# ENGINEERING BACTERIOPHAGE P22 AS A NANOMATERIAL

by

Alison Linsley O'Neil

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

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2013

©COPYRIGHT

by

Alison Linsley O'Neil

2013

All Rights Reserved

APPROVAL

of a dissertation submitted by

Alison Linsley O'Neil

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Dr. Trevor Douglas

Approved for the Department of Chemistry and Biochemistry

Dr. Mary Cloninger

Approved for The Graduate School

Dr. Ronald W. Larsen

# STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Alison Linsley O'Neil

April 2013

# ACKNOWLEDGEMENTS

I would like to acknowledge all of my fantastic mentors who have been key components to my success. To my first mentor at MSU, Dr. Sebyung Kang, I thank you for teaching me how to thoughtfully plan out an experiment. Dr. Jamie Snyder, without your guidance through the molecular biology black box, this entire dissertation could not have been possible. It is to you that I owe my expertise in this field. My committee members, Professors Brian Bothner and Valerie Copie, both of you have never let me get away with anything and have always pushed me to my best. Committee member Professor Peter Prevelige, it has been such an honor to have a P22 expert on my committee. Your sarcasm and humor always kept me on my toes. Lastly, with absolute gratitude, I thank my advisor, Professor Trevor Douglas. I have never worked for someone as open and receptive to scientific wonder. Whenever I was discouraged, you had a wonderful way of getting me excited to get back into the lab. Our fruitful brainstorms have taken the lab in a whole new direction- I cannot wait to see where you take it next.

Thank you to all of the Douglas Lab members, present and past, for your friendship, comradery, and support. We are such a great team! Thank you to all of my fantastic friends for being understanding when I was late because I got "stuck in the lab". Thank you to my family for your continued support through *all* of my wild adventures. Finally, many thanks to the beautiful state of Montana. Your mountains and rivers have truly expanded my mind and soul.

TABLE OF CONTENTS

1. INTRODUCTION .....1

Protein Containers and Encapsulation .....1
Encapsulation of Proteins and Enzymes .....5
Polymers Inside Protein Containers .....15
Inorganic Particles Inside Protein Containers .....24
Drug Delivery Using Protein Containers .....31
Building New Protein Cage Architectures .....34
Bacteriophage P22 .....38
Assembly Overview .....39
P22 Scaffold Protein .....43
P22 Portal Protein .....48
Overall Conclusions .....52
Research Directions .....53

2. GENETICALLY PROGRAMMED IN VIVO PACKAGING OF PROTEIN CARGO AND ITS CONTROLLED RELEASE FROM BACTERIOPHAGE P22 .....55

Contribution of Authors and Co-Authors .....55
Manuscript Information Page .....56
Introduction .....57
Results and Discussion .....58
Conclusion .....65
Experimental .....66

3. CO-CONFINEMENT OF FLUORESCENT PROTEINS: SPATIALLY ENFORCED COMMUNICATION OF GFP AND MCHERRY ENCAPSULATED WITHIN THE P22 CAPSID .....68

Contribution of Authors and Co-Authors .....68
Manuscript Information Page .....69
Abstract .....70
Introduction .....71
Results and Discussion .....73
Conclusion .....80
Experimental .....81

TABLE OF CONTENTS – CONTINUED

4. ENCAPSULATION WITHIN THE P22 CAPSID GREATLY ENHANCES THE STABILITY OF A PHOSPHOTRIESTERASE....87

Contribution of Authors and Co-Authors....87
Manuscript Information Page....88
Abstract....89
Introduction....90
Results and Discussion....92
Conclusion....99
Experimental....99

5. TEMPLATED ASSEMBLY OF ORGANIC-INORGANIC MATERIALS USING THE CORE SHELL STRUCTURE OF THE P22 BACTERIOPHAGE....104

Contribution of Authors and Co-Authors....104
Manuscript Information Page....105
Abstract....106
Results and Discussion....106

6. ENCAPSULATION OF ACTIVE PEPTIDES WITHIN THE P22 CAPSID....113

Introduction....113
MVIIA Cono-toxin Encapsulation....114
Introduction....114
Results and Discussion....115
Conclusions and Future Directions....119
Experimental....120
Poly-Tyrosine Encapsulation....123
Rationale....123
Preliminary Results....123
Conclusions and Future Directions....125
Experimental....125
Poly-Lysine Encapsulation....126
Rationale....126
Preliminary Results....127
Conclusions and Future Directions....128
Experimental....128
Overall Conclusions and Future Directions....128

TABLE OF CONTENTS – CONTINUED

7. EXPLOITING THE NATURAL ASYMMETRY OF BACTERIOPHAGE P22 FOR NANOMATERIALS DESIGN ..... 130

Background ..... 130
Introduction ..... 134
Engineering New Cysteine Residues in to the Portal Protein ..... 134
Results and Discussion ..... 136
Conclusion ..... 141
Experimental ..... 142
Two Vector Approach for Portal Incorporation ..... 146
Results and Discussion ..... 147
Conclusion ..... 150
Experimental ..... 150
Non-Natural Amino Acid Incorporation in to the Portal Protein ..... 153
Results and Discussion ..... 155
Conclusion ..... 159
Experimental ..... 159
Overall Conclusions and Future Directions ..... 165

8. CONCLUDING REMARKS ..... 167

REFERENCES CITED ..... 173

APPENDICES ..... 186

APPENDIX A: Supplemental Information for Chapter 2 ..... 186
APPENDIX B: Supplemental Information for Chapter 3 ..... 201
APPENDIX C: Alternative Version of Manuscript in Chapter 3 ..... 208
APPENDIX D: Supplemental Information for Chapter 4 ..... 229
APPENDIX E: Supplemental Information for Chapter 5 ..... 234

# LIST OF TABLES

Table Page

3.1. Summary of the Physical Characteristics of the P22 Encapsulation Systems.....76

# LIST OF FIGURES

Figure Page

1.1 The Euthanolamine Utilization Microcompartment ..... 2

1.2 Protein Containers Have Three Distinct Modifiable Surfaces ..... 4

1.3 Charge Complementation as an Encapsulation Strategy for Loading Lumazine Synthase Protein Containers ..... 6

1.4 Using a Coil-Coil Linking Motif to Load Proteins Inside the CCMV Capsid ..... 7

1.5 A Genetic Fusion Strategy to the Minor SV40 Coat Protein as an Encapsulation Signal for Protein Cargo ..... 8

1.6 Exploiting the Natural Affinity for the Inside of the Qβ Virus Capsid to an RNA Hairpin as an Aptamer for Encapsulating Proteins ..... 10

1.7 The pH Dependent Dissociation and Re-assembly of CCMV Allows for the Trapping of Single Enzymes Inside ..... 11

1.8 Genetic Fusion of the P22 Scaffold Protein to Alcohol Dehydrogenase Acts as an Encapsulation Signal and Templates Assembly of P22 Capsids ..... 14

1.9 A Thermo-Responsive Polymer can Drive the Assembly of CCMV ..... 16

1.10 Using a ssDNA Molecule and ssDNA-Fluorescent Molecule Hybrid to Drive CCMV Assemblages ..... 18

1.11 The P22 Capsid as Restrictive Vessel for ATRP ..... 20

1.12 Bi-Functional CCMV: PEG Labeling on the Exterior with Polystyrene Sulfonate Captured Inside ..... 22

1.13 Using a Photolabile Dendron, Magneto-Ferritin was Assembled into Larger Assemblies that Could be Dismantled with UV Exposure ..... 23

1.14 BMV Coat Monomers Collapse onto TEG-Coated Gold Nanoparticles Assembling into Capsids with a Gold Core ..... 26

LIST OF FIGURES-CONTINUED

Figure Page

1.15 The Natural Affinity of Ferritin for Iron Allows for in vitro Mineralization of Iron Oxide on the Interior of the Protein Cage .....30

1.16 Using an Engineered Interior Cysteine, High Numbers of Modified Taxol Molecules to be Encapsulated .....33

1.17 Scheme for Designing New Protein Architectures Through the Fusion of Distinct Geometry Forming Monomers.....35

1.18 Overview of Bacteriophage P22 Assembly .....40

1.19 The in vitro Accessible Morphologies of the P22 Capsid .....42

1.20 The Solution NMR Structure of the C-terminus of the P22 Scaffold Protein and the Full-Length Predicted Structure Model .....45

1.21 The Proposed Functional Domains of the P22 Scaffold Protein .....48

1.22 Asymmetric Cryo-EM Reconstruction of the P22 Phage .....49

1.23 Crystal Structure of the P22 Portal Complex with DNA Modeled in to the Portal Inner Chamber.....53

2.1 Plasmid Map, Color Photograph, and TEM Images of P22 GFP and P22 mCherry .....59

2.2 Illustration of the Translation Products of the P22 Encapsulation Assembler Plasmid and the Capsid Morphological Forms.....61

3.1 Plasmid Map and Cartoon of P22-Cat Encapsulation System.....72

3.2 Characterization of P22-Cat(0gly).....73

3.3 Fluorescence Spectra and Relative FRET Versus Spacer Length .....78

4.1 Characterization of P22-PTE .....94

4.2 Michaelis-Menten Plot of P22-PTE Kinetics .....95

LIST OF FIGURES-CONTINUED

Figure Page

4.3 P22-PTE Thermal Stability. ....97

4.4 P22-PTE Proteolytic and Lyophilization Stability ....98

5.1 Schematic Illustration of the P22 Truncated Scaffold and EA Constructs ....108

5.2 SEC and Prussian Blue Stained Gel of P22 SP-EA ....110

5.3 TEM Images of Iron Loaded P22 cages ....111

6.1 S500 Size Exclusion Chromatography of the P22-MVIIA ....116

6.2 Conjugation of TAT-FAM to M338C P22-MVIIA, SDS-PAGE and MALS ....117

6.3 Fluorescent Cell Imaging Results for the TAT-FAM P22-MVIIA-mCherry ....118

6.4 Scheme for Recursive PCR Strategy for Appending Peptide Sequences to the N-terminus of the Scaffold Protein ....121

6.5 S500 Size Exclusion Chromatography of the P22-$YAG_9$ and Iodinated Ultra-Centrifugation Virus Pellets ....124

6.6 S500 Chromatogran of P22 KA SP238 ....127

7.1 Crystal Structure of the Portal Complex Showing Monomer Orientation and the Location of the Natural Cysteines ....135

7.2 Crystal Structure of the Portal Complex Illustrating the Location of the Newly Engineered Cysteines ....135

7.3 Characterization of the N375C and T377C Portal P22 Capsids ....137

7.4 SDS-PAGE of the Fluorescein Labeled T377C Portal P22 Capsids ....137

7.5 Schematic of the Portal-Catalase Linking Strategy ....139

LIST OF FIGURES-CONTINUED

Figure Page

7.6 Characterization of the Portal-Catalase Linked Sample .....140

7.7 SDS-PAGE and Western Blots of the Catalase Linked Samples .....141

7.8 SEC of the Co-Transformation P22-Portal Constructs .....149

7.9 SDS-PAGE and Western Blots of the Co-Transformation P22-Portal Constructs .....150

7.10 Crystal Structure of the Portal Illustrating the Natural and New Methionines .....155

7.11 Characterization of P22-AHA .....158

# ABSTRACT

The precise architectures of viruses and virus-like particles are highly advantageous in synthetic materials applications. These nano-size compartments are perfectly suited to act as containers of designed cargo. Not only can these nano-containers be harnessed as active materials, but they can be exploited for examining the effects of *in vivo* "cell-like" crowding and confinement on the properties of the encapsulated cargo. The high concentration of many different types of mutually volume excluding macromolecules in the cell causes it to be a crowded and confining environment in which to carry out reactions. Herein, the molecular design of the bacteriophage P22 encapsulation system is described and utilized for the synthesis of active nanomaterials and to explore the effect of encapsulation on the entrapped proteins' properties. In the designed system, any gene can be inserted and results in the fusion of the insert to a truncated form of the P22 scaffold protein. This scaffold protein fusion templates the spontaneous *in vivo* assembly of P22 capsids and also acts as an encapsulation signal. Once encapsulated, we can examine how crowding and confinement affect inter-molecular communication and activity of the cargo molecules. The P22 system is unique in that the capsid morphology can be altered, without losing the encapsulated cargo, resulting in a doubling of the capsid volume. Thus, the encapsulated fusions can be examined at two different internal concentrations. The packaged capsids contain up to 300 copies of fusion and fill more than 24% of the internal volume with the internal concentration of the fusions in the millimolar range. Not only are these fusions densely and efficiently packaged, but they retain their activity. Described herein is the packaging of fluorescent proteins, enzymes, and active peptides. In all cases, it is shown that the enforced proximity via encapsulation greatly affects the fusions activity compared to the fusion free in solution. To expand the utility of the P22 capsid as a nanomaterial, the inherent asymmetry implored by the portal complex has also been exploited. The P22 encapsulation system has proved to be an effective and versatile vehicle for nanomaterials design.

# CHAPTER ONE

# INTRODUCTION

## Protein Containers and Encapsulation

Compartmentalization is a defining characteristic of life. In all domains of life compartments delineate self from the rest of the environment, individual cells from each other, and the localization of sub-cellular organelles. Through the use of lipid membranes, compartments have evolved to create local environments based on activity. The nucleus, an excellent example of this, harbors all the components for DNA maintenance, replication, and transcription inside a lipid membrane. The specific sequestration of these components inside a lipid-delineated compartment, creates a specialized, highly efficient, functional epicenter. More recently, it has been discovered that all domains of life also take advantage of protein assemblies as compartments for specialized activities. Protein-based containers are three-dimensional shells built out of many copies of one or more structural proteins. In nature they are used as houses for metabolic pathways, storage vessels, and places of protein folding and protection.

An example of these protein compartments is found in the bacteria *Escherichia coli* and *Salmonella enterica*, where the complex ethanolamine utilization microcompartment is a hollow shell built of four primary protein components (Figure 1.1) (*20*). This microcompartment sequesters the components necessary to metabolize ethanolamine (*20, 23*) while sequestering harmful side products of the reaction (*23*). Another bacterial example of a protein container is the carboxysome. This ~80

nanometer structure encapsulates the key enzymes in $CO_2$ fixation and, in turn, enhances the process. In mammals and plants, ferritins are made of 24 protein subunits creating a protein cage (*24*). Ferritins sequester, store, and deliver iron to maintain iron homeostasis (*24*). The pentagonal dodecahedral bovine pyruvate dehydrogenase complex is built out of 3 enzymes and compromises the largest enzyme complex known (*25*). This important complex is the link between glycolysis and the tricarboxcylic acid cycle in mammals (*25*).

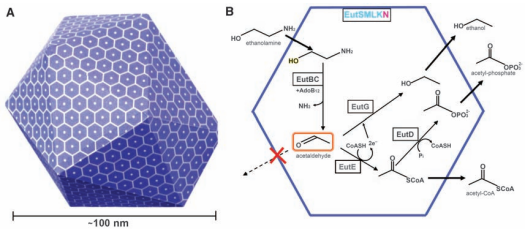



Figure 1.1: (A) A structural model of the bacterial ethanolamine utilization compartment. (B) A schematic of the complex chemistry that proceeds inside the compartment through the coupling of several enzymatic reactions. Boxed in orange, the harmful side product, formaldehyde, is sequestered inside the compartment to protect the cell. (*20*)

Another class of protein containers can be derived from viruses. Viruses, like protein containers, assemble from many copies of their coat proteins into precise hollow architectures. In their native state, viruses contain their nucleic acid cargo, shielding it from the exterior environment to safely deliver it to a host cell. However, many virus particles can be made devoid of their nucleic acid cargo while retaining their exact geometries. These hollow, non-infectious, protein containers are called virus like particles (VLPs).

Virus like particles, and other naturally occurring protein containers, confer excellent advantages for nanomaterials design and synthesis. As the assemblages are hollow containers, non-native cargo can be encapsulated through either a process of directed assembly around the cargo or through modification after the cage-like architecture has been formed. Through encapsulation within a container, the entrapped cargo can be protected from the exterior environment. Although permeable to most small molecules, naturally occurring protein containers are not penetrable to larger proteolytic enzymes, which could degrade the encapsulated cargo. Encapsulation of proteins has also been shown to confer protection from thermal perturbation, the lyophilization process, increase encapsulated cargos expression level and solubility, and in some cases enhance catalytic activity.

Sequestration within the nano-container can also lead to higher internal concentrations of the cargo than can be reached for the cargo free in solution. These high concentrations and in turn, the crowding of the cargo, can have significant influence on the activity and other properties of the cargo. When conducting multi-step reaction pathways within a nano-container, intermediates can be sequestered in high concentration such that the consecutive step can be enhanced due to high available substrate concentration. Through sequestering and crowding enzyme partners together in a protein container, the distance an intermediate substrate has to travel to reach the next step is shortened, leading to enhanced catalytic rate. In this way, pathways can be optimized as all components are spatially controlled and brought within the same vicinity.

By conducting reactions inside a protein container, it is also possible to sequester harmful reaction by-products and intermediates. When encapsulated, the harmful product has a shortened diffusion distance to find its consecutive processing step. Therefore it is more likely that the harmful product will get processed rather than diffuse out of the container to cause deleterious effects.

An additional advantage of protein nano-containers is their precise architectures and size. The precise assembly of protein containers, and in some cases the availability of the calculated structure, allows for the precise genetic engineering of the container for optimized use. As an example, cysteine molecular handles can be engineered to be on the interior, exterior, or at protein interfaces of the container for downstream use. The three distinct surfaces available for engineering in protein containers allows for diverse modifications and applications (Figure 1.2). The defined volume of the container can also allow for restrictive templating of inorganic materials and nano-particle growth.

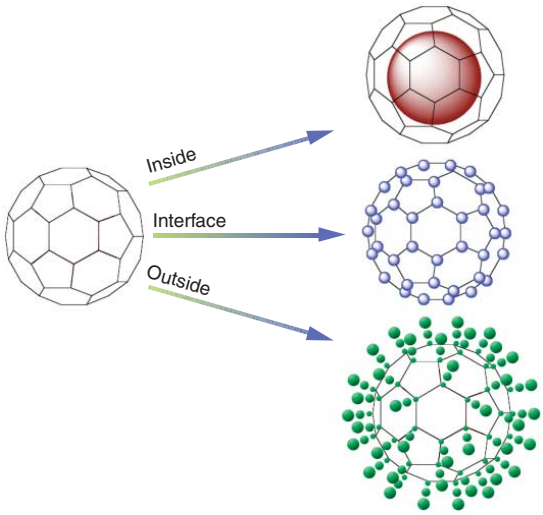



Figure 1.2: Protein containers are precise architectures built of repeating monomers. The use of protein containers allows for distinct engineering of the interior, the monomer interfaces, or the exterior surface for diverse functionalities. (*5*)

The nano-size and repetitive nature of the assembly also confers advantages for *in vivo* use of nanocontainers. The size of protein containers, ~10-100 nanometers versus ~10 microns for a mammalian cell, causes them to be processed differently than larger complexes. The small size allows for the containers to be engulfed by the cell via several different mechanisms. It has also been hypothesized that the repetitive nature of the assembly (the container is usually made of many copies of the same protein) causes a special response by the immune system that may be advantageous in nanomaterial design.

Encapsulation of Proteins and Enzymes

As discussed above, encapsulation within a protein container confers many desirable properties on the cargo. These traits seem most appealing for the designed encapsulation of protein cargos because the encapsulated proteins can be shielded from the exterior environment while maintaining their native function. It has also been shown that many enzymes are stabilized by crowding agents and hence, densely packaging enzymes in a protein container, may also confer these stabilizing effects. When designing new encapsulation strategies, many research groups have turned to fluorescent proteins as a cargo for demonstrating the 'proof of concept'. The expression and location of fluorescent proteins are easily tracked by their distinctive absorbance and emission spectra.

Fluorescent Protein Cargo. A simple strategy for encapsulation is to use complementary charge between the cargo and the interior interface of the capsid. Taking this approach, the Hilvert group first engineered the interior cavity of lumazine synthase to contain several negatively charged amino acids. By expressing this negatively charged lumazine synthase monomer together with a either a green fluorescent protein (GFP) or HIV protease, to which an added positively charged amino acid tail has been added, the charge complementarity caused the GFP or HIV protease to be sequestered inside the lumazine synthase cage (Figure 1.3). (*6, 26, 27*) By using this simple tagging system for encapsulation, the Hilvert group devised a strategy for the directed evolution of their lumazine synthase cages to be better sequestration vehicles. First, a random library was made of the lumazine synthase. The library was then expressed with the positively charged HIV protease. If the HIV protease was not sequestered into a lumazine synthase cage, the protease had deleterious effects on the bacteria that were expressing the proteins. Therefore, bacterial cultures that grew at a steady rate, despite the level of HIV

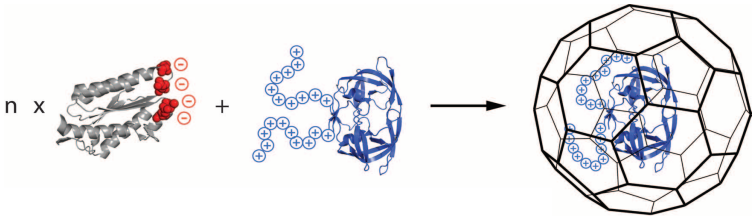



Figure 1.3: Charge complementarity can be used as a simple packaging scheme. By engineering the monomer unit of lumazine synthase with a negative interior surface, any positively charged protein (shown here HIV-protease) is effectively sequestered in the assembled lumazine synthase cage. (*6*)

protease that was expressed, were thus expressing an efficient sequestration vehicle. (*6*) Using a similar strategy, it could be possible to screen protein container libraries for other desirable traits.

A similar tagging strategy for encapsulation, used by the Cornelissen group, exploited coil-coil complementary tags on the interior surface of a plant virus coat protein and their cargo of interest (Figure 1.4) (*16*). Cowpea chlorotic mottle virus (CCMV) has

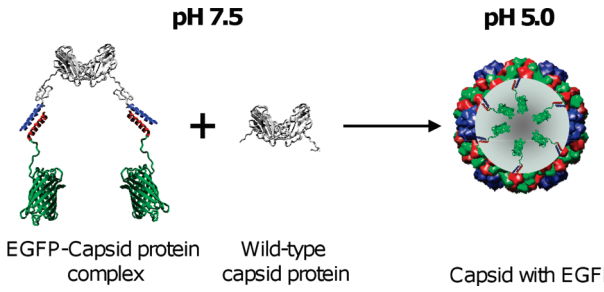



Figure 1.4: Through exploitation of a coil-coil interacting motif (red and blue helices in figure), cargo proteins (here EGFP) are anchored to the coat protein dimers of CCMV. By Adjusting the pH to 5.0, the dimer-cargo complexes will assemble with wild-type dimers to spontaneously assemble into CCMV VLPs containing the cargo. (*16*)

a pH dependent disassembly process. When purified from plants, CCMV contains its native RNA nucleic acid cargo. However, when the pH is raised to 7.5, the CCMV capsids disassemble into dimers releasing the RNA. (*16*) The solution of dimers can then be adjusted to low pH, and the CCMV capsids, devoid of nucleic acid, will reform into 28 nm, T=3 capsids. In the coil-coil scheme, an E-coil was genetically fused to the C-terminus of a 6xHis N-terminally labeled GFP cargo molecule. The GFP was then attached to a nickel resin column via its His tag and well washed. (*16*) Next, the CCMV

coat protein dimers (at pH 7.5) with an N-terminal genetically fused K-coil was flowed over the column allowing the E-coil tag and the K-coil tag to make a strong non-covalent linkage. The complex was then eluted from the nickel resin column. The CCMV dimer-GFP complex was then mixed with native dimers and the pH adjusted to 5, causing the spontaneous assembly of capsids containing GFP. By varying the ratio of dimer-GFP complex to native dimers, a controllable range of up to 15 GFPs per CCMV could be encapsulated. (*16*)

Perhaps a more elegant strategy for encapsulation is to harness naturally occurring signals that cause cargo proteins to be sequestered inside a protein container. Through exploitation of these naturally occurring systems, encapsulation is more directed as it does not rely on non-natural interactions. An additional advantage to natural strategies, is that encapsulation can occur *in vivo* as the container is being expressed and thus does away with time consuming purification of components before they are assembled.

