Two-Dimensional Protein Arrays: De Novo Design and Applications

James F. Matthaei

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington
2015

Reading Committee:
François Baneyx, Chair
David G. Castner
Cole A. DeForest

Program Authorized to Offer Degree:
Department of Chemical Engineering

University of Washington

Abstract

Two-Dimensional Protein Arrays: De Novo Design And Applications

James F. Matthaei

Chair of the Supervisory Committee:
Department Chair and Charles W.H. Matthaei Professor of Chemical Engineering
François Baneyx, Ph.D.
Department of Chemical Engineering

Biological building blocks that self-assemble into predetermined supramolecular structures hold enormous promise for the production of advanced materials, devices and systems. However, our ability to predict, engineer and control short- and long-range interactions in proteins is lagging and nature's most versatile building block remains underused. Two dimensional (2D) protein arrays are of considerable interest in bionanotechnology. For instance, the remarkable stability and unique photochromic properties of purple membrane patches have been exploited for optical information storage and processing, and the fact that archeal and bacterial S-layers can be patterned on surfaces or used as templates and display scaffolds has opened the door to an even broader range of applications. Yet, neither the geometry, nor the chemistry or assembly of these systems can be precisely controlled from the nano- to the mesoscale. To address these shortcomings, we computationally designed a family of proteins capable of robust self-assembly into 2D arrays by fusing monomers from symmetry-compatible oligomers and redesigning interfacial contacts between unit cells. We optimized self-assembly conditions to promote the formation of protein lattices that retain short-range order at the nanoscale but can grow over tens to hundreds of micrometers. Finally, we demonstrated that modification of the designed protomers with a C-terminal biotinylation tag does not interfere with self-assembly and supports lattice decoration with Avidin. With proof of concept established, we anticipate that these well-defined and genetically accessible arrays will prove useful for bionanotechnology applications ranging from tunable molds for inorganic mineralization to scaffolds for high-density display of enzymes and proteins.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank everyone who helped me along this long journey. In particular to François Baneyx who shaped me from a mathematician into an engineer. Thank you for continuing to believe in me throughout all of my projects, even the ones that failed. I will carry your mentorship and advice about what it means to conduct research at the PhD level with me for the rest of my life.

To the rest of my committee, Dave Castner, Cole DeForest, and Marco Rolandi, thank you for your support and insights into my project. Also to Dan Schwartz, you served on my committee for my prelim, masters and general exams and, if not for a trip across the world, would have made the final exam too. Thank you to my collaborators David Baker, Frank Dimaio, Possu Huang, Jeff Richards, Lilo Pozzo, Scott Lea and James Evans. I would like to especially thank Jeff for his insights about SAXS and James Evans for his inspiration.

A big thanks to everyone who has been a part of the Baneyx lab: Weibin Zhou, Brandon Coyle, Brent Nannenga, Miki Pueras, Brian Swift, Rosie Zhang, Jessica Soto-Rodriguez, Carolynn Grosh, David Chiu, Alex Thomas, Kannan Aravagirl, Wenlan Yang and Brittney Hellner. Weibin, thank you for all of the late nights you spent helping me in the lab and for your friendship. Brent, Brandon, Brian and Miki thanks for all the conversation whether it was about the details of an experiment or simply just to laugh.

Finally thank you to my family. To my beautiful, smart and nice wife Christie, without your support I would not have made it through this process. Your smile cheered me up every day, especially on the days when the experiments didn't work out. To Mom, Dad and Josh and the rest of my extended family – your help, support, and love were and are greatly appreciated. And last, but definitely not least, to my Grandpa Chuck and Grandma Helen – thanks for the introduction to chemical engineering and the support along the way. You both showed me that learning is a life long endeavor and simply put, you're the best.

DEDICATION

To my Grandfather:

a family-man, an engineer, a baker, a businessman, a philanthropist and a Husky.

Chapter 1 OVERVIEW AND INTRODUCTION TO TWO-DIMENSIONAL ARRAYS

1.1 INTRODUCTION

Self-assembly at the molecular and nanoscale has been a widely studied area of research for many decades [1-3]. Assembly at the molecular and nanoscale can be accomplished by two major methods: (1) bottom-up and (2) top-down. Bottom-up assembly can be summarized as combining small molecules to create complex structures. A naturally occurring example is the crystalline Surface layers (S-layer) of archaebacteria and bacteria, which are comprised of repeating monomeric proteins that entirely encapsulate the cell. Individual S-layer proteins are synthesized in a bottom-up process by the cell translation machinery, which strings together amino acids in a specific order as shown in Figure 1.1a,b. These proteins then self-assemble into larger arrays comprised of monomers as seen in Figure 1.1c. Figure 1.1d also shows that S-layers can be stripped from the cell wall and reassembled in a self-assembly bottom up process to be used as a template for nanotechnology. This method of self-assembly produces highly-ordered structures with nanometer precision over a micrometer or greater scale with very little effort. Top-down assembly involves the use of tools to cut, mill and shape materials into a desired pattern. Examples of this are nanolithography and MEMS fabrication techniques [4]. Because top-down processes always involve removing material, bottom-up assembly is considered greener and has the potential for large cost savings compared to top-down processes. Bottom-up assembly also has potential for controlling spacing and organization with atomic precision, a feat that is not yet possible with currently available top-down processes.

Biological building blocks that self-assemble -and not just assemble- into predetermined supramolecular structures hold enormous promise for the production of advanced materials,

devices and systems. DNA nanotechnology has dominated the field so far. With this approach, it is possible to arrange nucleic acids via strand exchange or DNA origami into well-defined one, two and three-dimensional nanostructures [5-7]. Moving up in complexity and diversity of the building block, tremendous strides have been made in enabling the fabrication of predetermined supramolecular structures with peptides and peptoids [8-12]. For instance, peptoids were recently engineered to form two-dimensional (2D) sheets that cover micron-distances while being only 3nm thick [12-14]. However, our ability to predict, engineer and control short- and long-range interactions with proteins is lagging behind [15]. This leaves nature's most versatile building block underused.

Naturally occurring proteins assembling into regular 2D arrays (e.g., S-layer proteins and the purple membrane patches of *Halobacteria*) exemplify the versatility and practicality that proteins have to offer for the production of predetermined supramolecular structures. The remarkable stability and unique photochromic properties of purple membrane patches have been exploited for optical information storage and processing [16, 17]. The most widely used 2D proteins are the archeal and bacterial Surface layer proteins (S-layers). S-layers can be harvested and used for many different applications [18] including ultrafiltration membranes [19, 20], matrices for defined binding of functional molecules [21, 22], templates for biomineralization or immobilization of metallic or semiconducting nanoparticles [23-27], etc. However, neither the geometry, nor the chemistry or assembly of these systems can be precisely controlled from nano- to mesoscale. Additionally, because 2D protein arrays are rare in nature, finding one exhibiting a desirable lattice spacing for a specific application is difficult.

Although the development of protein building blocks capable of self-assembling into well-defined architectures lags behind DNA nanotechnology, progress in the computational

design of protein-protein interfaces has accelerated over the past few years [28-31]. For example, interfacial metal coordination [32, 33], engineered disulfide bonds [34, 35], novel α-helical coiled-coil [36, 37] and β-strands [38], and fusions between domains derived from distinct oligomeric proteins [39-41] have all been exploited to create new quaternary structures. Even more exciting is recent success in using computational approaches to design interfaces and protein pairs in which weak, more natural, interactions collectively drive higher order structure formation with the possibility of precise control of position and orientation [42-47].

Here, we push the field forward by designing novel proteins capable of self-assembling into regular 2D protein arrays upon addition of an ionic trigger. More specifically, we set out to identify a naturally occurring protein that could be modified to produce a flat, 2D crystalline layer pierced by a single small pore. To this end, we redesigned the interface of symmetry-compatible oligomeric proteins so as to bring together unit cells into a tiling pattern held together by fusion joints and stabilized by multiple weak interactions. We show robust, controllable and scalable self-assembly into macroscopic 2D arrays exhibiting both short-range (nm) and long-range (10 to 100s of μm) order. In Chapter 6 we will show that this new designer lattice can be functionalized with avidin-FITC, a model protein conjugated to a fluorescent dye for visualization. These experiments lay the groundwork to further explore the potential of designed arrays as templates for inorganic mineralization, scaffolds for high-density enzyme display, and ultimately, as a new class of combined heterogeneous and biological catalysts.

Before diving into this exciting work, we will briefly review the work that has been conducted so far on 2D lattices, starting with naturally occurring two-dimensional proteins and moving on to DNA, peptides and peptoids. This section will be followed by a review paper by Baneyx and Matthaei entitled “Self-assembled two-dimensional protein arrays in

bionanotechnology: From S-layer to designer lattices" [15]. Here, we take a more in-depth look at the versatile S-layer protein and describe recent work focused on computational design of protein lattices. Chapter 3 will discuss computationally designed proteins for self-assembly at greater length. Next in Chapter 4 and 5, we describe the construction of our own 2D protein arrays, moving on to their functionalization in Chapter 6. Chapter 7 covers early work aimed at constructing a designer protein capable of biofabricating quantum dots and binding gold nanoparticles for plasmonic enhancement of the fluorescent signal. The final chapter consists of a paper describing the Department of Chemical Engineering *Distinguished Young Scholars Seminar Series* and its impact on graduate students professional development [48].

1.2 NATURALLY OCCURRING TWO-DIMENSIONAL PROTEINS

Proteins, nature's most versatile building blocks, are still underused for the production of artificial 2D arrays. This is largely due to the highly complex nature of proteins and the need for a thorough understanding of their structure. However, certain naturally occurring proteins capable of self-assembly into regular 2D structures exemplify the versatility and practicality that proteins can offer to the production of supramolecular structures. The two major types of naturally occurring 2D protein arrays come from the purple membrane patches of *Halobacteria* species [49] and the S-layer exoskeleton of nearly all archaea and many bacteria [50]. Both families of proteins have been studied for nearly half a century and have been shown to have many relevant uses in bionanotechnology [16, 51, 52].

1.2.1 *Purple membrane*

The purple membrane was first described in 1967 by Stoecken and coworkers [53]. It is found in patches of *Halobacteria* species and can make up to 80% of the cell membrane. The purple membrane is a 2D crystalline layer of the protein bacteriorhodopsin (bR) [16], consisting

of hexagonally packed bR trimers and lipids at a 1:3 weight ratio of lipids to bR (Figure 1.2). The purple membrane is about 5nm thick and can span up to 5μm as seen in Figure 1.2a. Bacteriorhodopsin has the unique ability to function as an outward light-driven proton pump [54].

The mechanism is well understood: bR consists of seven membrane-spanning helices (A, B, C, D, E, F, and G) with a molecule of the chromophore retinal bound to helix G via a Schiff base and reaching out to helix A in a ground, or all-*trans* state. The proton pump works as follows. First, the retinal absorbs a photon of light, protonates and isomerizes into a 13-*cis* state. This triggers the transfer of a proton to Aspartate-85. The deprotonated retinal straightens out and pushes against the F helix making it tilt. This is known as the 13-*cis* retinal neutral phase. As a consequence, the cytoplasmic side of bR opens up and Asparate-96 can be re-protonated by giving its proton to retinal. Meanwhile, Aspartate-85 transfers its proton through a framework of hydrogen bonded and water molecules past Arginine-82 to the outside of the cell. The complete process is highlighted in Figure 1.3 [55].

Even with such a complex process of conformational changes, purple membranes are remarkably stable. They can handle a range of pH from 0 to 12, high salt concentration (e.g., 3M NaCl), be exposed to sunlight in an air/oxygen environment for years, and survive exposure to over 80°C in water and 140°C in a dry form. Even when dried, the purple membrane's color and photochemical properties are conserved [16]. This stability lends the purple membrane to many applications in optical information storage and processing [16, 17].

1.2.2 *Surface layer protein*

S-layers are monomolecular and regularly ordered 2D assemblies of proteins or glycoproteins located on the outermost part of the cell envelope of many eubacteria and archaea

[56-60]. The smooth outer surface of the S-layer array is hydrophobic and carries a net neutral charge, whereas the corrugated inner surface tends to be hydrophilic and carries either a net negative or positive charge [18]. Individual proteins in these arrays vary in molecular mass from about 40 to 200-kDa and orient themselves into unit cells with dimensions ranging from 3 to 30nm [18]. Figure 1.4 shows that S-layer arrays have been found in oblique (p1, p2), square (p4), or hexagonal (p3, p6) lattice symmetries [60]**.** This diversity makes S-layers valuable systems for nanotechnology and biomimetic applications.

S-layers contribute to many cellular functions including protective coating, cell adhesion and surface recognition, molecular sieving, molecular and ion trapping, as well as scaffolding for enzymes, and virulence factors [57]. Multiple strategies have been developed and refined for harvesting S-layers from various organisms and to reassemble S-layer proteins into crystalline arrays under different conditions and at various interfaces (Figure 1.5) [61-66]. Several genes encoding S-layer proteins have been successfully expressed in *E. coli*, but the majority of S-layer proteins have not yet been cloned. Key parameters in the S-layer disassembly process include the presence and nature of chaotropic agents and chelating agents, the solution pH, and cation substitution. Reassembly appears primarily influenced by protein concentration, composition of the buffer solution, nature of the surface onto which the array is to be reassembled, and the concentration of the added divalent cation [18]. The exact reason why these divalent cations are important in 2D lattice formation is unclear, but it is most likely that they are involved in salt bridges required for intermolecular interactions between S-layer subunits [67, 68]. Any of these factors can cause large differences in the resulting assembly. For instance, changing the concentration and type of divalent cation used during reconstruction of the *Bacillus*

stearohermophilus S-layer lattice leads to transitions from sheets, to cylinders, to morphologically poorly defined structures, to unstructured precipitates [68].

Numerous reviews have been written on the technological applications of S-layers [22, 61, 62, 69-72]. Major applications include isoporous ultrafiltration membranes, supporting structure for functional lipid membranes, drug delivery, drug-targeting, vaccine applications, matrices for the defined binding of functional molecules, templates for biomineralization, or for the immobilization of metallic or semiconducting nanoparticles [18]. An example of supramolecular device fabrication is the biomineralization of CdS nanocrystals within the pores of the *Bacillus stearothermophius* NRS2004/3a variant 1 S-layer [25]. This study showed that the orientation of the array (i.e., whether the inner or outer surface was exposed to the solvent), played a key role in CdS mineralization suggesting that the electrostatic binding of Cd(II) ions is important in site-directed nucleation as shown in Figure 1.6 [25]. An alternative approach for creating inorganic nanoarrays is electrochemical nanofabrication. Here, the S-layer is used as a resist and a potential is applied to nucleate and grow a material of interest (e.g., Cu_2O, Ni, Pt, Pd, Co) within the pores of the S-layer [23, 73-75].

There are however drawbacks to S-layers. First, p4 and p6 S-layers have two pores as can be seen in Figure 1.4. This is undesirable when the properties of a nano-inorganic are dictated by its size (e.g., the emission wavelength of CdS nanocrystals). Second, because S-layers encapsulate cells they have a natural curvature, which depends on the morphology of the host organism [60]. Such curvature can interfere with formation of long-range order on flat surfaces. Third, S-layer outer surfaces are often rough (inner surfaces are typically smooth). This could become problematic if a layer-by-layer 3D assembly process is desired. Fourth, there is only one crystalline structures of an S-layer protein [76] and a few high resolution EM structures are

available [18, 77, 78]. This means a trial and error approach to most S-layer engineering. Finally, most S-layers have been harvested from their native host cells and few S-layer proteins have been cloned and expressed in *E. coli*.

Even with these drawbacks, S-layer proteins have proven to be a valuable tool in nanotechnology. The four decades of research and numerous applications of the S-layer are a testament to this fact. S-layer protein engineering has been mainly done by trial and error [79], but the possibilities that computation brings to the table could be very exciting [15]. A more in-depth look at S-layer proteins including a discussion of crystal structure, reassembly, and applications is presented in Chapter 2 Section 2.

1.3 TWO-DIMENSIONAL DNA STRUCTURES

DNA is normally organized as a double helix made up of two complementary strands. Each strand consists of polymerized nucleotides with a phosphate-deoxribose backbone and stacked H-bonded bases within the center of the helix (adenine [A] paired with thymine [T], and guanine [G] with cytosine [C]). The diameter of the double helix is about 2nm and base pairs are separated by 0.34nm and make a rotation every 3.5nm [80, 81].

A notable property of double stranded DNA is that it melts at high temperature and re-anneals as the temperature is brought back down. The material is normally thought of as flexible but double-stranded DNA comprising 150 base pairs or less acts more like a rigid rod [81]. Single stranded DNA whose persistence length is about 1nm is much more flexible than double stranded DNA and can bend at a junction [81]. Using these properties of DNA along with sequence symmetry minimization rules let Seeman create an immobile nucleic acid junction (the Holliday junction) in 1982 via strand exchange [82, 83]. He used four different hexadecanucleotides where eight of the nucleotides of strand 1 complement eight of the

nucleotides on strand 2 and the other eight nucleotides of strand 1 are complemented by eight nucleotides of the strands on stand 4. Each strand is designed this way, so that strand 1 will hybridize with strand 2 to become double stranded DNA then bend to hybridize with strand 4. Each of the four strands hybridize and bend in this fashion and create a Holliday junction which looks like a plus sign [83]. The original lattices were too flexible and unstable [80, 81, 83] but the problem was solved by using double-crossover tiles, which are single strands of DNA that act as a stabilizing tile to the desired shape, as shown in Figure 1.7a [84]. This was extended to triple-crossovers [85], and paranemic-crossovers [86] and the production of many new 2D structures and even nanotubes.

In 2006, Rothemund invented the DNA origami technology. In the original work, a large 7.4-kb single strand of DNA from the M13 bacteriophage genome was used for hundreds of shorter 'staple' strands to anneal and fold the scaffold into predicted arbitrary structures [87]. These shapes, on the order of 100nm, included a square, a star, and even a smiley face (Figure 1.7c). Scaling DNA origami beyond this size range remains difficult but almost any shape can be made [81]. DNA origami structures have a resolution of 4-6nm and have been used to immobilize heteroelements such as proteins and nanoparticles [88].

Both methods of creating two-dimensional DNA nanostructures can be used to immobilize peptides [89], proteins [90-92], nanoparticles [93-97] and carbon nanotubes [98] some of which give rise to enhanced properties, e.g. enzymatic activity [88, 99]. The technology is becoming so robust that it is starting to enter the third dimension [100, 101]. For instance, Anderson and coworkers have created a DNA box with a functional lid [102]. These results are a testament to how multiple decades of research can lead to major breakthroughs in creating new templates for nanotechnology applications.

However, Pinheiro et al. have highlighted the challenges of structural DNA nantochnology [88] which include scaling 2D DNA arrays from ~1 μm^2 to large areas; self-assembling origami tiles in the way that small motifs do; and using nucleic acid analogs for nanoconstruction [80].

1.4 PEPTIDES AND PEPTOIDS

Next on the complexity scale are peptides and peptoids. Peptides (and proteins) owe their chemical diversity to the use of one of 20 amino acids (Figure 1.8c) or even unnatural ones [103, 104]. On the other hands, peptoids have side chains appended to the amide nitrogen rather than to the α carbon (Figure 1.8a,b). Peptides can and peptoids must be made synthetically through liquid- or sold-phase synthesis. This allows for ease in changing the sequence but also has increased the cost of production.

Self-assembling peptides tend to form fibrilar shapes [105, 106]. For instance, α-helices can form long, stable structures [107] when hydrophobic (H) and polar (P) amino acids are arranged in a $(HPPHPPP)_n$ heptad repeat that allows α-helices to wrap around each other in a left handed direction. These peptides can be programed to pack in a "slipped arrangement" where the two α-helices coil-coil interaction is in an offset arrangement to promote growth into longer fibers (Figure 1.9a).

Amyloid-like structures can also be made with peptides. These structures are also fibrous but tend to be β-structured with β-sheets hydrogen-bonded to each other perpendicular to the long axis of the fiber (Figure 1.9b). Such structures are not as well understood as the α-helices described above but their formation is favored by using a high proportion of β-structure-favoring residues, aromatic side chains, and employing sequence patterns in which H and P residues

alternate as in $(HP)_n$ [106]. Peptides can also self-assemble in amphiphilic [108] or collagen-like assemblies [109, 110] but these all-fibrous structures do not really span two-dimensions.

Peptoids are synthetic compounds that were originally developed in the late-1980's by a small molecule drug-discovery program [111]. Their backbones are identical to that of peptides but side chains are attached to the amide nitrogen (Figure 1.8b). This causes two major changes: first the chirality from the α-carbon is eliminated [112], and second, intrachain hydrogen bonding is lost [113]. This means that peptoids should be more flexible than peptides, in part because tertiary backbone amides could more easily adopt both *cis* and *trans* confirmations [111, 113]. A change in side chain location means resistance to proteases [114, 115]. Also, the variable group dominates structural and chemical properties [116, 117]. The diversity of peptoids is vast because the variable group is chemically added through amine synthon via an S_N2 displacement reaction [117]. This mean any one of hundreds of variable groups can be added to peptoids instead of the twenty naturally occurring variable groups in amino acids.

There are many uses for peptoids including diagnostics [118-120], protein structure mimicry [121, 122], antimicrobial [123], biosensors [124, 125], biomimetics [126] and biomineralization applications, to name a few [127]. There are several reviews on peptoids [111, 128-131] [117] but we focus here on the self-assembly of peptoids described by Zuckerman and coworkers incuding the production of free-floating ultrathin two-dimensional crystal [132] and (2) homochiral superhelices [133]. In the latter work, partially charged amphiphilic diblock copolypeptoid were used to self-assemble superhelix structures with diameters of ~600nm that stretched from 2 to 40μm.

To produce 2D crystals, two complementary amphiphilic, ionic, and periodic peptoids were created. The first has as variable group that is aromatic and hydrophobic (N-(2-phenethyl)

glycine; Npe) flanked by a negatively charged variable group (N -(2-carboxyethyl) glycine; Nce), creating $(Nce–Npe)_{18}$ as shown in Figure 10a. The complimentary strand replaces the negatively charged variable group with a positively charged group (N-(2-aminoethyl) glycine; Nae), creating the polypeptoid $(Nae\text{-}Npe)_{18}$ shown in Figure 10b. When these two peptoids are mixed at equal amounts between 10 to 100μM in an aqueous solution, they self-assemble into large and thin 2D sheets where the peptoids are fully extended in a bilayer (Figure 10). The sheets align themselves by maximizing π–π interactions, maximizing electrostatic attraction and minimizing the exposure of hydrophobic groups [132]. The 2D sheets can immobilize fluorescently labeled streptavidin through a binding peptide [132] and can be made with a single peptoid strand [12]. A key intermediate in 2D sheet formation happens at the air-water interface [134], and further studies have been conducted at the air water interface [14]. The sheets that form are only ~3nm in height and can span ~100 microns.

These distances far surpass that of two-dimensional DNA structures. While the pegboard characteristics of DNA origami or the immobilization of nanoparticles by the S-layer have yet to be obtained by this peptoidal two-dimensional array, this is still a growing field

1.5 CONCLUSION

We have described here the usefulness of naturally occurring 2D proteins for technological applications. Yet the diversity of naturally occurring 2D lattice is small and they are difficult to assemble over long distances (micron to millimeter). Two-dimensional DNA structures are remarkable from a computational standpoint but lack in overall size and chemical diversity and their scale-up might be expensive (as will be that of peptide and peptoid-based structures). Although peptides hold promise for producing fibrous materials, they do not yield well-defined 2D shapes. Peptoids have been used to create large arrays that hold promise but

functionalization and application are still in the preliminary stages. However, there is really only one class and shape of 2D peptoids. In the end, naturally occurring 2D proteins still seem to be the best candidate for producing 2D arrays from a durability, cost and functionalization standpoint. For these reasons, we discuss S-layer proteins further in Chapter 2 and computational protein design in Chapter 3.

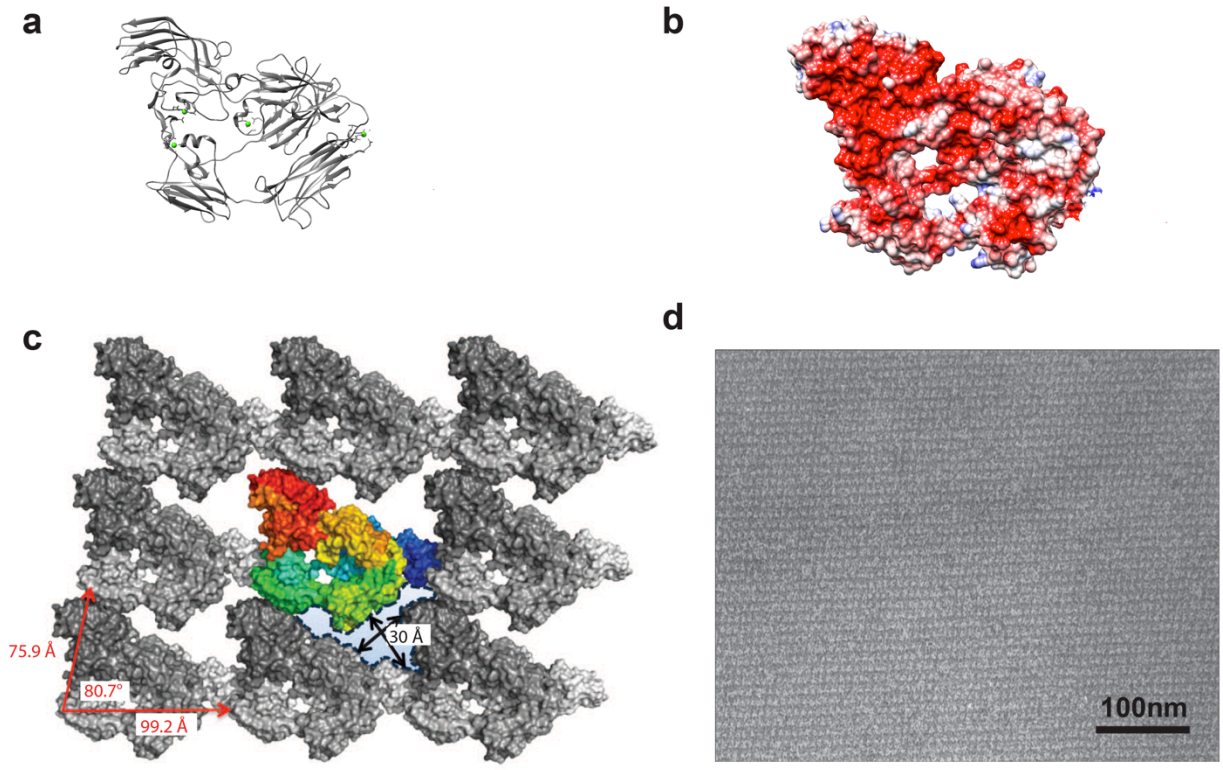

Figure 1.1. Bottom up and self-assembly. In a bottom up manner, the cell machinery assembles the series of amino acids that specifies a folded protein, which, upon addition of calcium ions, assembled into a S-layer protein (in this case that of *Geobacillus stearothermophilus* SbsB). Shown are the monomer ribbon structure **(a)** and molecular surface colored for, positive charge in red and negative in S **(b)** [pdb number 4AQ1]. SbsB self-assembles into an S-layer **(c)** [135] made up of 7.6nm by 9.9nm monomers that create 3nm asymmetric. The central multi-colored protein shows the 6 different domains present in SbsB .S-layers can grow to 1-2µm in size **(d)** [136].

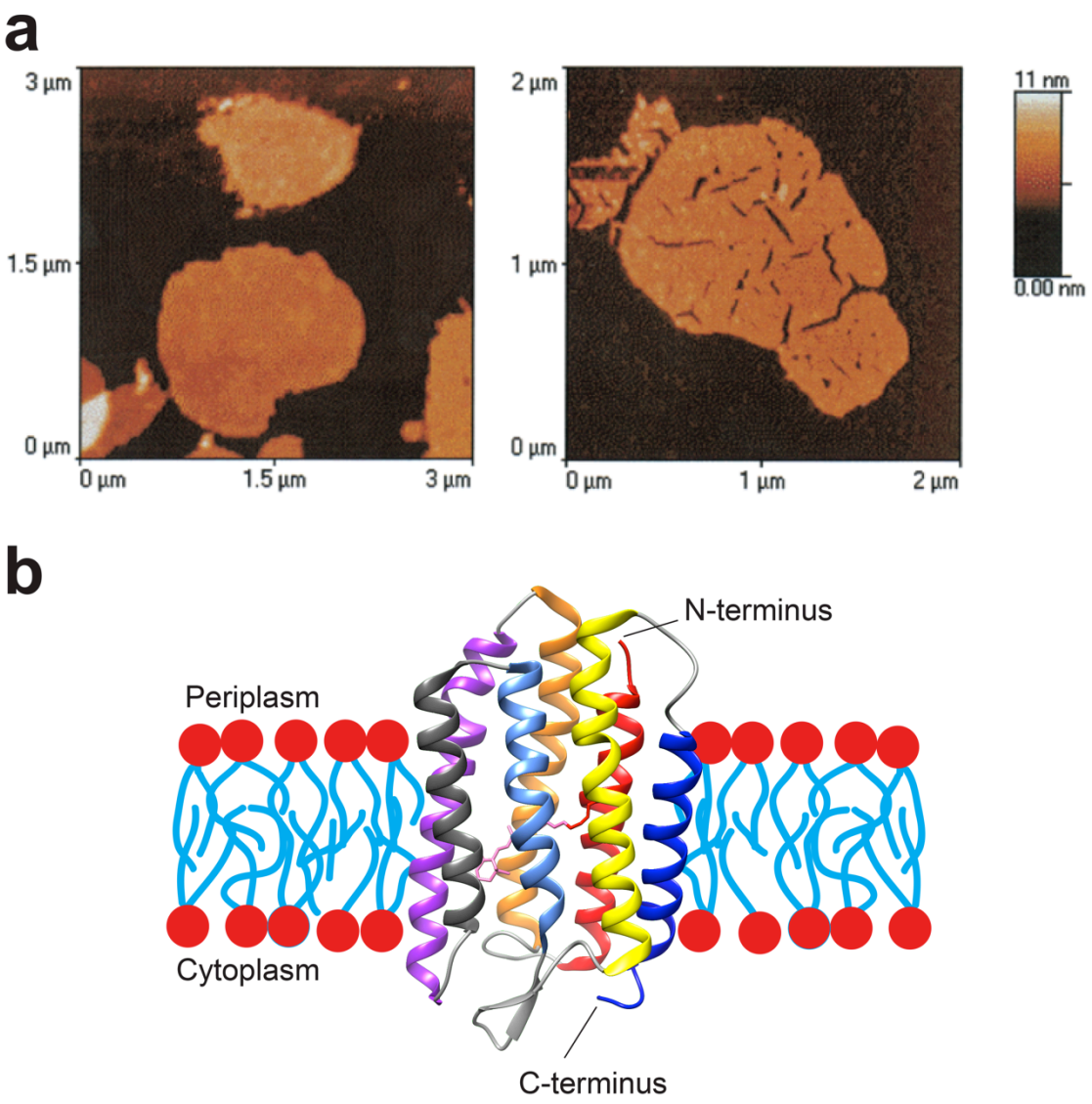

Figure 1.2. Structure of the purple membrane. The purple membrane is one of the rare 2D structures made of proteins. **(a)** These membranes can grow to micrometer scale, as visualized here by AFM [16]. **(b)** They are comprised of lipids and bR, (pdb file 1FBB) which works as a light-driven proton pump and consists of seven transmembrane helices. The retinal chromophore is shown in pink. The C-terminus points toward the cytoplasm and the N-terminus towards the periplasm.

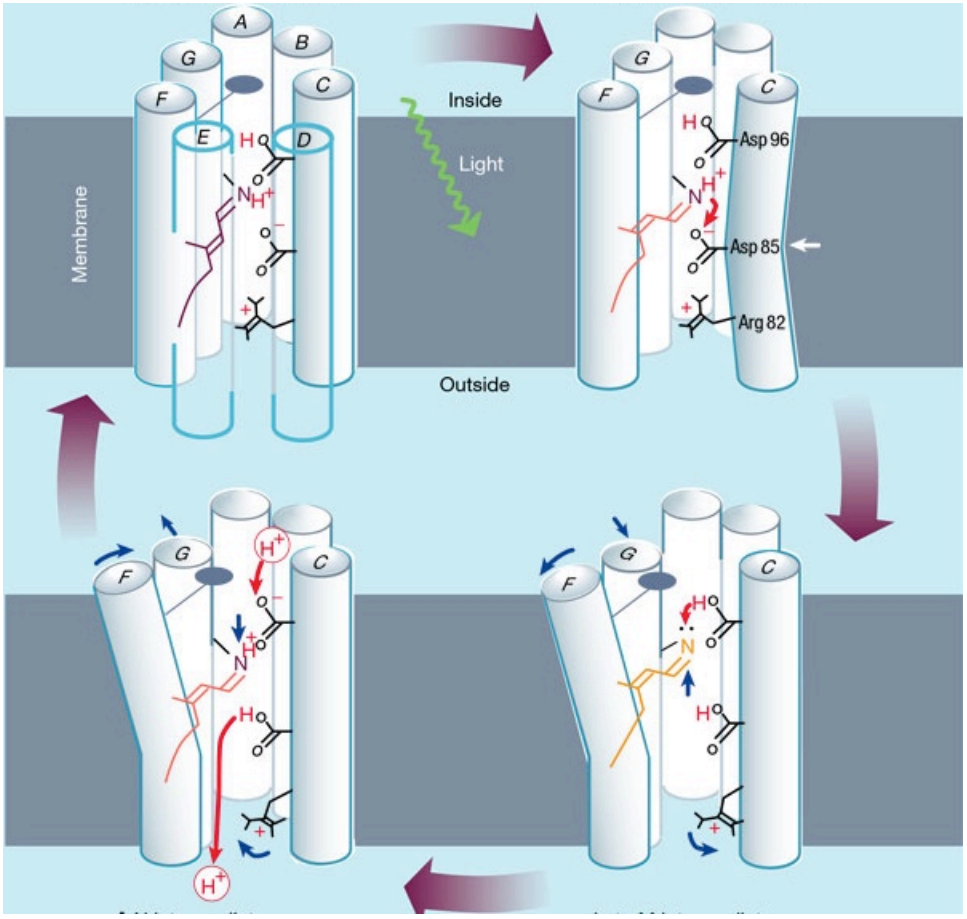

Figure 1.3. Bacteriorhodopsin proton pumping cycle. The cycle begins when the retinal chromophore absorbs a photon of light. Retinal transfers a proton to Asp 85 that hands it off to the outside of the cell. Retinal receives a proton from Asp 96 returning to ground state and Asp 96 takes a proton from the cytoplasm [55].

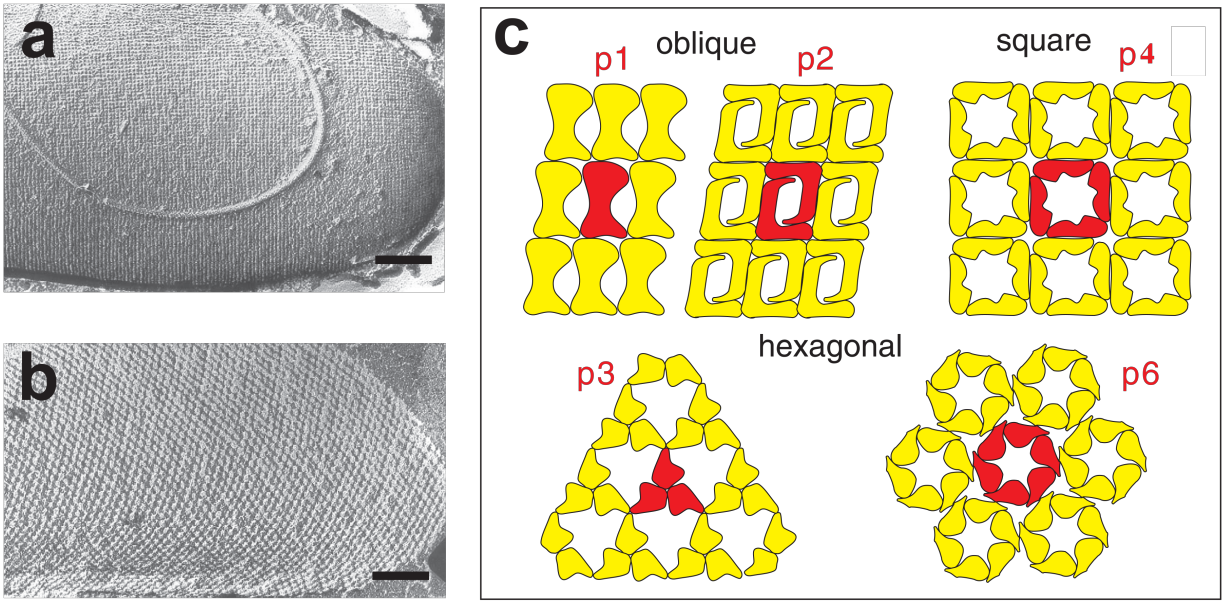

Figure 1.4. Lattice symmetry of S-layers. Freeze-etching preparation of whole cells of **(a)** *B. sphaericus*, showing a square S-layer lattice and; **(b)** *Th. Thermohydrosulfuricus* revealing a hexagonally ordered array. Bar in (a) 200nm and in (b) 100nm [60]. **(c)** Schematic drawing of different S-layer lattice types. The regular arrays exhibit either oblique (p1, p2), square (p4), or hexagonal (p3, p6) lattice symmetry. The morphological units (shaded in red) are composed of one, two, three, four or six identical subunits [61].

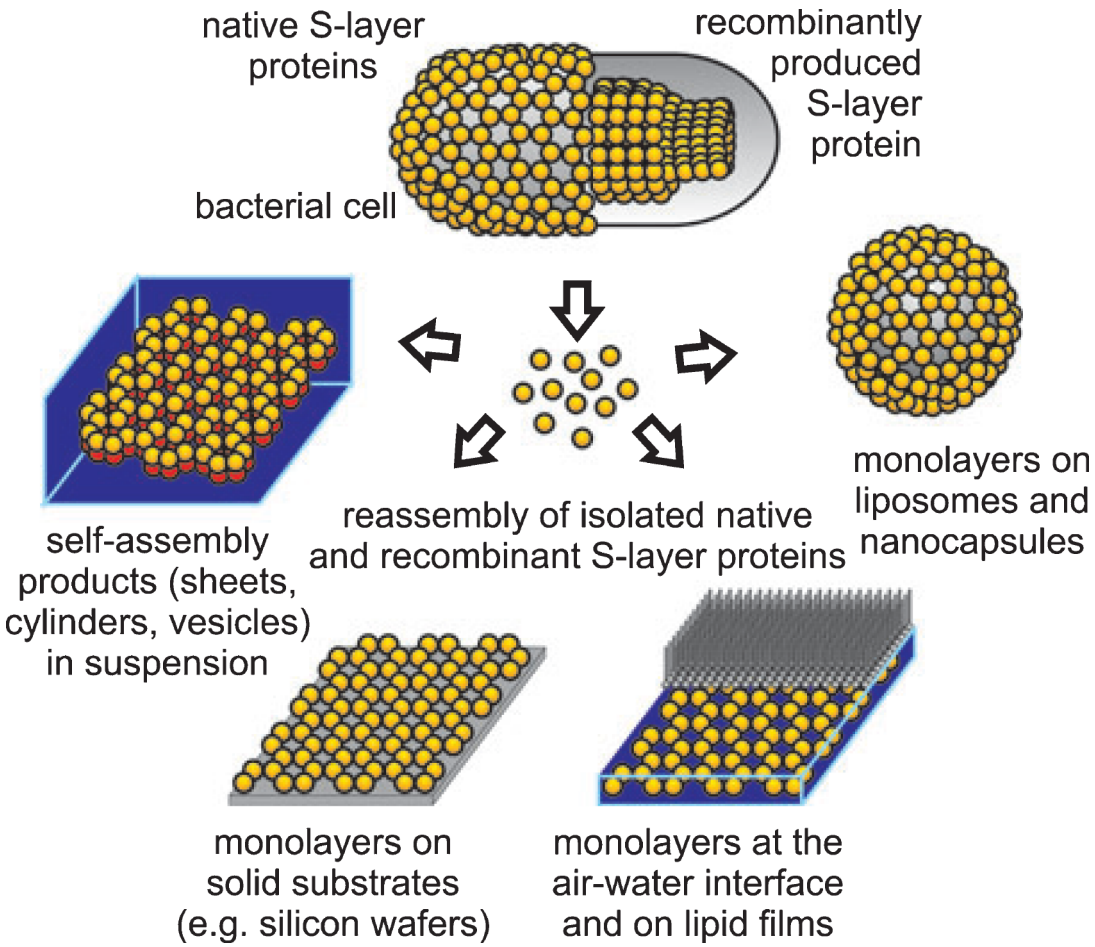

Figure 1.5. S-layer purification and reassembly. Schematic drawing of the isolation of native S-layer proteins from bacterial cells and the reassembly of native and recombinant S-layer proteins into crystalline arrays in suspension, on a solid support, at the air-water interface and on a planar lipid film, and on liposomes or nanocapsules. Adapted from reference [61].

