CONSTRUCTION OF NANOSCALE SUPRAMOLECULAR COMPLEXES USING

ENGINEERED PROTEIN CAPSIDS

by

Larry Sherman Grant III

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

December 2016

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of   **Larry Sherman Grant III**

has been approved by the following supervisory committee members:

Charles Dale Poulter	, Chair	**6/30/2016**
		Date Approved
Kenneth Woycechowsky	, Member	**5/24/2016**
		Date Approved
Jennifer Heemstra	, Member	**6/15/2016**
		Date Approved
Ilya Zharov	, Member	**6/23/2016**
		Date Approved
Saveez Saffarian	, Member	**7/1/2016**
		Date Approved

and by   **Cynthia Burrows**   , Chair/Dean of

the Department/College/School of   **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Most proteins assemble into oligomeric complexes. These supramolecular associations may confer many advantages to the substituents. Polyhedral capsids represent a common, highly symmetric nanoscale architecture in which multiple subunits self-assemble to form a hollow three-dimensional surface which often serve as molecular containers or platforms for multivalent display of ligands. Capsids can be tailored to serve in applications such as drug delivery, biocatalysis, and materials synthesis. In this dissertation, I present a body of work undertaken on the *Aquifex aeolicus* lumazine synthase (AaLS) capsid to expand our knowledge of supramolecular protein associations and to generate capsids with novel functions.

First, the construction and characterization of a novel nanoreactor is described. Using a previously established tagging system, an esterase was encapsulated in a laboratory-evolved variant of AaLS. Characterization of the purified complex shows an average loading of two esterases per capsid and an approximately 20-fold decrease in efficiency compared to the free esterase. This decrease is larger than most of the previously reported capsid-based nanoreactor systems which suggests that both the confinement molarity and the electrostatic environment of the capsid interior may significantly influence the kinetic parameters of guest enzymes.

Second, I utilize charge complementarity to decorate the exterior of an AaLS capsid variant with green fluorescent protein (GFP). A new interface was engineered by

negatively supercharging the five-fold symmetric capsid pores and appending a deca-arginine tag to the C-terminus of GFP. This interaction requires the engineered features of both binding partners and shows steep dependence on the buffer ionic strength, although it retains high affinity at physiological ionic strength. Thus, charge complementarity can provide a simple, powerful, and general method for designing protein associations *de novo.*

Finally, I expand upon previous work in which a redox switch was developed to control capsid assembly. The original switch relies on the formation of a disulfide-bonded adduct between a pentameric variant of AaLS and thiophenol. I explore alterations to the prosthetic group structure which reveal that the three-fold symmetric interface of the assembled capsid is highly plastic and can tolerate a range of different adduct sizes and shapes. These studies also identified two new disassembly switches, providing greater control over the supramolecular chemistry of the AaLS capsid.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AaLS	*Aquifex aeolicus* lumazine synthase
BsLS	*Bacillus subtilis* lumazine synthase
BsRS	*Bacillus subtilis* riboflavin synthase
CCMV	cowpea chlorotic mottle virus
DTT	dithiothreitol
DTNB	5,5'-dithiobis-(2-nitrobenzoate)
Dnase	deoxyribonuclease
E1	pyruvate dehydrogenase
E2	dihydrolipoamide acetyltransferase
E3	dihydrolipoamide dehydrogenase
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
EX	expanded bacteriophage p22 capsid
FPLC	fast protein liquid chromatography
GFP	green fluorescent protein
HEPB	hepatitis B virus
HIV	human immunodeficiency virus
IPTG	isopropyl β-D-1-thiogalactopyranoside
LC-MS/MS	liquid chromatography tandem mass spectrometry
MS2	bacteriophage MS2
Ni^{2+}-NTA	Nickel-Nitrilotriacetic acid
NTB	2-nitro-5-thiobenzoate
P22	bacteriophage P22
PAGE	polyacrylamide gel electrophoresis
PC	bacteriophage P22 procapsid
PDB	protein data bank
PDC	pyruvate dehydrogenase complex
PSBD	peripheral subunit binding domain
Qß	bacteriophage Qß
R_{10}	deca-arginine
RNA	ribonucleic acid
Rnase	ribonuclease
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
ScLS	*Saccharomyces cerevisiae* lumazine synthase
TEM	transmission electron microscopy
WB	wiffleball form of bacteriophage P22 capsid

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1

INTRODUCTION

1.1 Supramolecular Assembly of Proteins

1.1.1 Benefits of Oligomerization

Studies over the last two decades have consistently demonstrated that oligomerization plays a crucial role in a variety of cellular pathways and functions including structural maintenance,[1] cell division,[2] gene repair,[3] pathogen defense,[4] and apoptosis,[5] among others. The vast majority of proteins do not exist in their functional states as monomers. They assemble into supramolecular complexes containing two or more identical (homooligomer) or different (heterooligomer) subunits. Analysis of the Brenda enzyme database has shown that, for enzymes whose subunit composition is known, at most only one third are monomers.[6] The rest are multimers with dimers and tetramers being the most prevalent higher order associations. Analysis of structures deposited in the PDB yields similar results. In both databases, homoologimers significantly outnumber heterooligomers.

While supramolecular assembly in general confers a number of advantages over monomers such as greater stability,[7] resistance to proteases,[8] and enhanced enzymatic activity,[9] homooligomerization in particular is believed to provide additional benefits.[10] The use of a single, uniform building block reduces the minimum necessary size of an organism's genome. A smaller genome reduces resource expenditure during replication and is particularly important for viruses as they have a very limited space in which to pack their genetic material. Further, evolution of a single gene to achieve homooligomerization is likely to be faster and simpler than the same process used to achieve heterooligomerization of multiple gene products. Having smaller functional units is also more efficient in the sense that disposal of improperly formed subunits and subsequent

production of a replacement is much less energy and resource intensive.

Uniform building blocks also give rise to high levels of symmetry which have been shown to provide the most stable associations.[11] They achieve this stability in two ways. First, highly specific interfaces, such as those typically found between homooligomers, are stabilized by significant complementarity in the shape and charge of these interfaces. This specificity gives rise to strongly directional associations which, in turn, results in highly symmetric assemblies. Second, highly symmetric structures have a tendency to form closed complexes. These complexes provide a maximum number of subunit interfaces for a given volume and thus maximal stability.

In addition to improved stability, highly symmetric assemblies can benefit from avidity-enhanced receptor or ligand binding due to the multivalent display of the interacting domains on the surface of the complex.[12] Maintaining multiple copies of these domains spread symmetrically over the assembly surface greatly increases the statistical likelihood of a productive collision between them and their binding partners.

Despite the prevalence of supramolecular assemblies, the atomic level driving forces behind subunit association, and thus structure formation, are not completely understood. No generally applicable patterns of amino acid composition or secondary structural characteristics have been found which allow for consistently reliable prediction of the existence and function of protein-protein interfaces.[13] However, it has been noted that the majority of binding free energy is usually contributed by a minority of the residues at the binding interface.[14] Studying these hot spots, as they have come to be called, has greatly improved our understanding of how and why supramolecular complexes assemble. Further, these studies show promise for developing methods of predicting the existence

and function of interface associations.

1.1.2 The Structure of Protein Interfaces

Oligomerization interfaces are defined by their size, chemical composition, and surface complementarity.[15] The majority of protein-protein interfaces are composed of a single patch with a buried surface area of $1200 - 2000$ Å^2.[16] However, interfaces consisting of multiple patches with a total area of 7000 Å^2 have been reported.[17] Interfaces smaller than 1200 Å^2 typically correspond to transient associations. An area of $500 - 600$ Å^2 is generally held as the minimum surface area necessary to exclude solvent molecules from the binding interface.[14] Interestingly, the amount of free energy per unit surface area is largest for smaller interfaces, approximately 13 cal mol^{-1} Å^{-2}, and decreases linearly with total buried surface area until leveling off at a minimum of 4 cal mol^{-1} Å^{-2} at 2000 Å^2. This correlation suggests that smaller interfaces have a higher density of hot spots.[18]

Hot spots are regions of protein-protein interfaces containing residues which contribute a significant amount of energy to the free energy of binding of the interface. More specifically, these residues were originally identified as those whose mutation to alanine increased the change in free energy of binding by 2.0 kcal/mol or more.[14] Bulk solvent must be excluded from these regions, thus they are typically found clustered at the center of an interface. However, since most interfaces are flat, it would be difficult to achieve this sequestration without some method of screening solvent molecules at the perimeter of the interface. Studies have shown that the partially solvent exposed residues at the edges of an interface generally only make minor contributions to the free energy of binding.[14,19] It has been theorized that the primary function of these residues is to prevent

solvent molecules from disrupting interactions at the center of the interface. This theory has come to be known as the O-ring theory.

If the O-ring theory is true, then the residues found in hot spots should be significantly hydrophobic or capable of engaging in hydrogen bonding or salt bridge formation. The strength of these associations are enhanced by the decreased dielectric which results from solvent exclusion.[20] Analysis of the amino acid composition of hot spots has shown that they are highly enriched in tryptophan (trp), arginine (arg), and tyrosine (tyr) while also being deficient in leucine (leu), methionine (met), serine (ser), threonine (thr), and valine (val).[14,15] It has been posited that the reasons for these propensities are related to the ability of the amino acid side chains to engage in various types of stabilizing interactions as well as the entropic penalty paid for immobilization of the side chains at an interface.

The three most common amino acids found at hot spots (trp, arg, tyr) are all capable of engaging in hydrogen bonding and hydrophobic interactions (the arg side chain contains three methylene groups). The two aromatic side chains are also able to engage in π stacking interactions. Arg can interact with the aromatic side chains via cation $-\pi$ interactions, in addition to the five hydrogen bonds that it is capable of forming. The least abundant amino acids can primarily only engage in a single type of interaction via their side chain, either hydrophobic associations (leu, met, val) or hydrogen bonding (ser, thr). Further, immobilization of these side chains at the interface surface is likely to be highly entropically unfavorable. Being more rigid due to their aromaticity, the trp and tyr side chains are likely to be less disfavored. In addition, the extensive interactions available to the trp, tyr and arg side chains are probably more than enough to energetically offset the

entropic penalties of their immobilization. Interestingly, viral capsids have been shown to have fewer aromatic residues at their interfaces while having more uncharged polar residues.[21] This lack of aromatic residues and enrichment in polar ones may be due to the metastable nature of such capsids which requires them to be stable enough to withstand environmental challenges but not so strong that they cannot dissociate within their target cells.

1.1.3 Driving Forces Behind Interface Association

Protein-protein interactions are dominated by noncovalent associations such as hydrophobic interactions, hydrogen bonding, and salt bridge formation.[22] In particular, the burial of hydrophobic surface area is usually a major contributor to interface stability. This contribution is entropically driven due to the release of the ordered layer of water which surrounds the hydrophobic surface when it is solvent exposed. Van der waals contacts between hydrophobic residues of the adjacent subunits help to further stabilize the interaction and promote tighter packing of the interface. The degree of buried hydrophobic surface area can vary significantly, particularly between transient and obligate interfaces, but seems to reach an average of approximately 60%.[17,18] This average does not correlate with the size of the interface. It has been speculated that this average is an evolutionary constraint which prevents nonspecific aggregation of proteins. In addition to hydrophobic associations, polar interactions also play a significant role in the strength, stability, and specificity of protein-protein interfaces.

Hydrogen bonds are found in both polar and nonpolar regions of proteins due to the fact that the protein backbone is capable of engaging in this type of bonding. However, an

average of 76% of hydrogen bonds at an interface are between side chains. Individually, they are fairly weak bonds but what they lack in strength they make up for in numbers. The total number of hydrogen bonds is significantly associated with the interface's buried accessible surface area with an average of 1 hydrogen bond per 200 Å^2.[22,23] This ratio increases to 1 bond per 75 Å^2 when only polar area is considered.[17] It has been estimated that as few as 10 hydrogen bonds in a buried surface of 1500 Å^2 could result in a picomolar range K_D for the interaction.[24] Both inter- and intramolecular hydrogen bonds have been found to increase the strength and stability of protein-protein interactions. This enhancement is thought to be due to an increase in the enthalpy of binding that helps offset the entropic penalty resulting from immobilization of the interface.

Salt bridges are another polar interaction commonly associated with interface stability and specificity.[25] A significant pairing preference between oppositely charged side chains has been detected for residues at both homooligomeric and heterooligomeric interfaces.[26] An average of two salt bridges per interface has been estimated although the actual number observed can vary significantly.[25] A weak direct correlation between interface size and the number of salt bridge networks has been detected in some data sets. Despite their polar nature, salt bridges are often buried at protein-protein interfaces. This burial has likely evolved as a way to both strengthen the interaction by lowering the neighboring dielectric and prevent disruption of the salt bridge due to solvation of the interacting side chains.

These forces, hydrophobic and polar interactions, work together to maintain protein-protein interfaces. The specific nature of the association will dictate which of these forces is most prevalent at the binding site. Research has demonstrated that transient

associations tend to rely more upon polar interactions while obligate interfaces, which are more permanent in nature, typically better resemble a protein core with a significant degree of buried hydrophobic residues.[27,28] Regardless of the duration of the association, hydrogen bonding, salt bridge formation, and topological complementarity have been found to play a key role in the specificity of the interface.

1.2 Protein Capsids

1.2.1 Icosahedral Capsid Architecture

A capsid is a hollow, regular structure formed by multiple self-assembling protein subunits. Capsid assembly proceeds via a complex and systematic pathway in which monomers assemble into higher order oligomers that then come together to create the closed capsid. This process is highly concentration dependent and exhibits sigmoidal association kinetics in which assembly is slow until a critical concentration of building blocks is reached.[29, 192] At this point, production of whole capsids proceeds very quickly and finishes with few or no intermediate assembly states left over. These results strongly indicate that capsid assembly begins with a nucleation site, sometimes facilitated by binding of the monomers to scaffold proteins or, in the case of viruses, genomic material. The most frequently encountered form of the protein capsid is an icosahedron. These polygons account for the structure of approximately half of all viral capsids as well as a significant number of bacterial capsids.

The geometry of an icosahedron is that of an approximately spherical, enclosed structure with 20 triangular facets and 12 vertices on its exterior surface. Each facet may be subdivided into three smaller subunits, each with the same neighboring contacts. Thus,

perfect symmetry would require that icosahedral capsids assembled from homologous subunits must be built from 60 of them.[21] However, many icosahedral capsids are known to exist which are comprised of significantly more subunits. This observation gave rise to the theory of quasi-equivalence which states that increments of 60 identical subunits may be tiled across an icosahedral surface by arranging them into 12 pentamers and varying amounts of hexamers.[30] The number of hexamers increases with the number of subunits. Although not identical, the contacts made by subunits within the pentamers and hexamers are very similar, hence the term quasi-equivalence. The resulting structures can be described by their icosahedral triangulation number.

The triangulation number (T) is calculated from the following equation:

$$T = h^2 + hk + k^2 \qquad \text{(Equation 1.1)}$$

where h and k are positive integers corresponding to the number of steps along each axis of a two-dimensional plane required to move from one pentamer across the intervening hexamers to the next nearest pentamer. The total number of monomers in a capsid will be equal to 60T while the total number of hexamers is equal to 10(T-1). Since an icosahedron possesses 12 vertices and the pentamers are present only at the vertices, there are always exactly 12 pentamers in an icosahedral capsid. Heterooligomer capsids which obey the geometric constraints of an icosahedron despite having different monomer subunits are described with a pseudo-T number, abbreviated as P.

1.2.2 Viral Capsids

The most common examples of protein capsids found in Nature are viral capsids. The main function of these capsids is to condense and protect the virus' genetic material

as it is transported between host cells. Additionally, capsid proteins often play key roles in loading of the genome, specific cell targeting, and unloading of the capsid contents into the host cell. Thus proper assembly of the capsid is crucial to the viral lifecycle. These assemblies are usually highly symmetric with icosahedral capsids being among the most commonly found shapes.

The icosahedral viral capsids vary in size from $17 - 400$ nm primarily depending upon the size of the genome they must encapsulate.[31,32] The necessity of loading nucleotides has led virtually all of these capsids to have some degree of positive charge on their interior surface. In the case of RNA viruses, the capsid proteins often possess an N-terminal arginine rich motif that greatly enhances binding of nucleic acids.[33] Viruses have also evolved a functionalized exterior surface which often displays receptor binding proteins used to gain entry to highly specific target cells. The potential to exploit this built-in multivalent display system or to load the capsid with a variety of guest molecules makes viral capsids attractive protein engineering scaffolds. Among the most widely used viruses for nanotechnology applications are the cowpea chlorotic mottle virus (CCM)[34] and bacteriophages MS2,[35] Qß,[36] and P22[37].

CCMV is a plant virus which forms a 180 subunit $T = 3$ icosahedral capsid of approximately 30 nm diameter. It has been widely studied for use in nanotechnology applications due to its good overall stability and its inherent, reversible pH assembly switch. The capsid disassembles into 90 dimers at a pH of 7 and can be made to reform by lowering the pH to 5 (see section 1.3.1).[38]

Both of the bacteriophages MS2 and Qß form 180 subunit $T = 3$ icosahedral capsids from dimer building blocks, much like CCMV. They are also similar in size with diameters

of approximately 30 nm. Both have been shown to be thermostable and tolerant of mutations, making them good candidates for engineering.

The bacteriophage P22 capsid is composed of 420 monomers arranged as a T = 7 icosahedral capsid. This 58 nm capsid is referred to as the procapsid (PC). It undergoes expansion to the 64 nm mature capsid (EX) in response to DNA loading or heating at 65 °C for 10 minutes. Continued heating at 75 °C for at least 20 minutes causes the EX capsid to specifically dissociate the pentamers located at each vertex, thereby opening 10 nm pores and adopting the wiffleball (WB) form. This virus was initially selected for development as a nanotechnology scaffold due to its fairly large size and therefore high potential loading capacity.

1.2.3 Bacterial Microcompartments

Among the largest nonviral capsids are the bacterial microcompartments (BMC). These organelles are typically 80 – 200 nm in diameter and are assembled from families of BMC-domain bearing proteins.[39] These shell proteins form homohexamers which then associate in two-dimensional layers. Homopentamers lacking BMC-domains sit at the vertices of a rough icosahedral capsid and serve to direct the folding of the homohexamer sheets to form the faces of this capsid. These complex capsids have been found to contain as many as 8 different shell proteins. BMCs encapsulate several different enzymes in pathways whose intermediates are capable of easily diffusing out of the cell and/or are toxic in the cytoplasm. The archetypal BMC is the carboxysome.

The enzymes encapsulated within the carboxysome, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and carbonic anhydrase (CA), catalyze the key steps in

the carbon fixation pathway. CA converts bicarbonate (HCO_3^-) to CO_2 which RuBisCO then adds to ribulose-1,5-bisphosphate (RuBP) to form two molecules of 3-phosphoglycerate. RuBisCO is an inefficient enzyme[40,41] and if it is not co-encapsulated with CA, the CO_2 produced by the latter diffuses out of the cell before RuBisCO can bind it[42,43]. Further, RuBisCO is competitively inhibited by O_2. The carboxysome shell is permeable to HCO_3^- but not CO_2 or O_2. Thus, substrate is retained and concentrated to help offset RuBisCO's inherent inefficiency while also preventing exposure to an otherwise ubiquitous inhibitor.

The natural existence of BMCs serves as an excellent proof of concept that encapsulation of enzymes within a protein shell can lead to enhancements in individual enzyme activity or flux through the pathway as a whole. Most examples of engineered protein nanoreactors have exhibited mild decreases in overall catalytic efficiency although small improvements have been seen in some individual parameters. While BMCs are currently far more complex than the engineered nanoreactors produced to date, they demonstrate what may one day be accomplished when we have a much more advanced understanding of the intricate nature of protein-based supramolecular assemblies.

1.2.4 Catalytic Bacterial Capsids

Protein capsids often serve to group together several different enzymes in a pathway either through encapsulation or by acting as a scaffold. In addition to structural and container functions, capsids can possess catalytic activity of their own. They have been found to play significant roles in reactions related to detoxification, storage, immune response, energy metabolism, and biosynthetic pathways, among others.

Ferritin is a 24 subunit octahedral protein capsid found in all forms of life. In bacteria, it exists as a homooligomer while in higher order organisms, it can often be found as a heterooligomer composed of two separate chains.[44] It plays a critical role in iron metabolism where it serves to detoxify excess iron by converting Fe^{2+} to Fe_2O_3. This mineral is then stored in the capsid interior and released when iron levels are low. Sequestration of iron via this method has also been implicated to play a role in host immune response to bacterial pathogens.[45] Thus, this capsid employs catalysis of iron mineralization to maintain iron homeostasis and protect the cell from damaging free radicals and foreign organisms.

The pyruvate dehydrogenase complex (PDC) is composed of three separate enzymes which, together, catalyze the decarboxylation of pyruvate to form acetyl-CoA. This conversion links glycolysis and the citric acid cycle. The second enzyme in this series, dihydrolipoamide acetyltransferase (E2), forms a 24 subunit octahedral capsid in prokaryotes and Gram – negative bacteria[46] or a 60 subunit T = 1 dodecahedral capsid in eukaryotes and Gram – positive bacteria.[47] The first (pyruvate dehydrogenase, E1) and third (dihydrolipoamide dehydrogenase, E3) enzymes in the reaction cascade bind to the exterior surface of both capsid forms. The prokaryotic E2 capsid binds E1 and E3 via a peripheral subunit binding domain (PSBD) which is part of the E2 structure. The eukaryotic E2 capsid also employs a PSBD to bind E1 but uses a catalytically inert E3 binding protein to attach E3 to the capsid's outer surface. Thus, both forms of the PDC rely upon a capsid core to produce supramolecular complexes which enhance the overall flux through the reaction pathway.

Bacillus subtilis lumazine synthase (BsLS) and its associated riboflavin synthase

(BsRS) catalyze the penultimate and last steps in riboflavin synthesis, respectively.[48] Like E2 of the PDC, BsLS also forms a 60 subunit icosahedral capsid.[49] However, rather than decorating the outer surface, it encapsulates a trimeric form of BsRS, presumably through the use of an as yet unidentified recognition domain similar to the E2 PSBD. It is interesting that both the PDC and BsLS/BsRS complexes have evolved icosahedral capsids as part of their supramolecular assemblies but then functionalized different surfaces of the capsid. This difference is likely due to the unique evolutionary requirements imposed on each structure.

The BsLS active site is located on the interior surface of the capsid. This design allows for the accumulation of BsRS and its substrate in the capsid lumen. BsRS needs to function under conditions of low substrate concentration and is not subject to regulation by other enzymes. A high local concentration of substrate could allow the enzyme to function at overall substrate concentrations that would otherwise be too low for efficient catalysis to take place.[50] The rate of flux through the pathway also likely benefits from the colocalization of the enzymes catalyzing the last two steps via substrate shuttling from the BsLS active site to that of BsRS.

On the other hand, PDC activity is very tightly regulated by two enzymes which control phosphorylation of the E1 subunit.[51] These enzymes are bound to the complex via a lipoyl domain on the E2 subunit. Encapsulation of E1 would then require that these regulatory enzymes also be encapsulated. There may not be sufficient room within the capsid to allow for encapsulation of all of these proteins or, more likely, it was just simpler and more efficient to evolve external colocalization. Regardless of exactly why they evolved different approaches to supramolecular assembly, both LS and E2 serve as good

examples of the ways in which the various surfaces of protein capsids may be functionalized to enhance and control enzymatic activity.

1.3 Capsid Functionalization

1.3.1 Cargo Loading

One of the most common ways in which proteins produce supramolecular assemblies is via recognition domains encoded in the protein's primary structure.[52–55] These domains fold into specific secondary and tertiary structural elements which strongly associate with a complementary domain in the other binding partner. In some instances, binding occurs between a prosthetic group on one of the proteins and a binding site on the other protein. For example, the E1 kinase and dephosphorylase regulatory enzymes of the PDC bind to the lipoyl moieties on the E2 subunit. Capsids which naturally contain other proteins typically achieve encapsulation by using this recognition domain strategy.

Viruses often employ this strategy to bind their coat proteins to the scaffolding proteins which help the capsid assemble.[56,57] For example, of the 303 amino acids which comprise the P22 scaffold protein, only 15 near the C-terminus mediate binding to its capsid coat proteins during the initial nucleation and assembly stages.

Bacteria have also evolved capsids that employ this method of cargo loading. Encapsulins are $T = 1$ capsids which have been found to play a role in oxidative stress response by encapsulating a peroxidase and a ferritin-like protein. Loading of these enzymes is achieved via a 30 - 40 residue recognition domain located at their C-termini.[58] Unsurprisingly, higher order organisms have also evolved nanocontainers which function on the same principle.

Vaults are large, eukaryotic organelles whose function is not completely understood although they appear to be involved in pathogen defense. They have been found to encapsulate proteins expressing a 161 residue recognition domain, again, at the C-terminus.[59] Numerous studies have demonstrated that genetic fusion of these and other recognition domains to various proteins results in the efficient encapsulation of the chimera by the capsid specific to each recognition domain.

Similarly, an RNA aptamer-based approach was developed in which the Qß capsid protein mRNA was modified to encode a peptide binding sequence at one end and a packaging hairpin at the other.[60] Co-expression of the capsid proteins with enzymes fused to the targeting peptide resulted in encapsulation of those enzymes via association with the viral mRNA. Another, RNA-derived system employed MS2 mRNA covalently linked to a guest molecule to achieve capsid loading.[61]

Finally, two other encapsulation systems have been developed, both of which rely on charge complementarity. One is a derivative of the MS2 system described previously which takes advantage of the viral capsid's inherent positive charge to encapsulate guest molecules.[62] The other system uses an engineered bacterial capsid with a negatively charged interior surface (see section 1.5.2 for further information).[63] Both systems require tagging of the intended guest molecule with a short peptide ranging from 10 to 16 residues in length whose side chains possess a net charge opposite that of the interior of the intended container. These electrostatic encapsulation systems are less specific than the aforementioned recognition domain derived methods but they employ significantly smaller tags and offer a generally applicable means of achieving encapsulation.

Any capsid with a native net charge on its lumenal surface or capable of being engineered to possess such a charge should be able to encapsulate guests tagged with a complementarily charged peptide. As an alternative to tagging a guest molecule, supercharging of its surface has also been shown to produce efficient encapsulation.[64,65] However, supercharging is much more difficult to accomplish without disrupting the structure and function of the guest.

Finally, not all potential guest molecules are large. The ability to load capsids with small molecules is crucial to a number of their engineered functions, particularly as delivery vehicles. This loading is often achieved by covalently tethering the small molecule guests to specific residues on the interior surface of the capsid.[66,67] However, loading capacity is then limited by the number of available residues and the efficiency of the tethering reaction. Recent work with the heat shock protein G41C, Qß, and P22 capsids has shown that constrained growth of a polymer within the capsid lumen can greatly enhance loading of small molecules by providing a larger internal surface area and significantly more binding sites.[68–70]

While these methods have been very successful at achieving encapsulation, they do not provide an effective means of freeing the cargo molecules from the capsid, such as would be desired in a drug delivery vehicle or a nanoreactor whose products are too large to exit the capsid via its pores. Efforts to solve this problem have largely centered around developing capsid variants whose building blocks dissociate or assemble based on the solution conditions.

1.3.2 Capsid Assembly Switches

It has been shown that viral particles obtained by disassembling fully formed capsids can be induced to reassemble into capsids in the presence of a negatively charged polymer.[71] Presumably, this functions by mimicking the encapsulation of the virus' genomic material. However, these methods only provide control over assembly and reduce the available space for capsid loading. Nature has already done much of the heavy lifting to develop assembly switches in viral capsids. Owing to the need for virus' to deliver their genome into host cells, they are metastable supramolecular assemblies which can be disassembled by exposure to various environments within the cell. This disassembly is typically irreversible due to the release of encapsulated genetic material. However, several viral capsids have been found to disassemble and reassemble without their genomes *in vitro* via exposure to environmental conditions other than those usually seen *in vivo*. Few bacterial capsids have been found to possess inherent assembly switches, possibly due to the lack of an evolutionary need to disassemble.

1.3.2.1 Naturally Occurring Assembly Switches

The CCMV capsid is perhaps the best studied example of a reversible capsid assembly switch. It can be made to dissociate into dimers by raising the pH above 7 while simultaneously raising the ionic strength of the solution above 1 M.[38] These dimers can then be induced to reassemble by lowering the pH to approximately 5. Reassembly occurs at this pH for both high and low ionic strengths but at lower levels, capsid assembly is less efficient. The presence of DTT also decreases the reassembly efficiency, resulting in some partial assembly states.

The only reports of naturally occurring, reversible pH switches in bacterial capsids were observed in several DNA-binding proteins from starved cells (Dps)[72] and horse spleen apoferritin.[73] The Dps family of proteins possess ferroxidase activity and are members of the ferritin superfamily. They assemble into distorted icosahedral capsids which reversibly dissociate into dimers below pH 6.0 or above pH 7.5. This disassembly has been attributed to the disruption of salt bridges formed at the three-fold and two-fold interfaces.

Apoferritin is ferritin which does not possess bound iron atoms. Rather than assembling as an octahedral capsid, apoferritin is spherical. It was shown that this capsid partially dissociates to various degrees as the pH drops below 3 and that the spherical capsid can be reconstructed from some of these dissociation products, albeit with minor defects, upon return to physiological pH. Disassembly switches are far more common than reversible switches, particularly in viral capsids.

The T = 1 icosahedral foot and mouth disease virion (FMDV) capsid was found to dissociate around pH 5.[74] This disassembly is believed to be due to charge repulsion between a histidine residue upon protonation at lower pH and an α-helix dipole present at the two-fold pentamer-pentamer symmetry axis. In support of this hypothesis, mutation of the histidine to arginine destabilizes the capsid while mutation to aspartate resulted in improved acid stability. Another virus, hepatitis B (hepB), has been shown to possess two different disassembly switches.

A truncated form of the hepB capsid protein which is missing the last 34 amino acids of the C-terminus has been shown to assemble normally into a 240 subunit T = 4 capsid like the full length protein.[75] This capsid may be dissociated into dimers under nondenaturing conditions by exposure to 2 M urea at pH 9.5. Capsids can then be

reassembled by either lowering the pH to 7 in 250 mM ionic strength buffer or by increasing the ionic strength to 250 - 500 mM at pH 9.5. The presence of DTT significantly slows the reassembly process and decreases the total amount of capsid reformed. Despite the lower yield, slowing the reassembly process could result in higher loading of the capsid by guest molecules and merits further study.

The polyomavirus family possesses two members who display unique assembly switches in addition to the pH and ionic strength switches described above. These viruses form 360 subunit $T = 7$ capsids from 72 pentamers of a single protein (VP1) with the assistance of two other proteins (VP2, VP3) which interact with the interior of the capsid. VP1 alone is capable of assembling into capsids and other structures depending on the solution conditions.

Like hepB, the murine polyomavirus (MPV) and simian virus 40 (SV40) have been shown to possess an assembly switch in which capsid assembly can be induced via high ionic strength (> 1 M) or low pH (6.0 − 6.4).[76–78] MPV assembles into $T = 1$ icosahedral capsids under both conditions. SV40 also assembles into $T = 1$ capsids at high ionic strength but adopts a long tubular form at low pH. With SV40, these structures can also be produced by titrating VP1 pentamers with VP2 under various solution conditions.[79]

Metal induced assembly of MPV or SV40 can be achieved through the addition of 100 − 200 µM Ca^{2+} to the buffer solution at pH 7.2. Treatment with other divalent metal cations showed no assembly induction.[77,78] Further, this metal induced assembly was fully reversible by treatment with the calcium chelator EGTA under reducing conditions. Subsequent removal of the reducing agent and addition of calcium causes MPV to reform capsids although SV40 only partially reassembles.

1.3.2.2 Engineered Assembly Switches

Metal inducible capsid assembly has been successfully engineered into two separate protein capsids. Ferritin was modified by the mutation of four amino acids to histidines at its two-fold symmetry interface to produce two copper binding sites. Three more mutations were then introduced to the same interface in order to destabilize it sufficiently that the capsid did not assemble unless Cu^{2+} was present.[80] This switch was not reversible as removal of the copper ions via treatment with EDTA did not cause the capsid to disassemble.

The other capsid successfully engineered to possess a metal assembly switch was CCMV.[81] This switch was created by appending a six histidine tag to the N-terminus of the CCMV monomer which is located at the five-fold and quasi six-fold symmetry axis'. Addition of nickel to the capsid at pH 5 allowed it to remain assembled at pH 7.5 where the wild type capsid dissociates into dimers. Treatment of the mutant capsid with EDTA at this pH caused the capsid to disassemble, indicating that this switch is bidirectional.

