASSTIYADKDQIRAANDRSSSK
QLKQSSAQISSGPEGTSVSSKDRQYSNDKRSKSDAYVGRDGTVAYSNKDSEKTSRQSNTNY
ADQNSVRSDSAASDQTSKSYDRGYSDKNIVAHSSGSRGSQNQKSSSYRADKDGFSSSTNTK
SKFSSSNSVVETSDGASASRESSAEDTKSSNSNVQSDEKSASQSSSSRSSQESASYSSSSSSST
LSEDSSEVDIDLGNLGWWWNSDNKVQRAAGGATKSGASSSTQATTVSGADDSADSYTW
WWNPRRSSSSSSSASSSSSGSNVGGSSQSSGSSTSGSNARGHLGTVSSTGSTSNTDSSSKSA
GSRTSGGSSTYGYSSSHRGGSVSSTGSSSNTDSSTKNAGSSTSGGSSTYGYSSSHRGGSVSST
GSSSNTDSSTKSAGSSTSGGSSTYGYSSRHRGGRVSSTGSSSTTDASSNSVGSSTSGGSSTYG
YSSNSRDGSVSSTGSSSNTDSNSNSAGSSTSGGSSTYGYSSNSRDGSVSSTGSSSNTDSNSNS
AGSSTSGGSSTYGYSSNSRDGSVSSTGSSSNTDASTDLTGSSTSGGSSTYGYSSDSRDGSVSS
TGSSSNTDASTDLAGSSTSGGSSTYGYSSDCGDGSVSSTGSSSNTDASTDLAGSSTSGGSSTY
GYSSDSRDGSVSSTGSSSNTDASTDLAGSSTSGGSSTYGYSSNSRDGSVSSTGSSSNTDASTD
LTGSSTSGGSSTYGYSSSNRDGSVLATGSSSNTDASTTEESTTSAGSSTEGYSSSSHDGSVTS
TDGSSTSGGASSSSASTAKSDAASSEDGFWWWNRRKSGSGHKSATVQSSTTDKTSTDSAS
STDSTSSTSGASTTTSGSSSTSGGSSTSDASSTSSSVSRSHHSGVNRLLHKPGQGKICLCFENI
FDIPYHLRKNIGV

### *5.5.5 BP124NAB peptide*

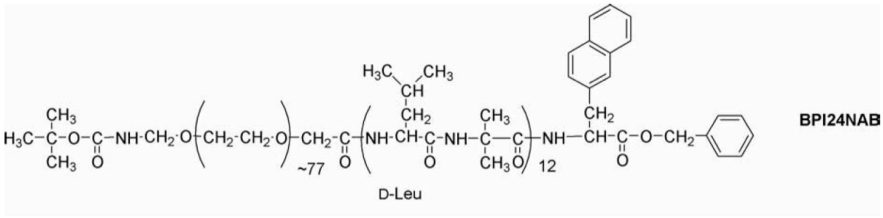



Figure A 7. Molecular structure of the amphiphilic helical peptide BP124NAB.[167]

### *5.5.6 Helical gelator*

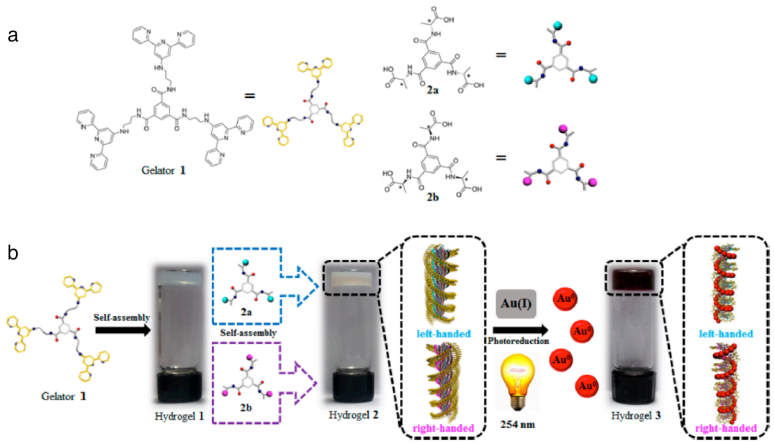



Figure A 8. Molecular structures of gelator 1 and chiral components 2a and 2b (a). Schematic of hydrogel formation, resulting in nanofibers with tunable helicity and left-handed or right-handed chirality (b).[168]