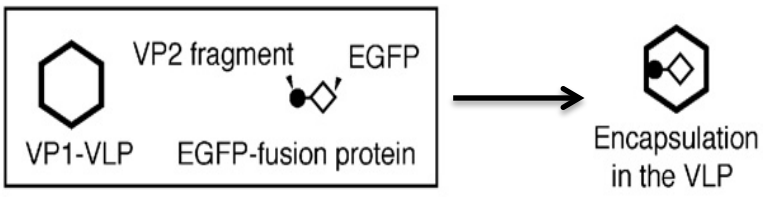



Figure 1.5: Through genetic fusion of the cargo protein (EGFP) to an interior component of the SV40 capsid (VP2), the cargo is encapsulated upon SV40 assembly. Adapted from Inoue, 2008. (*11*)

Using a genetic fusion strategy to an interior capsid protein, it was possible to encapsulate fluorescent proteins in the SV40 VLP (*11*). SV40 virus assembles from 72 pentamers of the major capsid protein VP1 and 72 copies of the minor coat proteins VP2.

When assembled, the minor coat proteins are arranged on the interior surface of the capsid through an interaction with VP1. Using a genetic fusion strategy, GFP was fused to the C-terminus of VP2 and expressed together with the major coat protein in insect cell culture. The capsids spontaneously assembled *in vivo* and the VP1-VP2 interaction was sufficient to encapsulate the VP2-GFP fusion (Figure 1.5). (*11*) An interesting outcome of this encapsulation method was that because the cargo was sequestered inside the VLP, the exterior surface was left unperturbed and the VLP was still able to bind and enter host cells. (*11*)

Using a different genetic fusion, Fan and colleagues were able to encapsulate GFP (and other proteins) into bacterial microcompartments (*28*). The authors reported that the N-terminal 18 amino acids from propionaldehyde dehydrogenase (a naturally occurring cargo enzyme) is the signaling peptide that causes this enzyme to be encapsulated in its associated microcompartment. This peptide was then fused to GFP, GST, and maltose-binding protein with the result being that when expressed together, these proteins were directed to the interior of the microcompartment. (*28*)

Viruses naturally package their nucleic acid in the interior of their capsid and this interaction can also be exploited for encapsulation. The native Qβ virus packages ssRNA through an interaction between its coat proteins and a specific ssRNA hairpin structure (*29*). The Finn group has exploited this naturally occurring interaction by creating a bifunctional ssRNA that at one end contains this specific hairpin and on the other end has an α-Rev aptamer, which has affinity to a Rev-tag (Figure 1.6). The Rev-tag was then genetically fused to the cargo protein of interest and thus, by expressing Qβ coat proteins,

the bi-functional ssRNA, and the Rev-tagged cargo proteins, these investigators were able to direct the encapsulation of several fluorescent proteins into the Qβ VLPs (Figure 1.6). (*29*) Additionally, a carbohydrate-based ligand to the CD22 cell receptor was attached to the exterior of the capsids. The fluorescent protein packaged and ligand labeled Qβ VLPs were shown to specifically bind cells through fluorescence activated

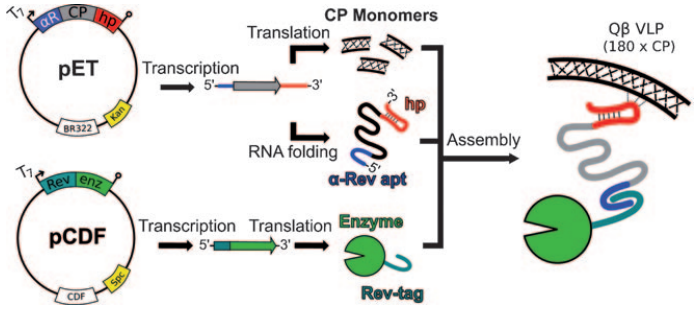



Figure 1.6: The VLP Qβ coat proteins have a natural affinity for a RNA-hairpin. A chimeric RNA molecule was engineered with this coat-binding hairpin and an α-Rev aptamer. Cargo proteins are expressed such that have a Rev-tag. All three components, Qβ coat proteins, chimeric RNA, and Rev-tagged cargo, are expressed together and result in an assembled Qβ VLP encapsulating the cargo. (*21*)

cell sorting (FACS) and confocal laser microscopy. (*29*)

Fluorescent proteins act as easily tracked 'proof of concept' cargos for protein containers. The strongly fluorescent engineered containers have applications in *in vivo* imaging and using a Forster Resonance Energy Transfer (FRET) pair, one can begin to probe protein communication within the protein container (*29, 30*). However, to create a catalytically active protein container, an active enzyme must be encapsulated.

Enzymes Encapsulation. In the last 20 years, the idea of enzyme kinetics in a crowded, more cell-like environment has spurred many studies into the effects and contributing factors of crowding. Through macromolecular crowding, it has been shown that the properties of some enzymes can be modulated. (*31, 32*) Interestingly, overall it appears that non-protein crowding agents (such as PEG) stabilize enzymes, while globular proteins (such as BSA) have a more destabilizing effect as crowding agents. (*33, 34*) By sequestering enzymes into the defined volume of a protein container, where internal concentrations can reach that of what is encountered in the cell (300 g/L), one can engineer a material that is well suited to study the effects of cell-like crowding on kinetics. Additionally, protein containers are applicable for this work because their porous nature allows small substrates to freely diffuse in and out of the container, engaging the encapsulated enzyme.

One of the first strategies for loading enzymes into VLPs was to exploit the pH dependent disassembly and re-assembly of CCMV discussed above (*8*). In this method, CCMV capsids were dissociated into dimers at pH 7.5. Next, the enzyme horseradish peroxidase (HRP) was added to the dimer solution. (*8*) The solution was then adjusted to pH 5 where the capsids spontaneously assembled, entrapping the HRP (Figure 1.7). This

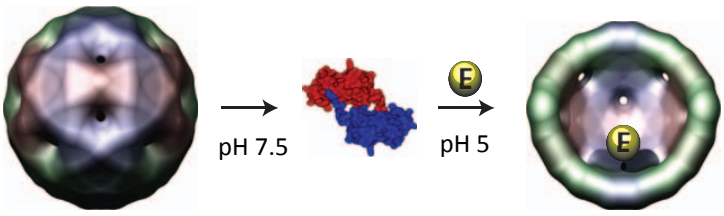



Figure 1.7: The CCMV capsid undergoes a reversible dissociation in to dimers at pH 7.5. By mixing the CCMV dimers with an enzyme cargo (E) and adjusting the pH to 5, cargo can be trapped in the assembling VLPs. Adapted from Comellas-Aragonés, 2007. (*8*)

method yielded capsids containing only a single HRP per capsid and allowed for single molecule kinetics to be studied. (*8*) Using a substrate that becomes fluorescent after encountering the HRP, the investigators showed particle specific activity and sequestration of the product within the capsid. (*8*)

Using the more directed coil-coil loading scheme for CCMV (described above, Figure 1.4), the enzyme lipase B from *Pseudozyma antartica* (PalB) was encapsulated resulting in an average of 1.3, 2, 3.5, or 4 enzymes per capsid (*35*). Standard Michaelis-Menten kinetics was carried out and it was observed that the encapsulated enzymes outperformed the non-encapsulated PalB. Additionally, it was reported that the reaction velocity of the encapsulated enzymes were inversely proportional to the number of encapsulated enzymes per capsid. (*35*) To investigate the role of crowding in this study, dimer-GFP complexes were added to the dimer-PalB reassembly solution resulting in capsids containing 1 PalB and 7 GFPs or 3 PalB and 7 GFPs, on average. The addition of more proteins that do not contribute to the substrate degradation (i.e. GFP) to the interior of the CCMV capsid was predicted to decrease the reaction velocity even more than seen in the initial study. (*35*) However, increasing the crowding in the CCMV only slightly decreased the rate. Therefore, crowding is not the only factor that limited the catalysis in this system. The investigators suggest that limits on diffusion may also play a role. (*35*)

Through exploitation of the Qβ RNA aptamer strategy described above (*29*), the encapsulation of peptidase E (PepE) and firefly luciferase was accomplished and a detailed kinetic analysis was described (*21*). Through this strategy 2-18 PepE and 4-8 luciferease enzymes could be loaded per particle. In the case of the PepE, the free

enzyme was only 3 times more efficient ($k_{cat}/K_M$) than the encapsulated PepE and the $K_M$ for both were very similar (210±20 μM for free versus 270±20 μM for the encapsulated PepE). (*21*) However, encapsulation did greatly increase the enzymes thermal stability and protected the enzyme from degradation from proteinase K. When the kinetics were examined for luciferase, the overall activity for free and encapsulated enzymes were very similar however the $K_M$ was significantly higher for the encapsulated enzyme. In this case, the capsid did not provide any thermal protection but did rescue the enzyme from hydrophobic adsorption to unblocked polystyrene plates. (*21*) Similar to what was seen in the case of PalB above (*35*), it was shown that capsids with 2 copies of luciferase outperformed capsids with 9 copies of luciferase. These investigators suggest that the difference is a result of increased crowding in the capsid, which causes the enzyme to have limited flexibility; ultimately affecting the rate. (*21*)

The Michaelis-Menten kinetics have been described for alcohol dehydrogenase D (AdhD) and CelB glucosidase (both from *Pyrococcus furiosus*) when encapsulated in the bacteriophage P22 VLP (*7, 36*). Through fusion of the enzymes with the P22 scaffold protein, the fusion was able to template assembly of monodisperse capsids and sequester the enzymes inside (Figure 1.8). Through this method, ~250 copies of the monomeric AdhD and 80 copies of the CelB monomer were encapsulated. In the case of AdhD, an internal concentration of 7 mM, the highest reported, was achieved. The high internal concentration, and in turn the high level of crowding, was shown to effect the kinetics of AdhD as compared to the enzyme free in solution. When encapsulated, the $k_{cat}$ was 7-fold lower than the free AdhD and the $K_M$ was also lowered upon encapsulation. A

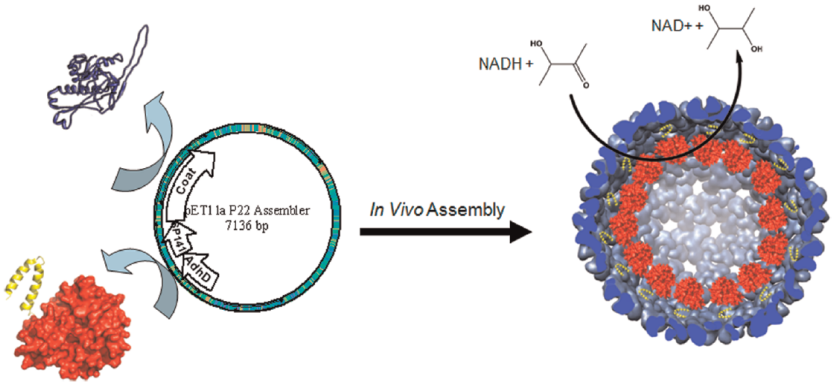



Figure 1.8: Schematic illustrating the genetic fusion of alcohol dehydrogenase (AdhD) with the P22 scaffold protein (SP141) and its concomitant expression with the P22 coat protein. The expression of the two proteins yields the spontaneous *in vivo* assembly of P22 VLPs that retain the AdhD catalytic activity. (*7*)

useful property of P22 is that the capsid can be controllably expanded to yield a capsid with twice the interior volume without losing any of the encapsulated enzyme. The investigators exploited this expanded P22 morphology to examine the kinetics of AdhD at two distinct concentrations but still entrapped within the P22 capsid. Overall, it was seen that as the concentration of AdhD increases in the capsid, the $K_M$ decreases. (*7*) In the case of CelB encapsulation, the P22 system was shown to be tolerant of the CelB tetramerization (which is necessary for activity) and the kinetics of the free enzyme were very similar to the encapsulated enzyme. An additional application was shown in that the packaged P22-CelB could be embedded in a polymeric gel and retain CelB activity, even after dehydration and subsequent rehydration. (*7, 36*)

Conclusions. Various systems have been demonstrated for the encapsulation of proteins inside protein containers. While some use simple tagging strategies, others harness naturally occurring affinities for the interior of the container. Upon

encapsulation of enzymes in protein containers, we and others, have observed varied effects to the enzymes kinetic profile. In some cases, there is little to no effect on the enzymes kinetics (*36, 37*), in another the overall rate seemed to be enhanced upon encapsulation (*38*), and in a third case the rate ($k_{cat}$) and the $K_M$ were decreased (*7*). The kinetic behavior of this promising new class of biomaterials is not yet predictable, as no trends are evident and the kinetic outcomes seem to be enzyme specific. However, encapsulation does bestow protection to the cargo enzyme.

Polymers Inside Protein Containers

Polymers are repetitive organic molecules that can occur naturally as in poly-sugars (starch), DNA, and polypeptides or synthetically as in plastics and acrylamide gels. Polymers come in all shapes (linear or branched), sizes (several monomers to hundreds), and applications (hydrophobic, hydrophilic, thermally responsive, etc.). The wide variety of polymers can be utilized to increase the functionality of protein containers. Polymers can be incorporated on the inside and outside of the containers conferring new properties and providing many new chemically addressable groups.

Polymers Used to Drive Assembly. An interesting application is to use polymers to drive the assembly of protein containers. Using a thermal-responsive elastin-like polypeptide (ELP) fused to the CCMV coat protein, the authors report being able to drive and modulate CCMV assembly (Figure 1.9) (*13*). The CCMV-ELP fusion could still undergo pH responsive assembly creating the native-like 28 nm icosahedra. However, an additional assemblage could be accessed by dialyzing the CCMV-ELP dimers against pH

7.5, 2.5 M NaCl buffer at room temperature, (below the ELP critical temperature, causing the ELP to collapse) resulting in non-native 18 nm icosahedral capsids. (*13*)

Using a passive CCMV loading scheme where the coat protein dimers are incubated with a negatively charged polymer (poly[(2-methoxy-5-propyloxy sulfonate)-

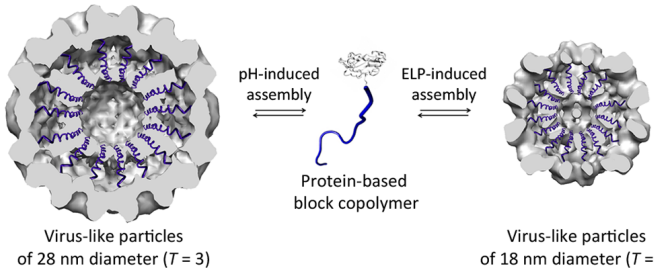



Figure 1.9: By linking a thermally responsive polymer (ELP, purple) to the CCMV coat protein, it is possible to drive assembly of a non-native species by changing the temperature (right). However, the polymer does not disrupt the pH sensitive reassembly to the native species (left). (*13*)

phenyl-ene vinylene], MPS-PPV) and then adjusted to pH 5 to induce capsid formation, even more non-native structures can be accessed (*39*). CCMV naturally packages its negatively charged ssRNA genome thus readily incorporates non-native negatively charged cargos on its interior. Through incorporation of MPS-PPV, T=1 18 nm particles were formed along with an aggregated form made of 2 or more 18nm CCMV particles coming together. The MPS-PPV has a unique emission spectrum depending on if it is in its extended or coiled form, and hence the investigators could establish which form the polymer was in in each capsid outcome. (*40*) As might be expected, the polymer was more coiled in the 18 nm free particles and more extended in the aggregate form. Additionally, the investigators showed that the polymer encapsulation stabilized the

capsids as evidenced by adjusting the pH to 7.5, at which the capsids should dissociate into dimers, however the structures were mostly maintained. (*40*)

DNA is a naturally occurring, regular, predictable polymer that is well suited, because of these properties, for nanomaterial synthesis. Through the engineering of a lipid-DNA fusion, DNA amphiphiles were created that assembled into micelles with a hydrophobic core and an anionic DNA corona (*41*). The negatively charged micelles induced CCMV coat proteins to assemble around them (based on the coat proteins positive charged inner surface) allowing for the entrapment of many small oligonucleotides that were part of the hydrophobic core. The templated capsids were largely 19 nm in diameter suggesting a T=2 structure with a small population of T=1 capsids. An additional functionality could be added to this general scheme by pre-loading the micelles with hydrophobic small molecules (*41*). As proof on concept, the investigators loaded two hydrophobic dyes into the hydrophobic core of the micelle and showed the co-elution of this dye with the assembled capsids from size exclusion chromatography (SEC). Another utility for this system is to hybridize the cargo of interest with a cDNA tag that is complimentary in sequence to the internal DNA of the micelle. To illustrate this, the investigators loaded a hydrophilic dye (which does not want to be incorporated into the hydrophobic core) using a hybridized cDNA tag. Again, the investigators showed entrapment by the co-elution of the dye and the templated CCMV capsid from SEC. (*41*) This DNA-micelle loading and templating scheme has many applications as a wide variety of molecules could be loaded in the micelles and

then entrapped by the CCMV coat proteins. However, additional experiments need to be performed on the stability of these particles *in vivo*.

Another nucleic acid VLP templating strategy exploits a ssDNA-chromophore stack hybrid to create CCMV nanotubes (Figure 1.10) (*1*). ssDNA (poly-thiamine) was first used to assemble long helical stacks of chromophoric molecules via complementary hydrogen bonds. This creates a linear polymer with a negatively charged DNA side and a fluorescent chromophore side. (*1*) The polymer is then used to template CCMV coat proteins, attracted to the negatively charged ssDNA, in to long tubes. Using this method, nanotubes up to 2 μm could be synthesized. Interestingly, this length is 2 orders of magnitude longer than the ssDNA template, suggesting either multiple templates coming together or an extension of the coat protein assemblage. The CCMV nanotubes have not been extensively studied, but they are thought to be rolled sheets of hexagonal assemblies. (*1*) In this case, perhaps the "rolled tube" extends past the template.

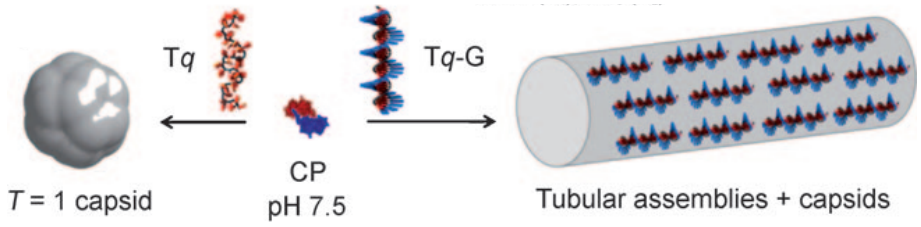



Figure 1.10: Disassociated CCMV coat protein dimers (CP) can be templated to spontaneously form different assemblies. Co-incubating CP with ssDNA (poly-thymine, Tq) molecules yields small closed shells. However a bi-functional rigid structure made of ssDNA bound to fluorescent molecules (Tq-G) templates CCMV tubes. (*1*)

Small Molecule Sequestration by Entrapped Polymers. Polymers are made up of many monomer units and hence have the capability of having many addressable pendant groups (i.e. *n* per monomer). Protein cages are also made of many repeating units and therefore can provide multiple initiator or anchoring sites for polymer growth (i.e. *n* initiator sites per coat protein). The combination of polymers and protein containers is an excellent way to sequester small molecules inside a container.

Using a step-wise click-chemistry based polymer growth inside small heat shock protein (sHSP) the investigators were able to make a polymer that cross-linked the sHSP monomers (*42*). This was achieved by first mutating a cysteine residue on the interior surface of sHSP. Next, the cysteine was reacted with a bromo-alkyne initiator. To the alkyne, a diazido amine was attached using copper (I) catalyzed cycloaddition of the azide to the alkyne. Then the azide was reacted with a trialkyne, creating a branching point. The trialkyne and diazido amine were then cycled until the cage was full. (*42*) The cross-linking of the polymer greatly increased the stability of the protein-cage. Additionally, the addressable groups (amines) on each of the incorporated diazido amines in the polymer were addressable by fluorescein isothiocyanate (FITC), showing loading of ~150 FITC per sHSP. (*42*) A similar system was described cycling with the trialkyne monomer an iron centered tri-phenanthroline complex (*43*). This method incorporates click chemistry polymer growth and coordination polymers into the interior of the constrained reaction vessel of sHSP. (*43*) This approach, when the polymer was loaded with gadolinium, has been studied and shown to have excellent properties as an MRI

contrast agent due to the high loading, water accessibility, and slow rotational correlation time. (*42*)

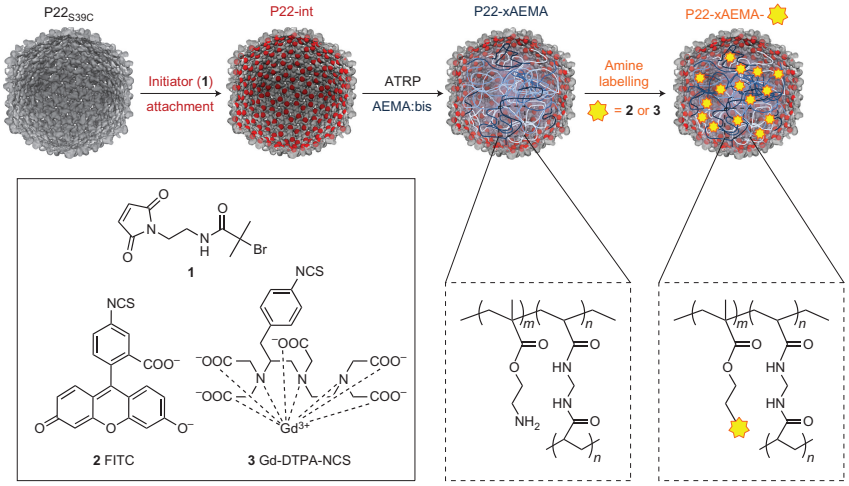



Figure 1.11: Schematic of templated and restricted ATRP inside the P22 VLP. First, the P22 capsid is engineered to contain an interior cysteine (S39C). The docking initiator is then covalently attached to this cysteine providing 420 polymer initiation sites. Next, polymer monomers and catalyst are added to begin polymer growth. The polymer monomers bear a pendent amine group, which was labeled with a fluorescein molecule (FITC) or an MRI contrast agent (Gd-DPTA). (*3*)

While click chemistry polymers are useful for their specific chemistry, they must be grown in step-wise manner. To fill a larger protein container would take many click and purification steps. To alleviate these issues, investigators describe the robust and directed growth of an atom transfer radical polymerization (ATRP) polymer inside the P22 VLP (Figure 1.11) (*3*). Using ATRP, an initiator site was attached to the interior of the protein cage as described above. Next, the protein cage-initiator complex was mixed with the monomer solution and the radical generating catalyst, starting polymer growth. In general, ATRP polymer growth is rapid, robust, and tolerant of many different kinds of monomers, leading to many different functionalities. (*3*) In this example, the polymer is

designed such that the amine pendent group of the monomer was accessible for the loading of large numbers of small molecules. The investigators show loading of fluorescein as proof of concept (and to quantitate the number of addressable groups) and gadolinium as a potential MRI contrast agent. (*3*)

Combining the advantages of quick polymerization using ATRP and the specificity of click chemistry, investigators used ATRP to polymerize the interior of Qβ VLPs with monomers bearing an azide moiety (*44*). The azide-containing polymer was addressed through click chemistry adding up to 650 alkyne-containing small molecules. (*44*) The advantage to this method is that the specific chemistry used for loading the small molecules of interest, results in no side reactions with the protein container itself. Additionally, for both the P22 and the Qβ methods, the VLP acts as a constrained reaction vessel containing the growing polymer.