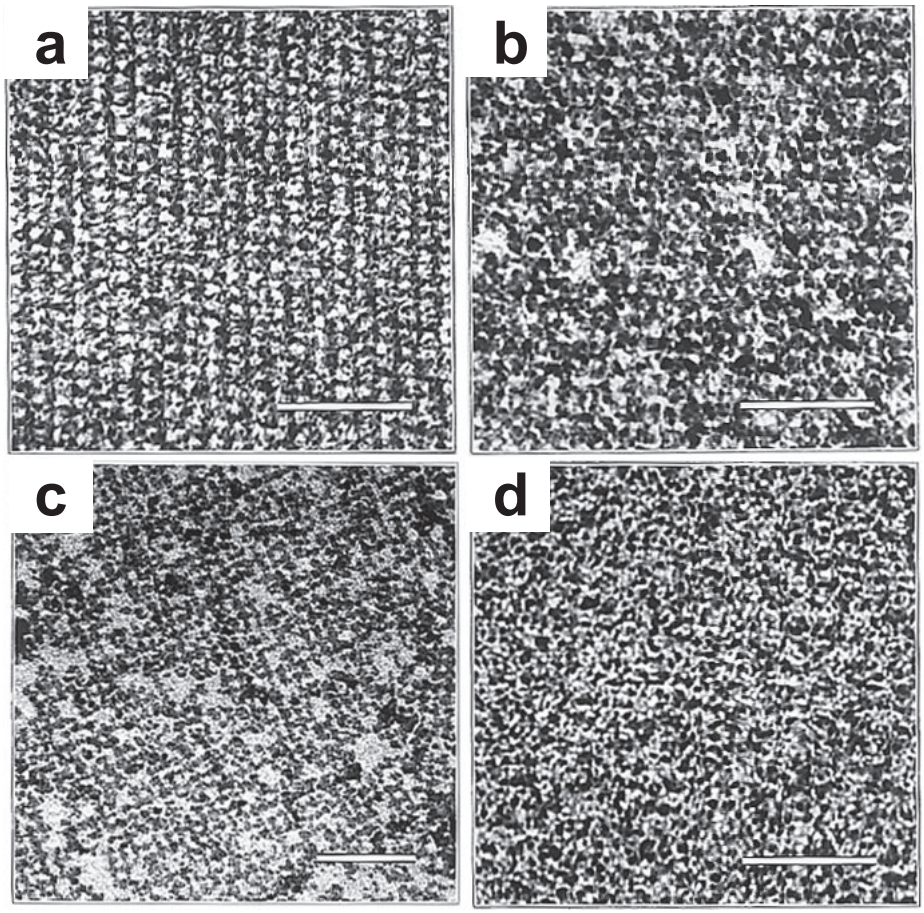

Figure 1.6. S-layer-templated mineralization. (a) TEM image of negatively stained self-assembled S-layer of *B. stearothermophilus* NRS 2004/3a variant 1; scale bar, 60nm. **(b)** As in (a) but without staining and after CdS mineralization of the exposed inner face. The TEM image shows an oblique periodic array of uniform 5nm-size CdS nanocrystals; scale bar, 60 nm. **(c)**. **(d)**, TEM micrograph of the mineralized outer face of an S-layer from *B. stearothermophilus* NRS2004/3a variant 1, showing CdS nanoparticles associated with the protein surface. The array shows weak periodicity (more readily observed by viewing the micrograph on its side); scale bar, 60 nm. Adapted from reference [25].

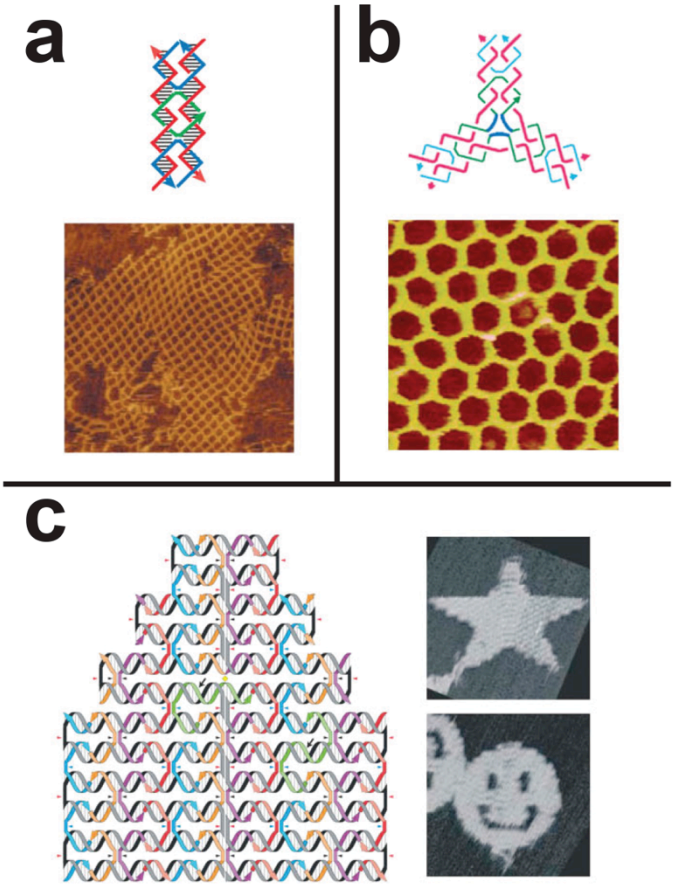

Figure 1.7. 2D DNA nanostructures. Examples of DNA strand exchange are found in (a) and (b). **(a)** is the example of the double-crossover tiles for stabilization. The main strands are shown in red and the double crossover tiles are shown in blue or green. The visualization of this pattern can be seen by AFM [84]. This theory was expanded on, **(b)** shows the more complex three-point star motif that assembles in a hexagonal lattice, visualized by AFM [137]. **(c)** is an example of DNA origami where a large single strand of DNA, red, is 'stapled' by small single strands of DNA. This stapling process can result in almost any arbitrary structure. This figure was adapted from references [81, 87].

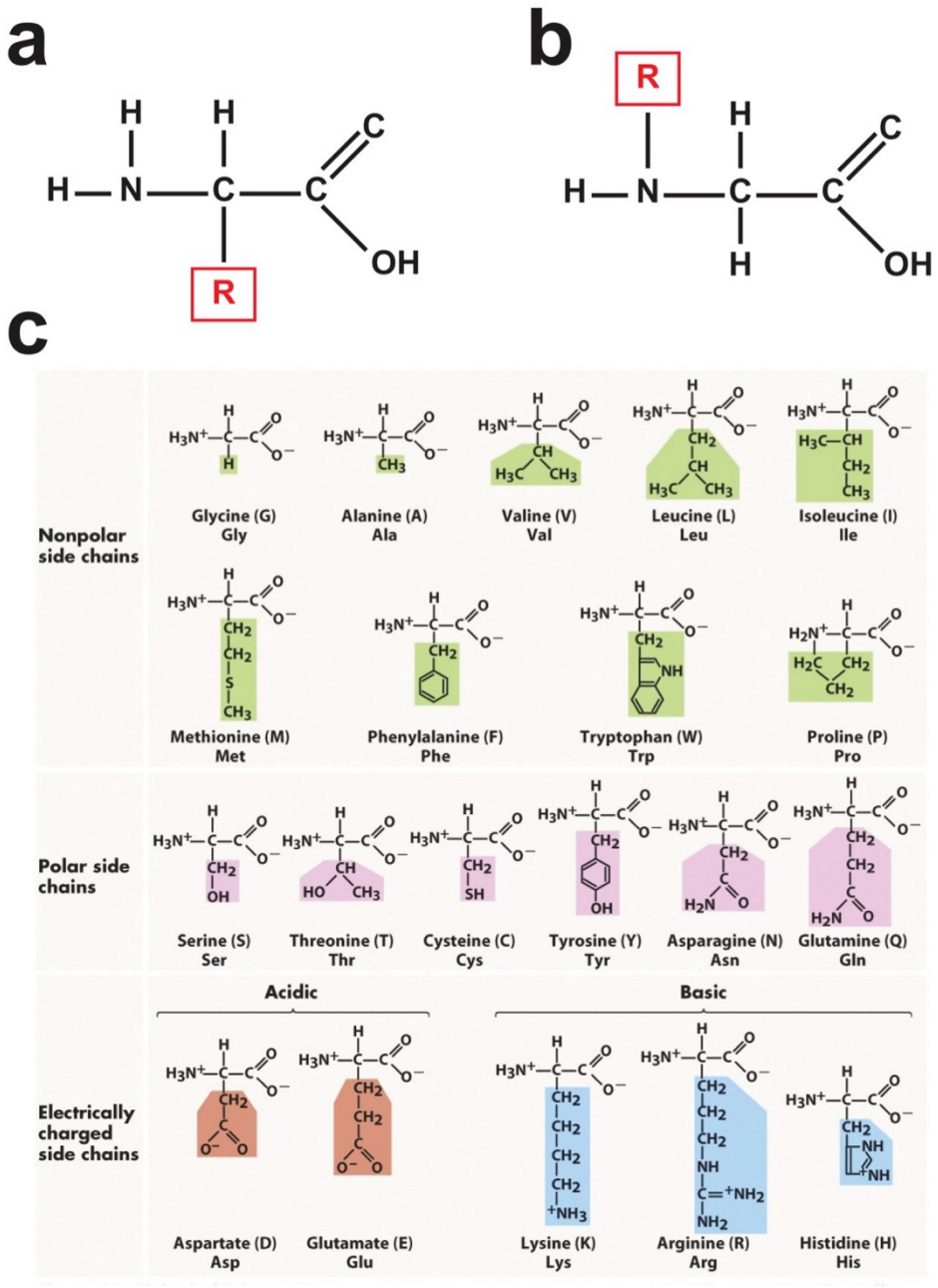

Figure 1.8. Peptides and peptoids. (a) Amino acids, the building blocks of proteins, are made up of an amino group linked to an α-carbon linked to a carboxyl group. The α-carbon is also linked a hydrogen atom and a side chain, R. **(b)** A peptoid monomer resemble amino acids but the variable group is appended to the amide nitrogen rather than to the α carbon. **(c)** The variable group for amino acids is limited to the 20 naturally occurring variable groups, but peptoids have access to hundreds of variable group chemistries [138].

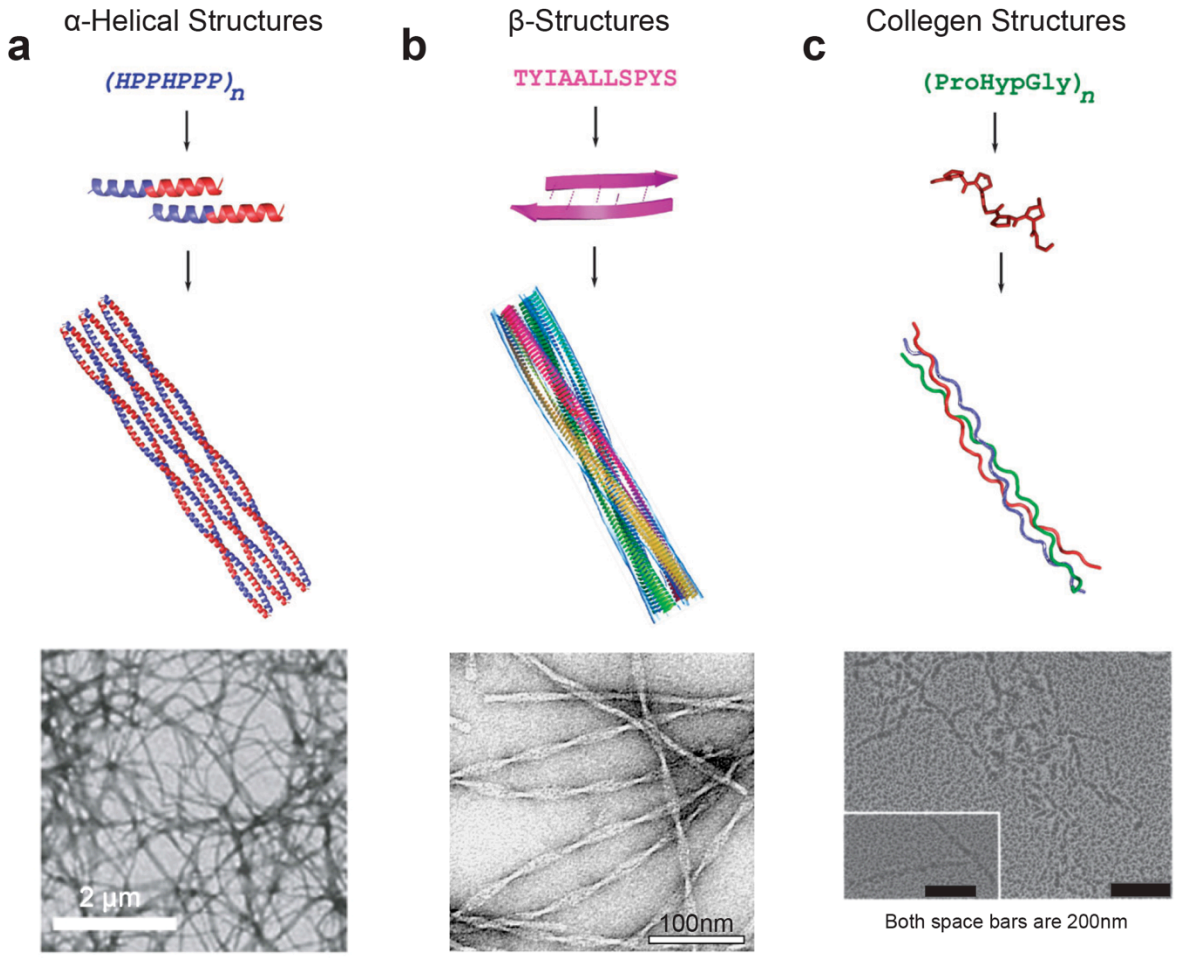

Figure 1.9. Peptides tend to form long fibrous structures. (a) Shows α-helices using the heptad repeat $(HPPHPPP)_n$ where H is hydrophobic residues and P is polar residues in order to induce stabilizing α-helices coiling in a staggered way. This staggering leaves a sticking end to induce further growth of the fiber. α-helical structures depicted here by TEM have a 2nm diameter and extend for several micrometers. This type of construct can also lead to structures ~50nm in diameter [139]. **(b)** Small β-sheets hydrogen bond to each other perpendicular to the long axis of the fiber. These fibers can twist around each other cause larger thicker strands. TEM images shows that β-structures tend to be more of a ribbon structure with a width of 2-4nm and extend for hundreds of nanometers [140]. **(c)** Collagen style assembly comes from three strands of ProHypGly, which form polyproline-type helices, coming together in right-handed triple helices. These helices will continue to assemble to form collagen-like fibers. TEM image shows this type of construction leads to fibers of about 20nm in length [109]. This figure was adapted from [105, 106].

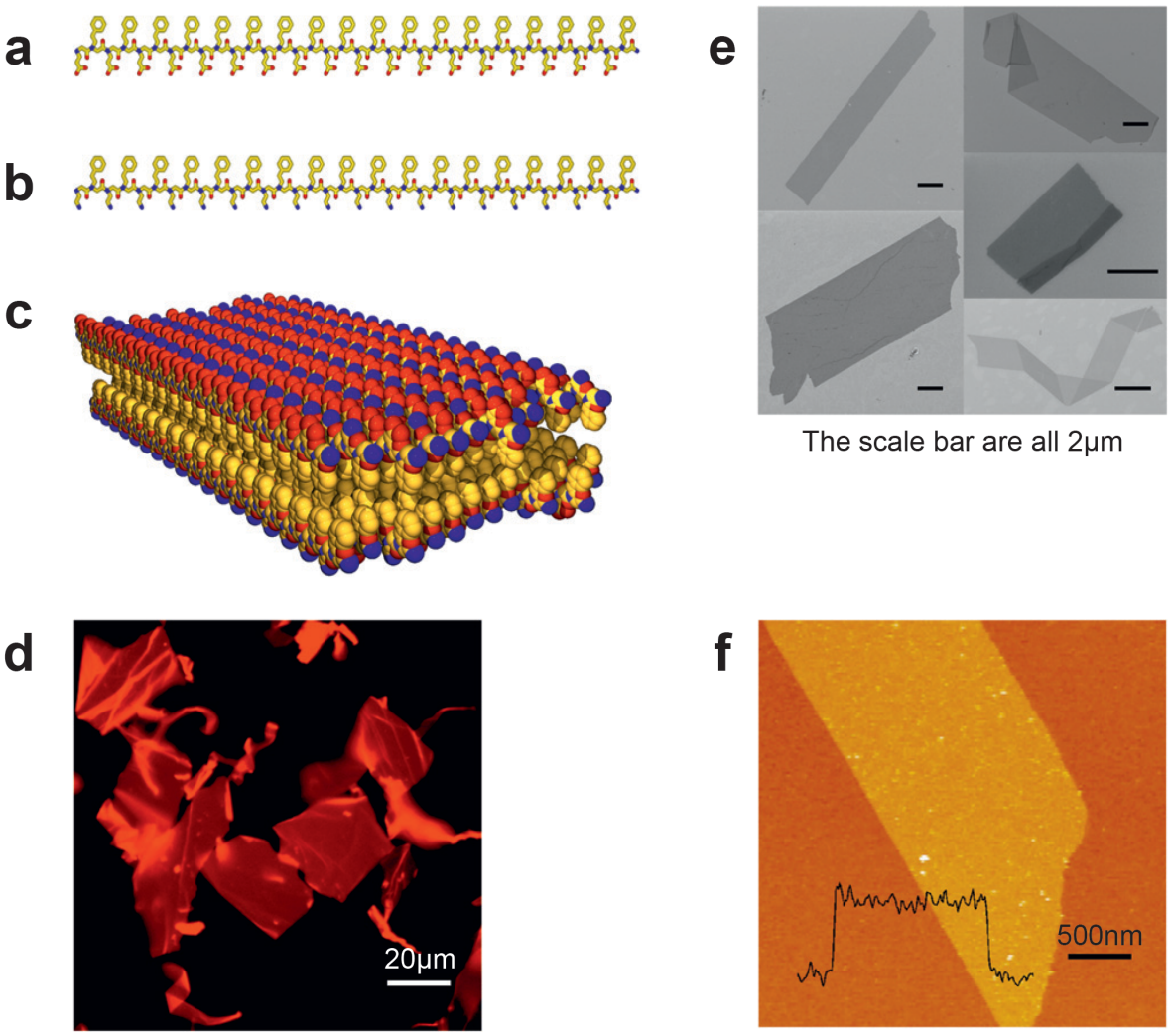

Figure 1.10. Free floating two-dimensional peptoid structures. Two peptoid where **(a)** is a negatively charged amphiphilic peptoid and **(b)** is a positively charge amphiphilic peptoid. When the two are mixed hydrophobic imteractions drive the formation of thin 2D arrays. These sheets can be visualized by fluorescent optical microscopy **(d)**, upon staining with Nile Red; SEM **(e)**, and AFM **(f)** images from sheets on silica (Z range: 20nm). This figure was adapted from [132].

Chapter 2 SELF-ASSEMBLED TWO-DIMENSIONAL PROTEIN ARRAYS IN BIONANOTECHNOLOGY: FROM S-LAYER TO DESIGNED LATTICES

2.1 ABSTRACT

Although the crystalline S-layer arrays that form the exoskeleton of many archaea and bacteria have been studied for decades, a long-awaited crystal structure coupled with a growing understanding of the S-layer assembly process are injecting new excitement in the field. The trend is amplified by computational strategies that allow for *in silico* design of protein building blocks capable of self-assembling into 2D lattices and other prescribed quanternary structures. We review these and other recent developments toward achieving unparalleled control over the geometry, chemistry and function of protein-based 2D objects from the nanoscale to the mesoscale.

2.2 INTRODUCTION

Biological building blocks that self-assemble into predetermined supramolecular structures are of considerable interest in bionanotechnology where an ability to control shape, size, geometry and surface chemistry is crucial to the production of advanced materials with tailored properties. Predictive control of shape has been particularly effective with nucleic acids where a variety of one, two and three-dimensional (3D) nanostructures have been produced via strand exchange and DNA Origami technologies [5]. Peptides and peptoids (polymers whose constituent monomers resemble amino acids but have side chains appended to the amide nitrogen rather than to the a carbon) have also been engineered to assemble into 2D structures [9, 11, 141]. Compared to these molecules, proteins offer a richer and more versatile structural,

chemical and functional palette that can be further expanded through rational design, selection and directed evolution.

Two-dimensional (2D) protein arrays are of particular interest in bionanotechnology because they allow for the high-density display of peptides and proteins in sensor, diagnostic and vaccine applications. They also enable the periodic organization (or templating) of inorganic particles with nanoscale control of position for plasmonic, opto-electronic, magnetic and catalytic applications. In nature, 2D protein arrays are only found in the purple membrane patches of *Halobacteria* species [49], and the surface (S-) layer exoskeleton of nearly all archaea and many bacteria [50]. Here we will not discuss the purple membrane – a crystalline assembly consisting of trimers of bacteriorhodopsin tightly packed in a lipid-containing hexagonal array – since its structure and potential for optical applications have been reviewed elsewhere [142, 143]. Instead, we focus this review on recent developments in our understanding of S-layer structure-function relationship and on progress in the computational design of entirely new kinds of protein arrays.

2.3 S-LAYER

2.3.1 *Structure*

S-layers are monomolecular lattices of (glyco)proteins that encapsulate certain bacteria and archaea and connect to the cell surface through one or several N-terminal glycan-binding domains. Their function ranges from protective coating, cell adhesion, surface recognition, molecular sieving and ion trapping, to scaffolding for enzymes and virulence factors [50, 144]. S-layers are 5-to-20 nm thick in bacteria and up to 70-nm thick in archaea, have a smooth, hydrophobic outer surface with net neutral charge, and a corrugated inner surface that tends to be hydrophilic and carries either a net negative or positive charge [145]. Individual S-layer proteins

have molecular masses between 40 and 200-kDa and form morphological units composed of one, two, three, four or six subunits which assemble with oblique (p1, p2), square (p4), or hexagonal (p3, p6) 2D rotational symmetries (Figure 2.1A-B) [145]. Center-to-center unit spacing ranges between ≈ 5 and 30 nm, and two or more classes of 2-to-6 nm pores typically perforate the array.

2.3.2 *Technological uses*

Crystalline patches of S-layer proteins can be stripped from bacteria and archaea via detergent extraction or by using other agents that disrupt their interaction with the cell wall, and directly used for practical applications [146]. In some cases, S-layer proteins can be expressed in heterologous hosts, such as *Escherichia coli,* unfolded by GuHCl or urea treatment, and re-assembled by dilution or dialysis [50]. Recrystallization from unfolded subunits is most reliably performed at the air-water interface using a Langmuir-Blodgett trough, but is also possible on the surface of zwitterionic lipids and certain technologically-relevant substrates such as silicon, carbon and metals. Reassembled S-layers are a mosaic of well-ordered domains that range in size from about 100 nm to 1-2 μm. However, interdomain dislocations and gaps are not uncommon. The reassembly process is influenced by protein concentration, buffer composition, identity of the surface or interface onto which the array is reassembled, and nature and concentration of added divalent cation, which can induce reassembly transitions from sheets, to cylinders, to morphologically poorly defined structures [147].

S-layers have been evaluated for myriad applications including ultrafiltration membranes, drug delivery systems, scaffolds for immunogen display, and substrates for the spatial organization of functional molecules, metals, and semiconducting nanoparticles (for recent reviews, see references [79]• [144, 146]). Because they have a large void content (30-70% porosity), they can also be used to template the synthesis of inorganic structures conformational

to the geometry of the pores. Examples include the precipitation of CdS nanoparticles within the pores of the *Bacillus stearothermophilus* S-layer via solution chemistry [148] and work from our own laboratories showing that the *Deinococcus radiodurans* and *Sporosarcina ureae* S-layers can be used as a resist to template the electrodeposition of a broad range of materials including Pt, Ni, Co and Cu_2O (Figure 2.1C-E) [149, 150]. Although such inorganic nanostructures hold promise for catalysis and opto-electronics applications, they are difficult to produce at large scales with the current S-layer "patch" technology. Furthermore, the presence of two or more types of pores in an array means that structures of distinct size and shape are also generated. This could be an enormous advantage if one could fill them with two separate materials. However, achieving this – let alone mineralizing crystalline materials within the confined environment provided by the pores — has so far proven elusive. Other S-layer idiosyncrasies (e.g. the fact that inner and outer faces have different topography and chemistry) may complicate the control of surface interactions and interfere with in-registry stacking, thus making the production of 3D structures challenging.

2.3.3 *Insights from the SbsB crystal structure*

Although Baumeister and coworkers produced out-standing low-resolution TEM reconstructions of S-layers in the mid-eighties [151, 152], X-ray crystallography attempts have long been thwarted by the difficulty of obtaining 3D crystals from proteins that evolved to self-assemble into 2D lattices [135]••. As a result, topological information has been obtained the hard way. For instance, in a bid to identify surface-exposed amino acids in the 98-kDa *Geobacillus stearothermophilus* SbsB protein, Howorka and coworkers created 75 cysteine substitution mutants and screened the solvent accessibility of these residues in both the monomeric, dimeric and assembled forms of the protein [153, 154]. This heroic effort established that amongst 23

residues that were highly accessible in the monomer, 8 were interfacial 10 were located on the inner face of the S-layer and 5 were on its outer face.

The long-awaited crystal structure of an S-layer protein (*G. stearothermophilus* SbsB) has validated this work and provided major new insights on the mechanism by which S-layers form [135]••. SbsB consists of an unresolved cell wall attachment domain followed by 6 immunoglobulin-like domains organized into a disk shaped like the Greek letter j [135]••. Four Ca^{2+} ions stabilize inter-domain and intra-domain contacts and are critical for quaternary structure acquisition (Figure 2.2A). Indeed, unfolded SbsB only reassembles into an array if the buffer is supplemented with calcium ions, and while individual domains retain secondary structure upon EDTA treatment, all quaternary structure is lost [135]••. In short, Ca^{2+} binding induces a structural transition that alters the conformation of SbsB monomers from extended to a j-shaped structure that is assembly-competent. The main contacts stabilizing the oblique (p1) lattice are between domains IV and VII of each monomer and domains II and IV of neighboring subunits. These contacts bury comparable surface area (450 $Å^2$ and 490 $Å^2$, respectively; Figure 2.2A). Interestingly, domain II, which orients the cell wall-binding domain I toward the cell side, appears to be able to adopt multiple conformations in order to better handle topological defects on the cell surface. Overall, the SbsB structure provides a good explanation for the essential role of calcium in the S-layer assembly process and valuable insights for S-layer engineering.

2.3.4 *Reassembly at surfaces*

Recent *in situ* AFM imaging studies of the assembly of *Lysinibacillus sphaericus* SbpA at interfaces [155]•• [156] have provided complementary insights on the mechanism by which S-layers nucleate and grow. SbpA is a 132-kDa protein that readily self-assembles into lattices exhibiting square (p4) symmetry and composed of tetrameric morphological units. As with SbsB,

there is evidence that SbpA protomers have different conformationsin monomeric and tetrameric states [157] and lattice self-assembly depends upon calcium addition. The assembly of SbpA on supported lipid bilayers is schematically depicted in Figure 2.2B and involves the following steps: (i) adsorption of monomers in an extended conformation from the bulk solution and onto the surface (step 1); (ii) nucleation of amorphous clusters that are uniform in height in a process that is slow because it requires significant adsorbed protein surface coverage (step 2); (iii) comparably rapid (5-10 min) rearrangement into more compact crystalline clusters composed of SbsB tetramers (step 3); and (iv) extension of the clusters in all directions of the plane by addition of tetramers at crystalline cluster edges with a preference for sites that have a large number of nearest neighbors (steps 3 to 5) [155]•• . Of particular interest, new tetramers were only detected at the perimeter of growing clusters. This suggests that interactions with the crystalline phase converts vicinal monomers and/or loosely folded oligomers (dimers and tetramers?) into compact tetramers that can be incorporated into the growing lattice. In a separate study of the reassembly of the same protein on mica [156], De Yoreo and coworkers were able to detect kinetically trapped clusters that are crystalline in nature but more loosely packed than the lower energy compact arrays that they eventually transform into (Figure 2.1B, steps 3' to 5) [156]. While these studies tell us little about the process of S-layer formation *in vivo*, they have strong practical implications. Indeed, understanding the mechanics of the crystallization process and how it is influenced by the interface will be key to the production of uniform structures exhibiting nanoscale periodicity and very long-range order.

2.3.5 *S-layer protein engineering*

Because no crystal structure of a S-layer protein was available until recently [135]••, most protein engineering efforts have been conducted by trial and error and have focused on a

limited number of candidates [79]•. A fragment identified by deletion analysis and spanning the first 1037 amino acids of mature *L. sphaericus* SbpA ($SbpA_{31-1068}$) [158], has been especially useful as an N-terminal fusion partner because it remains capable of self-assembling in a crystalline lattice of p4 symmetry when various peptides and proteins are fused to it [79]•. This has been exploited to build immobilization platforms, sensor heads and immunogens in which the fused detection moiety, ligand, enzyme or antigen/allergen is displayed with high density and, in a repetitive geometry on the reconstituted crystal [79]•. Interestingly, a fusion protein between $SbpA_{31-1068}$ and a short tag of sequence GSLCTPSRLEHHHHHH was recently reported to form bilayers of stacked monolayers [159]. When considering this result with the fact that truncation of 150 additional residues from the $SbpA_{31-1068}$ C-terminus converts the lattice space group symmetry from square to oblique [158], it becomes clear that modifications at the C-terminus can have quite unexpected consequences on quanternary structure.

2.4 BEYOND NATURE: ARTIFICIAL 2D PROTEIN ARRAYS

Although the development of protein building blocks that self-assemble into well-defined architectures lags behind DNA nanotechnology, progress in the computational design of protein-protein interfaces has accelerated over the past few years [160, 161]. For example interfacial metal coordination [162, 163], disulfide bonds [164], novel α-helical coiled-coil [165-167] and β-strands [168], and fusions between domains derived from distinct oligomeric proteins [169-171] have all been exploited to create new quaternary structures. Even more exciting is recent success in using computational approaches to design entirely new interfaces and orthogonal protein pairs in which weak interactions collectively drive higher order structure formation with the possibility of exquisite control over position and orientation [172-174] [45]••.

An attractive alternative to the engineering of natural S-layer proteins is therefore to computationally design protein building blocks that self-assemble into extended periodic arrays. One approach to do so is inspired by early work by the Yeates lab [170] and relies on the construction of genetic fusions between peptide chains derived from separate oligomeric proteins that exhibit rotational symmetry axes of equal order [175]. As an example, Sinclair *et al.* [171]• fused the Streptag octopeptide to the C-terminus of *E. coli* aminolevulinic acid dehydrogenase (ALAD; a tetrameric protein) and added Streptavidin to the purified protein to "stitch" a 2D lattice (Figure 2.3A). The same group fused the Lac Repressor-derived Lac21E and Lac21K peptides [176] to ALAD, and exploited the formation of coiled-coils to assemble a 2D lattice (Figure 2.3B). These groundbreaking results were mitigated by the fact that the resulting structures lacked long-range order and were as small as a few hundred nanometers for the Lac21 fusions and at most a few micrometers for Strepavidin-assembled lattices. Furthermore, the material produced was reportedly sensitive to heat and small molecule inhibitors.

An alternative pioneered by the Tezcan group relies on engineering metal coordination sites at the surface of proteins for cation-driven assembly of multidimensional structures [163]. Recently, Brodin and coworkers [162]• computationally redesigned the surface of cytochrome cb_{562} (cyt cb_{562}, a four-helix bundle heme protein) to allow for the formation of **C2**-symmetric dimers through a high affinity Zn^{2+}-binding site, while simultaneously allowing for dimer self-assembly in two perpendicular directions via a low affinity Zn^{2+}-binding site (Figure 2.3C). By controlling the Zn^{2+} to protein ratio and the protonation state of the zinc-coordinating histidine residues, the team showed that it was possible to form 2D nanoplates under conditions of slow nucleation (low pH and/or zinc) or helical nanotubes under conditions favoring fast nucleation events (high pH or low pH and high zinc concentration). Protein sheets were observed upon

reordering of precipitates after 5- 7 days of incubation. These sheets consisted of non-perforated squares or rectangles with characteristic dimensions varying between 5 and 10 μm. Whether this engineered protein (called RIDC3) can be further manipulated to incorporate guest peptides or fusion partners without losing its ability to assemble remains unclear, as is the robustness and scalability of the self-assembly process.

2.5 CONCLUSION AND THE ROAD AHEAD

The availability of an S-layer protein crystal structure combined with recent advances in computational structural biology has opened the door to a two-pronged approach for the production of next generation 2D protein arrays. More than four decades of S-layer work by Uwe Sleytr and other laboratories all but guarantees that the SbsB structure (and hopefully that of other S-layer proteins) will be exploited to optimize the location of fusion joints and help fine-tune existing designs for robustness and usefulness in real world settings. However, an incomplete understanding protomer folding and assembly pathway might frustrate complex redesign attempts.

Computationally designed protein arrays remain in their infancy but they offer tremendous opportunities for engineering any and all desirable features on a nearly blank slate. These opportunities include lattice periodicity and symmetry; location of insertion and fusion joints; number, size, shape and chemistry of the pores; and corrugation, composition and electrostatics of “top” and “bottom” surfaces. Exciting short-term prospects include the redesign of protein interfacial contacts to produce building blocks that self-assemble via collective interactions and with atomic-level accuracy, the design of lattices that exhibit arbitrarily specified properties with nanoscale periodicity through the use of multiple such blocks, and the construction of 3D structures made up of stacked 2D arrays.

With artificial arrays, it should also be possible to take full advantage of the promise of solid binding peptides (SBPs), which are short sequences of amino acids that have been selected by combinatorial display for an ability to bind to inorganic or synthetic materials [177, 178]. Insertion of SBPs at computationally defined, solvent-exposed locations should provide unparalleled control over array immobilization, orientation, and conceivably actuation, on virtually any surface(s) or interface(s) of technological interest. Because certain SBPs have the ability to promote inorganic mineralization [179], their incorporation within the architecture of custom-designed pores could be used to manufacture inorganic nanostructures of predictable size, shape, composition, and perhaps crystallinity over extended distances. Finally, it should be easier to access orthogonal chemistries for array derivatization by encoding non-natural amino acids in computationally designed protomers. (Although this could also be done with recombinant S-layer proteins, and especially when using new expression hosts [180], genetic instability and low-level expression issues might complicate the prospect).

Whether based on S-layers or computationally designed, 2D protein arrays face similar challenges. To make a technological difference, it will be necessary to reduce expression and purification costs and to develop robust technologies for the assembly and manipulation of very large lattices that retain order at the nanoscale. The size needed will of course depend on the application. For instance, while mosaicity should be acceptable in the case of catalytic and structural nanoarrays, long-range order in the hundreds of micrometers to millimeters might be desirable for photonic, plasmonic and magnetic applications. The production of such defect-free crystals will likely require a detailed understanding of how environmental and process conditions influence assembly and growth. It will also be important to identify application spaces where protein-based 2D structures offer unique advantages over those fabricated by chemical self-

assembly and top-down writing and patterning processes. It will be even more interesting to integrate these technologies in the pursuit of disruptive materials, systems and devices [181].

2.6 ACKNOWLEDGEMENTS

JFM gratefully acknowledges financial support from NIH through a Cancer Nanotechnology Training Grant (T32CA138312). This work was supported in part by the Office of Naval Research (BRC-11123566).

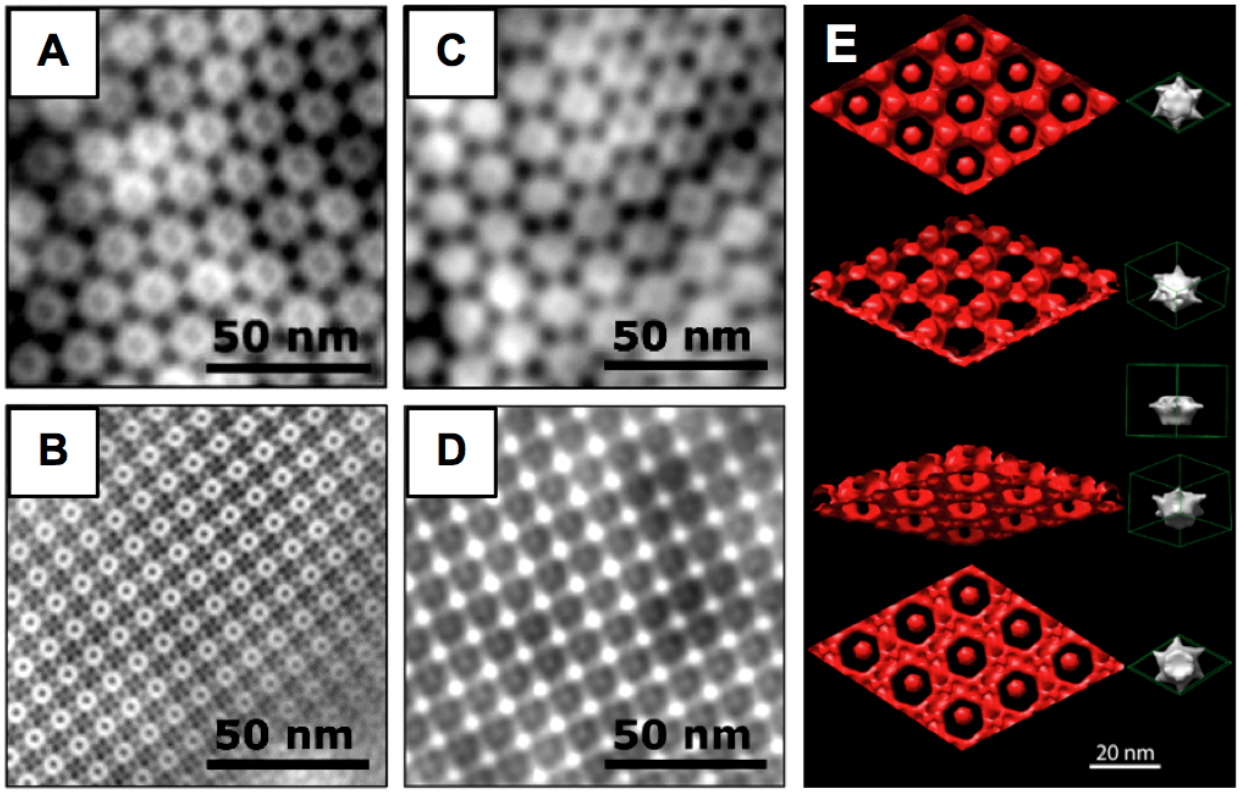

Figure 2.1. TEM images of negatively stained *D. radiodurans.* (A) and *S ureae* (**B**) S-layers and of the corresponding electrodeposited Cu_2O films (**C,D**). (**E**) TEM-based 3D reconstruction of nanostructured Cu_2O (red). Four different angles are shown along with a protein unit cell (right). Panel E is reprinted with permission from reference [149].

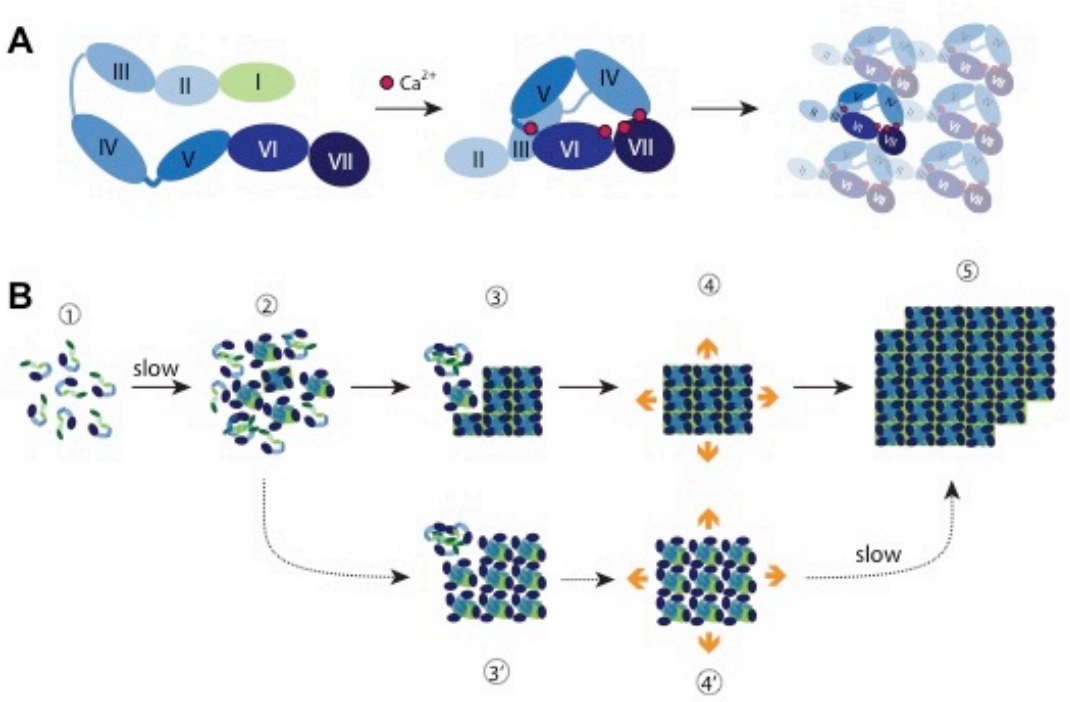

Figure 2.2. Building S-layers. (**A**) Schematic structure and assembly pathway of the *G. stearothermophilus* SbsB S-layer protein. The cell-wall attachment domain (light green) and six consecutive immunoglobulin-like domains (shades of blue) are schematically depicted in the extended monomer structure (left). The four calcium ions stabilizing intra- and inter-domain contacts are shown as red spheres in the assembly-competent monomer (middle). The cell-wall attachment domain is not resolved in the crystal structure and thus not shown. Structure of the oblique array formed upon assembly of compact monomers viewed from the outside of the cell (right). All drawings are based on reference [135]••. (**B**) Growth of *L. sphaericus* SbpA at interfaces. Monomers are believed to consist of three cell wall anchoring domains (green) followed by a calcium binding domain and three C-terminal immunoglobulin-like domains (shades of blue) [157]. Steps involved in S-layer crystallization on supported lipid bilayers are indicated with solid arrows and labeled 1 to 5. S-layers are viewed from their external side. Mica can stabilize a subpopulation of less compact but yet crystalline clusters (3' and 4') that eventually convert to the compact and lower energy form (step 5). All drawings based on references [155]•• and [156].