### *5.5.7 Cty c*

Theoretical pI = 9.59

Molecular weight = 11.7 kDa

Amino acid sequence:

MGDVEKGKKIFIMKCSQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGYSYTAANKNKGII
WGEDTLMEYLENPKKYIPGTKMIFVGIKKKEERADLIAYLKKATNE

### *5.5.8 Bovine serum albumin*

Theoretical pI = 5.82

Molecular weight = 69.3 kDa

Amino acid sequence:

MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFD
EHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPE
RNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYA
NKYNGVFQECCQAEDKGACLLPKIETMREKVLASSARQRLRCASIQKFGERALKAWSVA
RLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDNQDTISSKLKEC
CDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHP
EYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLG
EYGFQNALIVRYTRKVPQVSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLC

VLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQ
IKKQTALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVVSTQTA
LA

## 5.6 Experimental protocols

### *5.6.1 Preparation of chemically competent cells*[*]

1. Prepare TSA plates with appropriate antibiotics
2. For M15MA cells: use (final concentration) 35 mg/L kanamycin
3. Streak plates with cells. Incubate overnight at 37 °C.
4. Inoculate 2 mL starter culture of LB, supplemented with appropriate antibiotics (35 mg/L kanamycin for M15MA), with a single colony of *E.coli*. Grow overnight at 37 °C (300-350 rpm) in a shaking incubator
5. Subculture large-scale culture (100 mL) of LB, supplemented with appropriate antibiotics (35 mg/L kanamycin for M15MA), 200 µL of starter culture. Grow culture at 37 °C (300-350 rpm) in a shaking incubator until early log-phase of growth ($OD_{600} = 0.3 \pm 0.05$). Keep in mind that doubling times for cell density (i.e. $OD_{600}$) once log-phase is reached is approximately 20 minutes, thus requiring sometimes frequent measurements. Culture times are typically around 2 hrs.
6. Once the $OD_{600} = 0.3$ is achieved, immediately store culture in appropriate centrifuge container and incubate at 4 °C for 20 minutes by submerging closed container in ice water bath. While culture is cooling, equip centrifuge with appropriate rotor and pre-chill centrifuge by running at 4 °C, 4000 rpm for 20 minutes.
7. Pellet the culture by spinning at 4 °C, 4000 rpm for 10 minutes.
8. Ensure that a pellet has indeed collected and decant supernatant. Discard supernatant.
9. Add 33 mL of TB I (pre-chilled) to pellet. Resuspend pellet by gently swirling buffered slurry.
10. Incubate slurry in ice-water bath for 2 hrs.
11. With 30 minutes left for ice-water incubation, crush blocks of dry ice to obtain a rough bed of powder. Embed 1.5 mL microcentrifuge tubes in the bed of powder to pre-chill.
12. With 20 minutes left for ice-water incubation, pre-chill centrifuge with appropriate rotor as in step 1.
13. Pellet the slurry by running at 4 °C, 4000 rpm for 10 minutes.
14. Decant and discard supernatant.
15. Resuspend pellet in 11 mL of TB II (pre-chilled)

---

[*] http://violet.poly.edu/wiki/index.php5/Preparation_of_Chemically_Competent_E.coli_Cells

16. Aliquot slurry in 100 µL amounts amongst the pre-chilled microcentrifuge tubes. Return tubes to dry ice for rapid freezing of cell slurry aliquots.
17. Collect and store aliquots in a cryocase (pre-chilled) for storage at -80 °C freezer.
18. Transform cells with desired plasmid vector containing your gene of interest via heat shock.