Exterior Polymers. While containing a polymer inside the protein container may be advantageous, decorating the outside may also be useful. Exterior polymers can be used to shield the particles from the immune system or for adding exterior signals. As protein containers are made of many repeating units, they can be engineered to contain, or naturally contain, many exposed sites for polymer growth and attachment on their exterior.

Using the genome free MS2 VLPs naturally occurring exterior lysines, the exterior of the capsid was densely conjugated with PEG molecules (*45*). Using PEG- *N*-hydroxysuccinimide (NHS) esters, the investigators singly, doubly, and triply label (all in the same sample) capsid monomers on the exterior of the MS2 capsid. This

unprecedented labeling efficiency remarkably did not disrupt the MS2 capsid morphology or stability. The PEG decoration caused a 90% reduction in the binding of anti-MS2 antibodies. (*45*) Similarly, a CPMV VLP was decorated with PEG also via NHS-ester conjugation to exterior lysines (*46*). While this method did not yield as much labeling (0.5 PEG per capsid monomer), the primary immune response was decreased as compared to non-PEGylated CPMV. (*46*) This method of PEG decoration thus effectively shields nanoparticles from an immune response.

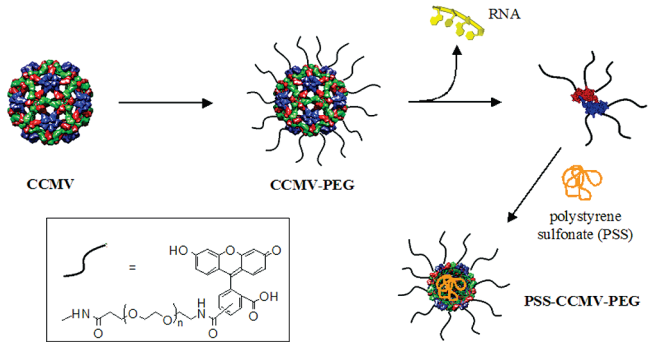



Figure 1.12: Using the CCMV VLP platform, authors show polymer decoration on the inside and outside of the VLP. First, the CCMV was labeled on the exterior with a PEG molecule. The PEG-CCMV was then dissociated into dimers with pH. Incubation of the CCMV-PEG dimers with a negatively charged polymer (PSS) caused their collapse into smaller than wild-type, T=1 18 nm, closed capsids. (*2*)

Investigators describe a two-polymer approach that creates stable CCMV particles with PEG polymer on the outside and PSS polymer on the inside of the VLP (Figure 1.12) (*2*). First, the lysines on the exterior of the CCMV capsids were decorated with NHS-ester-PEG functionalized with a fluorescein. It was reported that the PEG functionalized CCMV VLPs were thermally instable and dissociated into dimers over

time. The investigators hypothesized that the instability arose from steric hindrance of the PEG chains causing disruption of the protein-protein interactions of the coat. Next, the CCMV-PEG VLPs were dissociated into dimers by adjusting the pH to 7.5. (*2*) Polystyrene sulfonate (PSS) was then added to the dimers, assembling the capsids around the negatively charged PSS resulting in non-native T=1, 18 nm capsids. Interestingly, the PSS encapsulation rescued the CCMV-PEG VLPs from their instability. (*2*) This strategy for interior and exterior modification of CCMV with polymers illustrates the variety of functionalities polymers can bring to protein containers.

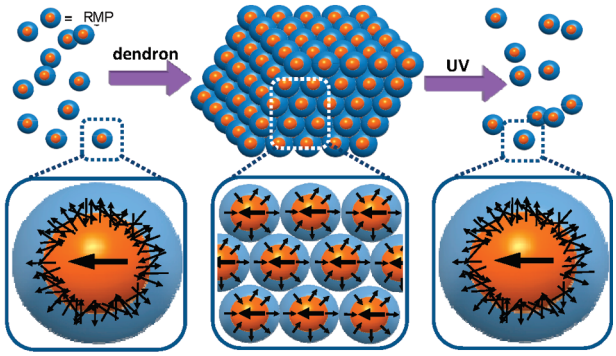



Figure 1.13: Using a photolabile dendron, the authors assemble larger ordered assemblies of magnetoferritin. Upon UV-irradiation, the dendrons degrade, releaseing the ferritin. The assemblage was shown to have different magnetic properties than the free mineralized ferritin. (*10*)

Polymer Directed Assemblages. Another interesting application for exterior polymers, is in the design of hierarchal assemblies with properties different from the dispersed particles. Using photolabile dendrons with terminal spermine groups (positively charged), the investigators made ferritin-magnetite (negatively charged) assemblies of 1.5-2 μm in diameter through ionic interactions (Figure 1.13) (*10*).

Interestingly, through SAXS analysis, the investigators showed that the dendron-initiated complexes had a face centered cubic lattice character which is the same lattice seen in pure magneto-ferritin crystals. Short exposure to UV light causes the degradation of the dendrons, releasing the ferritin nanoparticles into solution. This release causes a distinct change in the nanomagnetism between the free particles and the dendron-ferritin complexes. (*10*) This methodology results in a temporal and spatially controlled release mechanism.

Conclusions. Protein containers can act as constrained reaction vessels for polymer synthesis within them. The repetitive nature of protein container shells, yielding many identical sites encompassing the entire shell, make them excellent homogeneous anchoring sites for polymer growth on the inside or exterior of the container. Entrapped polymers hold the possibility for increased container stability and the addition of many addressable sites for further conjugation. Exterior polymers can be used to shield particles from the immune system and change the solubility of the protein container. The wide array of polymer monomers, each with their different functions, can be incorporated in to protein container design, greatly expanding the utility of protein container based nanomaterials.

## Inorganic Particles in Protein Containers

Similar to restrained polymer growth, protein containers can restrain and contain inorganic particle synthesis. This class of composite materials is inspired by naturally occurring ferritins. Ferritin is a small, 12 nm, protein cage that sequesters free iron ions

and mineralizes them into an iron oxide core. This protein mineralizes iron under ambient, *in vivo* conditions without the aid of enzymes or co-factors. To encapsulate inorganic particles in non-ferritin-like protein containers, nanomaterial scientists have taken two broad approaches: either templating the protein container assembly around a pre-formed particle or by redesigning the protein container such that inorganic particles can be synthesized inside.

Assembly Around a Pre-Formed Particle. An intriguing inorganic nanoparticle to encapsulate is paramagnetic iron oxide because of its usefulness as a MRI contrast agent. When encapsulated, the iron oxide loaded protein containers location can then be monitored through MRI. To encapsulate paramagnetic iron oxide in Brome mosaic virus (BMV), the metal particle was first coated with carboxyl-terminated PEGylated phospholipids (*47*). This polymer attached to the iron nanoparticle through the hydrophobic interactions of the polymer with the oleic acid tails present on the iron oxide particles. The BMV was then dissociated into stable dimers, mixed with the coated iron oxide, and the pH and buffer was adjusted to initiate assembly. The anionic nature of the polymer (mimicking the RNA negative charge of the native cargo) caused the protein dimers to collapse onto the metal particle and assemble into capsids. (*47*) Interestingly, using larger sized nanoparticles than the naturally occurring interior cavity of BMV resulted in different sized capsids, showing that the capsid can conform to a wider range of morphologies than seen in nature. (*47*) Exploiting this same approach, the investigators described the first example of the encapsulation of cubic iron oxide nanoparticles inside a protein container (*48*). The BMV-NP, in both cases, showed

excellent MRI imaging capabilities. In the case of the cubic cores, the investigators showed excellent transport in the plant vasculature. Additionally, MRI of actual plant tissue showed the presence of the BMV composites across different plant cell types. (*48*) The investigators suggested this nanomaterial for the high contrast imaging of different pathways in plants and for the delivery of therapeutic antimicrobial cores to infected plants. (*48*)

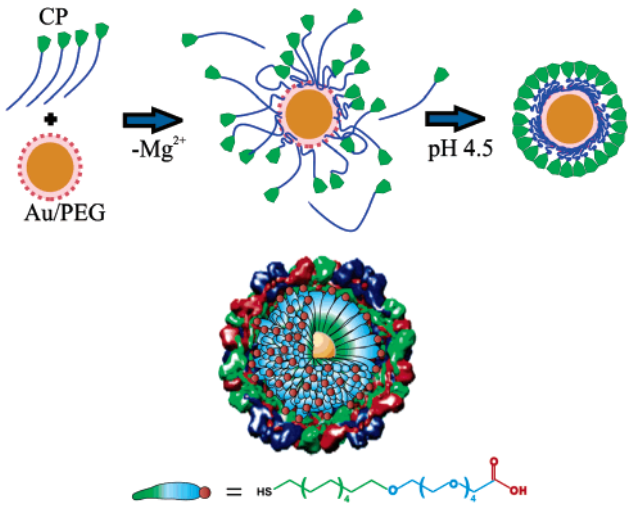



Figure 1.14: BMV coat proteins decorated with polymer (CP) are mixed with functionalized gold nanoparticles (Au/PEG). Adjusting the pH causes the collapse and ordering of the CP-PEG around the gold nanoparticle core forming a closed protein shell. (*9*)

The ease and broad utility of the PEG-decorated iron oxide encapsulation method by BMV, was exploited to also encapsulate pre-formed gold nanoparticles (Figure 1.14) (*49*). Using TEM image reconstruction, the authors examined how the gold core size affects the BMV VLP assembly. Again, it was shown that larger particles than the interior volume of BMV could be contained, accessing new BMV morphologies. The ability to crystalize BMV-Au not only lead to high-resolution reconstructions, but lead to

an interesting new optical behavior. (*49*) The regularity of the VLP allowed for close packing in the crystal and therefore the close proximity of gold nanoparticles, causing multipolar coupling and the appearance of plasmonic banding. The investigators suggested these BMV-Au crystals may be a new class of plasmonic metamaterials. (*49*)

In a later report using the same method, the investigators dig deeper into how the gold core dictates the BMV assembly outcome (*50*). This report suggested that the surface charge density of the core is critical for proper BMV assembly, not the core diameter alone. (*50*) In addition to iron oxide and gold nanoparticles, the investigators used this system to incorporate quantum dots (QD) into BMV, leading to minimal release of photoreduction products and improving the stability of the QDs to prolonged irradiation (*50*). Lastly, the investigators applied this concept to a CCMV mutant that has aberrant assembly (*51*). By introducing TEG-functionalized Au nanoparticles, the assembly was restored to a more native-like distribution. (*51*)

A method that mimics the natural RNA encapsulation strategy was used to encapsulate a range of inorganic nanoparticles into red clover necrotic mosaic virus (RCNMV) (*52*) (*53*). In this method, a thiol-modified DNA, that mimics the stem loop structure found in one of the RCNMV's bipartite RNA genomes (RNA-2), was first conjugated onto the pre-formed gold nanoparticles. Next, RNA-1, the other half of the genome, recognized this stem loop structure, forming the origin of assembly. The origin of assembly recruited coat protein, initiating the capsid assembly process around the preformed gold particle. The resultant nanocomposites incorporated 5 or 15 nm nanoparticles. (*52*) However, when the investigators tried to encapsulate 20 nm particles,

they did not get well-formed VLPs. This is in agreement with the natural interior diameter of RCNMV, 17 nm. (*52*) Using the same method, the investigators also showed the ability to encapsulate $CoFe_2O_4$ and CdSe nanoparticles ranging in size from 3-15 nm (*53*). As seen in the gold nanoparticle example, the authors could not encapsulate particles larger than 17 nm. (*53*) Therefore, unlike the BMV example above where larger nanoparticles could be encapsulated, the RCNMV system does not allow for this flexibility.

Conclusions. The templated assembly of protein containers around preformed inorganic nanoparticles has the advantage that high temperature (and other harsh conditions) can be used to synthesize the inorganic particles before protein is introduced. This encapsulation strategy, in some cases, also allows for the formation of new capsid morphologies through encapsulation of particles larger than the natural interior cavity. Thus, this method extends the utility of the container and gives insight into the protein containers assembly mechanism. However, the disadvantage to this system is that the protein containers must be dissociated and then reassembled, a non-trivial step that leads to product waste (left over coat protein monomers) and empty containers (dissociated pieces coming together on their own). To alleviate these issues, materials scientists have taken the alternative approach of synthesizing the inorganic particle within a pre-formed container.

Synthesis in a Pre-Formed Container. A simple method was used to synthesize monodisperse Prussian blue (PB) nanoparticles inside CCMV (*54*). CCMV VLPs were

incubated with precursor salts, which are negative in charge, and hence accumulate on the positive interior of the CCMV. Next, the sample was irradiated with a 405 nm laser causing the photoreduction of the precursors to PB nanoparticles. In this method, PB particles were made in the solution as well as entrapped in the VLP, but were easily separated from the VLP-PB with low speed centrifugation. (*54*) The CCMV VLP acted as a limiting container resulting in monodisperse PB particles 18 nm in diameter. The investigators then exploited CCMVs capability to form hexagonally packed monolayers on mica and hydrophilized graphite surfaces to show the utility of the PB magnetic nanoparticles for downstream applications, such as information storage devices or other magneto-optic applications. (*54*)

Using a similar co-incubation method, zinc-phthalocyanines (Pc) were encapsulated in CCMV VLPs (*40*). Two methods were employed: either reassembling CCMV coat dimers around the Pcs or co-incubating the Pcs with already formed T=3 VLPs. Interestingly, the reassembled VLPs assembled into T=1 capsids and encapsulated many more Pcs then the co-incubation method. The high concentration of Pcs in the reassembled VLPs caused the Pcs to aggregate into stacks, giving a unique optical property. The co-incubated preformed particles only entrapped 2 Pcs, causing only dimer formation, however the dimerization did trap the Pcs in the capsids. (*40*)

Utilizing the natural affinity of the interior cavity of ferritin for iron, the investigators used *in vitro* methods to mineralize magnetite cores inside the ferritin cage

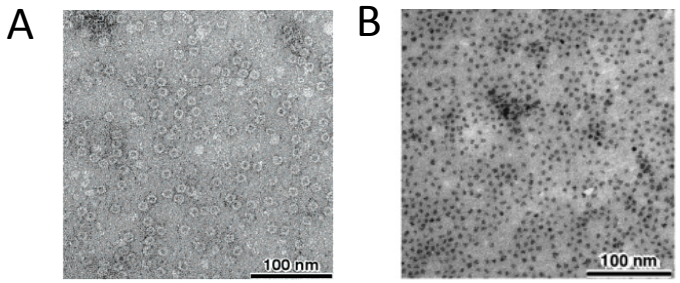



Figure 1.15: TEM micrographs of apo-ferritin (A) and ferritin that has been mineralized with iron oxide (B). The ferritin cage restricts the particle growth resulting in homogenous synthesis. Adapted from Uchida, 2006. (*14*)

(Figure 1.15) (*55*). Preformed apo-ferritin was subjected to a nitrogen atmosphere, 65°C, a constant pH of 8.5, iron salt, and hydrogen peroxide. During the course of the reaction, the iron salts were oxidized and the natural affinity of the interior of ferritin for iron, caused the iron to be mineralized inside the cage. An advantage of ferritin for biomedical use is that it is readily taken up by macrophages and hence has a natural targeting avidity to sites of inflammation, atherosclerosis, and other macrophage rich tissues. (*55*) In this study, the uptake of the ferritin-magnetite composites were evaluated in macrophage cells *in vitro* and assessed for their magnetic resonance imaging capabilities. The investigators showed that the ferritin-magnetite nanoparticles were readily taken up by the macrophages and provide excellent T2 magnetic resonance contrast. (*55*) The same material was taken *in vivo* in mice with artificially induced atherosclerotic plaques (*56*). Here, the investigators showed MRI images illustrating the contrast conferred to the plaques by the ferritin-magnetite nanoparticles. (*56*) Atherosclerotic plaques are difficult to image using MRI because their water content is very similar to their surrounding area. Thus, this ferritin-magnetite material is a promising new contrast agent to image macrophage rich tissues. Further engineering of the ferritin-magnetite material to display

RGD-peptides on the surface, allowed for specific targeting of melanoma cells *in vitro.* (*14*) The ferritin protein cage is a robust and versatile platform as it can be used to synthesize magnetic iron oxide on the interior and the exterior can be exploited for displaying targeting ligands.

Conclusions. The use of preformed protein cages for the synthesis and containment of inorganic nanoparticles has the distinct advantage of being a restrictive container for nanoparticle synthesis. By using protein containers as reaction vessels, it is possible to create very monodisperse sized particles of a controlled size. The containment of the inorganic nanoparticle in a protein container confers the advantages of biocompatibility, as seen in the above ferritin examples, and the added molecular handle of a protein coating. The protein coating can then be used to attach targeting moieties or PEGylated to modulate the immune response.

## Drug Delivery in Protein Containers

As has been discussed, protein containers confer many advantageous properties on their cargo. The internalized cargo is protected from the exterior environment, the cargo is sequestered in high concentration, and the protein container dually acts as molecular handle for further engineering and delivery of the cargo. An exciting application for protein containers is to use them for the delivery of drug molecules *in vivo*. In this case, the protein container can also protect non-diseased tissue from the possible harmful side effects of the drug molecules. Additionally, the protein container can be engineered to display targeting ligands for a specific disease states, such as cancer.

Heat shock protein (Hsp) was used to covalently entrap a chemotherapeutic agent and the release of the toxic drug could be controlled via pH (*57*). First, Hsp was genetically modified to contain a cysteine specifically on the interior cavity. Doxorubicin, a toxic cancer therapeutic, was then synthesized to contain a maleimide functionality such that it could be covalently linked to the new cysteine in the Hsp. The loaded Hsp protein cages retained their same monodisperse morphology and contained 24 doxorubicins per particle. Under low pH (pH 4-5) 50% of the drug was released from the Hsp. This low pH release mechanism mimics the environment the Hsp would encounter after lysosomal entry into cells. (*57*) The investigators additionally showed that this Hsp platform coulc be engineered to display on the exterior surface cancer cell targeting ligands (*58*). For targeting, the RGD peptide (known to target $\alpha_V\beta_3$ integrins) was engineered into the exterior loops of the Hsp coat protein and was shown to specifically target melanoma cells *in vitro*. (*58*) Thus, the Hsp platform is a promising anti-cancer therapeutic as it can entrap and controllably release a toxic drug molecule and this reactivity can be specifically directed to cancer cells.

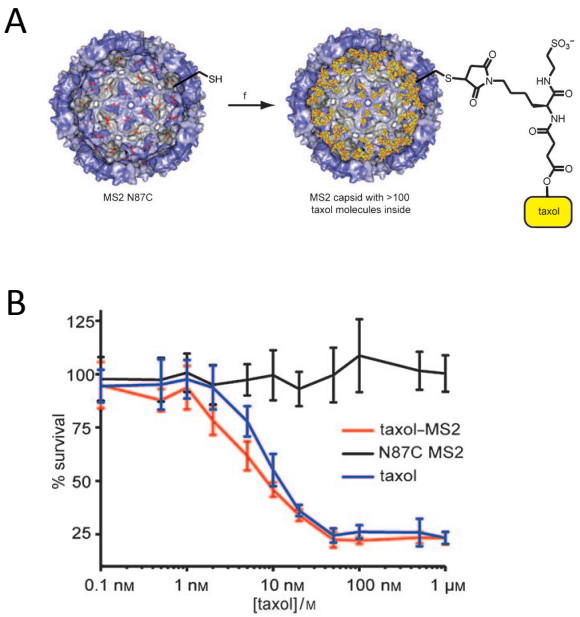



Figure 1.16: (A) The MS2 VLP was engineered with an addressable interior cysteine. This cysteine was then used to load the VLP with a cancer therapeutic, taxol. Over 100 drug molecules were loaded per VLP. (B) Administering the taxol loaded VLP to cells was as effective as administering the drug itself. However, encapsulation protects non-targeted cells from death and also greatly improved the solubility of the drug. Adapted from Wu, 2009. (*18*)

Using a molecular engineered approach that has been discussed above, where the protein cargo is genetically fused to a minor coat protein of the SV40 virus, the investigators show encapsulation of a pro-drug modifying enzyme, yeast cytosine deaminase (YCD) (*11*). As in the fluorescent protein example above, the packaged SV40 VLPs retain the ability to bind and enter host mammalian cells. By treating cells with the SV40-YCD VLP, cells became sensitive to the non-harmful pro-drug as the pro-drug is converted to 5-fluorouracil, which causes cell death. (*11*) This is an amazing example where a non-harmful agent can be delivered to the body and only cells that have taken up the enzyme containing nanoparticles, are sensitive to the drug.

Using an engineered MS2 VLP, a modified taxol molecule (a chemotherapy

agent) was linked to the interior cavity via a new cysteine functionality resulting in 110 taxol molecules per VLP (*18*). The synthesized linker between the taxol and capsid contained an ester-bond cleavable linker, controlling the release of the toxic taxol (Figure 1.16). The investigators showed that the half-life for the release of the taxol from the VLPs in fetal bovine serum (a mimic of *in vivo* conditions) was 1.9 days. The MS2-taxol VLPs were shown to have the same cellular toxicity as free taxol on *in vitro* cancer cells. (*18*) Another advantage of this system is the increased solubility of the taxol drug. Normally, the fairly insoluble taxol has to be delivered via a slow infusion with a toxic detergent. The synthesized linker increased the solubility of the taxol and through containment inside MS2, the taxol could be readily dispersed in phosphate buffer. (*18*)

Conclusions. The biocompatibility, sequestration of large numbers of drug molecules per container, and the shielding of drug molecules from the environment, are all advantages of using protein containers as drug delivery vehicles. The ability to use the protein container as a platform for the display of specific targeting ligands also greatly enhances the utility of these nanocomposites for directed therapeutic use. Combining the advantages of encapsulation, the ability to target cell-types, and the ability to control the release of a drug molecule, make protein containers very promising platforms for the engineering of therapeutic agents.