CCMV is also the only capsid so far for which a thermal assembly switch has been developed.[82] This switch was developed by genetically replacing the first 26 amino acids of the CCMV monomer with a hexahistidine tag (for easy purification) and a series of nine elastin-like polypeptide (ELP) repeats. ELP polymers have been shown to contract in response to heating.[83] Assembly of the CCMV-ELP mutant via the native CCMV pH switch or by heating briefly in high ionic strength buffer resulted in the formation of an ELP shell inside the CCMV capsid. Assembly by pH manipulation resulted in the standard $T = 3$ CCMV architecture while heating caused the dimers to come together in a $T = 1$ capsid.

To date, two bacterial capsids have been engineered with pH switches. The *Bacillus stearothermophilus* PDC E2 subunit forms a T = 1 capsid in which the N-terminal arm stabilizes neighboring subunits. Removal of the first 50 amino acids of the N-terminus sufficiently reduced the overall capsid stability such that protonation of naturally occurring histidines at the three-fold symmetry axis caused dissociation of the truncated capsid between pH 6 – 6.5 at low ionic strength and pH 7 – 7.5 at high ionic strength.[84] This disassembly was not shown to be reversible.

The other capsid to receive an engineered pH switch, formed by *Aquifex aeolicus* lumazine synthase, is discussed in detail in section 1.5.2. Briefly, mutations which disrupt an ionic network at the two-fold symmetry axis and introduce a set of three histidines at the three-fold axis allowed for disassembly of the capsid by lowering the pH from 8 to 5.7 and subsequent reassembly by raising the pH in the presence of PEG-3350.[85]

Examination of the various natural and engineered assembly switches reveals significant utility for histidines as key features in both pH and metal induced assembly switches. This functionality is due to the moderate pK_a of its imidazole moiety and the fact that this group is capable of chelating several different metal ions. Its side chain pK_a of approximately 6.0 allows for modulation of the residue's charge state at physiologically relevant pH ranges. These facts suggest that histidine containing sequences may be useful as a general way of engineering capsid assembly switches for both *in vitro* and *in vivo* applications.

1.3.3 Functionalization of the Capsid Exterior

The exterior surface of protein capsids have been extensively modified in nature. Given the symmetric character of most capsids, this modification often leads to the uniform, multivalent display of proteins and prosthetic groups. The primary benefit of such display is enhancement of binding to targets through avidity effects. Just as having many weak interactions at an individual binding site may lead to a very strong overall interaction, association of many identical binding sites between complexes greatly increases the effective affinity, or avidity, of these associations as a whole.

For example, many viruses display receptor binding proteins on the outside of the capsid as a way of targeting and entering specific cells. Studies have shown that increasing the number of receptor binding proteins on the capsid surface results in more efficient uptake of the virus by the cell.[86] Similarly, DNA viruses also use the outer capsid surface to gain entry to the cell nucleus.[87] Proteins recruited to the exterior of the viral capsid have been found to play critical roles in the formation of viral envelopes via lipid recruitment or in the induction of apoptosis in host cells.[88,89] In at least one case, a virus has evolved a triple layer of concentric capsid shells for which the inner layers presumably recruit the outer layers.[90]

Bacterial capsids also exhibit functionalization of their outer surface. Since these capsids do not need to exit the cell, their modifications usually address the capsid to various regions of the cell interior or promote supramolecular assembly around the capsid.[91,92] There is a general paucity of information regarding the exterior functionalization of bacterial capsids.

Inspired by nature, researchers have developed a variety of different methods of enhancing capsid utility via modification of the exterior surface. Covalent modification of different amino acid side chains has been demonstrated to allow for the uniform, multivalent display of proteins as well as several different prosthetic groups including biotin, chromophores, haptens, and PEG.[93,94]

In particular, PEG modification of the capsid surface shows significant promise in enhancing the ability to use protein capsids in industrial processes by improving the solubility of the capsid in various organic solvents while simultaneously imparting greater thermal stability.[93] It is also an effective method of preventing antibody recognition of the capsid which is a major problem associated with the use of viral capsids as drug delivery vectors.[95–98] The unbound end of the PEG chain may also be modified to serve as a binding site for further additions to the capsid exterior.

In some cases, permanent attachment of the guest to the capsid exterior may not be desirable. The use of genetic modifications such as hexahistidine tags and coiled − coil motifs allows for dissociable interactions between the capsid and a guest molecule.[99,100] These tags have been appended to the solvent exposed N- or C-terminus of different capsids to allow for the binding of heme groups or gold nanoparticles and should be broadly applicable to other binding interactions. This approach was also used to create a multilayered immobilized network of bound capsids. While this technique is a versatile and fairly straight-forward method of producing protein interaction sites, it is incompatible with capsids whose termini are not exposed at the surface to be modified or those which make extensive subunit contacts as these are typically important stabilizing associations. Insertion of the genetic modifications in flexible loop regions on the capsid surface has

seen some success, but this route carries a significant risk of destabilizing the capsid.[101] A method of noncovalently decorating the outer surface of the capsid without introducing any genetic modifications has been developed.

In a similar approach to that which has had much success in encapsulating guest molecules, a whole protein or domain capable of binding to the exterior surface of the capsid is fused to a guest molecule.[102] This fusion allows for display of the guest on the outside of the capsid without the need to alter the capsid itself. Many of these capsid binding proteins exist for several different capsids. Alternatively, the combination of this method with one or more of the previously described methods of decorating the capsid exterior could allow for the simultaneous multivalent display of several different guest molecules on the capsid surface.

Lastly, the work presented in Chapter 3 demonstrates a novel means of decorating the outer surface of a capsid by functionalizing the pores at the five-fold symmetry axis. Charge complementarity was employed as a means of binding a tagged guest molecule with the tag presumably threading through the pore.

1.4 Nanotechnology Applications of Protein Capsids

1.4.1 Medical Uses for Capsids

A significant problem associated with many medical treatments, particularly cancer treatments, is efficient delivery of the therapeutic or diagnostic agent to target tissues. Many current treatments are nonspecific and rely on large doses to achieve delivery of functional drug levels to the intended sites of action. This shotgun approach often harms otherwise healthy tissues and requires the use of significantly more of the therapeutic agent

than a well targeted delivery, which exacerbates the side effects and raises the cost of treatment.

Protein capsids have shown great promise for use as highly addressable drug and bio-imaging delivery vectors.[103–107] Loading of small molecules is often accomplished via covalant conjugation to specific amino acids on the interior surface of the capsid while attachment of cell specific receptors to the exterior results in highly efficient uptake in the target tissues. In general, encapsulation by a variety of molecules, including lipids and proteins, has been shown to reduce toxicity and enhance uptake of various drugs.[108–110]

In addition, sequestration of drug or bio-imaging compounds inside capsids has been shown to extend the half-life of these molecules *in vivo* to days or even as long as several months.[111] This longer half-life opens up the possibility of engineering slow release treatments which allow for sustained dosing or long-term metabolic studies without the need for multiple injections.

Capsids have also been used to create improved vaccines by enhancing their stability or through multivalent display of antigens on the exterior capsid surface.[112–114] Avidity effects improve binding of humoral and cellular immune components to the antigen which results in strong activation of B cells and T lymphocytes, ultimately producing high levels of specific antibodies. Viral capsid-based vaccines are currently showing significant success as a means of activating the immune system against HIV and members of the plasmodium genus of malaria causing parasites.[115,116]

1.4.2 Materials Synthesis

Production of monodisperse nanoparticles and polymers is a challenge that protein capsids have been used to address. Their uniform size and hollow interiors make capsids ideal as templates or nucleation sites for controlled biomineralization. Several different capsids, including CCMV and Qß, have been used as reaction vessels to control a variety of mineralization reactions.[117–119] However, owing to their inherent activities as ferroxidases and iron storage sites, ferritins have been the most extensively developed for this function.

The ferritin family of proteins are generally able to bind a wide variety of metals. This property has been exploited to produce several different metal nanoparticles, bio-imaging contrast agents, semiconductors and metal catalysts.[120] Ferritin's metal binding capabilities have also led to it being investigated as a component in biofuel cells as an electron transfer mediator.

In this capacity, ferritin picks up electrons from the enzyme diaphorase and transfers them to a polypyrrole polymer which then delivers the electrons to the fuel cell anode. Incorporation of ferritin as part of this reaction cascade resulted in an enhanced electron transfer efficiency which demonstrated its potential for use in future biofuel cell designs.[121]

As mentioned at the end of section 1.3.1, capsids have also been used to regulate the size of several different polymerization reaction products. Although, to date, these reactions have only been used to enhance small molecule loading of the capsids, it is possible that this method of polymer growth inside a nanocontainer could be used to produce small, monodisperse architectures for use in other nanotechnology applications.

1.4.3 Nanoreactors

The use of enzymes to replace expensive inorganic catalysts and harsh reaction conditions in industrial processes is highly desirable and has been the subject of intense research efforts. However, naked enzymes suffer from a number of problems which have impeded their wide-scale industrial use, including relatively short half-lives, thermal instability, intolerance to organic solvents, and poor scale-up efficiency. It has been demonstrated that protein capsids may be modified to withstand these inactivating conditions, thus encapsulation has the potential to confer these advantages to the guest enzymes. In addition, the observation that there exist naturally occurring nanoreactors which exhibit enhanced activity in comparison to the free enzymes has led to the speculation that encapsulation could improve reaction kinetics and thereby address the scale-up efficiency problem as well. Although promising, the science of nanoreactors is still very much in its infancy.

The majority of early capsid engineering work focused on drug delivery applications. The first engineered protein nanoreactor was reported in 2007. It was demonstrated that CCMV could be used to encapsulate an active horseradish peroxidase and that substrate molecules could diffuse into the capsid interior.[122] Less than six months later, SV40 was used to encapsulate cytosine deaminase and deliver it, in active form, to CV-1 cells.[123] Although this construct fits the definition of a nanoreactor, the focus of the experiments were still on drug delivery applications.

No further nanoreactor systems were reported until late 2010 when the Qß capsid was used to encapsulate two different enzymes with the specific goal of examining the effects of encapsulation upon enzyme activity.[60] From 2011 until the present, there has

been an average of two new nanoreactors reported per year with the majority being developed by the Douglas lab using the P22 capsid[62,124–130]. Although all of these recent studies have examined the encapsulated enzyme kinetics of their respective nanoreactors, the exact factors which influence these kinetics is still not well understood.

For a detailed discussion of encapsulated enzyme kinetics, please see section 2.3. Briefly, enzyme encapsulation seems to have either no significant effect upon the confinement k_{cat} ($k_{cat, conf}$) or decreases it by less than an order of magnitude. The confinement K_m ($K_{m, conf}$) is also typically unchanged or slightly decreased, leading to an overall confinement turnover number ($k_{cat, conf}/K_{m, conf}$) that is roughly an order of magnitude lower than that of the free enzyme. However, three of the reported nanoreactors exhibited significantly increased $K_{m, conf}$ values. Our own nanoreactor, as reported in Chapter 2, also exhibited aberrant behavior with the single largest decrease in $k_{cat, conf}$ of any reported system to date. These outliers demonstrate that, while significant progress has been made in the study of nanoreactors, there are still vast swathes of uncharted territory to explore in this field.

1.5 *Aquifex aeolicus* Lumazine Synthase as a Protein Engineering Scaffold

1.5.1 The AaLS Capsid

Aquifex aeolicus lumazine synthase (AaLS) is a 60 subunit T = 1 icosahedral capsid. The individual monomers are 154 amino acids long with a molecular weight of 16.7 kDa, yielding a capsid with a molecular weight of approximately 1 MDa. The capsid is best viewed as a dodecamer of pentamers due to the extensive monomer-monomer interactions within each pentamer (Fig. 1.1a). The wild-type capsid is approximately 16

nm in diameter with a shell thickness of around 3.5 nm, giving an internal diameter of roughly 9 nm. The shell is solid with the exception of moderately sized pores at each of its five-fold symmetry axes.[131]

AaLS catalyzes the penultimate step in riboflavin synthesis and the five-fold axial pores are likely portals for entry of substrate into and egress of products from the interior of the capsid, where the active sites are located. The pores are formed by the third α-helix of each monomer of the pentamer (residues 90-108). These helices possess five turns, are 27 Å long, and give the pore an hourglass shape with the narrowest point located approximately halfway through the pore at lysine 98 (Fig 1.1b). At this point, the pore is only 13.7 Å across, as measured between Cα atoms in the backbones of parallel residues. Previous work on a pentameric lumazine synthase from *Saccharomyces cerevisiae* (ScLS) has demonstrated that mutation of the solvent exposed residues lining the pore does not necessarily disrupt the association of the pentamer.[132] Due to the high sequence homology between ScLS and AaLS, the AaLS pores were expected to be equally robust.

Indeed, since *A. aeolicus* is a hyperthermophile, the entire AaLS capsid is extremely stable with a melting temperature of 120 °C. It possesses a buried surface area of 555,145 Å^2, of which 53.3% is hydrophobic. Such extensive sequestration of hydrophobic surface is a hallmark of highly stable protein – protein interfaces. Further, the total exterior surface area of the capsid is 276,102 Å^2, of which 60.3% is charged. This significant degree of exposed surface charge is indicative of a high number of ion pairs and, unsurprisingly, AaLS was found to have more ion pairs (17) than any other icosahedral lumazine synthase. In particular, a six amino acid ionic network was found to stabilize the pentamer-pentamer interface across three interacting monomers. The extraordinary stability and tolerance of
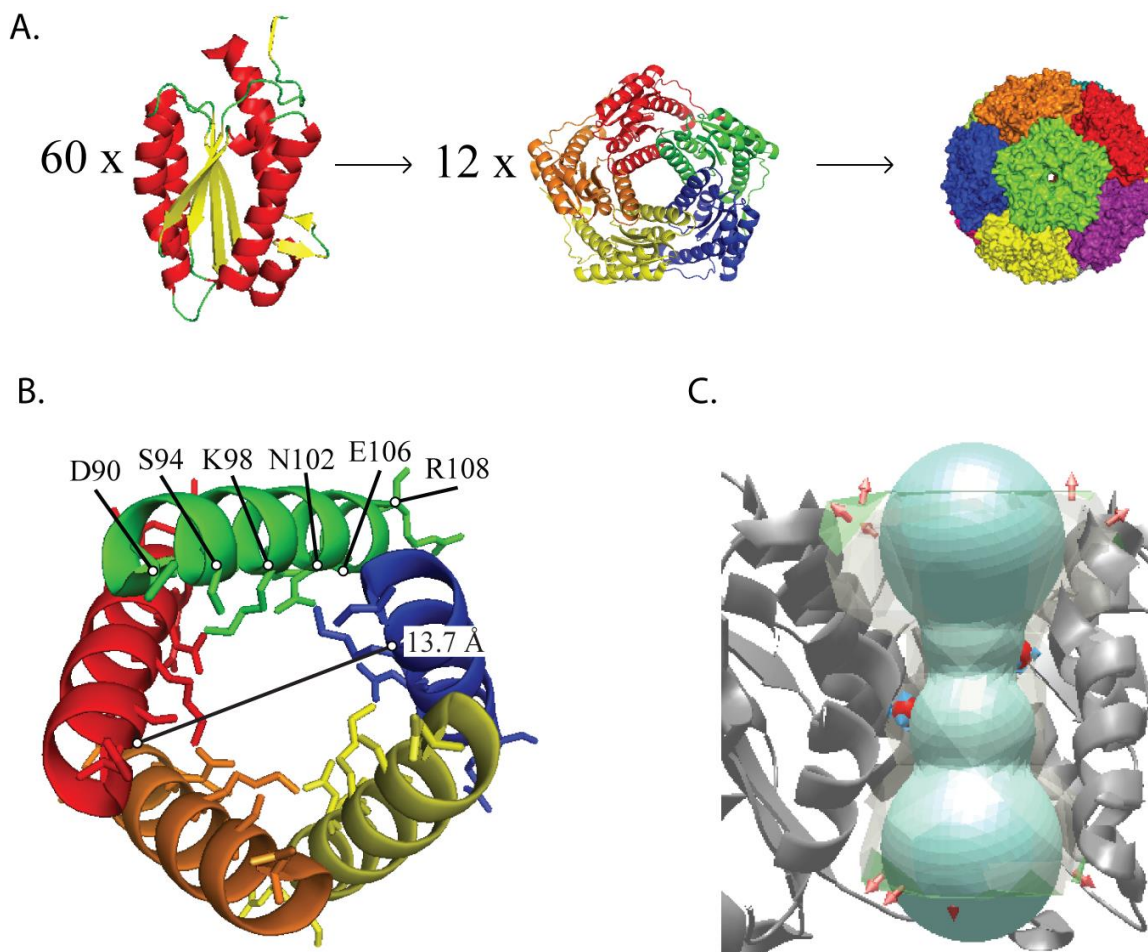
Figure 1.1 The AaLS capsid. A) Surface representation of the AaLS capsid (right, PDB ID: 1HQK). Pentameric (middle) and monomeric (left) building blocks are shown as ribbon diagrams extracted from the whole capsid structure. B) The five-fold axial pore viewed from the interior capsid surface. All labelled residues protrude into the lumen of the pore, except aspartate 108 which is located on the exterior surface of the pore. C) A volumetric representation of the pore lumen, viewed from the side, which demonstrates its slight hourglass shape. Image made with Mole 2.0.

the AaLS capsid to mutation makes it an ideal scaffold for use in a variety of protein engineering applications.

1.5.2 Engineered AaLS Variants

AaLS has been the subject of several studies designed to explore the potential of the capsid to be loaded with a variety of different guest molecules. The first of these studies introduced four mutations to the interior surface of the capsid (R83E, T86E, T120E, Q123E) designed to impart the lumenal surface with a significant negative charge. These mutations also caused the assembly state to shift from a 60 monomer T = 1 capsid to a 180 monomer T = 3 capsid with a diameter of approximately 29 nm. This variant was named AaLS-neg.[63] The mutant capsid was produced in order to demonstrate that appending a positively charged deca-arginine tag (R_{10}) to a guest molecule (GFP, in this case) and co-expressing both the capsid and the tagged guest in *E. coli* cells results in significant encapsulation of the tagged protein.

Subsequent work involved expanding on the idea of a charge complementarity based encapsulation system. The R_{10} tag was appended to HIV protease which was then co-expressed with AaLS-neg[133]. This protease is toxic to the cell when free in the cytoplasm, thus its encapsulation provided a significant growth advantage over cells expressing the protease alone or co-expressed with the wild-type capsid. The researchers were able to use this growth advantage in a directed evolution experiment to select for capsids demonstrating an enhanced ability to encapsulate the tagged protease. The most effective of the variants that they found, termed AaLS-13, showed a 5- to 10-fold greater encapsulation efficiency over AaLS-neg.

The AaLS-13 capsid possesses the following mutations, in addition to those found in AaLS-neg: D28G, R52C, T112S, V115D, A118D, R127C, and K131E. Only the latter two mutations alter residues which protrude into the lumen of the capsid. The other mutations occur at the various subunit interfaces. Overall, the mutations render each monomer significantly more negative which is likely the primary reason for the improved encapsulation efficiency observed with this capsid. EM imaging revealed a diameter of roughly 35 nm, indicating that the mutations also caused the capsid to expand to a 240 monomer $T = 4$ structure.

A follow-up experiment used positively supercharged GFP and AaLS-13 pentamers, which are commonly separated from fully assembled capsids during purification, to demonstrate the ability to control *in vitro* loading by varying the ratio of pentamers to guest molecules.[64] Interestingly, *in vitro* encapsulation could also be achieved using fully assembled capsids, suggesting the capsid is capable of at least partially disassembling to allow for guest loading. Further experimentation with a positively supercharged variant of ferritin resulted in encapsulation of between one and four ferritin capsids inside the AaLS-13 capsid, raising the possibility of creating a single nanoreactor with multiple internal reaction chambers.[65]

The Woycechowsky lab then designed a variant called AaLS-pos which possessed mutations designed to impart the interior surface of the capsid with a net positive charge (T86R, D90N, T120R, and E122R).[134] This capsid remained a 60 subunit $T = 1$ assembly like the wild-type. The purpose behind creating this mutant was to generate a nonviral protein capsid capable of binding cellular nucleic acids *in vivo* during protein production. It was shown that the purified AaLS-pos capsid was loaded with different RNA molecules

of various sizes up to 350 bases in length, suggesting a maximum size limit for loading of the capsid.

Finally, a method of halting capsid assembly at the pentamer stage and then introducing a redox switch to induce the pentamers to assemble into capsids was developed. It was found that disruption of the ionic network and hydrogen bonding sites at the two-fold symmetry axis (R40, H41) as well as a hydrophobic interaction site at the three-fold symmetry axis (I125) by introducing three mutations (R40E, H41E, and I125E) caused assembly to stop after the formation of pentamers.[135]

Based on these results, a variant comprised of a similar set of mutations (C37A, R40S, H41S, I125C) was made (AaLS-switch-red). The C37A and I125C mutations were introduced to remove an endogenous cysteine and introduce one at the three-fold symmetry axis, respectively. The pentameric AaLS-switch-red was treated with DTNB to form an NTB adduct at cysteine 125, thereby priming the sidechain for a thiol disulfide exchange reaction in which thiophenol replaced the NTB moiety. After formation of the AaLS-switch-thiophenol adduct (Switch-Ox), the modified pentamers were incubated with 10% PEG – 3350 for two days to induce capsid assembly.[85]

To further expand upon the usefulness of the AaLS capsid, a pH assembly switch was designed by once again introducing mutations at the two-fold (R40S, H41S) and three-fold (T120H, E122H, Q123H) symmetry axes to produce the AaLS-switch-pH variant[85]. The histidines at the three-fold axis induce dissociation of the capsid via charge repulsion upon protonation of the histidine side chain. At pH 8 AaLS-switch-pH exists primarily as a capsid of comparable size and morphology to the wild-type AaLS capsid. As the pH is decreased, the capsid disassembles. However, at pH 6.2 in phosphate buffer,

a capsid peak can still be observed in the SEC chromatogram and lowering the pH further results in protein precipitation. Dialysis of the capsid into citrate buffer at pH 5.7 or 6.2 yields only pentamers. Subsequent dialysis of these pentamers back into pH 8 phosphate buffer does not produce reassembled capsid. However, incubation of those pentamers with 10 % PEG-3350 for four days results in an 86% yield of reformed capsid.

AaLS is an extremely stable, highly manipulable protein scaffold. Its several engineered assembly switches and capsid variants possessing multiple different morphologies and charge states make it highly attractive for further development as a nanotechnology tool.

1.6 Dissertation Aims

The overarching goal of the work presented in this dissertation was to add to the body of knowledge regarding supramolecular assembly of protein molecules. AaLS was selected as a nonviral capsid scaffold due to its high stability and tolerance to mutations. The novel complexes described in the following chapters demonstrate the versatility and potential of protein capsids for nanotechnology applications.

Chapter 2 discusses the effects of encapsulation on enzyme kinetics. A previously engineered AaLS capsid variant was used to encapsulate an esterase *in vivo* using a charge complementarity-based tagging system. The various factors at play when encapsulating enzymes and the effects they have on reaction kinetics are not well understood. Here I report an unusual finding which I speculate may be a result of the electrostatic environment within the capsid, a potential influence which has been previously unconsidered when explaining the observed effects of encapsulation upon enzyme kinetics.

In Chapter 3, charge complementarity is once again employed, only this time it is used to decorate the outer surface of the capsid. Many capsids possess pores which span the width of the capsid shell. These pores often occur at symmetry axes. The pore surface has largely been ignored as a means of functionalizing the capsid. I demonstrate that capsid pores can be modified via mutagenesis to provide a simple and remarkably strong binding site to allow for the multivalent display of a guest molecule on the capsid's outer surface. These results have implications for the use of capsids as drug delivery vehicles and nanoreactors.

Finally, in Chapter 4, I build on previous work done in this lab which sought to engineer a redox-based switch for control over capsid assembly. The effects of the size and shape of various adducts on the ability of the capsid to assemble are explored and the results demonstrate the significant plasticity of the three-fold symmetric interface. Further, one adduct was found to disassemble completely upon lowering the pH, providing a combined redox-pH switch. These approaches show potential for enhancing control over capsid assembly which is a highly desirable function to have for any nanotechnology application of protein capsids.

Overall, this work has added new tools to the nanotechnology toolbox by highlighting the utility of the pore surface as a supramolecular assembly site and expanding understanding of redox and pH switchable control over capsid assembly. In addition, my work suggests that future exploration of the effects of the electrostatic environment within the capsid on encapsulated enzyme kinetics could significantly enhance our understanding of how and why encapsulation can influence protein interactions.

CHAPTER 2

EFFECTS OF ENCAPSULATION UPON THE KINETIC PARAMETERS

OF A CARBOXYLESTERASE

2.1 Introduction

Compartmentalization or encapsulation as a means of protecting molecules and controlling reactions is a strategy successfully employed by nature on several levels. At the microscale, cell organelles have evolved to protect critical components such as DNA (nucleus), to isolate otherwise dangerous molecules such as proteases (lysozome), and to precisely control reaction cascades (mitochondria). Natural encapsulation methods have been extremely effective at the nanoscale as well.

Viral capsids are perhaps the most well-known example of nanoscale compartmentalization. These capsids are significantly functionalized to protect and deliver their genomic material, often to highly specific targets. The fact that all viruses are comprised of some form of capsid is a strong indicator of how effective encapsulation can be at protecting and controlling reactive molecules. Other, nonviral, capsids have also proven useful to nature.

Nanoreactors such as the carboxysome or some fatty acid synthases have evolved to house the multiple enzymes necessary to carry out their functions.[136,137] By sequestering together these enzymes, the nanoreactors greatly enhance the overall rate of reaction of the cascade.[42,138] In the bacterium *Bacillus subtilis,* a nanoreactor has evolved in which lumazine synthase, the enzyme responsible for the penultimate step of riboflavin synthesis, encapsulates the next enzyme in the pathway, riboflavin synthase.[139,140] The constrained proximity of the sequentially active enzymes allows for rapid progression through the last two steps of the synthetic pathway.[50]

Given how successful encapsulation strategies have been in nature, it is no surprise that scientists from a wide variety of technical backgrounds have spent decades researching

how to improve upon existing capsids and one day hope to be able to generate them *de novo.* Protein capsids have shown significant promise for use as drug-delivery vehicles,[141-145] contrast agents for bioimaging,[111,146–149] biomineralization templates,[113,150-154] and synthetic nanoreactors.[60,62,65,122–124,126–130,133,155,156]

Aquifex aeolicus is an extremophilic bacterium first identified from samples collected near submerged volcanic vents around Sicily. As a thermophile, its proteins have evolved to maintain stability and function at high temperatures.[157] The lumazine synthase complex isolated from this organism (AaLS) exists as a 60 subunit protein capsid which has shown itself to be singularly appealing as an engineering scaffold due to its inherently high stability and corresponding tolerance toward mutation.[131]

The AaLS capsid forms as a dodecamer of pentamers that possesses pores at the five-fold symmetry axes which span the width of the capsid shell. These pores are thought to allow influx and efflux of small molecules through the capsid. This is evinced by the observation that the active site is located on the interior surface of the capsid. Therefore, it should be possible to engineer the capsid to selectively encapsulate a guest enzyme and to measure the difference in activity between the free and encapsulated forms. Indeed, previous work has shown that the capsid can be altered through directed evolution to achieve successful encapsulation of either deca-arginine tagged green fluorescent protein (GFP) or an identically tagged HIV protease through charge complementarity-based interactions.[63,64,133] The most effective of the evolved capsids, termed AaLS-13, possessed mutations imparting it with a significantly negatively charged interior surface.

Encapsulated GFP was seen to fluoresce at similar levels to the unencapsulated protein, suggesting that native folding is not disrupted by this method of encapsulation,

thus it was believed that the protease would be active within the capsid. However, no protease activity was observed in the encapsulated protease samples. This apparent inactivity may be due to inhibition of the protease by the capsid, but a reexamination of the properties of both proteins revealed a more likely culprit.

The protease is most active at a pH around 4, but the capsid precipitates at these pH levels; therefore, the experiments were carried out at pH 8, where the capsid is most stable. Unfortunately, HIV protease is only minimally active at this pH. Thus, although encapsulation was successfully achieved, HIV protease was not a suitable enzyme for measuring the effects of encapsulation on enzymatic activity.

In this chapter, we seek to demonstrate that an enzyme which is active at the optimum pH for capsid stability (pH 8) does, in fact, remain active upon encapsulation. Additionally, the kinetic parameters of the free and encapsulated enzymes will be examined to determine what effect, if any, sequestration within the capsid has upon these values.

The guest enzyme we chose, which exhibits maximum activity at pH 8, is a carboxylesterase isolated from another thermophilic bacterium, *Geobacillus stearothermophilus* (Est55, PDBID: 2OGT).[158,159] It is a promiscuous enzyme that acts on a variety of different ester substrates, including 4-nitrophenyl acetate (PNPA). This substrate was specifically chosen because it is small enough to enter the capsid and one of the cleavage products, nitrophenolate, absorbs strongly at 405 nm. This allows for colorimetric assessment of the reaction kinetics.

2.2. Results

2.2.1. Design of an Esterase Encapsulation System

Based on an existing deca-arginine tag (R_{10}) method of achieving encapsulation,[63] an R_{10} tag was appended to the C-terminus of Est55 to produce a mutant esterase (Est55-R_{10}) which was then coproduced in the presence of AaLS-13. It was expected that Est55-R_{10} would be electrostatically attracted to the assembling capsid *in vivo* which would result in the formation of a complex and encapsulation of the esterase (Figure 2.1). The complex should then exhibit a much higher degree of esterase activity than the capsid alone.

2.2.2. Encapsulation of the Esterase by the Capsid

Est55-R_{10} and AaLS-13 were over-produced in *E. coli* cells both individually and together. After purification, all proteins were analyzed via SEC. The activity of the eluted fractions obtained from the purification of AaLS-13 alone and coproduced with Est55-R_{10} was examined as an initial assessment of whether or not encapsulation had occurred. There is significantly more esterase activity in the capsid fractions obtained from the coproduced sample than in those from the capsid alone (Figure 2.2), indicating an association between the capsid and the tagged esterase.