### *5.6.2 Expression in XL1blue*

#### 5.6.2.1 Streaking plates

- Take 200 mg/ml ampicillin (stored at -20 °C), add 250 µl ampicillin to 250 mL tryptic soy agar, mix well.
    - 250 mL agar is enough for 11 plates (approx. 23 ml/plate)
- Take TSA from the autoclave and let it cool. If it's too hot it will degrade the ampicillin. Do not let it cool too much, however, since it solidifies quickly.
    - Once at the appropriate temperature, pour some into each of the 11 plates. Just enough to coat the bottom of the plate.
    - Let the agar cool and solidify on the plate
- Streaked plates with different bacteria (1 plate/type)
- *Put glycerol bacteria vials on dry ice while streaking plates*
- Take the loop for streaking and heat it until red twice, rinsing between each time with methanol. The last time just heat enough to evaporate remaining methanol.
    - Dip the tip in the frozen bacterial sample (glycerol stock) and swirl around a little.
    - Streak the plate, starting in one area and zig-zagging around the plate, rotating the plate in one hand.
    - Close the plate and keep it upside down.
    - Repeat for all bacteria types.
- Incubate plates overnight (between 12 and 16 h) at 37 °C, upside-down position.

#### 5.6.2.2 Starter cultures

- Make 3 x 1 L flasks of LB broth. For 1 L:
    - 1 L water + 10 g NaCl + 10 g tryptone + 5 g yeast extract
    - Autoclaved 3 flasks of LB broth and 6 centrifuge bottles & caps, rinsed.
- Get sterilized starter culture tubes (3/plate)
- Add 5 mL LB broth + 5 µl 200 mg/ml ampicillin in each tube
- Pick and add bacterial colony to each tube, flame tube and cap and close. Repeat for all tubes and bacteria types.
- Put starter culture tubes on shaker: 37 °C, 250 rpm. Leave them on shaker 12-16 h.

#### 5.6.2.3 Large scale expression

- Optical density measurements

  - Measure the optical density of the starter cultures to see which has the highest cell concentration at **600 nm**
    - Wait until ODs are > 1.0 to inoculate large scale (otherwise it'll take a really long time later)
  - Use starter cultures that have the highest ODs (within the same range)
- Inoculate 1 l flasks
  - Add 1 mL of 200 mg/ml ampicillin to 1 L broth (x2 for 2 flasks)
  - Add 3 mL of selected starter culture. Flame flask and foil. Label flask.
  - Put on shaker at 250 rpm, 37 °C
    - Will take 7 - 10 h to reach OD of 0.8 usually
- Induction
  - Induce when OD is > 0.8 (after around 7 hours)
  - Take 1 mL samples of the culture before induction (x2)
    - Store samples in -20 °C freezer
  - 1 mL of 200 mg/ml IPTG added to each flask for induction
  - Shake for 3 h at 37 °C at 250 rpm
  - Collect 1 mL sample from each flask post-induction. Store in -20 °C freezer
- Spin the cells
  - After induction, put approx. 400 mL in centrifuge bottles. (2 bottles/flask)
    - Cool down for 30 minutes in ice
  - Spin at 4 °C, 4000 rpm, 20 min
  - Discard LB, and resuspend the pellet in 40 mL 50 mM tris-HCl, 0.5 M NaCl, 20 mM imidazole, 6 M Urea pH 8 buffer and put in -80 °C freezer

### *5.6.3 Expression in M15MA with incorporation of azidohomoalanine*

#### 5.6.3.1 Reagent recipes

10x M9 stock solution (200 mL final volume prepared in a 500 mL autoclaved bottle)

12 g $Na_2HPO_4$
6 g $KH_2PO_4$
1 g NaCl
2 g $NH_4Cl$
Fill to 200 mL with $dH_2O$.

Note: Upon diluting 10X M9 to obtain 1X M9, autoclave 10X solution and bottle for 1X M9 prior to expression.

#### 5.6.3.2 Additional reagent preparation

0.1 M $CaCl_2$: 111 mg in 10 mL $dH_2O$

1.0 M $MgSO_4$: 1.20 g in 10 mL $dH_2O$

40 % w/v glucose monohydrate (requires stirring on magnetic stirrer; add this last): 40 g glucose + 100 mL $dH_2O$

34 mg/mL vitamin B

2 mg/mL IPTG: 2 g IPTG in 10 mL $dH_2O$

20 amino acids 1 mg/mL solution

19 amino acids solution (no methionine)

#### 5.6.3.3 Pre-induction

1. When $OD_{600}$ > 1, add 4 mL starter culture to:

   200 mL 1X M9
   200 µL Ampicillin (200 mg/mL)
   200 µL Kanamycin (35 mg/mL)
   200 µL Vitamin B (35 mg/mL)
   200 µL 1 M $MgSO_4$
   200 µL 0.1 M $CaCl_2$
   2 mL glucose (40 % w/v)
   8 mL 20 amino acids (1 mg/mL)
   *TOTAL VOLUME = 211 mL or 105.5 mL per tube*