## Building New Protein Architectures

Although nature has bequeathed many variations of protein containers, it is still valuable to create new protein architectures to access and tune new desirable properties.

Using diverse approaches, materials scientists have developed techniques to bring together protein building blocks to build new hierarchal assemblies. Inspired by nature, these new protein cages offer another tool in the design of functional biomaterials.

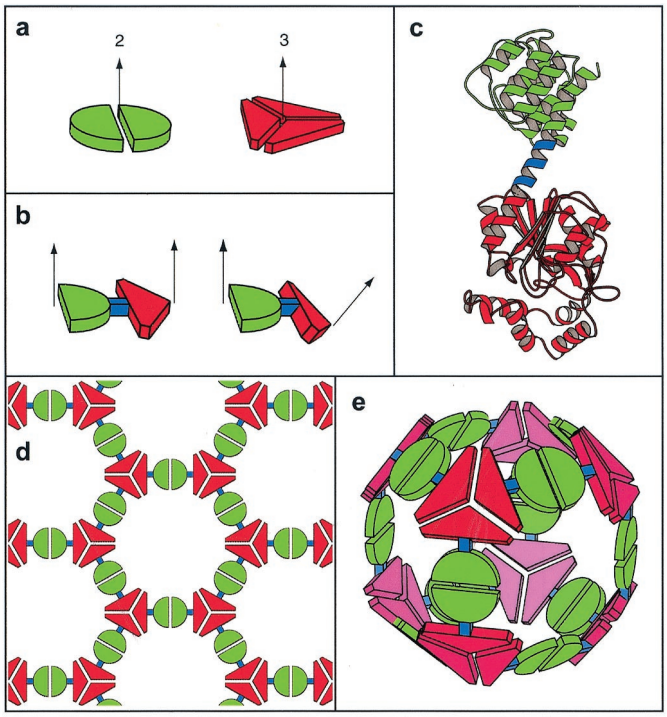



Figure 1.17: Schematic for the intelligent design of novel protein architectures. (A) Through protein mining of the protein databank, two proteins were chosen with either dimer (green) or trimer (red) capabilities. (B) The fusion of the monomers from the two chosen proteins with either a straight or tilted linker provide different nucleation sites for building larger architectures. (C) Ribbon diagram of the fusion of the dimer and trimer monomer. (D) The straight linker yields assembled layers while the kinked linker creates closed shells. The assembly outcomes are directly related to the geometry of the building block input. (*22*)

Using computational methods, proteins were chosen from the Protein Data Bank with desirable properties such that different geometries could be intelligently designed (Figure 1.17) (*22*). First, a protein with trimer and a protein with dimer characteristics were identified and then genetically fused together via a flexible linker. This protein

block, which formed dimers on one end, and trimmers on the other, assembled into stable cage like structures with a diameter of 15 nm. Next, using the same computational strategy, two proteins were chosen that both had dimerization domains. (*22*) These two proteins were fused together with a short alpha helix linker, giving the fused protein a twist in the middle. This engineered block assembled into filaments of end-to-end fusions. Using their design method, the investigators describe "rules" for creating several new classes of protein architectures. (*22*)

Exploiting a strong protein-protein interaction motif, the authors are able to assemble larger assemblies of their enzyme building block (*59*). To engineer their building block, a heterodimeric forming coil-coil unit (either A- or its binding partner A+) was fused to a trimeric enzyme (KDPG aldolase). Through the exploitation of the inherent symmetry of the enzyme (trimer forming), when the fusion is mixed with its coil-coil partner, the engineered fusion proteins self-assemble into several globular structures. Unexpectedly, a limited number of structures were obtained with a closed shell trimer of dimers being the most abundant. (*59*) Unlike the method described by Yeates and colleagues above (*19*), this methodology is more of a generalized approach that can be readily applied to any trimeric protein.

Using a metal as the coordination point of a modified version of the tetrameric cytochrome cb562, the authors describe a new protein cage assembly and its applications (*60*). The inherent zinc binding sites of the cytochrome were used as the linkage points to link the cytochrome building blocks together to form a cage like structure. The cage was then utilized to entrap an unstable 9 amino acid long micro-peroxidase fragment.

The resultant complex allowed for the MP9-cytochrome to be readily crystalized and for MP9, which has never been able to be crystalized, to be structurally characterized by x-ray diffraction. (*60*)

Conclusions. Although the creation of new nano-capsules is interesting, the process is still not yet generalized for easy use. The development of a "designer protein container" by order is still in the future. However, through these synthetic strategies, it is possible to increase the diversity of protein containers available for materials design.

Overall Conclusions

From these examples, the advantages of protein containers in nanomaterials synthesis are clear. Through protein encapsulation strategies, protein containers can become fluorescent packages for *in vivo* imaging and catalytic nanoreactors with desired functions. By exploiting polymer chemistry in conjunction with protein containers, new morphologies can be driven to assemble, many new addressable groups can be added for labeling and sequestration of small molecules, and by putting polymers on the exterior, protein containers can be shielded from the immune system.

Using protein containers as restrictive reaction vessels allows for the monodisperse, controlled size synthesis of inorganic nanoparticles. The high density of inorganic molecules inside the protein container can cause changes to the optical and magnetic properties of the inorganic molecules. Coated inorganic nanoparticles that mimic the charge of the natural nucleic acid cargo can be used to template the assembly of protein containers into native and non-native forms.

Another exciting application of protein containers is in drug delivery. Here, the protein cage has multiple functions: the exterior can be used for displaying targeting ligands, the inside is used to sequester high numbers of drug molecules, and the shell itself is used to protect non-diseased tissue from the toxic cargo within.

The utility of naturally occurring protein containers is so great that some are researching how to increase the breadth of nanocontainers available through synthetic building strategies. The area of protein containers is an ever growing and exciting topic of research. The large number of functionalities that can be added to one protein container causes them to be highly versatile platforms for many different kinds of nanomaterials design.

## Bacteriophage P22

The bacteriophage P22 was identified in 1952 and has played key roles in the development of molecular biology and in the understanding of viral assembly. The natural host of P22 is *Salmonella typhimurium* and the phage can be either lytic, where cells are lysed as a result of infection, or lysogenic, where the P22 genome is incorporated into the host's genome and passed to daughter cells. (*61*) P22 contains a nuclease and such was used to discover general transduction in *Salmonella* (*62*). The genome of P22 has been completely sequenced and 65 genes have been annotated (*63*). The availability of an *in vitro* assembly method has led to extensive studies into its assembly mechanism, and thus P22 being a model system.

Overview of P22 Assembly

The infectious virion of bacteriophage P22 spontaneously assembles into a T=7 icosahedral capsid in the cytoplasm of its host. The assembly progresses via two linear, independent pathways: the capsid formation and the tail machine assembly (*61*). A cartoon of the assembly is provided in Figure 1.18.

Characteristic of dsDNA phages, the P22 assembly has an intermediate procapsid step. The procapsid assembles from 415 copies of the 47 kDa coat protein and 100-300 copies of the 33 kDa scaffold protein. However, one of the 12 vertices is differentiated by a dodecameric ring-assemblage of 88 kDa (monomer molecular weight) portal protein. (*61*) The portal complex provides the site of DNA entry during genome packaging and exit during infection. Also assembling into the procapsid, are 12 copies of each of the three minor ejector proteins: gp7, gp16, gp20. Although their exact function and position on the procapsid assembly are unknown, they are imperative for proper genome delivery and thus successful infection. When assembled, the result is a coat protein exterior shell, encapsulated scaffold protein, and a portal complex (portal protein and the ejector proteins) at one asymmetric unit. (*61*)

After the procapsid is assembled, gp3 and gp2 attach to the portal protein complex forming the packaging and terminase complex (*12*). The packaging protein, gp3, recognizes a specific 22 base pair span of the concatemeric DNA (*64*).

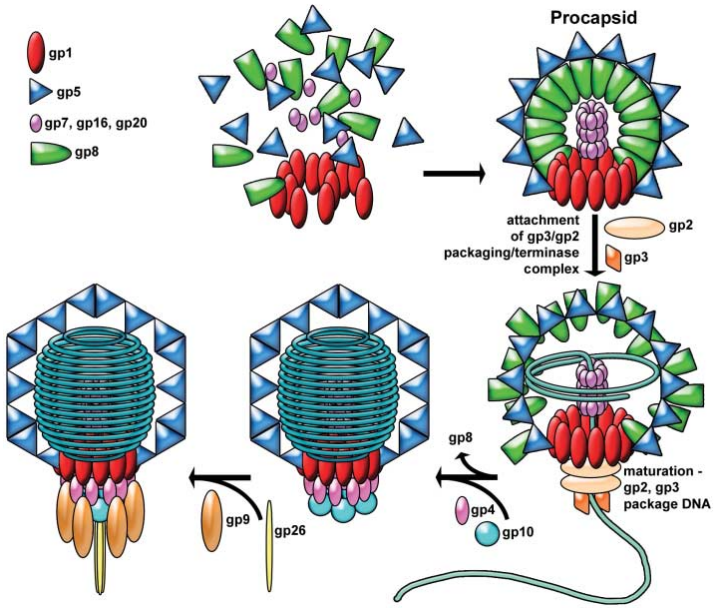



Figure 1.18: Cartoon schematic of the bacteriophage P22 assembly. Coat (gp5), scaffold (gp8), portal (gp1), and the ejector proteins assemble to form the procapsid. Once assembled, the terminase complex is added and DNA packaging ensues. DNA packaging forces the scaffold protein to exit and the coat proteins to mature into the expanded capsid form. When the head reaches a certain pressure from DNA packaging, the tail machine is assembled on to the capsid, resulting in the infectious phage. (*12*)

Once bound, the terminase protein, gp2, then cuts the circular dsDNA into a linear strand, allowing for the end to begin to be packaged, through the portal ring, into the procapsid. The genome is then packaged until approximately 103.5% of the genome is packaged. The actual P22 genome is only 41.7 kilobases (kb) but the phage packages ~43.5 kb. This "over-packing" is a phenomenon none as "headful" DNA packaging. (*64*) Once this critical packing density is reached, the gp2 protein cleaves the remaining DNA strand (*65*). Recently, gp2 was structurally characterized and a loop was identified that inhibited nuclease activity. When removed, nuclease activity was restored. Since this protein is known to be the nuclease in the packaging mechanism (i.e. makes the first and last cut), it is hypothesized that the movement of this loop caused by the initial binding of

DNA, and perhaps the "pressure signal" from packaging, is key to the proteins function. (*65*)

As the DNA is packaged into the procapsid, the scaffold protein exits the capsid. It is hypothesized that the highly flexible scaffold protein exits the capsid through the 2.5 nm pores at the center of each coat protein hexamer, however this has not been experimentally demonstrated. (*61*) As well as scaffold protein exit, the packaging of the DNA causes a global change in the capsid morphology. As the DNA is packaged, the coat proteins under go a structural change, causing the thick procapsid walls to thin, the capsid to become more angular, and the overall diameter of the particle to increase by 10%. (*61*) This morphological shift has been carefully characterized by cryo-electron microscopy (cryo-EM) structural studies, resulting in a model of how the coat protein changes from the procapsid to expanded form. (*66, 67*)

After DNA packaging and the concomitant expansion of the capsid head, the portal complex is closed by the sequential addition of multiple copies of gp4, gp10, and gp26. Now that a stable, packaged head is formed, the trimer-tailspike (gp9) is added, creating an infectious phage. (*61*)

*In vitro*, the only proteins needed to assemble a closed procapsid head, are the coat protein and the scaffold protein. Either purified separately and mixed, or expressing the two proteins together heterologously in *E. coli*, results in spontaneous assembly of wild-type looking procapsids, lacking the asymmetric portal complex. Portal protein can be expressed with coat and scaffold heterologously and be incorporated into one vertex of the capsid, like in the wild-type phage. However, if purified separately, and mixed with

purified coat and scaffold, portal is not incorporated (*68*). This has lead to the hypothesis that there may be some nucleic acid based signal that is need for portal incorporation.

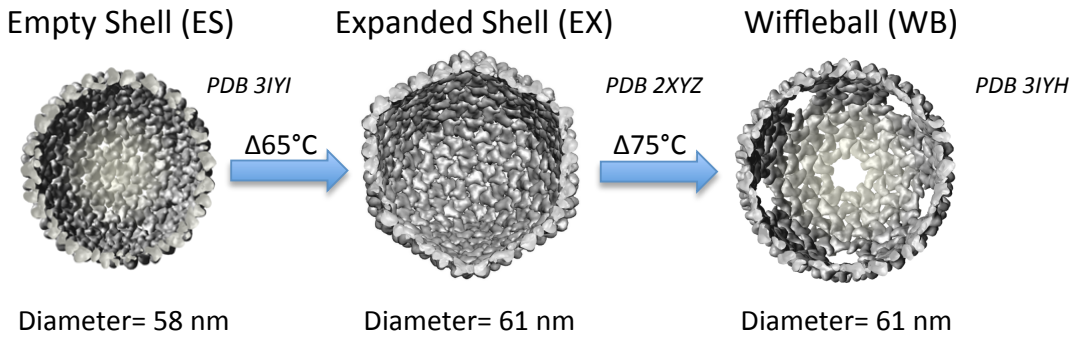



Figure 1.19: The P22 capsid can be assembled heterologously and altered *in vitro*. The procapsid, an assemblage of caot and scaffold protein, can be made into the empty shell via repeated treatment with 0.5M GuHCl. The procapsid or empty shell can be taken to the expanded morphology, which mimics the DNA packaged capsid morphology, by heating to 65°C for 15 minutes. Alternatively, all forms can be taken to the wiffle ball form, with its twelve 10 nm pores, by heating to 75°C for 15 minutes. Structures were made in Chimera (UCSF) from the Protein Data Bank coordinates (PDB numbers listed in figure).

Additionally, the expansion of the P22 capsid from DNA packaging can be mimicked *in vitro* (Figure 1.19) (*67*). Purified capsids can be heated to 65°C for ~15 minutes, and the capsid will expand and the scaffold protein will be released (*69*). An empty procapsid form can also be accessed *in vitro* through repeated treatments of the purified procapsids with 0.5 M guanidine hydrochloride (*70*). A non-biologically relevant morphology can be accessed by heating purified capsids to 75°C. In this form, the pentamer units (the points of most stress in the icosahedral lattice) are ejected, resulting in 12 apparent 10 nm pores (Figure 1.19). (*71*)

P22 Scaffold Protein

The 303 residue long scaffold protein of bacteriophage P22 is of much interest as it plays an important role in the proper assembly of capsids and the full-length protein's structure has been elusive. Another curious property of the scaffold protein is that it auto-regulates its own transcription (*72*). Through the use of the *in vitro* assembly method, where only scaffold and coat protein are mixed to spontaneously form procapsids resembling the *in vivo* procapsid, the scaffold's role in assembly has been highly studied. Additionally, the scaffold protein has been carefully dissected to expose its functional regions.

When coat protein is expressed alone, it assembles primarily into aberrant and unclosed shells (*73*). The scaffold protein plays a transient role in assembly, as it is released from the capsid as the DNA is packaged. The released scaffold can then be reused in up to 5 more assembly reactions. (*61*) *In vitro* analytical ultra centrifugation studies have shown that the scaffold protein forms a mixture of monomers, dimers, and tetramers in solution (*74*). A mutant was made to the scaffold protein that forced dimerization and this was shown to initiate assembly better than the wild-type scaffold protein (*74*). Later studies using FRET active dyes on the scaffold protein gives more evidence that the scaffold proteins are associated into dimers when bound to the coat protein (*19*). Therefore, it is hypothesized that the scaffold protein dimer is the active assembly unit.

It has also been shown through isothermal titration calorimetry experiments, that there are two distinct binding populations of scaffold protein to the procapsid (*75*). The

first, high affinity site, allowed for ~60 scaffolds per capsid. The second binding phase was difficult to fit exactly, due to low binding energies, but was distinctly different than the first. Interestingly, the covalent dimer scaffold mutant, only bound the first high affinity sites. (*75*) The hypothesis of two binding populations fits well with the more recent high-resolution cryo-EM structure of the P22 procapsid, in which the scaffold proteins occupy several different binding sites on the capsid (*76*). In the structure model, the scaffold proteins are shown binding in 6-mer clusters around the coat hexamer unit, 5-mer clusters around the coat pentamer unit, and clusters of three scaffolds at the coat three-fold axis (*76*).

*In vitro,* the scaffold protein can be removed from the procapsid by repeated washing of the procapsids with 0.5 M GuHCl (*70*). This process is reversible, as the scaffold protein will reenter the capsid when the guanidine is removed. The scaffold protein can also be released when procapsids are heated *in vitro* to achieve the expanded morphology. This morphology mimics the fully DNA packaged capsid found *in vivo.* However, once the capsid is in this mature-mimic form, scaffold protein loses the ability to bind to coat protein and hence does not reenter the expanded capsids. (*70*)

The determination of the full-length scaffold protein's structure has proven to be very difficult. The protein will not crystallize and the NMR spectrum shows a highly flexible, non-ordered spectra (*77*). However, efforts to solve the structure of the coat protein binding domain, the extreme C-terminal residues 264-303, has been successful (Figure 1.20) (*78*). This report shows the structure to be a helix-turn-helix motif stabilized by hydrophobic interactions between the two helices. On one face of the

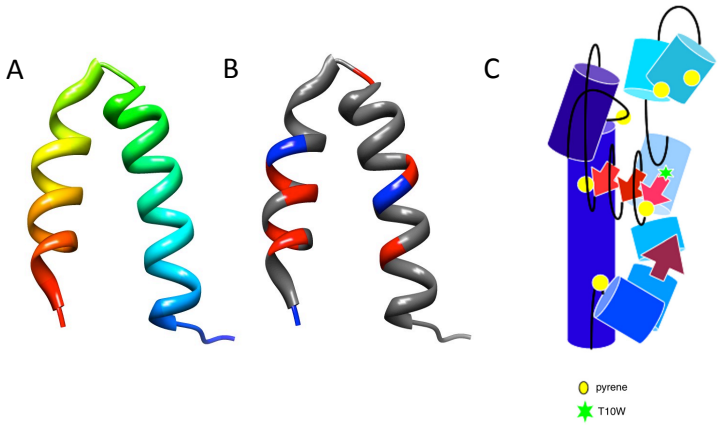



Figure 1.20: The solution NMR structure of the C-terminal end of scaffold protein along with the model for the full-length scaffold monomer. (A) NMR of the C-terminal residues 264-303 rainbow colored from the N-terminus (blue) to the C-terminus (red). (B) The same NMR structure showing the high proportion of positive amino acids in this section of the protein. Red = lysine, Blue = arginine. Images created in Chimera (UCSF) using the published coordinates from the Protein Data Bank (PDB: 2GP8) (C) The folded model of the scaffold protein proposed from secondary prediction calculations and FRET analysis by Padilla-Meier, et al. (*19*) Helices are shown as cylinders and sheets as arrows. The C-terminus is located at the top right hand corner as two cyan cylinders.

protein, there are a large percentage of positive residues (Figure 1.20). (*78*) It has been demonstrated through ionic strength competition assays, that the scaffold protein binds to the coat through electrostatics (*79*). Recently, it has also been shown through the determination of a sub-nanometer resolution structure of the P22 procapsid, that the scaffold protein does, in fact, interact with a negative patch on the inner surface of the coat protein (*76*).

By monitoring the interaction between the FRET pairs at different locations on the scaffold protein, Padilla and co-workers were able to tease out the conformational changes in scaffold protein during coat protein binding and its release (Figure 1.20C) (*19*). In this study, a pyrene molecule was covalently attached to various introduced cysteines (either at the termini, or in the central regions) and was used as the acceptor

molecule in the experiments. As a donor molecule, an engineered tryptophan was added at the N-terminus of the scaffold protein. Pyrene was used because it is not only a suitable acceptor for tryptophan, but when two pyrenes are in proximity, a characteristic emission spectrum is also produced. Using combinations of terminal labeling of the scaffold protein with pyrene and an engineered tryptophan in the middle region of the protein, the scaffold protein folding characteristics were described. (*19*) In solution, the N and C-termini were determined to be nearby, but upon binding to the procapsid, the termini are not proximal. Additionally, parallel dimerization of the scaffold proteins was confirmed through FRET analysis. Through the incorporation of secondary structure modeling with the FRET data and cysteine cross-linking experiments, the authors proposed a model for the scaffold protein as a monomer in solution (Figure1.20C). (*19*)

Taking an *in vitro* assembly strategy, Parker and co-workers revealed the functional domains of the scaffold protein in relation to assembly and scaffold protein packaging density (*80*). In this study, scaffold proteins encompassing amino acids 141-303 (SP303) and 141-292 (SP292) were compared. SP303, and not SP292, were able to assemble capsids that were the same size as wild-type phage *in vitro*. The last 11-amino acids are thus necessary for capsid assembly while the N-terminal 140 amino acids are not. To elucidate if SP292 could bind to coat proteins in the procapsid, the scaffold mutants were added to empty procapsid heads. Wild-type and SP303 could reenter and bind to the procapsids while SP292 could not. This is further evidence that the C-terminal residues are critical to coat protein binding. (*80*) Examining these same mutants using analytical ultracentrifugation, showed both were a mixture of monomers and

dimers (*81*). Wild-type scaffold also resides as a tetramer, thus the tetramer signal must be in the N-terminal region missing from the SP303 and SP292 mutants (i.e. in residues 1-141). (*81*) Another interesting finding from this study was that twice the number of molecules of the SP303 were incorporated into the *in vitro* assembled capsids than the wild-type scaffold protein. As SP303 is about half the length of wild-type, this suggests that the number of scaffold proteins incorporated is limited by the volume of the capsid, not the number of binding sites. (*80*)

Through examining the resultant phage from *in vivo* scaffold protein complementation studies, the different scaffold protein functional domains were further elucidated (figure 1.21) (*17*). In this method, a P22 infection is initiated in *Salmonella* with a phage lacking the scaffold protein gene. This infection is complemented with the gene for various pieces of the scaffold protein, and the resulting assemblies were analyzed. This study showed that the N-terminal 20 residues are non-essential to make infectious phage but truncating 57 of the N-terminal residues had deleterious effects.