SDS-PAGE was performed to confirm the presence of both proteins in the coproduced sample (Figure 2.3). A band corresponding to the calculated molecular weight of the tagged esterase (56.7 kDa) was observed at the expected position. The AaLS-13 capsid monomer (17.7 kDa) is known to migrate anomalously and is typically observed at a distance corresponding to a protein of approximately 18.5 kDa, which is where the largest and darkest band of the sample was observed. The bands assigned as AaLS-13 and
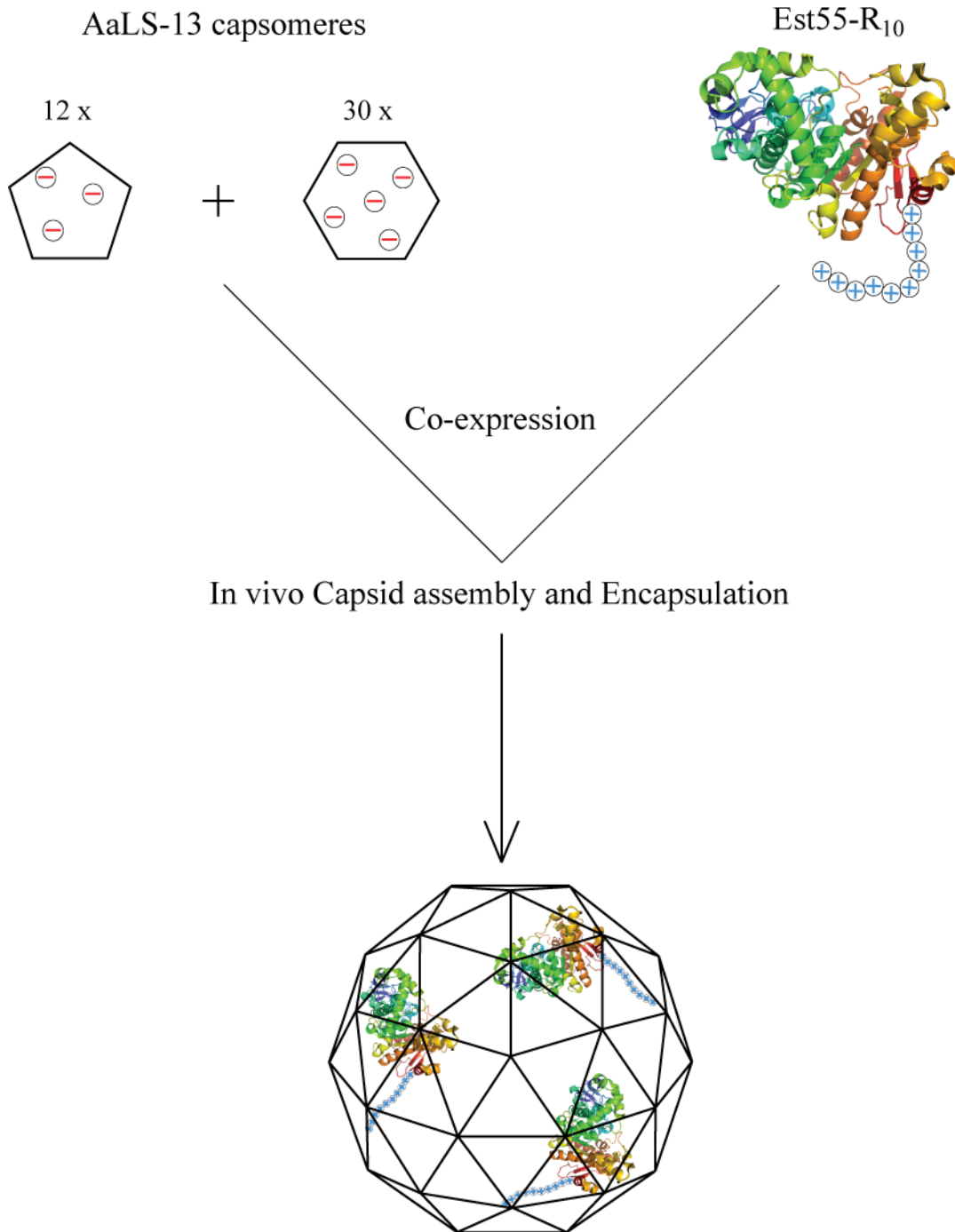
Figure 2.1. Schematic representation of the encapsulation of Est55-R$_{10}$ by AaLS-13. The positively charged deca-arginine tag is electrostatically attracted to the negatively charged interior surface of the capsid during capsid assembly, resulting in encapsulation of the tagged esterase.
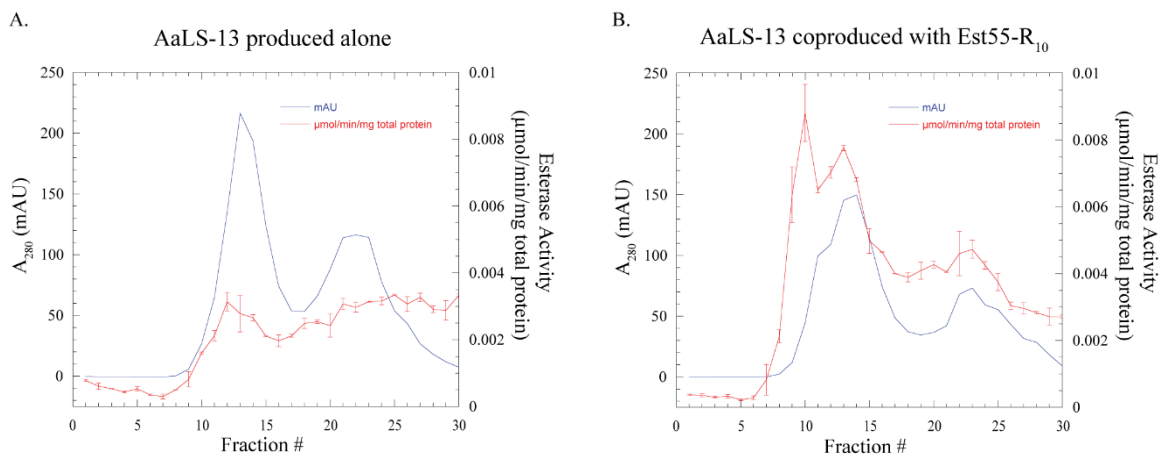
Figure 2.2. UV absorbance and esterase activity of SEC fractions from AaLS-13 capsid produced alone (A) or coproduced with Est55-R$_{10}$ (B). Fractions were obtained via SEC using a Sephacryl S-400 matrix. The capsid peak was defined as fractions 12-15.
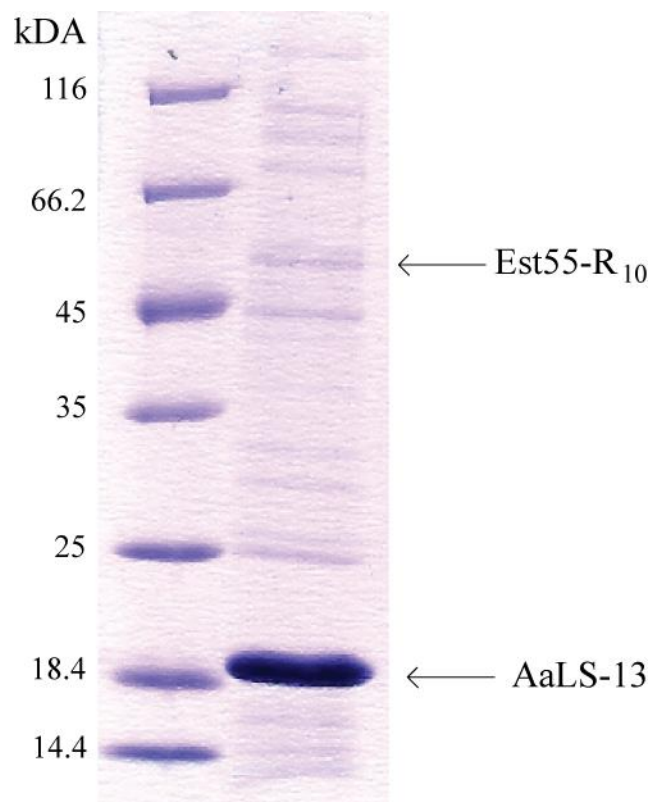
Figure 2.3. SDS-PAGE analysis of esterase encapsulation. After purification, a sample of AaLS-13 that had been coproduced with Est55-R$_{10}$ was analyzed by SDS-PAGE. The indicated bands were excised and identified via mass spectrometry.

Est55-R_{10} were excised and their identities confirmed by mass spectrometry. Several other bands were also excised and mass spectrometry analysis showed these were similar to the endogenous *E.coli* proteins previously shown to be encapsulated.[133]

2.2.3. Quantification of the Encapsulation of Est55-R_{10}

To quantify the amount of Est55-R_{10} and AaLS-13 present in the purified complex, the bands corresponding to these proteins were excised from the SDS-PAGE gel. After excision, protein was extracted from the gel and digested via protease treatment. The resulting peptide fragments were doped with a known quantity of two different ^{13}C labeled peptide standards, one whose sequence matched the capsid and one which matched the esterase. The amount of protein present in a given band was estimated by comparing the size of the unlabeled peptide peak to its labeled counter-part (Figure 2.4). This calculation yielded an average of 2.1 esterases encapsulated per capsid. From this value, we calculated a mass ratio of Est55-R_{10} : Total protein of 0.0284. Total protein is defined as the sum of the masses of one AaLS-13 capsid (240 monomers) and 2.1 Est55-R_{10}'s. With this information, it became possible to calculate the exact amount of Est55-R_{10} present in each encapsulation sample and thus to differentiate between the specific activity of the sample as a whole (esterase and capsid) and that of the encapsulated esterase alone.

2.2.4. Specific Activity of the Free and Encapsulated Esterase

The specific activities of untagged Est55 and Est55-R_{10} produced alone and in combination with AaLS-WT or AaLS-13 are shown in Table 2.1. The individual esterases have similar activities, although appending a deca-arginine tag to the C-terminus of Est55

A. AaLS-13 fragment (GLANLSLELR)

Unlabelled peptide fragment ^{13}C - labeled standard
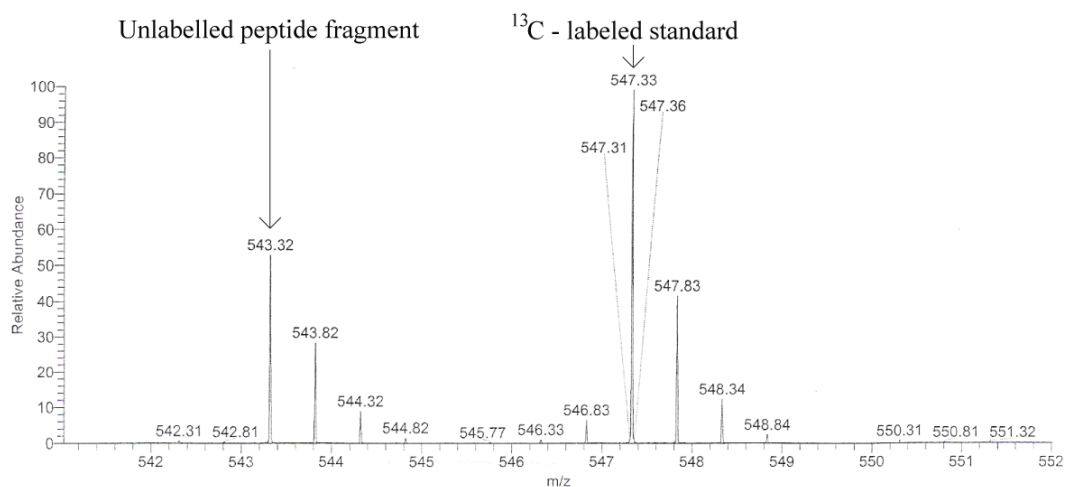
B. Est55-R$_{10}$ fragment (AMLQSGSGSLLLR)

Unlabelled peptide fragment ^{13}C - labeled standard
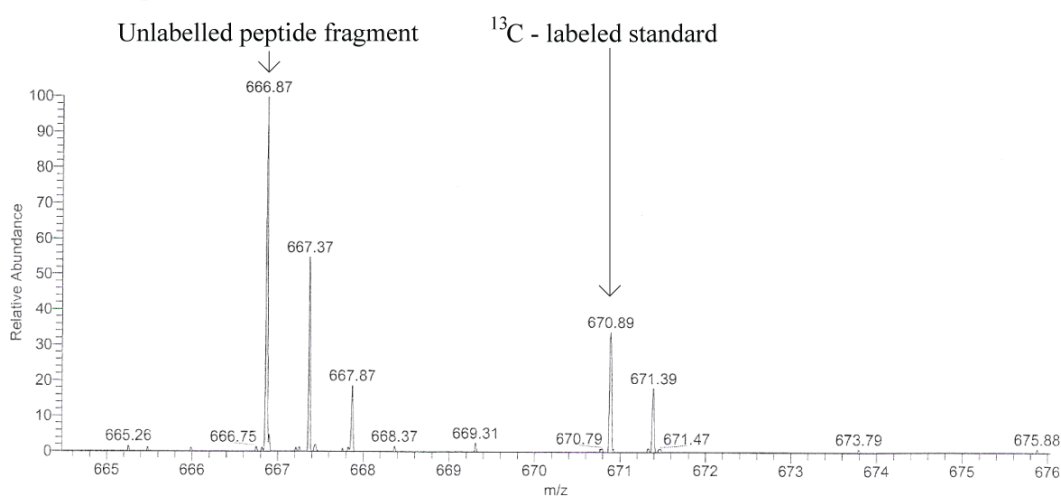
Figure 2.4 LC-MS/MS spectra of the excised SDS-PAGE bands. A known quantity of the ^{13}C – labeled standard was doped into each sample. The mass peaks of unlabeled peptide fragments resulting from proteolytic digest of AaLS-13 (a) or Est55-R$_{10}$ (b) as well as their respective ^{13}C labeled counter-parts are indicated by arrows. The amount of AaLS-13 and Est55-R$_{10}$ in the encapsulation complex was estimated by comparing the peak areas of the indicated unlabeled peptide fragments to the ^{13}C labeled standard.

Table 2.1 Specific activities of various combinations of capsid and esterase.

Protein(s)	Specific Activity (µmol/min/mg total protein)	Specific Activity[a,b] (µmol/min/mg esterase)
Est55	325 ± 13	325 ± 13
Est55-R_{10}	192 ± 35	192 ± 35
AaLS-WT	0.0045 ± 0.00024	n/a
AaLS-13	0.0040 ± 0.0019	n/a
Est55 + AaLS-13	0.053 ± 0.0003	n.d.
Est55-R_{10} + AaLS-WT	0.076 ± 0.047	n.d.
Est55-R_{10} + AaLS-13	0.29 ± 0.06	10 ± 2

[a]n/a = not applicable [b]n.d. = not determined

appears to decrease the specific activity of the resulting mutant by approximately 40%. This decrease in activity may be due to mechanical screening by the R_{10} tag as the C-terminus is located near the active site. Presumably, the tag is free to move around in solution in front of the active site and may deflect substrate prior to binding. Both the wild-type and mutant capsids show virtually no esterase activity by themselves.

Coproduction of Est55 with AaLS-13 or of Est55-R_{10} with AaLS-WT results in a pooled capsid peak with a specific activity that is approximately 10-fold higher than that of the capsid peaks from either capsid when they are produced alone, suggesting that some degree of encapsulation or association naturally occurs between these proteins. The specific activity of the capsid peak obtained from the coproduction of Est55-R_{10} with AaLS-13, however, is 100-fold higher than that of the capsid alone, indicating a much higher degree of association. This increased activity is likely due to charge complementarity between the cationic tag and the highly negatively charged interior of AaLS-13. The specific activity of the encapsulated esterase was determined to be approximately 20-fold lower than the unencapsulated form.

2.2.5. Michaelis-Menten Kinetics

In order to determine how encapsulation affects the kinetic parameters of the esterase, Michaelis-Menten kinetics were measured in order to determine what affect, if any, encapsulation had upon the activity of the esterase (Figure 2.5). The estimated confinement k_{cat} ($k_{cat, conf}$) for encapsulated Est55-R_{10} is approximately 30-fold lower than that of free Est55-R_{10} which suggests a diminished ability for each molecule of encapsulated enzyme to stabilize the transition state. The confinement K_m ($K_{m,conf}$) of the
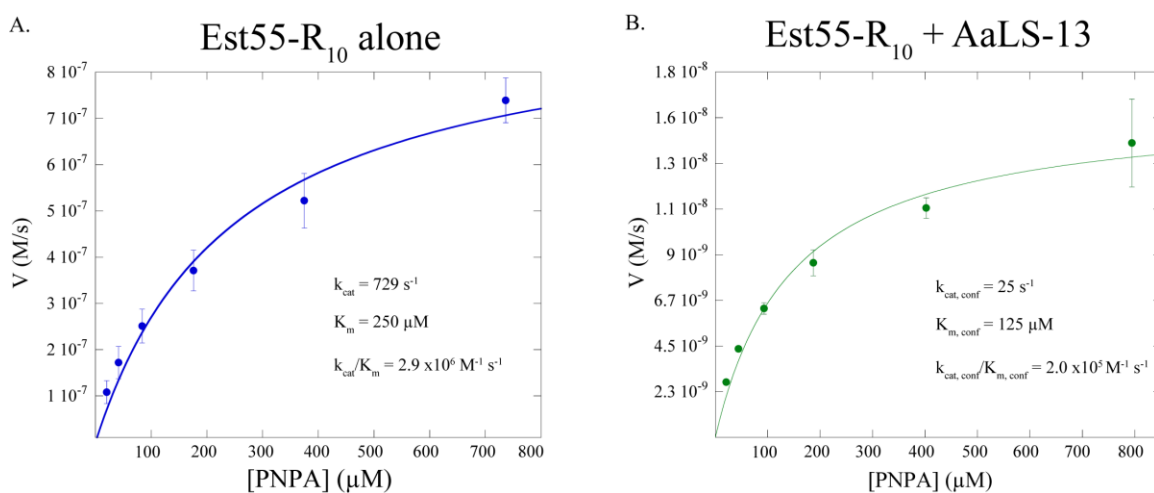
Figure 2.5. Michaelis-Menten kinetics of the free (A) and encapsulated (B) esterase. The initial velocity of PNPA hydrolysis was measured spectrophotometrically in buffer at 25 °C. Reactions were quenched by the addition of concentrated NaOH.

encapsulated esterase was also lower in comparison to the free esterase, with the encapsulated $K_{m, conf}$ being approximately half that of the free esterase. In terms of the turnover number or catalytic efficiency of the enzyme, the significant decrease observed in the k_{cat} of the encapsulated esterase is partially counterbalanced by a concurrent drop in the $K_{m, conf}$. This offset resulted in an encapsulated Est55-R_{10} efficiency that was only a single order of magnitude lower than that of the free enzyme.

2.3 Discussion

Encapsulation of enzymes within a proteinaceous shell allows for control over substrate access, increases the concentration of pathway intermediates, and can perturb the kinetic parameters of the confined enzymes. In this study, we found that the use of a charge complementarity-based loading method, previously developed to encapsulate an R_{10} tagged HIV protease in the same capsid, resulted in an average *in vivo* loading capacity of 2.1 Est55-R_{10} molecules per AaLS-13 capsid. The loading capacity of HIV protease-R_{10} was approximately seven-fold higher with an average of seven HIV protease-R_{10} dimers per capsid.[133] There are several factors that could potentially contribute to the elevated protease loading, including size, inherent pI, quaternary structure differences, and production levels between the two enzymes.

Est55-R_{10} exists as a 510 residue, approximately 57 kDa monomer with an estimated pI of 6.8 while HIV protease-R_{10} is a dimer composed of two separate 99 residue, 12 kDa monomers with an estimated $pI_{monomer}$ of 11.1. Thus, we can see that the esterase is significantly larger than the protease dimer and is also much more negatively charged. The sheer size difference between the two different enzymes may have contributed to the lower

loading of the esterase. Additionally, charge repulsion between the negatively charged lumenal capsid surface and the esterase surface may have decreased loading of the esterase. Charge neutralization between the same capsid surface and the protease may have enhanced protease loading. Avidity effects are likely to have played a significant role in the greater protease loading as the tagged dimer has 20 positive charges in close proximity while the tagged esterase has only 10 such charges.

Encapsulation of the esterase decreased $K_{m, conf}$ and $k_{cat, conf}$ by approximately 2-fold and 30-fold, respectively, with the overall catalytic efficiency ($k_{cat, conf}/K_{m, conf}$) decreasing by an order of magnitude. These effects are similar to those observed in native and other engineered encapsulation systems.

In nature, the lumazine synthase from *Bacillus subtilis* (BsLS) encapsulates a trimer of its own riboflavin synthase (BsRS) and thereby confines the last two steps in the riboflavin synthesis pathway together.[140] When examined individually, encapsulation actually increases the K_m values of both BsLS and BsRS for their respective substrates while also reducing their maximum velocities. However, a comparison of the rate of riboflavin production of the complex versus a mixture of the BsLS capsid and unencapsulated BsRS trimer indicates the complex produces riboflavin nearly twice as fast.[50] This rate enhancement is thought to be due to substrate channeling as result of the close proximity of the BsLS and BsRS active sites. Thus, even though their individual kinetic parameters appear to be functioning less efficiently, encapsulation of BsRS by BsLS allows for a significant enhancement to the overall rate of the pathway at the lower substrate concentrations found within the cell. A similar result is seen in the larger protein shells known as bacterial microcompartments (BMC). These 80-200 nm heterologous

associations of capsid subunits often encapsulate several enzymes from a single metabolic pathway.[136,160,161]

The carboxysome is one of the most extensively studied BMCs. It contains carbonic anhydrase (CA) and ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), the key enzymes in the carbon fixation pathway. Confinement of RuBisCO within the carboxysome does not significantly alter its reaction kinetics, presumably due to its inherent poor efficiency, but it does prevent competitive inhibition by O_2 via substrate gating since the capsid does not allow O_2 into the lumen.[42,162,163] CA, on the other hand, is approximately 3-fold faster in the cytosol than in the carboxysome. Experiments in which CA was produced in the cytosol significantly reduced RuBisCO activity because the CO_2 produced outside the capsid left the cell at a much faster rate than it diffused into the carboxysome.[138] Thus, the only way to effectively pair CA and RuBisCo is to co-encapsulate them, despite the reduction in activity suffered by CA.

In an effort to understand the effects of encapsulation upon enzyme kinetics, several novel encapsulation systems have been developed. All of these systems employ viral capsids,[60,62,123,130,155] except for one which used a bacterial capsid.[133] Several different loading methods have been used to achieve encapsulation of a guest protein, including covalent (genetic or chemical fusion to a naturally encapsulated molecule) and noncovalent (electrostatic) techniques. Regardless of the means used to load the capsid, the specific identity of the capsid or the guest enzyme, encapsulation of a single species had remarkably similar effects on the kinetic parameters of the enzyme.

For all of the experiments in which kinetic parameters were determined, encapsulation of the guest enzyme decreased k_{cat}, generally by less than an order of

magnitude. A few encapsulated enzymes showed no significant change in activity compared to their free forms. The $K_{m, conf}$, however, was typically unchanged or very slightly decreased in relation to the free enzyme. These decreases may have been the result of an accumulation of substrate within the capsid (M_{conf}) relative to the substrate concentration in the bulk solvent. Overall, encapsulation typically decreased the turnover number (k_{cat}/K_m) by approximately an order of magnitude.

Experiments regarding the effects of crowding agents on enzyme activity in solution have shown that crowding agents typically cause a significant decrease in K_m and have varying effects on k_{cat}, depending on the enzyme being studied, with the turnover number increasing significantly in most cases.[164] These results generally agree well with the observed effects of encapsulation in both naturally occurring and engineered nanoreactors. However, we noted three exceptions in which $K_{m, conf}$ was observed to increase as a result of engineered encapsulation.

In one of these cases, the increase was attributed to over-crowding.[60] A nanoreactor designed to use the bacteriophage Qß capsid to encapsulate firefly luciferase resulted in an increased $K_{m, conf}$ for both the substrate luciferin and the cofactor ATP. This increase scaled directly with the number of enzymes loaded per capsid.

In the second case, another bacteriophage capsid, P22, was used to encapsulate an alcohol dehydrogenase.[130] The researchers studied the effects of encapsulation for all three capsid forms and noted that, while the $K_{m, conf}$ of the substrate was decreased for all three forms, the $K_{m, conf}$ of the cofactor, NADH, was increased under certain conditions. No definite cause was determined for the results but crowding was suggested as a potential reason.

The third case of an increase in $K_{m, conf}$ was a result of experiments designed to determine the effects of the electrostatic nature of capsid pores on the ability of small molecules to enter and exit the capsid.[124] Another bacteriophage, MS2, was engineered to encapsulate an alkaline phosphatase (PhoA) by genetically appending a negative tag to the enzyme, much like the R_{10} tag we added to Est55. When the pores were engineered to have the same charge as the substrate, an almost 6-fold increase in $K_{m, conf}$ was observed with a concurrent 3-fold decrease in $k_{cat, conf}$. Pores with the opposite or neutral charge had little or no effect on the measured kinetics. These results demonstrate that the electrostatic nature of the pore can play a significant role in altering encapsulated enzyme kinetics.

Although it is known that the AaLS-13 capsid has a negatively charged interior surface,[133] no structure is available to elucidate the nature of its pores. For the encapsulation of Est55-R_{10}, it is possible that the pores are slowing the rate of substrate diffusion into the capsid, thereby lowering the $k_{cat, conf}$. If such substrate gating was responsible, we would expect to see a concurrent increase in $K_{m, conf}$ due to the decreased diffusion of the substrate into the capsid, as was previously observed with PhoA. However, both $K_{m, conf}$ and $k_{cat, conf}$ decreased in comparison to the unencapsulated parameters. This difference suggests that substrate gating is not playing a significant role in the kinetics of our nanoreactor system. Alternatively, we speculate that the electrostatic environment within the capsid lumen could alter the rate of release of product from the enzyme.

Electrostatic interactions have been shown to alter reaction rates through both substrate attraction[165,166] and intramolecular active site transfer in bifunctional enzymes or fusion proteins.[167] These effects are typically a rate enhancement achieved through favorable Coulombic interactions that promote substrate binding by directing the substrate

to the active site. Mutagenesis studies in which these types of interactions were abolished or rendered unfavorable via charge neutralization or reversal demonstrate the potential of electrostatic interactions to modulate protein activity.[168–170]

In the case of the AaLS-13 encapsulated esterase, it is plausible that the highly anionic nature of the capsid interior could electrostatically influence several different aspects of the catalytic mechanism such as inhibiting activation of the catalytic serine 194 or altering the stability of the transition state. It may also suppress release of the negatively charged acetate and nitrophenolate products from the esterase, yielding a lower V_{max} and thus a lower $k_{cat, conf}$. It is possible that a similar interaction is occurring in the nanoreactors engineered from viral capsids since they inherently possess highly positively charged lumenal surfaces to bind their genomes. Encapsulation systems which rely on fusion with the viral genome or scaffold proteins may neutralize some or all of the interior charges and thereby reduce or abolish their effects on the confined enzymes. In some cases, the ratio of genomic negative charge to capsid positive charged can exceed 1:1, meaning that use of viral genomic material as an encapsulation tag may even result in negative supercharging of the capsid interior.[171] These Coulombic interactions could potentially alter the $K_{m,conf}$ as well.

Consider two potential scenarios. The first scenario has a substrate and capsid interior that are complementarily charged while the second has both components possessed of the same charge. In the first scenario, the enzyme must compete with the capsid surface to bind substrate. This competition will likely cause an increase in $K_{m,conf}$. The second scenario may see a decrease in $K_{m,conf}$ because binding of the substrate, which is neutral overall but does possess inherent dipole moments, by the enzyme could reduce electrostatic

repulsion between the substrate and interior capsid surface.

Alternatively, if the substrate is capable of diffusing into the capsid at a faster rate than it diffuses out, then it is possible that the decreased $K_{m, conf}$ is due to an increase in the M_{conf} of the substrate resulting from the local accumulation of substrate molecules within the capsid. This increase in effective concentration also applies to the encapsulated esterase. Given our calculated loading capacity of 2.1 esterases per capsid, we estimate the M_{conf} of the esterase to be approximately 300 µM. Although we cannot calculate the substrate M_{conf}, the fact that the encapsulated esterase kinetics fit well to the Michaelis-Menten equation suggests it must be significantly higher than the M_{conf} of the esterase, so as not to violate the steady-state assumption.[172]

Further, these elevated concentrations may have led to a change in the intrinsic thermodynamic activity coefficients of the substrate and/or esterase. As the concentration of an electrolyte increases, its thermodynamic activity deviates from ideality. These deviations could greatly alter the kinetics of the reaction. In at least one case, the kinetic parameter alterations caused by a crowding agent were attributed to a change in activity coefficients.[173] Such changes could also be the result of alterations to the way the encapsulated enzyme or substrate are solvated in comparison to the bulk solvent.

There are fewer water molecules inside the AaLS-13 capsid than in the bulk solvent and, presumably, they are more highly ordered due to the significant negative charge of the lumenal surface.[174] Given that our encapsulation system relies upon charge neutralization, the tagged esterase is mostly likely very near the interior capsid surface and therefore may be surrounded by a higher degree of ordered water molecules than when it is floating free in the bulk water. This may significantly alter the way the enzyme is solvated which could,

in turn, alter its intrinsic activity.

Our work demonstrates that a charge complementarity-based encapsulation method does not necessarily abolish the function of the tagged guest enzyme. Further, we demonstrated that noncovalent encapsulation methods are a viable means of quickly and easily producing a novel nanoreactor system. This system has the potential for further functionalization through mutagenic alteration of unreactive interior capsid residues to those capable of easily undergoing covalent modification such as cysteine, lysine, or tyrosine. Such alteration potentially allows for the creation of heteroenzyme nanoreactors composed of both covalently and ionically loaded guest enzymes. Finally, this work serves to highlight the fact that electrostatic influences of the capsid upon the function of the encapsulated guest enzyme are poorly understood and merit further study.

2.4. Materials and Methods

2.4.1. Materials

All cell culture media and chemical reagents were purchased from Bio-rad, Fisher Scientific, Gold Biotechnology, or Pierce Biotechnology and used without further purification. *Pfu*-turbo DNA polymerase and *E. coli* cell strains BL21 (DE3) and XL1-Blue were purchased from Stratagene. T4 DNA ligase and all restriction endonucleases were purchased from New England Biolabs. The oligonucleotides used in this study were synthesized by the DNA/peptide synthesis Core Facility at the University of Utah.

2.4.2. Production and Purification of Crude Protein

All proteins were produced in $CaCl_2$-competent BL21 (DE3) *E. coli* cells. Transformation of the gene-bearing plasmids (pACYC-Est55, pACYC-Est55-R$_{10}$, pMG-AaLS-13) either individually or in combination was accomplished via heat shock at 42 °C for 2 minutes. The cells were then grown at 37 °C for 1 hour and plated on antibiotic-containing LB agar plates (50 µg/mL chloramphenicol (CAM) for Est55 or Est55-R$_{10}$, 100 µg/mL ampicillin (AMP) for AaLS-13, both antibiotics for *in vivo* coproduction). The plates were grown overnight in a 37 °C oven. A single colony was selected from the plate and used to inoculate a 7.5 mL antibiotic-containing LB culture which was then grown overnight at 37 °C with shaking at 250 rpm. An aliquot of this culture (1 mL) was then used to inoculate antibiotic-containing LB media (500 mL). Over-production of the desired protein was achieved by growing this large culture at 37 °C to an OD_{600} of 0.7, inducing production by adding IPTG to a final concentration of 0.2 mM, and then incubating the cultures at 30 °C for 20 hours at 250 rpm. Co-expressed cultures containing AaLS-13 and Est55-R$_{10}$ where incubated at 37 °C for 4 hours, postinduction. The cells were harvested via centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (6000 rpm for 15 minutes at 4 °C). Pellets were frozen at -80 °C until used.

Cell pellets were resuspended in 10 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). The cells were lysed by incubation with lysozyme (10 mg), RNase A (1.2 mg), and DNase I (20 µg) on ice for 1 hour followed by sonication at 45 µm amplitude for 3 minutes in 10 second on/off pulses using a Misonix ultrasonic liquid processor. The lysate was then clarified by centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (12000 rpm for 45 minutes at 4 °C).

2.4.3. Purification of AaLS Capsid

For purification of AaLS-WT, AaLS-13, or coproduction samples, the supernatant was loaded onto a Ni^{2+}-NTA resin which had been equilibrated with lysis buffer and incubated for 1 hour at 4 °C with gentle rocking. The column was then washed with lysis buffer, lysis buffer containing 40 mM imidazole, and lysis buffer containing 500 mM imidazole. The high imidazole fraction was then concentrated and either run on the FPLC immediately (AaLS-WT) or dialyzed into lysis buffer containing 5 mM EDTA overnight and then run on the FPLC (AaLS-13 and coproduction samples). The assembled capsid was isolated by loading the sample onto an ÄKTA FPLC system (GE Healthcare) equipped with a HiPrep 16/60 Sephacryl S-400HR column and run with lysis buffer containing 5 mM EDTA at 4 °C at a flow rate of 0.5 mL/min.

2.4.4. Purification of Est55 and Est55-R_{10}

For purification of Est55 or Est55-R_{10} produced in the absence of an AaLS variant, the clarified lysate was dialyzed into ion exchange buffer A (50 mM sodium phosphate, 20 mM NaCl, pH 8.0) overnight at 4 °C. The sample was then loaded onto the FPLC equipped with either an anion-exchange MonoQ 5/50 GL column (Est55) or a cation-exchange MonoS 5/50 GL column (Est55-R_{10}) which had been equilibrated with ion exchange buffer A. Bound protein was eluted using a 20 mL gradient ranging from 0% to 100% ion exchange buffer B (50 mM sodium phosphate, 1M NaCl, pH 8.0) run at 4 °C. SDS-PAGE with a Thermo Scientific Unstained Molecular Weight Marker was then used to determine which fractions contained the esterase. These fractions were pooled, concentrated, and then loaded onto a HiLoad 16/60 Superdex 75 PG column run with lysis

buffer at 4 °C. Est55R$_{10}$ production typically yielded 16 mg/L culture. Purity was assessed by SDS-PAGE.

2.4.5. Activity and Kinetics Assays

To determine the activity of the enzymes used in this study, the hydrolysis of p-nitrophenyl acetate to p-nitrophenolate and acetate was followed spectrophotometrically using a Hitachi U-3310 by observing the change in absorbance at 405 nm. Assays were conducted in triplicate using a final total protein concentration of 1.2 nM. Various concentrations of PNPA stock solutions were prepared in 95% ethanol such that the addition of 10 µL of PNPA stock and 4 µL of protein stock to 1986 µL of lysis buffer resulted in estimated final concentrations of 25, 50, 100, 200, 400, or 800 uM in the reaction cuvette. Actual substrate concentrations were determined by quenching the reaction with concentrated NaOH. After the addition of PNPA and protein, the reaction was monitored for 500 seconds with readings taken every 0.1 seconds. Using the slope of the linear portion of the resulting curve (Δabsorbance/s), the Beer-Lambert law, and the extinction coefficient of p-nitrophenol at 405 nm (17800 M^{-1}cm^{-1}), the reaction velocity (M/s) and specific activity (µmol/min/mg protein) were determined. Protein concentration was determined via Bradford assay.