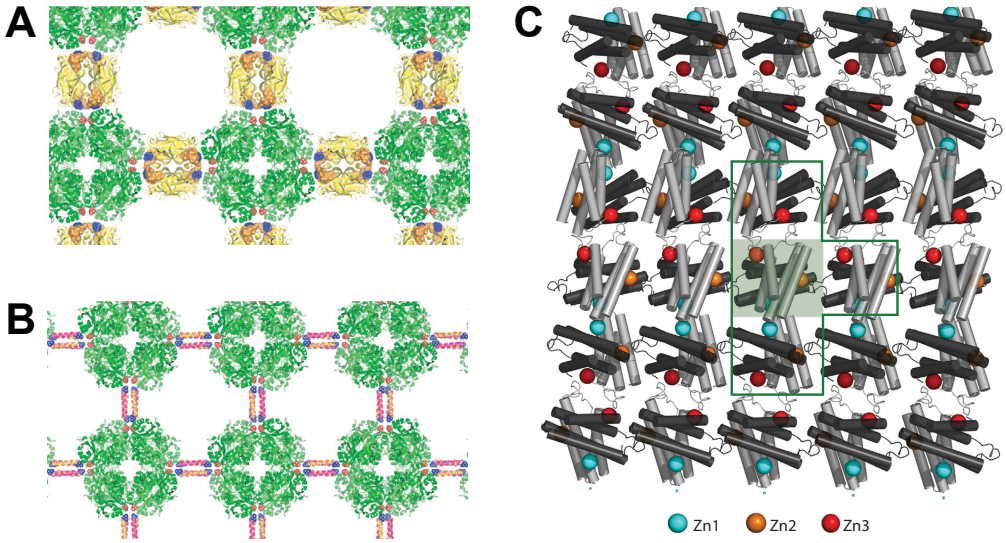

Figure 2.3. Schematic representation of 2D lattices. The lattices were generated by (**A**) fusing a Streptag to ALAD (green) and using Streptavidin (yellow) to stitch the structure; or (**B**) ALAD-Lac21E and ALAD-Lac21K fusions to assemble the array via coiled-coil interactions. Adapted from reference [171]• with permission. (**C**) Molecular arrangement of 2D sheets obtained with the RIDC3 derivative of cyt cb_{562} in the presence of zinc. The shaded green box contains a single **C2**-dimer. Three Zn coordination environments enable 2D self assembly. The Zn1 (blue) and Zn2 (orange) sites are formed by the high-affinity coordination motif and the Zn3 site (red) by the low-affinity coordination motif. Adapted from reference [162]• with permission.

Chapter 3 COMPUTATIONAL DESIGN OF PROTEINS FOR CO- AND SELF-ASSEMBLY

3.1 INTRODUCTION

The end of Chapter 2 highlights exciting work that aims to create novel, two-dimensional lattices with proteins [33, 41]. Several reviews have been written on the subject [31, 182-185] and there is general agreement that the field lags behind what has been achieved with DNA nanotechnology [5-7]. Computational design is, however, helping close the gap [28, 186-188]. Already, the Sinclair group has used a protein-fusion approach and the Tezcan laboratory has used metal-coordination along with advances in computation to design two-dimensional protein arrays. These approaches are two of the four strategies that have been exploited so far to induce protein assembly: the use of genetic fusions, the coordination of metal ions at interfaces, the use ligand-mediated assembly and the design of novel protein-protein interfaces [182]. This chapter starts with symmetry and design considerations, expands on the four approaches used to the design self-assembling protein structures, and discusses limitations associated with the extraction and purification of designed proteins.

King and coworkers, who performed seminal work in the design of interfaces that mediate protein-protein interaction, describe their goals as: "*The ultimate aim of these studies is to facilitate the design of protein nanomaterials with functions and characteristics that go beyond the capabilities of those assemblies found in nature and that are tailored to specific applications*" [182]. The following paragraphs will show that computational design of proteins for self- and co-assembly is indeed on that path.

3.2 SYMMETRY OF BUILDING BLOCKS

The most significant feature that a protein building block can possess is preexisting group symmetries [31]. Lai et al. state that a symmetry group "*is an exact mathematical idea that expresses the complete set of spatial operations that interrelate a set of individual components, in this case a set of structurally identical protein subunits in a 3D assembly*" [183]. Indeed, if a protein already has preexisting symmetry it will be easier to pack or assemble multiple copies with one another. If a protein does not have symmetry then interaction between subunits will have to be completely engineered. Oligomeric proteins are almost always symmetric and are therefore often used as building blocks for the production of self-assembing supramolecular structures [189].

A good starting point to understand symmetry in proteins is to start with cyclic symmetry, Cn. This involves n-fold rotation around, 360°, the proteins only symmetry axis C. If a protein can make it all the way around without having a reaching symmetry n is 1 and the protein is asymmetric. C1 is the case for most monomeric proteins. Most homodimers have C2 symmetry because they can rotate 180° to find symmetry. Symmetry for easy tiling of a 2-D pattern are C3 (120°), C4 (90°), or C6 (60°). The continuation of building these groups in 2-D patters leads to plane groups. Figure 3.1, shows the 17 different plane groups that can form 2D crystals. Of course proteins are not two-dimensional and can have more complicated symmetries of dihedral, tetrahedral, octahedral to name a few. These will not be covered in more detail because we are concerned with the 2D case.

3.3 FOUR APPROACHES TO ENGINEERING PROTEINS FOR ASSEMBLY

3.3.1 *Fusion of protomers*

This strategy relies on identifying two naturally occurring, and preferentially homo-oligomeric proteins, and of linking protomers from each oligomer to one another via a short linker. An example would be to fuse a monomer from a dimer (A_2) to that of a trimer (B_3) to create the high order structures depicted in Figure 3.2. Before this operation is conducted, however, it is critical to first perform computational symmetry testing on each oligomer to determine the location of the fusion joint so that the proper arrangement can be propagated.

The foundation for this approach was laid by the Yeates lab with the design of a self-assembling protein cage [40]. The process starts by a search of the Protein Data Bank (PDB) [190], a depository of known protein crystal structures, to identify dimers or trimmers that meet certain symmetry parameter [40]. Candidates are further screened for the presence of an N- or C-terminal α-helix so that a variable α-helical linker can be built between two monomers belonging to different oligomers. An example of this is tetrahedral cage construction, where the dimmer to trimmer fusion must create a 54.7° angle or the cage will not form. This can be seen in Figure 3.2a. Computer models are used to check the symmetry of a hypothetical fusion protein, steric hindrances of residue at the interface of a particular fusion protein, and if the fusion protein could satisfy the rules, like the tetrahedral cage angle constraint, they Padilla *et al.* set out for making the their self assembling structures [40]. In Padilla *et al.* study, the winners were Bromoperoxidase (a trimeric enzyme from *Streptomyces aureofaciens*) and the dimeric M1 matrix protein of the influenza virus. A 9 amino acids long linker (KALEAQKQK) derived from the long α-helix of ribosomal protein L9 was selected to create a 49-kDa fusion protein, between residues 1-276 of Bromoperoxide and 3-164 of influenza virus matrix protein M1, predicted to

self-assemble into a 16nm cage comprised of 12 protomers. The fusions were found to be highly soluble (up to 20mg/ml) and formed the prescribed structures as evidenced by TEM imaging (Figure 3.2). Two mutations were introduced to eliminate steric clashes in a separate study [191], reflecting the fact that designing for assembly is not as trivial and robust as advertised. In fact, while cages were also observed in the second study, TEM images reveal significant bending and twisting in the linker [191].

The approach taken by Sinclair et al. to produce arrays by genetically fusing protein assemblies possessing rotational symmetry axes of equal order [31] through a streptavidin-leucine zipper was described more extensively in Chapter 2 and is illustrated in Figure 3.2.

3.3.2 *Metal-directed protein assembly*

Metal-directed protein self-assembly uses metal-ligand synergy to control protein-protein interactions [192]. Metal-ligand coordination is a relatively simple phenomenon that involves the donation of electrons by the electron-rich ligand to the electron-poor metal center. Proteins, however, are not as simple as metal ligands because they feature extensive, heterogeneous surfaces that can interact with each other and with metal ions in unpredictable ways [192]. The Tezcan group addressed these difficulties by using a folded monomeric protein (cytochrome Cb562) as a building block [193]. The stability and rigidity of a folded protein ensured that the overall architecture of the protein was retained as its surface was modified [193]. Using a monomeric protein prevented a bias towards self-association. A minimal number of mutations transformed the non-self-associating surface to a self-associating surface that enabled Zn coordination [193]. The protein interfaces in the resulting Zn-directed tetramer were then redesigned to yield a protein architecture that could self-assemble in the presence or absence of metals [193].

3.3.3 *Ligand-mediated protein assembly*

Ligand-mediated approaches resemble metal-directed assembly in that the innate ability of proteins to bind a specific ligand are used to drive self-assembly [33]. For instance, two identical monomers can be brought into a homodimer by making use of a symmetrical ligand. Similarly, two different proteins can be linked to form a heterodimer by making use of an asymmetrical ligand specifying a region bound by each protein at each of its ends [194].

An example is the production of protein nanorings [195] where the symmetrical linker is dimeric methotrexate (MTX_2-C9) and the protein is wild-type dihydrofolate reductase (DHFR) from *Escherichia coli*. With no modification, a homodimer is produced [196]. However, when two DHFR monomers fused via a glycine linker are used, stable cyclic nanorings are formed [195] although the product of the reaction are heterogeneous and range in size between 8 and 20nm (Figure 3.3). This is the best example of ligand based protein assemble into larger structures.

3.3.4 *Design of novel protein-protein interfaces*

This strategy relies on computational approaches to design interfaces and protein pairs in which a collection of weak, more natural interactions, collectively drive higher order structure formation. An enormous advantage of this much more challenging approach is the possibility of achieving atomic-level control of position and orientation [42-47]. The process starts by selecting a target symmetrical group and an oligomeric protein that fulfills the requirement for that symmetrical group. Next, initial symmetric docking is performed [188], clashes in interfacial residues are manually identified, and interfaces are computationally redesigned to create low-energy, geometrically specific interfaces that drive assembly [182]. The Rosetta3 software suite allows for this type of experiments [188, 197].

The strategy was used with great success by King et al. in 2012 [45] in a landmark study where both octahedral and tetrahedral cage-like structures were produced. The overall design concept is illustrated in Figure 3.4: the cage consists of 8 oligomers and has a hollow core. In the case of the octahedral cage, the starting oligomer are homotrimers of which their are 271 in the PDB that fit this case. Thirty of those were symmetrically docked with Rosetta and 10 were selected to undergo computational interface redesign. In this fashion, 33 designs were produced for experimental characterization, each including an average of 9 mutations per monomer [45]. A protein called O3-33; was found to assemble into the prescribed octahedron with atomic level precision by cryo TEM reconstruction (Figure 3.5.)

3.4 EXPRESSION OF DESIGNER PROTEINS

Before a design can be tested for assembly, the protein must first be expressed and purified. A major impediment is the formation of large aggregates inside the host cell because proteins are designed to self-assemble [183]. Refolding from the inclusion body state can significantly complicate the purification train and the process is often inefficient and low yield. A possible approach to circumvent the problem would be to rely on multiple proteins to create a self-assembled structure. In this case, each individual protein can be produced and purified in isolation and self-assembly only occurs upon mixing [34, 183]. In King's work, only 13 of the 33 designed trimeric monomers were soluble. Thus nearly half of the computational designs were not experimentally tested. Clearly, it is important to keep the purification train in mind when designing a protein.

3.5 CONCLUSION

The world of computational design of protein assemblies is generating much excitement although only a few structures have been made. This field is on the cusp of leaving its infancy

behind and taking the next steps forward. As computer software, like Rosetta [198, 199], continues to expand along with advances in protein crystallization technology [200] modelers have more tools and proteins to use to create novel new computational structures. At the same time these structures are being observed, usually via cryo-TEM. Each new structure proven to assemble as predicted gives scientist and engineers a new tool and model to work with. Advances will come from the use of a hybrid approach, e.g. fusion approach mixed with a protein-protein interface design. New techniques for quick and efficient prescreening are being developed in order to accelerate feedback from experiment to computation. This will allow for models to be quickly redesigned, and tested and help gain more insight into the complex world of protein interaction. This type of feedback helped push DNA nanotechnology forward years ago and computational design for protein self-assembly should make a similar leap.

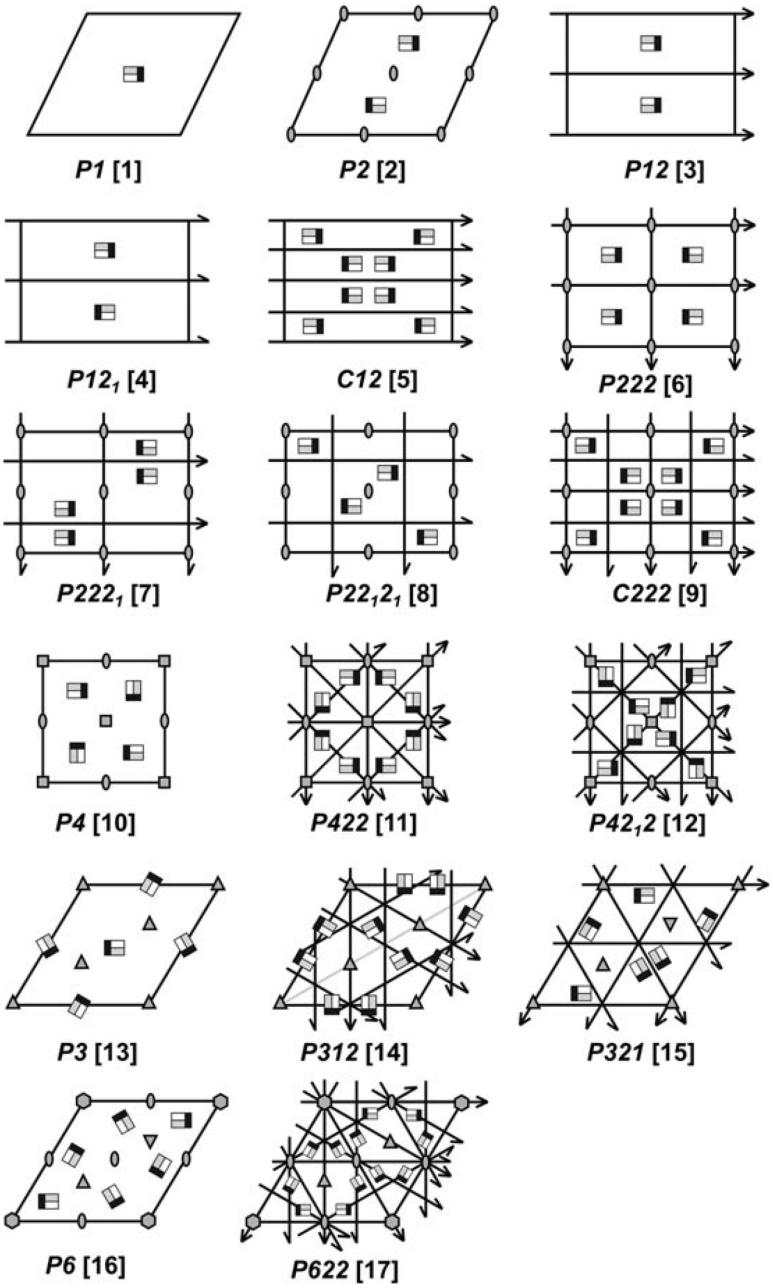

Figure 3.1. The 17 possible plane groups for 2D crystals. The black lines outline the unit cell, with full arrows indicating a 2-fold reflection and half arrows showing a 2-fold screw axis. The locations of rotations with orders of 2 (180°), 3 (120°), 4 (90°), and 6 (60°) are shown by ovals, triangles, squares, and hexagons, respectively. Typically shown as black, rotation symbols are shown filled in gray for clarity. For all panels the 2D space group number is indicated in brackets. Reprinted from reference [201].

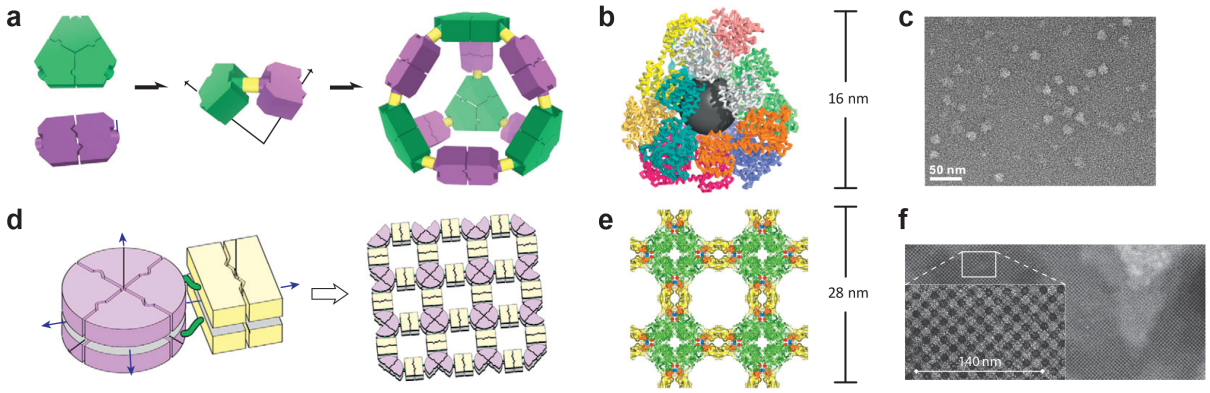

Figure 3.2. The fusion approach. (a) Protomers from naturally-occurring oligomers (a green trimer and a purple dimer) are fused to one another by a linker (yellow). This α-helical semi-rigid linker is flexible enough that during assembly, a 3D cage structure can be built **(b).** The production of the predicted structures are verified by cryo-TEM reconstructions and the cryo-TEM image is seen in **(c)** [191]. **(d)** Fusion approach relying on a common symmetry axis (blue arrows). Here a short linker (green), limits the degrees of freedom sampled by subunits from two different oligomeric proteins (pink and yellow). As the subunits come together, propagatation progresses on the blue arrow axis. When the fusion proteins assemble, a two-dimensional lattice should be created. **(e)** A molecular model proposed by Sinclair and coworkers with **(f)** showing the results of TEM analysis [41]. Adapted from reference [183].

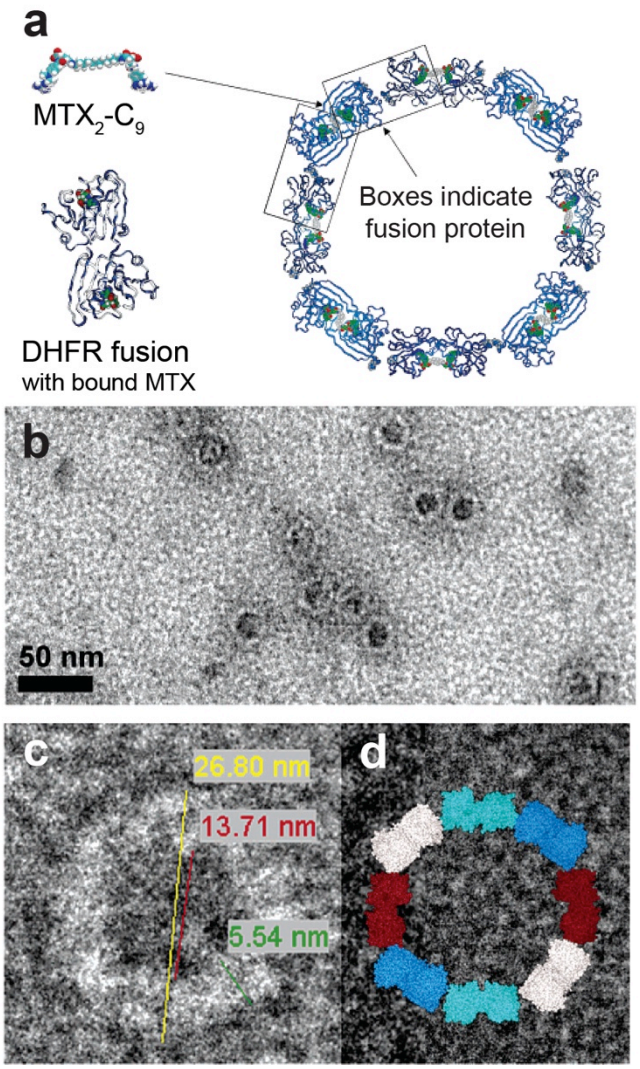

Figure 3.3. The ligand approach. Panel **(a)** shows the symmetrical ligand, MTX_2-C_9, which can bind two different DHFR proteins at once. DHFR was fused to itself by a single glycine with the expectation that a nanoring would be produced upon mixing of ligand and fusion protein (right). **(b)** Negative stain TEM image of nanorings. Uranyl acetate pools inside the nanorings. **(c)** Close up view of a nanoring that shows the expected spacing. **(d)** Space-filling model of an octameric ring scaled to match the TEM image. Adapted from reference [195].

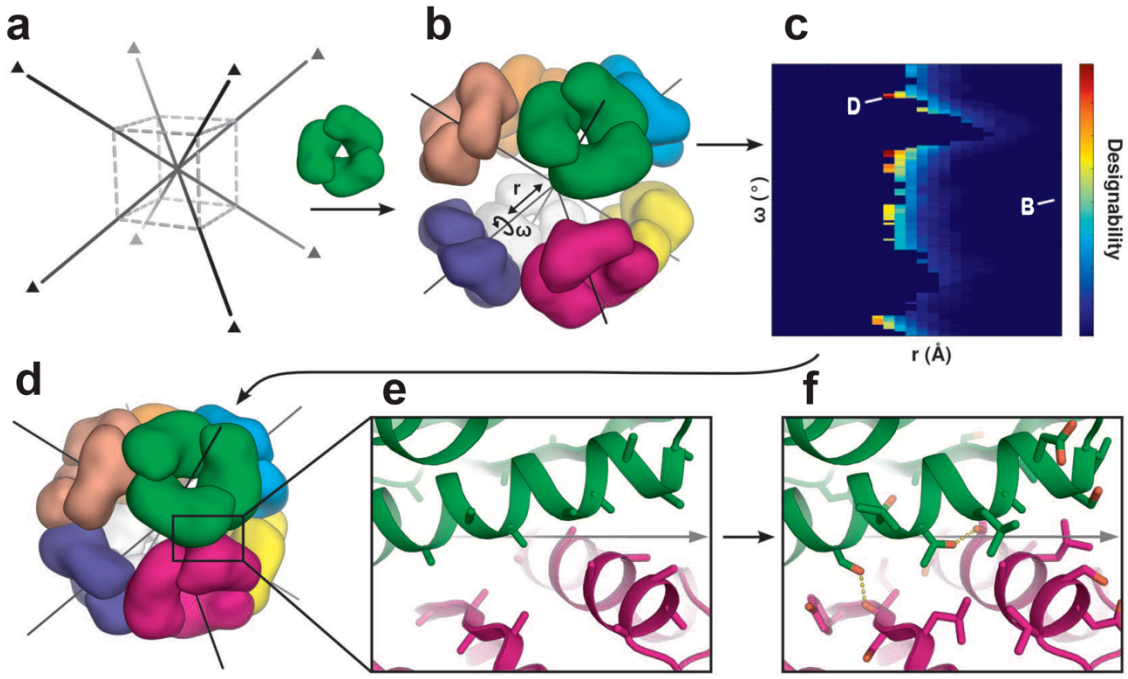

Figure 3.4. General concept for designing novel protein-protein interfaces. (a) A target symmetric architecture is first chosen. Octahedral point group symmetry is used in this example; the three- fold rotational axes are marked by triangles and shown as black lines throughout. The dashed cube is shown to orient the viewer. A symmetric oligomer which shares an element of symmetry with the target architecture, here a C3 symmetric trimer (green), is selected as a building block. **(b)** Multiple copies of the building block are symmetrically arranged in the target architecture by aligning their shared symmetry axes. The preexisting organization of the oligomeric building block fixes several (in this case four) rigid-body degrees of freedom (DOFs). The two remaining DOFs, radial displacement (r) and axial rotation (w) are indicated. **(c)** Symmetrical docking is performed by systematically varying the two DOFs (moves are applied symmetrically to all subunits) and computing the suitability of each configuration for interface design (red: more suitable; blue: less suitable). Points corresponding to the docked configurations in (B), in which the building blocks are not in contact, and (D), a highly complementary interface, are indicated. **(e)** Closer view of the interface in **(d)**. The interface lies on an octahedral two-fold symmetry axis shown as a gray line. In all steps before interface design, only backbone (shown in cartoon) and carbon beta (shown in sticks) atoms are considered. **(f)** Sequence design calculations are used to create low-energy protein-protein interfaces that drive self-assembly of the desired material. Designed hydrogen bonds across the interface are indicated by dashed lines. Reprinted from reference [45].

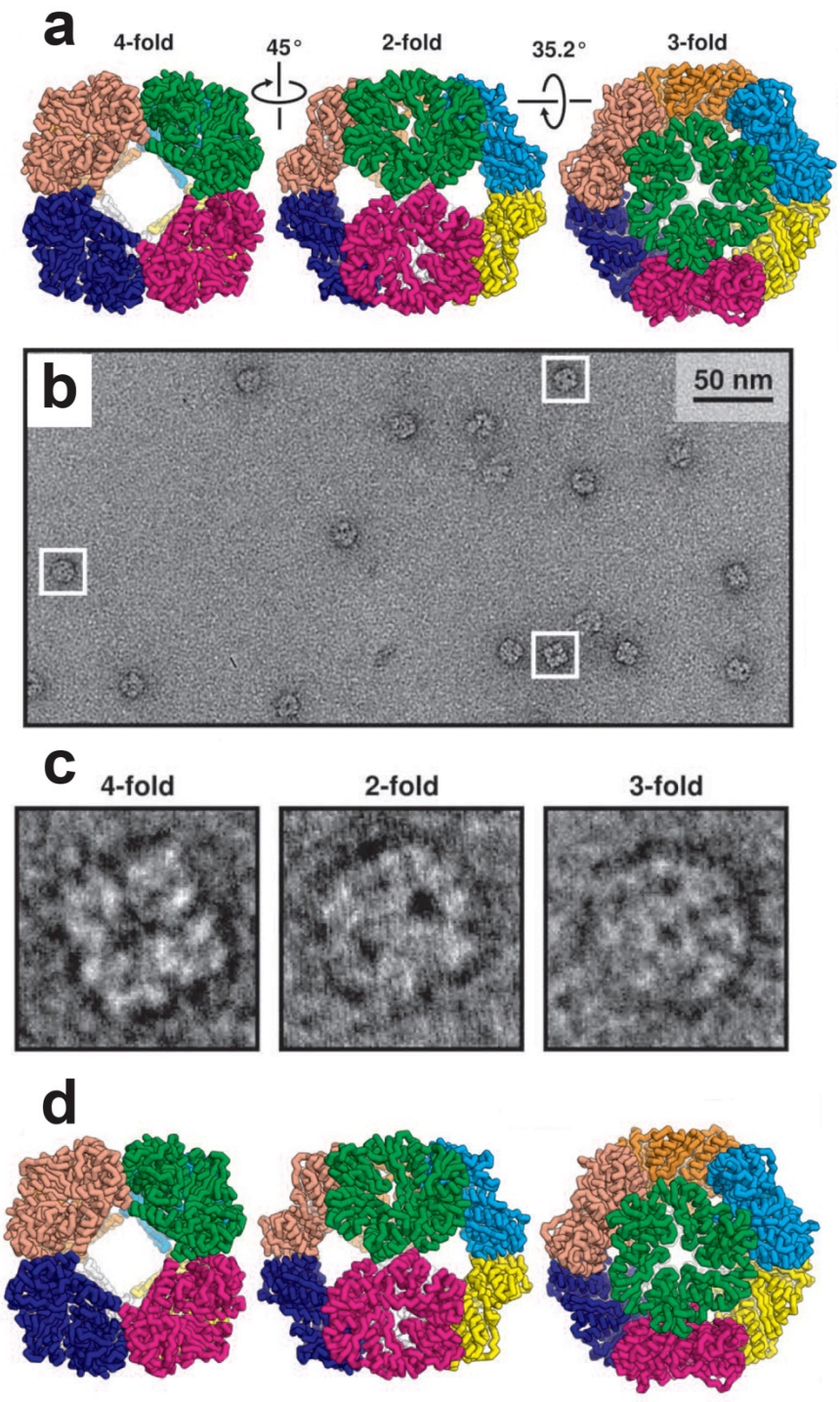

Figure 3.5. Structural validation of O3-33. (a) the ribbon structure model of O3-33. The first image shows 4-fold rotational axis then by moving the structure slightly the 2-fold rotational axis appears (middle). Finally on more move and the 3-fold rotational axis is shown (right). **(b)** TEM images of O3-33 and the white boxes show 4-fold, 2-fold and 3-fold rotational axes, which are enlarged below **(c)**. **(d)** The crystal structure of O3-33 from cryo-TEM reconstruction. Adapted from reference [45].

Chapter 4 DESIGNING TWO-DIMENSIONAL PROTEIN ARRAYS THROUGH FUSION OF MULTIMERS AND COMPUTATIONAL INTERFACE REDESIGN

4.1 INTRODUCTION

Building blocks that self-assemble into periodic two-dimensional (2-D) arrays with nanoscale precision are of considerable interest for the production of mesostructures with superior catalytic, transport, opto-electronic or biological properties[202-206]. Although the S-layers of archaea and bacteria are archetypal 2-D protein materials[50], and despite recent advances in the computational design of protein-based nanostructures[45, 165, 169, 207-209], proteins remain underused for the fabrication of synthetic 2-D lattices[162, 171]. Here, we combine fusion of oligomers with cyclic symmetry and computational redesign of packing interfaces to produce proteins that self-assemble into 2-D arrays upon addition of calcium ions. We use TEM, AFM, small-angle X-ray scattering and fluorescence microscopy to show that the designed lattices which are 5 nm high with *p*3 space group symmetry and 7.25 nm periodicity self-assemble into structures that can reach hundreds of micrometers. The versatile strategy, experimental methods, and hexagonal arrays described herein should prove valuable for the engineering of functional nanostructured materials in 2-D.

4.2 RESULTS AND DISCUSION

To engineer a protein monomer capable of robust self-assembly into regular 2-D arrays, we combined fusion of two copies of point-symmetric building blocks and computational interface redesign to promote intimate interactions at the newly introduced interface (Figure 4.1). This strategy is a departure from the streptavidin-leucine zipper approach of Sinclair *et al.*[171]

the metal-coordination approach of Tezcan and coworkers[162], and the helical-bundle-mediated approach of Lanci *et al.*[165].

We set out to identify a naturally occurring protein that could be modified to produce a flat, 2-D crystalline layer pierced by a single small pore. To this end, we evaluated the structure of oligomeric proteins[210, 211] deposited in the protein data bank (PDB) for the following design specifications: (*i*) C3, C4, or C6 rotational symmetry for easy tiling of a 2-D pattern; (*ii*) topologically smooth top and bottom faces to facilitate interactions with technological surfaces and interfaces including 2-D substrates; (*iii*) non-occluded central pore smaller than 5 nm to allow for future templating of inorganic mineralization[149]; (*iv*) flexible loops suitable for future insertion of solid binding peptides[212]; and (*v*) termini oriented so that the C-terminus of one subunit in an oligomer can be fused to the N- terminus of a subunit in a separate oligomer via a linker long enough to bridge the gap between unit cells but short enough that the fusion protein cannot exist in one unit cell (Figure 4.1c).

Rosetta's symmetric docking[213] was run on the subgroup of models meeting the above criteria and *S. typhimurium* STM4215 was selected as the most likely candidate because it exhibited all features specified above. This protein has unknown function but was crystallized by Minasov and coworkers under the Protein Structure Initiative (PDB accession number 2GJV). STM4215 is a homohexamer 4.8 nm high and 7.5 nm in diameter pierced by a ≈ 3 nm pore. Each 151-residue subunit coordinates a calcium ion and specifies three cysteines that do not form disulfides. All N- and C-termini lie on the more acidic "bottom" face which, like the "top" face is fairly flat (Figure 4.2a and Figure 4.6).

The 19 amino C-terminal amino acids (which are not resolved in the structure and presumably flexible) were eliminated to yield a 14.7-kDa protomer referred to as "T". In the

docked conformation, we determined that six amino acids were sufficient to connect neighboring subunits into a 2-D array while prohibiting linkage between two subunits of the same hexamer (Figure 4.1). The Rosetta software suite[214] was used to redesign the newly introduced interface (Figure 4.2b-c and Figure 4.8). Four Rosetta-predicted interfacial mutations (S5A, N12A, E16A and D98A) as well as a C99A substitution were introduced to eliminate steric hindrances and charge repulsions, allowing for tighter packing between adjacent hexamers. We refer to this resulting monomer as TM. Finally, a hexaglycine flexible linker was selected through Rosetta's loop modeling[215] to join the two TM protomers in a head to tail fashion and yield a TTM dimer predicted to be competent for self-assembly (Figure 4.2b-c and Figure 4.7).

The gene encoding wild type STM4215 was amplified from *S. typhimurium* DNA and the T, TM, TT and TTM variants constructed by standard molecular biology. All genes were placed under control of the T7 promoter and expressed in *E. coli* BL21(DE3). Whereas wild type (WT) and truncated (T) proteins were primarily soluble, introduction of interfacial mutations in a single subunit (TM), or tandem duplication of the protomer (TT) led to insolubility (Figure 4.2d). Not surprisingly, the TTM variant partitioned quantitatively in the insoluble cell fraction and was therefore purified from the aggregated state by denaturing gel filtration chromatography.

Far UV circular dichroism (CD) was used to monitor the conformation of TTM in the unfolded state and following refolding by rapid dilution. While the CD spectrum of urea-treated material was typical of an unfolded protein, refolded TTM exhibited distinct features in the 210-220 nm window with calculated α-helical and β-sheet content of 25% and 22%, respectively (Figure 4.3a). The protein's hydrodynamic diameter was about 6 nm as measured by dynamic light scattering and a 3.4 μM solution did not experience an increase in light scattering at 380 nm

when held at room temperature, as would have been expected if the protein were to spontaneously aggregate or self-assemble in a 2-D array.

Based on the observations that recrystallization of *Geobacillus stearothermophilus*[147] and *Lysinibacillus sphaericus*[216] S-layers requires calcium and that the STM4215 structure contains liganded Ca^{2+}, we tested calcium as a chemical switch for TTM polymerization. Figure 4.3b shows that addition of this cation caused an immediate increase in scattering but no gross misfolding as the CD spectrum of calcium-supplemented TTM remained similar to that of untreated protein (Figure 4.3a). Initial rates and extent of scattering depended on calcium and protein concentrations with apparent second order kinetics, and increased with temperature as would be expected from a typical nucleation and growth process (Figure 4.3b and Figure 4.8). Additionally, the smaller divalent cation Mg^{2+} was less efficient than Ca^{2+} at causing an increase in scattering while K^{+} was incapable of doing so (Figure 4.8c). Because the calcium binding region of native STM4215 is not part of the designed interface, it appears that, as is the case with *G. stearothermophilus* SbsB[147], Ca^{2+} binding induces a structural transition that makes TTM competent for assembly.

Previous 2-D protein design efforts have exclusively relied on electron microscopy to characterize designed lattices[162, 165, 171]. Inspired by Nam *et al*[11], we first assessed the size and distribution of self-assembly products following staining with Nile Red, a lipophilic dye binding to protein hydrophobic domains and self-assembled peptoid structures[11]. Fluorescence microscopy revealed the presence of polydisperse particles in the 1 to 10 μm range (Figure 4.4a) with occasional appearance of large (>100 μm) thin structures with sharp edges and folded-over domains that we take to be self-assembled TTM sheets (Figure 4.4b and Figure 4.9). Although

these structures were rare, spontaneous long-range ordering at these dimensions is unparalleled for any self-assembling protein, including S-layer proteins.

Short-range order was assessed by transmission electron microscopy (TEM) following negative staining of calcium-supplemented TTM samples with uranyl acetate. TEM images revealed periodic hexagonal patterns and fringes associated with the stacking of multiple sheets (Figure 4.4c). Consistent with design predictions, we found that the unit cell was 7.25 nm by FFT analysis of micrographs (inset). AFM imaging on silicon substrates further revealed that self-assembled TTM was nearly flat with kinked edges and a height of 4.9 nm (Figure 4.4d-e). The smaller topographic features present on self-assembled terraces had the same height (Figure 4.4e) and may correspond to the stacking of independently nucleated monolayers. They may also be explained by a scenario in which an unpaired TM "seed" from a TTM repeat twists above the growth plane of the original monolayer at a defect region (e.g., Figure 4.4e, arrow) to initiate a new growth event. We favor the former scenario because shortening the linker by two glycine residues –a change predicted by Rosetta to prevent out of plane excursions – did not discernably affect nanostructure (Figure 4.10) or long-range assembly (Figure 4.11).

To gain further structural information, we performed small angle x-ray scattering (SAXS) measurements on a His-tagged variant of STM415 and on calcium-treated TTM (Figure 4.5). For the wild type protein, a fit of the Guinier region of the SAXS profile yielded a radius of gyration of 34 ± 0.3Å which is in good agreement with the 30.9Å value calculated from the crystal structure when contributions from hexahistidine extensions are taken into account. Additionally, the scattering profile calculated from the 2GJV coordinates using CRYSOL[217] matched experimental data with deviation at high Q attributable to the presence of His tags (Figure 4.5a, solid line).

As seen from the lack of a Guinier region at low Q (Figure 4.5b), instrument resolution was insufficient to obtain the overall size of self-assembled TTM structures. However, high-Q scattering intensities of STM415 and calcium-treated TTM were comparable, confirming that the hexameric structure of the wild type protein is preserved in the assembly. Furthermore, deconvolution of the STM415 form factor from that of the assembly yielded a power-law dependence of -2.1 ± 0.1 which is characteristic of a 2-D object (Figure 4.5 and Figure 4.12). Finally, the scattering profile derived from the coordinates of the structure shown in Figure 4.2c was in very good agreement with experimental data with deviation a low Q indicating that the products of TTM self-assembly have a characteristic length exceeding that of the model (Figure 4.5b).

4.3 CONCLUSION

Taken together, the above experiments fully validate design predictions: TTM self-assembles into flat 2-D arrays of with *p*3 symmetry (the asymmetry introduced by the linker makes the overall design *p3* rather than *p6* symmetric), 7.25 nm lattice constant and 5 nm thickness that can grow to yield structures on the order of 100 μm. We expect that the TTM platform will be superior to S-layers for practical applications[218, 219] because its building blocks are fully specified and can be engineered and computationally tested for desired outcomes. Our design strategy should prove a valuable addition to the growing number of approaches being developed for the production of protein-based nanomaterials by enabling the construction of 2-D structures with exquisite control of space group symmetry, periodicity and chemistry.

4.4 METHODS

4.4.1 *Computational docking and design*

The 3D complex database[210] was used to identify 3-, 4-, and 6-fold complexes with cyclic symmetry to serve as building blocks for 2-D array design. The pool was narrowed down to structures containing a central pore of about 3 nm, and N- and C-termini accessible from the "outside" (e.g., located on the face furthest away from the helical axis). These requirements eliminated most native complexes, leaving only two scaffolds. The resulting structures were docked using Rosetta symmetric docking with a low-resolution score term favoring large numbers of backbone-backbone contacts across the interface and short distance between N- and C-termini across the interface. After docking, STM4215 was chosen for further design. Rosetta loop modeling[215] was used to identify ideal linker length, by building linkers of lengths 2-8 and identifying the fraction of trajectories with closed loops. This yielded an ideal loop length of 6, where virtually every trajectory resulted in a closed loop. Finally, we redesigned the interface for tighter subunit packing. Twelve key interface residues were manually selected, and these were allowed to maintain native identity or mutate to alanine. Multiple independent trajectories yielded four positions where mutation to alanine would allow for tighter packing. A fifth position, C99, was manually mutated to eliminate one cysteine residue from the designed protein and allow for future introduction of disulfides.

4.4.2 *DNA manipulation*

Genes encoding native STM4215 or a truncated (T) version lacking 19 C-terminal residues were amplified from *S. typhimurium* LT2 and inserted into the *Nde*I and *Bam*HI sites of pET-22b(+) (Novagen). The wild-type gene was moved to pET-28a(+) to create an N-terminal His-tagged version. The S5A, N12A, E16A, D98A and C99A mutations were introduced into T

to produce TM. The latter gene was isolated on a *Nde*I-*Bam*HI fragment, part of this product was digested with *Bst*NI to remove the stop codon and 32 bp from the 3' end, and *Nde*I-*Bam*HI and *Nde*I-*Bst*NI fragements were used with overlapping primers to produce TTM via PCR overlap extension[220]. A similar approach was used to build TT. The resulting plasmids, pSTM4215TTM-22 and pSTM4215TT-22, encode tandem repeats of TM or T joined by a Gly_6 flexible linker. TTM4, a TTM variant specifying a Gly_4 linker was similarly constructed.

4.4.3 *Protein expression, fractionation and purification*

BL21(DE3) transformants were grown at 37°C in LB medium supplemented with 100 μg/mL carbenicillin to mid-exponential phase and protein synthesis was induced with 0.5 mM IPTG. Cultures were fractionated into whole cell, soluble and insoluble fractions after 3hr of growth at 37°C as described[221]. For TTM purification, paste from 500 mL of culture was washed with 100 mL of buffer A (20 mM Tris-HCl, pH 8.5, 500 mM NaCl), resuspended in resuspended in 35 mL of buffer B (Buffer A with 5 mM EDTA, 0.1 mM PMSF, 0.1% NaN_3, 1 mM DTT) and disrupted by 6 cycles of sonication at 3 min intervals on a Branson Sonifier 450 operated at 30% duty. Insoluble material collected by centrifugation at 12,000*g* for 20 min was washed in 35 mL of buffer B supplemented with 2M Urea and 1% Triton X100 followed by two more washes with buffer B with 1% Triton and a final wash with Buffer B alone. Insoluble material (≈ 0.5 g) was unfolded by overnight incubation in 5 mL of buffer C (20 mM Tris-HCl, pH 8.5, 8 M urea, 0.5% 2-mercaptoethanol) and aggregated material removed by centrifugation. For purification, 1mL aliquots were fractionated on a Superdex 200 column (16 x 200 mm; GE Healthcare) developed at 0.5 mL/min in buffer C. TTM-rich fractions were concentrated to 68 μM (≈ 2 mg/mL) with 10-kDa cutoff Ultra-4 filters units (Millipore), divided into 50 μL aliquots and stored at -20^{o}C. Aliquots were refolded by rapid dilution in 950 μL of buffer A. Misfolded

proteins were removed by centrifugation and protein concentration was adjusted to 0.1 mg/mL 3.4 μM (0.1 mg/mL) using Ultra-4 units.