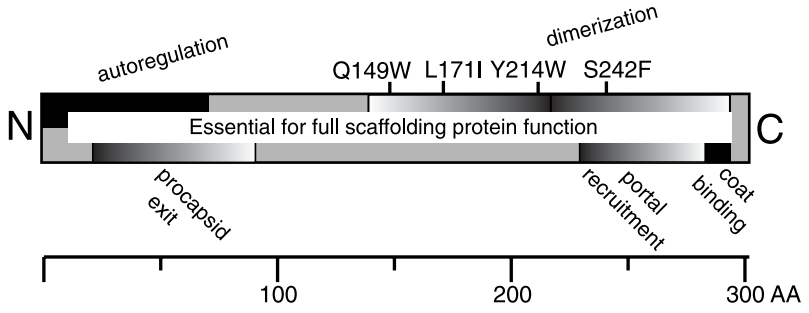



Figure 1.21: Using complementation studies, Weigele et al. determined the regions of the scaffold protein important to scaffold function, phage assembly and infectivity. The diagram covers all 3030 residues of the structure from left to right. Each section is labeled with its proposed function and the gradient demarcates where function is lost (dark) to where the outcome is yet to be determined (light grey). (*17*)

These virions had fewer copies of gp16 (needed for DNA injection), some scaffold protein was retained in the mature phage, and the encapsulated DNA molecule was shorter than in the wild-type phage. When only the C-terminal 66 residues were expressed, phage contained no portal protein. However, when the C-terminal 75 amino acids were expressed, portal was incorporated. Thus, the C-terminal 75 amino acids must regulate portal incorporation. (*17*) In fact, 12 scaffold proteins are shown to bind the portal protein in the recent sub-nanometer cryo-EM reconstruction of the P22 procapsid (*76*). Lastly, through the complementation studies, it was shown that the minimal coat-binding site are residues 280-294. (*17*)

Through its extensive study, the overall structure and functional domains of the P22 scaffold protein have been predicted. This is the best studied scaffolding protein in virology and researchers hope the insight gained from P22 assembly and maturation will lead to new classes of scaffold protein targeted anti-viral therapeutics (*77*).

P22 Portal Protein

The portal protein of bacteriophage P22 assembles into a ring-like structure from 12 monomers. Portal is essential for the infectivity of the phage as it is the docking site for the tail-machinery, which recognizes the host and injects the DNA (*61*). It has been shown that phage with more than one portal are non-infectious (*68*). It has also been postulated that the portal protein acts in gauging how much DNA to package into the capsid (*82*). The portal is extremely intriguing in that it is some how directed to incorporate into only one of the 12 five-fold vertices (illustration Figure 1.22A).

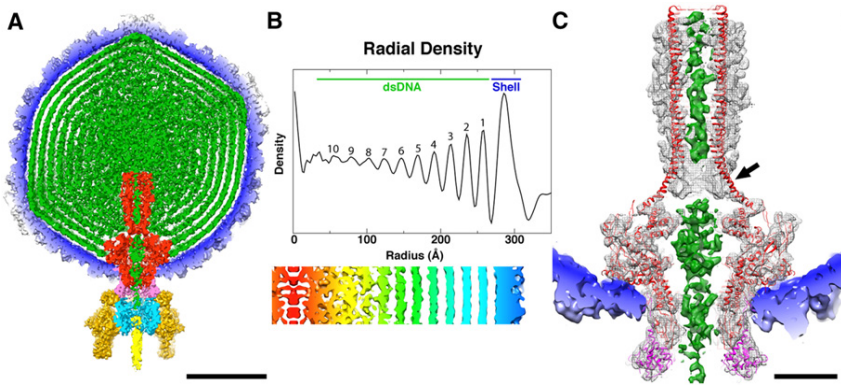



Figure 1.22: (A) Asymmetric cryo-EM reconstruction of the P22 phage complete with portal (red), tailspike (gold), needle (yellow), and packaged genome (green). (B) By graphing the density of a slice across the phage capsid, the highly ordered spacing of the internal genome can be seen. (C) In red, the determined crystal structure of the portal complex is overlaid on the cryo-EM density. Density from DNA in the portal channel can be seen in green. Adapted from Tang, et al. (2011). (*4*)

The most intriguing aspect of the P22 assembly mechanism, is how the single asymmetric unit is incorporated into the capsid. To study portal incorporation, Moore and colleagues created an *in vivo* system by which the infectious phage did not code for the portal protein (*68*). The host cells were then complemented with a controllable vector carrying the portal gene. In this way, the portal expression levels could be controlled by an external stimulus during the infection cycle. The authors utilized radiolabeled methionine in pulse-chase experiments to track portal protein's expression timing and incorporation into capsids. (*68*) Through these experiments, several important findings about portal protein were discovered. First, portal is not essential for the formation of closed, T=7 capsids. This supports the hypothesis that the portal is not the assembly nucleation site. Second, portal will not incorporate into pre-formed capsids either *in vitro* or *in vivo.* However, by expressing portal in the host prior to infection, the authors were

able to fully incorporate portal protein into the resultant phage. (*68*) These two findings suggest that the portal protein is inserted into the assembling phage, not after capsid assembly. However, the "portal pre-loading" experiments showed a direct correlation between greater over expressed portal and fewer infectious phage. Examining the resultant non-infectious phage, showed the incorporation of more than one portal per capsid. The multi-portal capsids were shown to be devoid of nucleic acid, and thus multiple portals results in a packaging defect. (*68*)

Exploiting the fact that over-expression of portal protein causes a severe decrease in infectious phage, Moore and co-workers developed a "portal over-expression escape" mutant (*83*). Through five passages of phage on the portal over-expressing host, isolates were identified that actually depended on the over-expression of portal for infectivity. All mutations from these experiments mapped to either the coat protein, the scaffold protein, or ORF69 (which has no known function). One isolate that conferred the escape phenotype contained only one single amino acid change which mapped to the scaffold protein (G287E). This scaffold mutant was able to restrict the amount of portal incorporated. From this direct evidence that scaffold can regulate portal, it is suggested that the scaffold protein is responsible for proper portal incorporation into the capsid. (*83*)

Portal monomers will assemble into their dodecameric structure when purified and mixed at high concentrations *in vitro* (*84*). The portal ring assembly *in vitro* was slower than for portal assembly in phage, thus it was predicted a structural change must occur to allow for oligomerization. The structural transitions of the portal monomer upon

polymerization have been characterized by far UV- circular dichroism (CD), thermal denaturation, fluorescence, limited proteolysis and Raman spectroscopy. (*84, 85*)

Upon oligomerization, CD indicates an increase in α-helicity and the thermal denaturation studies shows an added structural core present in the dodecamer that is not present in the monomer (*84*). Additionally, limited proteolysis and fluorescent binding assays indicate the N-terminus of the monomer is meta-stable, as it is readily digested and its structure is easily perturbed by the binding assays. (*84*) By Raman spectroscopy, the N-terminal residues had a large amount of change in their side-chain environments, indicating this region is part of inter-subunit interactions. (*85*) More recently, the structure of the portal complex has been determined, from both x-ray crystallography and cryo-EM reconstructions, to contain an N-terminal alpha-helix rich barrel directed to the center of the capsid (Figure 1.23 and Figure 1.22) (*4, 15*). The structures have good agreement with each other and seem to corroborate the biochemical data just discussed.

In the determined structures, the N-terminal tails of each monomer are alpha helices that wrap around each other to form the barrel (Figure 1.23). This barrel has never been seen before in structure calculations and thus must be meta-stable. The density of this structural motif was contributed to the internal proteins gp7, gp16, and gp20 (see figure 1.18) but we now know this density is from the extended barrel of the N-terminus of the portal complex. Deletions to the portal N-terminus result in proper DNA packaging but a loss in infectivity. It is hypothesized that the barrel may act as spooling mechanism for the tightly wound DNA and assist in DNA ejection. However, it is still not known how the barrel structure changes upon packaging and ejection. (*4, 15*)

Although there are now some indications of how the portal is incorporated into one vertex, the exact mechanism is unknown. The determination of the portal complexes structure should lend itself to the further elucidation of the incorporation mechanism. The recent discovery of the interesting alpha-helix barrel structure will assist in the modeling of how DNA is so tightly wound and contained within the capsid.

Overall Conclusions

The bacteriophage P22 has been instrumental in the development of assembly and DNA packaging models in the field of virology. Not only has P22 contributed to the advancement of virology, but the phage has been exploited for materials use as well. Immobilized P22 phage has been exploited for the specific capture of bacteria as a removal and sensing system (*11*) and the great stability of the phage has lent itself for being used as an agent for the assessment of microbial risk on surfaces (*86*). Additionally, the capsid head of P22 has been re-engineered as a nanomaterial for the encapsulation of protein cargo and a constrained reaction vessel for iron oxide nanoparticles and addressable polymers (*7, 36, 87*). The tailspike protein can also be re-engineered to display targeting and functional peptides (*88, 89*). Overall, bacteriophage P22 is a remarkable virus with many applications.

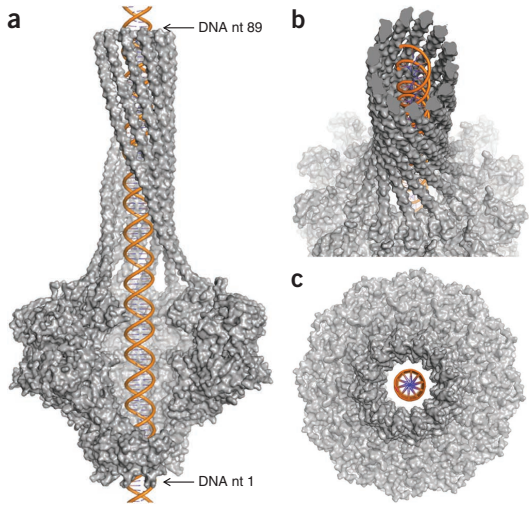



Figure 1.23: Crystal structure of the P22 portal complex with DNA modeled inside the barrel. (A) Front view with some of the portal monomers peeled away revealing the dsDNA inside. (B) Tilted view looking down from the top of the barrel, showing the extreme N-termini of each of the 12 monomers. (C) View looking straight down the barrel of the portal complex. (*15*)

## Research Directions

The bacteriophage P22 has inspired this work with its many natural properties desirable for nanomaterials design. These include a precise assembly mechanism, a naturally occurring encapsulation signal, controllable capsid morphologies, and the controlled incorporation of an asymmetric unit. In the current work, all of these characteristics are exploited and redesigned for the synthesis of the P22 nanomaterial. The scaffold protein, which templates the assembly of P22 capsids and is sequestered inside the capsid, was the first exploited component of P22. Discussed in chapters 2 and 3, the fusion of fluorescent proteins, either singly or as concatamers, to the P22 scaffold protein results in high-copy number sequestration of the cargo and the accurate templating of P22 capsids. In chapter 3, the communication of a protein FRET pair is

shown to be modulated by the capsid morphology. In chaper 4, the same strategy is employed to encapsulate a mesophilic enzyme, phosphotriesterase, inside the capsid conferring greatly enhanced stability to the enzyme. Chapters 5 and 6 demonstrate the utility of short active peptides to the scaffold protein. Lastly, chapter 7 describes strategies for incorporating the asymmetric portal protein complex into the P22 nanomaterial. The robust nature of the P22 system, and its great versatility, lends itself to a wide variety of P22 nanomaterials.

# CHAPTER TWO

## GENETICALLY PROGRAMMED IN VIVO PACKAGING AND CONTROLLED RELEASE OF PROTEIN CARGO FROM BACTERIOPHAGE P22

### Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: Alison O'Neil

Contributions: Writing of manuscript, molecular design, protein purification and characterization

Co-Author: Courtney Reichhardt

Contributions: Assisted in molecular biology and experiments

Co-Author: Benjamin Johnson

Contributions: Performed the HPLC-MALS data analysis.

Co-Author: Peter E. Prevelige

Contributions: Provided the original P22 genes, manuscript editing, and coordinated the project.

Co-author: Trevor Douglas

Contributions: Obtained funding, coordinated the project, and assisted with the design of the experiments. Discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Authors: Alison O'Neil, Courtney Reichhardt, Benjamin Johnson, Peter E Prevelige, Trevor Douglas

Journal: Angewandte Chemie International Edition

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
__x_ Published in a peer-reviewed journal

Publisher: Wiley-VCH Verlag GmbH & Co

Issue in which manuscript appears: Volume 50, Issue 32, (2011) pp. 7425-7428

# CHAPTER 2

# GENETICALLY PROGRAMMED IN VIVO PACKAGING AND CONTROLLED RELEASE OF PROTEIN CARGO FROM BACTERIOPHAGE P22

## Introduction

Viruses have evolved exquisite selectivity for complex hierarchical assembly of components directed towards the programmed packaging and sequestration of cargo molecules. In addition, protein-based containers are increasingly recognized for their ability to act as isolated intracellular compartments with unique and selective catalytic functionality. (*20, 23, 90*) The inspiration drawn from these systems has provided blueprints for synthetic encapsulation approaches based on programmed interactions in self-assembling nano-container systems. Heterologously expressed protein cage architectures are excellent candidates for such engineered systems, based on their self-assembly from a limited number of protein subunits into symmetrical, highly monodisperse structures, the ability to package or release the cargo, and the relative ease of large scale production and purification. This approach has been used with icosahedral viral capsids to encapsulate enzymes, other proteins, and metal cores. (*8, 11, 21, 38, 49, 90*) Furthermore, a wide range of other protein cage architectures have been used to incorporate active sites and/or fluorescent proteins and to develop an *in vivo* selection for packaging based on the cellular toxicity of free (un-packaged) proteases. (*6, 91-93*)

The design of effective, widely applicable, nano-containers for programmed cargo encapsulation requires a capsid structure with sufficient internal volume for the cargo of

interest to be loaded in high copy-number, physical stability over a wide range of conditions, and the incorporation of molecular recognition for both specific packaging of the cargo and programmed assembly. A number of approaches have been developed to achieve programmed encapsulation but in all cases the copy-number of the encapsulated cargo has been relatively modest. (*11, 16, 26, 94, 95*) Here we report an approach for the programmed encapsulation of gene product fusions, in very high copy number, using the capsid and scaffold protein machinery from the bacteriophage P22 in a heterologous expression system using fluorescent protein encapsulation as proof-of-concept. Having established the methodology for genetically programmed self-assembly and encapsulation, the well designed multiple cloning site described below will allow us to easily substitute a wide range of gene products into a fusion site for programmed encapsulation.

## Results and Discussion

Our genetically engineered system uses the *Salmonella typhimuriam* bacteriophage P22, which assembles from 420 copies of the coat protein subunit into an icosahedral capsid with the assistance of 100-330 co(*66, 73*)pies of a scaffold protein (SP)The C-terminus of the SP interacts with the coat protein and is necessary for self-assembly while the N-terminus can be severely truncated or mutated with little to no effect on assembly. (*17, 80*) Our system genetically fuses cargo proteins to a truncated form of the P22 scaffold protein, which acts as a template for the capsid assembly as well as a specific encapsulation signal for the cargo. In this way, additional space within the

capsid from the truncation of the SP, allows the cargo to be assembled *in vivo* with high packaging efficiency. This method does not alter the coat protein in any way and eliminates the need for affinity tags, chemical linkers, or bridging ligands and perhaps avoids mis-assembled virus particles. Instead, the scaffold protein-cargo fusion naturally provides the necessary affinity to the interior of the capsid while templating assembly around it.

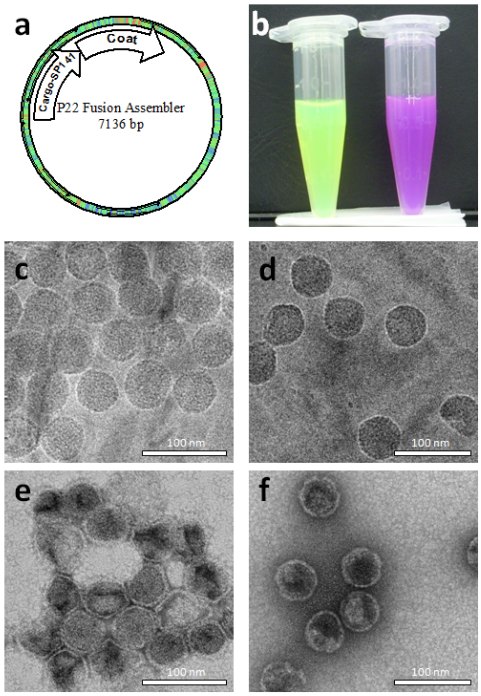



Figure 2.1. (a) Plasmid map illustrating the gene order of the cargo-scaffold protein fusion and P22 coat protein. (b) Purified P22 procapsids containing either EGFP -$SP_{141}$ (left) or mCherry-$SP_{141}$ (right). (c) Cryo-TEM image of P22 procapsids encapsulating EGFP -$SP_{141}$ or (d) mCherry-$SP_{141}$. (e) TEM images of P22 EGFP -$SP_{141}$ heated to 65°C and (f) 75°C.

Through genetic manipulation of the P22 assembler plasmid (*96*), a vector was designed containing the gene for P22 coat protein and a truncated variant of the wild-type

(WT) scaffold protein encoding for amino acids 141-303($SP_{141}$). A thrombin cleavage site and a new multiple cloning site (MCS) was included upstream of the $SP_{141}$ gene. (Figure 1a, Appendix A.1) This new MCS allowed for the insertion of a series of fluorescent protein genes (EGFP and mCherry (*97*)) so as to create N-terminal fluorescent protein-$SP_{141}$ fusions. In these protein fusions, the thrombin cleavage site is located between the $SP_{141}$ and the fluorescent protein and acts as both a flexible linker sequence and an accessible cleavage site to release the fluorescent protein from the $SP_{141}$ as needed.

As expected, the P22 fluorescent protein- $SP_{141}$ procapsids were indistinguishable from the WT P22 procapsids during the expression and purification process except for their brilliant color indicating the expression, encapsulation, and proper folding of the fluorescent protein (Figure 1 b). On average, the expression yields were 150 mg of encapsulated fluorescent cargo P22 per liter of E. coli culture by way of a single step sucrose cushion ultracentrifugation purification. Assembly of the capsids from the SP141 –fusion vector was confirmed by dynamic light scattering (DLS) which yielded hydrodynamic radii ($R_H$ of 25 ± 2 nm and 24.5 ± 1.5 nm for the EGFP and mCherry constructs) and negative stain transmission electron microscopy (TEM) revealed particles of 22.5 ± 1 nm and 24.5 ± 1.5 nm in radius for the two constructs. Additionally, cryo-TEM confirmed the assembly of homogenous sized capsids ranging in radius from 24 nm to 25 nm for both constructs (Figure 2.1 c,d). Masses corresponding to the coat protein (46596 Da) and $SP_{141}$–fusions (45515 Da for EGFP-$SP_{141}$ and 45994 Da for mCherry-$SP_{141}$) were confirmed by ESI-MS and by SDS-PAGE Appendix A.2-3).

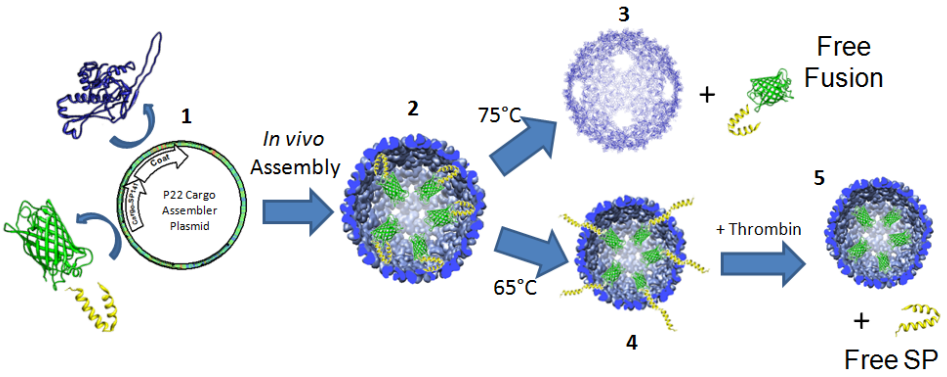



Figure 2.2: Illustration of the translation products from the P22 cargo assembler plasmid (coat protein and the cargo-scaffold protein fusion, 1) and how these spontaneously assemble *in vivo* into a packaged P22 procapsid (2). Heating the procapsids to 75°C results in a wiffle ball capsid morphology and the release the cargo (3). Heating to 65°C accesses the expanded capsid morphology (4) exposing the engineered thrombin site between the cargo and the scaffold protein. This site can then be cleaved to release the scaffold protein but retain the cargo inside the capsid (5).

Using a preparative sephacryl S-500 size exclusion chromatography (SEC) column, with the resolution to separate any aggregate from the large P22 viral capsids, the visible absorbance corresponding to the fluorescent protein fusion co-purified with the capsids (Figure S4). Using HPLC-size exclusion chromatography coupled to multi-angle light scattering (MALS) ($\lambda$ = 662 nm, outside the absorbance of EGFP and mCherry), the diameter and total molecular weight of the constructs were determined (Figure Appendix A.5). The molecular mass determination of particles in solution by MALS is based on the physical property that larger particles scatter more light. If the concentration and the dn/dc of the scattering element are known, then the Mw is proportional to the amount of light scattered. Particle sizes were similar to the non-fusion ($SP_{141}$ only) control capsids (average $R_g$ of 22.5 nm ± 1 for EGFP-$SP_{141}$, 22.5 nm ± 0.8 for mCherry-SP141, and 23.3 ± 0.3 nm for the SP141). Substantial mass increases were

observed for the fusion encapsulated capsids when compared to the WT P22 capsids. By MALS the WT P22 control capsids had an average Mw of 21.7 MDa while the EGFP-$SP_{141}$ capsids increased to Mw of 32.4 MDa and to 30.53 MDa for mCherry-$SP_{141}$ capsids. Based on the MALS data (with an estimated 5% error) approximately 281 EGFP and 233 mCherry fusions are contained within each P22 capsid. This correlated well with estimates based on UV-Vis spectra (~345 EGFP and ~150 mCherry fusions per capsid).

It has been shown that P22 undergoes a structural maturation from a spherical procapsid containing SP to an angular icosahedral mature phage 64 nm in diameter (a 10% increase), lacking scaffold protein and packed with dsDNA (*71, 73*). This structural transition can be mimicked *in vitro* by gently heating the heterologously expressed procapsids to 65°C, which expands the capsid and releases the scaffold protein. Further heating to 75°C releases the pentamer unit from the icosahedron creating an extremely stable "wiffle ball" structure in which the shell is transected by 12 apparent 10-nm pores, but which is otherwise identical to the mature phage (*71*).

The P22 $SP_{141}$-fusion capsids show the same temperature stability as WT P22 and undergo the same heat induced morphological changes, which we have taken advantage of in the product (Figure 2.1e-f, Appendix A.4c-d, 2). When heated to 65ºC to the expanded shell form, which lacks scaffold protein in the WT P22, the scaffold protein-fusion was still constrained inside the capsid as evidenced by the fluorescent signal of the fusion protein, which co-eluted with the P22 capsid by SEC (S-500 column) (Figure Appendix A.4). By agarose non-denaturing gel, the sample population of both fusions is

approximately 50% shifted to the expanded form (Figure Appendix A.6). WT P22 also shows this dual population, when heated to 65ºC, by agarose gel (Figure Appendix A.6). By MALS, both fusion capsids retain their cargo upon heating to 65ºC and the Mw decreased only slightly from 32.4 MDa to 31.8 MDa and from 30.53 MDa to 29.5 MDa for EGFP-$SP_{141}$ and mCherry-$SP_{141}$ containing capsids respectively. The Rg of the EGFP- $SP_{141}$ capsid increased by ~11% (22.5 to 25.2 nm) and the mCherry-$SP_{141}$ capsid increased by an average ~8.5% (22.5 to 24.3 nm), indicating that the capsids undergo expansion. However, the addition of the fusion protein to the SP141 makes it too large to escape the capsid via the normal route and these fluorescent cargos remain trapped inside the capsid during the expansion and are stable to 65°C.