2.4.6. Identification and Quantification of Encapsulated Est55-R$_{10}$

Direct detection of Est55-R$_{10}$ in coproduced samples was initially accomplished using SDS-PAGE. Samples showing the presence of both Est55-R$_{10}$ and AaLS-13 were then sent to the University of Utah proteomics core facility for further analysis by

LC-MS/MS. Samples were divided into 50 µL aliquots and treated with 5 µL 1% ProteaseMax (Promega) for 5 minutes before being vortexed and diluted with 15 µL of 50 mM ammonium bicarbonate (ABC). Finally, 30 µL of 20 ng/µL trypsin/lys-C solution in 50 mM ABC was added to the sample which was then thoroughly mixed and incubated for 5 hours prior to injection on an LTQ-FT (Thermoelectron) mass spectrometer.

Identification of esterase and capsid proteins was confirmed by comparison to sequences in the NCBI database and the previously determined sequence of the AaLS-13 capsid.[133] Quantification of both proteins was accomplished by doping samples with known amounts of ^{13}C-labeled peptide fragment standards (Table 2.2). The standards contained eight ^{13}C atoms each. The concentration of each standard was determined by amino acid analysis and by absorbance at 205 nm. Amino acid analysis was conducted by the Molecular Structure Facility at UC Davis.

Table 2.2. Sequences of ^{13}C labeled peptide fragments used for quantification of the capsid and esterase from coproduced samples. Molecular weights given for the +2 charge state.

Parent Protein	Peptide Fragment Sequence	Molecular Weight (Da)	Extinction Coefficient at 205 nm (M^{-1} cm^{-1})
Est55-R$_{10}$	AMLQSGSGSLLLR	1333.5	36940
AaLS-13	GLANLSLELR	1087.2	26770

CHAPTER 3

MULTIVALENT DISPLAY OF A PROTEIN ON THE EXTERIOR SURFACE OF A

BACTERIAL PROTEIN CAPSID VIA CHARGE COMPLEMENTARITY

3.1 Introduction

Nearly all proteins associate into multimeric assemblies in their active form.[6] The highly symmetric nature of many of these supramolecular complexes typically results in the multivalent display of one or more proteins on the outer surface of the complex.[12] When these proteins serve to bind other assemblies or large molecules via multiple interfaces, this display results in a significant increase in the effective affinity, or avidity, between them. Due to their potential to augment functions which require binding to a large molecule or surface, avidity enhancing effects are one of the most common benefits of supramolecular assembly.

Viral capsids are composed of many copies of one or more proteins, usually symmetrically arranged as hollow polyhedrons. This architecture imparts significant stability and protection to their genomes while affording an excellent platform for the multivalent display of receptor binding proteins. The avidity effects of this display greatly improve the virus' ability to target and enter specific cells,[175–177] although this can also be a problem for the virus as the same display improves the ability of antibodies to bind to its surface. Enhancement of avidity can also be beneficial to catalysis.

Cellulosomes are a collection of carbohydrate-active enzymes (CAZymes) produced by bacteria which digest plant cell wall polysaccharides.[178] Multiple copies of these enzymes are linked to a display platform, known as a scaffoldin, through docking proteins which display K_D values in the nanomolar range.[179] This assembly is significantly more efficient at degrading native cellulose than the monomeric CAZymes.[180] Researchers have already capitalized on the discovery of these very selective, high affinity complexes to engineer novel cell-surface display scaffolds, tagging methods for protein purification,

biosensors, and fusion proteins designed to impart new catalytic function to the complex.[181]

As our understanding of how to manipulate and alter the supramolecular assembly of proteins becomes more sophisticated, so too do the applications of this knowledge. Among recent advances, a protein scaffold of unknown native function was computationally converted into a potent inhibitor of hen egg lysozyme,[182] the *Methanoccocus maripaludis* group-II chaperonin capsid opening and closing reaction was redesigned from a chemically driven process to a photoreactive one through the covalent modification of engineered cysteine pairs by azobenzene-dimaleimide,[183] and a cytochrome (cb_{562}) bearing no relation to any known metallo-ß-lactamase was manipulated into a fully functional member of that enzyme class through the use of metal templated interface redesign.[184] In this chapter, we seek to further expand the variety of available supramolecular assembly techniques by functionalizing the pores of a bacterial protein capsid.

Aquifex aeolicus lumazine synthase (AaLS) is an extremely stable protein capsid making it attractive as a scaffold for supramolecular assembly. Previous work has generated mutant versions of the capsid with a variety of different properties including changes in assembly state, size, and interior surface charge.[63,133–135] A mutant with an overwhelmingly positive lumenal surface has been shown to encapsulate RNA via charge complementarity during *in vivo* assembly.[134] In another study, it was demonstrated that a capsid variant with a negatively charged interior surface could be loaded with positively charged guest molecules, also driven by charge neutralization.[63,64,133]

On the basis of this previous work, we propose a novel means of noncovalently decorating the outer surface of the capsid via charge complementarity between a

deca-arginine (R_{10}) tagged guest protein and a mutant AaLS capsid engineered to have negatively charged five-fold axial pores. This negative pore surface serves as a means of inducing the positively charged tail of the R_{10} tagged protein to thread through and bind to the pore, thus providing a multivalent presentation of the tagged protein on the outer surface of the capsid (Figure 3.1). We report here a proof-of-principle that capsid pores, which have been previously ignored as supramolecular assembly interfaces, may be functionalized through mutagenic manipulation of their electrostatic surface.

3.2. Results

3.2.1. Development of a Charge Complementarity Based Method
of Decorating the Outer Surface of the AaLS Capsid

The wild-type AaLS five-fold axial pore only possesses a net charge of -5 so we sought to mutate residues lining the pore to increase its negative character. The initial design of the negatively super-charged pore called for four of the five residues whose sidechains protrude into the pore to be mutated to aspartates (S94, K98, N102, E106) (Figure 1.1b). The first position (D90) is an aspartate in the wild-type sequence and therefore did not need to be mutated. A sixth residue, R108, was also mutated to aspartate, despite the fact that its sidechain does not enter the pore itself. This was done in the hopes that the residue, which protrudes into the bulk solvent on the outer surface of the pore, would enhance attraction of the R_{10} tag to the pore surface and facilitate threading of the tag through the pore. The capsid variant containing these five mutations was labeled AaLS-6D and possessed a negatively super-charged pore with an estimated -25 charge. In order to make any association between the tagged protein and the capsid easily detectable via fluorescence,
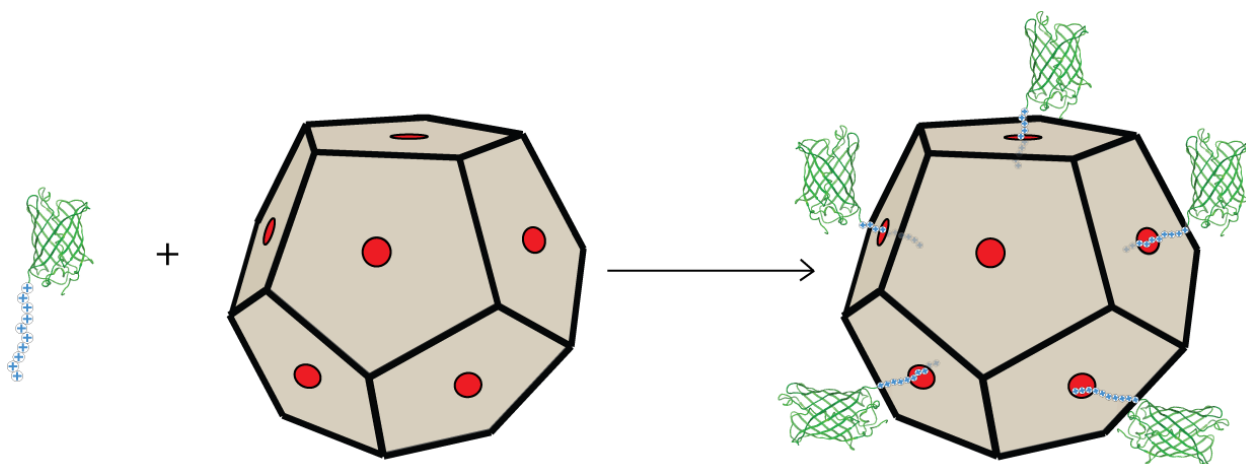
Figure 3.1. Schematic representation of the assembly of the tag-pore complex. Upon mixing, charge complementarity will allow the deca-arginine tag on GFP (blue pluses) to thread through the negatively supercharged pores (red circles) resulting in a uniform, multivalent display of GFP on the outer surface of the capsid. Center pore left empty for clarity.

GFP was selected for use as the R_{10} tagged protein (GFP-R_{10}).

3.2.2. Estimated Packing Density of the Tag-Pore Complex

Using the program Mole 2.0, which is specifically designed to measure the volume of protein cavities, and the crystal structure of the AaLS capsid, we estimated the volume bounded by the backbones of the pore helices to be approximately 3733 $Å^3$ (Figure 1.1c). Using previously reported amino acid volumes.[185] We estimated the volume of a single aspartate sidechain to be 59 $Å^3$. Given that there are five aspartates per helix and five helices per pore, a total of 25 aspartate sidechains protrude into each pore with a total volume of 1475 $Å^3$. If the R_{10} tag adopts an extended conformation upon threading through the pore, then only seven arginine residues will be required to span the pore. Again using previously reported values, we estimate the total volume of the sidechains and backbone to be approximately 1575 $Å^3$. Thus, the combined volume of the aspartate sidechains and the tag residues threading through the pore is 3050 $Å^3$ yielding a packing density of 0.80. This packing density is slightly above the maximum crystallographic packing density of 0.74 but not unreasonable given our simplified volume estimates and suggests excellent contact between the tag and pore.

3.2.3. Fluorescence Detectable in AaLS-6D Capsid Peak Fractions

Following Incubation with GFP-R_{10}

In order to test whether or not GFP-R_{10} associates with AaLS-6D differently from the wild-type capsid, the following mixtures were produced and analyzed: GFP + AaLS-6D, GFP-R_{10} + AaLS-WT, and GFP-R_{10} + AaLS-6D. Individual components

where purified in high ionic strength running buffer (525 mM), combined at a molar ratio of 72:1 GFP:capsid, then dialyzed overnight at 4 °C into 275 mM ionic strength running buffer. The components of each mixture were then separated via SEC and the fluorescence emission at 507 nm and UV absorption at 280 nm measured for each fraction eluted from the column (Figure 3.2). Each of the various mixture components was also run over the SEC individually to allow for comparisons between the single component and mixture chromatograms.

The wild-type and mutant capsid peaks both elute at the same volume, suggesting that they are the same size and therefore, the mutations introduced to the pore did not significantly alter the quaternary structure of the capsid. This conclusion is further supported by TEM (Figure 3.3). The size of the wild-type and mutant capsids was estimated at 17.4 ± 3.6 nm and 16.2 ± 2.4 nm, respectively. The 6D capsid does seem to be less stable, though, as it exhibits a small second peak which corresponds to the volume at which lesser oligomerization states of the capsid elute.

GFP-R_{10} elutes off of a sephacryl S400 column significantly later than the untagged variant and with a much broader peak. This elution volume is actually greater than the column volume and is likely due to interactions between the inherently sticky R_{10} tag and the column matrix. When mixed with either capsid, a peak was seen for each GFP variant at the same volume as when that variant was run individually (Fig 3.2a and 3.2b). All controls showed identical fluorescence patterns to the individually run GFP variant (data not shown). No fluorescence was visible to the naked eye in the capsid peak fractions from either control; however, a very small amount of fluorescence was detected via fluorimeter in the later capsid fractions of the GFP + AaLS-6D mixture. This low fluorescence is due

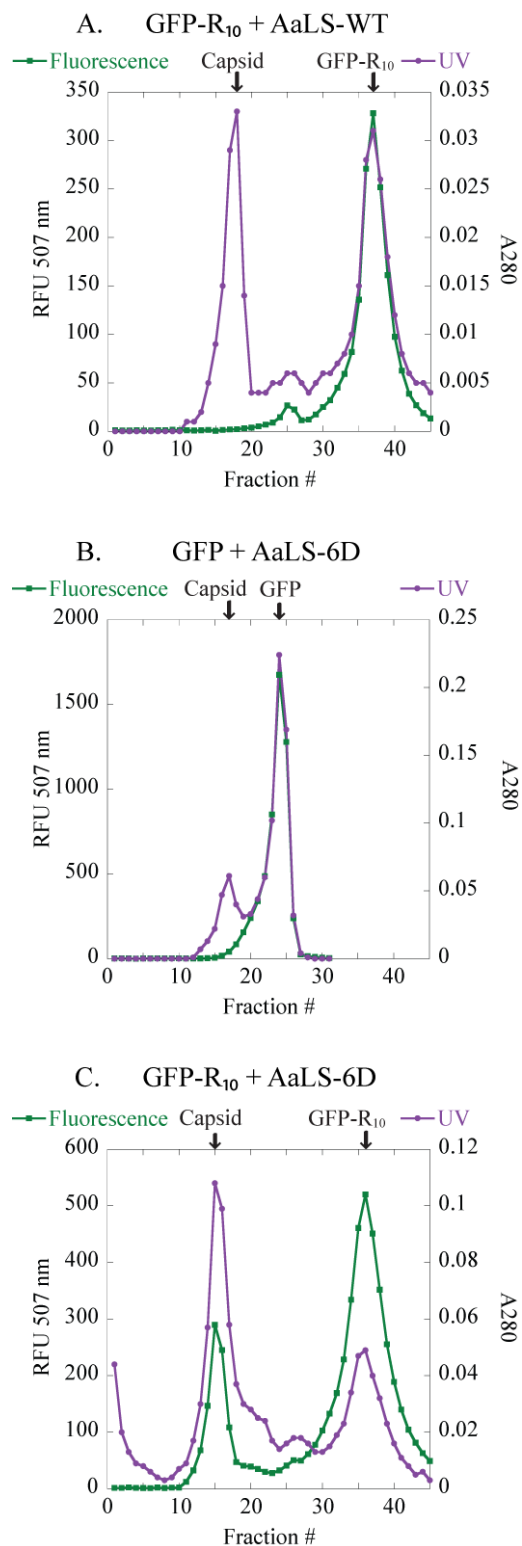Figure 3.2. Fluorescence and A_{280} measurements of SEC fractions from different GFP and capsid combinations. Samples were mixed at a ratio of 72:1 GFP:capsid and run in 20 mM Tris, 250 mM NaCl, 5 mM EDTA, pH 8 buffer. The position of the 60 subunit capsid and GFP variant are indicated by arrows.
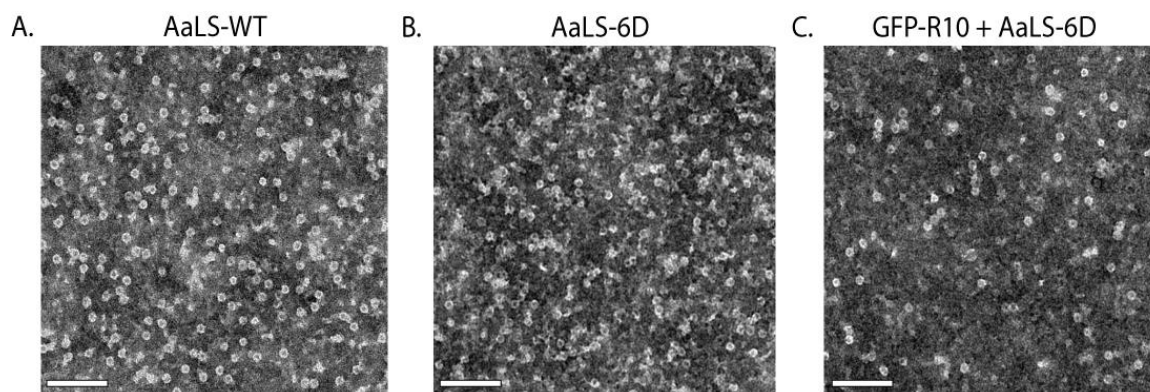
Figure 3.3. TEM images of the wild-type AaLS (A), AaLS-6D (B), and GFP-R$_{10}$ + AaLS-6D complex (C). Scale bars are 100 nm.

to the inability of the column used to obtain baseline separation of these two proteins. When both GFP-R_{10} and AaLS-6D were present, the capsid peak was brightly fluorescent to the naked eye and eluted slightly earlier than the mutant capsid alone. These are both good indicators of an association between the two components.

3.2.4. SDS-PAGE of Capsid Peak Fractions Shows Significant GFP-R_{10} Co-Elution with the AaLS-6D Capsid

To determine what proteins were present in the capsid peak fractions of the various mixtures, an SDS-PAGE was performed on these fractions. The SDS-PAGE of the GFP- R_{10} + AaLS-WT fractions shows only a single band for the capsid peak fractions of the expected size for the AaLS monomer (Figure 3.4). The GFP-R_{10} peak fractions also show a single band of the expected size. The GFP + AaLS-6D capsid peak fraction, in keeping with the fluorescence and UV results, shows a tiny amount of GFP present, relative to the amount of capsid protein. Only one band, corresponding to GFP, is observed in the GFP peak fractions. The capsid peak fractions for GFP-R_{10} + AaLS-6D clearly show bands for both GFP-R_{10} and AaLS-6D. The GFP peak fractions show two bands as well. This pair of bands is probably a result of GFP-R_{10} interacting with the lesser oligomerized forms of AaLS-6D. These findings agree well with the fluorescence and UV absorption data to support the idea that there is a significant association between the deca-arginine tagged GFP and AaLS-6D.

Figure 3.4. SDS-PAGE analysis of pooled SEC fractions from the capsid and GFP. Capsid (C) and GFP (G) peak fractions from the control and experimental mixes shown in Figure 3.2 are labeled. Lane 1 contains the molecular weight ladder. The expected positions of GFP and AaLS are indicated.
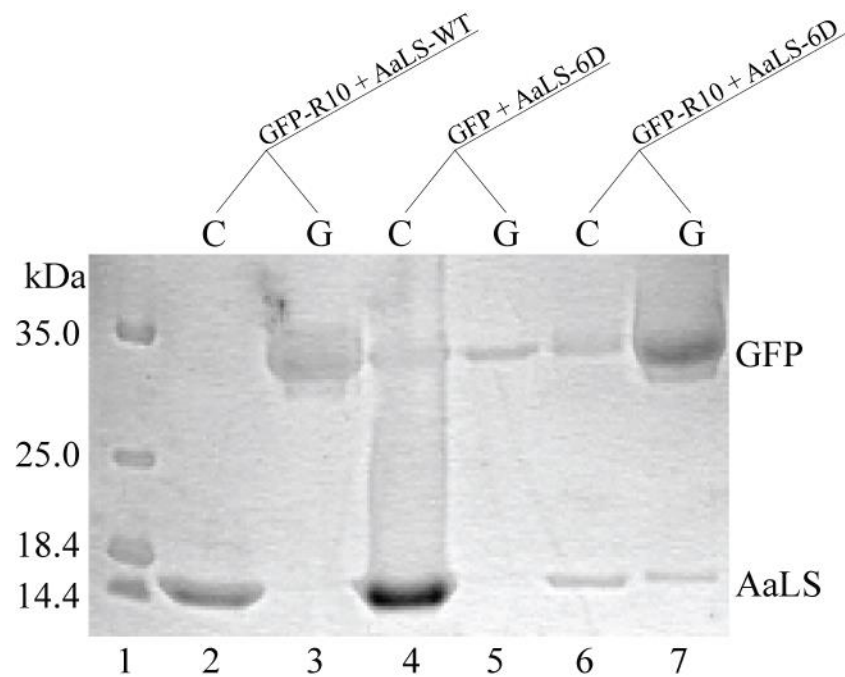
3.2.5. Gel Mobility Shift Assay Shows a Band Present only in the

GFP-R$_{10}$ + AaLS-6D Mixture

A native agarose gel shift mobility assay was performed with each GFP and AaLS variant run individually as well as in the three mixtures specified earlier. The individual GFP variants ran well and produced a single, well-defined band as viewed by both fluorescence and coomassie brilliant blue staining (Figure 3.5). The capsid variants also produced well-defined single bands. The mutant capsid appears to move further than the wild-type capsid, possibly due to the extra negative charge on the AaLS-6D variant (-25 per pore, -300 per capsid).

The GFP + AaLS-6D mixture behaved as expected with the individual components separating and moving exactly as their respective solo lanes did. The GFP-R$_{10}$ + AaLS-WT mixture, however, showed a smeared fluorescence band which extended approximately half the length of the smear observed by staining with no clearly defined bands. This smear suggests weak interactions exist between AaLS-WT and GFP-R$_{10}$ in the mobility shift assay running buffer that do not maintain cohesion under the higher ionic strength buffer conditions used during SEC. For comparison, the mobility shift assay running buffer has an ionic strength of 89 mM while the SEC running buffer varied in ionic strength from 125 mM to 475 mM. These interactions may be occurring between the weakly anionic wild-type pore and the tag, but a visual inspection of the structure reveals eight negatively charged residues (three Asp, five Glu) per monomer on the exterior surface of the capsid. Thus, nonspecific interaction between the R$_{10}$ tag and the outer capsid surface is another plausible explanation. The fact that the same fluorescent smearing is observed from the mixture of GFP-R$_{10}$ + AaLS-6D suggests the latter explanation is most likely. The
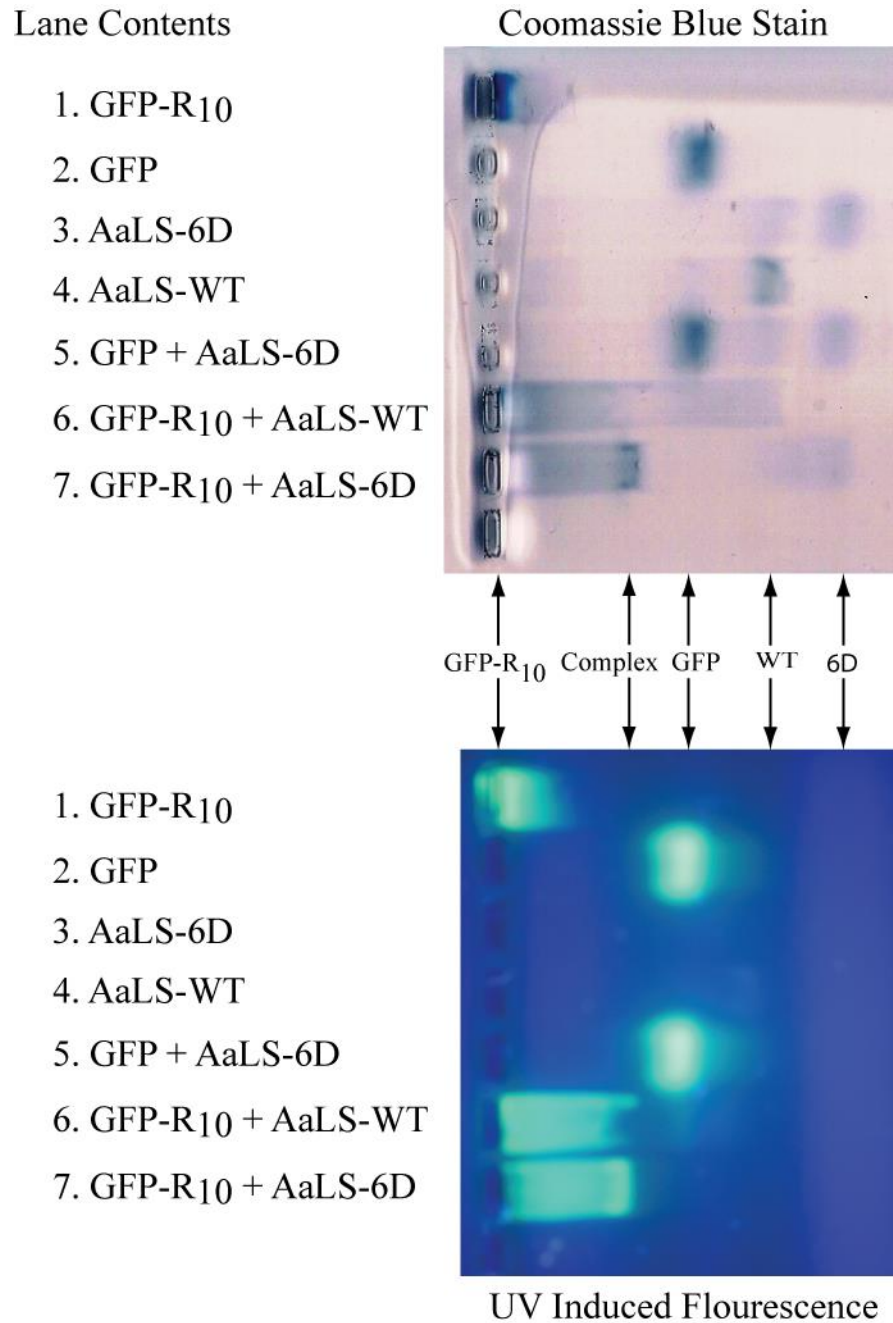
Figure 3.5. Native agarose gel mobility shift assay of the various GFP + capsid mixtures and their individual components. The positions of the individual proteins and the complex on both gels are indicated by arrows.

GFP-R_{10} + AaLS-6D lane also shows a clearly visible band at the end of the fluorescent smear. No other bands are present in the lane and this band is significantly shifted relative to both the GFP-R_{10} and AaLS-6D bands alone, indicating the cationic tag and anionic pore are able to interact with each other.

3.2.6. Stoichiometry and Ionic Strength Dependence of the Tag-Pore Complex

The architecture of the capsid gives rise to a single pore at the five-fold symmetry axis for a total of 12 pores per capsid. Therefore, we expected to see, at most, a ratio of twelve GFPs per capsid. However, at lower buffer ionic strengths, more than double that ratio was observed (Figure 3.6). The closest ratios to the expected 12:1 value were observed between ionic strengths of 225 mM to 275 mM. At higher buffer ionic strengths, the ratio dropped steeply until virtually no fluorescence was detected at 475 mM.

Given that the interior surface of both the wild-type and mutant capsid has a mildly negative net charge (-1 per monomer, -60 per capsid), it is possible that the extra GFPs have been encapsulated. This possibility seems extremely unlikely as the capsid peak fraction from the GFP-R_{10} + AaLS-WT control did not exhibit any measurable fluorescence (Figure 3.2a). Additionally, previous work has shown that a larger AaLS variant with a significantly more negative interior surface only encapsulates three or four GFP-R_{10} molecules and that encapsulation of an R_{10} tagged guest does not occur when the components are mixed after individual purification.[63] It is possible that more than one R_{10} tag could be threading through the pore. However, this seems extremely unlikely as our calculations show there is only enough volume for a single tag. It may be possible, though, for more than one tag to interact with just the surface residues of a single pore.
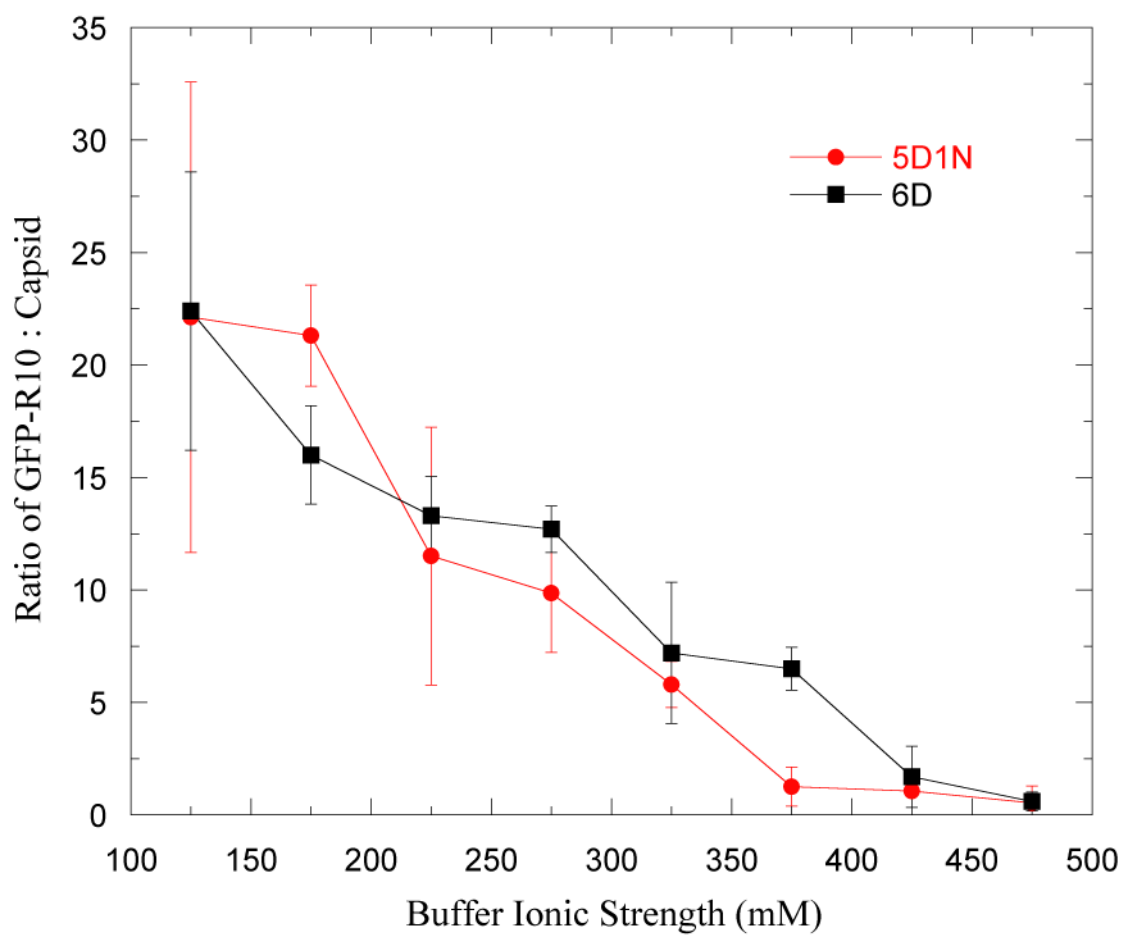
Figure 3.6. Ratio of GFP-R$_{10}$ per capsid at various buffer ionic strengths. Complexes were isolated via SEC. AaLS-5D1N showed a decreased ratio compared to AaLS-6D at most ionic strengths.

In an effort to reduce the amount of nonspecific binding observed, a D108N mutation was introduced to AaLS-6D, creating AaLS-5D1N. By converting the aspartate to an asparagine, we hoped to abolish any binding occurring on the outer edge of the pore. This capsid variant retained the ability to bind the tagged GFP. The ratio of GFP:capsid did decrease at most ionic strengths; however, it also appears that the ionic strength at which a 12:1 ratio of GFP:capsid was reached dropped from 275 mM with AaLS-6D to 225 mM with AaLS-5D1N (Figure 3.6).

Regardless of how the tag(s) may be interacting with the pore, examination of the elution behavior of the GFP-R_{10} + AaLS-6D capsid on a SEC column using various ionic strengths of running buffer also suggests that the tagged GFP is decorating the surface of the capsid. Since SEC separates molecules by their hydrodynamic volume, if the bound GFP-R_{10}'s were being encapsulated, we would expect the capsid to elute at the same volume irrespective of the number of bound guests because there would be no expected change in the hydrodynamic volume of the complex in comparison to the capsid alone.