4.4.4 *Characterization*

CD spectra were recorded using 400 μL of 7 μM samples on an Aviv 420 spectrometer at 23°C. Scattering spectra were recorded at 380 nm on a Beckman Coulter DU640 spectrophotometer at the indicated temperatures, proteins and salt concentrations. SAXS experiments were performed on an Anton Paar SAXSess instrument with line collimation using a Cu-K_α (λ=1.54 Å) source. Protein samples (2 mg/mL) of His_6-STM4215 purified by Ni-NTA chromatography or TTM were incubated for 24h at 4^{o}C in the presence of 1 mM $CaCl_2$. Samples (250 μL) were measured in 1 mm quartz cuvette for > 4h using line collimation. Intensity was integrated as a function of pixel position and converted to I vs. Q using known sample to detector distance. Absolute scale calibration was performed using water as a standard. Scattering patterns were calculated using CRYSOL[217] and convoluted with the known beam profile to account for instrumental broadening. Calcium-treated TTM samples (50 μL) supplemented with 1 μM Nile Red (Sigma) were imaged on a Nikon Eclipse using 510-560 nm excitation. For TEM, samples (5 μL) were deposited on carbon-coated grids, washed 4 times, stained with 2.5 μL of 1-3% uranyl acetate for 30 s, washed, dried and imaged on a FEI Tecnai G2 F20 at 120 keV. AFM images were acquired in tapping mode on a Bruker Dimension Icon using a Tetra 15 cantilever (K-tek). Samples (50 μL) were deposited on acetone-cleaned silicon substrates for 2 min and washed 3 times before image acquisition.

4.5 ACKNOWLEDGMENTS

We are grateful to Dr. Yongdong Liu for help with initial purification and re-assembly experiments. Part of this work was conducted at the UW Molecular Analysis Facility, a member

of the NSF-NNIN network and at EMSL, a DOE Office of Science User Facility at PNNL. This work was supported by NSF award BBBE 1401835 (to FB) and DOE BES DE-SC0005153 (to LDP). JFM acknowledges financial support from NIH through a Cancer Nanotechnology Training Grant (T32CA138312). DB and FM acknowledge financial support from DTRA and AFOSR.

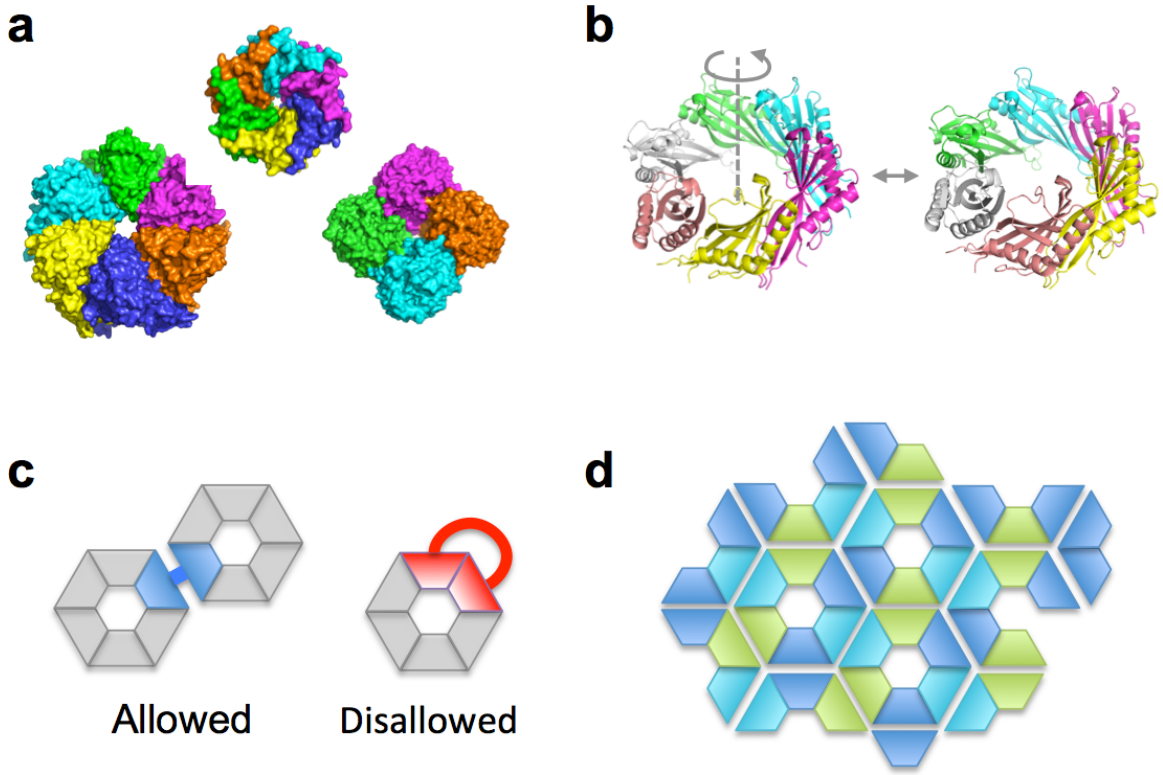

Figure 4.1. Lead candidate identification. **(a)** Screen the PDB for oligomeric proteins with **C3**, **C4**, or **C6** symmetry, flat faces and pore size < 5nm. **(b)** Conduct low-resolution Rosetta symmetric docking with 2D lattice symmetry. **(c)** Build flexible linker that will link two subunits from different oligomers and **(d)** redesign the interface to promote lattice self-assembly.

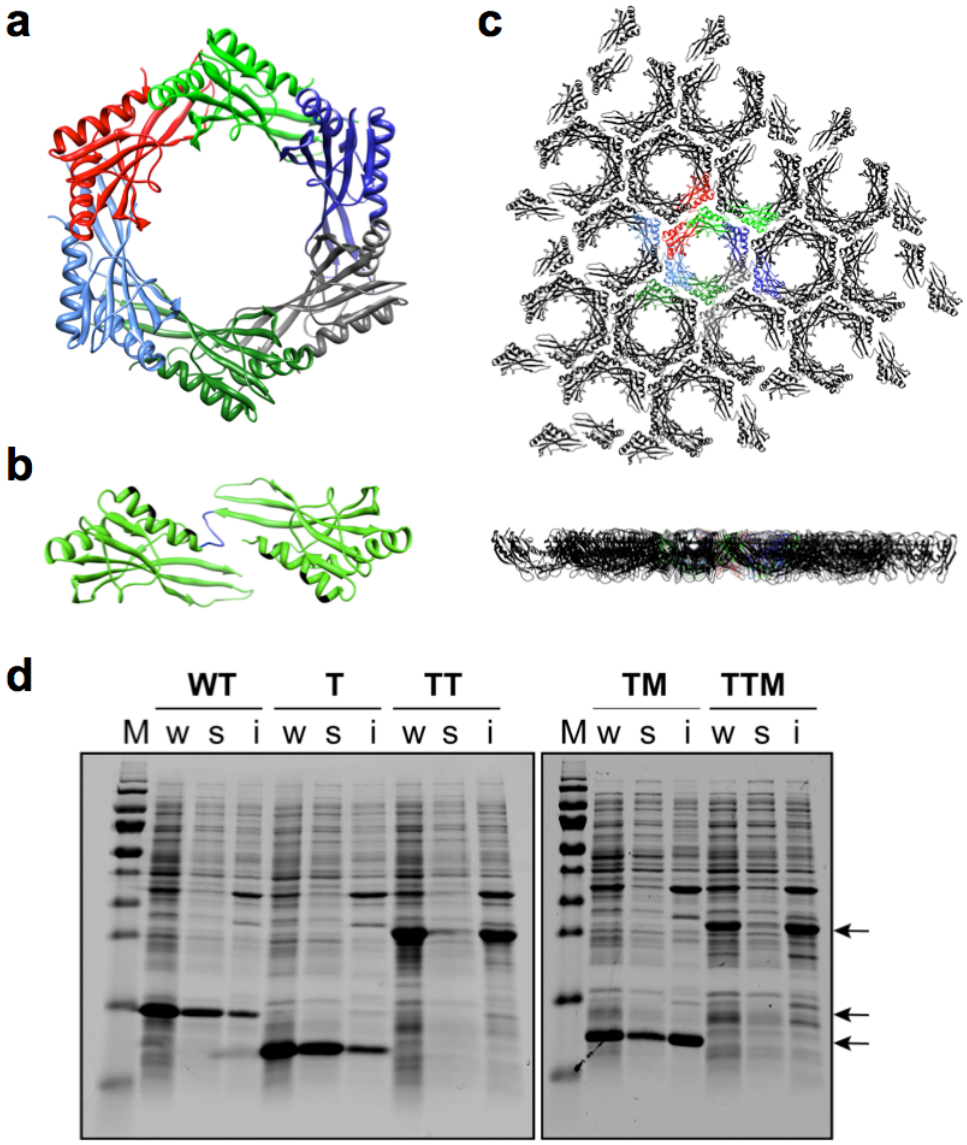

Figure 4.2. Engineering *S. typhimurium* STM4215 for 2D self-assembly. **(a)** Ribbon structure of wild type STM4215. **(b)** and of the TTM dimer linked by a flexible linker (blue) and incorporating Rosetta-predicted interfacial mutations (black). **(c)** Top and side views of the expected tiling pattern with individual TTM dimers shown in different colors. **(d)** Subcellular localization of wild type (WT) and variant STM4215. Whole cell (w), soluble (s) and insoluble (i) fractions are shown. Lane M contains markers. Arrows from top to bottom denote the migration position of TT/TTM, WT and T/TM.

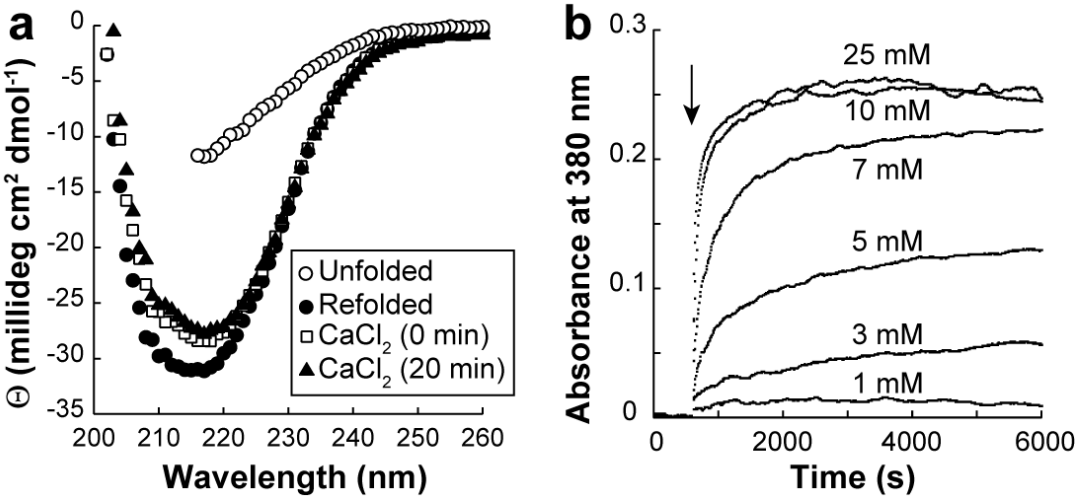

Figure 4.3. Calcium-induced self-assembly. (a) Far UV CD spectra of urea-unfolded TTM (open circles), refolded TTM in the absence of additive (closed circles) or immediately after (open squares) or 20 min after (closed triangles) addition of 5 mM $CaCl_2$. **(b)** Influence of the $CaCl_2$ concentration on TTM (3.4 μM) scattering at 380 nm and 23°C. The arrow denotes the time of $CaCl_2$ addition.

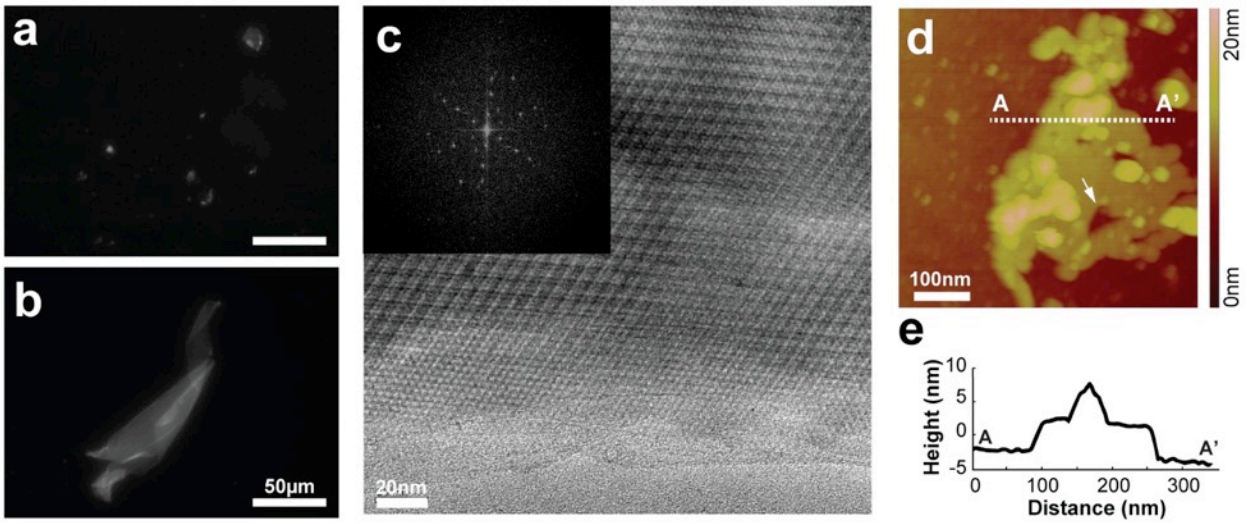

Figure 4.4. Characterization of TTM self-assembly at long- and short-range. (a-b) Fluorescence microscopy images of Nile Red-stained TTM structures obtained after incubating 1.7 µM protein with 1 mM $CaCl_2$ for 2h at 23°C. **(c)** TEM image of uranyl acetate-stained TTM structures obtained after incubating 3.4 µM of protein with 10 mM $CaCl_2$ at 23°C. The inset is a FFT of the entire image showing crystalline hexagonal packing with 7.25 nm periodicity. AFM image **(d)** and height scan along the AA' line **(e)** of TTM structures obtained after incubating 0.34 µM protein with 1 mM $CaCl_2$ for 2h at 23°C. Steps from the silicon substrates along AA' are 5.1 nm, 4.6 nm, 4.7 nm and 4.8 nm. Different protein and calcium concentrations were selected to optimize results for each technique but images shown are representative under the range of conditions used.

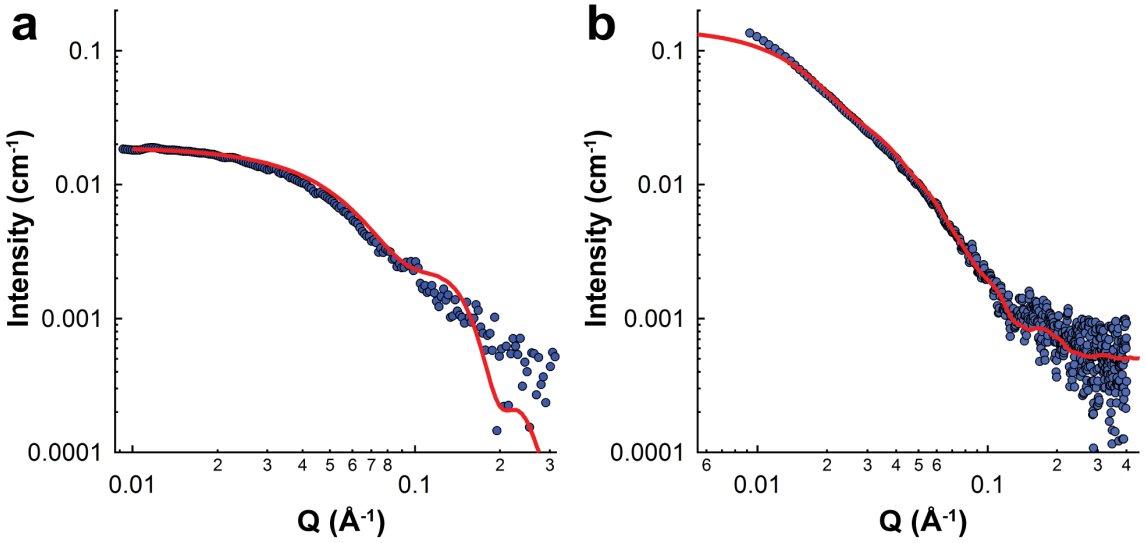

Figure 4.5. SAXS analysis of wild type STM415 and self-assembled TTM. Smeared scattering intensity versus wave vector plots for a variant of *S. typhimurium* STM415 fitted with an N-terminal His-tag **(a)** and self-assembled TTM **(b)**. Blue circles correspond to data and solid red lines to calculations. See text for details.

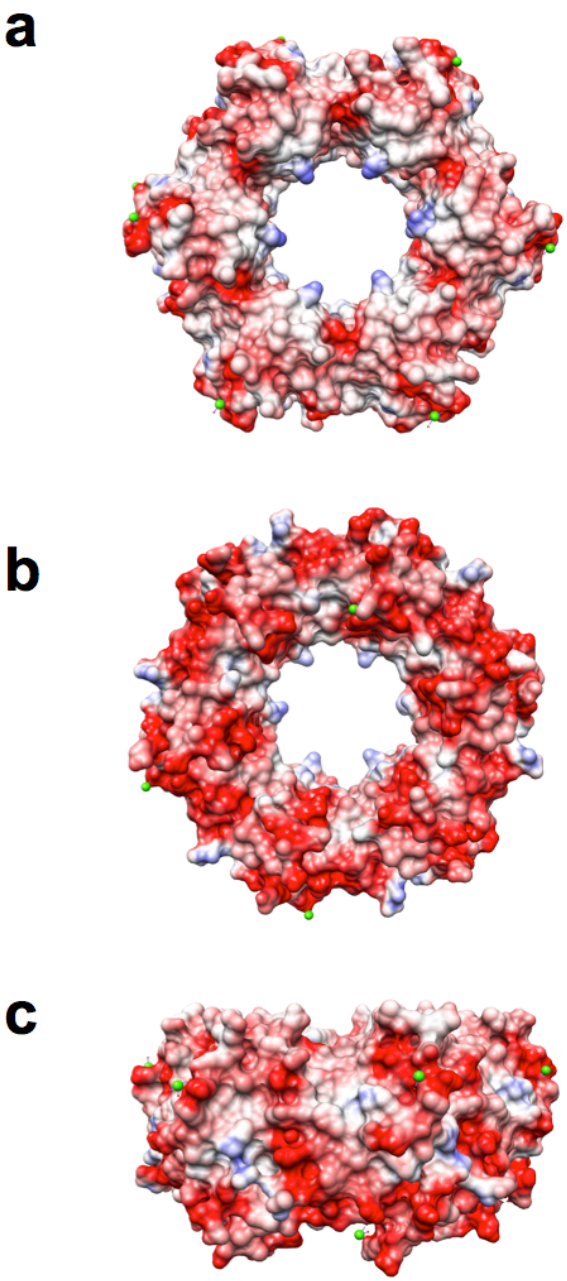

Figure 4.6. Molecular surface of *S. typhimurium* STM4215. Electrostatic coloring of the top **(a)**, bottom **(b)** and side **(c)** surfaces surfaces is between -10*kT* (red) and +10*kT* (blue). Calcium ions are shown as green spheres.

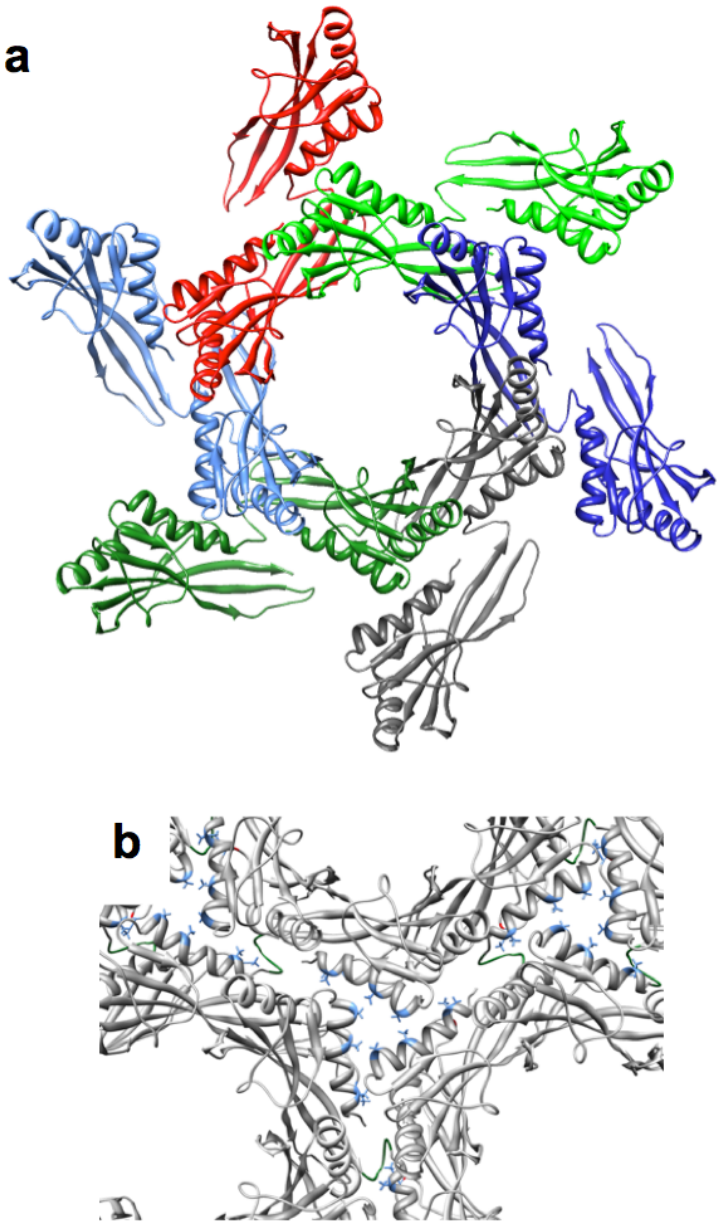

Figure 4.7. Engineering 2D self-assembly. (a) Ribbon structure of a TTM hexamer with unpaired TM "seeds". **(b)** Close-up view of interfacial mutations at the junction of three assembled hexamers.

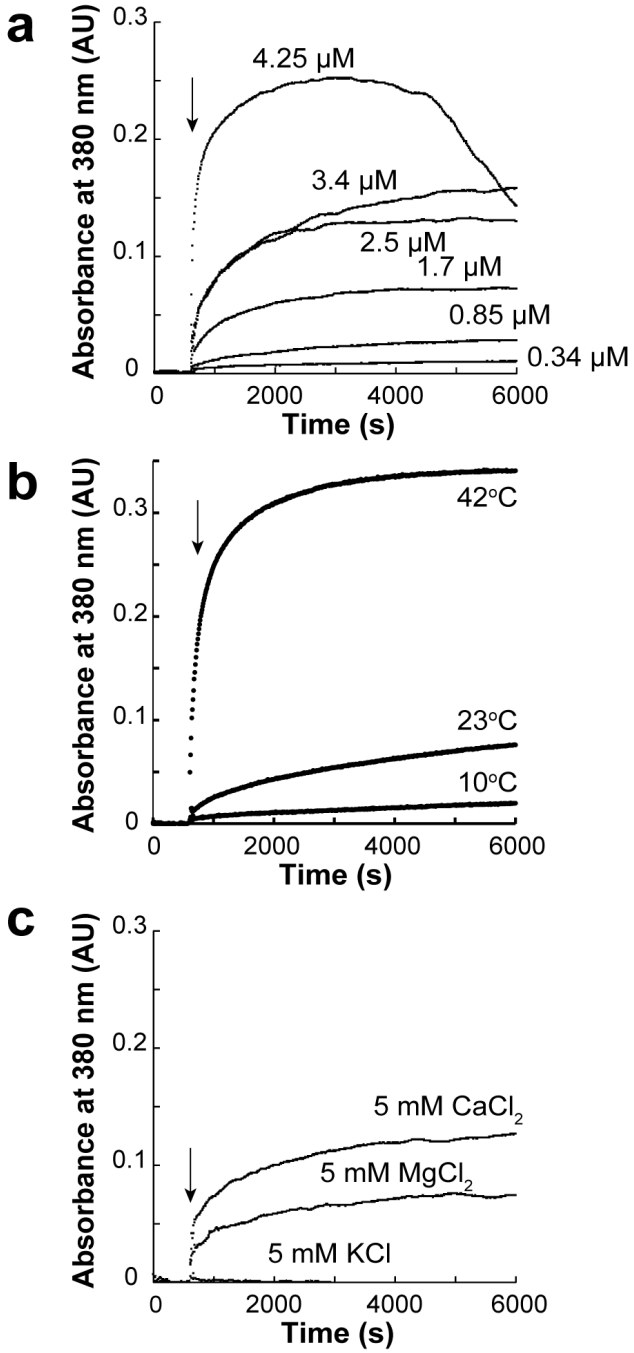

Figure 4.8. Influence of protein concentration, temperature and ion identity on TTM assembly. (a) Solutions of TTM at the indicated concentrations were supplemented with 5 mM $CaCl_2$ (arrow) and absorbance was recorded at 380 nm and 23°C as a function of time. **(b)** TTM (3.4 μM) held at the indicated temperatures was supplemented with 5 mM $CaCl_2$ (arrow) and absorbance was recorded as above. **(c)** TTM (3.4 μM) was supplemented the indicated ions (arrow) at 23°C and absorbance was recorded as above.

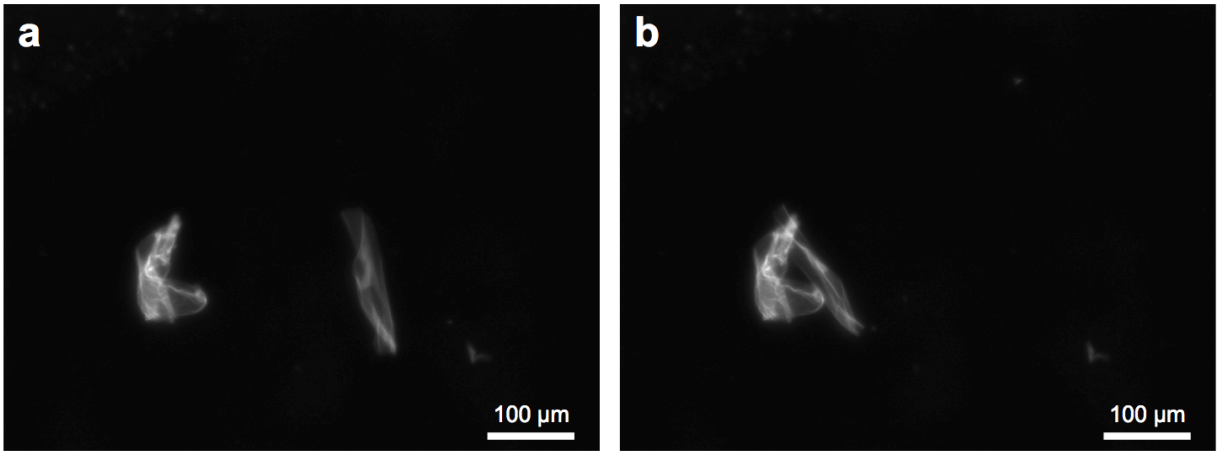

Figure 4.9. Collision of two free-floating self-assembled TTM structures. (a-b) Sequential fluorescence microscopy images of two large TTM sheets stained with Nile Red.

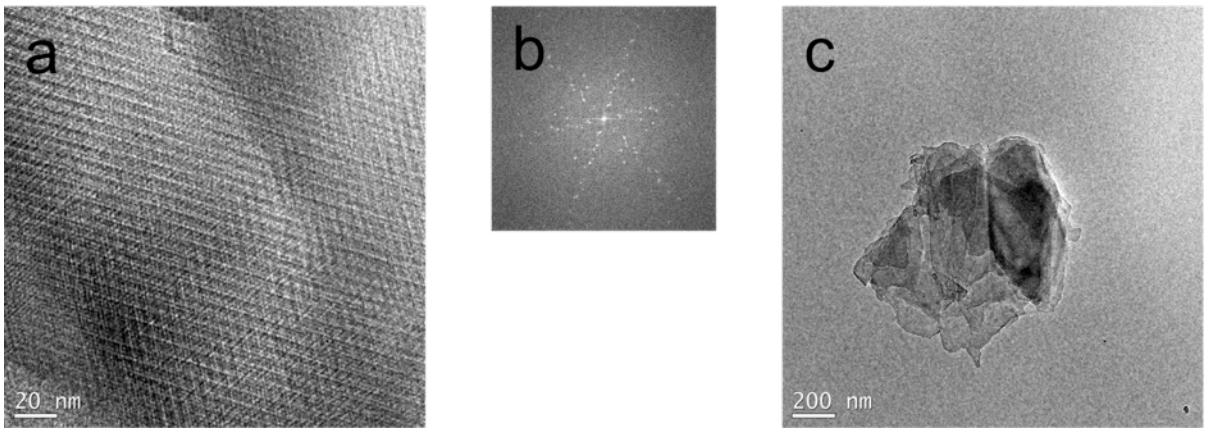

Figure 4.10. TEM images of calcium-treated TTM4. (a) TEM image of uranyl acetate-stained TTM4 structures obtained after incubating 1.7 μM of the Gly_4 TTM variant with 1 mM $CaCl_2$ at 4°C. The FFT of the image **(b)** reveals a hexagonal lattice with 7.25 nm unit cell. **(c)** Low resolution image of the structure in (**a**).

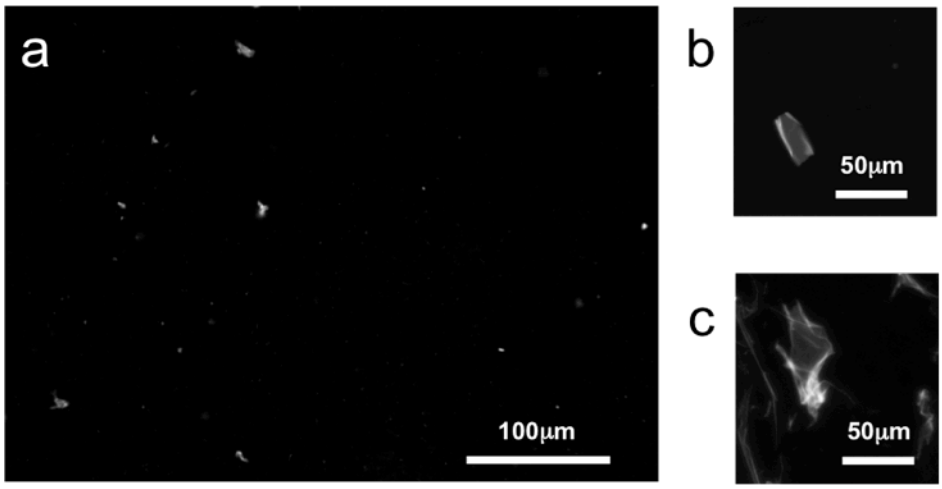

Figure 4.11. Fluorescent microscopy images of self-assembled TTM4 structures. Images were acquired after incubating 3.4μM of TTM with 1mM $CaCl_2$ at 23°C for 1hr and Nile Red staining. **(a)** Typical field showing self-assembled structures in the 1-10 μm range. **(b-c)** Larger arrays obtained under the same conditions.

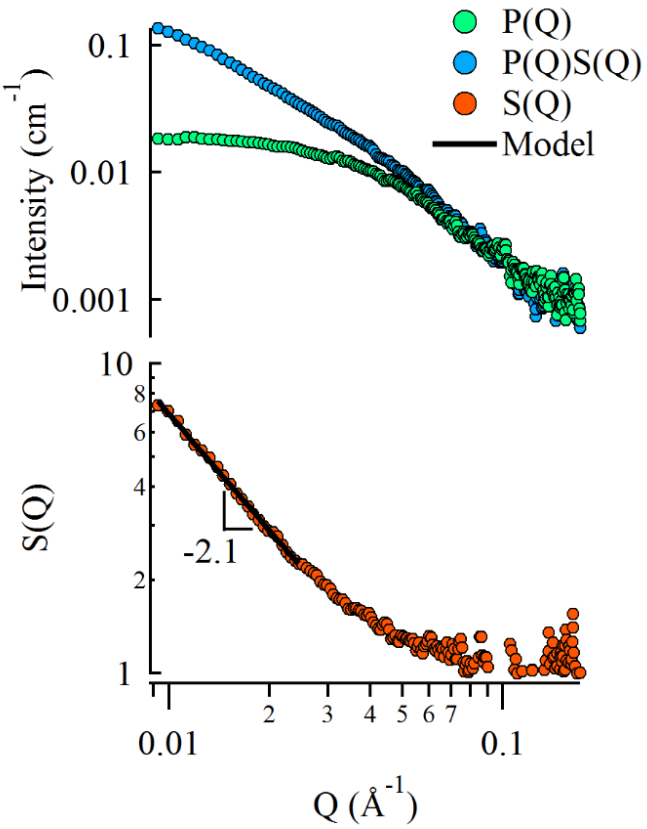

Figure 4.12. Deconvolution of the form factors of self-assembled TTM. The top panel shows the superimposition of unscaled scattering data from His-tagged STM415 (green circles) and calcium-treated TTM (blue circles). Overlap at high Q indicates that the form factor of the hexamer is preserved in the self-assembly. The bottom panel shows the ratio of assembly scattering to hexamer scattering (orange circles). A fit of the slope at low Q reveals a power law of -2.1 ± 0.1 which is consistent with a 2D object.

Chapter 5 TTM VARIANTS: PURIFICATIONS AND OTHER FACTS

5.1 INTRODUCTION

This chapter is a continuation of the work presented in Chapter 4 and helps paint a fuller picture of TTM two-dimensional arrays. Variants of TTM, TTM4 and TTMAB are also described here as are new purification schemes to improve purity and yield.

5.2 TT AND TTM: DIFFERENCES IN ASSEMBLY

TT is a tandem repeat protein linked by a hexaglycine linker with all native residues predicted to prevent interfacial packing intact. We compared the assembly kinetics of 3.4μM TT or TTM in the presence of 10mM calcium chloride and at 45°C by monitoring the change in absorption at 380nm over time, as in Chapter 4 [222]. Figure 5.1 shows that TT monomers polymerize at a faster rate than TTM. After about 25 minutes of assembly the TT reaction solution becomes cloudy and a white precipitate becomes visible. Indeed, the initial rise in scattering if followed by a noisy behavior due to the formation and eventual sedimentation of large structures. TTM, on the other hand, exhibits the expected assembly behavior characterized in Chapter 4.

SEM was used to visualize the structures formed by TT and TTM. SEM images were collected following deposition of a 5-10 nm thick Au-Pd film. Figure 5.2 shows that, under these experimental conditions, the non-mutated tandem repeat (TT) formed a porous 3D network of interconnected chains and nodules. By contrast, TTM samples appeared to polymerize into flatter structures that nevertheless still contained holes and exhibited topography (Figure 5.3). Clearly, eliminating electrostatic repulsion and steric hindrances through interfacial mutations has a profound effect on the topology of the final structure.

Both TT and TTM are isolated from the insoluble fraction of cell lysates. To understand whether monomers assemble into well-ordered structures in vivo or form large aggregates, thin sections of cells expressing TT or TTM were examined by TEM. Briefly, at 3 hours post induction, cells were fixed first with glutaraldehyde and then with osmium tetroxide, and stained with uranyl acetate. A dehydration process and resin infiltration followed before a thin section of cells was sliced and examined under the TEM. Figure 5.4 shows that both TT and TTM appear to form large aggregates that migrate to the poles of the cells. Although the resolution is poor, these structures appear amorphous rather than crystalline.

5.3 TTM VIARITANTS

5.3.1 *TTM4 variant*

TTM4 was briefly mentioned in Chapter 4 and will be talked about in more detail here. The length of the polyglycine linker connecting the TM repeats was originally specified to be six residues with the expectation that it would be long enough to span the gap to next unit cell in the lattice, while remaining short enough to prevent the two connected TM domains from lying in the same unit cell. This concept is shown in Figure 4.1c. The TEM data from the TTM assembly showed the presence of staked monolayers, presumably formed as a drying artifact but which may originate from growth in the third dimension because the linker is long enough to allow for a TTM lobe to act as a nucleation point above or below the 2D growth plane. TTM4 was built with a tetraglycine linker to preclude this possibility (Figure 5.5a).

After purifying TTM4 in the same manner as TTM, assembly kinetics were studied. Figure 5.5b shows how the scattering of a 3.4 μM solution of TTM4 depends on calcium concentration. Results are similar to those obtained with TTM (Figure 4.3), although the overall extent of aggregation seems to be slightly smaller at high calcium concentrations for TTM4.

Using the method of initial rates, we found that, like TTM, TTM4 exhibits apparent second order polymerization kinetics with respect to $CaCl_2$. The dependency on protein concentration was not examined for TTM4 because of the high protein cost of these experiments. Nile Red staining images (Figure 4.11) revealed that the structures produced with TTM4 were similar to those obtained with TTM and TEM (Figure 4.10) and confirmed hexagonal packing with a spacing of 7.25nm.

Overall, these data indicate that TTM4 behaves almost identically to TTM. The kinetics of assembly can be controlled by multiple parameters including temperature, cation concentration, and protein concentration. These parameters lead to identical hexagonal packing as seen by TEM.

5.3.2 *TTMAB variant*

A new design proposed by Dr. Possu Huang in the Baker laboratory, is based on an A-B interaction instead of an A-A interaction in the current model. The 130 residue long A motif will contain four new mutations that include the removal of the remaining cysteine residues by introduction of the C47V and C118A substitutions, a R89L mutation that will remove positively charged arginine side chains that point directly at each other at the unit cells interface, and a stabilizing A98F mutation at position 98. The final change to the A motif is the deletion of the last two amino acids (I131 and A132). A more rigid linker (GTGL) was used to connect the A domain to the B motif, which carried the C181V and C252V mutations along with the insertion of a tryptophan at position 264. The inserted tryptophan formed a hydrophobic cluster with A98F mutation in the A motif. This hydrophobic contact was designed to make the lattice more rigid and promote 2D growth. These changes are illustrated in Figure 5.6.

The gene encoding this protein was cloned into pET-29b(+) in order to make use of a C-terminal His-tag. In order to remove the His-tag, a TEV cleavage site was encoded at the 3' end of the DNA sequence. The Tobacco Etch Virus protease, TEV, recognizes the a linear epitope, ENLYFQ(G/S), in a protein (in our case an S is used instead of a G). The TEV protease cuts the epitope between the Q and S with high activity. However, TEV cleavage sites are most useful when inserted at the N-terminus of a protein so that only one amino acid is left behind. With a C-terminal insertion, an additional nine amino acids (GSTENLYFQ) are introduced at the C-terminus, six from the TEV epitope three from the spacer. This is not ideal but should not be detrimental to assembly.

TTMAB-His was expressed and isolated from the insoluble cell fraction after lysis. Inclusion bodies were washed, denatured and the protein was purified under denaturing conditions by Ni-NTA chromatography (Figure 5.7a). After refolding by dialysis, AcTEV, a commercially available TEV protease containing a His-Tag, was used to cleave the His-Tag to yield TTMAB (Figure 5.7b). AcTEV and the small amount of uncleaved TTMAB-His were removed by another round of Ni-NTA chromatography. Pure TTMAB was recovered in the flow-through.

Next the kinetics of assembly were examined by light scattering using 3.4μM protein with 5mM calcium chloride at 23°C. Figure 5.7c shows that TTMAB-His polymerization is very rapid upon addition of calcium ions. As with TT, a visible precipitate forms at the bottom of the cuvette. The scattering pattern also matches that of TT, except that the absorbance is roughly 30% higher for TTMAB-His than TT. It should nevertheless be noted that the TTMAB-His reaction was performed at 23°C verses 45°C for TT and with half the calcium concentration. Both of these changes should slow down assembly. This leads to the conclusion that TTMAB-

His is forming large aggregates, which is confirmed by fluorescence microscopy of Nile Red stained structures (Figure 5.7d). Thus, the His-tag appears detrimental to TTMAB assembly.

When the axes are adjusted in Figure 5.7c, the TTMAB assembly rate resembles those of TTM and TTM4 at 3mM calcium. This can be seen in Figure 5.7e. Fluorescence microscopy indicates that TTMAB forms smaller particles with few larger, ~25μm, structures (Figure 5.7f). TEM imaging was performed on 1.7μM protein samples incubated with 1mM calcium chloride at 4°C for 18 hours. Although Figure 5.8 shows hexagonal periodicity stacked layers still appear to form. The lattice spacing calculated from the image FFT is 7.06nm perhaps because a more ridged linker, extra mutations to remove electrostatic repulsion and steric hindrances, and the hydrophobic cluster leads to tighter packing of unit cells in assembled TTMAB.

In short, TTMAB is another variant that can assemble into a hexagonal array over micrometers. The C-terminal domain is accessible and a 9 residue-scar left after TEV protease cleavage does not interfere with assembly but a His-tag does.

5.3.3 *Proposed model: circular permutation*

Thus far, design has focused on fusing the C and N terminals of the STM4215 gene to one another. Another option proposed by the Baker lab is a circular permutation where the gap between the unit cells is bridged twice (Figure 5.9). This will allow for a more tightly packed lattice to form and should help prevent 3D growth. The N- and C- terminus are still located on the "back" side of the protein and both seem to be accessible. A large, flexible loop is created on the "front" side that could be used to insert a solid binding peptide. This could allow for a dual functionalization to this lattice where each side has its own unique function. Dr. Possu Huang, of the Baker group, is still refining this model.

5.4 NEW PURIFICATION SCHEMES

Purification from the aggregated state is time consuming with long inclusion body cleanup and low yield denaturing SEC chromatography steps. His-tagging proved useful for TTMAB purification but the TEV cleavage site is better at located at the N-terminal. This section describes the fusion of a His-tag at the N-terminus of TTM variants that is cleavable with either thrombin or the TEV protease.

5.4.1 *pET-28a(+)*

pET-28(+) was selected because it encodes a N-terminal His-tag followed by a short SSG linker and a thrombin cleavage site. The thrombin protease cleaves between R and G in the amino acid sequence LVPRGS, leaving a GS fused to the N-terminus of the desired protein. A third amino acid, H, is encoded at the 5' end of pET-28(+).H is added because MH can encode for an *Nde*I site, which is the restriction site that genes encoding all TTM variants have at their 5' end. This allowed for easy insertion of genes for WT, T, TM, TT, and TTM into the pET-28a(+). The resulting variants were named His-WT, His-T, His-TM, His-TT and His-TTM.