Upon heating to 75ºC to form the wiffle-ball structure it was found that most of the fluorescent signal associated with the P22 capsid fraction had disappeared and was found in a later eluting fraction on SEC (Figure Appendix A.4). This suggests that the fluorescent cargo is free to diffuse out of the wiffle-ball form of the P22 capsid. The average Mw of these samples is severely reduced suggesting the capsid has lost most of the pentamer units and the encapsulated cargo (Mw for GFP-$SP_{141}$ capsid is 20.8 MDa, Mw for mCherry-$SP_{141}$ capsid is 22.69 MDa, predicted Mw of WT P22 capsid only with no pentamers is 17 MDa). Additionally, the MALS data shows a slightly increased average radius ($R_g$ = 26.2 nm for EGFP-$SP_{141}$ and $R_g$ = 25.2 nm for mCherry-$SP_{141}$) suggesting the total sample population has shifted to an expanded capsid form. A non-denaturing agarose gel also confirms the total population morphology shift to the wiffle-ball form (Figure Appendix A.6) and TEM confirms the capsids are intact (Figure 2.1f,

Appendix A.4d). This data indicates that the capsids with internal fusions are still able to undergo the structural transition to the wiffle-ball form and that the holes in this morphology of P22 are large enough for the $SP_{141}$-fusion to leave the capsid.

To release the fluorescent protein cargo from the $SP_{141}$, the fusion capsids were first heated to 65°C to change the capsid to the expanded form which disrupts the interaction between scaffold protein and the capsid. The expanded fusion capsids were then treated with excess thrombin overnight. Unheated, thrombin treated samples had the same diameter and Mw as non-thrombin treated P22 $SP_{141}$-fusion procapsids (thrombin treated EGFP-$SP_{141}$ capsids had Mw of 33.1 MDa and $R_g$ of 22.6 nm while thrombin treated mCherry-$SP_{141}$ had Mw of 31.5 MDa and $R_g$ of 22.5 nm). The heated and thrombin treated P22 $SP_{141}$-fusion capsids exhibited a diameter corresponding to the expanded morphology and displayed a Mw loss equivalent to the loss of the $SP_{141}$ but still retaining the fluorescent protein (heat and thrombin treated EGFP-$SP_{141}$ capsids had Mw of 29.6MDa and $R_g$ of 23.3 nm while heat and thrombin treated mCherry-$SP_{141}$ had Mw of 29.7 MDa and $R_g$ of 23.5 nm). Chromatographs of the heated and thrombin treated samples showed co-elution of the fluorescent signal with the capsid, further illustrating the retention of the fluorescent cargo inside the capsid (Figure Appendix A.4). The release of the fluorescent protein from scaffold protein was additionally confirmed by SDS-PAGE (Figure Appendix A.7).

## Conclusion

In conclusion, we have demonstrated the implementation of a genetically based system for the production of a versatile nano-container that can direct the packaging of cargo through use of a programmed assembly system. Through our system, in vivo encapsulation of engineered cargo in high copy number can be achieved easily and reproducibly. Furthermore, by exploiting the naturally occurring morphologies of the P22 phage, which appear to be unaffected by the fusion protein encapsulation, we have shown the ability to control the release of the cargo from the capsid container (see Table Appendix A.1 for summary). This demonstrated control of capsid morphology may also allow for directed substrate access; a necessity for encapsulated enzyme systems. The added advantage of retaining the cargo while releasing the SP fusion appendage, may prove useful for substrate access issues and oligomerization of enzymes in this system. Using the designed P22-cargo assembler plasmid with introduced multiple cloning sites, it should also be possible to express and encapsulate a wide range of gene products using the scaffold protein as both a mediator of programmed capsid assembly and a cargo-encapsulation director. The degree of control over the release of and access to cargo in either the fusion or non-fusion form makes this a truly versatile platform and is a promising starting point for further programmed encapsulation using genetic manipulation as our synthetic approach.

## Experimental Section

To create the fusion assembler plasmid, standard PCR was used to amplify the gene for the truncated scaffold protein (AA 141-303), adding a 5' NdeI and 3' BamHI recognition sequence, a new multiple cloning site (MCS) and a thrombin recognition site (product organization: 5'-NdeI-SacI-NcoI-*thrombin site*-SP141 gene-STOP-BamHI-3'). This product was inserted to replace the full-length scaffold protein in the WT P22 assembler plasmid(*96*). The gene for either EGFP or mCherry was cloned with the stop codons omitted and ligated into the new MCS resulting in a genetic fusion. The plasmid was transformed into XL-2Blue ultra competent cells (Agilent Technologies, Santa Clara, CA) and colonies were screened for insert. The confirmed sequence fusion assembler plasmid was subsequently transformed into BL21 *E. coli* (EMD Chemicals) for protein expression.

Transformed BL21 *E. coli* were grown to an OD600nm = 0.6, induced with 1mM IPTG, and grown an additional 4 hours. Cells were then lysed with DNase, RNase, lysozyme (Sigma Aldrich) and sonication. Clarified lysates were then subjected to ultra centrifugation through a 35% w/v sucrose cushion. Virus pellets were resuspended in PBS pH 7.6 and further purified over a sephacryl (S-500) size exclusion column (GE Healthcare).

For HPLC-multi angle light scattering analysis, P22 capsid samples were separated via size exclusion chromatography (WTC-0200S column, Wyatt Technologies) and monitored using both a UV-vis detector (Agilent) and a refractive index detector (Wyatt) to determine concentration. Multi angle light scattering was detected with a Dawn 8

(Wyatt). Average Mw and Rg were determined from a fit to the data using a Zimm plot in the Astra software (Wyatt).

For more details, please see Appendix A.

# CHAPTER THREE

## CO-CONFINEMENT OF FLUORESCENT PROTEINS: SPATIALLY ENFORCED COMMUNICATION OF GFP AND MCHERRY ENCAPSULATED WITHIN THE P22 CAPSID

### Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Alison O'Neil

Contributions: Molecular design, designed and carried out the experiments and wrote the manuscript.

Co-Author: Gautam Basu

Contributions: Assisted in designing experiments and was the major contributor of the data analysis

Co-Author: Peter E. Prevelige

Contributions: Provided the original P22 genes, manuscript editing.

Co-author: Trevor Douglas

Contributions: Obtained funding, coordinated the project, and assisted with the design of the experiments. Discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Authors: Alison O'Neil, Gautam Basu, Peter E. Prevelige, Trevor Douglas

Journal: Biomacromolecules

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
__x_ Published in a peer-reviewed journal

Publisher: American Chemical Society
Issue in which manuscript appears: Volume 13, Issue 12, (2012) pp. 3902-3907

# CHAPTER THREE

## CO-CONFINEMENT OF FLUORESCENT PROTEINS: SPATIALLY ENFORCED COMMUNICATION OF GFP AND MCHERRY ENCAPSULATED WITHIN THE P22 CAPSID


### Abstract

The precise architectures of viruses and virus-like particles are proving to be highly advantageous in synthetic materials applications. Not only can these nano-containers be harnessed as active materials, but they can be exploited for examining the effects of in vivo "cell-like" crowding and confinement on the properties of the encapsulated cargo. Here we report the first example of inter-molecular communication between two proteins co-encapsulated within the capsid architecture of the bacteriophage P22. Using a genetically engineered three protein fusion between the P22 scaffold protein, GFP, and a red fluorescent protein (mCherry), we were able to direct the encapsulation of the genetic fusion when co-expressed with P22 coat protein. These self-assembled P22 capsids are densely packaged, occupying more than 24% of the available volume. To probe this crowded nano-environment, we examined the distance dependence of FRET inside the capsid compared to the same tethered FRET pairs free in solution. The P22 system is unique in that the capsid morphology can be altered, without losing the encapsulated cargo, resulting in a doubling of the capsid volume. Thus we have additionally examined the encapsulated fusions at two different internal concentrations. Our results indicate that FRET is sensitive to the expansion of the capsid

and encapsulation enforces significant inter-molecular communication, increasing FRET by 5 fold. Additionally, the enhancement of the encapsulated FRET signal elucidates the distance dependence of FRET as compared to the unencapsulated FRET pairs. This P22 co-encapsulation system is a promising platform for studying crowding, enforced proximity, and confinement effects on communication between active proteins.

## Introduction

Viruses and virus-like particles (VLPs) have been appreciated by a wide range of disciplines for their bottom-up self-assembly to form highly symmetric and monodisperse architectures and for their chemical and genetic plasticity. (*1, 5, 90, 98*) These versatile bio-compatible vessels have been used to encapsulate a range of catalytic centers, addressable polymers, drugs, and active proteins such as enzymes and fluorescent proteins. (*3, 6-8, 29, 36, 40, 43, 48, 93, 94, 99*) VLP encapsulation has the ability to direct high density packaging of cargo and an extension of this approach is the directed co-encapsulation of interacting partners. The allowed, or enforced, communication between partners is of considerable interest to develop and understand spatial and concentration dependent interactions in crowded nano-scale environments. (*90*) Protein-protein communication can be probed by Förster resonance energy transfer (FRET), which has a strong distance dependence ($1/r^6$). (*100*) Additionally, the power of FRET has been realized to probe structural characteristics of virus particles. (*19, 101*) By comparing the FRET communication between proteins encapsulated within a VLP, or

free in solution, we have a sensitive tool for measuring the effects of spatial confinement on their interactions.

We have developed a system by which genetically fused FRET partners, GFP and mCherry, separated by oligomeric glycine spacers, can be encapsulated within the P22 bacteriophage capsid with high local concentration (Figure 3.1). Based on the naturally occurring scaffold protein templated assembly of the bacteriophage P22 capsid, our system genetically fuses the protein cargo of interest to a truncated version of the scaffold protein resulting in spontaneous *in vivo* assembly of P22 capsids containing many copies of the cargo-scaffold fusion. In addition, gently heating the P22 capsids to 65°C results in a well-characterized expansion of the capsid, which nearly doubles the interior volume while maintaining the encapsulated fusions. (*67*) The P22 encapsulation system described here is thus ideally suited to study inter-molecular communication in crowded nano-environments. Our system not only specifically packages the desired cargo in high copy number and high local concentration, but the container is controllably modular, allowing us to investigate the same communicating pairs in two distinctive environments.

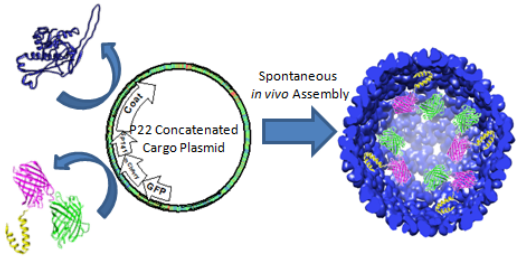



Figure 3.1: A representation of the P22 encapsulation system. The genes for GFP, mCherry, and truncated P22 scaffold protein ($SP_{141}$) were expressed as one genetic fusion. This triple fusion was co-expressed with P22 coat protein, which results in the spontaneous self-assembly into the P22 particles containing ~100 copies of the fusion.

## Results and Discussion

We have constructed a triple protein fusion, consisting of the truncated scaffold protein, GFP, and mCherry catenated together as a single poly-protein (GFP-mCherry-SP). The P22 scaffold protein directs the assembly of T=7 icosahedral capsids when co-expressed with coat protein in an *E. coli* expression system. Building on previous work demonstrating the encapsulation of individual fluorescent proteins in P22 (*102*), we have

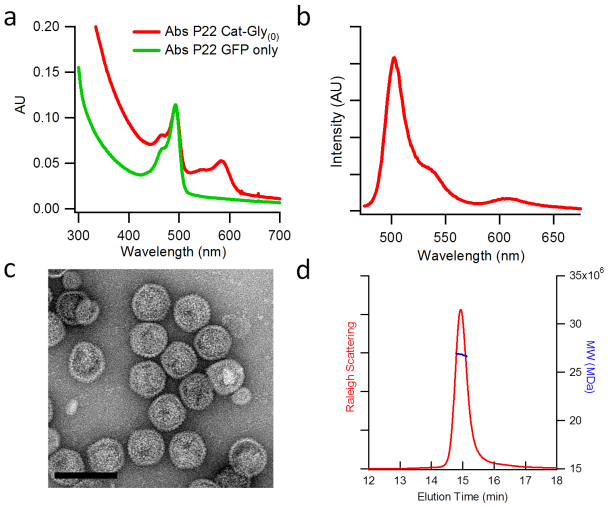



Figure 3.2: Characterization of the P22-Cat-Gly$_{(0)}$ construct. (a) Absorbance spectra of GFP and GFP-mCherry encapsulation inside P22 showing the characteristic GFP (495nm) and mCherry (585nm) absorbance. (b) Emission spectrum of the purified P22-Cat-Gly$_{(0)}$ construct showing emission from both GFP (505nm) and mCherry (610nm). (c) TEM of purified P22-Cat-Gly$_{(0)}$ (scale bar is 100 nm). (d) Size exclusion chromatography and MALS of the monodisperse P22-Cat-Gly$_{(0)}$ showing the scattering (red) and the corresponding Mw (blue) of the elution profile.

expanded this bio-engineered approach to co-encapsulate GFP and mCherry, at an exact 1:1 ratio, in high copy number, into a single capsid. The resulting genetic construct was sequenced and expressed in *E. coli* and the particles purified by ultracentrifugation and

size exclusion chromatography prior to their physical characterization. Importantly, the triple fusion of GFP-mCherry-SP is still able to template the assembly of the icosahedral P22 capsids (P22-Cat) with high fidelity.

Once encapsulated in the P22 capsids, the fluorescent proteins maintain their characteristic absorbance and emission profiles. This indicates that the two fluorescent proteins are each assembled into their fully functional form. The purified P22-Cat showed absorption peaks at 495nm for GFP and 585nm for mCherry (*97*) as well as the expected emission at 505nm for GFP and 610nm for mCherry (*97*) (Figure 3.2a,b).

The purified P22-Cat capsids were found to be extremely homogenous. The P22-Cat was characterized by transmission electron microscopy (TEM) and found to have a diameter of 47± 4 nm and to be morphologically indistinguishable from other recombinant P22 capsids (Figure 3.2c). (*7, 102*) Measurement of the particle size using HPLC-size exclusion chromatography coupled to multiangle light scattering (MALS) revealed particles with a radius of gyration ($R_g$) of 22.7nm and by dynamic light scattering the particles had hydrodynamic radii ($R_H$) of 30.4 nm (Table 3.1). From the MALS data an average particle Mw of 27.5±0.16 MDa was obtained, which corresponds to 111 ± 2 GFP-mCherry-scaffold protein fusions per capsid (Table 3.1). The Mw profile across the SEC elution profile revealed a monodisperse population highlighting the homogeneous packing of the fusion proteins (Figure 3.2d) within each particle. The capsids exhibit dense packaging with occupancy of greater than 24% of the capsid interior (see Supporting Information for calculation).

The P22-Cat capsids retained the characteristically plastic morphological transformation from procapsid (PC) to expanded (EX) forms upon heating to 65ºC, which results in a doubling of the internal volume and provides a unique opportunity for investigating the effects of confinement on the behavior of the encapsulated cargo proteins. The transformation was probed by non-denaturing agarose gel electrophoresis where co-migration of the fluorescence and coomassie protein stain confirm that the EX capsids retain the fluorescent protein cargo (Figure Appendix B.3). This was further confirmed by the co-elution of both the GFP and mCherry absorbance, together with the EX capsid, when analyzed by size exclusion chromatography (Figure Appendix B.2).

To probe the effects of enforced communication between cargo proteins within our viral capsids, we engineered a series of glycine spacers between the co-encapsulated GFP and mCherry catenated fusions and exploited FRET to probe their interactions. The resulting fusions have 6, 12, or 18 glycine residues between the FRET donor and acceptor. These constructs were grown and purified in the same manner as P22-Cat-$Gly_{(0)}$ and resulted in the spontaneous *in vivo* assembly of icosahedral P22 capsids containing both GFP and mCherry. All glycine spacer samples showed the characteristic GFP and mCherry absorbance and emission profiles (Figure Appendix B.4, 3.3a). These samples were also extremely homogeneous in size and cargo loading when analyzed by TEM, DLS, and MALS (Figures Appendix B.5, B.6, and Table 3.1). Additionally, the P22 capsids packaged with glycine spacer fusions could access the expanded morphology while retaining their packaged fluorescent cargo (Figure Appendix B.3).

GFP and mCherry are an established FRET pair where GFP is the donor and mCherry is the acceptor, with a characteristic FRET length of 55 Å (50% FRET efficiency). (*103*) Under favorable FRET conditions, mCherry is expected to exhibit fluorescence even when only GFP is exclusively excited. Upon selectively exciting GFP

Table 3.1: Summary of the physical characteristics of the P22 encapsulated systems. Samples were measured for homogeneity, size, and total particle molecular weight using high-performance liquid chromatography coupled to multi-angle light scattering (MALS), dynamic light scattering (DLS), and transmission electron microscopy (TEM). All samples were run in triplicate and reported values are averages of the three runs.

| Sample | Fusion Mw (Da) | $R_g$ (nm, MALS) | $R_H$ (nm, DLS) | Radius (nm, TEM) | Total Mw (MDa) | # cargo |
|---|---|---|---|---|---|---|
| P22 GFP PC | 45515 | 22.5 | 25±2 | 22.5 ±0.5 | 32.4 | 282 |
| P22 mCherry PC | 45994 | 22.5 | 24.5±2.5 | 22.5±0.5 | 30.5 | 238 |
| P22 Cat PC | 71582 | 22.7 | 30.3 ±2.5 | 23.5±2 | 27.5±0.16 | 111 |
| P22 Cat EX | | 25.0 | | | 26.9±0.12 | |
| P22 Cat_$Gly_{(6)}$ PC | 72167 | 22.3±0.07 | 34.6±3.2 | 22±2 | 28.4±0.01 | 123 |
| P22 Cat_$Gly_{(6)}$ EX | | 24.0±1.4 | | | 27.3±0.62 | |
| P22 Cat-$Gly_{(12)}$ PC | 72751 | 23.4±1.2 | 35.3±2.6 | 26.5±1 | 27.0±0.23 | 103 |
| P22 Cat-$Gly_{(12)}$ EX | | 24.4±1.4 | | | 25.7±0.94 | |
| P22 Cat-$Gly_{(18)}$ PC | 73336 | 24.9±0.12 | 34.8±4 | 30±2 | 26.4±0.74 | 93 |
| P22 Cat-$Gly_{(18)}$ EX | | 24.8±0.1 | | | 26.4±0.77 | |

$R_g$ = Radius of gyration, $R_H$ = Radius of hydration

(450 nm) mCherry fluorescence was observed for all GFP-mCherry catenated fusions in PC, EX and free states showing that, all GFP and mCherry catenated fusions are FRET active (Figure 3.3a). The standard technique for quantifying FRET efficiency is to compare the donor fluorescence in the molecule of interest (in presence of the acceptor) with donor fluorescence in a suitable control molecule (where the acceptor is absent). For our catenated molecules, where molecular environment is crucial to the study, this would demand constructing controls where mCherry is subtly altered such that only its chromophoric property is destroyed but correct folding and packaging remained

unchanged. Instead, as elaborated below, we follow a simple method of extracting relative FRET efficiencies in the catenated molecules.

The fluorescence intensity of GFP ($l_{ex}$: 450 nm; $l_{em}$: 505 nm) is given by:

$$I_{ex:450}^{em:505} \propto \frac{k_F^{GFP}}{k_F^{GFP} + k_{IC}^{GFP} + k_{ISC}^{GFP} + k_{FRET}} \qquad (1)$$

where $k_F^{GFP}$, $k_{ISC}^{GFP}$, $k_{IC}^{GFP}$ and $k_{FRET}$ are the rate constants for intrinsic fluorescence, intersystem crossing, internal conversion and FRET. Similarly, the efficiency of FRET, which is proportional to mCherry fluorescence ($l_{em}$: 610 nm), is given by:

$$I_{ex:450}^{em:610} \propto \frac{k_{FRET}}{k_F^{GFP} + k_{IC}^{GFP} + k_{ISC}^{GFP} + k_{FRET}} \qquad (2)$$

Combining Eq. (1) and (2) one obtains,

$$k_{FRET}^{*} \propto \left( I_{ex:450}^{em:610} / I_{ex:450}^{em:505} \right) \qquad (3)$$

Eq. 3 assumes that fluorescence yield of mCherry is identical in all GFP-mCherry constructs being compared. In reality, this may not be true and therefore the rate constant $k_{FRET}^{*}$ is only an apparent rate constant. Upon normalization with mCherry fluorescence yield ($\phi_{ex:585}^{mCherry}$), $k_{FRET}^{*}$ can be related to the real FRET rate constant $k_{FRET}$ as:

$$k_{FRET} \propto \left( k_{FRET}^{*} / \phi_{ex:585}^{mCherry} \right) \qquad (4)$$

The relative values of fluorescence intensity ratios ($k_{FRET}^{*}$) described in Eq. (3) are effectively mCherry fluorescence intensities ($I_{ex:450}^{em:610}$) in the constructs when the spectra are normalized at GFP fluorescence peak ($I_{ex:450}^{em:505}$). $I_{ex:450}^{em:505}$-normalized (to 100) fluorescence spectra ($\lambda_{ex}$ = 450 nm) of the constructs are shown in Figure 3.3a with the mCherry fluorescence spectra in the inset. Clearly, energy transfer from the donor (GFP)

to the acceptor (mCherry) is dependent on the molecular environment (i.e. capsid morphology and packing density) as the most energy transfer is observed in PC, followed by EX and free state. There is very little dependence on glycine spacer lengths in the three molecular environments.