What was actually observed was a slight but continual shift toward an earlier elution volume as the ionic strength of the running buffer was decreased (Figure 3.7). This shift is indicative of an increase in the hydrodynamic volume and is in keeping with what would be expected if the tag were binding to the pore with the guest on the outer surface of the capsid and an inverse relationship between the number of guest molecules and the ionic strength of the running buffer. This relationship is evinced by the increasing fluorescence intensity of the capsid peak as the ionic strength of the buffer decreases. It also supports the idea that nonspecific binding is occurring on the surface of the capsid because there
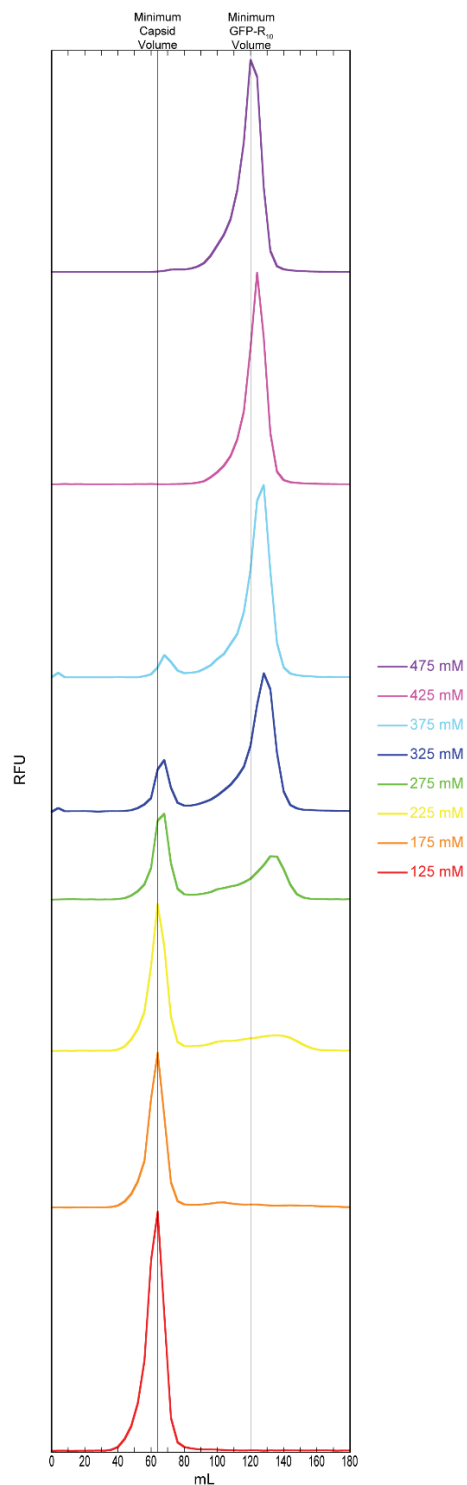
Figure 3.7 Relative fluorescence of GFP-R$_{10}$ + AaLS-5D1N complex and GFP-R$_{10}$ SEC peaks at different buffer ionic strengths. The GFP- R$_{10}$ peak elutes later as the buffer ionic strength decreases due to increased interaction between the tag and the column matrix. Ionic strength was controlled via sodium chloride concentration and includes the 20 mM tris buffer. Minimum elution volumes are shown to illustrate peak creep.

was no observed leveling off of the fluorescence intensity and earlier elution shifting of the capsid peak.

3.2.7. K_D Determination via Fluorescence Anisotropy

Fluorescence anisotropy was used to determine the strength of the complex association. Briefly, a fluorophore excited by plane polarized light will fluoresce in the same plane. The faster the fluorophore rotates in solution, the lower the intensity of the polarized emissions. The magnitude of the intensity difference between a fluorophore free in solution and one bound to another molecule is called the fluorescence anisotropy. By measuring the observed anisotropy at a constant concentration of fluorophore in the presence of different concentrations of a molecule that binds to the fluorophore, it becomes possible to estimate the K_D of the association.

The results of several anisotropy experiments using either AaLS-6D or AaLS-5D1N are shown in Figure 3.8. The estimated K_D's for both are statistically identical, suggesting that the mutation converting the aspartate on the outer edge of the pore to an asparagine in AaLS-5D1N did not significantly affect binding of the tag to the pore. Further, the nanomolar concentration of the K_D indicates an extremely strong affinity between the R_{10} tag and the negatively charged pore.

3.2.8. Stoichiometry and K_D Determination via Isothermal Titration Calorimetry

In an effort to further corroborate these data, we collaborated with Stign von Dongen at Radboud University Nijmegen to perform isothermal titration calorimetry

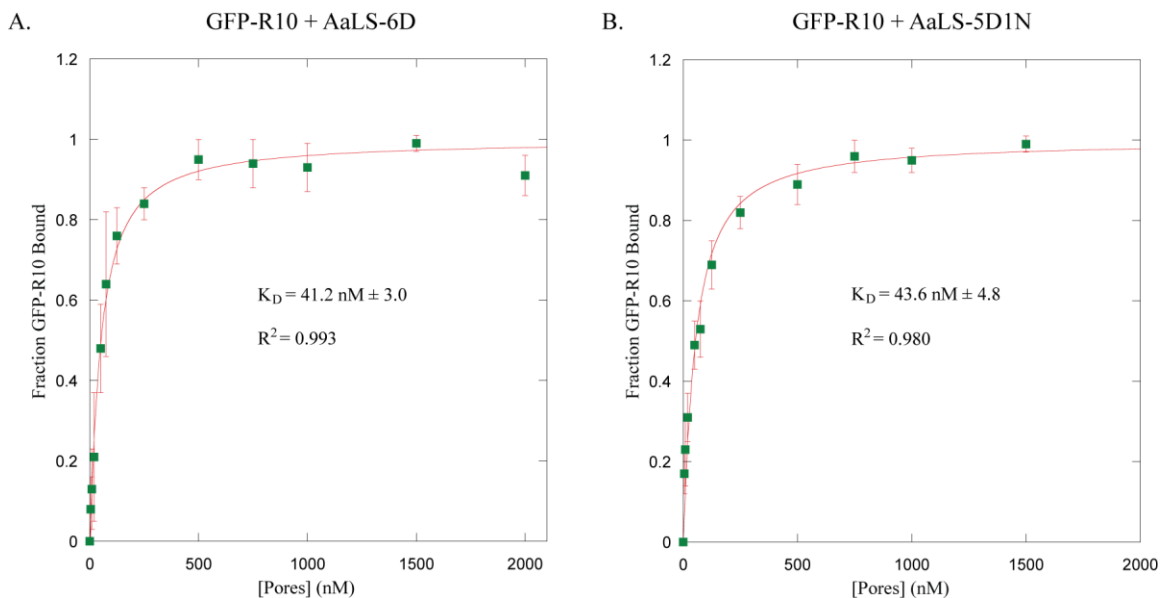Figure 3.8. Affinity of GFP-R_{10} for AaLS variants. Fraction of GFP-R_{10} bound at various pore concentrations of AaLS-6D (A) or AaLS-5D1N (B) as determined by fluorescence anisotropy measurements. Experiments were conducted at room temperature with buffer ionic strength held constant at 275 mM. The [pores] is defined as the protein monomer concentration * 12 pores/capsid * 1 capsid/ 60 monomers.

measurements on the tag-pore complex (Figure 3.9).[1] Initial calculations yielded an estimated K_D of 114 nM and a GFP:capsid ratio of 6.7:1. However, our collaborators used spectroscopic methods to determine their protein concentrations. More specifically, they used A_{280} measurements with an extinction coefficient calculated via the ExPASy ProtParam tool to determine their capsid concentrations and A_{480} measurements with an empirically determined extinction coefficient to determine their GFP-R_{10} concentrations. This method results in capsid and GFP-R_{10} concentrations approximately 1.9 and 1.3 times higher, respectively, than the concentrations determined by Bradford assay, the method that we used exclusively for protein concentration determination. When the ITC measurements are adjusted for these differences, the GFP:capsid ratio becomes 13:1 and the calculated K_D becomes 47 nM.

Thermodynamic parameters derived from these data indicate approximate changes in enthalpy and entropy of -4.2 kcal/mol and 5.3 kcal/mol, respectively. These calculations suggest that the binding of the tag to the AaLS-5D1N pore is entropically driven, most likely by freeing water molecules bound to the interior of the otherwise empty pore.

3.3. Discussion

A variant of the AaLS capsid with a negatively super-charged five-fold axial pore and a GFP with a C-terminal deca-arginine tag were shown to form a novel association with each other via charge complementarity between the anionic pore and the cationic tag. This complex demonstrated strong binding with a K_D in the low nanomolar range. Such an

---

[1] Due to technical difficulties, only one ITC measurement was conducted.
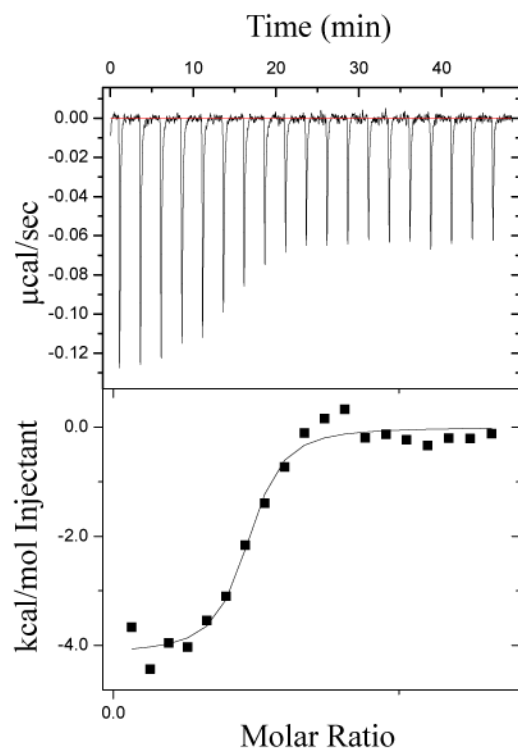
Figure 3.9. Thermogram for the binding of GFP-R$_{10}$ to AaLS-5D1N.

extremely tight binding affinity is on par with those determined for antibody-antigen interactions.[186]

Antibodies often undergo a process called affinity maturation in which successive generations show stronger antigen affinities, usually by one or two orders of magnitude. Structural studies have shown that this improved affinity can be due to an increase in the rigidity and complementarity of the interface with a corresponding increase in the number of hydrogen bonds.[187] This degree of interaction requires significant desolvation between the interfaces to prevent water from disrupting the hydrogen bonding network and salt bridges.

Salt bridges are regularly found to be buried at protein-protein interfaces where they are often stabilizing and contribute significantly to the specificity of the interaction.[188] The strength of attraction between the two components, and therefore of the bridge itself, can be calculated from Coulomb's law which dictates that the strength increases as the dielectric constant of the environment decreases.[189] The estimated packing efficiency of the tag-pore complex indicates a tight fit, very near the crystallographic maximum packing density. Given this projection, the fact that the association appears to be entropically driven, and arginine's ability to form up to five hydrogen bonds, it seems likely that there will be little or no water in the interface, yielding an extensive network of high strength salt bridges and hydrogen bonds.

Although it is strong, the association we generated is not highly specific. Any positively charged particle can associate with a negatively charged one. The architecture of the pore does impart some specificity as only molecules capable of fitting into the pore lumen will be able to bind there. However, the relatively simple hourglass surface of the

pore means a fairly wide variety of narrow, elongated molecules or small globular ones would be able to fit into the pore. This lack of specificity could be the source of the greater than 12:1 stoichiometries observed at low ionic strengths.

It is also possible that the tag is only interacting with the surface of the pore, rather than threading through it as envisioned (Figure 3.10a). The hourglass shape of the pore causes it to flare wider at the ends, which could allow multiple tags to interact with just a few aspartates at the top of the pore. However, interacting like this would likely result in a significantly weaker dissociation constant than was observed. The fact that conversion of aspartate 108 to asparagine did not change the observed K_D suggests that nonspecific interactions do not play a significant role in formation of the tag-pore complex at higher buffer ionic strengths.

Alternatively, at lower ionic strengths, the R_{10} tag may be able to replace the positive residue in some or all of the 17 salt bridges found on the AaLS capsid surface. The smeared fluorescence seen in the gel mobility shift assay (Figure 3.5 lane 6) indicates some degree of association between GFP-R_{10} and the wild-type capsid. However, at the higher ionic strength of the SEC running buffer, there is no sign of any interaction between the two (Figure 3.2a). This supports the idea that there is some degree of nonspecific interaction between the R_{10} tag and the surface of the capsid in low ionic strength buffer.

The sensitivity of the complex to ionic strength indicates that the association should be tunable based on the buffer conditions and the amino acid content of the tag and pore. Fewer positive and/or negative charges should result in a weaker complex which dissociates at an ionic strength lower than 375 mM, as we observed. Given the fact that the ionic strength of blood and that of the intracellular environment differ by a fairly significant
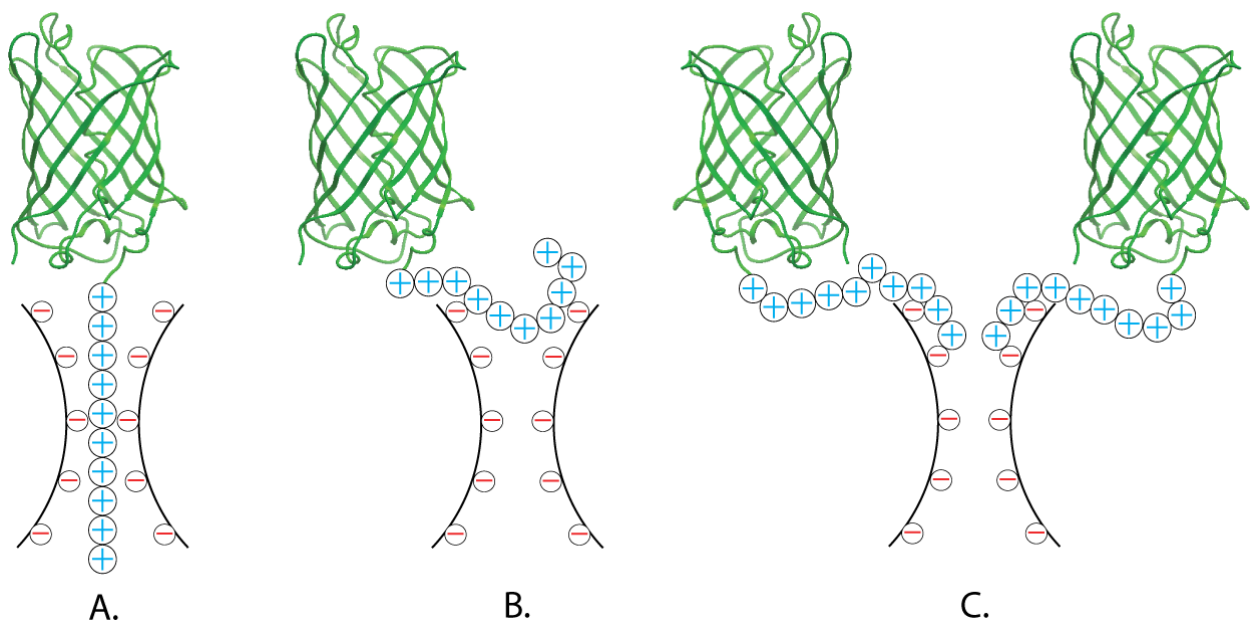
Figure 3.10. Potential binding modes between the deca-arginine tag and the pore. A) Threaded through the pore as envisioned. B) Interacting only with the surface aspartates of the pore. C) Multiple tags partially interacting with the surface aspartates of the pore.

amount (150 mM vs 250 mM, respectively), this enhances the potential use of the capsid as a drug delivery vehicle.

It may be possible to reduce the strength of the tag pore complex via mutagenesis such that transition from the blood to the cell interior causes the complex to dissociate, thereby exposing the interior of the capsid and allowing delivery of small molecule drugs. Further, the multivalent display of receptor binding proteins to address the capsid to specific cell types would be easily achievable by appending the tag to these proteins. Thus, the problems of specific delivery and drug release may both potentially be addressed via a single, novel protein-protein interface.

In addition, control over supramolecular assembly and multivalent display is a highly desirable trait for such nanotechnology applications as antigen display in vaccines, control over nanoreactor activity, and production of novel nanoscale architectures. Our method of functionalizing the AaLS capsid pore surface is compatible with previously reported surface engineering techniques such that the simultaneous multivalent display of several different guest proteins may be possible. The use of a charge complementarity-based tagging system to functionalize capsid pores should be generally applicable to other capsids or tunnel bearing structures as a means of engineering novel supramolecular assemblies.

3.4. Materials and Methods

3.4.1. Materials

All cell culture media and chemical reagents were purchased from Bio-rad, Fisher Scientific, Gold Biotechnology, or Pierce Biotechnology and used without further

purification. BL21 (DE3) and XL1-Blue *E. coli* cell strains and *Pfu*-turbo DNA polymerase were purchased from Stratagene. T4 DNA ligase and all restriction endonucleases were purchased from New England Biolabs. The oligonucleotides used in this study were synthesized by the DNA/peptide synthesis Core Facility at the University of Utah.

3.4.2. Mutagenesis

In order to make AaLS-6D, site-directed mutagenesis PCR was carried out using pMG-AaLS as the template plasmid for the mutagenic primers EH10 and EH11 (Table 3.1). A 51 µL reaction volume was made by mixing the following: nanopure water (42 µL), 10x reaction buffer (5 µL), 10 ng/µL template DNA (1 µL), 250 ng/µL forward and reverse primers (0.5 µL each), dNTP mix (1 µL), and 2.5 U/µL *Pfu*-turbo DNA polymerase (1 µL). The reaction was carried out using a Mastercycler Personal Thermocycler (Eppendorf). The polymerase was activated (95 °C for 30 seconds) prior to initiation of the mutagenesis reaction which was carried out for 18 cycles of denaturation (95 °C for 30 seconds), annealing (55 °C for 1 minute), and primer extension (68 °C for 12 minutes). The mixture was then cooled to 37 °C and incubated for 1 hour with 1 µL DpnI (10 U/µL) to remove the template plasmid.

The product plasmid was transformed via heat shock (42 °C for 1 minute) into CaCl$_2$ competent XL1-Blue cells. These cells were then grown on 100 µg/mL AMP LB-agar plates overnight at 37 °C. Colonies were selected, grown in 6 mL LB media containing AMP, and plasmid DNA extracted via a QIAprep Spin Miniprep Kit (Qiagen) as per the manufacturer's protocol. The presence of the desired mutations was confirmed by DNA sequencing conducted by the University of Utah DNA sequencing core facility. A similar

Table 3.1 List of mutagenic primers used to generate the capsid variant, AaLS-5D1N. Mutagenic bases are bolded.

Primer ID	DNA Sequence (5' → 3')
EH10	CATTTCGATTATATCGCC**GA**TGAAGTTTCA**GA**T GGCCTCGCG **GA**CCTTTCATTAGA**C**CTA**GA**TAAACCTATCACCTTCGG
EH11	CCGAAGGTGATAGGTTTA**TC**TAG**G**TCTAATGAAAGG**T**C CGCG AGGCC**A**T**C**TGAAACTTCA**TC**GGCGATATAATCGAAATG
KINGf	GCGGACCTTTCATTAGACCTA**A**ATAAACCTATCACCTTC
KINGr	GAAGGTGATAGGTTTAT**T**TAGGTCTAATGAAAGGTCCGC

protocol to the one described above was used to produce pMG-AaLS-5D1N from pMG-AaLS-6D and the primers AaLS-KINGf and AaLS-KINGr.

3.4.3. Production and Purification of Crude Protein

All proteins were produced in $CaCl_2$-competent BL21 (DE3) *E. coli* cells. Transformation of the gene-bearing plasmids (pMG-AaLS (hexahistidine tagged wild-type capsid), pMG-AaLS-6D, pMG-AaLS-5D1N, pACYC-GFP, or pACYC-GFP-R_{10}) was accomplished via heat shock at 42 °C for 2 minutes. The cells were then grown at 37 °C for 1 hour and plated on antibiotic-containing LB agar plates (50 µg/mL chloramphenicol (CAM) for pACYC plasmids, 100 µg/mL ampicillin (AMP) for pMG plasmids). The plates were grown overnight in a 37 °C oven. A single colony was selected from the plate and used to inoculate a 7.5 mL antibiotic-containing LB culture which was then grown overnight at 37 °C with shaking at 250 rpm. An aliquot of this culture (1 mL) was then used to inoculate antibiotic-containing LB media (500 mL). Over-production of the desired protein was achieved by growing this large culture at 37 °C to an OD_{600} of 0.7, inducing production by adding IPTG to a final concentration of 0.2 mM, and then incubating the cultures at 30 °C for 20 hours at 250 rpm. The cells were harvested via centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (6000 rpm for 15 minutes at 4 °C). Pellets were frozen at -80 °C until used.

Cell pellets were resuspended in 10 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). The cells were lysed by incubation with lysozyme (10 mg), RNase A (1.2 mg), and DNase I (20 µg) on ice for 1 hour followed by sonication at 45 µm amplitude for 3 minutes in 10 second on/off pulses using a Misonix ultrasonic liquid

processor. The lysate was then clarified by centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (12000 rpm for 45 minutes at 4 °C). Unless stated otherwise, a Bradford assay was used to determine all protein concentrations.

3.4.4. Purification of AaLS Capsid Variants

For purification of AaLS-WT, AaLS-6D, or AaLS-5D1N, the supernatant was loaded onto a Ni^{2+}-NTA resin which had been equilibrated with lysis buffer and incubated for 1 hour at 4 °C with gentle rocking. The column was then washed with lysis buffer, lysis buffer containing 40 mM imidazole, and lysis buffer containing 500 mM imidazole. The high imidazole fraction was then concentrated. The assembled capsid was isolated by loading the sample onto an ÄKTA FPLC system (GE Healthcare) equipped with a HiPrep 16/60 Sephacryl S-400HR column and run with tris running buffer (20 mM Tris, 500 mM NaCl, 5 mM EDTA, pH 8) at 4 °C at a flow rate of 0.5 mL/min.

3.4.5. Purification of GFP Variants

Prior to dialysis, all GFP and GFP-R_{10} samples were treated with saturated ammonium sulfate in lysis buffer to a final concentration of 40% v/v ammonium sulfate. These samples were incubated at room temperature for 1 hour and then clarified by centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (10,000 rpm for 15 minutes at 4 °C). The supernatant was retained and the target protein was further purified from the crude sample as described in section 2.4.4. The untagged GFP variant was loaded onto a MonoQ 5/50 GL column while GFP-R_{10} was loaded onto a MonoS 5/50 GL column. For both ion exchange columns, bound protein was eluted using a 20 mL

gradient ranging from 0% to 100% ion exchange buffer B run at 4 °C. Fractions that fluoresced green to the naked eye were pooled and then loaded onto a HiLoad 16/60 Superdex 75 PG column run with tris running buffer at 4 °C.

3.4.6. Assembly and Isolation of the Tag-Pore Complex

Mixtures of GFP and capsid were prepared in tris running buffer by combining a GFP variant with a capsid variant to final concentrations of 1 mg/mL and 0.5 mg/mL, respectively. This results in a 72-fold molar excess of GFP (6-fold compared to the pores). The components were added in the following order: tris running buffer, GFP, capsid. The sample was produced in a 2.5 mL volume and dialyzed overnight at 4 °C, with gentle stirring, into working buffer (identical to tris running buffer but with the [NaCl] varying from 100 mM to 450 mM) of the desired ionic strength. The ionic strength of the buffer was controlled by manipulating the concentration of NaCl. The sample was then loaded onto an ÄKTA FPLC system (GE Healthcare) equipped with a HiPrep 16/60 Sephacryl S-400HR column and run with the same ionic strength tris buffer that the sample was previously dialyzed into. Fluorescence measurements were taken for all fractions eluted from the column.

3.4.7. Native Agarose Gel Mobility Shift Assay

A 2% w/v agarose gel was prepared using running buffer consisting of 89 mM Tris, 89 mM Boric acid, at pH 8.4. The sample loading buffer consisted of running buffer, 3 mM bromophenol blue, and 20% v/v glycerol. Samples were prepared by combining a 10 µL aliquot of a 1mg/mL protein sample with 10 µL of loading buffer. The entire 20 µL

preparation was loaded into a single well on the gel which was run at 100 V for 70 minutes. Gels were cast and run in a Bio Rad Mini-Sub Cell GT using a Bio Rad PowerPac Basic. All gels were initially visualized via UV-induced fluorescence on a UVP Benchtop Transilluminator to locate GFP in the samples. They were then stained overnight with Coomassie Brilliant Blue to locate all proteins in each sample.

3.4.8. Fluorescence Measurements

All fluorescence and fluorescence anisotropy measurements were conducted on a Hitachi F-7000 Fluorescence Spectrophotometer with the sensitivity set to 450 V. The polarizer attachment was used for all fluorescence anisotropy measurements. Fluorescence anisotropy measurements were carried out in 20 mM Tris, 250 mM NaCl, 5 mM EDTA, pH 8 buffer. Samples were prepared immediately prior to analysis and allowed to equilibrate for 15 minutes at room temperature, in the dark. The initial sample consisted of 20 nM GFP-R_{10} and 5 µM AaLS pores (0.4 µM capsid) in a total volume of 2 mL. The fluorescence anisotropy of this sample was measured and the concentration of AaLS pores was then diluted by removing 1 mL of the sample and replacing it with 1 mL of a 20 nM GFP-R_{10} solution in the same buffer. This cycle of measurement and dilution was repeated ten more times. Data were fit to Equation 3.1 with the following parameters: B is the fraction of bound ligand, L_T is the total concentration of ligand, R_T is the total concentration of receptor (AaLS capsid pores), and K_D is the dissociation constant. The fraction bound was estimated as the measured anisotropy over the maximum observed anisotropy.

$$B = [L_T + K_D + R_T - ([L_T + K_D + R_T]^2 - 4L_T R_T)^{1/2}]/2R_T \qquad \text{(Equation 3.1)}$$

3.4.9. Isothermal Titration Calorimetry

All ITC experiments were carried out using a Microcal Auto-iTC200 (GE Healthcare). Samples were analyzed in 20 mM Tris, 250 mM NaCl, 5 mM EDTA, pH 8 buffer at 25 °C. A 400 µL aliquot of 57 µM 5D1N monomer was titrated via 19 injections (2 µL each) of 93 µM GFP-R_{10}. Results were analyzed using Origin 6.0 (Microcal) and fitted using a nonlinear single site binding model. The concentrations of all proteins used for ITC measurements were determined spectrophotometrically as discussed in section 3.2.8.

3.4.10 Transmission Electron Microscopy

In order to visualize the mutant capsid alone and in complex with GFP-R_{10}, 10 µL of the each sample was applied to separate 200 square mesh, formvar coated copper grids. Samples were allowed to adhere to the grid for 1 minute before the excess was wicked away with sterile filter paper. The grids were then stained for 1 minute using 2% phosphotungstic acid (pH 8.0). Excess stain was wicked away with sterile filter paper. Sample images were obtained on a Hitachi 125 keV H-7100 Transmission Electron Microscope equipped with a Gatan Orius SC1000 slow scan 4kX2.6k CCD camera and analyzed using ImageJ software (National Institutes of Health).

CHAPTER 4

FURTHER EXPLORATION OF AN ENGINEERED REDOX SWITCH AND

DEVELOPMENT OF A NOVEL pH SWITCH FOR CAPSID ASSEMBLY

4.1 Introduction

Polyhedral capsids represent one of the more common multimeric protein associations. The building blocks of such structures typically self-assemble in a hierarchical fashion to form symmetric, hollow shells that can serve as molecular containers.[11,21] The functions of protein capsids include protection of encapsulated guest molecules, sequestration of toxins, and enhancing the speed and efficiency of metabolic pathways. Thus, there has been significant interest in harnessing these structures for use in nanotechnology. In particular, control over capsid assembly and disassembly is highly desirable as it allows for precise loading and unloading of guest molecules.

The capsid formed by *Aquifex aeolicus* lumazine synthase (AaLS) provides an attractive platform for nanotechnology. AaLS forms a dodecahedral capsid from 60 identical subunits, which can be viewed as a dodecamer of pentamers.[131] Previously, our lab reported the generation of a pentameric variant of AaLS (AaLS-Switch-Red), which contains point mutations near the two-fold (R40S and H41S) and three-fold (I125C) symmetry axes of the wild-type capsid.[135] The mutations along the two-fold axis disrupt a hydrogen bonding network while the mutation at the three-fold axis replaces a hotspot residue within a hydrophobic cluster. These regions are part of the pentamer-pentamer interface within the capsid. The aforementioned mutations disrupt the noncovalent interactions which stabilize this interface, resulting in the formation of individual pentamers rather than full capsids.

AaLS-Switch-Red can be induced to form wild-type-like capsids via disulfide bond formation between the engineered cysteine (at position 125) and thiophenol (Figure 4.1) to give AaLS-Switch-Ox.[85] The increase in nonpolar surface area provided to the protein by

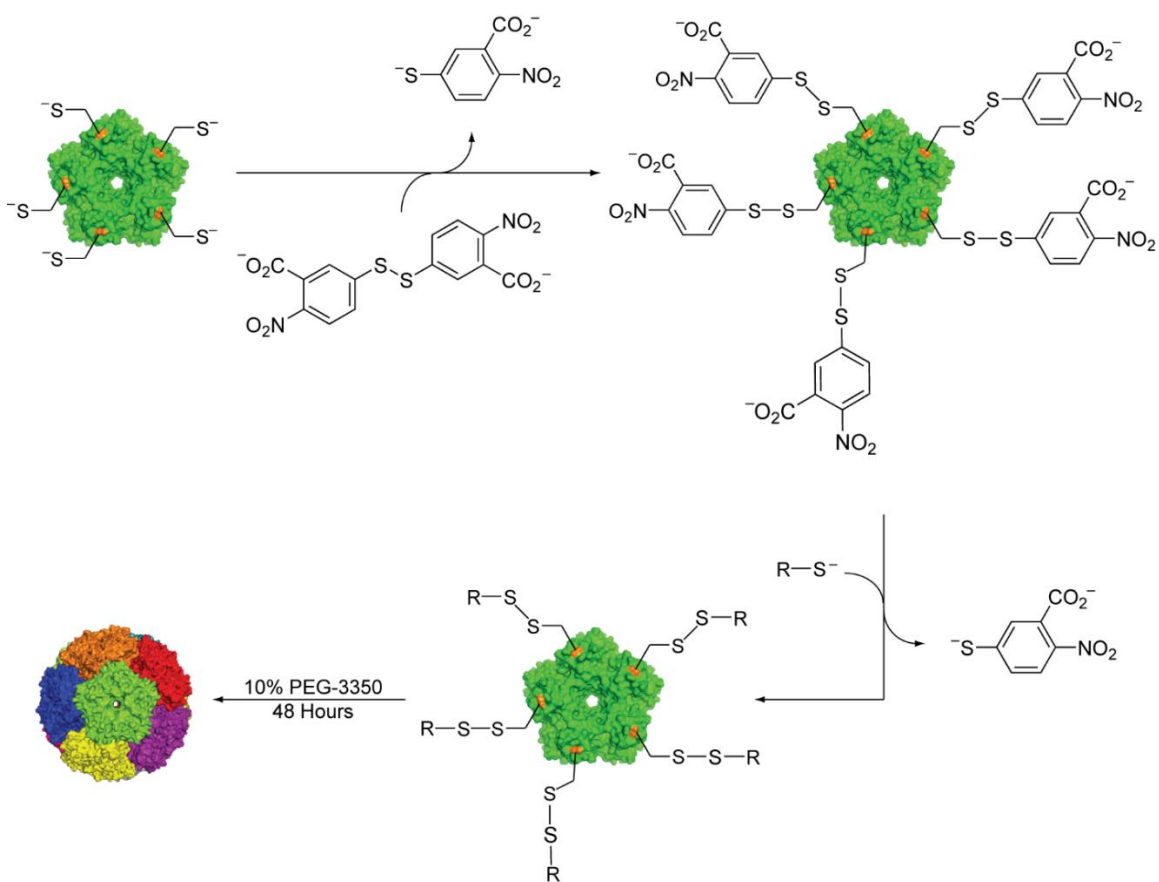Figure 4.1. General schematic of the thiol disulfide exchange reaction used to induce capsid assembly. First, the engineered cysteines (orange) on each pentamer were primed by reaction with DTNB to form the AaLS-Switch-NTB adduct (top right). Next, the NTB moiety was replaced with one of the thiols shown in Table 4.1, denoted as R, and the Switch-thiol adduct was incubated with PEG-3350 to produce assembled capsids.

this modification presumably allows for the reconstitution of the hydrophobic cluster at the three-fold symmetry axis and drives capsid assembly by AaLS-Switch-Ox.

Here, we expand upon this work by examining how capsid assembly is affected by the size and shape of the thiol adduct. Interestingly, we find that capsid assembly by AaLS-Switch-Ox can tolerate a wide range of aliphatic and aromatic prosthetic groups, suggesting that this hot spot in the pentamer-pentamer interface of the capsid is quite plastic. In addition, we use this disulfide exchange system as a means of combining both pH and redox sensitive switches for further control over capsid assembly and disassembly.

4.2. Results

4.2.1. Formation of the Pentamer-Thiol Adducts

Formation of the pentamer-thiol adducts was accomplished via a two-step disulfide exchange reaction in which the Switch-Red pentamer was initially treated with DTNB to generate an adduct via disulfide bond formation between cysteine 125 of each monomer and NTB (Switch-NTB). This initial adduct formation was typically accomplished with greater than 90% efficiency. The Switch-NTB adduct was then reacted with the desired thiol to form the final Switch-Thiol adduct via displacement of the NTB moiety by the added thiol. Reaction efficiency was monitored via UV-spectroscopy by measuring the absorbance of the solution at 412 nm. The majority of these disulfide exchange reactions showed greater than 80% efficiency (Table 4.1).