The expression of these proteins was similar to proteins without His-tags (Figure 5.10a). Purification of soluble variants (His-WT, His-T and His-TM) was straightforward and gave high yields of protein. The His-WT was used in Chapter 4 for SAXS experiments and can be seen in Figure 4.5. However, none of the soluble proteins have the ability to create large sheets.

The His-TTM expression and inclusion body clean up procedures were the same as with TTM. Several different Ni-NTA purifications were tried but a hybrid purification approach was found to give the best results. While some contamination was present (Figure 10b) the yield was high enough to try cleaving the His-tag with thrombine. However, only about 33% of the protein was cleaved and no improvement was achieved with different buffers, temperatures and times.

Moreover, during the purification step to remove His-TTM from cleaved His-TTM all products stuck to the Ni-NTA resin. At this point the overall yields of cleaved His-TTM were deemed unsuitable for further experimentation.

5.4.2 *pCWHM100*

A new plasmid pCWHM100, was created, using a pET-22b(+) variant in which the pelB leader was removed and replaced with a DNA sequences encoding a His-tag, a GGGS linker and a TEV cleavage site. Thus, after TEV cleavage, a single S will be present at the end of TTM. The pCWHM100 plasmid is designed so that any gene with an *Nco*I restriction site at the 5'end will be in frame with the TEV cleavage site. All of the restriction sites in the multiple cloning region of pET-22b(+) remain accessible. To account for the possibility that the N-terminus of TTM is not easily available, for TEV cleavage, we built another plasmid, pCWHM500, which is identical to pCWHM100 except that the TEV cleavage site is followed by a flexible GGGS linker. Thus, with this plasmid, a GGGGS sequence will be appended to the N-terminus of the desired protein.

Expression of HisTEV-TM, HisTEV-TTM, HisTEV-TTM4 and HisTEV-TTMAB was at the same level as proteins lacking HisTEV N-terminal sequences (Figure 11a). HisTEV-TTM was selected to determine if this system worked better than the His-thrombin system.

HisTEV-TTM was produced and inclusion bodies cleaned as described but 6M guanidine hydrochloride instead of 8M urea was used to solubilize aggregated material. The hybrid Ni-NTA purification provided near-homogeneous HisTEV-TTM (Figure 5.11). HisTEV-TTM was mixed with AcTEV and cleavage at 30°C for 3h gave greater than 75% yields, perhaps because the AcTEV was old and had lost some activity. Nevertheless, this far exceeds what we previously achieved with His-thrombin of His-TTM.

Early results with HisTEV-TTM are very favorable. This purification scheme is easy and leads to high purity and yields. This new purification train holds the promise for exploring and realizing the promise of the TTM platform.

5.5 CONCLUSIONS

In this chapter, we showed that the interfacial mutations predicted by Rosetta are critical for 2D assembly as illustrated by the fact that TT produces large aggregates in reconstitution experiments and in spite of the fact both proteins appear to form inclusion bodies *in vivo*. Two other variants, TTM4 and TTMAB, were also shown to produce hexagonal arrays but did not seem to have an advantage over TTM. A new plasmic, pCWHM100 will allows for the rapid and high yield purification of TTM variants through Ni-NTA chromatography and TEV cleavage. This availability of a simple and rapid purification train should enable full realization of the TTM platform potential in bionanotechnology.

5.6 MATERIALS AND METHODS

5.6.1 *DNA manipulations*

5.6.1.1.1 *TTM4*

The gene encoding TM, the truncated and mutated version of STM4215 (Chapter 4), was excised from pSTM4215TM-22 by double digestion with *Nde*I and *Bam*HI. Half of the DNA was saved as "STM4215TM Back" for later splicing of the second STM4215TM into the STM4215TTM4 gene. The other half was digested with *Bst*NI to remove the stop codon and 32 bp from the 3' end of the TM gene. This product was named "STM4215TM Front".

Two "artificial" genes, A and B, were created to allow for an easy sequence of GENE splicing to occur, by which, the STM4215TTM gene was created [220, 223]. Gene A will

overlap STM4215TM Front by 22 base pairs and gene B by 18 base pairs. Gene B will overlap STM4215TM Back by 27 base pairs. This amount of overlap is ideal for GENE splicing.

Two primers were annealed to created gene A and gene B. The annealing solution contained 30mL of 10mM of the forward and reverse primer for the desired gene, 1x PCR Rxn Buffer –$MgCl_2$ (Invitrogen), 10mM $MgCl_2$ (Invitrogen), which was placed in a microcentrifuge tube. The tube was place in a boiling water bath for 10 minutes, after which, the heating source was removed and the solution and the water bath were allowed to cool to room temperature. The products purity and concentration was checked by low-melting-point agarose electrophoresis. The primers for gene A were 5'-GCAAACTTCTGCCGTTACGCCCTGGATATGACCGCCAGCACCCTGTTTATCGCGGGTGG-3' and 5'-CCACCCGCGATAAACAGGGTGCTGGCGGTCATATCCAGGGCGTAACGGCAGAAGTTTGC-3' and gene B were 5'-CCTGTTTATCGCGGGTGGAGGTGGAGGTGGGATGGAAACATTAGCTGTTATACATACC -3' and 5'-GGTATGTATAACAGCTAATGTTTCCATCCCACCTCCACCTCCACCCGCGATAAACAGG -3'.

Heckman's 2007 paper describes the outline for the GENE splicing process [220]. The first step is creating the AB gene by splicing together gene A and gene B. Here equal parts of gene A and B, a forward primer for gene A, 5'-GCAAACTTCTGCCGTTACGCCCTGG-3', a reverse primer for gene B (Breverse), 5'-GGTATGTATAACAGATAATGTTTCCATCCC-3', nucleotides, pfu ultra buffer, and pfu ultra were combined for PCR. The PCR products were separated with low-melting-point agarose electrophoresis and the AB gene was recovered with a

gel extraction kit. Next, the STM4215TM Front was spliced to AB using the same process. The primers used were STMforward primer, 5'-TATATACATATGGAAACATTATCTGTTATAC-3', and Breverse. This product was spliced with STM4215TM Back using primers STMforward and Treverse, 5'-TTAAGGATCCCTATCACG CGATAAACAGGG-3'. After the desired product was purified, it was digested with NdeI and BamHI, ligated into pET-22b(+) and transformed into Top10. The plasmid was named pSTM4215TTM4-22. After verification via sequencing, the plasmid was transformed into BL21(DE3).

5.6.1.1.2 *TTMAB*

The TTMAB gene was ordered through Genescript. The gene was placed on a pET-29b(+) plasmid using the restriction site of *Nde*I and *Xho*I. This allows the 3' His-tag on the plasmid to be used. A C-terminal TEV cleavage site was also constructed into the genetic code in order to remove the His-tag by TEV protease. The plasmid was transformed into Top10 and BL21(DE3) strains.

5.6.1.1.3 *pET-28a(+)*

The following plasmids, pSTM4215-22, pSTM4215-T-22, pSTM4215-TM-22, pSTM4215-TT-22, pSTM4215-TTM-22, were purified by and digested by *Nde*I and *Bam*HI. The genes were ligated into pET-28a(+) and transformed into Top10 and BL21(DE3) strains. This process created a plasmid that encodes for a series of N-terminally His-tagged proteins named His-WT, His-T, His-TM, His-TT, and His-TTM. The resulting plasmids are named pHis-STM4215-28, pHis-STM4215-T-28, pHis-STM4215-TM-28, pHis-STM4215-TT-28, pHis-STM4215-TTM-28.

5.6.1.1.4 *pCWHM100, pCWHM500 and pHisTEV-TTM*

The creation of pCWHM100 started with these two primers 5'-CGTATACATATGGGCAGCAGCCACCACCACCACCACCACGGCGGCGGCTCTGAAAA CTTATACTTCCAGTCCATGGACTGACTGGGATCCTATACG -3' and 5'-CGTATAGGATCCCAGTCAGTCCATGGACTGGAAGTATAAGTTTTCAGAGCCGCCGCC GTGGTGGTGGTGGTGGTGGCTGCTGCCCATATGTATACG -3', which were annealed to create a DNA cassette encoding for a His-tag followed by a flexible linker to a TEV cleave sight (MGSSHHHHHHGGGSENLYFQSM) and *Nde*I and BamHI sites. The annealing reaction consisted of 30mL of forward and reverse primer (at 10 μM concentration) 1x PCR Rxn Buffer –$MgCl_2$ (Invitrogen), and 10mM $MgCl_2$ (Invitrogen). The mixture was boiled for 10 min and allowed to cool to room temperature. The cassette was digested with *Nde*I and *Bam*HI and the fragment cloned into the same sites of pET-22b(+) creating pCWHM100.

The same process was followed to create pCWHM500, except the cassette was ligated into pCWHM100, the restriction sites of *Nde*I and *Nco*I were used and the primers were 5'-CGTATACATATGGGCAGCAGCCACCACCACCACCACCACGGCGGCGGCTCTGAAAA CTTATACTTCCAGGGTGGTGGTGGTTCCATGGACGTACGT-3' and 5'-ACGTACGTCCATGGAACCACCACCACCCTGGAAGTATAAGTTTTCAGAGCCGCCGCC GTGGTGGTGGTGGTGGTGGCTGCTGCCCATATGTATACG -3'. This plasmid, pCWHM500, encodes for the same amino acid sequence as pCWHM100 with a 4 amino acid flexible linker following the TEV cleavage site, (MGSSHHHHHHGGGSENLYFQSGGGGM). Both pCWHM100 and pCWHM500 were transformed into Top10.

To construct HisTEV variants of TM, TTM, TTM4 and TTMAB the plasmids encoding for these genes were isolated. The TM, TTM, TTM4 genes were amplified by the following

primers 5'- ATAGCGATCCATGGAAACATTAGCTG- 3' and 5'- GCTCGAATTCGGATCCCTATCA -3'. The forward primer changes the *Nde*I site to an *Nco*I while the reverse was set just downstream of the stop codons. For TTMAB the following primers were used were 5'- ATATATATCCATGGAAACCCTGGCCG -3' and 5'- TATATAGGATCCCTATCATGCGATCCACAGCGTTG-3'. The forward primer made the same change a *Nde*I site to an *Nco*I. The reverse primer, however, added a stop codon before a *Bam*HI site position to remove the section that encodes for C-terminal TEV cleave site and His-tag so that the TTMAB was terminated as the model was designed. The PCR products were run onto an agarose gel by electrophoresis and the desired products were removed and digest by *Nde*I and *Bam*HI along with pCWHM100. These products were run back onto an agarose gel and the desired inserts, TM (~400bp), TTM, TTM4 and TTMAB (~800bp), were ligated in ~5.5 kbp backbone from pCWHM100. These products were transformed into Top10 and verified by sequencing resulting plasmids pHisTEV-TM, pHisTEV-TTM, pHisTEV-TTM4, and pHisTEV-TTMAB. All were then transformed into BL21(DE3).

5.6.2 *Expression, cell lysis, and clean of the inclusion bodies*

BL21(DE3) transformants were grown at 37°C in LB medium supplemented with proper antibiotic (100 µg/mL carbenicillin for pET-22(b)+ and pCWHM100 variants or 50µg/mL kanamycin pET-28(b)+ or pET-29b(+)) to mid-exponential phase and protein synthesis was induced with 0.5 mM IPTG. 3 hours post induction, the cells were harvest by centrifugation (3,000g) for 10min. Paste from 500 mL of culture was washed with 100 mL of buffer A (20mM Tris-HCl, pH 8.5, 500mM NaCl), resuspended in 35 mL of buffer B (Buffer A with 5mM EDTA, 0.1mM PMSF, 0.1% NaN_3, 1mM DTT) and disrupted by 6 cycles of sonication at 3 min intervals on a Branson Sonifier 450 operated at 30% duty. Note that for all Ni-NTA purification,

EDTA was left out of the buffers. Insoluble material collected by centrifugation at 12,000*g* for 20 min was washed in 35 mL of buffer B supplemented with 2M Urea and 1% Triton X100 followed by two more washes with buffer B with 1% Triton and a final wash with Buffer B alone. The clean pellets were store at -20°C to await purification.

5.6.3 *Purification and protein prep*

5.6.3.1 TTM4

TTM4 shares the same purification train as TTM. For more information please see Chapter 4.

5.6.3.2 TTMAB-His

The clean pellet from the inclusion body cleaning procedure was denatured with 20mM Tris, 8M urea, 500mM NaCl, 10mM imidazole at pH 8.3 for 4 hours. The denatured protein placed in the centrifuge at 10,000g for 20min and the supernatant was loaded onto a Ni-NTA agarose column developed in the buffer describe above at a flow rate of 0.5 mL/min. The column was subjected to a 50mM imidazole wash prior to elution of TTMAB-His with 250mM imidazole. The protein was dialyzed four times overnight against 20mM Tris pH 8.3. Aggregates were removed by centrifugation (10,000g) and stored at -20°C. To remove the His-tag, AcTEV (Life Technologies) was added to TTMAB-His with provided buffers and the cleavage took place at 30°C for 2 hrs. The protein mixture was then run over a Ni-NTA agarose column where the flow through, which contains TTMAB, was kept while AcTEV and the small amounts of un-cleaved TTMAB-his remained on the column. TTMAB was subjected to dialysis, centrifuged (10,000g) to remove aggregates, placed in aliquots and stored at -20°C at a concentration of ~0.5mg/ml.

5.6.3.3 pET-28a(+)

For His-WT, His-T and His-TM, which are all soluble, after cell lysis by sonication and following centrifugation at 10,000*g* for 20 minutes to remove insoluble material, clarified extracts were supplemented with 10mM imidazole and loaded onto a Ni-NTA agarose column developed in 20mM Tris, 500mM NaCl, 10mM imidazole at pH 8.3 at a flow rate of 1 mL/min. The column was subjected to a 50mM imidazole wash prior to elution of the desired protein with 200mM imidazole. The protein was dialyzed 3 times overnight against 20mM Tris pH of 8.3. Aggregates were removed by centrifugation (10,000g) and the proteins were stored at 20 °C at a concentration of ~1mg/ml.

A hybrid purification scheme was found to work the best for the insoluble proteins, *e.g*. His-TTM. The clean pellet from the insoluble fraction was denatured with 20mM Tris pH 8.3, 8M urea, and 500mM NaCl for 4 hours. The denatured protein was placed in the centrifuge at 10,000g for 20min and the supernatant was loaded onto a Ni-NTA agarose column developed in the buffer describe above at a flow rate of 0.5 mL/min. The column was subjected to 20mM Tris pH 6.3, 8M urea, and 500mM NaCl to wash contaminating protein off the column. The column was return to the original buffer and then subjected to 20mM Tris pH 8.3, 500mM NaCl. After greater than 15 column volumes had passed the column was subject to a 50mM imidazole wash prior to elution of the desired protein with 200mM imidazole. The protein was dialyzed 3 times overnight against 20mM Tris pH of 8.3. Aggregates were removed by centrifugation (10,000g) and the proteins were stored at 20°C at a concentration of ~0.5mg/ml. Thrombine CleanCleave^{TM} Kit (Sigma-Aldrich) was used for the cleavage of the His-tag.

5.6.3.4 HisTEV-TTM

Only HisTEV-TTM has been purified so far. The same hybrid purification was used except the 8M urea was replaced by 6M guanidine hydrochloride because 6M guanidine

hydrochloride solubilized more of the clean pellet. During the spin to remove aggregates from the purified protein at end of the purification, all the protein crashed out. The protein was then subject to 20mM Tris pH 8.3, 6M guanidine and 500mM NaCl and then subjected to another round of dialysis. This time when aggregates were removed by centrifugation (10,000g) no visible pellet showed up. The proteins was placed into aliquot and stored at 20 °C at a concentration of ~0.5mg/ml. Right now AcTEV (Life Technologies) is being used to test the accessibility of cleaving the His-tag from TTM.

5.6.4 *Absorption spectrophotometry*

For absorption spectrophotometry, samples were prepared in a quartz cuvette by aliquoting TTM (usually at a 3.4mM final concentration) in a 500mL final volume. The cuvette was next placed on the thermostrated stage of a Beckman Coulter DU640 spectrophotometer and samples were allowed to equilibrate at the indicated temperature for at least 10 minutes. The kinetics of apparition of scattering species (presumably self-assembled protein sheets) was monitoring the absorbance at 380nm with respect to time. $CaCl_2$ (usually 12.5mL of a 205mM stock for a final concentration of 5mM) was added as indicated.

5.6.5 *SEM*

Samples for the SEM were prepared by adding 250μL of 3.4μM TTM or TT to a microcentrifuge tube. Samples were placed in a heating block at the desired temperature. Samples were allowed to incubate at the desired temperature for over ten minutes before 6.25μL of stock calcium chloride was added to bring the final concentration of calcium to the predetermined amount. Samples were then allowed to assemble for two hours. Next 25μL of sample was placed on a clean silicon wafer for five minutes. The sample was wicked away by placing a KimWipe at the corner of the sample. Then, the wafer was dried with nitrogen gas.

Next the silicon was dipped into five different falcon tubes containing 40mL of ultra pure water before the wicking and drying procedure was repeated. Samples had 5-10nm of Au-Pd deposited on them so that images could be seen in the FEI Sirion SEM.

5.6.6 *TEM of TTM variants*

TEM sample prep was the same as describe in Chapter 4.

5.6.7 *Thin sectioning of cells for TEM*

After 3 hours of induction, cells harboring pSTM4215-TT-22 and pSTM4215-TTM-22 were pelleted by centrifugation at 3,000g for 10min. Primary fixation happened by resuspending pelleted cells in 2.5% glutaraldehyde (10ml) and incubating for 5 min by rocking. Cells were pelleted as describe above and the supernatant was discarded. Then cells were resuspended in 1.5mL of 2% glutaraldehyde, 0.1M sodium cacodylate with a pH of 7.2 in a microcentrifuge tube and rocked a for 5 minutes before being refrigerated overnight. The next day cells were pelleted and twice rinsed in phosphate buffer pH 7.5 before being placed in 0.1M sodium cacodylate pH 7.2 and mixed for 10 minutes. Cells were pelleted again and the supernatant was discarded. Secondary fixation started with the resuspension of cells in 1ml of 1% osmium tetroxide and incubate at 4°C for 1 hour. This was followed by two rinses with phosphate buffer. The cells were then resuspened in 0.1M sodium cacodylate pH 7.2 for 5min and were pelleted again. Next cells were rinsed in water three times. This was followed by en bloc staining in 1% uranyl acetate solution and kept at 4°C overnight. Cells were pelleted, washed in water and pelleted again. Dehydration of the cells happened by repeated pelleting and resuspending for 10 min on ice or later at -20°C in increasing amounts until cells were in 100% ethanol. Then cells were resupended in 100% acetone for 10 minutes at room temperature and pelleted. Finally, the cell

were embedded in a resin and dried. Then thin sections were cut and examined under a JEOL 1230 Transmission Electron Microscope with an accelerating voltage of 120keV.

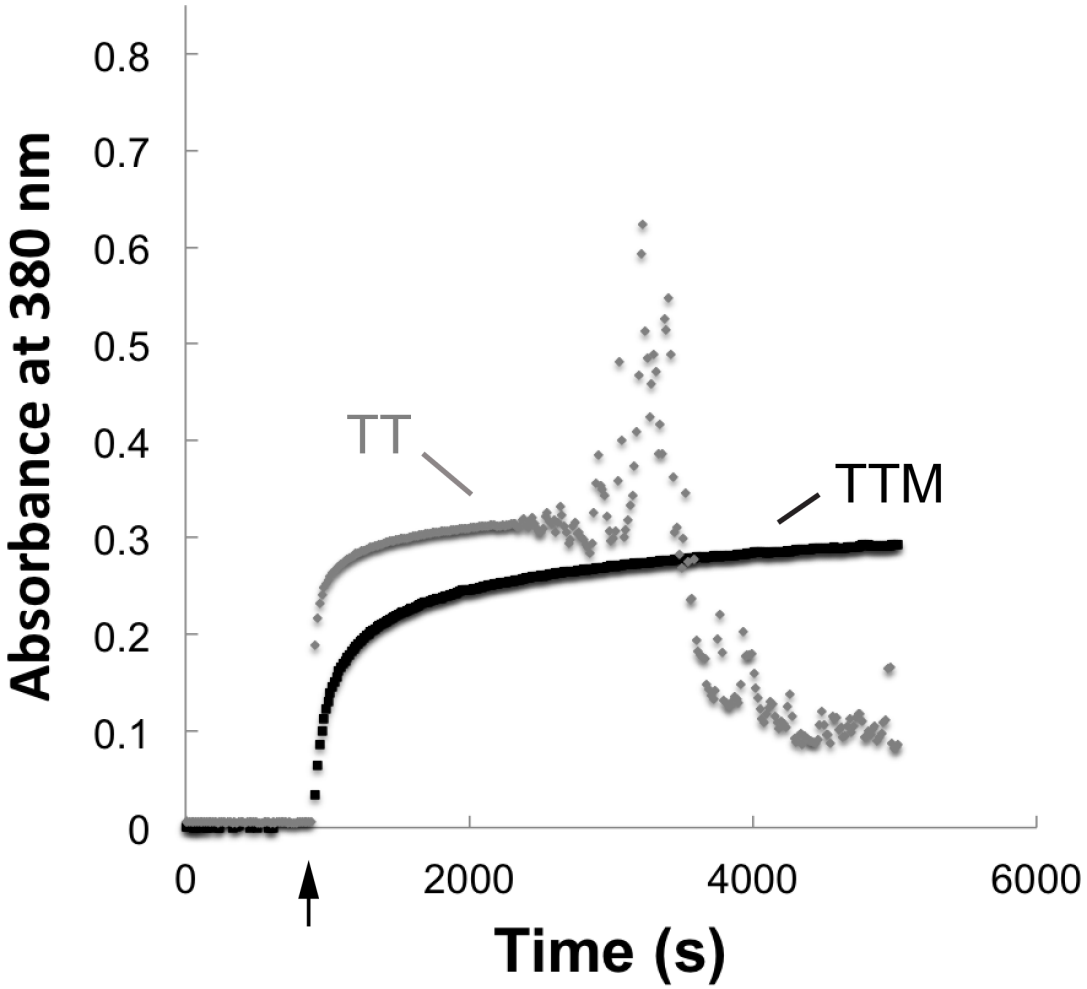

Figure 5.1. Difference in assembly kinetics between TT and TTM. 3.4μM of TT or TTM was incubated with 10mM $CaCl_2$ at 45°C and light scattering was monitored over time at 380 nm. The arrow marks the time of $CaCl_2$ addition.

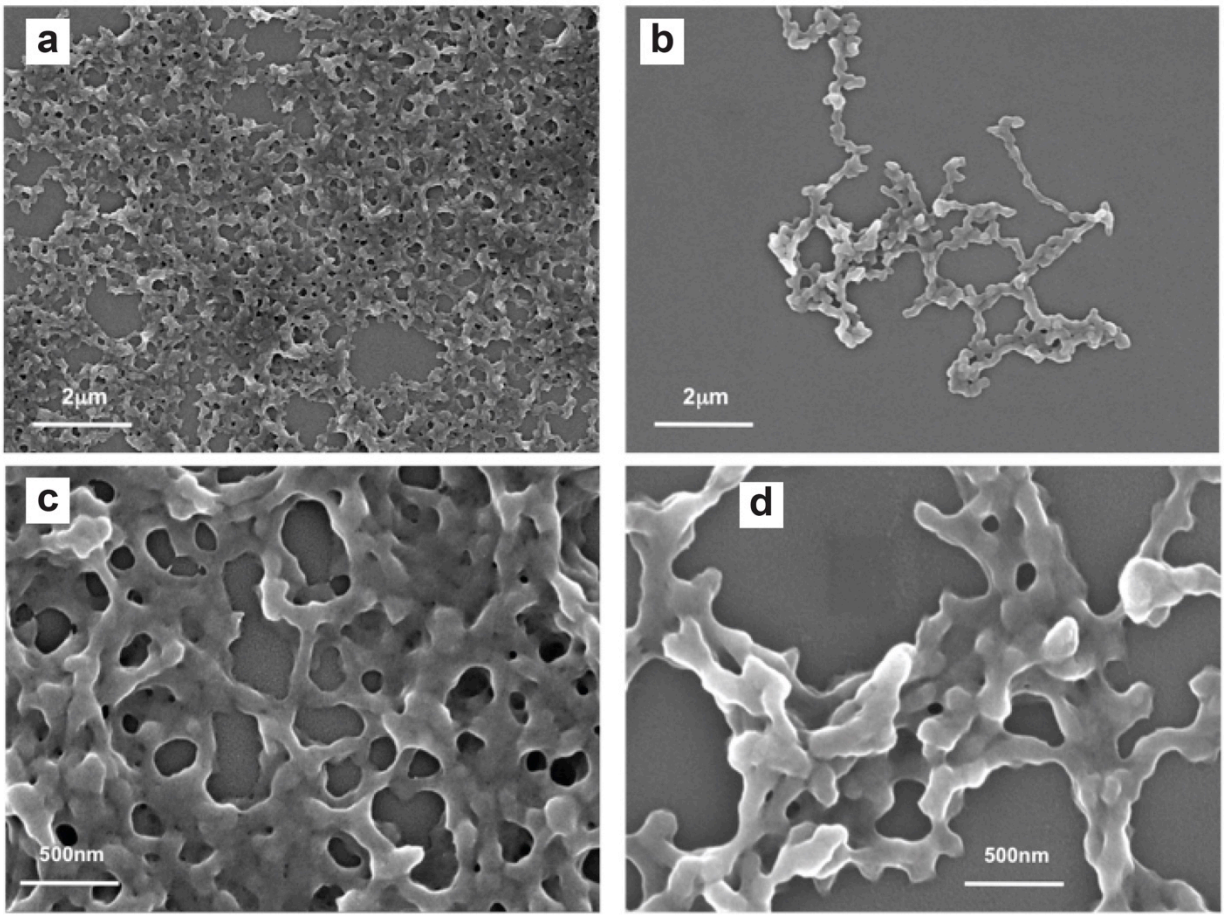

Figure 5.2. Visualization of self-assembled TT structures by SEM. 3.4μM TT was incubated at 45°C for two hours with 10mM $CaCl_2$ before being deposited onto a silicon wafer and coated with a 5-10nm thin film of Au-Pd.

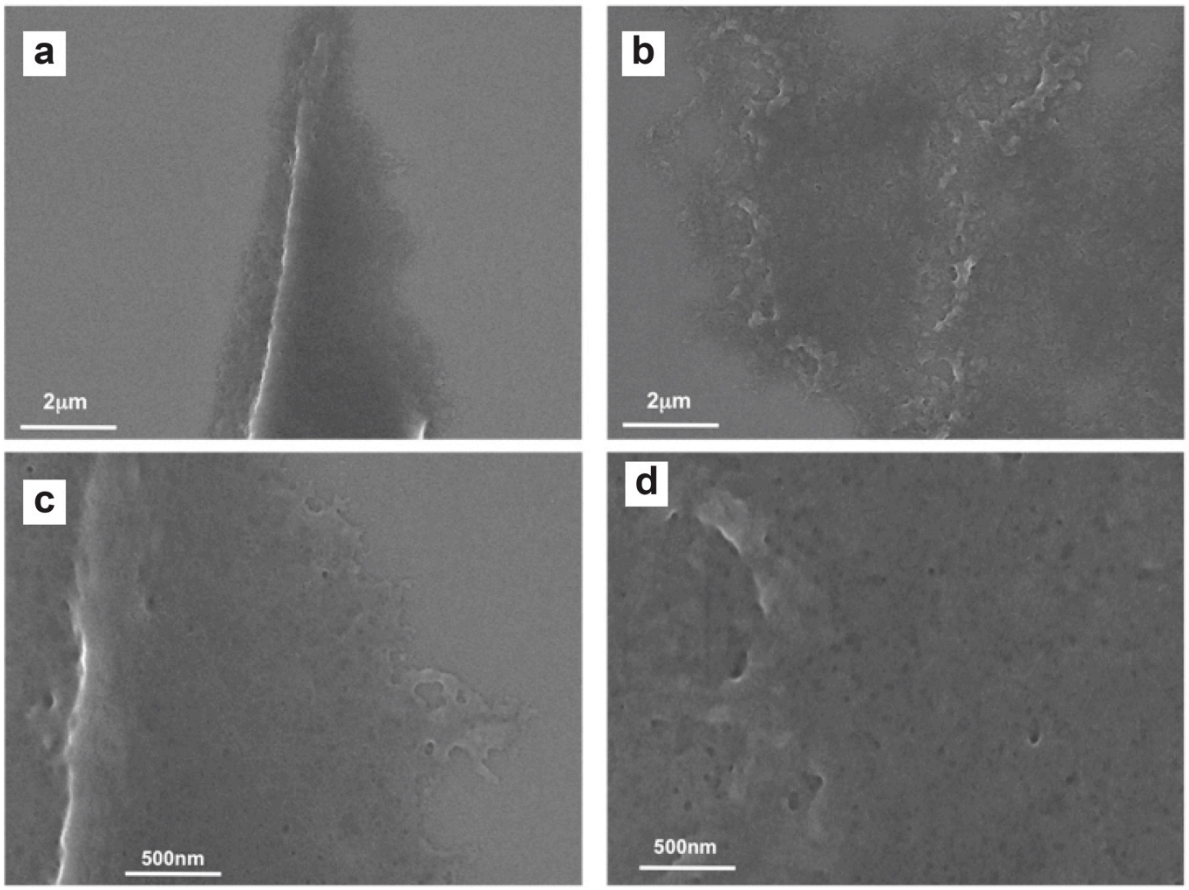

Figure 5.3. Visualization of self-assembled TTM structures by SEM. 3.4μM TT was incubated at 45°C for two hours with 10mM $CaCl_2$ before being deposited onto a silicon wafer and coated with a 5-10nm thin film of Au-Pd.

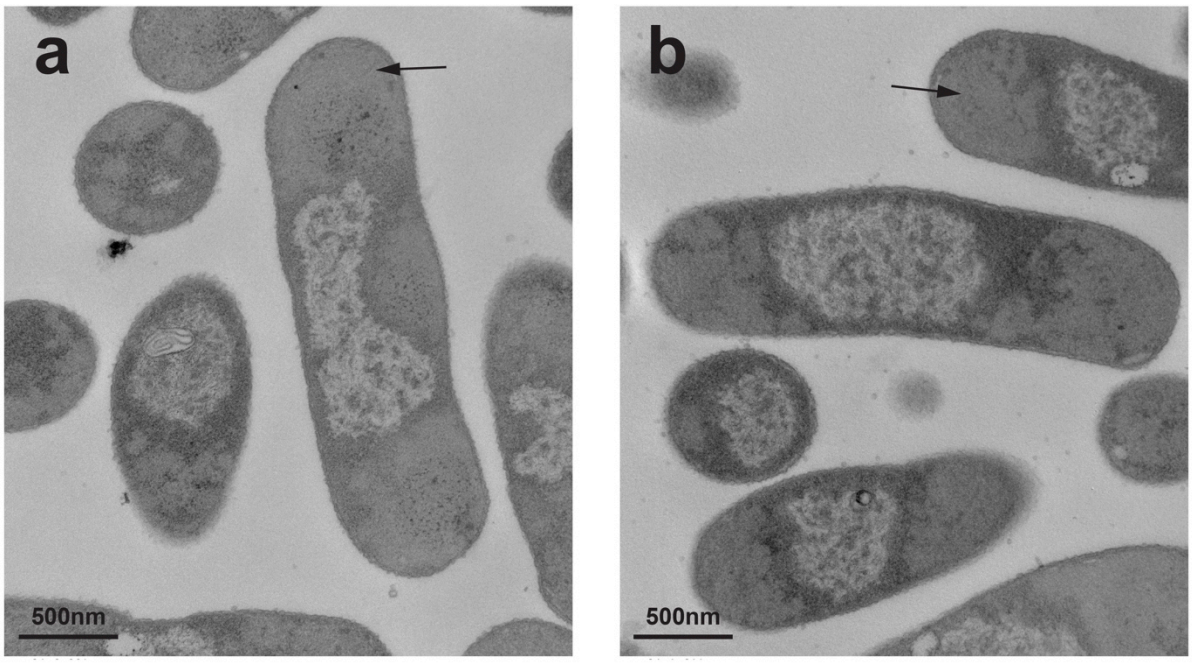

Figure 5.4. TEM images of thin sections of *E. coli* BL21(DE3) cell overexpressing TT or TTM. Cells were grown at 37°C for 3h post induction. **(a)** TEM image of cells producing TT. **(b)** TEM image of cells producing TTM. Arrows point to likely inclusion bodies at the poles of the cell.

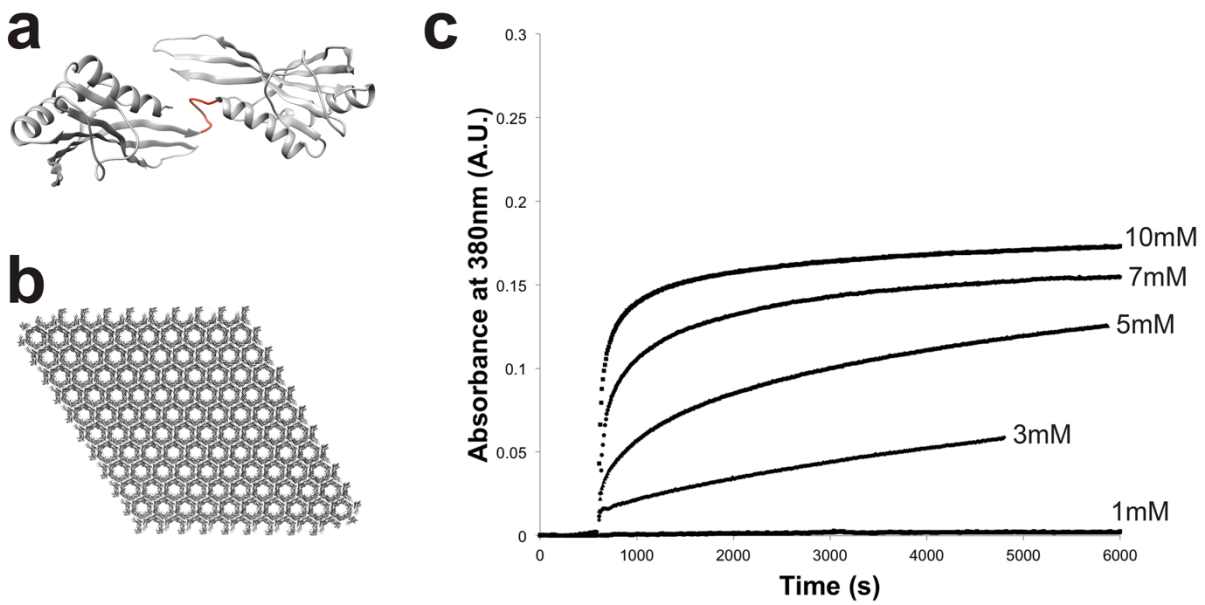

Figure 5.5. TTM4 kinetics of assembly with respects to calcium concentration. (a) Ribbon structure of TTM4. The flexible linker (red) was shrunk from 6 to 4 glycine residues. **(b)** The TTM4 monomer should assemble into the 2D hexagonal structure proposed for TTM. **(c)** Influence of the $CaCl_2$ concentration on the scattering of a 3.4 µM solution of TTM4 at 23°C. At 600 seconds $CaCl_2$ was added to start the polymerization process.

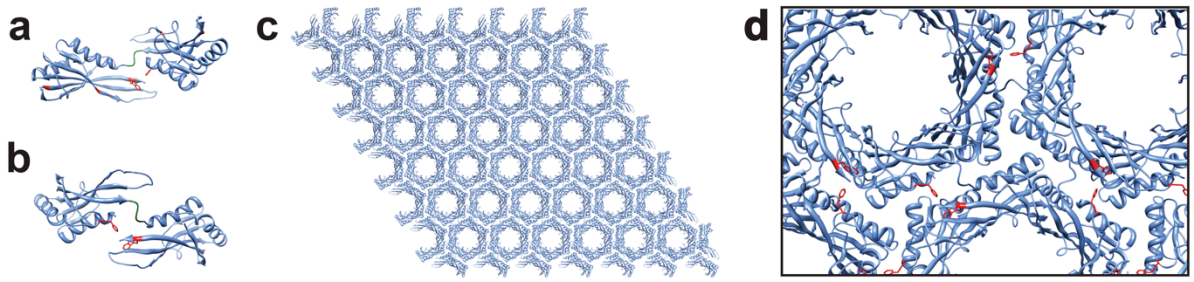

Figure 5.6. TTMAB Model. Ribbon structures of TTMAB in front **(a)** and back orientations **(b)**. The rigid linker GTGL is shown in green. The hydrophobic cluster that is formed by the mutation of A98F and the insertion of 264W is shown in red. Panel **(c)** shows the 2D array that should be formed with TTMAB. **(d)** A close up of the packing of the 2D array.

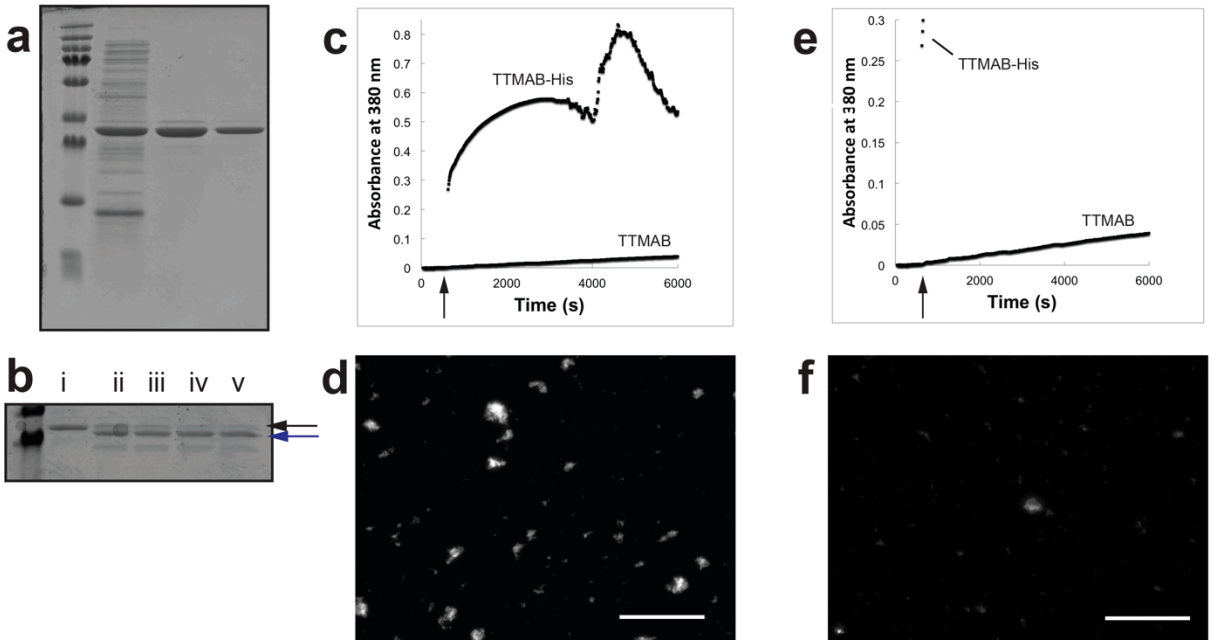

Figure 5.7. TTMAB characterization. (a) Ni-NTA purification of TTMAB-His. **(b)** AcTEV cleavage of TTMAB-His. (i) Intact TTMAB-His (top arrow) and cleaved TTMAB (bottom arrows) are shown after (i) 0 min, (ii) 30min, (iii) 1h, (iv) 2h, and (v) 3h incubation with AcTEV at 30°C. **(c,e)** Effect of the His-tag on assembly. TTMAB-His and TTMAB (3.4μM) were incubated with 5mM $CaCl_2$ at 23°C and scattering was monitored at 380 nm. Panel (e) is a zoomed in version of early time points (c). Fluorescence microscopy images of Nile Red-stained TTMAB-His **(d)** and TTMAB **(f)** prepared as in (c).

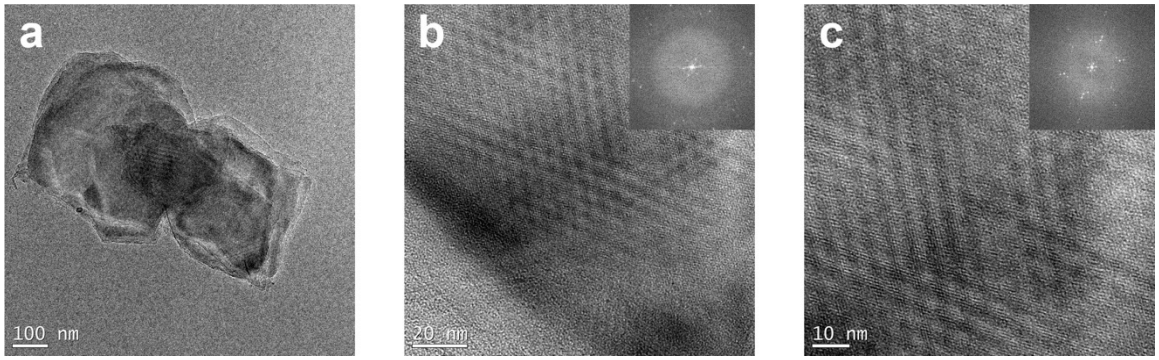

Figure 5.8. TEM images of calcium-treated TTMAB. (a) Low resolution TEM image of uranyl acetate-stained TTMAB structures obtained after incubating 1.7 μM TTMAB with 1 mM $CaCl_2$ at 4^{o}C for 18hrs. **(b,c)** Higher resolution images and the FFT images (inserts) reveal a hexagonal lattice with 7.06 nm unit cell.