2. Sterilize flask and cap over flame
3. Leave on shaker for 6 hours at 350 rpm at 37 °C
4. Measure $OD_{600}$ = 1 (take 1 mL pre-sample and store in - 20 °C)
5. Transfer culture to centrifugal bottles and spin down at 4000 rpm, 15 min, 4 °C
6. Repeat the following steps x3 to remove any & all amino acids for introduction of AHA in place of methionine
7. Centrifuge bottles, dump the supernatant
8. Add 150 mL 0.9 % cold NaCl solution, resuspend
9. Weigh, centrifuge at 4000 rpm, 15 min, 4 °C
10. Remove all supernatant, resuspend, repeat.

#### 5.6.3.4 Induction

Volumes are adjusted for 2 small scale expressions

1. x3: Prepare auxotrophic media (no methionine)

   200 mL 1X M9
   200 µL Ampicillin
   200 µL Kanamycin
   200 µL Vitamin B
   200 µL 1 M $MgSO_4$

200 µL 0.1 M $CaCl_2$
8 mL 19 Amino acids (no methionine)
2 mL 40 % glucose
TOTAL VOLUME = 211 mL

2. Add 100 mL prepared auxotrophic media (19 aa, see step 1 above) to pellet in centrifuged bottles. Shake/vortex to resuspend pellet completely.
3. Sterilize flasks over flame
4. Starve the cells of methionine by shaking them at 350 rpm, 15 min, 37 °C
5. Measure $OD_{600}$
6. Add 200 µL of 200 µg/mL IPTG and 57.6 µL 100 mg/mL AHA to each flask
7. Induce at 300 rpm, 37 °C, 3 h
8. Goal: obtain OD that is 2x that of pre-induction
9. Take 1 mL and store in - 20 °C for gel.
10. Transfer to centrifuge bottles and spin for 4000 rpm, 15 min, 4 °C

*5.6.4 Purification*

5.6.4.1 Lyse the cells

1. Take frozen lysed cells and thaw until liquid
2. Lyse cells via sonication for 1 minute (Pulse 5 sec on/ 5 sec off; Time: 1 min; Amplification: 35%)
3. Spin at 4 °C, 4000 rpm for 50 min

5.6.4.2 Ni-NTA binding

1. 50 mL conical tubes for Ni beads (1/pellet)
2. Ni-NTA slurry stored in 4 °C fridge, common shelf. Shake well before use.
3. Add 5 mL Ni-NTA slurry (5 mL slurry/40 mL lysate) and 25 mL lysis buffer to each conical tube
4. Spin beads: 4000 rpm, 4 °C, 15 min. Pour out supernatant.
5. Add supernatant from lysed cells to the Ni-NTA containing-tubes. Close tightly and shake to resuspend beads.
6. Put on rotator in 4 °C fridge overnight (or at least for 3 h)
7. Spin the mixture 20 minutes, 4 °C, 2000 rpm.
    a. Take a 1 mL aliquot of the supernatant and save for SDS-PAGE.
    b. Discard remaining supernatant.

c. Add 10 mL of lysis buffer [50 mM tris-HCl, 0.5 M NaCl, 20 mM imidazole, 6 M urea pH 8 buffer] and resuspend the beads

5.6.4.3 Purification by denaturing conditions - imidazole gradient

Washing and elution achieved with a buffer solution of progressively increasing concentration of imidazole. Protein dissociates from the Ni-NTA through competitive binding of the beads with imidazole.