Mass spectrometry was used to confirm the presence of the intended adduct (Table 4.2). The spectra for all tested samples showed a peak corresponding to the expected size and, with three exceptions, this was also the major peak. The Switch-Oct sample displayed

Table 4.1 Formation efficiency and capsid yield of the various capsid generating adducts. Overall capsid yield was calculated by dividing the mass of protein obtained from the SEC capsid elution peak by 4 mg of starting material. Adjusted capsid yield was calculated by dividing the mass of capsid protein by the product of 4 mg and the adduct yield. n.d. = not determined

AaLS-Switch Variant	Adduct (R = C125 Sidechain)	Adduct Yield (%)	Overall Capsid Yield (%)	Adjusted Capsid Yield (%)
n - Butyl		96	46	47
t - Butyl		94	30	31
Pent		73	24	33
TMP		92	51	55
Hex		70	33	48
CHex		77	44	58
Hept		93	46	49
Oct		89	61	68
Ox		100	32	32
Bz		n.d.	41	n.d.
2TP		69	11	16
4TP		59	11	19
mOx		91	37	41
eOx		81	16	20

Table 4.2 Mass spectrometry results for the different AaLS-Switch variants.

AaLS-Switch Variant	Calculated Mass (Da)	Observed Mass (Da)
Red	16786	16787
NTB	16985	16983
n-Butyl	16874	16880
t-Butyl	16874	n.d.
Pent	16888	16889
TMP	16931	16855
Hex	16902	16899
CHex	16900	16899
Hep	16917	16919
Oct	16931	16998
Ox	16894	16898
mOx	16908	16910
eOx	16922	16926
2TP	16895	16881
4TP	16895	16898
Bz	16876	16878
Bz†	16924	16925

† AaLS-Switch-Bz after 24 hour peroxide treatment. The calculated mass assumed that all three methionines in the protein where oxidized to sulfoxides.

a spectrum with many peaks, including a major peak corresponding to the expected mass plus 67 Da. This additional mass is very close to the expected mass of three sodium atoms and is likely due to residual salt that was not properly cleared from the sample. The major peaks for the Switch-TMP and Switch-2TP samples were 76 and 14 Da lower than anticipated, respectively. It is not immediately clear what these peaks indicate about their respective samples as the missing masses do not correspond to any probable adduct fragments.

4.2.2. Verification of Assembly of Pentamer-Thiol Adducts into Capsids

After thiol treatment, the resulting adducts were dialyzed to remove excess thiol reagent and assembled via incubation with PEG-3350 for two days. The procedure used in this study was a scale-up of the previously reported assembly method using Switch-Ox[85]. The original protocol called for a 1 mL reaction volume and gave quantitative assembly after two days while the protocol reported here employed a 4 mL reaction volume and did not achieve quantitative assembly for any of the tested thiol adducts. The reason for this nonlinear behavior is not immediately apparent. Assembled capsids were separated from lower order aggregates and unassembled pentamers via SEC (Figure 4.2). Fully formed capsids typically eluted at 69-70 mL (Table 4.3). None of the thiols shorter than four carbons in length assembled into capsids (Figure 4.2 c,d,e).

TEM images of the capsid elution peaks were obtained to visually verify that they had assembled into capsids and to determine their sizes (Figure 4.3). The assembled Switch-thiol capsids were comparable in size to the wild-type capsid (Table 4.3). Mass
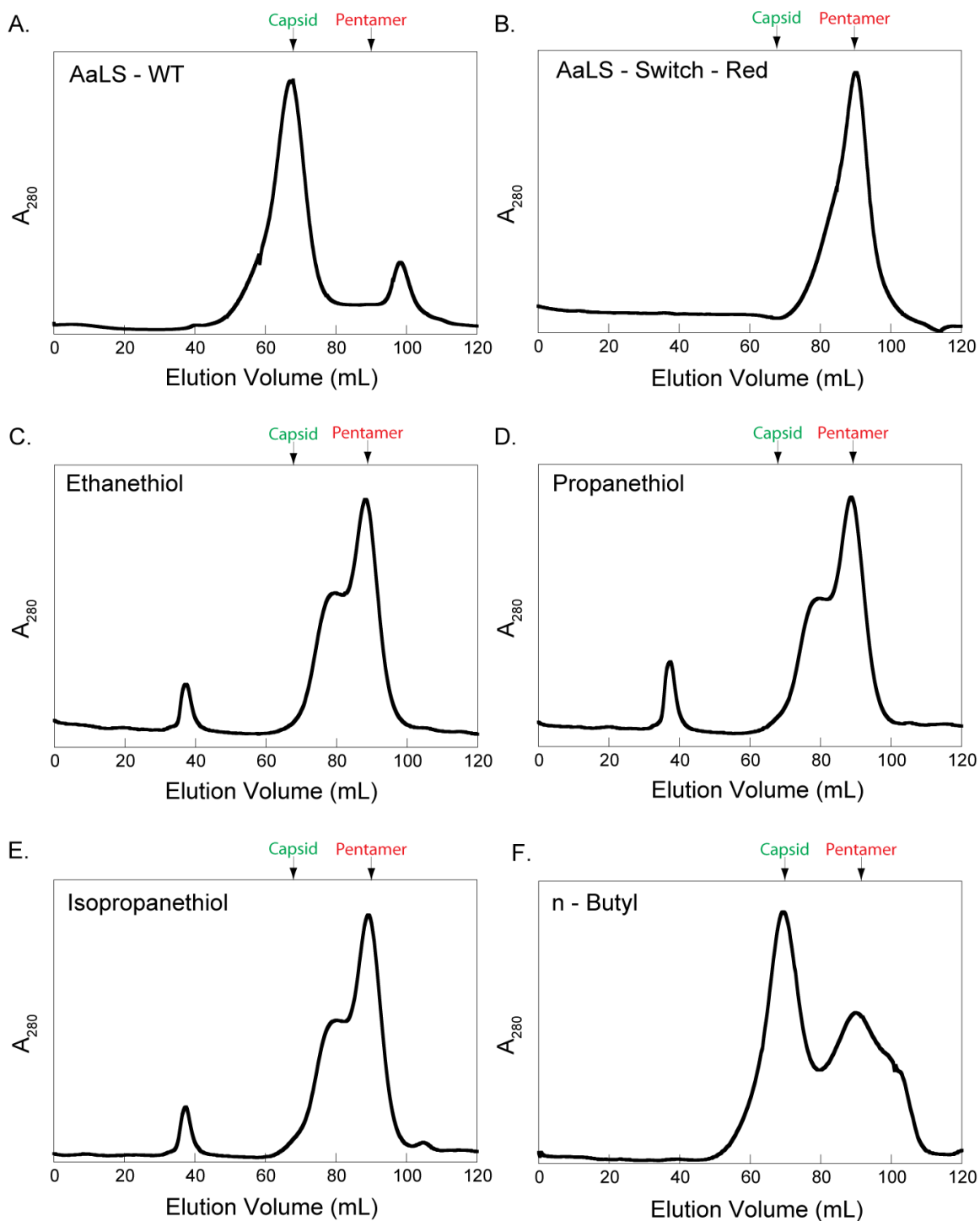
Figure 4.2. SEC chromatograms of all tested thiol assemblies. AaLS-WT and AaLS-Switch-red are shown as capsid and pentamer references, respectively. Data for the latter was collected by H. Chen. Expected elution volumes are denoted for capsid (green) and pentamer (red).
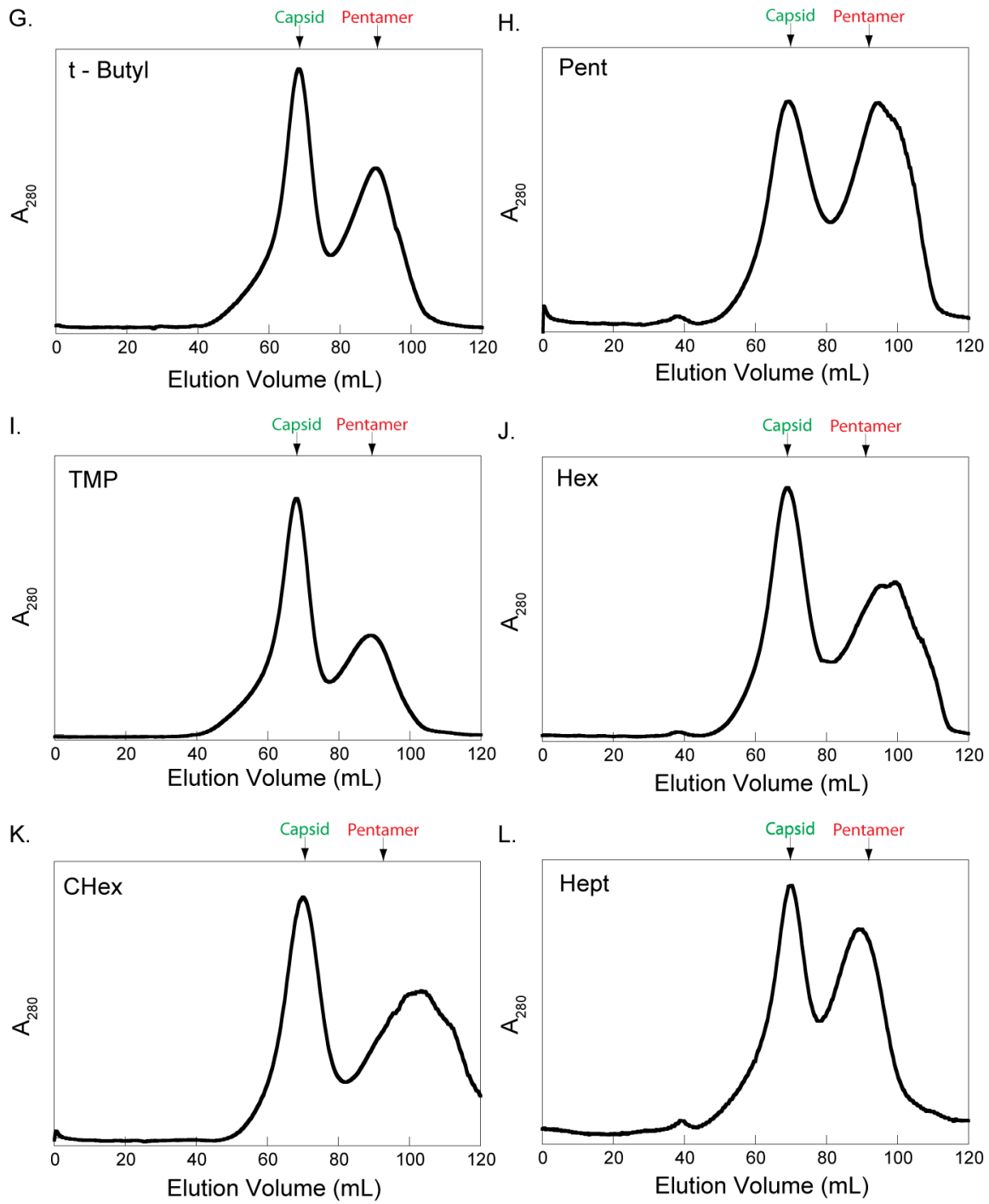
Figure 4.2. Continued.
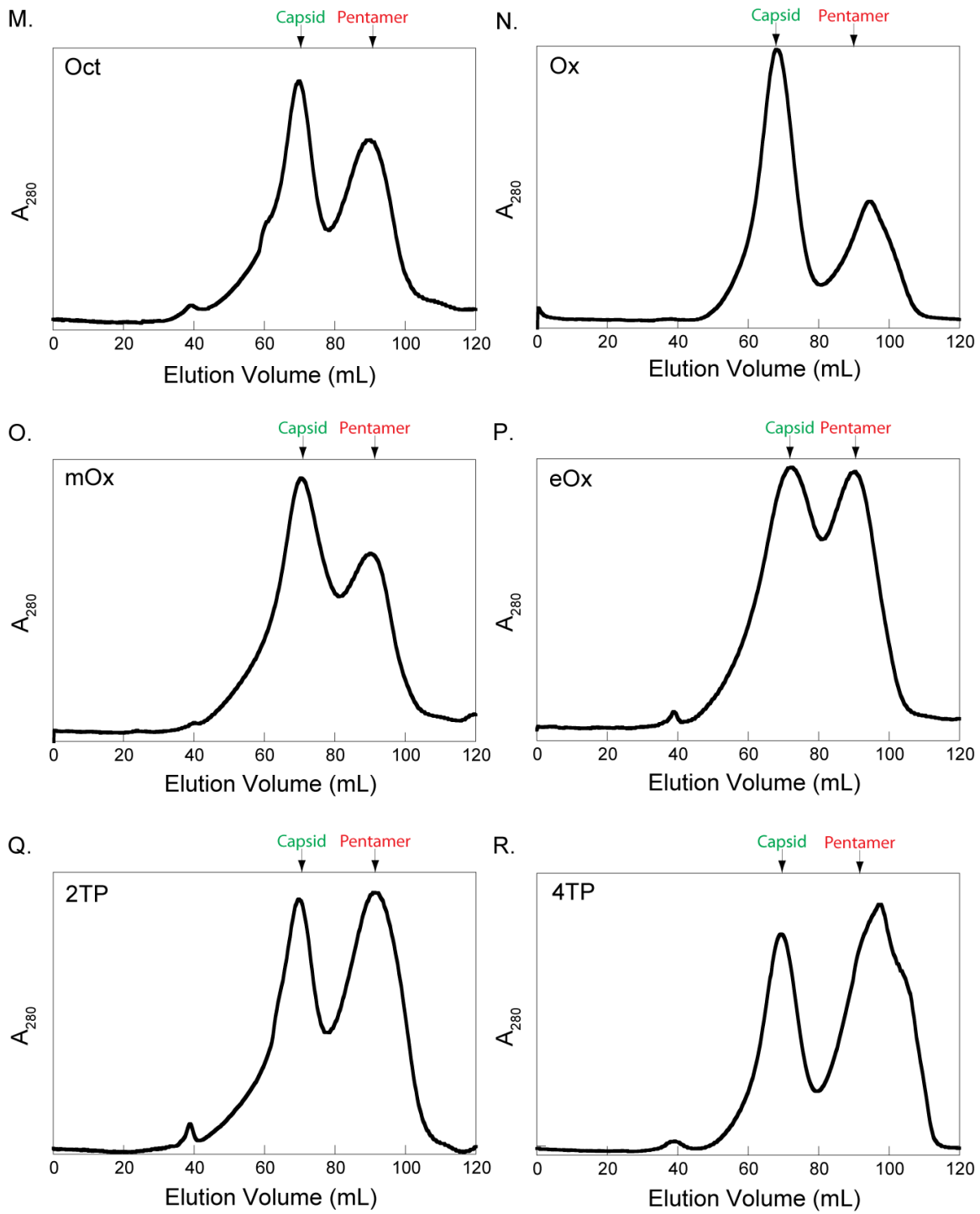
Figure 4.2. Continued.

Table 4.3 Size analysis of the different thiol capsids.
n.d. = not determined

AaLS-Switch Variant	SEC Elution Volume (mL)	Diameter (nm)	Number of Particles
n-Butyl	69	14 ± 2	168
t-Butyl	69	n.d.	-
Pent	69	16 ± 2	183
TMP	68	15 ± 2	125
Hex	69	16 ± 1	151
Chex	70	14 ± 2	87
Hep	70	13 ± 1	116
Oct	70	15 ± 2	107
Ox	69	16 ± 2	77
mOx	70	15 ± 2	71
eOx	73	14 ± 1	134
2TP	70	16 ± 2	199
4TP	69	18 ± 2	347
Bz	71	15 ± 2	300

A. n - Butyl

B. Pent

C. TMP

D. Hex

E. CHex

F. Hept
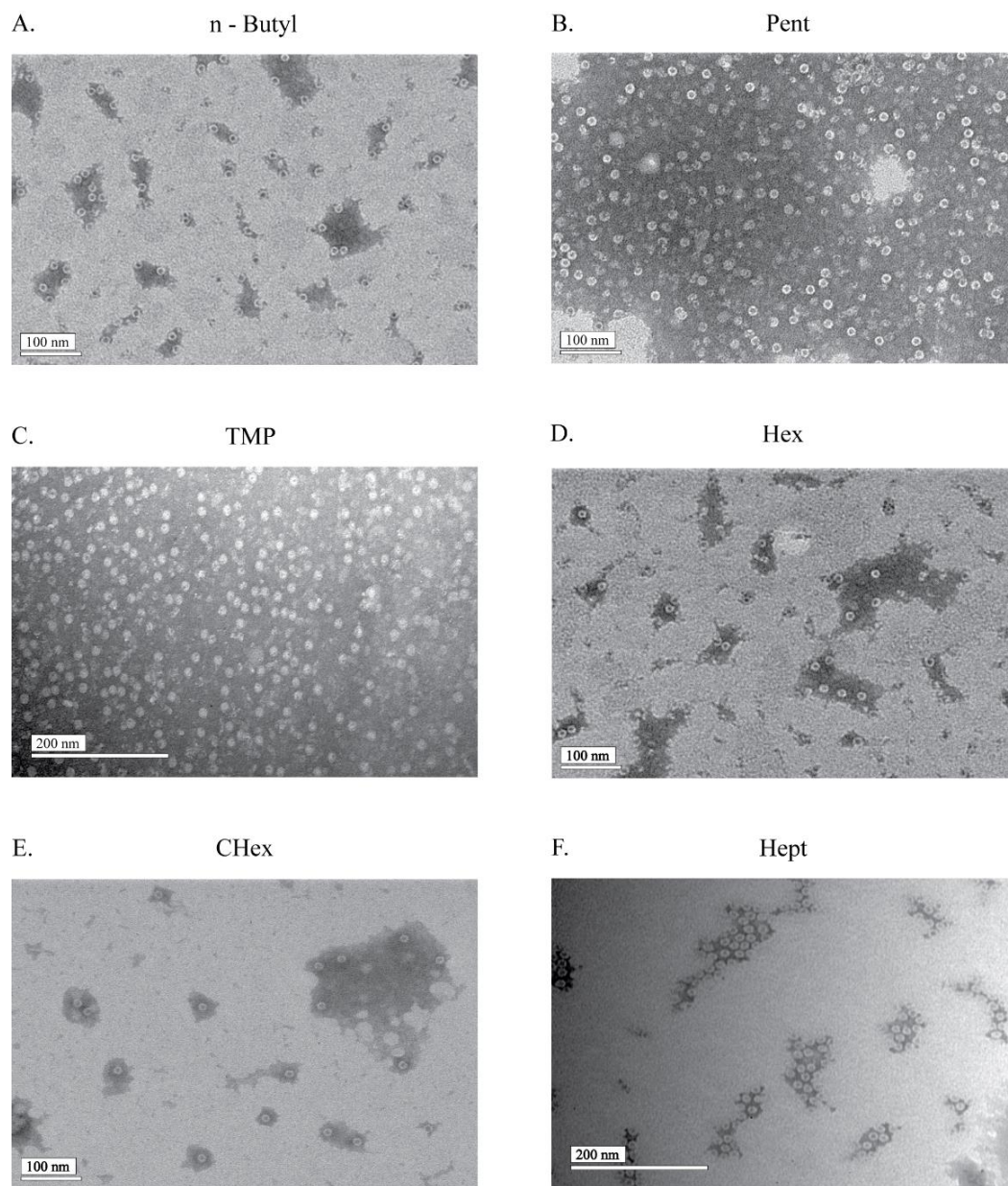
Figure 4.3. TEM images of assembled capsids. Scale bars are as indicated.

G.   Oct

H.   Ox

I.   mOx

J.   eOx

K.   2TP

L.   4TP

M.   Bz

Figure 4.3. Continued.

spectrometry was used to confirm the existence of the thiol adduct in samples showing capsid assembly (Table 4.2).

4.2.3. Attempted Disassembly of the Switch-Thiol Capsids via DTT Treatment

In an effort to generate a redox controlled assembly, several attempts were made to induce dissociation of the capsids back into pentamers via treatment with the reducing agent DTT. None of the capsids so treated showed quantitative dissociation of the capsid to pentamers in phosphate buffer (Figure 4.4). Several of the capsids showed minimal amounts of disassembly (Figure 4.4 c, d, e, i, j), but only the Switch-t-Butyl and Switch-eOx capsids showed significant conversion to pentamer when exposed to DTT in phosphate buffer (Figure 4.4 b, k). Dialysis of the thiophenol capsid into nanopure water followed by treatment with 10 mM DTT for 24 hours and subsequent dialysis back into phosphate buffer resulted in the complete disassembly of the capsid to pentamers (Figure 4.5). The nature of this disassembly suggests that the phosphate buffer may be interfering with the ability of DTT to free the thiol adduct. Therefore, we dialyzed various thiol capsids into citrate buffer (Figure 4.6) and attempted to once again disassemble them using DTT.

Upon initial dialysis into citrate buffer, both Switch-eOx and -2TP showed a broadened and delayed capsid peak suggesting an assembly state somewhere between pentamer and capsid (Figure 4.6 k, l). Surprisingly, Switch-TMP appeared to completely disassemble into pentamers (Figure 4.6 d). The pentamers elute at a significantly later volume in citrate buffer (110 mL) than they do in phosphate buffer (90 mL). DTT treatment in citrate buffer did not appear to have a significant effect on any of the tested capsids (data
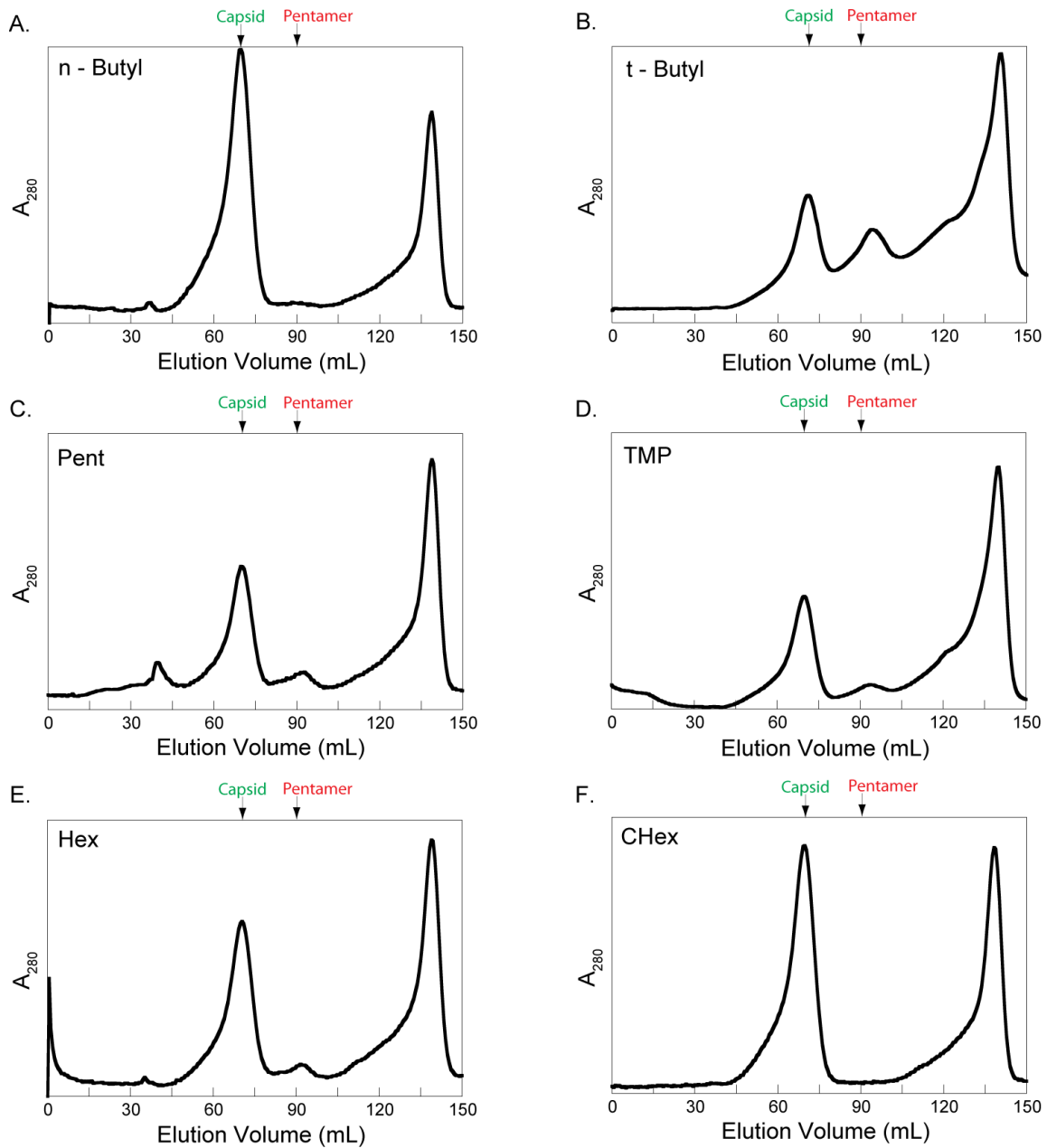
Figure 4.4 SEC chromatograms of capsid samples treated with DTT in phosphate buffer. Expected e-lution volumes are denoted for capsid (green) and pentamer (red).