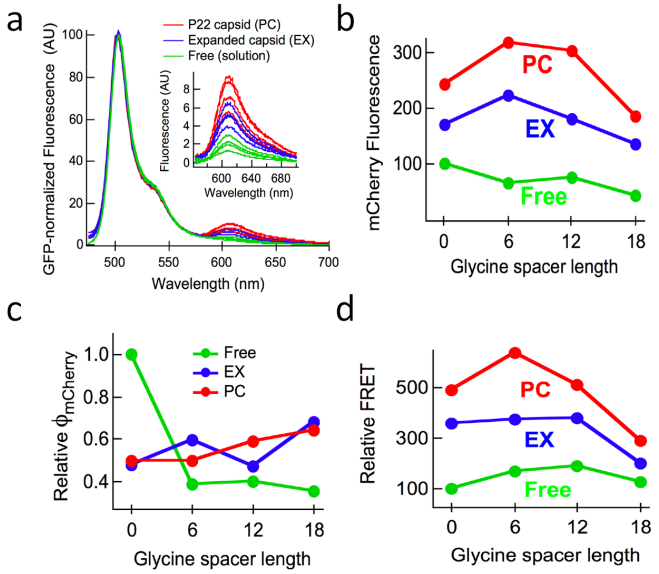



Figure 3.3: (a) Fluorescence spectra ($\lambda_{ex}$ = 450 nm) of GFP-$(Gly)_n$mCherry constructs in free, PC and EX states, normalized at GFP emission (505 nm). FRET-induced mCherry fluorescence spectra (610 nm) are shown in the inset after removal of GFP fluorescence tail by subtracting fluorescence spectra of appropriate controls lacking mCherry. (b) mCherry fluorescence intensities (from inset of panel a) in GFP-$(Gly)_n$mCherry constructs as a function of $(Gly)_n$ spacer. The fluorescence of GFP-$(Gly)_0$mCherry in the free state is normalized to 100. (c) Relative fluorescence yields of mCherry ($\lambda_{ex}$ = 585 nm; $\lambda_{em}$ = 610 nm) in GFP-$(Gly)_n$mCherry constructs as a function of $(Gly)_n$ spacer. The fluorescence yield of mCherry in GFP-$(Gly)_0$mCherry in free state is normalized to 1.0. (d) Relative FRET in GFP-$(Gly)_n$mCherry constructs as a function of $(Gly)_n$ spacer. Relative $k_{FRET}$ was estimated by dividing mCherry fluorescence of panel (b) by mCherry fluorescence yields of panel (c).

In order to obtain relative values of $k_{FRET}$ from mCherry fluorescence ($I_{ex:450}^{em:610}$) intensities, relative fluorescence yields of mCherry $\phi_{ex:585}^{mCherry}$ were estimated for the four

constructs in all three molecular environments (Figure 33.c). $\phi_{ex:585}^{mCherry}$ values are fairly independent of glycine spacer length in the PC and the EX environments, and similar to each other. In the free state, $\phi_{ex:585}^{mCherry}$ is also independent of glycine spacer length, and slightly less than that in PC and the EX states. However, for the catenated GFP-mCherry (without a glycine spacer) in the free state, $\phi_{ex:585}^{mCherry}$ is about two-fold higher. It is possible that flexible glycine linkers provide additional pathways for mCherry quenching, intrinsically unavailable in GFP-$(Gly)_0$-mCherry. Such fluorescence deactivation pathways may become available to mCherry in GFP-$(Gly)_0$-mCherry inside the capsid, but not in the less crowded solution state.

FRET measurements were sensitive to changes in the capsid morphology, from procapsid to expanded, and the distance dependence of FRET in the series of glycine linkers became more evident when measured in the encapsulated form. Relative values of $k_{FRET}$, obtained by normalizing $k_{FRET}^{*}$ (Figure 3.3b) with $\phi_{ex:585}^{mCherry}$ (mCherry emission, Figure 3.3c) are shown in Figure 3.3d. The $k_{FRET}$ is enhanced from the free in solution state by 5-fold in the most crowded PC and increased by 3-fold in the EX capsid. This demonstrates enforced communication through encapsulation. Interestingly, the PC shows the most linker length dependence on $k_{FRET}$. While the fluorescence emission from mCherry may have other relaxation pathways, the FRET signal from the encapsulated glycine spacer fusion samples follows a trend that correlates with the distance between donor and acceptor. In the EX and unencapsulated free protein forms of the glycine spacer samples, $k_{FRET}$ is relatively constant across linker lengths. The enhanced $k_{FRET}$ in the procapsid form, the most densely packaged form where the

scaffold proteins are interacting with the capsid wall, is most likely due to the enforced proximity of multiple fusions and thus suggests that both inter- and intra-molecular communication is occurring between the encapsulated fluorescent proteins. Additionally, dense encapsulation may force catenated pairs to bend in their flexible linker region and be closer than they would be free in solution. Once in the expanded form, the volume doubles and some of the enforced proximity is relieved and inter-molecular communication, as measured by $k_{FRET}$, is diminished. However, even in the EX form the $k_{FRET}$ between GFP and mCherry occurs at a level higher than seen in their free in solution counterparts, indicating some residual inter-molecular communication.

## Conclusion

We have successfully demonstrated the first example of co-encapsulation of two interacting proteins and enforced their molecular communication within the confined environment of a virus-like particle. Not only does the P22 scaffold protein-mCherry-GFP fusion fold correctly and maintain the fluorescence of the proteins, but the scaffold protein accurately templates the assembly of P22 capsids around the triple catenated protein fusion. The resulting capsids are homogeneous in size with dense packaging on the order of ~100 fusions per capsid, occupying roughly a quarter of the interior capsid volume.

The encapsulated FRET pairs experience intra- as well as inter-molecular FRET due to their enforced confinement inside the capsid. The fluorescence spectra of the fusion proteins show a clear difference between encapsulated and free in solution with

the encapsulated samples showing significantly enhanced FRET over their free in solution counterparts. Additionally, the FRET signal is much higher in the procapsid form and becomes more bulk-like in the expanded capsid. Interestingly, the most crowded PC state exhibited the most dependence of FRET on the linker length between the two chromophores, unlike the capsid-free proteins, which show little distance dependence.

Crowding and confinement have a profound effect on cellular processes and might make macromolecular interactions quite different from the same reaction occurring in dilute solution. Our P22 encapsulation system is a promising platform to study these effects as crowding is inherent to this programmed self-assembly and the confinement volume is definable, predictable, and modular with the capsids ability to be expanded without the loss of encapsulated cargo. These results highlight the possibility for chemical communication of co-encapsulated enzymes performing product-substrate exchange inside a viral capsid.

## Experimental Section

### P22 Catenated Fusion Assembler Plasmid Design

The background plasmid for all constructs was the pET11a-mCherry-SP141- P22 assembler plasmid (Novagen).(*102*) To create the catenated fusion assembler plasmid, standard PCR was used to amplify the gene for EGFP and add a 5' XmaI (Fwd primer: aa**cccggg**aaggggtgaaggaagtaatgaagatc) and a 3' SacI (Rvs primer:

aa**gagctc**cttgttcttctcaggcagcagaga) recognition sequences. This product was inserted upstream of the mCherry gene in the aforementioned assembler plasmid.

Next, to create the different linker lengths, a duplex DNA strand was ordered from Integrated DNA Technologies (IDT Inc., Coralville, IA) which coded for 6 glycines with SacI restriction sites on both ends. (aaa**gagctc**GGTGGCGGAGGGGGTGGA**gagctc**aaa). The SacI site, which is between the EGFP and mCherry genes, was cut and the digested 6xGly DNA ligated in. The ligated plasmid was transformed into XL-2Blue ultra competent cells (Agilent Technologies, Santa Clara, CA) and colonies were screened via single pass sequencing for the incorporation of one, two, or three of these spacers. The confirmed sequence fusion plasmids were subsequently transformed into BL21 *E. coli* (EMD Chemicals) for protein expression.

Unencapsulated Catenated Fusion Design

Using the catenated fusion assembler plasmids described above, the gene for the EGFP-6gly$_{(n)}$-mCherry-SP141 fusion was amplified using standard PCR adding a 5' SacI (Fwd primer aa**ggatcc**GCaaggggtgaaggaagtaatgaagatc) and a 3' XhoI (Rvs primer aa**ctcgag**ttatcggattccttaaagttttgcctttagcttgcgg) recognition site. These genes were ligated into a linearized pET-Duet-1 vector (Novagen) such that the protein product has a N-terminal 6xHis tag. Ligated plasmids were transformed into BL21 *E. coli* (EMD Chemicals) and colonies were screened via fluorescence of EGFP and mCherry using a bench-top fluorescent scope (Nikon Eclipse E600, mercury lamp power supply).

Protein Purification

Transformed BL21 (DE3) *E. coli* were grown to an $OD_{600nm}$ = 0.6 and then induced with 1mM IPTG. The culture was allowed to grow an additional 4 hours at 37°C with vigorous shaking. The bacteria were separated from the media by centrifugation at 3700 x rpm for 15 minutes. The cell pellets were resuspended in 50mM phosphate buffer (100mM NaCl, pH 7.2) and frozen at -20°C overnight. The next day, the cell pellets were thawed and incubated with DNAse (20 mg/ml), RNAse (30 mg/ml), and lysozyme (15 mg/ml) (Sigma-Aldrich) for 30 minutes at room temperature. Cells were lysed further by sonication (40% duty cycle, 5 minutes three times with a 2 minute cool down between) on ice. The cell debris was removed via centrifugation at 5000 x rpm for 45 minutes.

For the P22 encapsulated samples, clarified lysate (25 mL) was layered on top of a 35% (w/v) sucrose cushion (5 mL) and centrifuged at 48,000 rpm for 50 minutes in an ultra centrifuge (Sorvall WX Ultra). The resulting virus pellet was resuspended in 50mM phosphate buffer (25mM NaCl pH 7.0) and spun at 17,000 x g for 20 minutes to remove any particulates and lipid. The virus was then dialyzed against the aforementioned buffer overnight.

The unencapsulated free catenated fusion clarified lysate was loaded onto a HisTrap HP Column (Amersham) using an AKTA FPLC (GE Healthcare). Samples were washed with 20mM imidizole in 20mM phosphate, 500mM NaCl, pH 8. Protein was eluted with 500mM imidizole and dialyzed against 50mM phosphate, 25mM NaCl, pH 7.0 extensively.

Size Exclusion Chromatography

Protein samples were additionally purified via SEC. Encapsulated samples were centrifuged at 17,000 rpm for 10 minutes and the supernatant was loaded on to a preparative S-500 Sephadex (GE Healthcare) 80 mL column using an AKTA Pharmacia FPLC. The flow rate was 1 mL/minute using 50mM phosphate buffer (25mM NaCl pH 7.0) and the elution volume was 150 mL. Protein (280 nm) and the chromophore (490 nm for EGFP and 585 nm for mCherry) wavelengths were monitored simultaneously. (Figure Appendix B.1)

To achieve the expanded morphology, capsids were heated for 15 minutes at 65°C and were centrifuged at 17,000 rpm for 10 minutes. The supernatant was loaded on to a preparative Superose 6 (GE Healthcare) 30 mL column using an AKTA Pharmacia FPLC. The flow rate was 0.5 mL/minute using 50mM phosphate buffer (25mM NaCl pH 7.0) and the elution volume was 30 mL. Protein (280 nm) and the chromophore (490 nm for EGFP and 585 nm for mCherry) wavelengths were monitored simultaneously. (Figure Appendix B.2)

Free, unencapsulated, samples, after elution from the His-Trap column and dialysis, were centrifuged at 17,000 rpm for 5 minutes and the supernatant was loaded on to a preparative Superose 6 (GE Healthcare) 30 mL column using an AKTA Pharmacia FPLC. The flow rate was 0.5 mL/minute using 50mM phosphate buffer (25mM NaCl pH 7.0) and the elution volume was 30 mL. Protein (280 nm) and the chromophore (490 nm for EGFP and 585 nm for mCherry) wavelengths were monitored simultaneously. Fractions containing both the GFP and mCherry absorbance were collected and pooled.

Size Exclusion Chromatography -
Multi Angle Light Scattering

P22 capsid samples were separated by HPLC (Agilent 1200) size exclusion chromatography (WTC-0200S column Wyatt Technologies) at a flow rate of 0.7 mL/min. The mobile phase was 50mM phosphate buffer (100mM NaCl, 200ppm sodium azide, pH 7.2). Samples (25µL) were injected on the column and run for 30 minutes and elution was monitored using UV-vis detector (Agilent), a refractive index detector (Wyatt) and light scattering was detected with a Dawn 8 (Wyatt). Average Mw and Rg were determined from a fit of the data using a Zimm plot in the Astra software from Wyatt.

Dynamic Light Scattering

Dynamic light scattering was measured at 90º with a 661nm diode laser (Brookhaven 90Plus particle size analyzer) and the correlation function was fit using a non negatively constrained least squares analysis to yield a hydrodynamic radius ($R_H$) for the particles.

Agarose Gel

Samples were mixed with loading dye (40% glycerol, bromophenol blue) and loaded into the wells of a 0.8% agarose slab gel. Gels were run in 40mM tris 5mM acetate 1mM EDTA pH 8.2 at 65 V for three hours at room temperature. Unstained gels were illuminated with wide UV on a UVP MultiDoc-It Digital Imaging System and then subsequently stained with coomassie blue and destained using acetic acid and methanol. Stained gels were also imaged using a UVP MultiDoc-It Digital Imaging System. (Figure Appendix B.3)

Fluorescence Emission Scans

All samples were brought to the same EGFP 490nm absorbance on a UV spectrometer. 200µl of these samples were loaded into a quartz cuvette with a 1cm pathlength into a Cary Eclipse (Varian) Fluorescence Spectrometer. Samples were excited at 450nm, with a 5nm excitation slit width, and a fluorescence emission scan from 450nm to 700nm was taken. Fluorescence measurements were taken at 90° with a 5nm emission slit width. The PMT detector voltage was set to 600v and the scan rate set to 600nm/min.

Acknowledgements

This research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Science and Engineering DE-FG02-07ER46477. The authors would also like to thank Candace Goodman at Montana State University for many insightful conversations on fluorescence and Cindy Rodenburg at University of Alabama, Birmingham for the TEM images.

# CHAPTER FOUR

# ENCAPSULATION WITHIN THE P22 CAPSID GREATLY IMPROVES THE STABILITY OF A PHOSPHOTRIESTERASE

## Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Alison O'Neil

Contributions: Writing of manuscript, molecular design, protein purification and characterization

Co-Author: Peter E. Prevelige

Contributions: Provided the original P22 genes, manuscript editing, and coordinated the project.

Co-author: Trevor Douglas

Contributions: Obtained funding, coordinated the project, and assisted with the design of the experiments. Discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Authors: Alison O'Neil. Peter E. Prevelige, Trevor Douglas

Journal: Advanced Functional Materials

Status of Manuscript:
____ Prepared for submission to a peer-reviewed journal
__X_ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
___ Published in a peer-reviewed journal

Publisher: Wiley-VCH Verlag GmbH & Co

Date Submtted: February 15, 2013

# CHAPTER FOUR

# ENCAPSULATION WITHIN THE P22 CAPSID GREATLY IMPROVES THE STABILITY OF A PHOSPHOTRIESTERASE


## Abstract

Virus like particles, and other naturally occurring protein containers, are excellent building blocks for nanomaterials design and synthesis. Here we exploit a directed assembly and encapsulation approach to sequester multiple copies of a phosphotriesterase (PTE) within the capsid of bacteriophage P22. Phosphotriesterase, from *Brevundimonas diminuta*, is an intriguing enzyme as it is highly active against a wide range of harmful insecticides and nerve agents such as Soman and Sarin. However, difficulty in expressing large quantities of the active recombinant enzyme has limited efforts to scale-up its use. Additionally, as a mesophilic enzyme its low heat tolerance and its susceptibility to proteases makes it a less than ideal candidate as a practical bioremediation tool. Through encapsulation of the PTE in the P22 capsid, we demonstrate a greatly enhanced thermal tolerance of the enzyme, maintaining activity to 60°C. Additionally, the P22 capsid confers protection of the enzyme from proteases as well as stabilizing the enzyme to dessication. Thus, our engineered P22 encapsulation system greatly enhances the stability of the mesophilic phosphotriesterase and results in a robust and active nanoparticle reactor.

## Introduction

Viruses have evolved robust shells that function to transport and protect their nucleic acid cargo from degradation. These capsids, devoid of their nucleic acid cargo, have been exploited for their protective properties and precise, spatially controlled assembly, for nanomaterials design and synthesis. These assembled protein-based capsids have been utilized to entrap a wide range of metals, polymers, and proteins with beneficial effect.(*3, 7, 21, 27, 36, 41*) Through encapsulation in virus like particles (VLPs), the well-defined volumes of these architectures can be exploited for cargo occupation, restricted and defined particle growth, and unique interfacial chemistry at the high surface area of the capsid. The cargo, once sequestered inside the capsid, is forced into proximity with its neighbors in this crowded, confined space. This confinement and crowding may alter the cargo protein properties when compared to the unencapsulated cargo.(*7, 30, 38*)

Upon encapsulation of enzymes in protein containers we, and others, have observed varied effects to the enzyme kinetics. In some cases, there is little to no effect on the enzymes kinetics,(*36, 37*) in another the overall rate seemed to be enhanced upon encapsulation,(*38*) and in a third case the rate ($k_{cat}$) and the Km were decreased.(*7*) This promising new class of biomaterials is not yet predictable as no trends are evident and the kinetic outcomes seem to be enzyme specific.

Viral capsids are very robust and can be utilized to protect non-native cargo, analogous to their natural nucleic acid protection, in a biomimetic approach to developing

new active nanomaterials. Many virus like particles (VLPs) assemble into precise architectures, can be stable for long periods of time, and some can withstand the dramatic pH changes.(*104*) In the case of the bacteriophage P22 capsid, this VLP is soluble to high concentrations (30mg/mL), highly resistant to proteases, maintains the capsid assembly across a wide pH (pH 4-10) and temperature (-80 to 85°C) range, can be lyophilized and rehydrated, and can withstand high concentrations of organics. By exploiting this capsid as a nanocontainer for directed encapsulation of an enzymatic cargo, we hope to bestow these qualities on the enzyme and use this advantageously towards the engineering of novel active nanocatalysts.

We have investigated the encapsulation of the highly active phosphotriesterase (PTE) from *Brevundimonas diminuta,* which has been shown to break down organophosphates, including chemical warfare agents and commercial insecticides.(*105*) The homodimeric, mesophilic enzyme has been shown to contain a bi-nuclear zinc active site, which can be replaced with cobalt to generate an even more active enzyme.(*106, 107*) The high catalytic activity, and toxic substrate remediation, makes PTE an intriguing candidate for large-scale bioremediation. However, large-scale isolation of the enzyme for practical applications has been greatly hindered due to the difficulty in expressing active enzyme.(*108*) When over-expressed in *E. coli*, the recombinant PTE was found largely in an aggregated inactive form and active enzyme made up only ~0.1% of the total cell mass. (*108*) Additionally, as a mesophilic enzyme, PTE has low heat tolerance (loses activity above 45°C) when free in solution or immobilized on a resin. (*109*) Therefore, as promising as a bioremediation tool PTE maybe, the poor expression

yields and low heat tolerance makes it a less applicable to industrial use. These properties make PTE a prime candidate for protein expression rescue and thermal stability enhancement via incorporation into our synthetic bio-materials approach. Through the encapsulation of PTE within the P22 capsid, we expand the usefulness of this enzyme by increasing the active protein yield in *E. coli* and by greatly increasing the thermal tolerance of PTE. Our engineered system confers additional advantages by protecting the mesophilic enzymes from proteolytic degradation and lyophilization.

## Results and Discussion

### P22-PTE Characterization

The bacteriophage P22 requires the presence of a scaffolding protein (SP) and coat protein (CP) for heterologous assembly of T=7 capsids. As we have previously reported, genetic fusions of a cargo gene to a truncated SP allow the directed encapsulation of a gene product cargo, in high copy number, within the P22 capsid. The gene for phosphotriesterase (PTE) from *Brevundimonas diminuta* was cloned upstream of a truncated form of the P22 scaffold protein (SP) as an N-terminal fusion (PTE-SP). The PTE-SP fusion was able to template the assembly of P22 capsids *in vivo* while entrapping the PTE-SP inside (Figure 4.1a). By transmission electron microscopy, the P22 capsids have an interesting, slightly irregular, morphology (Figure 4.1b, see Figure Appendix D.1 for more examples). In our experience with the P22 encapsulation system, non-perfect capsids are obtained when the scaffold-cargo fusion expression is low relative to CP expression. With the PTE-SP fusion, *E. coli* growth conditions had to be optimized to

obtain consistent, single shelled, capsids, consistent with poor expression of the PTE enzyme itself as seen by others, (*108*) and relatively poor expression of the PTE-SP fusion. Even when growth conditions are optimized, we observed lower than expected cargo encapsulation numbers (~40 compared to ~300 for a similarly sized fusion, EGFP-SP141(*102*)) and more mis-formed capsids by size-exclusion chromatography (Figure Appendix D.2). This again suggests poor expression of the PTE-SP fusion, consistent with what has been reported. (*108*)

To further understand the role that the scaffold protein expression plays in the capsid assembly, a construct expressing only the P22 coat protein was made. An identical expression and was used and the resulting "coat protein only" materials were subjected to size-exclusion chromatography to assess the relative amount of properly assembled capsids. The SEC profile revealed a significant peak that eluted early and corresponds to a population that is mis-assembled into a wide range of structures. A slightly later eluting peak corresponds to well formed (T=7) particles. This is profile is similar to what was observed with the P22-PTE construct grown under non-optimized conditions. In both cases the ratio of the early mis-assembled/aggregate peak to the later properly assembled peak, was much higher than what is seen in constructs where the scaffold fusions are well expressed and present in high copy number during the assembly process (Figure Appendix D.3). With the poorly expressing (or no) scaffold protein constructs, the mis-assembled and assembled peaks are not perfectly resolved by SEC, indicating a heterogeneous material.

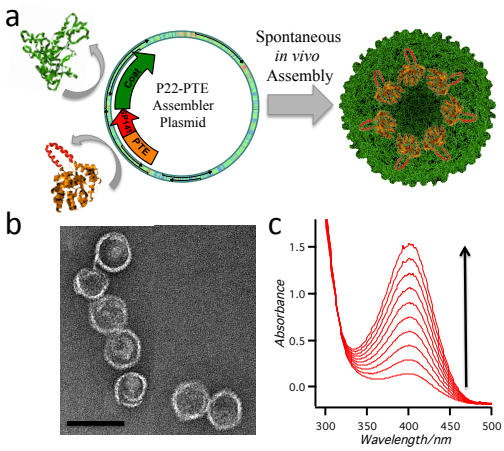



Figure 4.1: Through the P22 encapsulation system, phosphotriesterase (PTE) is specifically sequestered on the interior of the P22 capsid and maintains its catalytic activity. a) Cartoon representation illustrating the plasmid design on the left and the spontaneously assembled product on the right. b) TEM image of the P22-PTE capsids. Bar represents 100nm. c) As the P22-PTE cleaves the paraoxon substrate, a p-nitro-phenol group is released which can be monitored as an increase in the absorbance at 405nm.

When examining the TEM images of the poor expressing or no scaffold protein constructs, many multiple shelled, irregular capsids can be seen as well (Figure Appendix D.4). These results indicate that the scaffold protein fusion expression level is central to the proper templating of our P22 nanomaterial assembly.