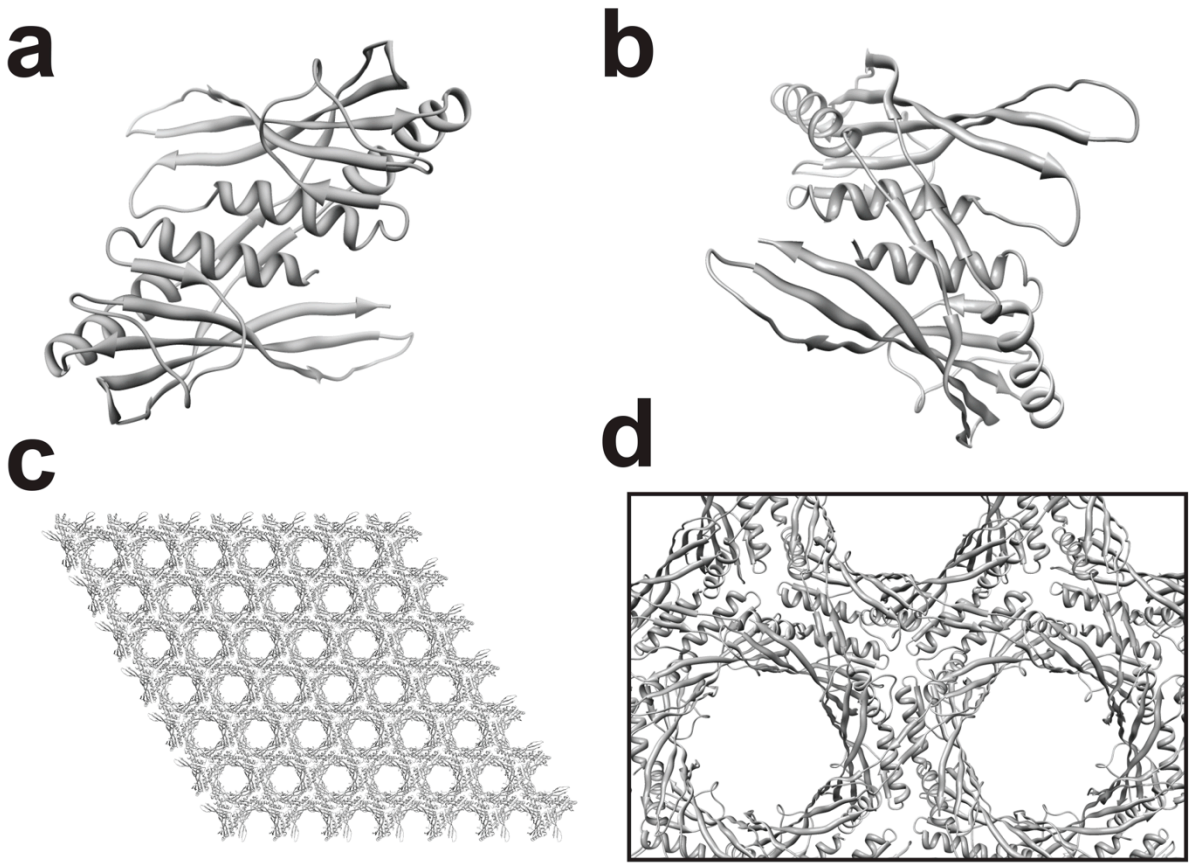

Figure 5.9. Circular permutation model. (a) Ribbon structure of circular permutation model, TTMCP, in the "front" compared to the "back" **(b)**. **(c)** is the 2D array that should be formed from TTMCP. **(d)** A close up of the packing of the 2D array.

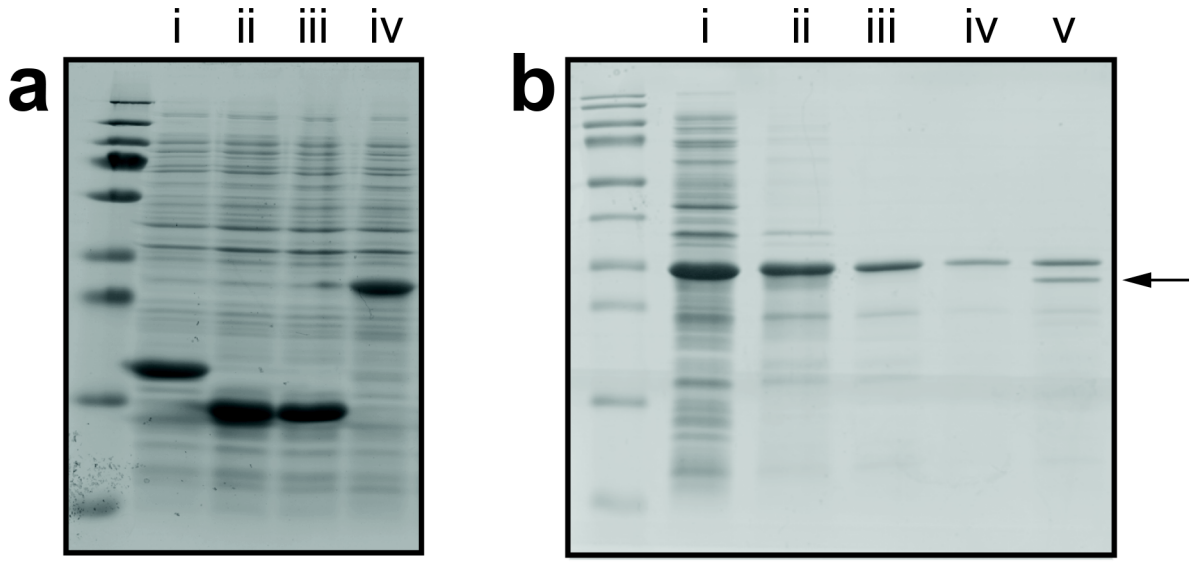

Figure 5.10. Purification scheme: His-TTM. (a) Whole cell fractions of His-WT (i), His-T (ii), His-TM (iii), and His-TT (iv) 3hr post induction at 37°C. **(b)** Purification train for His-TTM. Whole cell fraction 3hr post induction (i); Cleaned inclusion bodies (ii); Purified His-TTM (iii,iv) and thrombin cleavage of His-TTM (v). The arrow shows where the migration position of the cleaved product. Only about a third of the protein was cleaved.

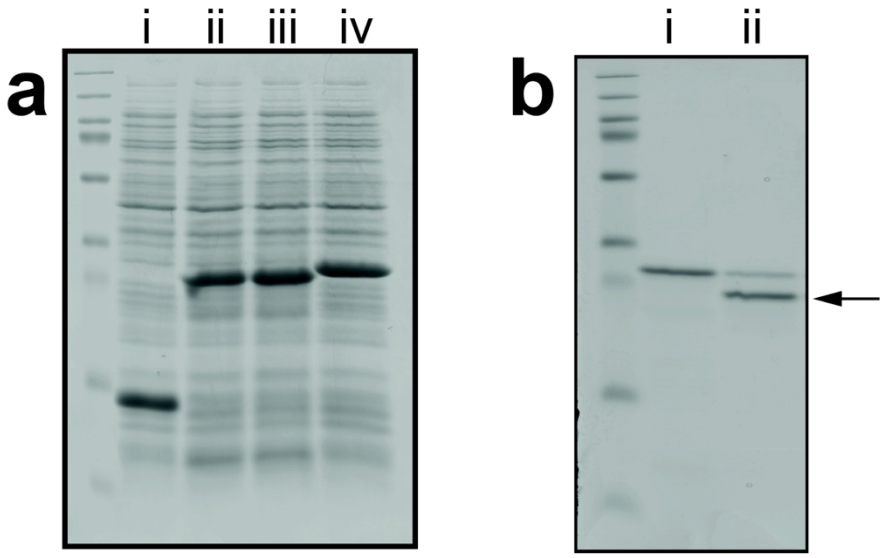

Figure 5.11. Purification scheme: HisTEV-TTM. (a) Whole cell fraction of HisTEV-TM (i), HisTEV-TTM (ii), HisTEV-TTM4 (iii), and HisTEV-TTMAB (iv) 3h post induction at 37°C. **(b)** Purified HisTEV-TTM (i). (ii) This protein is cleaved by AcTEV and the resulting cleavage gives S-TTM. Over 75% of HisTEV-TTM was cleaved. The arrow shows where S-TTM migrates. TTM (not shown) migrates to the same position as S-TTM.

Chapter 6 FUNCTIONALIZATION OF THE TTM ARRAY

6.1 INTRODUCTION

Self-assembly is of great importance to nanotechnology because it is one the most effective strategies to fabricate highly ordered nanostructures [1-3]. It is also a hallmark of many biological systems [224]. For example, Surface layer (S-layer) proteins, self-assemble into crystalline two-dimensional arrays of oblique, square and hexagonal symmetries with pore sizes ranging from 2 to 8nm. S-layers have been used for applications ranging from ultrafiltration membranes, to attachment layers for biosensor design, to templates for the synthesis of ordered metal and metal oxide nanoarrays [18]. There are a finite number of S-layer arrays and only one has been crystalized so far [76] and a few high resolution EM structures are available [18, 77, 78], which means tailoring these arrays is done by a lot of trial and error [153, 154].

DNA has also proven very useful for making functionalized, tailored nano-arrays [5-7]. The drawback to these arrays is their characteristic length scale is typically inferior to 1μm, which limits practical applications. There has also been significant progress in the development of programmed supramolecular structures based on peptoids (polymers whose constituent monomers resemble amino acids but have side chains appended to the amide nitrogen rather than to the α carbon) [12-14]. However, there is currently only one peptoid chemistry available.

Although the development of protein blocks capable of self-assembling into well-defined architectures lags behind DNA nanotechnology, progress in the computational design of protein-protein interfaces has accelerated over the past few years [28-31]. The functionalization of these structures has yet to come to fruition.

Here, we describe Avidin functionalization of the computationally designed protein array of Chapter 4 by fusing a biotin acceptor peptide (*aka* AviTag) to TTM monomers. In the

resulting fusion protein (TTM-Avi) the 15 amino acid long GLNDIFEAQKIEWHE AviTag, whose lysine residue can be biotinylated *in vivo* or *in vitro* by the *E. coli* enzyme biotin ligase (BirA), is linked to the C-terminus of TTM through a flexible GSGGS linker. We show that TTM-Avi, can be biotinylated with purified BirA [225], demonstrate subsequent 2D array formation upon calcium addition, and show that Avidin can be conjugated to the array to an extent that is modulated by varying the ratio of TTM-Avi to wild-type TTM.

This work lays the foundation for the production of TTM arrays that can be functionalized with nanoparticles or support inorganic biomineralization through the fusion of solid binding peptides. Towards these goals, we describe here the construction of TTM and TTM4 variants modified with C-terminal sequences binding to silica/sp^3-hybridized carbon (Car9) [226, 227], sp^2-hybridized carbon (Car15) [227], and silver (Ag4) [228].

6.2 RESULTS AND DISCUSSION

To endow TTM with a functionalization handle, an AviTag was fused to the C-terminus of the monomer through standard microbiology techniques. A flexible GSGGGS linker was selected to separate the tag from the plane of the array and avoid interfering with the assembly process. The resulting fusion protein was called TTM-Avi (Figure 6.1a,b). An Avi-tagged variant of TTM4 (Chapter 5) was also constructed and named TTM4-Avi, as were His-tagged versions of the fusions (His-TTM-Avi and His-TTM4-Avi). All proteins expressed at high levels in *E. coli* (Figure 2a) but only TTM4 and TTM4-Avi were purified for this work (Figure 2b).

To biotinylate Avi tagged proteins, a His-tagged version of biotin ligase BirA was overexpressed in *E. coli* and purified to near-homogeneity by Ni-NTA chromatography (Figure 6.2c). Next TTM4-Avi and BirA were mixed in the presence of ATP and biotin and the ligase

was removed with Ni-NTA magnetic beads (Figure 6.2d). The biotinylated TTM4-Avi was dialyzed to remove all small molecules and the resulting protein was used for array assembly.

We conducted three trials in which biotinylated TTM4-Avi was added at a 0%, 5% and 10% molar ratio to TTM4, polymerization was triggered by addition of 1mM $CaCl_2$, and assembly was allowed to proceed for one week at 4°C. These lattices were incubated with FITC-labeled Avidin-FITC and imaged by fluorescence microscopy using channels specific for FITC and Nile Red. As expected from the results of Chapter 4, Figure 4.4, all arrays were efficiently labeled with Nile Red (Figure 6.3a-c) Also as expected, no fluorescence was detected in the FITC channel for samples lacking biotinylated TTM4-Avi (Figure 6.3d) while there was faint fluorescence in the 5% array and significant fluorescence in the 10% TTM-Avi array. Next, Nile Red images were colored red and FITC images colored blue before being superimposed. Figure 6.3g,h,i confirmed co-localization of fluorescence signals and a progressive increase in purple with the amount of TTM4-Avi added. We conclude that biotinylated TTM4-Avi incorporates in a dose-dependent manner in growing TTM4 arrays and that the biotin is available for binding by Avidin.

6.3 CONCLUSION

Here we have shown that it is possible to functionalize TTM two-dimensional arrays with other proteins by making use of a TTM4-Avi monomer. The amount of avidin bound to the array surface is seems to be related to the amount of biotinylated TTM4-Avi incorporated in the lattice. This is but the first step in functionalization experiments that may also provide insights into array formation. An example, polymerization could be initiated with TTM alone and growing crystals supplemented with an excess of biotinylated TTM-Avi at a later time. If growth is isotropic,

incubation with FITC-labeled Avidin should yield crystals whose periphery alone fluoresces in the FITC channel.

The creation Ag4, Car9 and Car15 derivative of TTM should also prove valuable for practical applications. Because the AviTag is accessible, we expect that the same will be the case for solid binding peptides. TTM-Ag4 lattices should be useful to organize silver or gold nanoparticles that are easily imaged by SEM and to direct noble metal mineralization from precursor ions. The carbon binding Car9 and Car15 tags may help with TEM studies since it proved difficult to get large crystals to lay flat and at high density on carbon-coated TEM grids. This can also be investigated with Car9 and Nile Red staining since Car9 binds to microscope slides [226].This is an exciting time for the functionalization of TTM two-dimensional arrays.

6.4 MATERIALS AND METHOD

6.4.1 *DNA manipulation*

6.4.1.1 Construction of pSTM4215-TTM-X-22 and pSTM4215-TTM4-X-22

Plasmids pSTM4215-TTM-22 and pSTM4215-TTM4-22 specify the genes encoding TTM and TTM4 flanked by *Nde*I and *Bam*HI restriction sites. The two stop codons located upstream of the *Bam*HI site were removed by PCR-amplified using primer pair 5’-GAAGGAGATATACATATGG-3’ (pET22STMforward) and 5'-TATATAGGATCCCGCGATAAACAGGG -3' (TTMXreverse). Amplified fragments were digested with *Nde*I and *Bam*HI and cloned into the same sites of pET-22b(+) (Novagen). The resulting plasmids (pSTM4215-TTMX-22 and pSTM4215TM4X-22) encode TTM and TTM4 variants with a 26 residues-long C-terminal extension of sequence GSEFELRRQACGRTRAPPPPPLRSGC. In these constructs, the *Bam*HI restriction site (which specifies the GS dipeptide) is maintained while the two stop codons are removed.

6.4.1.2 Construction of pSTM4215-TTM-Avi-22 and pSTM4215-TTM4-Avi-22

Primers 5'-TATATAGGATCCGGTGGCGGTGGCGGTCTGAACGACA TCTTCGAGGCTCAGAAAATCGAATGGCACGAATAAGAATTCTATATA-3' and 5'-TATATAGAATTCTTATTCGTGCCATTCGATTTTCTGAGCCTCGAAGATGTC GTTCAGACCGCCACCGCCACCGGATCCTATATA-3' were annealed to create a DNA cassette encoding a GGGS linker, the Avi-tag sequence (GLNDIFEAQKIEWHE) and a stop codon flanked by *Bam*HI and *Eco*RI sites. The annealing reaction consisted of 30μL of forward and reverse primer (at 10 μM concentration) 1x PCR Rxn Buffer –$MgCl_2$ (Invitrogen), and 10mM $MgCl_2$ (Invitrogen). The mixture was boiled for 10 min and allowed to cool to room temperature. The cassette was digested with *Bam*HI and *Eco*RI and the fragment cloned into the same sites of pSTM4215-TTMX-22 and pSTM4215TM4X-22 to yield pSTM4215-TTM-Avi-22 and pSTM4215TM4-Avi-22. These plasmids encode TTM-Avi and TTM4-Avi followed by a GSGGGS linker and the AviTAG.

6.4.1.3 Construction of pSTM4215-His-TTM-Avi-28 and pSTM4215-His-TTM4-Avi-28

Plasmid pET-28a(+), pSTM4215-TTM-Avi-22 and pSTM4215TM4-Avi-22 were digested with *Nde*I and *Eco*RI, run onto a agarose gel by electrophoresis and the ~880 bp inserts ligated within the ~5.3 kbp fragments backbone. The resulting plasmids, pSTM4215-His-TTM-Avi-28 and pSTM4215-His-TTM4-Avi-28, encode His-tagged versions of TTM-Avi and TTM4-AVi.

6.4.1.4 Construction *of pSTM4215-TTM-Car9-22, pSTM4215-TTM-Car15-22, pSTM4215-TTM-Ag4-22, pSTM4215-TTM4-Car9-22, pSTM4215-TTM4-Car15-22, and pSTM4215-TTM4-Ag4-22*

These plasmids were constructed as described for pSTM4215-TTM-Avi-22 and pSTM4215-TTM4-Avi-22 except that the DNA cassettes specified Car9 [226, 227], Car15 [227]

and Ag4 [228] and that 2 stop codons were used instead of one in pSTM4215-TTM-Avi-22. The following primers pairs (Invitrogen) were used: for Car9 5'-TATATAGGATCCGGTGGTGGTGGTGACAGTGCTCGCGGGTTTAAAAAGCCTGGGAAGCGGTAGTGAGAATTCTATATA-3' and 5'- TATATAGAATTCTCACTACCGCTTCCCAGGCTTTTTAAACCCGCGAGCACTGTCACCACCACCACCGGATCCTATATA-3'; for Car15 5'- TATATAGGATCCGGTGGTGGTGGTCGGACTTACCTTCCGCTCCCATGGATGGCCGCGCTGTAGTGAGAATTCTATATA-3' and 5'-TATATAGAATTCTCACTACAGCGCGGCCATCCATGGGAGCGGAAGGTAAGTCCGACCACCACCACCGGATCCTATATA-3'; for Ag4 5'- TATATAGGATCCGGTGGTGGTGGTAACCCGAGCAGCCTGTTTCGCTATCTGCCGAGCGATTAGTGAGAATTCTATATA-3' and 5'- TATATAGAATTCTCACTAATCGCTCGGCAGATAGCGAAACAGGCTGCTCGGGTTACCACCACCACCGGATCCTATATA-3'. Plasmids were named pSTM4215-TTM-Car9-22, pSTM4215-TTM-Car15-22, pSTM4215-TTM-Ag4-22, pSTM4215-TTM4-Car9-22, pSTM4215-TTM4-Car15-22, and pSTM4215-TTM4-Ag4-22.

6.4.1.5 pET21a-BirA

Plasmid pET21a-BirA was purchased through Addgene (Plasmid 20857) and encode BirA followed by a His tag [225].

6.4.2 *Expression and purification*

TTM4 and TTM4-Avi were expressed and purified as described in Chapter 4.

BirA purification was adapted from Howarth and Ting [225]. Briefly, overnight inocula (10 mL) of BL21(DE3) cells harboring pET21a-BirA were used to inoculate 500 mL of LB supplemented with 100 μg/mL carbenicillin. Cultures were grown at 37 °C to $A_{600} \approx 0.5$,

transferred to 30 °C for 10 min and protein synthesis was induced by addition of 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG). After 3 h of incubation, cells were harvested by centrifugation at 3,000*g* for 20 min and the paste was stored at -20°C overnight. Cells were resuspended in 35mL of 20mM potassium phosphate, pH 8.3, 500mM NaCl , 0.1% NaN3, 0.1mM PMSF and 1mM DTT and disrupted by 6 cycles of sonication at 3 minute intervals on a Branson Sonifier 450 operated at 30% duty on ice. Following centrifugation at 12,000x*g* for 20 min to remove insoluble material, clarified extracts were supplemented with 10mM imidazole and loaded onto a Ni-NTA agarose column developed in 20mM potassium phosphate, 500mM NaCl, 10mM imidazole at pH 8.3 at a flow rate of 1 mL/min. The column was subject to a 50mM imidazole wash prior to elution of BirA with 200mM imidazole. The protein was dialyzed 3 times overnight against 20mM potassium phosphate, pH of 8.3. Aggregates were removed by centrifugation (10,000g) and the supernatant's concentration was adjusted to 50 μM. The protein was aliquoted and stored at -80 °C.

6.4.3 *Biotinylation and removal of BirA*

Biotinylation reactions were conducted by mixing TTM4-Avi (11 μM), BirA (2 μM), ATP (10 mM) and Biotin (50 μM) in 20mM Phosphate Buffer, pH 8.3 (final volume of 500 μL) and incubating the mixture for ~45min at 30°C [225, 229]. The solution was supplemented with 10mM imidazole and 25 μL of Ni-NTA magnetic beads (Thermo Scientific) were added. After 30min of mixing with gentle agitation at 23 ^{o}C, the beads where removed with a magnet. The supernatant was was dialyzed 3 times overnight against 20mM Tris with a pH of 8.3, aliquoted at 0.1–0.2mg/mL concentration and stored at -20°C.

6.4.4 *Characterization*

Samples were prepared as in Chapter 4, expect that different ratios of TTM4 and biotinylated TTM4-Avi were mixed at the indicated molar ratios. Array assembly was triggered by adding of 1 mM of $CaCl_2$ to 3.4 μM of protein in a final reaction volume of 50μL. After 7 days of incubation at 4^{o}C and temperature, 2 μL (80nM) of Avidin-FITC (Life Technologies) was added and samples were incubated at 23^{o}C for 15min with out shaking.. Next 250μL of 20mM Tris-HCl, pH 8.3, 500mM NaCl was added. The solution was subjected to low speed (100*g*) centrifugation and 250μL of supernatant was pipetted and discarded. The wash step was repeated and the remaining 50μL was stained with Nile Red (1μM final concentration; Sigma). 2D crystals were imaged on a Nikon Eclipse using 460-500 nm (FITC) or 510-560 nm (Nile Red) excitation.

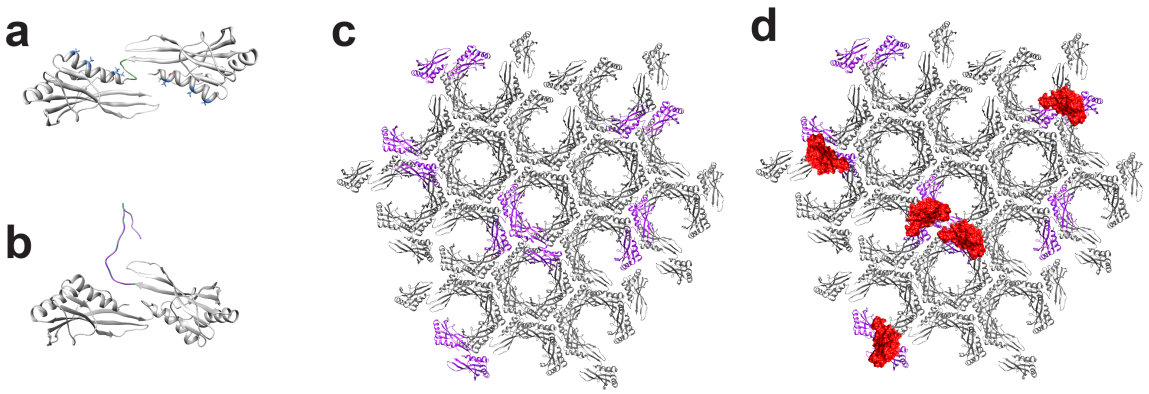

Figure 6.1. Decoration of 2D arrays ribbon structure of TTM. (**a**) and TTM-Avi **(b).** The Avi tag is schematically shown in blue. **(c)** Top view of possible TTM/TTM-Avi mixed arrays where TTM-Avi protomers are shown in purple and TTM protomers in grey. **(d)** Binding of Avidin, red, (PDB 2AVI) to some of the Avi-tags of the array shown in panel (**c**).

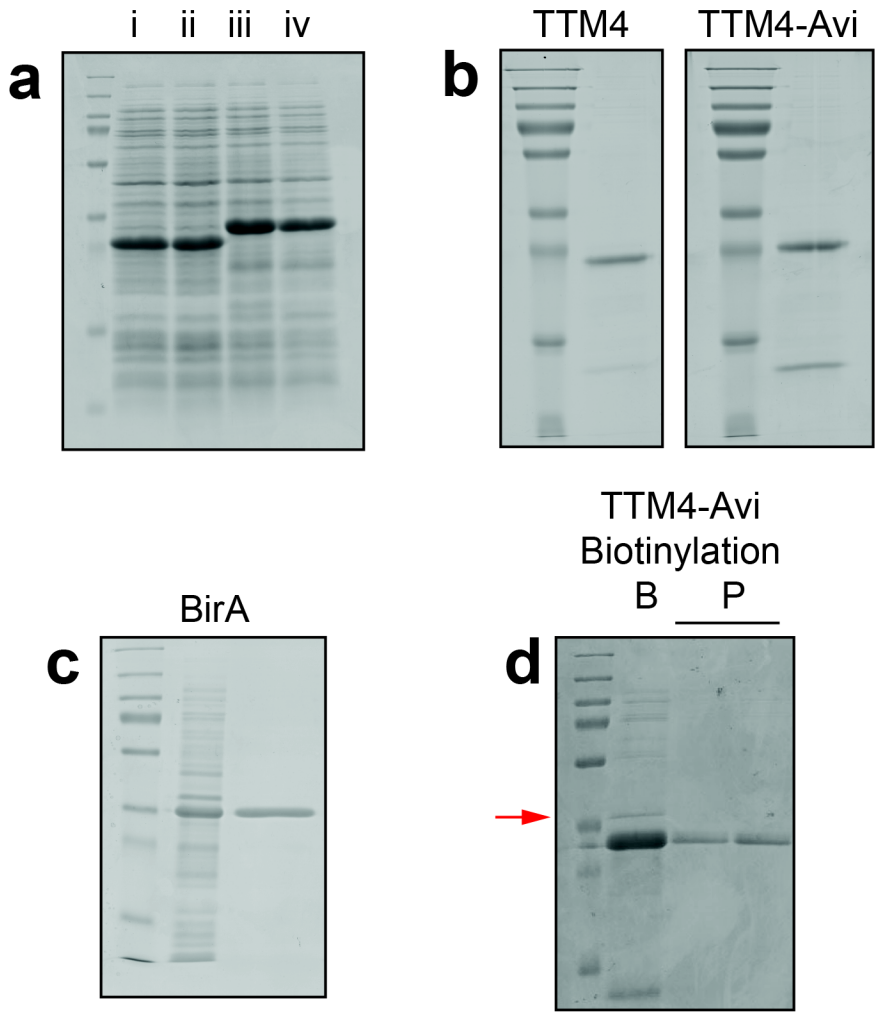

Figure 6.2. Biotinylation of TTM4-Avi. (a) Whole cell fractions showing high level expression of TTM-Avi (i), TTM4-Avi (ii), His-TTM-Avi (iii), and His-TTM4-Avi (iv) after 3h of post-induction growth at 37°C. **(b)** Purified TTM4 and TTM4-Avi. The difference in migration position is due to the presence of the 2.2kD Avi-tag. **(c)** Purified His-BirA from Ni-NTA purification. **(d)** Biotinylated TTM4-Avi (Lane B). The red arrow shows the position of His-BirA. The sample was mixed with Ni-NTA Magnetic Beads to remove His-BirA. The resulting purified and biotinylated TTM4-Avi are in lanes P.

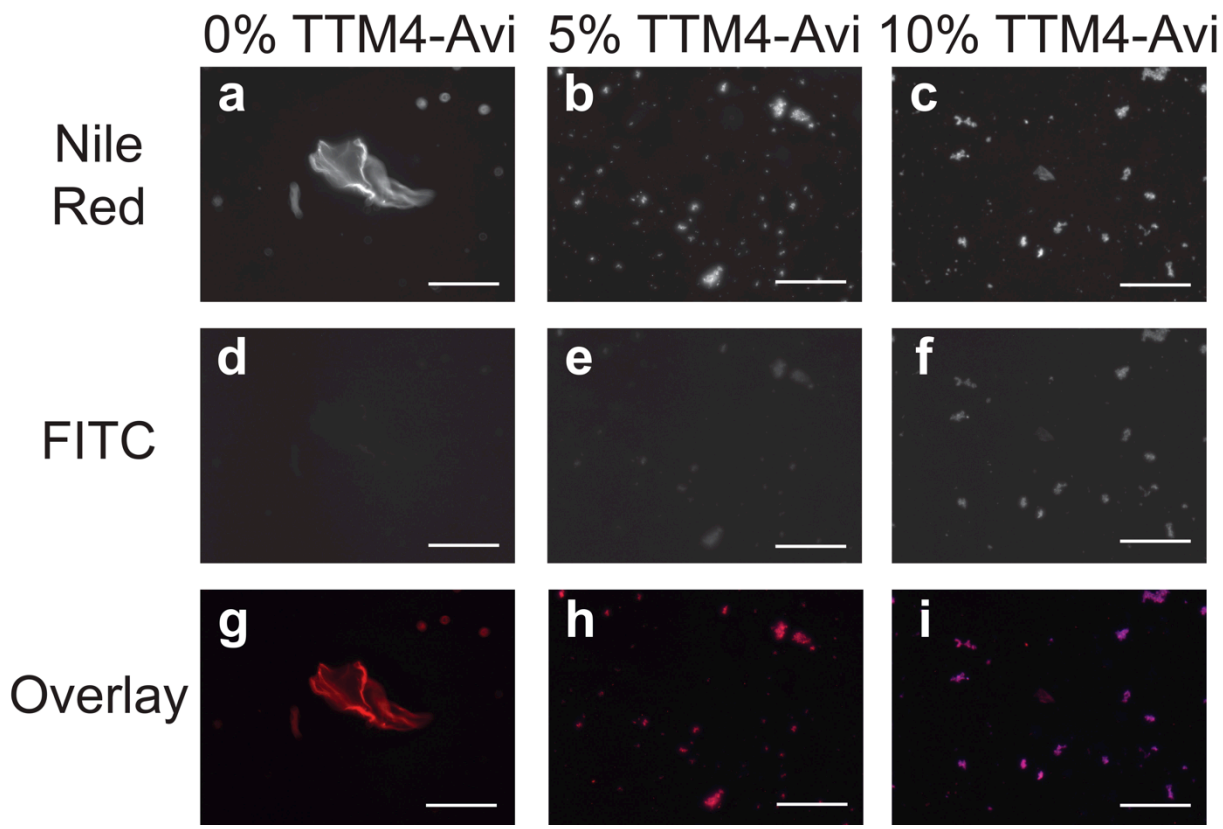

Figure 6.3. Formation of pure TTM4 and mixed TTM4-Avi/TTM4 arrays. Biotinylated TTM4-Avi was mixed with TTM4 at 0% or 5% or 10% TTM4-Avi (molar basis) to form mixed arrays. The arrays were incubated with 80nM of FITC-labeleled Avidin which binds to biotin. Structures were washed and subjected to Nile Red staining. The top row **(**panels **a,b,c)** show fluoresce images of Nile Red stained arrays prepared with TTM4-Avi and 0%, 5%, or 10% TTM4. The middle row **(d,e,f)** show fluorescence in the FITC channel. The bottom row **(g,h,i)** is an overlay with Nile Red images colored red and FITC images colored blue. Overlap is denoted by purple color. The scale bars are 100μm.

Chapter 7 A LITTLE BIT BRIGHTER: USING GOLD NANOPARTICLES TO ENHANCE QUANTUM DOT EMISSION

7.1 INTRODUCTION

Quantum dots (QDs) have been growing in popularity over organic dyes and fluorescent proteins, namely for their use in bioimaging [230-232]. This popularity is due to the desirable optical properties such as broad absorption spectra, large Stokes shift, narrow emission peaks, high photostability, and resistance to photobleaching [233]. Normally QDs are synthesized in harsh nonbiologically friendly conditions but the Baneyx lab has developed a robust biofabrication for zinc sulfide QDs [234-236]. This process relies on a designer protein that contains a ZnS-binding peptide, CT43, that caps the ZnS nanocrystals before they reach a critical size and precipitate out of solution [234]. These QDs have been doped with manganese to change their color from blue to orange and increase emission intensity [236]. This increase in emission intensity is related to an increase in the quantum yield, which is problematically low in traditional quantum dots.

Another area of research gathering attention is fluorophore enhancement and quenching by plasmon resonance from Nobel metals, e.g. silver and gold [237-239]. If a fluorophore is placed ~4 nanometers away from a metal capable of generating surface plasmons then the surface plasmons can help increase the brightness of the fluorophore. Fluorescence is quenched, however, if the metal is closer than ~4 nanometers [237, 240]. This specific spacing requirement can be achieved through use of a protein scaffold in combination with, for example, an Ag4 peptide. The Ag4 peptide, which was designed to bind to sliver [228], has been placed in a protein scaffold where the Ag4 peptide could control protein orientation when binding silver nanoparticles [241] and the growth habit of silver particles using electrochemical deposition

[242]. Therefore the Ag4 peptide can immobilize silver nanoparticles in a protein scaffold. A dual binding protein would be an ideal spacer such that the requisite ~4nm scale is achieved and the plasmon resonance can increase fluorophore emissions.

This was accomplished by fusing the Ag4 peptide to the C-terminus of the designer protein, BB-TrxA::CT43. The dodecapeptide Ag4 , NPSSLFRYLPSD, was spaced by a SSSGGG linker to BB-TrxA::CT43. The length of the linker can be elongated or shortened to maximize the enhancement of the quantum dots. This new protein should hopefully have the ability to mineralize ZnS quantum dots, immobilize antibodies and immobilize gold nanoparticles for enhancement. This enhancement scheme can be seen in Figure 7.1. Here we report on the ability of the new fusion protein BB-TrxA::CT43-Ag4 ability to form QDs and the slight enhancement seen.

7.2 RESULTS AND DISCUSION

First BB-TrxA::CT43-Ag4 and the control protein BB-TrxA-Ag4 were created using standard molecular biology techniques. These proteins, along with two other control proteins, BB-TrxA::CT43 and MBP-Ag4, were expressed and purified as described in previous publications by the Baneyx laboratory [234, 241]. High purity was achieved with the two new proteins, BB-TrxA::CT43-Ag4 and BB-TrxA-Ag4 (Figure 7.2). Now single-pot biofabrication of ZnS QDs can be done.

Because ZnS QDs doped with Mn were found to have a higher quantum yield than just the ZnS QDs biofabricated with BB-Trx::CT43, we decide to synthesize Mn-doped ZnS QDs [236]. The doping percent was 7.5% Mn during synthesis of QD. All four proteins made QDs (Figure 7.3), but the emission peaks for BB-TrxA::CT43-Ag4, BB-TrxA-Ag4, and MBP-Ag4 were all well below that seen by the control BB-TrxA::CT43. In fact, the two new fusion proteins, BB-TrxA::CT43-Ag4 and BB-TrxA-Ag4, were about 50% the emission of BB-

TrxA:CT43. BB-TrxA::CT43-Ag4 and BB-TrxA-Ag4 have emission peaks reminiscent of undoped ZnS QDs [234]. A reason for this might be the general variability in the biofabrication process. This QD biofabrication process is robust but the more times an engineer makes the QD the less variability he has from batch to batch. While the marker of a successful quantum dot is subjective, the QDs described above reached an emission of 50 or more and were therefore considered a success.

The ability to make QDs from MBP-Ag4 was an unexpected result but the QDs resulting from MBP-Ag4 have about 25% more emission than those made from BB-TrxA-Ag4 and BB-TrxA:CT43. This emission, however, is still 40% less than the standard QDs made from, BB-TrxA::CT43. This seems to imply that the Ag4 peptide can mineralize QDs but more controls will need to be run to prove this.

The main point to these experiments was to see if enhancement could be achieved via a plasmon resonance. Previous, unpublished data shows that the Ag4 peptide has a twofold higher affinity to bind to gold than silver (Figure 7.4). Gold nanoparticles, therefore, may further enhance the QDs described above.

A gold colloid solution was made up at a concentration of ~1mM and should have a particle size of about 2nm. This solution was slowly added to the QDs and the emission peaks were observed to see if enhancement or quenching occurred. The results are shown in Figure 7.5, where the QDs emission signal was normalized. There was about a 20% enhancement of BB-TrxA::CT43-Ag4 with the addition up to 6μM of gold colloid. The QDs were enhanced by even with small amounts of gold nanoparticle (0.25μM). The BB-Trx-Ag4 seems unaffected by the addition of gold nanoparticles. This may be due to unbound BB-Trx-Ag4 in solution, which can bind to the gold nanoparticles therefore capping the particles so they have no effect on the

system. This is also true for BB-TrxA::CT43-Ag4. The emission of BB-TrxA::CT43 QDs remained unchanged until the gold nanoparticles reached a concentration of 11μM at which point the emission was reduced by 12%. This may be because the gold nanoparticles destabilized the QDs or because the nanoparticles adhere to the proteins capping the QDs in such a way that quenching was observed, perhaps under ~4nm. The MBP-Ag4 QDs were unstable in the presence of, or were quickly quenched by, gold nanoparticles. No matter the mechanism the emission of the MBP-Ag4 QDs dropped by 55%.

7.3 CONCLUSION

We demonstrated that BB-TrxA-Ag4, BB-TrxA::CT43, and MBP-Ag4 are all capable of biofabricating QDs but they are not as good as BB-TrxA::CT43. The emission data is 50% less for BB-TrxA::CT43-Ag4 than for BB-Trx::CT43. The most exciting thing is that BB-TrxA::CT43-Ag4 had ~20% enhancement with the addition of gold nanoparticles. All these experiments were a first trial case and should be redone with tighter control over the formation of QD and gold colloid. It is possible to see enhancement of these biofabricated QDs with the addition of gold nanoparticles whose immobilization is controlled by a gold binding peptide.

7.4 MATERIALS AND METHODS

7.4.1 *DNA manipulation*

For this project, site directed mutagenesis were applied twice to both the pT-BBTrxCT43 and pT-BBTrx plasmids to add an *Avr*II site at the end of the BBTrxCT43 or BBTrx genes and an *Afl*II site slightly down stream in the MCS. QuickChange Site Directed Mutagenesis Kit (Stratagene) was used according to the manufacturer's instructions. The primers for the *Afl*II mutagenesis were 5'- GCCATCTAGAGTCGACCTTAAGTAATCGTACAGGGTAG-3' and

5’- CTACCCTGTACGATTACTTAAGGTCGACTCTAGATGGC-3’. For the *Avr*II site, the site directed mutagenesis primers were 5’- CCTCGACGCTAACCTAGGCTAGCTGGCCATCTAG -3’ and 5’- CTAGATGGCCAGCTAGCCTAGGTTAGCGTCGAGG-3’. The *Avr*II site mutagenesis came after the *Afl*II site had been created. After each round of site directed site directed mutagenesis, plasmid verified by sequencing and stored in Top10 cell. After both rounds of site directed mutagenesis the following plasmids were created pT-BBTrxCT43-X and pT-BBTrx-X, which contained *Avr*II and *Afl*II restriction sites.

Primers 5'-TATATACCTAGGGAGCAGCAGCGGTGGCGGTAACCCGAGCAGCCTGTTTCGCTATCTGCCGAGCGATTGATGACTTAAGTATATAT-3' and 5'-ATATATACTTAAGTCATCAATCGCTCGGCAGATAGCGAAACAGGCTGCTCGGGTTACCGCCACCGCTGCTGCTCCCTAGGTATATA-3' were annealed to create a DNA cassette encoding a SSSGGG linker, the Ag4 sequence (NPSSLFRYLPSD) and two stop codons flanked by *Avr*II and *Afl*II sites. The annealing reaction consisted of 30mL of forward and reverse primer (at 10 μM concentration) 1x PCR Rxn Buffer –$MgCl_2$ (Invitrogen), and 10mM $MgCl_2$ (Invitrogen). The mixture was boiled for 10 minutes and allowed to cool to room temperature. The cassette was digested with *Avr*II and *Afl*II and the fragment cloned into the same sites of pT-BBTrxCT43-X and pT-BBTrx-X to yield pT-BBTrxCT43-Ag4 and pT-BBTrxCT43-Ag4. These plasmids encoded BB-TrxA::CT43 and BB-TrxA followed by a SSSGGG linker and the Ag4 peptide.

7.4.2 *Protein expression and purification*

Protein expression and purification for BB-TrxA:CT43-Ag4 and BB-TrxA-Ag4 used the same protein purification train of BB-TrxA:CT43 [234]. The expression and purification of MBP-Ag4 is found in Sengupta *et al.*'s work [241].

7.4.3 *QD biofabrication*

The synthesis protocol of Mn-doped ZnS QDs can be found in previous publications from the Baneyx laboratory [236]. The same molar amount (2μM) of protein of BB-TrxA:CT43-Ag4, BB-TrxA-Ag4, BB-TrxA:CT43, and MBP-Ag4 were used in the QD assembly.

7.4.4 *Gold colloid synthesis*

The gold colloid synthesis by reduction of chloroauric acid with THPC was accomplished using the procedures described in Shi *et al.*'s work [243]. This procedure produced gold nanoparticles of 1-2nm.

7.4.5 *Fluorescence spectrophotometry*

Photoluminescence emission spectra were recorded from 300 to 800nm using 1mL of sample on a Hitachi F4500 fluorescence spectrophotometer with excitation at 280nm and slit widths set at 2.5nm. The wavelength region corresponding to the second-order diffraction peak of the excitation light was omitted [236].

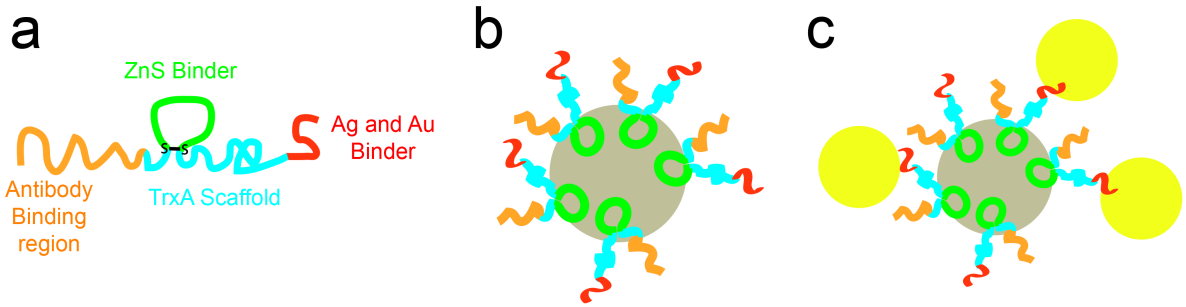

Figure 7.1. Brighter QD scheme. (a) Shows the regions tri-functional designer protein BB-Trx::CT43-Ag4. The ZnS binder is created by the dodecapeptide CT43. The Ag and Au binder is the dodecapeptide Ag4. The antibody-binding region is created by a repeat of the antibody-binding domain of *S. aureus* protein A (BB). **(b)** The ZnS binding region should cap ZnS nanoparticles to create ZnS quantum dots while the Au and Ag binding and antibody-binding domains is still freely available. **(c)** Gold nanoparticles should be able to be immobilized by the Ag4 peptide and the plasmon resonance of the nanoparticles will enhance the QD yield.