Figure 4.4. Continued

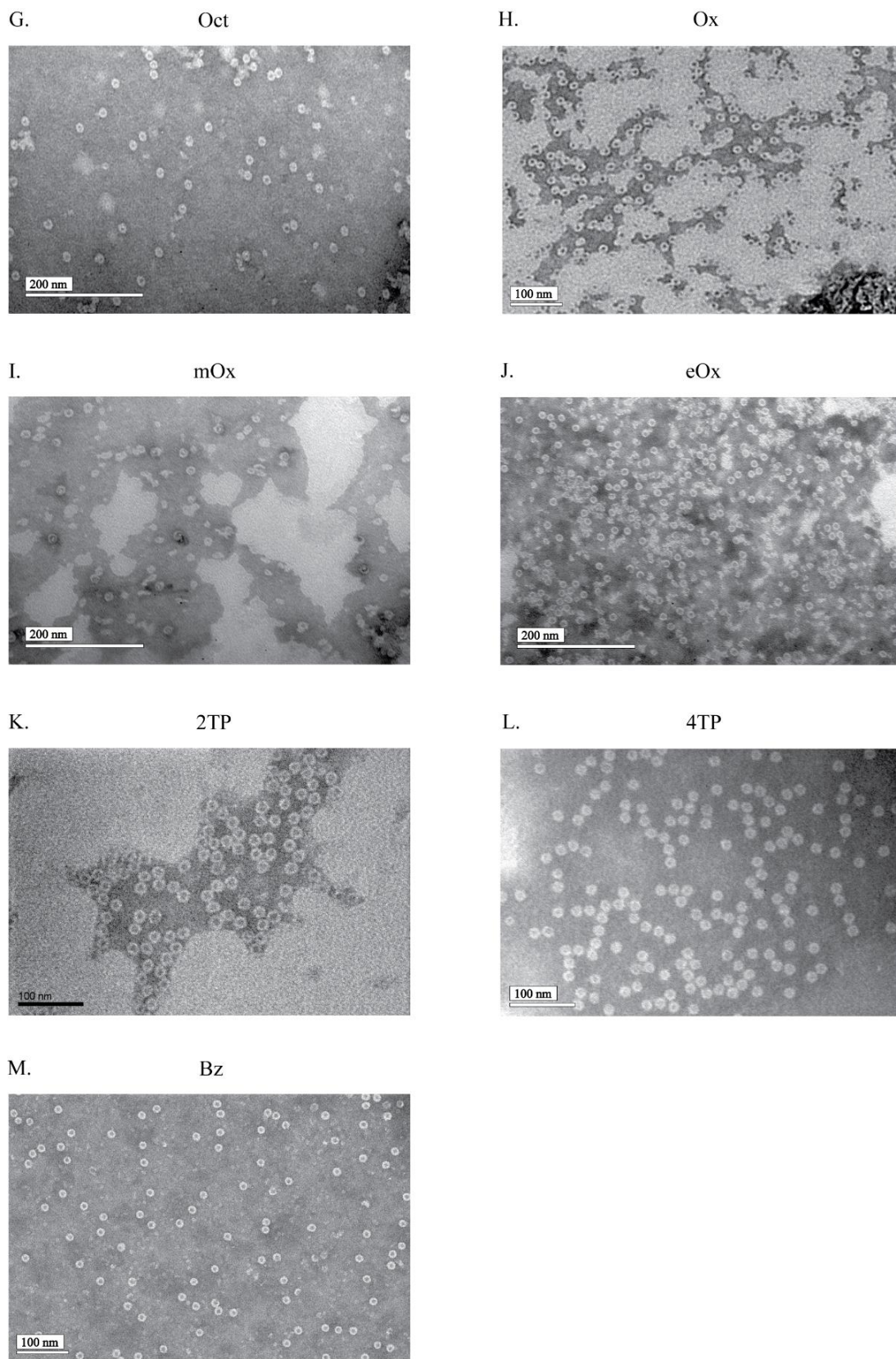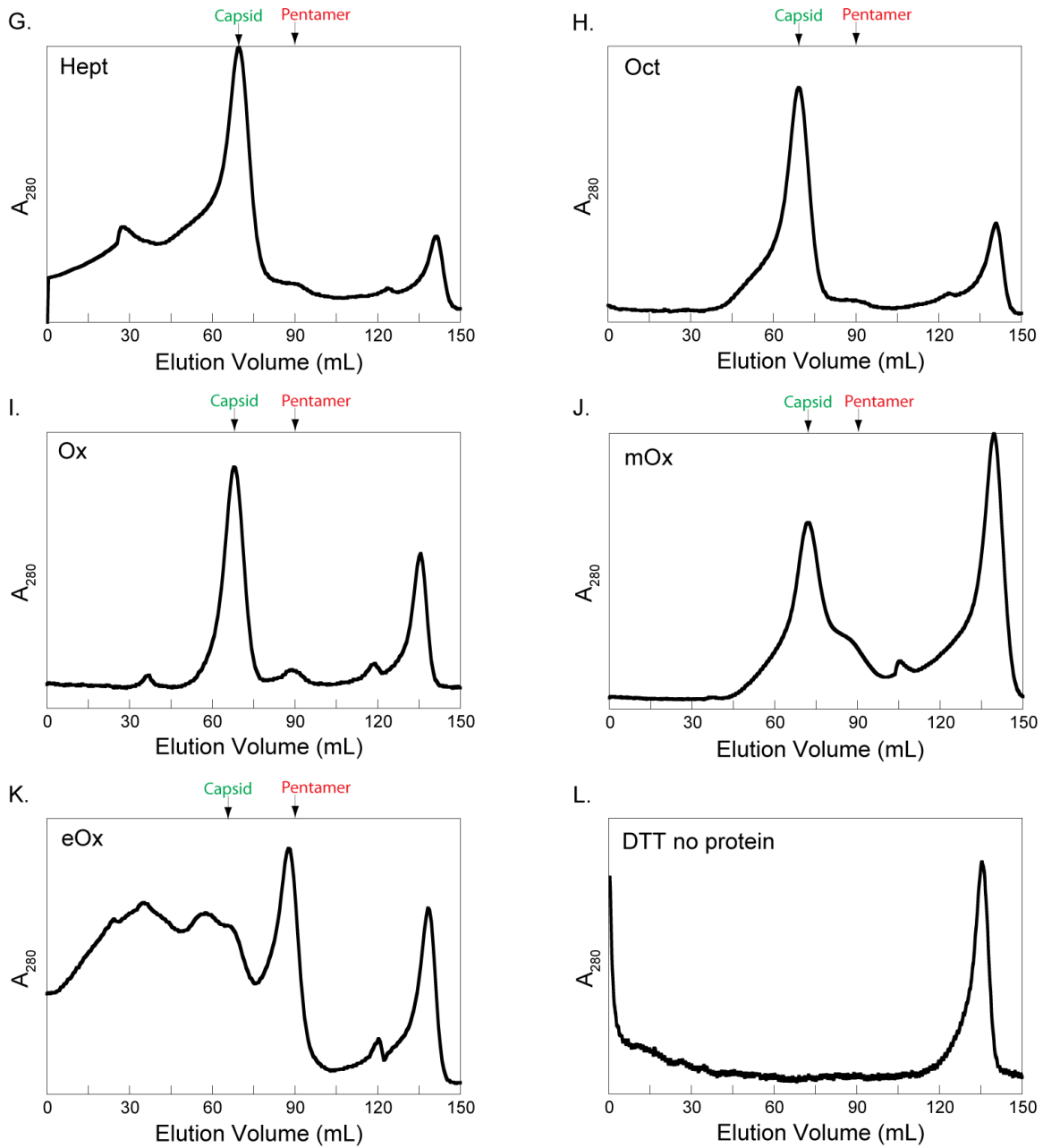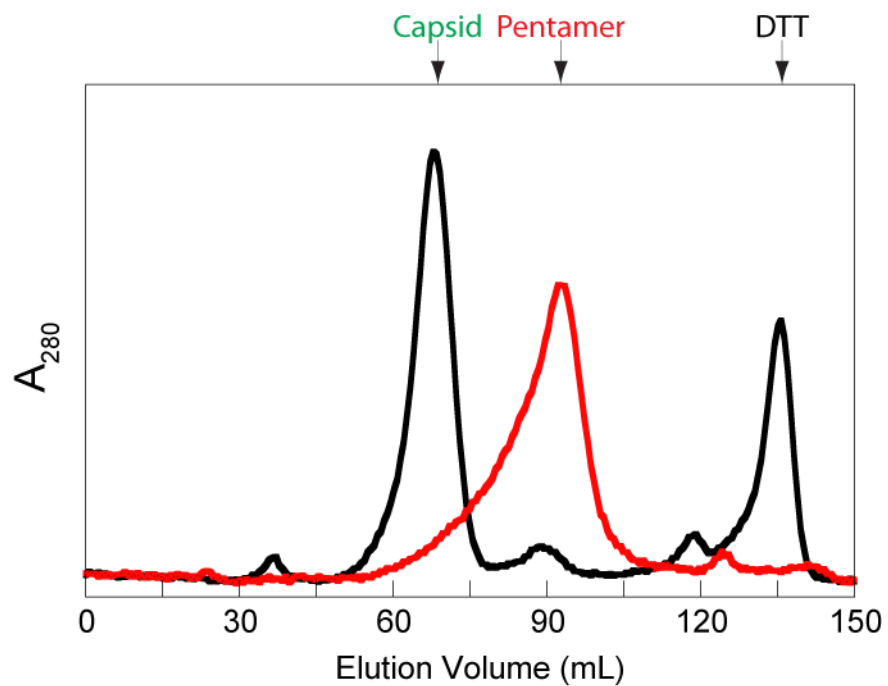Figure 4.5 Size analysis of DTT treated Switch-Ox. Switch-Ox capsid was either treated with DTT in phosphate buffer (black) or dialyzed into water, treated with DTT, then dialyzed back into phosphate buffer (red). Expected elution volumes are denoted for capsid (green) and pentamer (red).
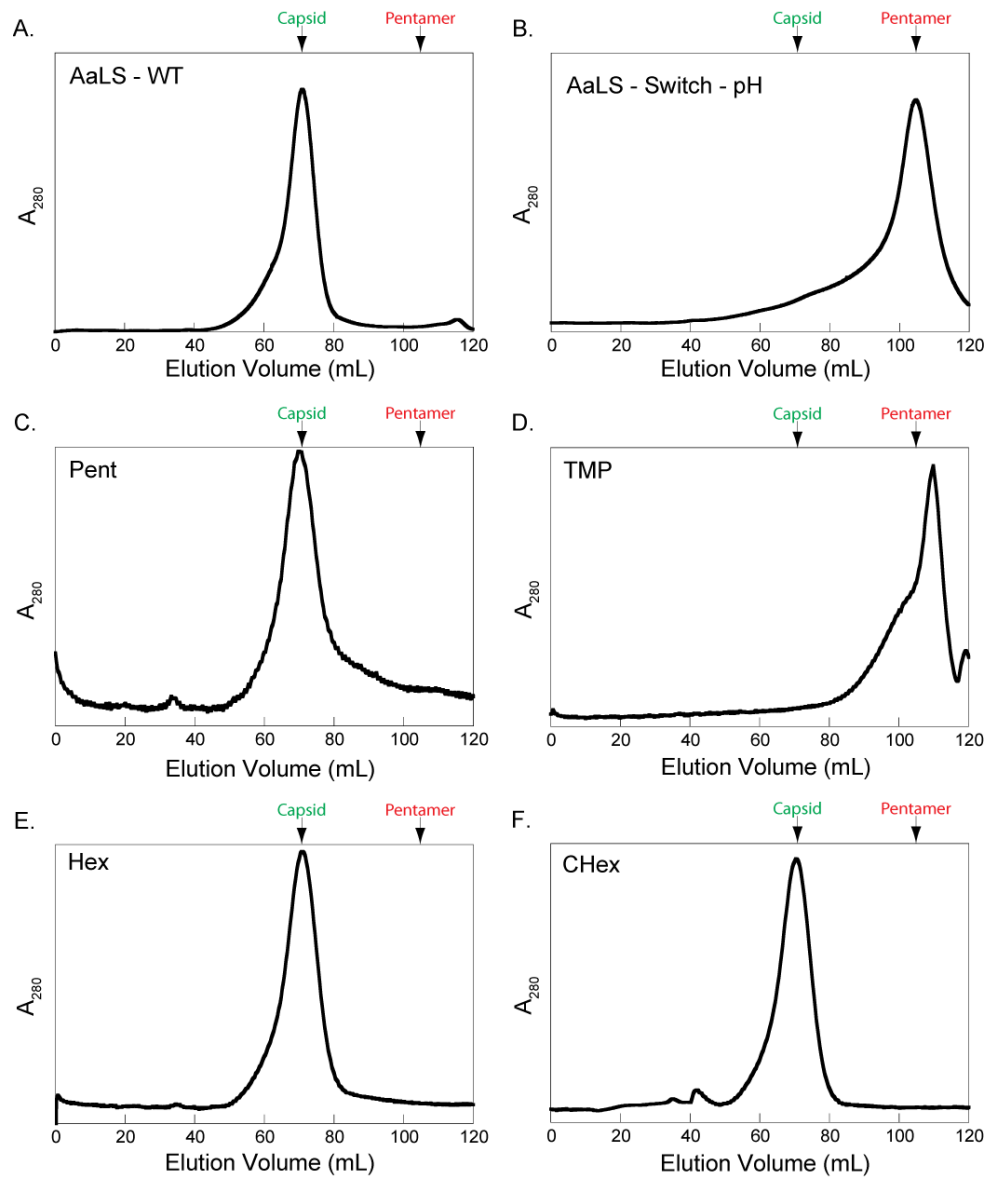
Figure 4.6. SEC chromatograms of capsid samples dialyzed into citrate buffer. AaLS-WT and AaLS-Switch-pH are shown as capsid and pentamer representatives, respectively. Switch-pH data collected by H. Chen. Expected elution volumes are denoted for capsid (green) and pentamer (red).
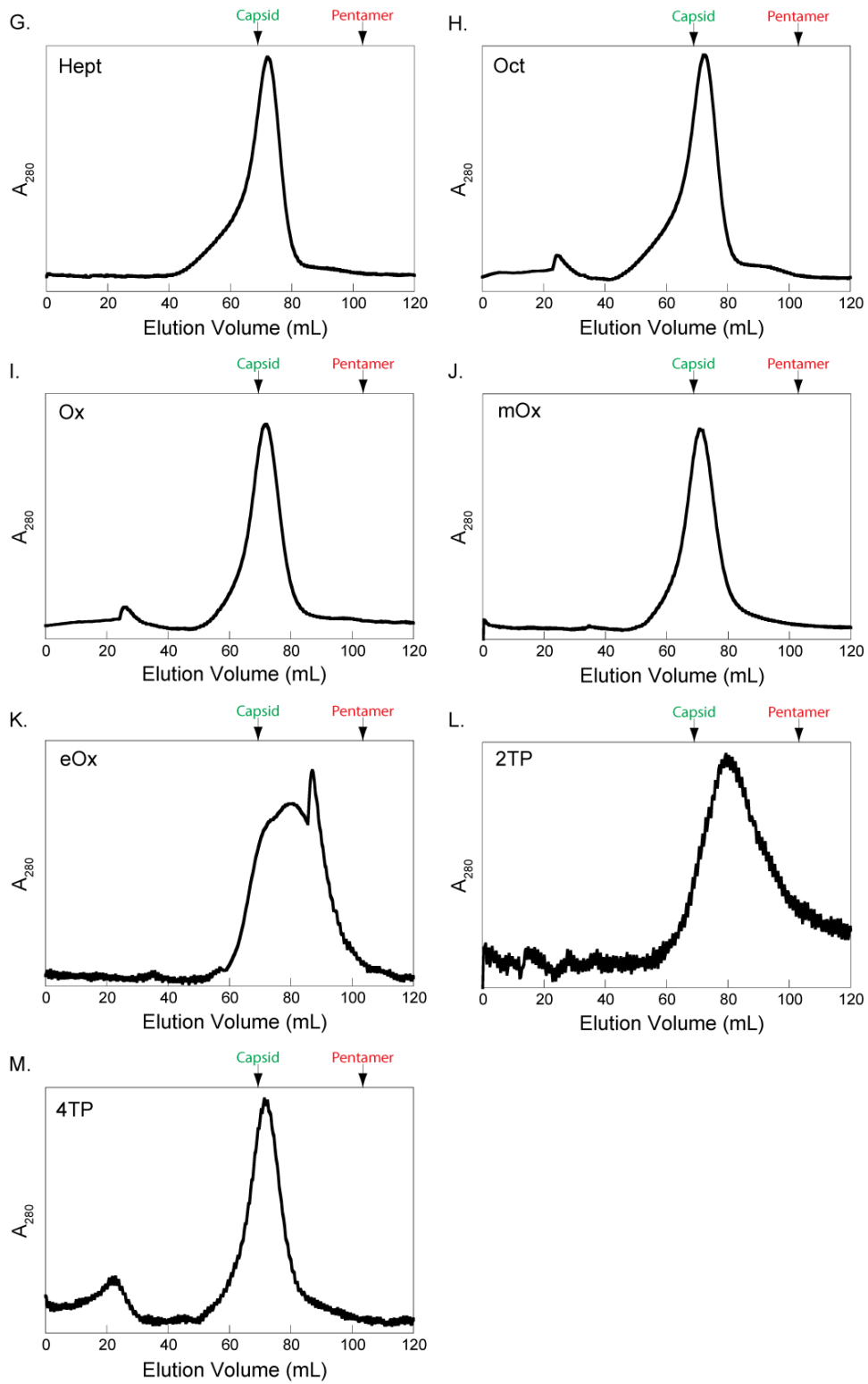
Figure 4.6. Continued.

not shown).

4.2.4 Development of a pH Inducible Disassembly Switch

Previous work has demonstrated that mutation of three residues at the three-fold symmetry axis to histidines (T120H, E122H, and Q123H), in addition to two mutations which destabilize the pentamer interfaces at the two-fold symmetry axis (R40S and H41S), results in a variant of AaLS which assembles into a capsid at pH 8 but reversibly disassembles to pentamers when the pH is lowered (AaLS-Switch-pH).[85] We attempted to mimic this approach by covalently bonding either 2- or 4-thiopyridine (2TP and 4TP, respectively) to the C125 residue. Thiopyridines were chosen for their structural similarity to thiophenol and their pK_a values. Formation of the 4TP adduct was the least efficient of all tested thiols followed by Switch-2TP. The thiopyridines also had the lowest yield of fully formed capsid.

4.2.5 Attempted Design of an Oxidative Switch for Capsid Disassembly

In an effort to make a redox switch which remains conjugated to the AaLS-Switch protein in both the pentameric and full capsid forms, a benzyl group was conjugated to C125 using benzyl bromide as the starting reagent. Since the benzyl group does not contain a sulfur atom, it is incapable of forming a disulfide bond. Instead, a bromine radical leaves which forms a benzyl radical that then attacks the cysteine thiol to form a thioether. This thioether may be reversibly oxidized to a sulfoxide. It was hoped that the steric repulsion caused by the addition of an oxygen to the thioether would be sufficient to induce disassembly of the assembled capsid.

The Switch-Bz adduct assembles into capsid like the majority of the tested thiols. The thioether was oxidized to a sulfoxide using hydrogen peroxide (H_2O_2). After a 24-hour treatment with H_2O_2, the sample and untreated control were analyzed via SEC (Figure 4.7). The traces for each were identical, showing no dissociation of the capsid due to H_2O_2 exposure.

Mass spectrometry analysis exhibited a large peak corresponding to the size expected if the two endogenous methionine sidechains and the cysteine adduct were oxidized to sulfoxides, demonstrating that the intended reaction did take place (Table 4.3). However, the second largest peak, corresponding to the mass of the unmodified Switch-Bz adduct, had an intensity of roughly half that of the major peak, suggesting that a significant amount of the available protein did not react. Interestingly, there were no peaks corresponding to the sizes expected for the reaction of only one or two of the available sulfides.

4.3. Discussion

A minimum adduct size of five nonhydrogen atoms (one sulfur and four carbons) were found to be necessary in order to induce capsid assembly. The reduced or NTB-adduct forms of the protein show no higher order structures beyond pentamers while conjugation to two or three carbon thiols resulted in a mixture of pentamers, intermediate aggregation states, and over-assembled aggregations. Use of linear thiol adducts ranging from four to eight carbons in length, two different branched chain thiols, or cyclic thiols containing six or more carbons all yielded capsids in similar sizes, suggesting that the three-fold symmetry axis is a highly plastic interface capable of adapting to a variety of structural
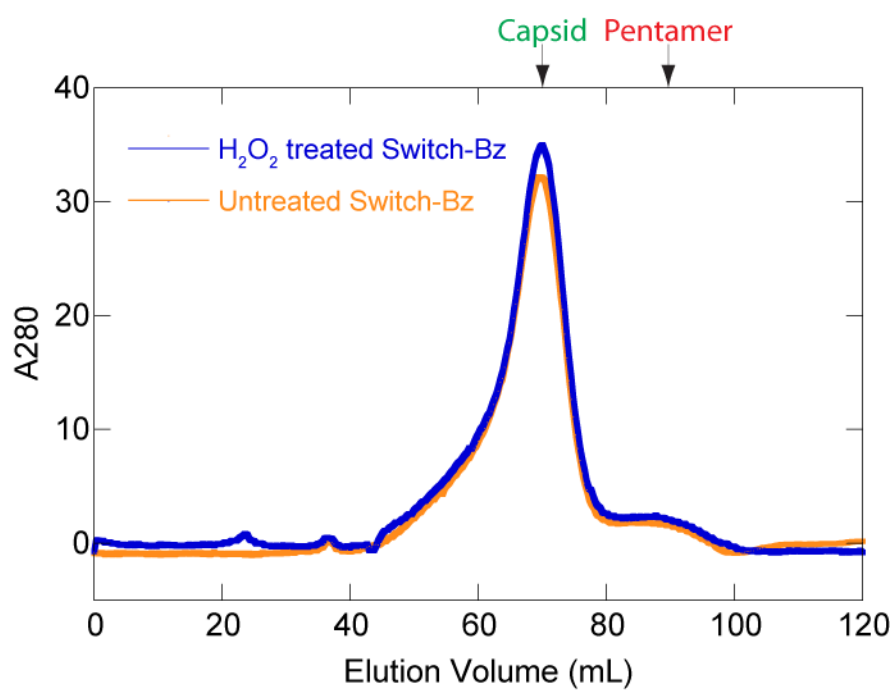
Figure 4.7. SEC chromatogram of untreated Switch-Bz capsid and the same capsid treated with hydrogen peroxide for 24 hours. Analysis by mass spectrometry indicates that the intended reaction took place but the SEC trace shows no difference compared to the control.

modifications.

In general, as the formation efficiency of the Switch-adduct complex increased, so did the capsid yield (Figure 4.8 b, d). At the highest levels of efficiency, there should be a greater number of pentamers in which all five cysteines have been modified and thus fewer assembly dead ends, which will translate into a higher yield of capsid. Linear regression suggests that the minimum adduct yield must be greater than 37% in order to form capsid (Figure 4.8 b). This corresponds to an average of 1.9 modified cysteines per pentamer. However, given the extremely low amounts of capsid obtained for adduct formation yields of less than 70% (an average of 3.5 modified cysteines per pentamer), this value may be a more realistic minimum. Precipitation was frequently observed during adduct formation. This loss of product could be exacerbating the uncertainty in this estimate by artificially deflating the minimum adduct yield. No clear trends emerged when examining adduct or capsid yields as a function of adduct molar mass (Figure 4.8 a, c, e).

In addition to improving our understanding of how the structure of the hydrophobic adduct influences capsid assembly, we also attempted to produce both a redox and a pH switch for control over said assembly. In pursuit of this goal, the various capsids were treated with DTT to see if it would be able to free the thiol adducts and thereby induce disassembly of the capsid. Unfortunately, this approach was not successful. However, the fact that small amounts of pentamer were observed for several of the treated capsids coupled with the fact that the thiophenol capsid was successfully disassembled in nanopure water with DTT means further study with varying combinations of reducing agents and buffers is warranted. The other attempt at a redox switchable capsid, through the use of
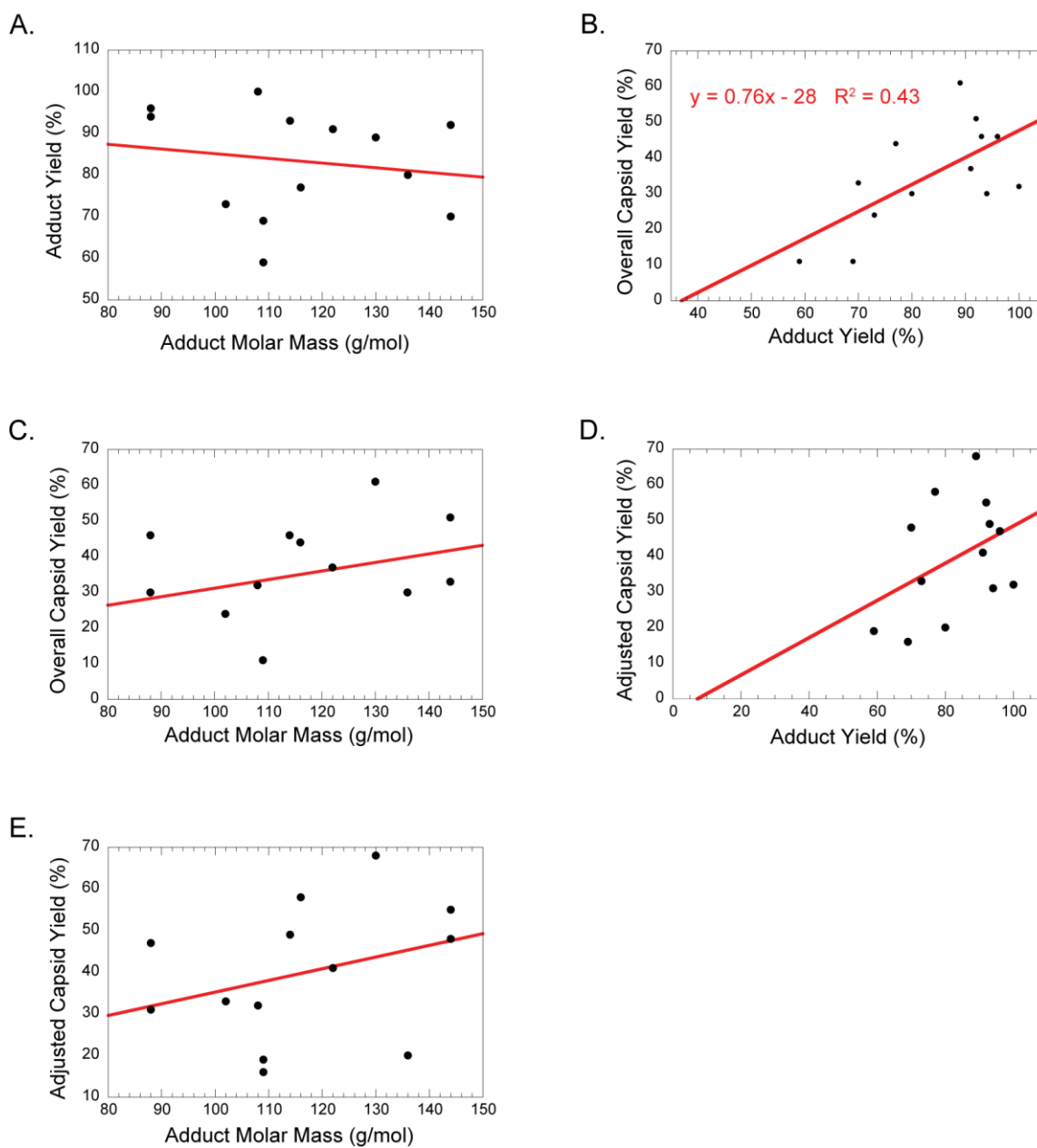
Figure 4.8. Analysis of adduct and capsid yields. No strong trends are evident when examining adduct or capsid yield as a function of adduct mass. A weak correlation between adduct yield and overall capsid yield was observed (B).

H_2O_2 to oxidize the thioether bond in the Switch-Bz variant, was unsuccessful and the evidence does not suggest that this approach is worth pursuing further.

Attempts to engineer a pH switch for control of capsid assembly using thiopyridines were also unsuccessful. However, the Switch-2TP capsid did exhibit unusual behavior in that it appears to partially disassemble in pH 5.7 citrate buffer. This may be useful if these lower order aggregates allow access to the capsid interior and/or dialysis back into phosphate buffer causes the capsid to reassemble. Alternatively, it is possible that the capsid has contracted in diameter without disassembling. Changes in capsid size based on pH have previously been observed in viral capsids and BsLS but not AaLS-WT.[190,191] Switch-eOx exhibited similar behavior to Switch-2TP so it may also be a potentially pH switchable capsid.

Finally, we found that the Switch-TMP capsid completely dissociated into pentamers upon dialysis into citrate buffer at pH 5.7. AaLS-Switch-pH has previously been shown to dissociate into pentamers reversibly, suggesting that the Switch-TMP pentamers may be able to reform capsid as well. In addition to pH, the identity of the buffer appears to play a role in the assembly state of the thiol-modified AaLS-Switch protein. Thus, exploring different buffers may lead to new redox- and pH-based assembly switches.

The work reported here expands our understanding of how to reengineer protein interfaces for controllable supramolecular assembly. The ability to efficiently pack a diverse set of hydrophobic sidechains, as demonstrated by the variety of thiol adducts which assemble into capsid, indicates the AaLS three-fold axis is a highly plastic interface capable of adjusting to significant alteration. This adaptability is further evinced by our success at engineering a new pH switch for capsid disassembly, Switch-TMP, and the

potential shown by several of the other variants. Additionally, our methods may serve as a generally applicable means of engineering assembly switches into other capsids with robust interfaces.

4.4. Materials and Methods

4.4.1. Materials

All cell culture media and chemical reagents were purchased from Bio-rad, Fisher Scientific, Gold Biotechnology, or Pierce Biotechnology and used without further purification. BL21 (DE3) and XL1-Blue *E. coli* cell strains and *Pfu*-turbo DNA polymerase were purchased from Stratagene. DTT was purchased from Research Products International. Butanethiol was purchased from Acros. All other thiol reagents where purchased from Sigma-Aldrich. All thiol reagents were used without further purification.

4.4.2. Production and Purification of AaLS-Switch-Red

Protein was produced in $CaCl_2$-competent BL21 (DE3) *E. coli* cells. Transformation of the gene-bearing plasmid (pMG-AaLSNoHis-switch-red[85]) was accomplished via heat shock at 42 °C for 2 minutes. The cells were then grown at 37 °C for 1 hour and plated on antibiotic-containing LB agar plates (100 µg/mL ampicillin). The plates were grown overnight in a 37 °C oven. A single colony was selected from the plate and used to inoculate a 7.5 mL AMP-containing LB culture which was then grown overnight at 37 °C with shaking at 250 rpm. An aliquot of this culture (1 mL) was then used to inoculate AMP-containing LB media (500 mL). Over-production was achieved by growing this large culture at 37 °C to an OD_{600} of 0.7, inducing production by adding IPTG

to a final concentration of 0.2 mM, and then incubating the culture at 30 °C for 20 hours at 250 rpm. The cells were harvested via centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (6000 rpm for 15 minutes at 4 °C). Pellets were frozen at -80 °C until used.

Cell pellets were resuspended in 10 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). The cells were lysed by incubation with lysozyme (10 mg), RNase A (1.2 mg), and DNase I (20 µg) on ice for 1 hour followed by sonication at 45 µm amplitude for 3 minutes in 10 second on/off pulses using a Misonix ultrasonic liquid processor. The lysate was then clarified by centrifugation in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor (12000 rpm for 45 minutes at 4 °C).

AaLS-Switch-Red was purified from the crude protein sample as previously described[85]. Briefly, the supernatant was heated for 5 minutes in a water bath at 60 C. The heated sample was then centrifuged in a Sorvall RC 6+ centrifuge with a Fiberlite F10 6x500y rotor at 12,000 rpm for 45 minutes at 4 °C. This supernatant was then dialyzed overnight into ion exchange buffer A at 4 °C. The freshly dialyzed sample was then loaded onto the FPLC equipped with an anion-exchange MonoQ 5/50 GL column which had been equilibrated with ion exchange buffer A. Bound protein was eluted using a gradient ranging from 0% to 100% ion exchange buffer B run at 4 °C. The flow-through fractions were then pooled and reinjected on the MonoQ column. The same gradient as the initial injection was used for elution. SDS-PAGE with a Thermo Scientific Unstained Molecular Weight Marker was then used to determine which fractions contained the pentamer. These fractions were pooled, concentrated, and then loaded onto a HiPrep 16/60 Sephacryl S-300HR column run with lysis buffer at 4 °C.

4.4.3. Preparation of AaLS-Switch-NTB from AaLS-Switch-Red

Purified AaLS-Switch-Red was concentrated to approximately 1 mg/mL in lysis buffer B (50 mM sodium phosphate, 200 mM NaCl, pH 8.0) and divided into five equal aliquots (typically 1-5 mL). The volume of a 20 mM 5,5'-dithiobis-(2-nitrobenzoate) (DTNB) solution in lysis buffer B needed to achieve a final concentration of 600 µM DTNB was calculated and divided into five even aliquots. One aliquot of the protein solution was dripped into one aliquot of the DTNB solution at a rate of approximately one drop/s with gentle stirring. This process was repeated for the other four aliquots of each solution and then all five were allowed to stir in the dark at room temperature for 1 hour. In order to remove excess DTNB, the aliquots were pooled and dialyzed into ion exchange buffer A over the course of three days with the buffer being refreshed once per day. After dialysis, the protein concentration was measured by Bradford assay.

4.4.4. Assessment of the Efficacy of NTB Conjugation

Ideally, since there is a single cysteine per monomer, treatment of AaLS-Switch-Red with DTNB should result in the quantitative formation of an AaLS-Switch-Red-2-nitro-5-thiobenzoate adduct (AaLS-Switch-NTB). NTB can be cleaved from the AaLS-Switch-NTB via treatment with dithiothreitol (DTT) and free NTB has an extinction coefficient of 14,150 $M^{-1}cm^{-1}$ at 412 nm; therefore, it is possible to colorimetrically assess the degree of NTBylation of AaLS-Switch-NTB.

An aliquot of 99 µL of AaLS-Switch-NTB solution was combined with 1 µL of 1 M DTT and allowed to equilibrate for 1 hour at room temperature. The absorbance at 412 nm of the resulting solution was measured and used to determine the concentration of

free NTB in solution. Typical batches gave a ratio of mol NTB per mol protein of 0.97 ± 1.1, indicating nearly quantitative yields of adduct formation.

4.4.5. Production of Other Thiol Adducts from AaLS-Switch-NTB

A disulfide exchange reaction was conducted in order to replace NTB with another thiol. A 10-fold excess of the thiol replacement was added to 4 mL of 1 mg/mL AaLS-Switch-NTB and incubated at room temperature with gentle stirring for 1 hour. Typically, the solution turned bright yellow immediately upon thiol addition. An assessment of the efficiency of NTB replacement was made by measuring the absorbance of the reaction mixture at 412 nm and calculating the moles of NTB freed per mole of monomer. The sample was then dialyzed into ion exchange buffer A to remove excess unreacted thiol and free NTB. Typical capsid yield was approximately 35% or 1.4 mg of protein from a standard starting batch of 4 mL of 1 mg/mL Switch-NTB (Table 4.1).

4.4.6. Assembly of AaLS-Switch-Thiol Capsids and SEC Analysis

of Assembly State

Upon removal from dialysis, the AaLS-Switch-Thiol samples were treated with 25% w/v PEG-3350 in lysis buffer B to a final PEG percentage of 10%. The mixture was then allowed to stir gently at room temperature, in the dark, for two days. At the end of this incubation, it was very common to see a small amount of fluffy, white precipitate present in the reaction mixture. The sample was then filtered over a 0.2 µm Whatman Puradisc filter and loaded into a 30 kDa MWCO Vivaspin concentrator. Removal of PEG-3350 from the sample was accomplished by repeated sessions of concentration via centrifugation with

a Fiberlite F10 6x500y rotor at 8,220 rpm for 10-20 minutes at 4 °C followed by dilution with buffer lacking PEG-3350. This cycle of concentration and dilution was maintained until the estimated remaining percentage of PEG-3350 was less than 0.01%. Higher concentrations of PEG-3350 (0.1-1%) were found to give anomalous SEC traces. Final volumes were typically 1 – 4 mL.

The assembly states of the various AaLS-Switch-Thiol samples were assessed using a HiPrep 16/60 Sephacryl S-400HR column injected with 1-4 mL of sample and run with lysis buffer B at 4 °C. Elution volumes of the samples were compared to the volumes of the AaLS-WT capsid and a pentameric variant, AaLS-R40S/H41S/I125S.

4.4.7. Transmission Electron Microscopy

In order to visualize the assembled capsid, 10 µL of the various AaLS-Switch-Thiol samples were applied to 200 square mesh, formvar coated copper grids. Samples were allowed to adhere to the grid for 1 minute before the excess was wicked away with sterile filter paper. The grids were then stained for 1 minute using 2% phosphotungstic acid (pH 8.0). Excess stain was wicked away with sterile filter paper. Sample images were obtained on a Hitachi 125 keV H-7100 Transmission Electron Microscope equipped with a Gatan Orius SC1000 slow scan 4kX2.6k CCD camera and analyzed using ImageJ software (National Institutes of Health).

4.4.8. Disassembly of AaLS-Switch-Thiol Adducts

In an effort to disassemble the AaLS-Switch-Thiol capsids, 1 mg/mL samples in lysis buffer B were treated with 10 mM DTT overnight at room temperature. The assembly

state of these capsids after DTT treatment was assessed as described in section 4.4.6. Attempts to disassemble the capsid by altering the buffer composition and pH were also conducted. A phosphate buffer was not used for low pH applications because the AaLS capsid has been shown to precipitate in phosphate buffer at a pH lower than 6.

Samples were dialyzed into citrate buffer (50 mM sodium citrate, 200 mM NaCl, pH 5.7) and their assembly state analyzed as described in section 4.4.7. Those still determined to be capsid where then treated with DTT as described above and their assembly state analyzed as described in section 4.4.6.

4.4.9. Formation of the AaLS-Switch-Bz Capsid

Since AaLS-Switch-Bz was to be formed via attack of the cysteine thiol on benzyl bromide, AaLS-Switch-NTB would not have been a productive starting reagent. Thus, AaLS-Switch-Red was used. A 10-fold excess of 0.4 M benzyl bromide (6 µL) was added to 4 mL of 1 mg/mL AaLS-Switch-Red and incubated at room temperature with gentle stirring for 1 hour. The reaction progress could not be colorimetrically monitored due to the lack of NTB. The sample was then dialyzed into ion exchange buffer A to remove excess unreacted benzyl bromide.

4.4.10 Hydrogen Peroxide Treatment of the AaLS-Switch-Bz Capsid

A 2 M H_2O_2 stock solution in lysis buffer B was prepared and used immediately. This was added to a 0.35 mg/mL solution of AaLS-Switch-Bz to a final concentration of 10 mM. The reaction was incubated for 24 hours, at room temperature, in the dark, with gentle stirring. The reaction was quenched by adding methionine to a final concentration

of 20 mM. The sample was then immediately injected onto a HiPrep 16/60 Sephacryl S-400HR column run with lysis buffer B at 4 °C.

4.4.11 Mass Spectrometry of the Switch-Thiol Adducts

Mass spectrometry was performed by the University of Utah Chemistry Department Mass Spectrometry Lab using a Waters LCT XE Premier ToF mass spectrometer. Samples were extensively dialyzed into nanopure water over the course of three days prior to being submitted for analysis.

4.4.12 Treatment of Switch-Ox with DTT in Water

A 1 mg/mL sample of AaLS-Switch-Ox was dialyzed into nanopure water overnight at 4 °C with gentle stirring, in the dark. After dialysis, 1 M DTT was added to the sample to a final concentration of 10 mM. The sample was incubated at room temperature overnight. The DTT treated sample was then dialyzed into lysis buffer B overnight at 4 °C with gentle stirring, in the dark. The assembly state of the AaLS-Switch-Ox capsid was then assessed as described in section 4.4.6.

REFERENCES

(1) Strale, P.-O.; Duchesne, L.; Peyret, G.; Montel, L.; Nguyen, T.; Png, E.; Tampe, R.; Troyanovsky, S.; Henon, S.; Ladoux, B.; Mege, R.-M. *J. Cell Biol.* **2015**, *210* (2), 333–346.

(2) Umbreit, N. T.; Miller, M. P.; Tien, J. F.; Ortolá, J. C.; Gui, L.; Lee, K. K.; Biggins, S.; Asbury, C. L.; Davis, T. N. *Nat. Commun.* **2014**, *5*, 4951.