1. Prepare 500 mL of purification buffer
2. Prepare solutions of 100 mM, 200 mM, 500 mM and 1 M imidazole using 50 mM tris-HCl, 0.5 M NaCl, 20 mM imidazole, 6 M urea pH 8 buffer
    a. 15 mL 100 mM imidazole: add 82 mg imidazole to 15 mL purification buffer
    b. 30 mL 200 mM imidazole: add 184 mg imidazole to 30 mL purification buffer
    c. 50 mL 500 mM imidazole: add 490 mg imidazole to 50 mL purification buffer
    d. 80 mL 1 M imidazole: add 1 g imidazole to 80 mL purification buffer
3. Setup gravity filtration columns and respective flow-through collection tubes. Wash columns thoroughly with $dH_2O$before/after use
4. Combine both samples from each protein type in one column. Load mixture onto column, let it go through column (Flow Through) and collect sample
5. Wash residual beads from tube with 100 mL purification buffer (20 mM imidazole)
    a. Wash 1: first 50 mL
    b. Wash 2: second 50 mL
6. Elutions: After washing with 100 mL buffer, begin pouring in imidazole solutions starting with 100 mM imidazole
    a. Elution 1: 5 mL of 100 mM imidazole solution
    b. Elution 2: 5 mL of 200 mM imidazole solution
    c. Elution 3: 5 mL of 200 mM imidazole solution
    d. Elution 4: 5 mL of 500 mM imidazole solution
    e. Elution 5: 5 mL of 500 mM imidazole solution
    f. Elution 6: 5 mL of 500 mM imidazole solution
    g. Elution 7: 10 mL of 1 M imidazole solution
    h. Elution 8: 10 mL of 1 M imidazole solution
    i. Elution 9: 10 mL of 1 M imidazole solution
7. In all, there should be supernatant, flow-through, 2 washes, and 9 elutions for each protein

*5.6.5 Dialysis*

1. Dialyze into 50 mM PB pH 4, 8, or 10. Volume of each bucket is 2 L.

    bucket 1: 3 M urea (2 hours)
    bucket 2: 1.5 M urea (2 hours)
    bucket 3: 0.75 M urea (2 hours)

bucket 4: 0 M urea (3 hours)
bucket 5: 0 M urea (3 hours)
bucket 6: 0 M urea (3 hours)

2. Cut the desired amount of dialysis tubing and presoak in buffer.
3. Fold tubing and clip. Fill with protein and fold and clip again to close. Attach foam floating devices to suspend.
4. Put a stir bar in the bucket and add the protein. Let stir and dialyze in 4 °C fridge.
5. When dialysis is complete, remove proteins from dialysis tubing, put into 50 mL conical vials.
6. Measure concentration via BCA Enhanced Protocol, Appendix 5.6.6.

### *5.6.6 BCA, Enhanced Protocol*

#### 5.6.6.1 Preparation of Standards

Table A 9. Preparation of standards for BCA.

| Dilution Scheme for Enhanced Test Tube Protocol (Working Range = 5–250µg/mL) | | | |
|---|---|---|---|
| Vial | Volume of Diluent (µL) | Volume and Source of BSA (µL) | Final BSA Concentration (µg/mL) |
| A | 700 | 100 of Stock | 250 |
| B | 400 | 400 of vial A dilution | 125 |
| C | 450 | 300 of vial B dilution | 50 |
| D | 400 | 400 of vial C dilution | 25 |
| E | 400 | 100 of vial D dilution | 5 |
| F | 400 | 0 | 0 = Blank |

#### 5.6.6.2 Preparation of the working reagent (WR)

1. Turn on the water bath (to 60 °C) so it is heated by the time you finish reagent preparation.
2. Use the following formula to determine the total volume of WR necessary:

   (# Standards + # unknowns) x (# replicates) x (1.43 mL of WR per sample) = total volume WR
3. Prepare WR by mixing 50 parts of BCA Reagent A with 1 part of BCA reagent B (50:1, Reagent A:B)

#### 5.6.6.3 Test-tube procedure

1. Use 1.5 mL epitubes
2. Pipette 71.4 µL of each standard, as well as 71.4 µL of each unknown to be measured, into an appropriately labeled 1.5 mL epitube.
3. Add 1.43 mL of the WR to each tube and mix well by inverting several times.
4. Cover and incubate tubes at 60 °C for 30 minutes in the water bath.