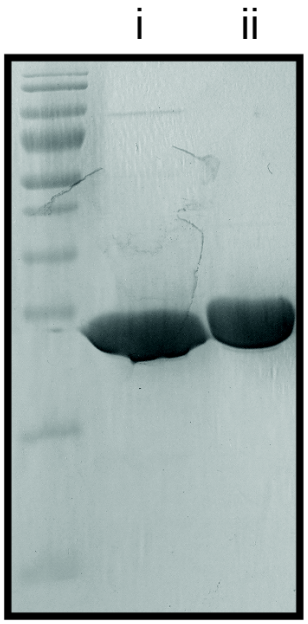

Figure 7.2. Purified BB-TrxA::CT43-Ag4 and BB-TrxA-Ag4. SDS minigels show that both BB-TrxA::CT43-Ag4 (i) and BB-TrxA-Ag4 (ii) are recovered with high purity following ion exchange chromatography.

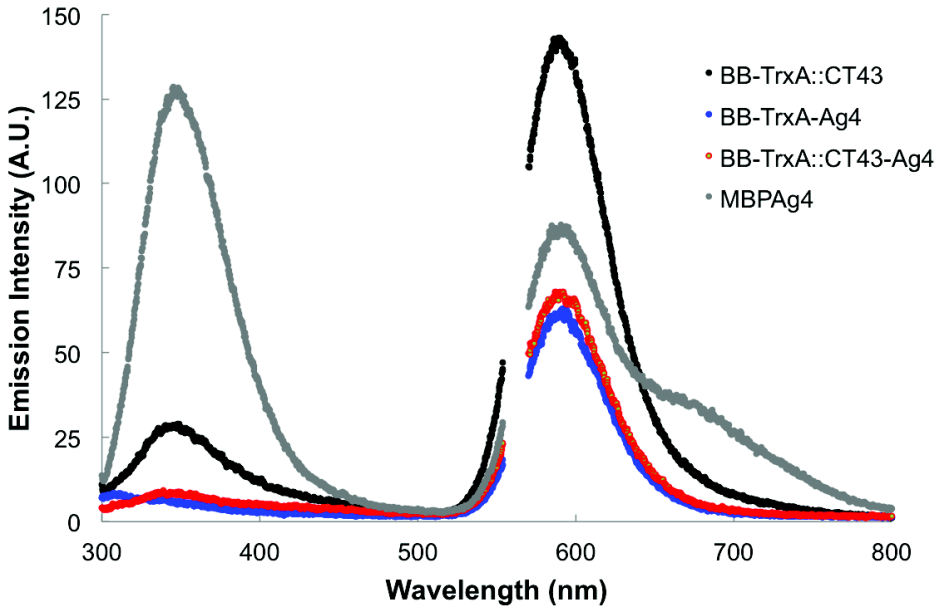

Figure 7.3 Mn-dope ZnS QDs biofabrication. Emission spectra of QDs mineralized in the presence of 7.5% Mn by BB-TrxA::CT43, BB-TrxA-Ag4, BB-TrxA::CT44-Ag4, or MBP-Ag4 when excited at 280nm.

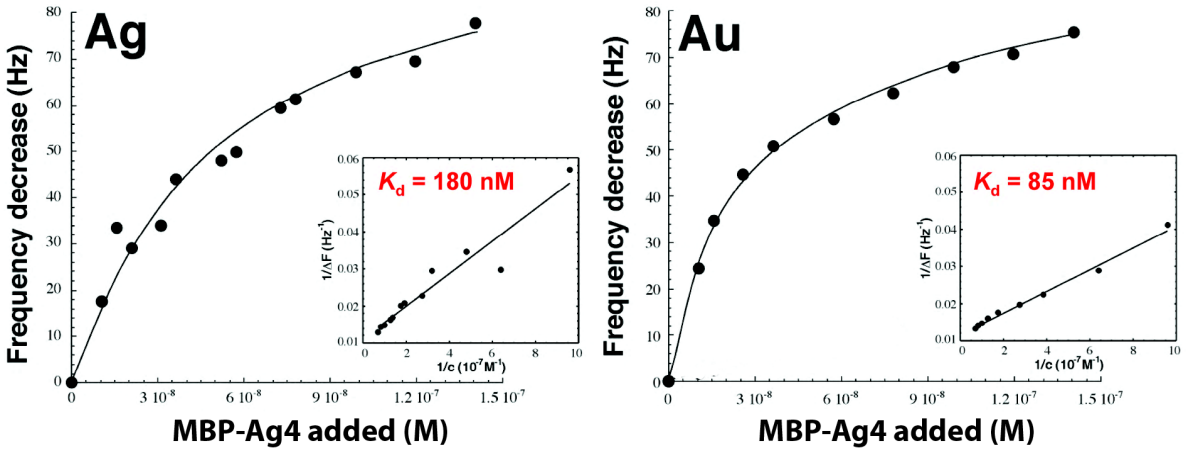

Figure 7.4 Ag verses Au QCM for MBP-Ag4. Quartz Crystal Microbalance experiments on silver (Ag) and gold (Au) to calculate the dissociation constant (K_d) for MBP-Ag4. These experiments reveal a twofold higher increase in MBP-Ag4 affinity for gold versus silver.

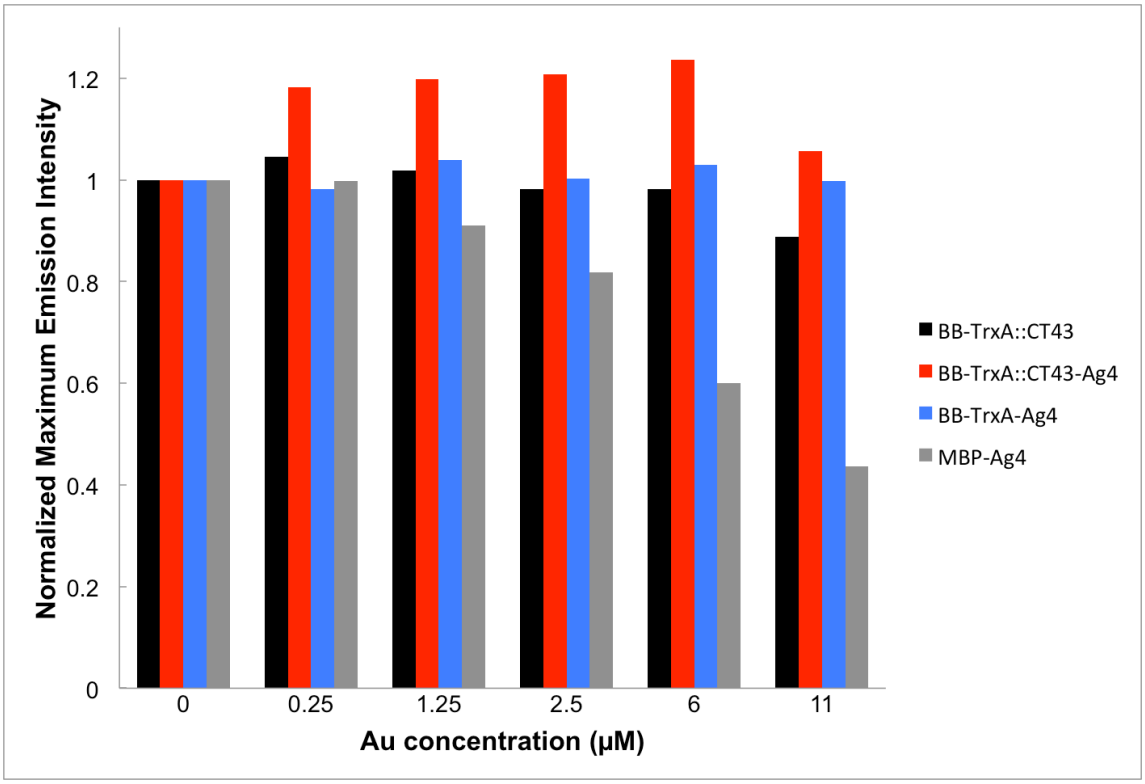

Figure 7.5 QD enhancement by immobilization of gold nanoparticles. Gold nanoparticles were slowly added to a solutions of QDs biofabricated by BB-TrxA::CT43, BB-TrxA-Ag4, BB-TrxA::CT43-Ag4 and MBP-Ag4. The emission intensity at 590nm was normalized Mn-dope QD for each protein without any gold. A 20% increase in emission can be seen for BB-TrxA::CT43-Ag4.

Chapter 8 A NEW GRADUATE LEVEL SEMINAR TO PREPARE STUDENTS FOR THE NEXT STEP IN THEIR CAREERS

Kelly Fleming, James Matthaei, Jim Pfaendtner*

8.1 ABSTRACT

Many new PhDs are not prepared for a competitive interview process in seeking a first job. The University of Washington (UW) designed the Distinguished Young Scholars Seminar (DYSS) to rectify this problem with three goals: simulate a visit associated with many interviews and a one hour seminar; make UW graduate students aware of where they fit into the job-seekers landscape, and expose UW graduate students to peer evaluation decision making processes that governs funding decisions.

8.2 INTRODUCTION

To would-be employers in academia and industry, a Ph.D. in engineering signifies not only the ability to conduct independent research, but also the skills to communicate the significance of research findings through written publications and oral presentations. Given the central role of a graduate student's advisor in the pursuit of a Ph.D., the acquisition of effective oral and written communication skills is often not addressed at the departmental level. For example, a brief survey of graduate curricula of the 50 top ranked chemical engineering Ph.D. programs in the US revealed only 8 classes that meet the criteria of professional development of oral/written communication. The importance of addressing these skills and helping graduate students prepare for the transition to professional life has been long noted in other fields, particularly the biological sciences[244-247]. Many Ph.D. graduates want to pursue a career academia, however, with the increasing number of doctoral graduates, a career in academia has become more competitive and difficult to obtain, therefore more professional development

should be provided too graduate students [248]. The National Institutes of Health (NIH) published a report in 2011 detailing areas of need in the development of well-rounded Ph.D. students and post-doctoral researchers [249, 250]. One of the key areas identified is professional development, but the report does not outline a path to reach this goal [249]. This leaves current graduate students in science and engineering with the "usual suspects": seminars, colloquia and professional societies.

Seminars and colloquia are a frequently used tool to help with professional development. Key benefits include exposing graduate students to disparate areas of research found outside their home departments and the opportunity to observe experienced professionals communicate the significance and impact of their research findings. At seminars students often have the opportunity to network and interact with external visitors, which can be beneficial irrespective of the specific research interests of the departmental guest. While the majority of speakers at graduate seminars in chemical engineering are faculty from other departments, visitors from industry or national labs are also common, providing the additional benefit of giving students exposure to career opportunities outside of academia. The aforementioned benefits are important components of a holistic graduate education in engineering, yet the nature of the seminar (i.e., a very experienced professional communicating to trainees) ensures that certain aspects will be missed (e.g., a trainee is unable to effectively self-assess their own abilities when comparing to someone with a decade or more of professional experience). Furthermore, before their first job interviews most graduate students gain presentation experience only in the form of conference presentations lasting 20 minutes or less or informal group meeting style presentations. Giving graduate students and post-doctoral researchers the opportunity to prepare for and provide a

formal presentation of their research is another area of improvement that needs to be addressed in the professional development of engineers.

In 2011, graduate students at the University of Washington (UW) in the Chemical Engineering department developed a new type of seminar to help both its own students and well-qualified peers bridge the gap between limited presentation opportunities and applying as a candidate for permanent academic positions or a career in industry or government. The Distinguished Young Scholars Seminar (DYSS) has three major goals 1) simulate the daylong departmental visit associated with many interviews capped by a one hour seminar talk given by rising stars in chemical engineering, 2) help our own graduate students at the UW become aware of where they fit into the chemical engineering landscape and 3) expose our own graduate students to the panel style peer evaluation decision making process that governs many funding decisions. All facets of the seminar are implemented and executed by the UW graduate students themselves to continue to broaden the idea of professional development. The remainder of this article is an overview of our process for developing and running this new seminar, including key changes we have made in the early stages as well as an assessment of the effectiveness of the new seminar.

8.3 MOTIVATION

8.3.1 *Organization and logistics*

To ensure that the seminar has the most value to the graduate students at the UW, it is entirely organized and run by graduate students from its inception in 2011. Each year an organizing team of 3-4 students, including one lead organizer, are selected by a faculty member overseeing the whole process. After the first year, the faculty served mostly in a limited advisory capacity due to the strong feeling of ownership by the students. Participating students are given credit towards

departmental requirements for teaching assistantship service. This credit encourages students to apply and helps offset any interruptions in the students' research progress. The organization team is responsible for advertisement, the selection process, and the execution of the seminar. They begin meeting in March to update the website (**http://depts.washington.edu/acesche/dysss**), and update and send out promotional material trying to reach as many eligible trainees as possible. In April a grading rubric is composed and a selection committee is organized. Fliers calling for applications are mailed to many chemical engineering programs around the country and to specific individuals that are presenting at the "Meet the Faculty Candidate" poster session at the American Institute of Chemical Engineers Annual Meeting. The professional networks of our faculty, graduate students, and alumni of the DYSS program are also leveraged in order to maximize exposure.

To amplify the quality of talks and depth of research in the seminars, the DYSS eligibility requirements were set at graduate students within 12 months of defense of their chemical engineering Ph.D. or postdoctoral scholars in chemical engineering. Emphasis is placed on early career post-docs (i.e., people in their first or the start of their 2nd post-doc). Anyone who has accepted a permanent full time position is not eligible to apply. These priorities establish a foundation of professional development for the DYSS (i.e., the goal is not to further reward those who have already achieved a prestigious position). The application consists of one letter of recommendation, a research abstract, and a two page C.V. Seven speakers and one alternate are chosen by the selection committee. In its inaugural year, 2011, there were 85 applicants, with 47 in 2012, and 60 in 2013.

A selection committee of about fifteen students is headed by the lead organizer. The organizing team strives for diversity on the selection committee across research groups,

background, and experience (although emphasis is placed on allowing more senior graduate students to serve on the panel). Following the organization of an NSF-style panel the goal is to have each application scored by three reviewers with each committee member needing to read fewer than 10 applications. Each application has one lead reviewer who will summarize the application if/when it is discussed by the full panel. The selection committee receives approximately 10-14 days to finish their reviews and enter them into an online system to assist the leader organizer who is responsible for reviewing and statistically organizing the scores.

At the selection panel the lead organizer facilitates the discussion and describes the panel process. The panel begins with a discussion of the goals for their selection process. For example, should research area be a factor?, Or should the quality of their application be the only criteria? Should there be an advantage for a strong presentation history? The sheer number of applicants means that only the top 20-30 applications can be actively deliberated at the panel and reviewed together as a group. However, prior to "retiring" the bottom tier of applicants any reviewer is given the opportunity to bring out an application and make the case that it should be reviewed or further discussed. The deliberation proceeds with a lead reviewer offering a brief summary followed by the entire committee reaching a consensus decision about lumping application into various categories (i.e., yes/maybe/no). At the end of the 1st round the "yes" applicants are ranked and the top 7 or 8 scholars are selected.

After the seven speakers are selected, the lead organizer plans the week-by-week schedule for the seminar. Each individual visit is focused primarily on maximizing contact between graduate students and the visitors. The day is organized like a typical seminar visit (lab tours and visits with research groups, meals and a seminar). There are a limited number of meetings with faculty, and the graduate students at UW have the opportunity to partake in all the

social and academic activities. One member of the organizing team is responsible for soliciting feedback and organizing the judging of all the talks (all seminar attendees can participate in the judging), culminating in the selection of a winner or best talk award for that year's DYSS.

8.3.2 *Improvements on the selection process*

Many improvements were made to the DYSS from its inaugural year. The most influential changes were made to the selection process and criteria. The first selection rubric was strongly based on faculty input following similar guidelines to those used in the review and selection of NSF Graduate Research Fellowship Program (GRFP) award winners. Three separate areas were judged: letters of recommendation, a C.V., and an abstract of the proposed research presentation. Originally, each piece was broken into subcategories. For example, the abstract was broken down into the following three categories in the inaugural year:

1. Topic of interest: does the audience care?
2. Introduction to topic: does the author describe why it's important?
3. Clarity: is the research conveyed in a way that is easy to understand?

Each of these subcategories was graded on a scale of excellent, very good, good, fair, and poor. A grade of "good" translated to "This person has a strong application, invite if possible." While this first rubric was a strong foundation, we found it was open to subjective interpretation, resulting in inconsistent scoring between reviewers. For example, we noted that a reviewer with a preference toward biological research might score research on solar cells with a ranking of "good," while a reviewer with a background in solar cells might score it as "excellent." The faculty advisors to the selection committee were able to use this and other similar observations to provide context to the UW grad students about challenges and competitiveness in the selection of fellowship winners and grant recipients.

In its inaugural year, the selection panel of graduate students felt that there was sufficient information contained within the application materials to base the selection of recipients only on quantitative metrics of quality. At multiple points in the first round of selection, the selection panel and faculty advisors discussed the priorities of the selection committee and clarified the mechanism by which applicants would be selected. Because of the strong sense of priority in quantitative metrics, the selection panel was somewhat unprepared with how to resolve an issue that faculty face constantly: how do you resolve or quantify quality to a fine detail with limited information? We do not have record of the exact numbers from the first year but the approximate breakdown was along the following lines: 1) the top three to four applicants were truly extraordinary in every possible metric and there was universal excitement about their selection, 2) the next six applications were excellent and there was near universal excitement about their selection, 3) approximately ten of the remaining discussed applicants were all very good and the committee was fine acknowledging their strong contribution and passing on the application. Seasoned faculty will be familiar to this concept – namely that there are always too many qualified people for available slots. On the other hand, many graduate students trained in engineering and quantitative sciences are unprepared for this harsh reality. The graduate students, with quantitative evaluation goals in mind, spent a long time passionately discussing the middle tier of candidates and eventually reached a near universal consensus about the slate of candidates to be invited for the summer. An unintended consequence of this process was that other departmental and institutional values (e.g., the value of the DYSS speaker in broadening participation in STEM) were neglected in favor of traditional metrics of quality (e.g., publications or strong letters).

As an example of how the DYSS strongly benefits the students in the home department,

we will briefly detail how the selection process went through a dramatic transformation in the second and third years. First, early in the process the student selection committee met with Dr. Joyce Yen, the Program/Research Manager of The University of Washington's ADVANCE Center for Institutional Change. The ADVANCE program was resourced to help the student organizers answer two questions: 1) how can we make the rubric less objective and easier to use for the entire panel? and 2) how can we ensure we simultaneously promote core institutional values like diversity while taking steps to eliminate any unconscious bias in the selection process? As we discuss below, question 1 and part of question 2 (the institutional values issue) were addressed through improvements to the rubric and panel system whereas limiting any unintended bias was addressed through intervention between our ADVANCE Center and the selection panel early in the selection process.

To improve the rubric and selection process, we have taken the following steps. Committee members now read a document on scoring applications, written by the Women in Science & Engineering Leadership Institute at the University of Wisconsin-Madison [250]. At the first meeting of the entire selection committee, Dr. Yen and student participants in the NSF-Funded UW College of Engineering Promoting Equity in Engineering Relationships (PEERS) program visit with the selection committee to discuss research on the impact of unconscious bias and issues likely to exacerbate such bias, e.g. time constraints, stress, fatigue, and unclear instructions [251-255]**.** Examples of gender, race, sexual orientation bias are shared with the selection committee. For example, UW ADVANCE and the PEERs students presented multiple studies showing both men and women consistently rate women candidates lower than those of men, even when the credentials (C.V.s, resumes, etc.) are identical [256, 257]. Diving deeper into this issue, the selection committee learned that men frequently are given argentic descriptors

(e.g. assertive, confident, aggressive, ambitious, independent) whereas women are given communal descriptors (e.g. helpful, nurturing, agreeable, interpersonal); communal descriptors have a negative relationship with getting hired or receiving awards [258].

The ADVANCE discussion concluded with the real-life example of the NIH Director's Pioneer Award. In the first year of the award the NIH fell into inherent bias traps and failed to select a single woman for the award. A 2005 paper in the Journal of Women's Health discusses biases in the NIH selection process: time pressure on the evaluator, absence of face-to-face discussion with applicants, ambiguity of performance criteria, emphasis on self-promotion, weight given to letters of recommendation and the need for finalist to make a formal, in-person presentation in which the individual, and not his or her science, was the focus of evaluation [259]. Once this bias trend and procedural shortcomings were brought to light, the evaluation process was modified. Over the next five years, an average of roughly 29% of the NIH Director's Pioneer Award went to women, a percentage that is consistent with the representation of women in the field.

For many (possibly most) of our graduate students, this discussion is their first exposures to these complex issues that engineering professionals will face throughout their careers. While we are not specifically tracking the spread of awareness or changing attitudes on gender, diversity, and bias, it is our strong hope and belief that amongst our graduate students ideas and attitudes are improving and that these experiences will make a lasting impression. For example, the faculty advisors to the DYSS have noted anecdotally that students not at all involved in the DYSS selection have come to them to speak with interest and passion about many of these issues.

There were two major outcomes from these interventions and investment of the NSF

ADVANCE and UW PEERS programs. First, for the second and third DYSS rounds our selection committee unanimously concluded that persons who are underrepresented in chemical engineering (women and underrepresented minorities) should be given preference in tiebreaker situations. This preference acknowledges the effect of implicit bias on the experiences of persons from underrepresented groups, that is they are more likely experience more barriers to entry, achievement, and recognition due to implicit bias. Thus, what may look like a tie is not. To underscore this strategy, each year the selection committee meets with ADVANCE/PEERS for a presentation on implicit bias in the weeks before the graded rubrics are due. Second the rubric is now broken into more sub-categories with specific guidelines (the 2013 rubric is shown in Figure 8.1). Points are allotted to each section, with examples of how many points are deserved for specific aspects that are included in the application. Additionally, panel members are given anonymous previous applications as examples of excellent applications and ones that need improvement. This gives a more broad view of the spectrum of applications, as each scorer cannot review every application received. We strongly believe that the selection is now more cohesive and consistent in selecting the best candidates that excel both in quantitative metrics and also more broadly represent our department and institutional values.

8.3.3 *Speaker profiles: where have they gone since being selected?*

As of January 2014, of the 2011 cohort of speakers for the Distinguished Young Scholars Seminar, 63% (5/8) have faculty positions. Because the speakers consisted of a mixture of postdoctoral scholars and graduate students, they were in different stages in their careers in terms of finding a permanent faculty position. Some of the graduate students have accepted postdoc positions and will go on to become faculty. 75% of the seminar speakers who were postdoctoral scholars at the time of the seminar are currently professors, and 25% are still at the same post-

doc position. 33% of the DYSS speakers who were graduate students when they gave seminar are current professors, 33% are full time researchers, and 33% are post-doctoral researchers.

As of January 2014, of the 2012 cohort of speakers, 43% (3/7) have faculty positions, with one of them being featured in "Forbes 30 Under 30" [260]. Based on entrants at the AIChE "Meet the Faculty Candidates Session", we note that many of the 2013 participants are currently looking for academic jobs. We believe the success of the DYSS alumni shows the impact a unique professional development opportunity can have on an early career chemical engineer, and also that the DYSS organizers have developed an effective series of activities for selecting top scholars in the field.

8.3.4 *Seminar assessments*

To assess the potential impacts of the new seminar, different surveys were given to three populations of students and professionals involved in the seminar series.

1. Graduate students who attended the seminar
2. Graduate students who served on the selection committee
3. The DYSS speakers from 2011

The speakers gave short answer responses to a series of questions about their visits and whether or not they thought the seminar helped them prepare for their future in academia. All seven of the speakers responded indicating that they thought the seminar was a valuable experience to their faculty interview process. Similarly, graduate students who served on the committee were given a questionnaire with short answer responses. The survey indicated that the committee members thought it was a valuable experience to their graduate career and professional development, and all of them agreed that the seminar should be continued the next year.

As a whole, the UW graduate students who attended the seminar found the overall effectiveness of DYSS to be extraordinarily positive and impactful, with over 90% of the 2013 participants registering “agree” or “strongly agree” to all questions. Students who attended also stated that it has helped them prepare for future career steps. 100% of students in attendance indicated that they would like to continue the DYSS in future years. Figure 8.2-8.4 show selected results from they surveys given to the UW graduate students who attended the seminar from each of the years that the DYSS has taken place.

8.4 CONCLUSIONS: BENEFITS TO UW AND FUTURE IMPROVEMNTS

The University of Washington strives to provide the best opportunities to prepare its students for careers after graduation. In developing this seminar, we hope to provide a competitive advantage for our students when applying for a position in academics or research.

Through a competitive application and selection process, UW students and post-docs attending DYSS see top researchers from their peer group and get exposure to cutting edge research. Additionally they are given examples of effective presentation methods from competitive applicants before they apply for prestigious positions themselves. The applications received annually show a range of students, most of whom have high impact papers. By serving on the panel, members of the selection committee see impactful research being done outside of the UW. This helps to extend knowledge of the chemical engineering field, as well as open ideas for their own research and skills in the immediate future. These benefits are extended to all of the graduate students in the chemical engineering department as well as any researchers from other departments that choose to attend.

DYSS organizers are provided extended professional development opportunities. Students on the selection committee see firsthand the scoring process and its challenges for

similar application processes, like fellowships and job positions. Inside experience will help provide them with knowledge about what increases their own applications' competitiveness.

DYSS teaching assistants who host the seminar speakers during their visits give them laboratory tours and short research talks. This provides an opportunity for the student to talk to the speaker and get personal advice from them on how to become a leader in the chemical engineering field. Host students also get the opportunity to recognize and aide in the professional development for emerging leaders in the chemical engineering field.

Applicants who are chosen as speakers also greatly benefit from the seminar. They gain national recognition as well as invaluable experience to prepare them for faculty interviews. During their visit, they are given critical feedback from the UW faculty about their seminar and the opportunity to ask detailed and candid questions about the job interview process. They are given an advantage over their competitors when interviewing for any position.

Improvements have been, and will continue to be made, in order to ensure the seminar reflects the interests of the UW Department of Chemical Engineering. The scoring process to select the seven speakers continues to evolve to meet the goals of the students. This includes evolving how the research topics are selected, and how implicit bias is taken into consideration when scoring is taking place. In addition to improving the scoring process, advertisement for applications, and advertisement for attendance across campus continues to be improved. In the end, UW created and continues to produce a seminar series that increases the professional development of its students along with top researchers across the US.

8.5 ACKNOWLEDGMENTS

UW Department of Engineering for funding, UW ADVANCE and Dr. Joyce Yen, Jeff Richards, Chris Wolcott, L.D. Pozzo, IkechuKwu Nwaneshiudu, Brandon Coyle, Qing Shao, Kayla

Vanous, and Tyler House for helping organize and execute DYSS.

Category	5 - Excellent	3 - Good	1 - Needs Work
Communication One of the most important skills for any successful researcher is the ability to communicate. Rate the application based on how well it shows the applicant is able to communicate. Look for presentations, talks, teaching experiences in the CV and the quality of the submitted abstract. Does the letter of recommendation mention anything special about the applicants presentation skills? Try to rate the applicant in this section based on how good of a presentation you would expect them to give if they were invited.	**CV:** - Has many presentations at different national conferences - Has invited talks - Has teaching awards **Letter of Rec:** - Recommender praises applicant for clarity in presentations, giving specific examples - Recommender praises applicant as a teacher, giving examples of applicants mentorship **Abstract:** - Abstract is excellently written and leaves you wanting to know more - Abstract is clear and easy to follow even for someone not in the field	**CV:** - Has presented at national conferences - Has teaching experience **Letter of Rec:** - Recommender mentions applicant's presentation ability - Recommender mentions applicant's teaching ability **Abstract:** - Abstract is well written and contains no significant spelling or grammar errors - Abstract is easy to follow, but a nonexpert may need to read it more than once to understand it	**CV:** - Has presented at atleast one conferences - Mentions mentoring undergraduate researchers **Letter of Rec:** - Recommender mentions applicant's presentation experiences - Recommender mentions student as a teacher or mentor in lab **Abstract:** - Abstract is well written but contains some errors - Abstract is difficult to understand for someone not in the field.
Scholarship Scholastic achievements are a measure of the impact the applicant has had so far in their career. Primarily this comes about through published papers. How many first author papers has the applicant written? Are they in high impact journals? Have they been cited? Is their research topic significant and of broad interest? Do they have any patents? Try to rate the applicant in this section based on the quality of their research.	**CV:** - Has a papers published in a variety of high impact journals - Papers have been cited numerous times **Letter of Rec:** - Recommender praises applicant as a researcher and highlights their ability to come up with novel ideas **Abstract:** - Abstract communicates the importance of the research which has been done - Abstract show cases its broad impacts on the scientific community	**CV:** - Has atleast one paper published in a high impact journal - Has numerous papers in a variety of journals - Papers have been cited a few times **Letter of Rec:** - Recommender highlights the applicant's research talents in solving problems **Abstract:** - Abstract cites appropriate literature to show it's place in the bigger research picture - Abstract shows the impact it has had and mentions future research directions	**CV:** - Has few first author papers - Majority of papers are published in a single journal **Letter of Rec:** - Recommender praises applicant's technical abilities **Abstract:** - Abstract highlights the important conclusions from the work done - Abstract shows its impact within its own specialty
Achievement What impact has the applicant had so far? How have they been recognized for their success? Do they show a commitment to the broader scientific and chemical engineering community? Look for awards, outreach, and service and recognition to the broader community. Try to rate the applicant in this section based on their success outside of publishing papers.	**CV:** - Has won awards for presenting, teaching, writing, entrepreneurship, etc. - Has outreach activities on CV not directly related to their research - Shows leadership **Letter of Rec:** - Recommender praises the applicant's work ethic and leadership abilities - Recommender highlights extracurricular activities of applicant **Abstract:** - Abstract cites papers from applicant in high impact journals - Abstract mentions any awards won for the work which will be presented	**CV:** - Has won some awards for presenting, teaching, writing, entrepreneurship, etc. - Has many outreach activities on CV - Is actively involved in scientific societies **Letter of Rec:** - Recommender praises the applicant's work ethic and self motiviation **Abstract:** - Abstract mentions papers where it has been cited	**CV:** - Has few outreach activities on CV not directly related to their research - Is a member of atleast one scientific society **Letter of Rec:** - Recommender praises the applicant's work ethic and ability to succeed

Figure 8.1. DYSS grading rubric. The 2013 DYSS grading rubric is broken into three categories with clearly defined criteria for each part. Each of these categories is broken down into recommended scoring areas.

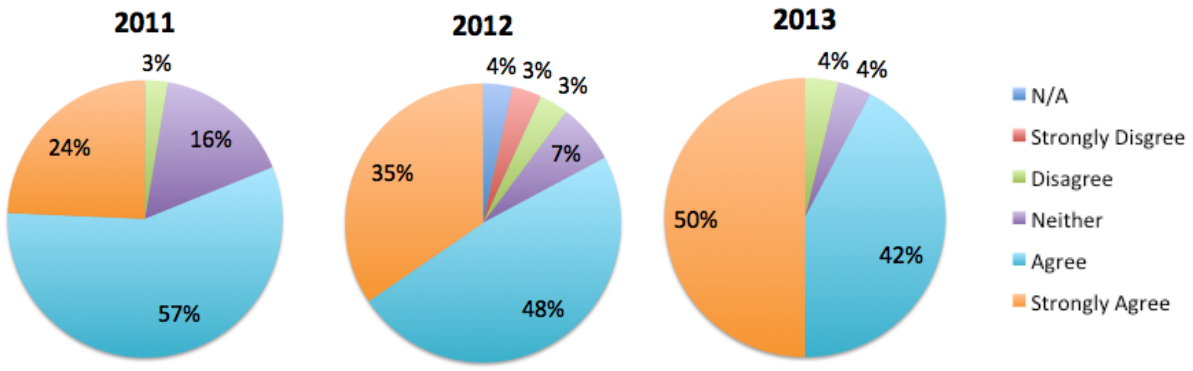

Figure 8.2. Summary of responses the statement 1. The DYSS helped me to calibrate my own abilities relative to top national candidates.

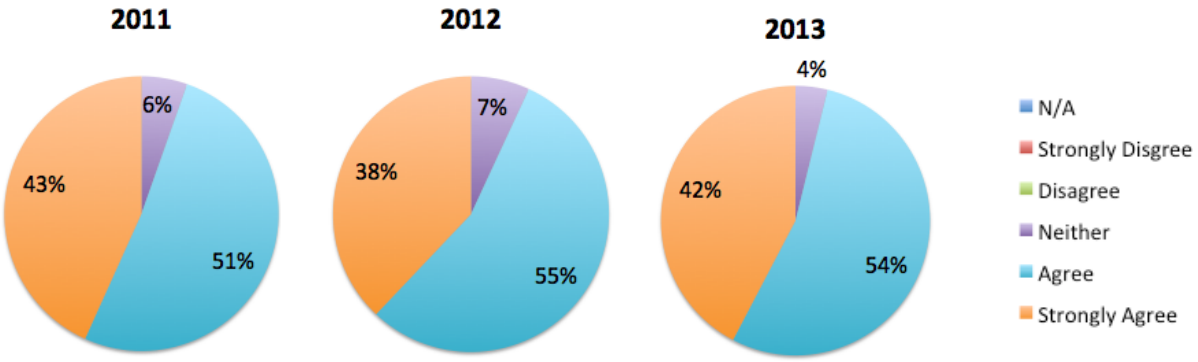

Figure 8.3. Summary of responses to the statement 2. The DYSS was a valuable way for our department to serve the discipline of chemical engineering nationally

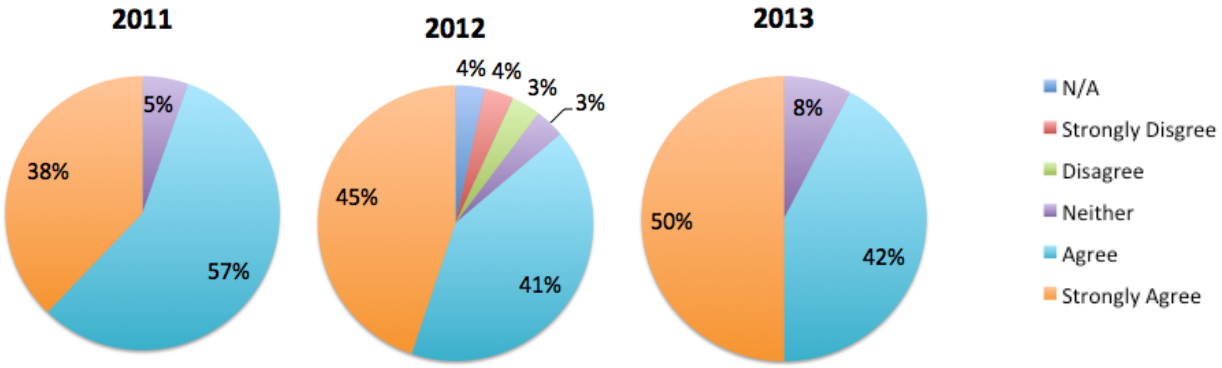

Figure 8.4. Summary of responses to the statement 3. The DYSS was a valuable experience for UW Chemical Engineering graduate students

BIBLIOGRAPHY

[1] Whitesides GM, Mathias JP, Seto CT. MOLECULAR SELF-ASSEMBLY AND NANOCHEMISTRY - A CHEMICAL STRATEGY FOR THE SYNTHESIS OF NANOSTRUCTURES. Science. 1991;254:1312-9.
[2] Whitesides GM, Grzybowski B. Self-assembly at all scales. Science. 2002;295:2418-21.
[3] Mann S. Self-assembly and transformation of hybrid nano-objects and nanostructures under equilibrium and non-equilibrium conditions. Nature Materials. 2009;8:781-92.
[4] Xia YN, Whitesides GM. Soft lithography. Annual Review of Materials Science. 1998;28:153-84.
[5] Seeman NC. Nanomaterials based on DNA. Annual review of biochemistry. 2010;79:65-87.
[6] Wang ZG, Ding B. DNA-based self-assembly for functional nanomaterials. Adv Mater. 2013;25:3905-14.
[7] Rothemund PW. Folding DNA to create nanoscale shapes and patterns. Nature. 2006;440:297-302.
[8] Zhang SG. Fabrication of novel biomaterials through molecular self-assembly. Nature Biotechnology. 2003;21:1171-8.
[9] Woolfson DN, Mahmoud ZN. More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials. Chem Soc Rev. 2010;39:3464-79.
[10] Yoo B, Kirshenbaum K. Peptoid architectures: elaboration, actuation, and application. Current opinion in chemical biology. 2008;12:714-21.
[11] Nam KT, Shelby SA, Choi PH, Marciel AB, Chen R, Tan L, et al. Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. Nat Mater. 2010;9:454-60.
[12] Kudirka R, Tran H, Sanii B, Nam KT, Choi PH, Venkateswaran N, et al. Folding of a single-chain, information-rich polypeptoid sequence into a highly ordered nanosheet. Biopolymers. 2011;96:586-95.
[13] Sanii B, Kudirka R, Cho A. Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets - Journal of the American Chemical Society (ACS Publications). Journal of the …. 2011.
[14] Robertson EJ, Oliver GK, Qian M, Proulx C, Zuckermann RN, Richmond GL. Assembly and molecular order of two-dimensional peptoid nanosheets through the oil-water interface. Proceedings of the National Academy of Sciences of the United States of America. 2014;111:13284-9.
[15] Baneyx F, Matthaei JF. Self-assembled two-dimensional protein arrays in bionanotechnology: from S-layers to designed lattices. Current Opinion in Biotechnology. 2014;28:39-45.
[16] Hampp N. Bacteriorhodopsin as a photochromic retinal protein for optical memories. Chemical Reviews. 2000;100:1755-76.
[17] Cartailler J-P, Luecke H. X-ray crystallographic analysis of lipid-protein interactions in the bacteriorhodopsin purple membrane. Annual Review of Biophysics and Biomolecular Structure. 2003;32:285-310.
[18] Kumar CSSR. Biomimetic and Bioinspired Nanomaterials: Weinheim: Wiley-VCH; 2010.
[19] Schuster B, Pum D, Sara M, Braha O, Bayley H, Sleytr UB. S-layer ultrafiltration membranes: A new support for stabilizing functionalized lipid membranes. Langmuir. 2001;17:499-503.

[20] Weigert S, Sara M. SURFACE MODIFICATION OF AN ULTRAFILTRATION MEMBRANE WITH CRYSTALLINE-STRUCTURE AND STUDIES ON INTERACTIONS WITH SELECTED PROTEIN MOLECULES. J Membr Sci. 1995;106:147-59.
[21] Moll D, Huber C, Schlegel B, Pum D, Sleytr UB, Sara M. S-layer-streptavidin fusion proteins as template for nanopatterned molecular arrays. Proceedings of the National Academy of Sciences of the United States of America. 2002;99:14646-51.
[22] Sara M, Pum D, Schuster B, sleytr UB. S-layers as patterning elements for application in nanobiotechnology. Journal of Nanoscience and Nanotechnology. 2005;5:1939-53.
[23] Allred DB, Sarikaya M, Baneyx F, Schwartz DT. Electrochemical Nanofabrication Using Crystalline Protein Masks. Nano Letters. 2005;5:609-13.
[24] Allred DB, Sarikaya M, Baneyx F, Schwartz DT. Bacterial surface-layer proteins for electrochemical nanofabrication. Pergamon-Elsevier Science Ltd; 2007. p. 193-9.
[25] Shenton W, Pum D, Sleytr UB, Mann S. Synthesis of cadmium sulphide superlattices using self-assembled bacterial S-layers. Nature. 1997;389:585-7.
[26] Bergkvist M, Mark SS, Yang X, Angert ER, Batt CA. Bionanofabrication of ordered nanoparticle arrays: Effect of particle properties and adsorption conditions. Journal of Physical Chemistry B. 2004;108:8241-8.
[27] Dieluweit S, Pum D, sleytr UB. Formation of a gold superlattice on an S-layer with square lattice symmetry. Supramolecular Science. 1998;5:15-9.
[28] Karanicolas J, Kuhlman B. Computational design of affinity and specificity at protein-protein interfaces. Curr Opin Struct Biol. 2009;19:458-63.
[29] Lanci CJ, MacDermaid CM, Kang SG, Acharya R, North B, Yang X, et al. Computational design of a protein crystal. Proc Natl Acad Sci U S A. 2012;109:7304-9.
[30] Gradiar H, Jerala R. Self-assembled bionanostructures: proteins following the lead of DNA nanostructures. Journal of nanobiotechnology. 2014;12:9.
[31] Sinclair JC. Constructing arrays of proteins. Current Opinion in Chemical Biology. 2013;17:946-51.
[32] Salgado EN, Ambroggio XI, Brodin JD, Lewis RA, Kuhlman B, Tezcan FA. Metal templated design of protein interfaces. Proc Natl Acad Sci U S A. 2010;107:1827-32.
[33] Brodin JD, Ambroggio XI, Tang C, Parent KN, Baker TS, Tezcan FA. Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. Nature Chemistry. 2012;4:375-82.
[34] Ringler P, Schulz GE. Self-assembly of proteins into designed networks. Science. 2003;302:106-9.
[35] Usui K, Maki T, Ito F, Suenaga A, Kidoaki S, Itoh M, et al. Nanoscale elongating control of the self-assembled protein filament with the cysteine-introduced building blocks. Protein Science. 2009;18:960-9.
[36] McAllister KA, Zou HL, Cochran FV, Bender GM, Senes A, Fry HC, et al. Using alpha-helical coiled-coils to design nanostructured metalloporphyrin arrays. J Am Chem Soc. 2008;130:11921-7.
[37] Zaccai NR, Chi B, Thomson AR, Boyle AL, Bartlett GJ, Bruning M, et al. A de novo peptide hexamer with a mutable channel. Nat Chem Biol. 2011;7:935-41.
[38] Stranges PB, Machius M, Miley MJ, Tripathy A, Kuhlman B. Computational design of a symmetric homodimer using beta-strand assembly. Proc Natl Acad Sci U S A. 2011;108:20562-7.