(3) Miyata, T.; Yamada, K.; Iwasaki, H.; Shinagawa, H.; Morikawa, K.; Mayanagi, K. *J. Struct. Biol.* **2000**, *131* (2), 83–89.

(4) Yee, C. S. K.; Sanal, O.; Chou, J. S.; Geha, R. S.; Ayvaz, D.; Aytekin, C.; Akarsu, A. N. *J. Allergy Clin. Immunol.* **2014**, *133* (2), AB94.

(5) Westphal, D.; Kluck, R. M.; Dewson, G. *Cell Death Differ.* **2014**, *21* (2), 196–205.

(6) Schomburg, I.; Chang, A.; Ebeling, C.; Gremse, M.; Heldt, C.; Huhn, G.; Schomburg, D. *Nucleic Acids Res.* **2004**, *32* (Database issue), D431–D433.

(7) Mateu, M. G. *J. Mol. Biol.* **2002**, *318* (2), 519–531.

(8) Nacken, W.; Kerkhoff, C. *FEBS Lett.* **2007**, *581* (26), 5127–5130.

(9) Zeiler, E.; List, A.; Alte, F.; Gersch, M.; Wachtel, R.; Poreba, M.; Drag, M.; Groll, M.; Sieber, S. A. *Proc. Natl. Acad. Sci.* **2013**, *110* (28), 11302–11307.

(10) Nishi, H.; Hashimoto, K.; Madej, T.; Panchenko, A. R. *Prog. Mol. Biol. Transl. Sci.* **2013**, *117*, 3–24.

(11) Goodsell, D. S.; Olson, A. J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153.

(12) Dennis, J. W. *Trends Biochem. Sci.* **2015**.

(13) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (1), 13–20.

(14) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280* (1), 1–9.

(15) Yan, C.; Wu, F.; Jernigan, R. L.; Dobbs, D.; Honavar, V. *Protein J.* **2008**, *27* (1), 59–70

(16) Lo Conte, L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285* (5), 2177–2198.

(17) Bahadur, R. P.; Chakrabarti, P.; Rodier, F.; Janin, J. *Proteins* **2003**, *53* (3), 708–719.

(18) Chen, J.; Sawyer, N.; Regan, L. *Protein Sci.* **2013**, *22* (4), 510–515.

(19) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. B* **2007**, *111* (10), 2697–2706.

(20) Keskin, O.; Ma, B.; Nussinov, R. *J. Mol. Biol.* **2005**, *345* (5), 1281–1294.

(21) Bahadur, R. P.; Rodier, F.; Janin, J. *J. Mol. Biol.* **2007**, *367* (2), 574–590.

(22) Janin, J.; Bahadur, R. P.; Chakrabarti, P. *Q. Rev. Biophys.* **2008**, *41* (2), 133–180.

(23) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204* (1), 155–164.

(24) Janin, J. *Biochimie* **1995**, *77* (7-8), 497–505.

(25) Xu, D.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **1997**, *10* (9), 999–1012.

(26) Glaser, F.; Steinberg, D. M.; Vakser, I. A.; Ben-Tal, N. *Proteins* **2001**, *43* (2), 89–102.

(27) Mintseris, J.; Weng, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (31), 10930–10935.

(28) Ansari, S.; Helms, V. *Proteins* **2005**, *61* (2), 344–355.

(29) Mateu, M. G. *Arch. Biochem. Biophys.* **2013**, *531* (1-2), 65–79.

(30) Caspar, D. L.; Klug, A. *Cold Spring Harb. Symp. Quant. Biol.* **1962**, *27*, 1–24.

(31) Allan, G. M.; McNeilly, F.; Cassidy, J. P.; Reilly, G. A. C.; Adair, B.; Ellis, W. A.; McNulty, M. S. *Vet. Microbiol.* **1995**, *44* (1), 49–64.

(32) Xiao, C.; Chipman, P. R.; Battisti, A. J.; Bowman, V. D.; Renesto, P.; Raoult, D.; Rossmann, M. G. *J. Mol. Biol.* **2005**, *353* (3), 493–496.

(33) Bayer, T. S. *RNA* **2005**, *11* (12), 1848–1857.

(34) Speir, J. a; Munshi, S.; Wang, G.; Baker, T. S.; Johnson, J. E. *Structure* **1995**, *3* (1), 63–78.

(35) Valegård, K.; Liljas, L.; Fridborg, K.; Unge, T. *Nature* **1990**, *345* (6270), 36–41.

(36) Golmohammadi, R.; Fridborg, K.; Bundule, M.; Valegård, K.; Liljas, L. *Structure* **1996**, *4* (5), 543–554.

(37) King, J.; Botstein, D.; Casjens, S.; Earnshaw, W.; Harrison, S.; Lenk, E. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1976**, *276* (943), 37–49.

(38) Lavelle, L.; Michel, J.-P.; Gingery, M. *J. Virol. Methods* **2007**, *146* (1-2), 311–316.

(39) Chowdhury, C.; Sinha, S.; Chun, S.; Yeates, T. O.; Bobik, T. a. *Microbiol. Mol. Biol. Rev.* **2014**, *78* (3), 438–468.

(40) Andersson, I.; Backlund, A. *Plant Physiol. Biochem.* **2008**, *46* (3), 275–291.

(41) Savir, Y.; Noor, E.; Milo, R.; Tlusty, T. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (8), 3475–3480.

(42) Dou, Z.; Heinhorst, S.; Williams, E. B.; Murin, C. D.; Shively, J. M.; Cannon, G. C. *J. Biol. Chem.* **2008**, *283* (16), 10377–10384.

(43) Cai, F.; Menon, B. B.; Cannon, G. C.; Curry, K. J.; Shively, J. M.; Heinhorst, S. *PLoS One* **2009**, *4* (10), e7521.

(44) Andrews, S. C.; Arosio, P.; Bottke, W.; Briat, J. F.; von Darl, M.; Harrison, P. M.; Laulhère, J. P.; Levi, S.; Lobreaux, S.; Yewdall, S. J. *J. Inorg. Biochem.* *47* (3-4), 161–174.

(45) Ong, D. S. T.; Wang, L.; Zhu, Y.; Ho, B.; Ding, J. L. *J. Endotoxin Res.* **2005**, *11* (5), 267–280.

(46) Izard, T.; Aevarsson, A.; Allen, M. D.; Westphal, A. H.; Perham, R. N.; de Kok, A.; Hol, W. G. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (4), 1240–1245.

(47) Vijayakrishnan, S.; Kelly, S. M.; Gilbert, R. J. C.; Callow, P.; Bhella, D.; Forsyth, T.; Lindsay, J. G.; Byron, O. *J. Mol. Biol.* **2010**, *399* (1), 71–93.

(48) Ladenstein, R.; Fischer, M.; Bacher, A. *FEBS J.* **2013**, *280* (11), 2537–2563.

(49) Ritsert, K.; Huber, R.; Turk, D.; Ladenstein, R.; Schmidt-Bäse, K.; Bacher, A. *J. Mol. Biol.* **1995**, *253* (1), 151–167.

(50) Kis, K.; Bacher, A. *Journal of Biological Chemistry*. 1995, 16788–16795.

(51) Patel, M. S.; Nemeria, N. S.; Furey, W.; Jordan, F. *J. Biol. Chem.* **2014**, *289*, 16615–16623.

(52) Weng, J.-H.; Hsieh, Y.-C.; Huang, C.-C. F.; Wei, T.-Y. W.; Lim, L.-H.; Chen, Y.-H.; Ho, M.-R.; Wang, I.; Huang, K.-F.; Chen, C.-J.; Tsai, M.-D. *Biochemistry* **2015**, 151001070035009.

(53) Morlacchi, P.; Robertson, F. M.; Klostergaard, J.; McMurray, J. S. *Future Med. Chem.* **2014**, *6* (17), 1909–1926.

(54) Yang, J.; Srinivasan, A.; Sun, Y.; Mrazek, J.; Shu, Z.; Kickhoefer, V. A.; Rome, L. H. *Integr. Biol.* **2013**, *5* (1), 151–158.

(55) Lee, H.-J.; Zheng, J. J. *Cell Commun. Signal.* **2010**, *8*, 8.

(56) Cortines, J. R.; Weigele, P. R.; Gilcrease, E. B.; Casjens, S. R.; Teschke, C. M. *Virology* **2011**, *421* (1), 1–11.

(57) Singer, G. P.; Newcomb, W. W.; Thomsen, D. R.; Homa, F. L.; Brown, J. C. *J. Virol.* **2004**, *79* (1), 132–139.

(58) Sutter, M.; Boehringer, D.; Gutmann, S.; Günther, S.; Prangishvili, D.; Loessner, M. J.; Stetter, K. O.; Weber-Ban, E.; Ban, N. *Nat. Struct. Mol. Biol.* **2008**, *15* (9), 939–947.

(59) Kickhoefer, V. A.; Garcia, Y.; Mikyas, Y.; Johansson, E.; Zhou, J. C.; Raval-Fernandes, S.; Minoofar, P.; Zink, J. I.; Dunn, B.; Stewart, P. L.; Rome, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (12), 4348–4352.

(60) Fiedler, J. D.; Brown, S. D.; Lau, J. L.; Finn, M. G. *Angew. Chemie - Int. Ed.* **2010**, *49* (50), 9648–9651.

(61) Wu, M.; Brown, W. L.; Stockley, P. G. *Bioconjug. Chem.* **1995**, *6* (5), 587–595.

(62) Glasgow, J. E.; Capehart, S. L.; Francis, M. B.; Tullman-Ercek, D. *ACS Nano* **2012**, *6* (10), 8658–8664.

(63) Seebeck, F. P.; Woycechowsky, K. J.; Zhuang, W.; Rabe, J. P.; Hilvert, D. *J. Am. Chem. Soc.* **2006**, *128* (14), 4516–4517.

(64) Wörsdörfer, B.; Pianowski, Z.; Hilvert, D. *J. Am. Chem. Soc.* **2012**, *134* (2), 909–911.

(65) Beck, T.; Tetter, S.; Künzle, M.; Hilvert, D. *Angew. Chemie Int. Ed.* **2015**, *54* (3), 937–940.

(66) Hooker, J. M.; Datta, A.; Botta, M.; Raymond, K. N.; Francis, M. B. *Nano Lett.* **2007**, *7* (8), 2207–2210.

(67) Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *Angew. Chemie Int. Ed.* **2009**, *48* (50), 9493–9497.

(68) Abedin, M. J.; Liepold, L.; Suci, P.; Young, M.; Douglas, T. *J. Am. Chem. Soc.* **2009**, *131* (12), 4346–4354.

(69) Hovlid, M. L.; Lau, J. L.; Breitenkamp, K.; Higginson, C. J.; Laufer, B.; Manchester, M.; Finn, M. G. *ACS Nano* **2014**, No. Xx, 8003–8014.

(70) Lucon, J.; Qazi, S.; Uchida, M.; Bedwell, G. J.; LaFrance, B.; Prevelige, P. E.; Douglas, T. *Nat. Chem.* **2012**, *4* (10), 781–788.

(71) Minten, I. J.; Ma, Y.; Hempenius, M. A.; Vancso, G. J.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Org. Biomol. Chem.* **2009**, *7* (22), 4685–4688.

(72) Ceci, P.; Ilari, A.; Falvo, E.; Giangiacomo, L.; Chiancone, E. *J. Biol. Chem.* **2005**, *280* (41), 34776–34785.

(73) Kim, M.; Rho, Y.; Jin, K. S.; Ahn, B.; Jung, S.; Kim, H.; Ree, M. *Biomacromolecules* **2011**, *12* (5), 1629–1640.

(74) Ellard, F. M.; Drew, J.; Blakemore, W. E.; Stuart, D. I.; King, A. M. *J. Gen. Virol.* **1999**, *80 (Pt 8)*, 1911–1918.

(75) Wingfield, P. T.; Stahl, S. J.; Williams, R. W.; Steven, A. C. *Biochemistry* **1995**, *34* (15), 4919–4932.

(76) Salunke, D. M.; Caspar, D. L.; Garcea, R. L. *Cell* **1986**, *46* (6), 895–904.

(77) Chuan, Y. P.; Fan, Y. Y.; Lua, L. H. L.; Middelberg, A. P. J. *J. R. Soc. Interface* **2010**, *7* (44), 409–421.

(78) Kanesashi, S. -n. *J. Gen. Virol.* **2003**, *84* (7), 1899–1905.

(79) Kawano, M.; Inoue, T.; Tsukamoto, H.; Takaya, T.; Enomoto, T.; Takahashi, R.; Yokoyama, N.; Yamamoto, N.; Nakanishi, A.; Imai, T.; Wada, T.; Kataoka, K.; Handa, H. *J. Biol. Chem.* **2006**, *281* (15), 10164–10173.

(80) Huard, D. J. E.; Kane, K. M.; Tezcan, F. A. *Nat. Chem. Biol.* **2013**, *9* (3), 169–176.

(81) Minten, I. J.; Wilke, K. D. M.; Hendriks, L. J. A.; van Hest, J. C. M.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Small* **2011**, *7* (7), 911–919.

(82) Van Eldijk, M. B.; Wang, J. C.-Y.; Minten, I. J.; Li, C.; Zlotnick, A.; Nolte, R. J. M.; Cornelissen, J. J. L. M.; van Hest, J. C. M. *J. Am. Chem. Soc.* **2012**, *134* (45), 18506–18509.

(83) Van Eldijk, M. B.; McGann, C. L.; Kiick, K. L.; van Hest, J. C. M. *Elastomeric Polypeptides*; Deming, T., Ed.; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; Vol. 310.

(84) Dalmau, M.; Lim, S.; Wang, S.-W. *Biomacromolecules* **2009**, *10* (12), 3199–3206.

(85) Chen, H.-N. Deconstruction and Reconstruction of a Protein Capsid. Ph.D. Dissertation [Online], The University of Utah, Salt Lake City, UT, August **2014**. http://content.lib.utah.edu/cdm/ref/collection/etd3/id/3139 (accessed Oct 15, 2015)

(86) Kurre, P.; Morris, J.; Miller, A. D.; Kiem, H. P. *Gene Ther.* **2001**, *8* (8), 593–599.

(87) Fay, N.; PantÃ©, N. *Front. Microbiol.* **2015**, *6*, 467.

(88) Yong, S.; Scott, J. L.; Stahelin, R. V. *The FASEB Journal;* **2013**, *27*.

(89) Mortola, E.; Noad, R.; Roy, P. *J. Virol.* **2004**, *78* (6), 2875–2883.

(90) Nason, E. L.; Rothagel, R.; Mukherjee, S. K.; Kar, A. K.; Forzan, M.; Prasad, B. V. V.; Roy, P. *J. Virol.* **2004**, *78* (15), 8059–8067.

(91) Kim, E.; Lee, S.; Mian, M. F.; Yun, S. U.; Song, M.; Yi, K.-S.; Ryu, S. H.; Suh, P.-G. *FEBS J.* **2006**, *273* (4), 793–804.

(92) Brautigam, C. A.; Wynn, R. M.; Chuang, J. L.; Chuang, D. T. *J. Biol. Chem.* **2009**, *284* (19), 13086–13098.

(93) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2005**, *127* (11), 3718–3723.

(94) Zhao, X.; Chen, L.; Luckanagul, J. A.; Zhang, X.; Lin, Y.; Wang, Q. *Chembiochem* **2015**, *16* (9), 1279–1283.

(95) Wang, Q.; Raja, K. S.; Janda, K. D.; Lin, T.; Finn, M. G. *Bioconjug. Chem. 14* (1), 38–43.

(96) Kovacs, E. W.; Hooker, J. M.; Romanini, D. W.; Holder, P. G.; Berry, K. E.; Francis, M. B. *Bioconjug. Chem. 18* (4), 1140–1147.

(97) O'Riordan, C. R.; Lachapelle, A.; Delgado, C.; Parkes, V.; Wadsworth, S. C.; Smith, A. E.; Francis, G. E. *Hum. Gene Ther.* **1999**, *10* (8), 1349–1358.

(98) Steinmetz, N. F.; Manchester, M. *Biomacromolecules* **2009**, *10* (4), 784–792.

(99) Prasuhn, D. E.; Kuzelka, J.; Strable, E.; Udit, A. K.; Cho, S.-H.; Lander, G. C.; Quispe, J. D.; Diers, J. R.; Bocian, D. F.; Potter, C.; Carragher, B.; Finn, M. G. *Chem. Biol.* **2008**, *15* (5), 513–519.

(100) Servid, A.; Jordan, P.; O'Neil, A.; Prevelige, P.; Douglas, T. *Biomacromolecules* **2013**, *14* (9), 2989–2995.

(101) Schmidt, U.; Rudolph, R.; Bohm, G. *Protein Eng. Des. Sel.* **2001**, *14* (10), 769–774.

(102) Schwarz, B.; Madden, P.; Avera, J.; Gordon, B.; Larson, K.; Miettinen, H. M.; Uchida, M.; LaFrance, B.; Basu, G.; Rynda-Apple, A.; Douglas, T. *ACS Nano* **2015**, *9* (9), 9134–9147.

(103) Obermeyer, A. C.; Capehart, S. L.; Jarman, J. B.; Francis, M. B. *PLoS One* **2014**, *9* (6), e100678.

(104) Shukla, S.; Ablack, A. L.; Wen, A. M.; Lee, K. L.; Lewis, J. D.; Steinmetz, N. F. *Mol. Pharm.* **2013**, *10* (1), 33–42.

(105) Koho, T.; Ihalainen, T. O.; Stark, M.; Uusi-Kerttula, H.; Wieneke, R.; Rahikainen, R.; Blazevic, V.; Marjomäki, V.; Tampé, R.; Kulomaa, M. S.; Hytönen, V. P. *Eur. J. Pharm. Biopharm.* **2015**.

(106) Aljabali, A. A. A.; Shukla, S.; Lomonossoff, G. P.; Steinmetz, N. F.; Evans, D. J. *Mol. Pharm.* **2013**, *10* (1), 3–10.

(107) Galaway, F. A.; Stockley, P. G. *Mol. Pharm.* **2013**, *10* (1), 59–68.

(108) Boman, N. L.; Bally, M. B.; Cullis, P. R.; Mayer, L. D.; Webb, M. S. *J. Liposome Res.* **2008**.

(109) Sadhukha, T.; Prabha, S. *AAPS PharmSciTech* **2014**, *15* (4), 1029–1038.

(110) Frank, L. A.; Contri, R. V.; Beck, R. C. R.; Pohlmann, A. R.; Guterres, S. S. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **2015**, *7* (5), 623–639.

(111) Valero, E.; Fiorini, S.; Tambalo, S.; Busquier, H.; Callejas-Fernández, J.; Marzola, P.; Gálvez, N.; Domínguez-Vera, J. M. *J. Med. Chem.* **2014**, *57* (13), 5686–5692.

(112) Tyler, M.; Tumban, E.; Peabody, D. S.; Chackerian, B. *Biotechnol. Bioeng.* **2014**, *111* (12), 2398–2406.

(113) Wang, G.; Cao, R.-Y.; Chen, R.; Mo, L.; Han, J.-F.; Wang, X.; Xu, X.; Jiang, T.; Deng, Y.-Q.; Lyu, K.; Zhu, S.-Y.; Qin, E.-D.; Tang, R.; Qin, C.-F. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (19), 7619–7624.

(114) Rodríguez-Limas, W. a; Sekar, K.; Tyo, K. E. *Curr. Opin. Biotechnol.* **2013**, *24* (6), 1089–1093.

(115) Gu, L.; Li, Z. C.; Krendelchtchikov, A.; Krendelchtchikova, V.; Wu, H.; Matthews, Q. L. *PLoS One* **2013**, *8* (3), e60347.

(116) Mizutani, M.; Iyori, M.; Blagborough, A. M.; Fukumoto, S.; Funatsu, T.; Sinden, R. E.; Yoshida, S. *Infect. Immun.* **2014**, *82* (10), 4348–4357.

(117) Lee, Y. J.; Yi, H.; Kim, W.-J.; Kang, K.; Yun, D. S.; Strano, M. S.; Ceder, G.; Belcher, A. M. *Science* **2009**, *324* (5930), 1051–1055.

(118) Lee, S.-Y.; Royston, E.; Culver, J. N.; Harris, M. T. *Nanotechnology* **2005**, *16* (7), S435–S441.

(119) Young, M.; Willits, D.; Uchida, M.; Douglas, T. *Annu. Rev. Phytopathol.* **2008**, *46*, 361–384.

(120) He, D.; Marles-Wright, J. *N. Biotechnol.* **2015**, *32* (6), 651–657.

(121) Shin, K. M.; Kim, S. I.; So, I.; Kim, S. J. *Electrochim. Acta* **2009**, *54* (16), 3979–3983.

(122) Comellas-Aragonès, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N. A. J. M.; Rowan, A. E.; Christianen, P. C. M.; Maan, J. C.; Verduin, B. J. M.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *Nat. Nanotechnol.* **2007**, *2* (10), 635–639.

(123) Inoue, T.; Kawano, M. A.; Takahashi, R. U.; Tsukamoto, H.; Enomoto, T.; Imai, T.; Kataoka, K.; Handa, H. *J. Biotechnol.* **2008**, *134* (1-2), 181–192.

(124) Glasgow, J. E.; Asensio, M. a; Jakobson, C. M.; Francis, M. B.; Tullman-Ercek, D. *ACS Synth. Biol.* **2015**, 150420173609002.

(125) Patterson, D.; Edwards, E.; Douglas, T. *Isr. J. Chem.* **2015**, *55* (1), 96–101.

(126) Patterson, D. P.; Schwarz, B.; Waters, R. S.; Gedeon, T.; Douglas, T. *ACS Chem. Biol.* **2014**, *9* (2), 359–365.

(127) Patterson, D. P.; Schwarz, B.; El-Boubbou, K.; van der Oost, J.; Prevelige, P. E.; Douglas, T. *Soft Matter* **2012**, *8* (39), 10158.

(128) Patterson, D. P.; LaFrance, B.; Douglas, T. *Chem. Commun. (Camb).* **2013**, *49* (88), 10412–10414.

(129) Patterson, D. P.; McCoy, K.; Fijen, C.; Douglas, T. *J. Mater. Chem. B* **2014**, *2* (36), 5948.

(130) Patterson, D. P.; Prevelige, P. E.; Douglas, T. *ACS Nano* **2012**, *6* (6), 5000–5009.

(131) Zhang, X.; Meining, W.; Fischer, M.; Bacher, a; Ladenstein, R. *J. Mol. Biol.* **2001**, *306* (5), 1099–1114.

(132) Woycechowsky, K. J.; Seebeck, F. P.; Hilvert, D. *Protein Sci.* **2006**, *15* (5), 1106–1114.

(133) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D. *Science* **2011**, *331* (6017), 589–592.

(134) Lilavivat, S.; Sardar, D.; Jana, S.; Thomas, G. C.; Woycechowsky, K. J. *J. Am. Chem. Soc.* **2012**, *134* (32), 13152–13155.

(135) Chen, H. N.; Woycechowsky, K. J. *Biochemistry* **2012**, *51* (23), 4704–4712.

(136) Yeates, T. O.; Crowley, C. S.; Tanaka, S. *Annu. Rev. Biophys.* **2010**, *39*, 185–205.

(137) Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolásek, B.; Ban, N. *Science* **2007**, *316* (5822), 254–261.

(138) Price, G. D.; Badger, M. R. *Plant Physiol.* **1989**, *91* (2), 505–513.

(139) Bacher, A.; Baur, R.; Eggers, U.; Harders, H. D.; Otto, M. K.; Schnepple, H. *J. Biol. Chem.* **1980**, *255* (2), 632–637.

(140) Schott, K. *J. Biol. Chem.* **1990**, *265* (21), 12686–12689.

(141) Deo, V. K.; Kato, T.; Park, E. Y. *Mol. Pharm.* **2015**, *12* (3), 839–845.

(142) Molino, N. M.; Wang, S. W. *Curr. Opin. Biotechnol.* **2014**, *28* (Figure 1), 75–82.

(143) Anand, P.; Neil, A. O.; Lin, E.; Douglas, T.; Holford, M. *Nat. Publ. Gr.* **2015**, *10021* (April), 1–10.

(144) Lipinski, D. M.; Reid, C. A.; Boye, S. L.; Peterson, J. J.; Qi, X.; Boye, S. E.; Boulton, M. E.; Hauswirth, W. W. *Hum. Gene Ther.* **2015**.

(145) ElSohly, A. M.; Netirojjanakul, C.; Aanei, I. L.; Jager, A.; Bendall, S. C.; Farkas, M. E.; Nolan, G. P.; Francis, M. B. *Bioconjug. Chem.* **2015**, *26* (8), 1590–1596.

(146) Li, K.; Nguyen, H. G.; Lu, X.; Wang, Q. *Analyst* **2010**, *135* (1), 21–27.

(147) Liepold, L.; Anderson, S.; Willits, D.; Oltrogge, L.; Frank, J. A.; Douglas, T.; Young, M. *Magn. Reson. Med.* **2007**, *58* (5), 871–879.

(148) Farkas, M. E.; Aanei, I. L.; Behrens, C. R.; Tong, G. J.; Murphy, S. T.; O'Neil, J. P.; Francis, M. B. *Mol. Pharm.* **2013**, *10* (1), 69–76.

(149) Datta, A.; Hooker, J. M.; Botta, M.; Francis, M. B.; Aime, S.; Raymond, K. N. *J. Am. Chem. Soc.* **2008**, *130* (8), 2546–2552.

(150) Zhou, Z.; Bedwell, G. J.; Li, R.; Prevelige, P. E.; Gupta, A. *Sci. Rep.* **2014**, *4*, 3832.

(151) Maity, B.; Fujita, K.; Ueno, T. *Curr. Opin. Chem. Biol.* **2015**, *25*, 88–97.

(152) Bedwell, G. J.; Zhou, Z.; Uchida, M.; Douglas, T.; Gupta, A.; Prevelige, P. E. *Biomacromolecules* **2015**, *16* (1), 214–218.

(153) Uchida, M.; Kang, S.; Reichhardt, C.; Harlen, K.; Douglas, T. *Biochim. Biophys. Acta - Gen. Subj.* **2010**, *1800* (8), 834–845.

(154) Shenton, W.; Mann, S.; Cölfen, H.; Bacher, A.; Fischer, M. *Angew. Chem. Int. Ed. Engl.* **2001**, *40* (2), 442–445.

(155) Minten, I. J.; Claessen, V. I.; Blank, K.; Rowan, A. E.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Chem. Sci.* **2011**, *2* (2), 358.

(156) Goldberg, M.; Langer, R.; Jia, X. *J. Biomater. Sci. Polym. Ed.* **2007**, *18* (3), 241–268.

(157) Deckert, G.; Warren, P. V; Gaasterland, T.; Young, W. G.; Lenox, A. L.; Graham, D. E.; Overbeek, R.; Snead, M. A.; Keller, M.; Aujay, M.; Huber, R.; Feldman, R. A.; Short, J. M.; Olsen, G. J.; Swanson, R. V. *Nature* **1998**, *392* (6674), 353–358.

(158) Liu, P.; Ewis, H. E.; Tai, P. C.; Lu, C. D.; Weber, I. T. *J. Mol. Biol.* **2007**, *367* (1), 212–223.

(159) Ewis, H. E.; Abdelal, A. T.; Lu, C. D. *Gene* **2004**, *329* (1-2), 187–195.

(160) Held, M.; Quin, M. B.; Schmidt-Dannert, C. *J. Mol. Microbiol. Biotechnol.* **2013**, *23* (4-5), 308–320.

(161) Yeates, T. O.; Thompson, M. C.; Bobik, T. a. *Curr. Opin. Struct. Biol.* **2011**, *21* (2), 223–231.

(162) Rae, B. D.; Long, B. M.; Badger, M. R.; Price, G. D. *Microbiol. Mol. Biol. Rev.* **2013**, *77* (3), 357–379.

(163) Heinhorst, S.; Williams, E. B.; Cai, F.; Murin, C. D.; Shively, J. M.; Cannon, G. C. *J. Bacteriol.* **2006**, *188* (23), 8087–8094.

(164) Norris, M. G. S.; Malys, N. *Biochem. Biophys. Res. Commun.* **2011**, *405* (3), 388–392.

(165) Meltzer, R. H.; Thompson, E.; Soman, K. V; Song, X.-Z.; Ebalunode, J. O.; Wensel, T. G.; Briggs, J. M.; Pedersen, S. E. *Biophys. J.* **2006**, *91* (4), 1302–1314.

(166) Folcarelli, S.; Battistoni, A.; Falconi, M.; O'Neill, P.; Rotilio, G.; Desideri, A. *Biochem. Biophys. Res. Commun.* **1998**, *244* (3), 908–911.

(167) Elcock, A. H.; Huber, G. A.; McCammon, J. A. *Biochemistry* **1997**, *36* (51), 16049–16058.

(168) Tao, Z.; Zhang, Z.; Grewer, C. *J. Biol. Chem.* **2006**, *281* (15), 10263–10272.

(169) Merickel, A.; Kaback, H. R.; Edwards, R. H. *J. Biol. Chem.* **1997**, *272* (9), 5403–5408.

(170) Koh, J. T. *Chem. Biol.* **2002**, *9* (1), 17–23.

(171) Ting, C. L.; Wu, J.; Wang, Z.-G. *Proc. Natl. Acad. Sci.* **2011**, *108* (41), 16986–16991.

(172) Schnell, S. *FEBS J.* **2014**, *281* (2), 464–472.

(173) Jiang, M.; Guo, Z. *J. Am. Chem. Soc.* **2007**, *129* (4), 730–731.

(174) Iglic, A.; Gongadze, E.; Bohinc, K. *Bioelectrochemistry* **2010**, *79* (2), 223–227.

(175) Watashi, K.; Wakita, T. *Cold Spring Harb. Perspect. Med.* **2015**, *5* (8).

(176) Banerjee, D.; Liu, A. P.; Voss, N. R.; Schmid, S. L.; Finn, M. G. *Chembiochem* **2010**, *11* (9), 1273–1279.

(177) Wiethoff, C. M.; Nemerow, G. R. *Virology* **2015**, *479-480*, 591–599.

(178) Fontes, C. M. G. A.; Gilbert, H. J. *Annu. Rev. Biochem.* **2010**, *79*, 655–681.

(179) Cameron, K.; Najmudin, S.; Alves, V. D.; Bayer, E. A.; Smith, S. P.; Bule, P.; Waller, H.; Ferreira, L. M. A.; Gilbert, H. J.; Fontes, C. M. G. A. *J. Biol. Chem.* **2015**, *290* (21), 13578–13590.

(180) Resch, M. G.; Donohoe, B. S.; Baker, J. O.; Decker, S. R.; Bayer, E. A.; Beckham, G. T.; Himmel, M. E. *Energy Environ. Sci.* **2013**, *6* (6), 1858.

(181) Hyeon, J. E.; Jeon, S. D.; Han, S. O. *Biotechnol. Adv.* **2013**, *31* (6), 936–944.

(182) Procko, E.; Hedman, R.; Hamilton, K.; Seetharaman, J.; Fleishman, S. J.; Su, M.; Aramini, J.; Kornhaber, G.; Hunt, J. F.; Tong, L.; Montelione, G. T.; Baker, D. *J. Mol. Biol.* **2013**, *425* (18), 3563–3575.

(183) Hoersch, D.; Roh, S.-H.; Chiu, W.; Kortemme, T. *Nat. Nanotechnol.* **2013**, *8* (12), 928–932.

(184) Song, W. J.; Tezcan, F. A. *Science (80-.).* **2014**, *346* (6216), 1525–1528.

(185) Simpson, R. J. Proteins and Proteomics: A Laboratory Manual http://www.proteinsandproteomics.org/home_1.html (accessed Oct 11, 2015).

(186) Vaughan, T. J.; Williams, A. J.; Pritchard, K.; Osbourn, J. K.; Pope, A. R.; Earnshaw, J. C.; McCafferty, J.; Hodits, R. A.; Wilton, J.; Johnson, K. S. *Nat. Biotechnol.* **1996**, *14* (3), 309–314.

(187) Cauerhff, A.; Goldbaum, F. A.; Braden, B. C. *Proc. Natl. Acad. Sci.* **2004**, *101* (10), 3539–3544.

(188) Wong, E. T. C.; Na, D.; Gsponer, J. *PLoS Comput. Biol.* **2013**, *9* (8), e1003192.

(189) Bosshard, H. R.; Marti, D. N.; Jelesarov, I. *J. Mol. Recognit. 17* (1), 1–16.

(190) Wang, L.; Lane, L. C.; Smith, D. L. *Protein Sci.* **2001**, *10* (6), 1234–1243.

(191) Wilts, B. D.; Schaap, I. A. T.; Schmidt, C. F. *Biophys. J.* **2015**, *108* (10), 2541–2549.

(192) Katen, S.; Zlotnick, A. *Methods Enzymol.* **2009**, *455*, 395–417.