5. Set spectrophotometer to 562 nm and zero the instrument on a cuvette filled only with $dH_2O$.
6. Pour contents of sample into spectrophotometric cuvette and measure.*

*5.6.7 Running SDS-Page gels*

5.6.7.1 Pre-post gels

1. Prepare the samples.
    a. If frozen, thaw pre/post samples until liquid.
    b. Spin pre/post samples for 3 minutes at 13600 rpm.
    c. Discard supernatant.
    d. Add 50 µl lysis buffer. Resuspend by shaking.
    e. Add 5 µl SDS dye.
    f. Heat at ~100 °C for ~20 min in heat block.
    g. Add 15 µl of each sample to gel lane.
2. Run the gels.
    a. Use stacking & separating buffer for 12 % protein gels
    b. Make and run gel with 8 tooth comb
    c. After loading (7 µl for ladder, 15 µl all other samples), turn on for 10 minutes at 85 V, then increase to 200 V for 45 min.
    d. Stain with staining solution, microwave for 45 s
    e. Destain with destaining solution, cover with 2 Kimwipes and microwave for 2 minutes. Repeat. Put on the shaker and leave overnight in destaining solution.
    f. Take gel pictures the following day.

5.6.7.2 Purification gels

1. Samples to be loaded in the gel are 9 µl protein + 3 µl SDS dye.
2. Follow Step 2 in Section 5.6.7.1for instructions on running the gel.

*5.6.8 Chemical synthesis of L-azidohomoalanine*

Flame dry all glassware. Reaction described by Link *et al.*[254] is scaled up in this section by 1.802 (= 1.77 g/3.19 g) to adjust for an increased amount of Boc-dab.

**Day 1 [Total time = 4.5 h]**

1. Distillation of triflic anhydride:
    a. Add 970 mg of $P_2O_5$ to 15.44 g $Tf_2O$ *slowly* in the fume hood. If $Tf_2O$ contains water, this will result in a highly exothermic reaction.
    b. Heat in an oil bath at 95 C until all liquid has been distilled. Condenser reflux water should be at 10 C.

---

* Important: be sure to take all measurements *within 10 minutes* of removing from water bath.

2. Add 9.5 g sodium azide to 23.4 mL $dH_2O$in round bottom flask. Cover and put in an ice bath.
3. Add 4.87 mL triflic anhydride to the flask. Remove from the ice bath and stir the mixture at room temperature for 2 h.
4. In another round bottom flask, combine 46.9 mL $dH_2O$and 93.7 mL methanol. Add 3.2 g Boc-dab, 3.03 g $K_2CO_3$, and 56.36 mg $CuSO_4 \cdot 5H_2O$.
5. Transfer reaction mixture to separatory funnel and extract it with 14.4 mL methylene chloride. Shake the separatory funnel by inverting it and swirling, occasionally venting to release gas buildup. Collect the bottom organic layer, set aside. Extract the aqueous layer with another 14.4 mL methylene chloride, mix, and separate.
6. Neutralization: Use $Na_2CO_3$ to neutralize any remaining triflic acid. Combine the methylene chloride layers and was in a separatory funnel with 45 mL saturated $Na_2CO_3$. Collect the methylene chloride layer (bottom), which contains triflic azide.
7. Add the methylene chloride/triflic azide solution dropwise to the flask from step 4. Stir overnight at room temperature.

**Day 2 [Total time = 6 h]**

1. Remove methanol and methylene chloride by rotary evaporation at 40 °C leaving behind about 45 mL of an aqueous solution. This solution was bright blue and pH ~ 9.3.
2. Add 6 N HCl slowly to the solution until the pH reaches 6. Note that upon reaching pH ~ 8.5 the solution turned green.
3. Prepare phosphate buffer: 6.80 g $KH_2PO_4$ (0.25 M) + 11.41 g $K_2HPO_4 \cdot 3H_2O$ (0.25 M) in 200 mL $dH_2O$. Adjust pH to 6.2. Add 180 mL of the phosphate buffer to the green solution. Readjust pH to 3 with 6 N HCl.
4. Extract the aqueous layer twice with 2 x 180 mL ethyl acetate.
5. Combine ethyl acetate layers and wash in separatory funnel with 90 mL saturated NaCl. Collect ethyl acetate layers and add 9 g drying agent ($MgSO_4$) to remove any water from the organic layer. Leave for 30 minutes. Remove $MgSO_4$ by gravity filtration through filter paper. Remove ethyl acetate by rotary evaporation at 40 °C. Result was 5 g of oil (protected AHA).
6. Remove from rotary evaporator; add 18 mL concentrated HCl. Stir at room temp for 1 h.
7. Dilute the solution with 162 mL $dH_2O$. Store in the 4 °C fridge overnight.