[39] Lai YT, Cascio D, Yeates T. Crystal structure of a 16 nm, half-megadalton protein cage designed by fusing symmetric oligomeric domains. Protein Sci. 2012;21:115-6.
[40] Padilla JE, Colovos C, Yeates TO. Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proceedings of the National Academy of Sciences of the United States of America. 2001;98:2217-21.
[41] Sinclair JC, Davies KM, Vénien-Bryan C, Noble MEM. Generation of protein lattices by fusing proteins with matching rotational symmetry. Nature Nanotechnology. 2011;6:558-62.
[42] Grueninger D, Treiber N, Ziegler MOP, Koetter JWA, Schulze MS, Schulz GE. Designed protein-protein association. Science. 2008;319:206-9.
[43] Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: A structural classification of protein complexes. PLoS Comput Biol. 2006;2:1395-406.
[44] Karanicolas J, Com JE, Chen I, Joachimiak LA, Dym O, Peck SH, et al. A De Novo Protein Binding Pair By Computational Design and Directed Evolution. Mol Cell. 2011;42:250-60.
[45] King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, Andre I, et al. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. Science. 2012;336:1171-4.
[46] Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch EM, et al. Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. Science. 2011;332:816-21.
[47] King NP, Bale JB, Sheffler W, McNamara DE, Gonen S, Gonen T, et al. Accurate design of co-assembling multi-component protein nanomaterials. Nature. 2014;510:103-+.
[48] Fleming KL, Matthaei JF, Pfaendtner J. A New Graduate-Level Seminar to Prepare Students for the Next Step in Their Careers. Journal of Chemical Engineering Education. 2015 ;49:29-36.
[49] Blaurock AE, Stoeckenius W. Structure of the purple membrane. Nature: New biology. 1971;233:152-5.
[50] Sara M, Sleytr UB. S-layer proteins. J Bacteriol. 2000;182:859-68.
[51] Pum D, SLEYTR UB. Reassembly of S-layer proteins. 2014:1-16.
[52] Pum D, Neubauer A, Gyorvary E, Sara M, sleytr UB. S-layer proteins as basic building blocks in a biomolecular construction kit. Nanotechnology. 2000;11:100-7.
[53] Stoecken.W, Rowen R. A MORPHOLOGICAL STUDY OF HALOBACTERIUM HALOBIUM AND ITS LYSIS IN MEDIA OF LOW SALT CONCENTRATION. J Cell Biol. 1967;34:365-&.
[54] Oesterhe.D. PURPURIC MEMBRANE FROM HALOBACTERIUM-HALOBIUM. Hoppe-Seylers Zeitschrift Fur Physiologische Chemie. 1972;353:1554-.
[55] Kühlbrandt W. Bacteriorhodopsin—the movie. Nature. 2000;406:569-70.
[56] Beveridge TJG, L L. Surface layers of bacteria. Microbiol Rev. 1991;55:684-705.
[57] Sleytr UB, Beveridge TJ. Bacterial S-layers. Trends Microbiol. 1999;7:253-60.
[58] Sleytr UB, Messner P, Pum D, Sara M. CRYSTALLINE BACTERIAL-CELL SURFACE-LAYERS. Mol Microbiol. 1993;10:911-6.
[59] Messner P, Sleytr UB. CRYSTALLINE BACTERIAL CELL-SURFACE LAYERS. Adv Microb Physiol. 1992;33:213-75.
[60] Sleytr UB, Sara M, Pum D, Schuster B. Characterization and use of crystalline bacterial cell surface layers. Prog Surf Sci. 2001;68:231-78.
[61] Sleytr UB, Egelseer EM, Ilk N, Pum D, Schuster B. S-Layers as a basic building block in a molecular construction kit. Febs Journal. 2007;274:323-34.

[62] Sleytr UB, Huber C, Ilk N, Pum D, Schuster B, Egelseer EM. S-layers as a tool kit for nanobiotechnological applications. FEMS Microbiol Lett. 2007;267:131-44.
[63] Schaffer C, Novotny R, Kupcu S, Zayni S, Scheberl A, Friedmann J, et al. Novel biocatatysts based on S-Layer self-assembly of Geobacillus Stearothermophilus NRS 2004/3a: A nano biotech nologicat approach. Small. 2007;3:1549-59.
[64] Sara M, Sleytr UB. S-layer proteins. J Bacteriol. 2000;182:859-68.
[65] Bahl H, Scholz H, Bayan N, Chami M, Leblon G, GulikKrzywicki T, et al. Molecular biology of S-layers. Fems Microbiol Rev. 1997;20:47-98.
[66] Sidhu MS, Olsen I. S-layers of Bacillus species. Microbiology-(UK). 1997;143:1039-52.
[67] Pum D, Sleytr UB. ANISOTROPIC CRYSTAL-GROWTH OF THE S-LAYER OF BACILLUS-SPHAERICUS CCM-2177 AT THE AIR/WATER INTERFACE. Colloid Surf A-Physicochem Eng Asp. 1995;102:99-104.
[68] Messner P, Pum D, Sleytr UB. CHARACTERIZATION OF THE ULTRASTRUCTURE AND THE SELF-ASSEMBLY OF THE SURFACE-LAYER OF BACILLUS-STEAROTHERMOPHILUS STRAIN NRS 2004/3A. Journal of Ultrastructure and Molecular Structure Research. 1986;97:73-88.
[69] Ilk N, Egelseer EM, Ferner-Ortner J, Kupcu S, Pum D, Schuster B, et al. Surfaces functionalized with self-assembling S-layer fusion proteins for nanobiotechnological applications. Colloid Surf A-Physicochem Eng Asp. 2008;321:163-7.
[70] Pum D, Sleytr UB. The application of bacterial S-layers in molecular nanotechnology. Trends in Biotechnology. 1999;17:8-12.
[71] Sleytr UB, Egelseer EM, Ilk N, Messner P, Schaffer C, Pum D, et al. Prokaryotic Cell Wall Components: Structure and Biochemistry Nanobiotechnological Applications of S-Layers. Berlin: Springer-Verlag Berlin; 2010.
[72] Sleytr UB, Schuster B, Pum D. Nanotechnology and biomimetics with 2-D protein crystals. IEEE Eng Med Biol Mag. 2003;22:140-50.
[73] Allred DB, Sarikaya M, Baneyx F, Schwartz DT. Bacterial surface-layer proteins for electrochemical nanofabrication. Pergamon-Elsevier Science Ltd; 2007. p. 193-9.
[74] Presenda A, Allred DB, Baneyx F, Schwartz DT, Sarikaya M. Stability of S-layer proteins for electrochemical nanofabrication. Colloid Surf B-Biointerfaces. 2007;57:256-61.
[75] Allred DB, Cheng A, Sarikaya M, Baneyx F, Schwartz DT. Three-dimensional architecture of inorganic nanoarrays electrodeposited through a surface-layer protein mask. Nano Letters. 2008;8:1434-8.
[76] Baranova EE, Fronzes RR, Garcia-Pino AA, Van Gerven NN, Papapostolou DD, Péhau-Arnaudet GG, et al. SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. Nature. 2012;487:119-22.
[77] Pavkov T, Egelseer EM, Tesarz M, Svergun DI, Sleytr UB, Keller W. The structure and binding behavior of the bacterial cell surface layer protein SbsC. Structure. 2008;16:1226-37.
[78] Baranova E, Fronzes R, Garcia-Pino A, Van Gerven N, Papapostolou D, Péhau-Arnaudet G, et al. SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. Nature. 2012.
[79] Ilk N, Egelseer EM, Sleytr UB. S-layer fusion proteins - construction principles and applications. Curr Opin Biotechnol. 2011;22:824-31.
[80] Seeman NC. Nanomaterials Based on DNA. In: Kornberg RD, Raetz CRH, Rothman JE, Thorner JW, editors. Annual Review of Biochemistry, Vol 792010. p. 65-87.

[81] Yang D, Campolongo MJ, Tran TNN, Ruiz RCH, Kahn JS, Luo D. Novel DNA materials and their applications. Wiley Interdisciplinary Reviews-Nanomedicine and Nanobiotechnology. 2010;2:648-69.
[82] Seeman NC. NUCLEIC-ACID JUNCTIONS AND LATTICES. Journal of Theoretical Biology. 1982;99:237-47.
[83] Kallenbach NR, Ma RI, Seeman NC. AN IMMOBILE NUCLEIC-ACID JUNCTION CONSTRUCTED FROM OLIGONUCLEOTIDES. Nature. 1983;305:829-31.
[84] Winfree E, Liu FR, Wenzler LA, Seeman NC. Design and self-assembly of two-dimensional DNA crystals. Nature. 1998;394:539-44.
[85] LaBean TH, Yan H, Kopatsch J, Liu FR, Winfree E, Reif JH, et al. Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes. Journal of the American Chemical Society. 2000;122:1848-60.
[86] Shen ZY, Yan H, Wang T, Seeman NC. Paranemic crossover DNA: A generalized Holliday structure with applications in nanotechnology. Journal of the American Chemical Society. 2004;126:1666-74.
[87] Rothemund PWK. Folding DNA to create nanoscale shapes and patterns. Nature. 2006;440:297-302.
[88] Pinheiro AV, Han DR, Shih WM, Yan H. Challenges and opportunities for structural DNA nanotechnology. Nature Nanotechnology. 2011;6:763-72.
[89] Williams BAR, Lund K, Liu Y, Yan H, Chaput JC. Self-assembled peptide nanoarrays: An approach to studying protein-protein interactions. Angewandte Chemie-International Edition. 2007;46:3051-4.
[90] Park SH, Yin P, Liu Y, Reif JH, LaBean TH, Yan H. Programmable DNA self-assemblies for nanoscale organization of ligands and proteins. Nano Letters. 2005;5:729-33.
[91] Lund K, Liu Y, Lindsay S, Yan H. Self-assembling a molecular pegboard. Journal of the American Chemical Society. 2005;127:17606-7.
[92] Sacca B, Meyer R, Erkelenz M, Kiko K, Arndt A, Schroeder H, et al. Orthogonal Protein Decoration of DNA Origami. Angewandte Chemie-International Edition. 2010;49:9378-83.
[93] Le JD, Pinto Y, Seeman NC, Musier-Forsyth K, Taton TA, Kiehl RA. DNA-templated self-assembly of metallic nanocomponent arrays on a surface. Nano Letters. 2004;4:2343-7.
[94] Sharma J, Ke Y, Lin C, Chhabra R, Wang Q, Nangreave J, et al. DNA-tile-directed self-assembly of quantum dots into two-dimensional nanopatterns. Angewandte Chemie-International Edition. 2008;47:5157-9.
[95] Sharma J, Chhabra R, Liu Y, Ke YG, Yan H. DNA-templated self-assembly of two-dimensional and periodical gold nanoparticle arrays. Angewandte Chemie-International Edition. 2006;45:730-5.
[96] Zheng J, Constantinou PE, Micheel C, Alivisatos AP, Kiehl RA, Seeman NC. Two-dimensional nanoparticle arrays show the organizational power of robust DNA motifs. Nano Letters. 2006;6:1502-4.
[97] Ding B, Deng Z, Yan H, Cabrini S, Zuckermann RN, Bokor J. Gold Nanoparticle Self-Similar Chain Structure Organized by DNA Origami. Journal of the American Chemical Society. 2010;132:3248-+.
[98] Maune HT, Han S-p, Barish RD, Bockrath M, Goddard WA, III, Rothemund PWK, et al. Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. Nature Nanotechnology. 2010;5:61-6.

[99] Wilner OI, Weizmann Y, Gill R, Lioubashevski O, Freeman R, Willner I. Enzyme cascades activated on topologically programmed DNA scaffolds. Nature Nanotechnology. 2009;4:249-54.
[100] Goodman RP, Heilemann M, Doose S, Erben CM, Kapanidis AN, Turberfield AJ. Reconfigurable, braced, three-dimensional DNA nanostructures. Nature Nanotechnology. 2008;3:93-6.
[101] Douglas SM, Dietz H, Liedl T, Hogberg B, Graf F, Shih WM. Self-assembly of DNA into nanoscale three-dimensional shapes. Nature. 2009;459:414-8.
[102] Andersen ES, Dong M, Nielsen MM, Jahn K, Subramani R, Mamdouh W, et al. Self-assembly of a nanoscale DNA box with a controllable lid. Nature. 2009;459:73-U5.
[103] Noren CJ, Anthonycahill SJ, Griffith MC, Schultz PG. A GENERAL-METHOD FOR SITE-SPECIFIC INCORPORATION OF UNNATURAL AMINO-ACIDS INTO PROTEINS. Science. 1989;244:182-8.
[104] Liu CC, Schultz PG. Adding New Chemistries to the Genetic Code. In: Kornberg RD, Raetz CRH, Rothman JE, Thorner JW, editors. Annual Review of Biochemistry, Vol 79. Palo Alto: Annual Reviews; 2010. p. 413-44.
[105] Woolfson DN, Ryadnov MG. Peptide-based fibrous biomaterials: some things old, new and borrowed. Current Opinion in Chemical Biology. 2006;10:559-67.
[106] Woolfson DN, Mahmoud ZN. More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials. Chemical Society Reviews. 2010;39:3464-79.
[107] Woolfson DN. The design of coiled-coil structures and assemblies. AdvProtein Chem. 2005;70:79-+.
[108] Hartgerink JD, Beniash E, Stupp SI. Self-assembly and mineralization of peptide-amphiphile nanofibers. Science. 2001;294:1684-8.
[109] Kotch FW, Raines RT. Self-assembly of synthetic collagen triple helices. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:3028-33.
[110] Koide T. Triple helical collagen-like peptides: Engineering and applications in matrix biology. Connective Tissue Research. 2005;46:131-41.
[111] Zuckermann RN. Peptoid Origins. Biopolymers. 2011;96:545-55.
[112] Rosales AM, Murnen HK, Kline SR, Zuckermann RN, Segalman RA. Determination of the persistence length of helical and non-helical polypeptoids in solution. Soft Matter. 2012;8:3673-80.
[113] Rosales AM, Murnen HK, Zuckermann RN, Segalman RA. Control of Crystallization and Melting Behavior in Sequence Specific Polypeptoids. Macromolecules. 2010;43:5627-36.
[114] Miller SM, Simon RJ, Ng S, Zuckermann RN, Kerr JM, Moos WH. PROTEOLYTIC STUDIES OF HOMOLOGOUS PEPTIDE AND N-SUBSTITUTED GLYCINE PEPTOID OLIGOMERS. Bioorganic & Medicinal Chemistry Letters. 1994;4:2657-62.
[115] Miller SM, Simon RJ, Ng S, Zuckermann RN, Kerr JM, Moos WH. COMPARISON OF THE PROTEOLYTIC SUSCEPTIBILITIES OF HOMOLOGOUS L-AMINO-ACID, D-AMINO-ACID, AND N-SUBSTITUTED GLYCINE PEPTIDE AND PEPTOID OLIGOMERS. Drug Development Research. 1995;35:20-32.
[116] Zuckermann RN. THE CHEMICAL SYNTHESIS OF PEPTIDOMIMETIC LIBRARIES. Current Opinion in Structural Biology. 1993;3:580-4.
[117] Sun J, Zuckermann RN. Peptoid Polymers: A Highly Designable Bioinspired Material. Acs Nano. 2013;7:4715-32.

[118] Gao CMY, A. Y.; Wang, X.; Magdangal, E.; Salisbury, C.; Peretz, D.; Zuckermann, R. N.; Connolly, M. D.; Hansson, O.; Minthon, L. Aβ40 Oligomers Identified as a Potential Biomarker for the Diagnosis of Alzheimer's Disease. PLoS One 2010.
[119] Yam AY, Wang X, Gao CM, Connolly MD, Zuckermann RN, Bleu T, et al. A Universal Method for Detection of Amyloidogenic Misfolded Proteins. Biochemistry. 2011;50:4322-9.
[120] Vanderstichele H, Kodadek T. Roadblocks for integration of novel biomarker concepts into clinical routine: the peptoid approach. Alzheimers Res Ther. 2014;6:8.
[121] Kirshenbaum K, Barron AE, Goldsmith RA, Armand P, Bradley EK, Truong KT, et al. Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure. Proceedings of the National Academy of Sciences of the United States of America. 1998;95:4303-8.
[122] Lee BC, Zuckermann RN, Dill KA. Folding a nonbiological polymer into a compact multihelical structure. Journal of the American Chemical Society. 2005;127:10999-1009.
[123] Chongsiriwatana NP, Patch JA, Czyzewski AM, Dohm MT, Ivankin A, Gidalevitz D, et al. Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides. Proceedings of the National Academy of Sciences of the United States of America. 2008;105:2794-9.
[124] Holub JM, Jang HJ, Kirshenbaum K. Clickity-click: highly functionalized peptoid oligomers generated by sequential conjugation reactions on solid-phase support. Organic & Biomolecular Chemistry. 2006;4:1497-502.
[125] Holub JM, Garabedian MJ, Kirshenbaum K. Peptoids on steroids: Precise multivalent estradiol-peptidomimetic conjugates generated via azide-alkyne 3+2 cycloaddition reactions. Qsar & Combinatorial Science. 2007;26:1175-80.
[126] Lee B-C, Chu TK, Dill KA, Zuckermann RN. Biomimetic nanostructures: Creating a high-affinity zinc-binding site in a folded nonbiological polymer. Journal of the American Chemical Society. 2008;130:8847-55.
[127] Chen C-L, Qi J, Zuckermann RN, DeYoreo JJ. Engineered Biomimetic Polymers as Tunable Agents for Controlling CaCO3 Mineralization. Journal of the American Chemical Society. 2011;133:5214-7.
[128] Rosales AM, Segalman RA, Zuckermann RN. Polypeptoids: a model system to study the effect of monomer sequence on polymer properties and self-assembly. Soft Matter. 2013;9:8400-14.
[129] Lau KHA. Peptoids for biomaterials science. Biomaterials Science. 2014;2:627-33.
[130] Horne WS. Peptide and peptoid foldamers in medicinal chemistry. Expert Opinion on Drug Discovery. 2011;6:1247-62.
[131] Dohm MT, Kapoor R, Barron AE. Peptoids: Bio-Inspired Polymers as Potential Pharmaceuticals. Current Pharmaceutical Design. 2011;17:2732-47.
[132] Nam KT, Shelby SA, Choi PH, Marciel AB, Chen R, Tan L, et al. Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. Nature Materials. 2010;9:454-60.
[133] Murnen HK, Rosales AM, Jaworsk JN, Segalman RA, Zuckermann RN. Hierarchical Self-Assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices. Journal of the American Chemical Society. 2010;132:16112-9.
[134] Sanii B, Kudirka R, Cho A, Venkateswaran N, Olivier GK, Olson AM, et al. Shaken, not stirred: collapsing a peptoid monolayer to produce free-floating, stable nanosheets. Journal of the American Chemical Society. 2011;133:20808-15.

[135] Baranova E, Fronzes R, Garcia-Pino A, Van Gerven N, Papapostolou D, Pehau-Arnaudet G, et al. SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. Nature. 2012;487:119-22.
[136] Mader C, Huber C, Moll D, Sleytr UB, Sara M. Interaction of the crystalline bacterial cell surface layer protein SbsB and the secondary cell wall polymer of Geobacillus stearothermophilus PV72 assessed by real-time surface plasmon resonance biosensor technology. Journal of Bacteriology. 2004;186:1758-68.
[137] He Y, Chen Y, Liu HP, Ribbe AE, Mao CD. Self-assembly of hexagonal DNA two-dimensional (2D) arrays. Journal of the American Chemical Society. 2005;127:12202-3.
[138] Freeman S. Biological Science. Upper Saddle River, NJ: Pearson Prentice Hall; 2005.
[139] Smith AM, Banwell EF, Edwards WR, Pandya MJ, Woolfson DN. Engineering increased stability into self-assembled protein fibers. Advanced Functional Materials. 2006;16:1022-30.
[140] Fishwick CWG, Beevers AJ, Carrick LM, Whitehouse CD, Aggeli A, Boden N. Structures of helical beta-tapes and twisted ribbons: The role of side-chain interactions on twist and bend behavior. Nano Letters. 2003;3:1475-9.
[141] Kudirka R, Tran H, Sanii B, Nam KT, Choi PH, Venkateswaran N, et al. Folding of a single-chain, information-rich polypeptoid sequence into a highly ordered nanosheet. Biopolymers. 2011;96:586-95.
[142] Cartailler JP, Luecke H. X-ray crystallographic analysis of lipid-protein interactions in the bacteriorhodopsin purple membrane. Annual review of biophysics and biomolecular structure. 2003;32:285-310.
[143] Hampp N. Bacteriorhodopsin as a photochromic retinal protein for optical memories. Chehm Rev. 2000;100:1755-76.
[144] Sleytr UB, Huber C, Ilk N, Pum D, Schuster B, Egelseer EM. S-layers as a tool kit for nanobiotechnological applications. FEMS microbiology letters. 2007;267:131-44.
[145] Howorka S. Rationally engineering natural protein assemblies in nanobiotechnology. Curr Opin Biotechnol. 2011;22:485-91.
[146] Pum D, Toca-Herrera JL, Sleytr UB. S-layer protein self-assembly. Int J Mol Sci. 2013;14:2484-501.
[147] Messner P, Pum D, Sleytr UB. Characterization of the ultrastructure and the self-assembly of the surface layer of Bacillus stearothermophilus strain NRS 2004/3a. Journal of ultrastructure and molecular structure research. 1986;97:73-88.
[148] Shenton W, Pum D, Sleytr UB, Mann S. Synthesis of cadmium sulphide superlattices using self-assembled bacterial S-layers. Nature. 1997;389:585-7.
[149] Allred DB, Cheng A, Sarikaya M, Baneyx F, Schwartz DT. 3D architecture of inorganic nanoarrays electrodeposited through an S-layer protein mask. Nano Lett. 2008;8:1434-8.
[150] Allred DB, Sarikaya M, Baneyx F, Schwartz DT. Electrochemical nanofabrication using crystalline protein masks. Nano Lett. 2005;5:609-13.
[151] Wildhaber I, Baumeister W. THE CELL-ENVELOPE OF THERMOPROTEUS-TENAX - 3-DIMENSIONAL STRUCTURE OF THE SURFACE-LAYER AND ITS ROLE IN SHAPE MAINTENANCE. Embo Journal. 1987;6:1475-80.
[152] Wildhaber I, Santarius U, Baumeister W. 3-DIMENSIONAL STRUCTURE OF THE SURFACE PROTEIN OF DESULFUROCOCCUS-MOBILIS. Journal of Bacteriology. 1987;169:5563-8.

[153] Howorka S, Sara M, Wang YJ, Kuen B, Sleytr UB, Lubitz W, et al. Surface-accessible residues in the monomeric and assembled forms of a bacterial surface layer protein. Journal of Biological Chemistry. 2000;275:37876-86.
[154] Kinns H, Howorka S. The surface location of individual residues in a bacterial S-layer protein. Journal of Molecular Biology. 2008;377:589-604.
[155] Chung S, Shin SH, Bertozzi CR, De Yoreo JJ. Self-catalyzed growth of S-layers via an amorphous-to-crystalline transition limited by folding kinetics. Proc Natl Acad Sci U S A. 2010;107:16536-41.
[156] Shin SH, Chung S, Sanii B, Comolli LR, Bertozzi CR, De Yoreo JJ. Direct observation of kinetic traps associated with structural transformations leading to multiple pathways of S-layer assembly. Proc Natl Acad Sci U S A. 2012;109:12968-73.
[157] Horejs C, Gollner H, Pum D, Sleytr UB, Peterlik H, Jungbauer A, et al. Atomistic structure of monomolecular surface layer self-assemblies: toward functionalized nanostructures. ACS Nano. 2011;5:2288-97.
[158] Huber C, Ilk N, Rünzler D, Egelseer EM, Weigert S, Sleytr UB, et al. The three S-layer-like homology motifs of the S-layer protein SbpA of Bacillus sphaericus CCM 2177 are not sufficient for binding to the pyruvylated secondary cell wall polymer. Mol Microbiol. 2005;55:197-205.
[159] Shin SH, Comolli LR, Tscheliessnig R, Wang C, Nam KT, Hexemer A, et al. Self-assembly of "S-bilayers", a step towards expanding the dimensionality of S-layer assemblies. ACS Nano. 2013;7:4946-53.
[160] Karanicolas J, Kuhlman B. Computational design of affinity and specificity at protein-protein interfaces. Curr Opin Struct Biol. 2009;19:458-63.
[161] Mandell DJ, Kortemme T. Computer-aided design of functional protein interactions. Nature chemical biology. 2009;5:797-807.
[162] Brodin JD, Ambroggio XI, Tang C, Parent KN, Baker TS, Tezcan FA. Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. Nature chemistry. 2012;4:375-82.
[163] Salgado EN, Ambroggio XI, Brodin JD, Lewis RA, Kuhlman B, Tezcan FA. Metal templated design of protein interfaces. Proc Natl Acad Sci U S A. 2010;107:1827-32.
[164] Medina-Morales A, Perez A, Brodin JD, Tezcan FA. In vitro and cellular self-assembly of a Zn-binding protein cryptand via templated disulfide bonds. J Am Chem Soc. 2013;135:12013-22.
[165] Lanci CJ, MacDermaid CM, Kang SG, Acharya R, North B, Yang X, et al. Computational design of a protein crystal. Proc Natl Acad Sci U S A. 2012;109:7304-9.
[166] McAllister KA, Zou H, Cochran FV, Bender GM, Senes A, Fry HC, et al. Using alpha-helical coiled-coils to design nanostructured metalloporphyrin arrays. J Am Chem Soc. 2008;130:11921-7.
[167] Zaccai NR, Chi B, Thomson AR, Boyle AL, Bartlett GJ, Bruning M, et al. A de novo peptide hexamer with a mutable channel. Nature chemical biology. 2011;7:935-41.
[168] Stranges PB, Machius M, Miley MJ, Tripathy A, Kuhlman B. Computational design of a symmetric homodimer using beta-strand assembly. Proc Natl Acad Sci U S A. 2011;108:20562-7.
[169] Lai YT, Cascio D, Yeates TO. Structure of a 16-nm cage designed by using protein oligomers. Science. 2012;336:1129.

[170] Padilla JE, Colovos C, Yeates TO. Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proc Natl Acad Sci U S A. 2001;98:2217-21.
[171] Sinclair JC, Davies KM, Venien-Bryan C, Noble ME. Generation of protein lattices by fusing proteins with matching rotational symmetry. Nature nanotechnology. 2011;6:558-62.
[172] Grueninger D, Treiber N, Ziegler MO, Koetter JW, Schulze MS, Schulz GE. Designed protein-protein association. Science. 2008;319:206-9.
[173] Karanicolas J, Corn JE, Chen I, Joachimiak LA, Dym O, Peck SH, et al. A de novo protein binding pair by computational design and directed evolution. Molecular cell. 2011;42:250-60.
[174] Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch EM, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science. 2011;332:816-21.
[175] Lai YT, King NP, Yeates TO. Principles for designing ordered protein assemblies. Trends Cell Biol. 2012;22:653-61.
[176] Fairman R, Chao HG, Lavoie TB, Villafranca JJ, Matsueda GR, Novotny J. Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu(-)-Lys(+) ion pair interactions. Biochemistry. 1996;35:2824-9.
[177] Baneyx F, Schwartz DT. Selection and analysis of solid-binding peptides. Current Opinion in Biotechnology. 2007;18:312-7.
[178] Sarikaya M, Tamerler C, Jen AKY, Schulten K, Baneyx F. Molecular biomimetics: nanotechnology through biology. Nature Materials. 2003;2:577-85.
[179] Coyle BL, Zhou W, Baneyx F. Protein-aided Mineralization of Inorganic Nanostructures: Caister Academic Press, 32 Hewitts Lane, Wymondham Nr 18 0ja, Uk; 2013.
[180] Knobloch D, Ostermann K, Roedel G. Production, Secretion, and Cell Surface Display of Recombinant Sporosarcina ureae S-Layer Fusion Proteins in Bacillus megaterium. Applied and Environmental Microbiology. 2012;78:560-7.
[181] Rothbauer M, Küpcü S, Sticker D, Sleytr UB, Ertl P. Exploitation of S-layer anisotropy: pH-dependent nanolayer orientation for cellular nanopatterning. ACS Nano. 2013;9:8020-30.
[182] King NP, Lai YT. Practical approaches to designing novel protein assemblies. Current Opinion in Structural Biology. 2013;23:632-8.
[183] Lai YT, King NP, Yeates TO. Principles for designing ordered protein assemblies. Trends in Cell Biology. 2012;22:653-61.
[184] Zhang J, Zheng F, Grigoryan G. Design and designability of protein-based assemblies. Current Opinion in Structural Biology. 2014;27:79-86.
[185] Bozic S, Doles T, Gradisar H, Jerala R. New designed protein assemblies. Current Opinion in Chemical Biology. 2013;17:940-5.
[186] Mandell DJ, Kortemme T. Computer-aided design of functional protein interactions. Nature Chemical Biology. 2009;5:797-807.
[187] Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, et al. ROSETTA3: AN OBJECT-ORIENTED SOFTWARE SUITE FOR THE SIMULATION AND DESIGN OF MACROMOLECULES. In: Johnson ML, Brand L, editors. Methods in Enzymology, Vol 487: Computer Methods, Pt C2011. p. 545-74.
[188] DiMaio F, Leaver-Fay A, Bradley P, Baker D, Andre I. Modeling Symmetric Macromolecular Structures in Rosetta3. Plos One. 2011;6.
[189] Goodsell DS, Olson AJ. Structural symmetry and protein function. Annual Review of Biophysics and Biomolecular Structure. 2000;29:105-53.

[190] Abola EE, Sussman JL, Prilusky J, Manning NO. Protein data bank archives of three-dimensional macromolecular structures. In: Carter CW, Sweet RM, editors. Macromolecular Crystallography, Pt B1997. p. 556-71.
[191] Lai Y-T, Tsai K-L, Sawaya MR, Asturias FJ, Yeates TO. Structure and Flexibility of Nanoscale Protein Cages Designed by Symmetric Self-Assembly. Journal of the American Chemical Society. 2013;135:7738-43.
[192] Salgado EN, Radford RJ, Tezcan FA. Metal-Directed Protein Self-Assembly. Accounts of Chemical Research. 2010;43:661-72.
[193] Salgado EN, Ambroggio XI, Brodin JD, Lewis RA, Kuhlman B, Tezcan FA. Metal templated design of protein interfaces. Proceedings of the National Academy of Sciences of the United States of America. 2010;107:1827-32.
[194] Fegan A, White B, Carlson JCT, Wagner CR. Chemically Controlled Protein Assembly: Techniques and Applications. Chemical Reviews. 2010;110:3315-36.
[195] Carlson JCT, Jena SS, Flenniken M, Chou TF, Siegel RA, Wagner CR. Chemically controlled self-assembly of protein nanorings. Journal of the American Chemical Society. 2006;128:7630-8.
[196] Carlson JCT, Kanter A, Thuduppathy GR, Cody V, Pineda PE, McIvor RS, et al. Designing protein dimerizers: The importance of ligand conformational equilibria. Journal of the American Chemical Society. 2003;125:1501-7.
[197] Song YF, DiMaio F, Wang RYR, Kim D, Miles C, Brunette TJ, et al. High-Resolution Comparative Modeling with RosettaCM. Structure. 2013;21:1735-42.
[198] Kaufmann KW, Lemmon GH, DeLuca SL, Sheehan JH, Meiler J. Practically Useful: What the ROSETTA Protein Modeling Suite Can Do for You. Biochemistry. 2010;49:2987-98.
[199] Rohl CA, Strauss CEM, Misura KMS, Baker D. Protein structure prediction using rosetta. Methods Enzymol. 2004;383:66-+.
[200] Nannenga BL, Shi D, Leslie AGW, Gonen T. High-resolution structure determination by continuous-rotation data collection in MicroED. Nature Methods. 2014;11:927-30.
[201] Nannenga BL, Iadanza MG, Vollmar BS, Gonen T. Overview of Electron Crystallography of Membrane Proteins: Crystallization and Screening Strategies Using Negative Stain Electron Microscopy. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2001.
[202] Baneyx F, Matthaei JF. Self-assembled two-dimensional protein arrays in bionanotechnology: from S-layers to designed lattices. Curr Opin Biotechnol. 2014;28:39-45.
[203] Butler SZ, Hollen SM, Cao L, Cui Y, Gupta JA, Gutierrez HR, et al. Progress, challenges, and opportunities in two-dimensional materials beyond graphene. ACS Nano. 2013;7:2898-926.
[204] Colson JW, Dichtel WR. Rationally synthesized two-dimensional polymers. Nature chemistry. 2013;5:453-65.
[205] Palma CA, Samori P. Blueprinting macromolecular electronics. Nature chemistry. 2011;3:431-6.
[206] Whittell GR, Hager MD, Schubert US, Manners I. Functional soft materials from metallopolymers and metallosupramolecular polymers. Nat Mater. 2011;10:176-88.
[207] Fletcher JM, Harniman RL, Barnes FR, Boyle AL, Collins A, Mantell J, et al. Self-assembling cages from coiled-coil peptide modules. Science. 2013;340:595-9.
[208] King NP, Bale JB, Sheffler W, McNamara DE, Gonen S, Gonen T, et al. Accurate design of co-assembling multi-component protein nanomaterials. Nature. 2014;510:103-8.

[209] Lai YT, Reading E, Hura GL, Tsai KL, Laganowsky A, Asturias FJ, et al. Structure of a designed protein cage that self-assembles into a highly porous cube. Nature chemistry. 2014;6:1065-71.
[210] Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: a structural classification of protein complexes. PLoS computational biology. 2006;2:e155.
[211] DiMaio F, Leaver-Fay A, Bradley P, Baker D, Andre I. Modeling symmetric macromolecular structures in Rosetta3. PloS one. 2011;6:e20450.
[212] Baneyx F, Schwartz DT. Selection and analysis of solid-binding peptides. Curr Opin Biotechnol. 2007;18:312-7.
[213] Andre I, Bradley P, Wang C, Baker D. Prediction of the structure of symmetrical protein assemblies. Proc Natl Acad Sci U S A. 2007;104:17656-61.
[214] Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. Methods Enzymol. 2011;487:545-74.
[215] Qian B, Raman S, Das R, Bradley P, McCoy AJ, Read RJ, et al. High-resolution structure prediction and the crystallographic phase problem. Nature. 2007;450:259-64.
[216] Teixeira LM, Strickland A, Mark SS, Bergkvist M, Sierra-Sastre Y, Batt CA. Entropically driven self-assembly of Lysinibacillus sphaericus S-layer proteins analyzed under various environmental conditions. Macromolecular bioscience. 2010;10:147-55.
[217] Svergun D, Barberato C, Koch MHJ. CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr. 1995;28:768-73.
[218] Ilk N, Egelseer EM, Sleytr UB. S-layer fusion proteins--construction principles and applications. Curr Opin Biotechnol. 2011;22:824-31.
[219] Pum D, Sleytr UB. The application of bacterial S-layers in molecular nanotechnology. Trends Biotechnol. 1999;17:8-12.
[220] Heckman KL, Pease LR. Gene splicing and mutagenesis by PCR-driven overlap extension. Nature Protocols. 2007;2:924-32.
[221] Schneider EL, Thomas JG, Bassuk JA, Sage EH, Baneyx F. Manipulating the aggregation and oxidation of human SPARC in the cytoplasm of *Escherichia coli*. Nature Biotech. 1997;15:581-5.
[222] Teixeira LM, Strickland A, Mark SS, Bergkvist M, Sierra-Sastre Y, Batt CA. Entropically Driven Self-Assembly of Lysinibacillus sphaericus S-Layer Proteins Analyzed Under Various Environmental Conditions. Macromol Biosci. 2010;10:147-55.
[223] Chen R, Lu JJ, Wang HF. Rapid and Efficient Gene Splicing Using Megaprimer-Based Protocol. Molecular Biotechnology. 2008;40:224-30.
[224] Karsenti E. Self-organization in cell biology: a brief history. Nature Reviews Molecular Cell Biology. 2008;9:255-62.
[225] Howarth M, Ting AY. Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. Nature Protocols. 2008;3:534-45.
[226] Coyle BL, Baneyx F. A Cleavable Silica-Binding Affinity Tag for Rapid and Inexpensive Protein Purification. Biotechnology and Bioengineering. 2014;111:2019-26.
[227] Coyle BL, Rolandi M, Baneyx F. Carbon-Binding Designer Proteins that Discriminate between sp(2)- and sp(3)-Hybridized Carbon Surfaces. Langmuir. 2013;29:4839-46.
[228] Naik RR, Stringer SJ, Agarwal G, Jones SE, Stone MO. Biomimetic synthesis and patterning of silver nanoparticles. Nature Materials. 2002;1:169-72.

[229] Czerwinski M, Krop-Watorek A, Wasniowska K, Smolarek D, Spitalnik SL. Construction of an agglutination tool: recombinant Fab fragments biotinylated in vitro. New Biotechnology. 2009;26:215-21.
[230] Zrazhevskiy P, Sena M, Gao XH. Designing multifunctional quantum dots for bioimaging, detection, and drug delivery. Chemical Society Reviews. 2010;39:4326-54.
[231] Medintz IL, Uyeda HT, Goldman ER, Mattoussi H. Quantum dot bioconjugates for imaging, labelling and sensing. Nature Materials. 2005;4:435-46.
[232] Jamieson T, Bakhshi R, Petrova D, Pocock R, Imani M, Seifalian AM. Biological applications of quantum dots. Biomaterials. 2007;28:4717-32.
[233] Chan WCW, Nie SM. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. Science. 1998;281:2016-8.
[234] Zhou W, Schwartz DT, Baneyx F. Single-Pot Biofabrication of Zinc Sulfide Immuno-Quantum Dots. Journal of the American Chemical Society. 2010;132:4731-8.
[235] Zhou WB, Baneyx F. Biofabrication of ZnS:Mn luminescent nanocrystals using histidine, hexahistidine, and His-tagged proteins: A comparison study. Biochemical Engineering Journal. 2014;89:28-32.
[236] Zhou W, Baneyx F. Aqueous, protein-driven synthesis of transition metal-doped ZnS immuno-quantum dots. ACS nano. 2011;5:8013-8.
[237] Chen Y, Munechika K, Ginger DS. Dependence of Fluorescence Intensity on the Spectral Overlap between Fluorophores and Plasmon Resonant Single Silver Nanoparticles. Nano Letters. 2007;7:690-6.
[238] Munechika K, Smith JM, Chen Y, Ginger DS. Plasmon line widths of single silver nanoprisms as a function of particle size and plasmon peak position. Journal of Physical Chemistry C. 2007;111:18906-11.
[239] Chen Y, Munechika K, Plante IJL, Munro AM, Skrabalak SE, Xia Y, et al. Excitation enhancement of CdSe quantum dots by single metal nanoparticles. Applied Physics Letters. 2008;93:3.
[240] Chen Y, Munechika K, Ginger DS. Bioenabled nanophotonics. Mrs Bulletin. 2008;33:536-42.
[241] Sengupta A, Thai CK, Sastry MSR, Matthaei JF, Schwartz DT, Davis EJ, et al. A genetic approach for controlling the binding and orientation of proteins on nanoparticles. Langmuir. 2008;24:2000-8.
[242] Grosh CD, Schwartz DT, Baneyx F. Protein-Based Control of Silver Growth Habit Using Electrochemical Deposition. Crystal Growth & Design. 2009;9:4401-6.
[243] Shi WL, Sahoo Y, Swihart MT, Prasad PN. Gold nanoshells on polystyrene cores for control of surface plasmon resonance. Langmuir. 2005;21:1610-7.
[244] Vanderford NL. Broadening PhD curricula. Nat Biotechnol. 2012;30:113-4.
[245] Kerr E. Too many scientists, too few academic jobs. Nat Med (N Y, NY, U S). 1995;1:14-.
[246] Schuster SM. BAMBED commentary: Post‚ÄêPhD education. Biochem Mol Biol Educ. 2009;37:381-2.
[247] Rethinking aspects of the training system. Nat Struct Mol Biol. 2001;8:587.
[248] Investing in the Future: Strategic Plan for Biomedical and Behavioral Research Training. U.S. Department of Health and Human Services, National Institute of Health, National Institute of General Medical Sciences; 2011.
[249] Mervis J. NIH report urges greater emphasis on training for all graduate students. Science. 2011;331:525-.

[250] Fine EaJH. Reviewing Applicants: Research on Bias and Assumptions. In: University of Wisconsin-Madison WiSELIUo, editor.2012.
[251] Russ T, Simonds C, Hunt S. Coming out in the classroom... An occupational hazard?: The influence of sexual orientation on teacher credibility and perceived student learning. Communication Education. 2002;51:311-24.
[252] Tierney J. Do You Suffer From Decision Fatigue? The New York Times2011.
[253] Danziger S, Levav J, Avnaim-Pesso L. Extraneous factors in judicial decisions. Proceedings of the National Academy of Sciences. 2011;108:6889-92.
[254] Biernat M, Manis M. Shifting standards and stereotype-based judgments. Journal of personality and social psychology. 1994;66:5.
[255] Devine PG, Plant EA, Amodio DM, Harmon-Jones E, Vance SL. The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice. Journal of personality and social psychology. 2002;82:835.
[256] Wenneras C, Wold A. Nepotism and sexism in peer-review. Women, Science, and Technology: a reader in feminist science studies. 2001:46-52.
[257] Steinpreis RE, Anders KA, Ritzke D. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. Sex roles. 1999;41:509-28.
[258] Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: Agentic and communal differences. Journal of Applied Psychology. 2009;94:1591.
[259] Carnes M, Geller S, Fine E, Sheridan J, Handelsman J. NIH director's pioneer awards: Could the selection process be biased against women? Journal of Women's Health. 2005;14:684-91.
[260] Heiman C. 30 Under 30 - The Future of ENergy is Nuclear. Forbes New York City2012.