**Day 3 [Total time = 6-7 h]**

1. Add 45 g Dowex resin to a large glass chromatography column
2. Wash the resin with:
    1. 3 x 180 mL 1 N $NH_4OH$
    2. 1 x 7 L $dH_2O$
    3. 1 x 180 mL 1 N HCl
    4. 1 x 7 L $dH_2O$

   Eluent should be slightly acidic (pH 6-7).

3. Add solution from Day 2 step 7 to the column. Collect the flow through and add it to the column once more.
4. Wash the column with 10.8 L $dH_2O$, or until the pH of the eluent remains constant (pH 6-7). Test the pH frequently with Litmus paper.
5. Elute the amino acid from the column by adding 720 mL 1 N $NH_4OH$.
6. Remove the aqueous $NH_4OH$ using a rotary evaporator at 65 °C.
7. Dissolve the product obtained in the previous step in a minimal amount of $dH_2O$in a 100 mL round bottom flask.
8. Add acetone drop wise until the solution starts to cloud. Swirl the mixture in a 75 °C water bath until the mixture is clear again. Cover the flask and store it in the fridge overnight.

**Day 4 [Total time = 1 h]**
Collect the crystalline AHA by gravity or vacuum filtration on filter paper.

## 5.7 Reagent recipes

- **Ampicillin, IPTG, APS aliquots**
  - Prepare 1 mL aliquots of 200 mg/ml AMP/IPTG/APS
  - Ampicillin stock is stored on the common shelf of the 4 $^0$C fridge
  - IPTG stock is stored on the top shelf of -20 $^0$C freezer
  - APS is stored at room temperature on the dry chemicals shelf
  - Add MQ to 2 g AMP/IPTG/APS to reach a total volume of 10 ml
  - Makes 10 x 1 mL aliquots of 200 mg/ml concentration
  - Store at -20 °C
- **CAM aliquots**
  - 0.35 g CAM in 10 mL of non-denatured ethanol
  - Makes 10 x 1 mL aliquots of 35 mg/ml concentration
  - Store at -20 °C
- **KAN aliquots**
  - 0.35 g KAN in 10 mL MQ
  - Makes 10 x 1 mL aliquots of 35 mg/ml concentration
  - Store at -20 °C
- **Phosphate buffer, pH 8**
  - sodium phosphate monobasic buffer (0.2 M)
  - sodium phosphate dibasic buffer (500 ml, 0.2 M)
    - 14.2 g sodium phosphate dibasic anhydrous in 500 mL MQ
  - sodium phosphate buffer (600 ml, 0.1 M, pH 8)
    - 300 mL MQ + 284.1 mL dibasic + 15.9 mL monobasic
  - Autoclave
- **Purification buffers**
  - Lysis buffer: 20 mM imidazole, 50 mM Tris-HCl, 500 mM NaCl, 6 M urea
  - Wash buffers (N = # of proteins to purify)
    - Wash 1: 20 mM imidazole, volume = 50 mL x N
    - Wash 2: 20 mM imidazole, volume = 50 mL x N

    - Elution buffers (N = # of proteins to purify)
        - Elution 1: 100 mM imidazole, volume = 5 mL x N
        - Elution 2-3: 200 mM imidazole, volume = 10 mL x N
        - Elution 4-6: 500 mM imidazole, volume = 15 mL x N
        - Elution 7-9: 1 M imidazole, volume = 30 mL x N
- **Tryptic soy agar**
    - 250 mL TSA for 10 plates
        - 10 g TSA + 250 mL MQ. Autoclave.
        - When TSA has cooled to around body temperature, add 250 µl of 200 mg/ml ampicillin
- **Luria bertani broth**
    - 250 mL LB
        - 2.5 g tryptone + 1.25 g yeast extract + 2.5 g NaCl + 250 mL MQ. Autoclave.
    - N x 1 L flasks of LB broth. For 1 l:
        - 1 l MQ + 10 g NaCl + 10 g tryptone + 5 g yeast extract
- **65 % glycerol**
    - Add 13 g glycerol to 7 g MQ
    - Autoclave
    - Store in 4 °C
    - After every use, flame the bottle & cap before returning to refrigerator
- **Glycerol stocks**
    - Use selected starter cultures to make stock solutions in glycerol
    - 500 µL 65 % glycerol (stored on common shelf, 4 °C fridge) + 500 µL starter culture.
    - Label:
        - Vector/gene
        - Host
        - Mutation
        - Date
        - Initials
    - Store in -80 °C freezer
- **Separating gel (for 2 gels)**
    - 1 mL MQ
    - 5 mL separating Buffer
    - 4 mL 30% Acrylamide (29:1)
    - 100 µL APS
    - 5 µL TEMED
- **Stacking gel (for 2 gels)**
    - 4.34 mL Separating Buffer
    - 0.68 mL 30 % Acrylamide (29:1)
    - 50 µL APS
    - 5 µL TEMED
- **Separating Buffer Solution**
    - 300 mL 3M TrisCl/SDS solution
    - 120 g (95 ml) Glycerol
    - 60 mL MQ

- **Stacking Buffer Solution**
  - 143 mL 3M TrisCl/SDS solution
  - 357 mL MQ
- **3M TrisCl/SDS solution**
  - Dissolve 182 g of tris base in 300 mL MQ
  - pH to 8.45 with 37 % (1N) HCl
  - Dilute to 500 mL
  - Add 1.4 g SDS
- **Tricine sample buffer (2x)**
  - 2 mL 4x TrisCl/SDS pH=6.8
  - 2.4 mL glycerol
  - 0.8 g SDS
  - 0.31 g DTT
  - 2 mg Coomassie Blue G-250
  - MQ to 10 ml
- **Transformation buffer I (TB-I)**
  - 30 mL potassium acetate (60 mM final)
  - 40 mL 1 M MnCl2 (80 mM final)
  - 50 mL 1 M RbCl (100 mM final)
  - 75 mL 15% (v/v) glycerol
  - Adjust volume to 500 mL with MilliQ H20
  - Filter sterilize
  - Store up to 1 year at 4 °C
- **Transformation buffer II (TB-II)**
  - 5 mL 1 M MOPS, pH 7.0 (10 mM final concentration)
  - 5 mL 1 M RbCl (10 mM final concentration)
  - 50 mL 1 M $CaCl_2$ (100 mM final concentration)
  - 75 mL 15% (v/v) glycerol
  - Adjust volume to 500 mL with MilliQ H20
  - Filter sterilize
  - Store up to 1 year at 4 °C
- **4x TrisCl/SDS pH = 6.8**
  - Dissolve 6.05 g of tris base in 40 mL MQ
  - pH to 6.8 with 37 % (1N) HCl
  - Add MQ to 100 mL total volume
  - Add 0.4 g SDS
- **Gel Drying Solution**
  - 40 % Methanol (1600ml)
  - 10 % Glycerol (400 ml)
  - 7.5 % Acetic Acid (300 ml)
  - Water to 4 L (1700 ml)
- **Staining Solution**
  - 2.5 g Coomassie Brilliant Blue
  - 450 mL methanol
  - 100 mL glacial acetic acid

    - Water to 1 L
- **Destaining Solution**
    - 300 mL methanol
    - 400 mL acetic acid
    - Water to 4 L
- **50x TAE (for 1 L)**
    - 242 g tris base
    - 57.1 g glacial acetic acid
    - 100 mL 0.5 EDTA pH = 8.0
        - To dilute to 0.5x TAE: 40 mL 50x TAE, bring to 4 L with MQ
        - To dilute to 0.5x TAE: 40 mL 50x TAE, bring to 2 L with MQ
- **6x Gel Loading Buffer**
    - Version 1:
        - 25 mg Bromophenicol Blue
        - 25 mg Xylene Cyanol
        - 4g Sucrose
        - MQ to 10 ml
    - Version 2:
        - 0.25 % Bromo phenico Blue
        - 0.25 % Xylene Cyanol
        - 15 % Ficoll Type 4000
        - 120 mM EDTA
- **1 % Agarose with ethidium bromide**
    - 1 g agarose
    - 100 mL 1xTAE buffer
    - 6 µL ethidium bromide
- **DNA ladder (50 bp or 1 kb)**
    - 50 µL ladder
    - 45 µL 0.5 % TAE buffer
    - 100 µL loading dye