However, despite these challenges, after the optimization of growth conditions and size-exclusion purification, the P22 encapsulation system still produces homogeneous material with good yields (~150 mg of purified material per liter of *E. coli* grown or ~5g cell paste). Analysis of the material by HPLC-SEC coupled to quasi-elastic and multi-angle light scattering detectors, the purified capsids were found to be homogeneous in size with a radius of hydration ($R_H$) of 30.57±0.12nm and a radius of gyration ($R_g$) of 27.87±0.06nm. The ratio of $R_g/R_H$ for empty P22 is 0.95, which is characteristic of an empty sphere, where as with P22-PTE this ratio is 0.91 suggesting

packaging of the cargo enzyme. From MALS, the average total molecular weight for the P22-PTE capsids was 21.8±0.56 MDa as compared to 19.65 MDa for empty P22, indicating packaging of 40±10 fusions per capsid (see Appendix D for calculation).

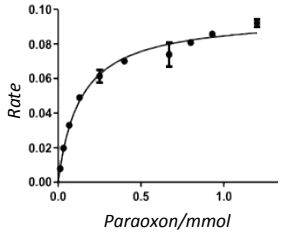



Figure 4.2: P22-PTE kinetics were measured using paraoxon as a substrate between 6.7µM -1.2mM. The plotted rates were fit to a Michaelis-Menten model ($R^2$ = 0.99) to determine a $V_{max}$ of 0.096±0.003 and a $K_m$ of 136±15µM.

The enzymatic activity of the encapsulated PTE was monitored at room temperature across 11 substrate concentrations ranging from 6.7 µmol-1.2 mmol. Paraoxon was us ed as the substrate because its enzymatic degradation product gives a characteristic absorbance at 405nm. For each substrate concentration, the rate of paraoxon hydrolysis was determined based on the change in absorbance at 405 nm over 5 minutes (Figure 4.1c). Fitting this data to a Michaelis-Menten model gave a $K_M$ of 136±15 µmol and a $k_{cat}$ of 10.7±0.5 $s^{-1}$ (Figure 4.2). The reported values of $K_M$ and $k_{cat}$, for the unencapsulated cobalt substituted enzyme are 100 µmol and 4870 ±90 $s^{-1}$ respectively. (*105*) Encapsulation of enzymes in protein containers has been observed to produce varied effects on the kinetics of the enzyme. In the case of P22-PTE, the enzymes Km is slightly increased and the turnover ($k_{cat}$) severely decreased. This observation may be linked to the poor expression yields of the enzyme fusion, and in turn perhaps poor folding of the enzyme. This would mean that not all of the enzymes encapsulated are

actually functional or are distorted such that they do not function as well as the PTE-fusion free in solution. However, like classical enzymology, characterizing the number of active proteins within a purified preparation is difficult. Considering the P22-PTE as a multi-component nanoreactor, with 40 active enzymes per capsid, the per capsid turnover rate is 400 $s^{-1}$.

P22-PTE Enzyme Stability Enhancement

The native and recombinant PTE free in solution is temperature sensitive and is reported to lose all activity after heating to 60ºC. However, the temperature sensitivity of the PTE enzyme is greatly improved upon encapsulation within the P22. P22-PTE and unencapsulated PTE were heated to a range of temperatures (up to 60ºC) for 20 minutes, cooled to room temperature, and then assayed for activity. The free PTE lost significant activity upon heating, maintaining only 45% activity when heated to 45°C, while the P22-PTE maintained more than 90% of its original activity (Figure 4.3). After heating to 55ºC, roughly 15% of the activity of the free PTE survived while the encapsulated PTE retained greater than 70% of activity. After heating to 60°C, the free enzyme had lost almost all activity while the P22-PTE was still more than 40% active. Through sequestration of the enzyme inside the P22 capsid, PTE has a much higher thermal tolerance. This is a promising enhancement of the properties for future materials use.

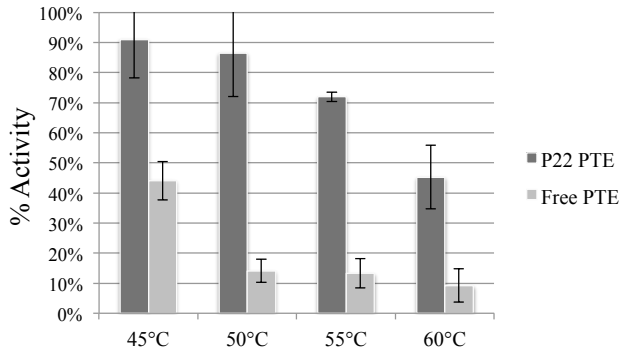



Figure 4.3: Encapsulation of phosphotriesterase within the P22 capsid greatly decreases its thermal sensitivity. Samples were heated at the indicated temperatures for 20 minutes, cooled to room temperature and assayed for activity. While the unencapsulated enzyme quickly loses activity with heat, the encapsulated enzyme maintains >70% activity to 55°C. At 60°C, the unencapsulated enzyme has lost almost all activity while the encapsulated maintains ~45% activity.

The P22 capsid confers an additional level of protection to the encapsulated PTE in that it shields the enzyme from access to proteolytic digestion. P22-PTE and the unencapsulated PTE were treated with excess trypsin for either one hour or over-night at room temperature and then assayed for activity. After one-hour trypsin treatment, the unencapsulated PTE lost 67% of its activity while the activity of the encapsulated PTE remained unchanged (Figure 4.4). Overnight (18 hours) trypsin treatment reduced the unencapsulated PTE activity to 19% of the untreated while the P22 encapsulated PTE maintained greater than 90% of its original activity (Figure 4.4). This enhancement of the nzyme stability makes the P22-PTE a significantly better candidate for remediation applications. If PTE were used to remediate insecticide-contaminated wastewater, the enzyme would likely encounter many proteases from an environmental sample. By protecting the enzyme inside the P22 capsid, which is porous enough for small substrate molecules, but restrictive of larger proteases, PTE is better suited for industrial use.

Additionally, when the P22-PTE and the free enzyme were lyophilized and then rehydrated, the encapsulated enzyme retained 91% activity while the free enzyme lost all activity (Figure 4.4). For long-term storage and use, it may be advantageous to be able to freeze dry active P22-PTE material. Through encapsulation within the P22 capsid, this storage method is now compatible with the PTE enzyme. It is also noteworthy that during the several months of this study, the P22-PTE stored in solution at 4°C did not lose any measureable activity. However the free-PTE was unstable in solution and slowly precipitated under the same conditions and exhibited inconsistent enzyme activity. The encapsulation of PTE in the P22 capsid thus confers many desirable materials properties to the enzyme making it a more advantageous candidate for incorporation into larger scale materials synthesis.

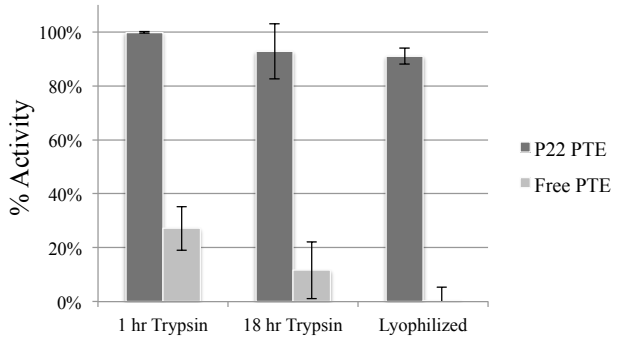



Figure 4.4: Encapsulation of the phosphotriesterase (PTE) within the P22 capsid confers heightened resistance to proteolytic degradation and loss of activity after rehydration. Treatment with trypsin for either one hour or overnight greatly decreased the activity of the unencapsulated PTE while the encapsulated PTE maintained activity with either length of treatment. Upon lyophilization and subsequent rehydration, the unencapsulated enzyme lost all activity while the encapsulated enzyme retained 91% of its activity.

Conclusions

## Conclusion

By genetically fusing the PTE to the truncated SP of P22 the system is programmed for the templated assembly and encapsulation of the PTE within the assembled P22 capsid. Thus, neither the scaffold protein directed assembly of P22 capsids, nor the PTE enzymatic activities are disrupted. The capsid acts as a protective shell for the PTE and upon encapsulation, the PTE exhibits greatly enhanced thermal stability, is resistant to proteolytic degradation, and can withstand the process of lyophilization and rehydration with little loss of activity. This approach takes advantage of the natural catalytic activity of the PTE enzyme against harmful nerve agents and insecticides and the stabilizing effects of the P22 encapsulation in the construction and design of a promising new material. The enhancement of the stability of the PTE enzyme now makes it a more promising material for industrial use.

## Experimental

### P22 Phosphotriesterase Fusion
#### Assembler Plasmid Design

The background plasmid for the design was the pET-Duet vector (Novagen). To create the fusion, first the gene that encodes amino acids 1-141 of P22 scaffold protein were amplify using standard PCR such that a 5' EcoRI (Fwd primer: aa**gaattc**cctggtgccgcgcggcagctgtcgcagcaatgccgt) and a 3' SacI (Rvs primer: aa**gagctc**ttatcggattcctttaagtt) recognition sequences were added. The digested product was inserted into MCS1 of the pET-Duet vector between the EcoRI and SacI sites.

Similarly, the gene for phosphotriesterase (PTE) was amplified off a vector graciously given from Prof. Frank Raushel (Texas A and M University) adding a 5' BamHI site (Fwd primer: aa**ggatcc**aggcgatcggatcaatacc) and a 3' EcoRI site (Rvs primer: aa**gaattc**aatgacgcccgcaaggtcgg). This gene was also cloned into MCS1 between the BamHI and EcoRI sites such that an N-terminal 6xHis tag was added to the PTE gene. Lastly, the gene for the P22 coat protein was amplified to add a 5' NdeI site (Fwd primer: aa**catatg**gctttgaacgaaggt) and a 3' XhoI site (Rvs primer: aa**ctcgag**ttacgcagtctgacc) and ligated into MCS2 of the same pET-Duet vector.

The ligated plasmid was transformed into XL-2Blue ultra competent cells (Agilent Technologies, Santa Clara, CA) and colonies were screened via single pass sequencing for the incorporation of the three genes. The confirmed sequence fusion plasmids were subsequently transformed into BL21 *E. coli* (EMD Chemicals) for protein expression.

Protein Purification

Transformed BL21 (DE3) *E. coli* were grown to an $OD_{600nm}$ = 0.6 at 30°C. The culture was then induced with IPTG (0.1mmol) and $CoCl_2$ (1mmol). The culture was allowed to grow an additional 18 hours at 30°C with vigorous shaking. The bacteria were separated from the media by centrifugation at 4500 x rpm for 10 minutes. The cell pellets were resuspended in HEPES (50 mmol) buffer pH 8.5 and frozen at -20°C overnight. The next day, the cell pellets were thawed and incubated with DNAse (20 mg/ml), RNAse (30 mg/ml), and lysozyme (15 mg/ml) (Sigma-Aldrich) for 30 minutes at room temperature. Cells were lysed further by sonication (40% duty cycle, 5 minutes

three times with a 2 minute cool down between) on ice. The cell debris was removed via centrifugation at 14500 x rpm for 25 minutes.

For the P22 encapsulated samples, clarified lysate (25 mL) was layered on top of a 35% (w/v) sucrose cushion (5 mL) and centrifuged at 48,000 rpm for 50 minutes in an ultra centrifuge (Sorvall WX Ultra). The resulting virus pellet was resuspended in 50mmol HEPES buffer pH 8.5 and spun at 17,000 x g for 20 minutes to remove any particulates and lipid. The virus was then dialyzed against the aforementioned buffer overnight.

The unencapsulated free phosphotriesterase fusion was purified out of the ultra centrifugation supernantant produced in the step above. The clarified supernatant was loaded onto a HisTrap HP Column (Amersham) using an AKTA FPLC (GE Healthcare). Samples were washed with imidazole (20 mmol) in phosphate (20 mmol, 500 mmol NaCl, pH 8). Protein was eluted with imidazole (250 mmol) and dialyzed against HEPES buffer (50 mmol, pH 8.5) extensively.

Size Exclusion Chromatography

Protein samples were additionally purified via SEC. Encapsulated samples were centrifuged at 17,000 rpm for 10 minutes and the supernatant was loaded on to a preparative S-500 Sephacryl (GE Healthcare) 80 mL column using an AKTA Pharmacia FPLC. The flow rate was 1 mL/minute using HEPES buffer (50 mmol, pH 8.5) and the total elution volume was 150 mL. Fractions were taken from the later half of the 60mL elution peak.

Size Exclusion Chromatography -
Multi Angle Light Scattering

P22 capsid samples were separated by HPLC (Agilent 1200) size exclusion chromatography (WTC-0200S column Wyatt Technologies) at a flow rate of 0.7 mL/min. The mobile phase was phosphate buffer (50 mmol, 100 mmol NaCl, 200 ppm sodium azide, pH 7.2). Samples (25μL) were injected on the column and run for 30 minutes and elution was monitored using UV-vis detector (Agilent), a refractive index detector (Wyatt), light scattering, and quasi-elastic light scattering was detected with a Dawn 8 (Wyatt). Average Mw, Rg, and $R_H$ were determined from a fit of the data using a Zimmol plot in the Astra software from Wyatt.

Enzymatic Rate Determination

To determine the $K_M$ and $V_{max}$ of the encapsulated PTE, the sample was diluted to a total protein concentration of 0.198 mg/mL and subjected to 11 concentrations of the substrate paraoxon from 6.7 μmol-1.2 mmol. Acetonitrile was kept at a constant 8% of the total reaction mixture and HEPES (50 mmol, pH 8.5) was used as the buffer. The absorbance of the paraoxon degradation product was monitored at 405 nm over five minutes in triplicate for each substrate concentration. The change in Abs 405 nm, was plotted against time and the slope of the linear portion of the graph was determined to be the initial rate of each sample. The triplicate-averaged rates were then plotted against the substrate concentrations and fit to a Michaelis-Menten kinetics model. From this fit, $V_{max}$ ($5.66x10^{-6}$) and $K_M$ (136μmol) were determined. Using the method previously described [1], the concentration of the PTE enzyme only in the encapsulated sample was

calculated to be $5.3x10^{-7}$M. By dividing the $V_{max}$ by the encapsulated enzyme concentration, the kcat was determined to be 10.67 $Ms^{-1}$.

Enzyme Sensitivity Assays

For the heat sensitivity assay, encapsulated and free PTE were heated to 45°C, 50°C, 55°C, or 60°C for twenty minutes and then allowed to cool to room temperature before assayed. For the proteolysis assay, encapsulated and free PTE were treated with excess trypsin (10 units trypsin per mg of protein) for either 1 or 18 hours at room temperature, in triplicate, before being assayed against an untreated sample. Additionally, encapsulated and free PTE were quickly frozen by immersion in liquid nitrogen and then subjected to lyophilization to complete dryness. The samples were then rehydrated in HEPES (50 mmol, pH 8.5) to the starting volume and assayed for activity. All treated and untreated samples, in triplicate, were assayed for activity with paraoxon (300 μmol) by measuring the change in A405nm over five minutes. Rates were calculated as described above and compared to the unheated sample average to report a percent activity.

Acknowledgements

The authors would like to thank Professor Frank M. Raushel (Texas A and M University) for supplying the phosphotriesterase gene. This research was supported by a grant from the National Science Foundation (BMAT-1104876). While conducting this research, Ms. O'Neil held an American Fellowship from the AAUW.

# CHAPTER FIVE

# TEMPLATED ASSEMBLY OF ORGANIC-INORGANIC MATERIALS USING THE CORE SHELL STRUCTURE OF THE P22 BACTERIOPHAGE

## Contributions of Authors and Co-Authors

Manuscript in Chapter 5

Author: Courtney Reichhardt

Contributions: Designed and carried out the experiments and wrote the manuscript.

Co-author: Masaki Uchida

Contributions: Assisted in mineralization and manuscript preparation.

Co-author: Alison O'Neil

Contributions: Assisted in experimental and molecular design

Co-author: Rui Li

Contributions: Took the TEM images

Co-author: Peter E. Prevelige, Jr

Contributions: Obtained funding and coordinated the project.

Co-author: Trevor Douglas

Contributions: Obtained funding, coordinated the project, and assisted with the design of the experiments. Discussed the results and edited the manuscript at all stages.

Manuscript Information Page

Authors: Courtney Reichhardt, Masaki Uchida, Alison O'Neil, Rui Li, Peter E. Prevelige, and Trevor Douglas

Journal: Chemical Communications

Status of the manuscript:
____ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-reviewed journal
____ Accepted by a peer-reviewed journal
_X__ Published in a peer-reviewed journal

Publisher: Royal Society of Chemistry
Issue in which manuscript appears: Volume 47, Issue 22, (2011), pp. 6326-6328

# CHAPTER 5

# TEMPLATED ASSEMBLY OF ORGANIC-INORGANIC MATERIALS USING THE CORE SHELL STRUCTURE OF THE P22 BACTERIOPHAGE

## Abstract

Biomimetic chemistry offers new approaches to supramolecular materials synthesis and assembly. We have demonstrated that an assembled viral protein cage, comprising an organic core–shell structure, can be used as a template for the size constrained synthesis of Fe2O3. Particle nucleation is directed by the inner scaffold protein layer, while the size constraints are determined by the outer capsid layer.

## Results and Discussion

Viruses are ubiquitous in nature and provide inspiration for the directed assembly, packaging, and targeting of macromolecular materials (*5*). The supramolecular structures of viral capsids and other protein cages, together with the range of available shapes and sizes, have made these biological assemblies useful and practical as platforms for biomimetic nano- materials synthesis. We have previously demonstrated that viral and other protein cage assemblies, ranging in size from 9 to 30 nm, can serve as organic containers for the encapsulation of inorganic materials (*110*). Utilizing electrostatic interactions between the viral capsid and the precursor ions, high local concentrations can be achieved at the protein interface, which can be used to initiate formation of mineral particles in a spatially selective manner on the interior cavity of the protein capsid (*110*-

*113*). By changing the electrostatics of the protein at this interface, through genetic design, we can direct the properties of the encapsulated mineral (*114, 115*). In the present work, we have engineered a multicomponent virus-based architecture that readily catalyzes the oxidative mineralization of Fe(II) precursors to form homogeneously sized iron oxide nanoparticles encapsulated within the cage-like architecture.

Utilizing viral capsids as synthetic containers for the encapsulation of nano minerals is a mimic of mineral formation in biological systems (biomineralization), specifically inspired by the iron oxide mineralization in ferritin (*114, 115*). Ferritin natively stores iron as ferrihydrite, rendering it inert and sequestering it from the outside environment. Ferritin directs ions to its interior via electrostatic gradients between its interior and exterior surfaces (*116*). We have previously shown that it is possible to alter the electrostatics of protein cage architectures to facilitate controlled mineral formation on the interior viral cavity using a mutant form of cowpea chlorotic mottle virus (CCMV) in which several basic residues were replaced with anionic glutamic acid residues, mimicking the electrostatic gradient found in ferritin (*110*).

In the present work, we expand on the use of the noninfectious, recombinantly expressed viral capsids from the bacteriophage P22 as supramolecular templates to create core–shell structures (*90, 117, 118*) comprising protein layers surrounding a nu- cleated and encapsulated nanoparticle of $Fe_2O_3$. P22 assembles from 420 identical copies of a coat protein with the assistance of a variable number of a 303 residue scaffold protein (100 to 300 copies are encapsulated per capsid). The C-terminal region (65 residues) of

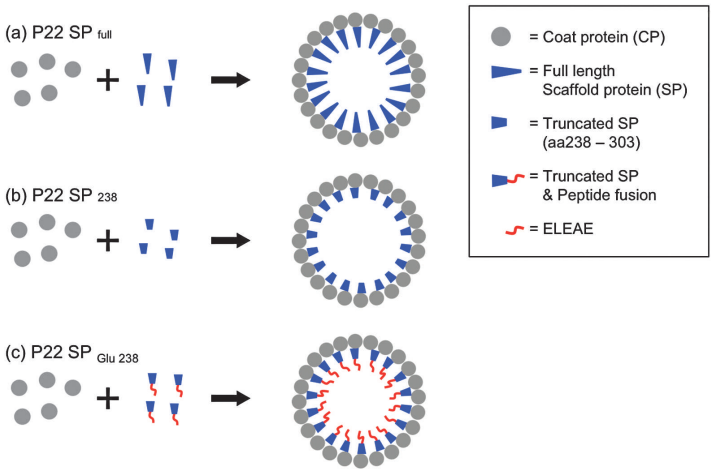



Figure 5.1: Schematic illustration of three P22 cage architectures that we have utilized in this study. (a) P22 coat protein (CP) is assembled with full-length scaffold protein (SP). (b) CP is assembled with a truncated SP which consist of amino acid position 238 to 303. (c) A peptide sequence ELEAE is genetically introduced to the C-terminus of the truncated SP.

the scaffold protein is required for the templated assembly of the capsid and the N-terminal region of the scaffold protein can be truncated without loss of assembly (*17, 79, 80, 119*). This protein-based core shell structure can be altered by removal of the scaffold protein, by treatment with guanidine-HCl, leaving a stable capsid structure, known as the empty shell (ES). Alternatively, N-terminal fusions to the SP can be engineered to direct unique functionality to the interior of the capsid (Figure 5.1). Our aim was to mimic the mineral sequestering abilities of the iron storage protein ferritin in a system with a much larger interior volume (approximately 216 times larger) by genetically fusing a polyanionic peptide to the SP, which is localized to the interior of the viral cavity. The SP is useful as a platform for genetic modification because modifications and deletions at the N-terminus of the protein still allow for assembly of the overall P22 cage, and also the SP can be removed, which allows a negative control

construct to be easily obtained. The larger interior cavity of P22 allowed us to synthesize larger size-constrained nanoparticles of iron oxide, which may be useful for many materials and bio-medical applications.

We explored the use of the viral capsid P22 as a platform for mineralization by examining the mineralization capabilities of four constructs: P22 with full length SP (SPfull), P22 with a truncated SP ($SP_{238}$), P22 with truncated SP having a fusion with an glutamic acid repeat ($SP_{Glu238}$), and P22–ES (where the SP or SP-fusion was removed from the capsid to leave an empty shell 'ES'). The fusion peptide in the P22 $SP_{Glu238}$ construct consists of a polyanionic repeat of glutamic acid (ELEAE) attached to the N-terminus of the SP. This peptide repeat was chosen in order to mimic the electrostatic gradient between the interior and exterior environments found in the native ferritin system, which facilitates the formation of $Fe_2O_3$ within the cage interior. These forms of the P22 capsid, with the SP variants, readily assembled into architecture that are essentially identical to heterologously expressed P22 as measured by TEM, size exclusion chromatography and dynamic light scattering (Appendix E).

These P22 constructs were suspended in buffer (100 mM MOPS, 37.5 mM $Na_2SO_4$, pH 6.8) at a concentration of 6 mg/mL and treated with $Fe^{2+}$ and allowed to air oxidize. $Fe^{2+}$ was added to the P22 solution at a rate of 166 Fe per capsid per minute to a final ratio of 20,000 Fe ions per capsid. In the reaction of P22-ES a large precipitate of $Fe_2O_3$ was formed and no P22 capsid material could be detected in solution after the reaction. In the other reactions, where P22 contained all three SP variants, rust colored solutions were formed, and either no precipitation or very little precipitation was

observed in each reaction. These homogeneous solutions suggest that the oxidative hydrolysis of $Fe^{2+}$ was confined to the interior cavity of the P22 capsid only when the SP

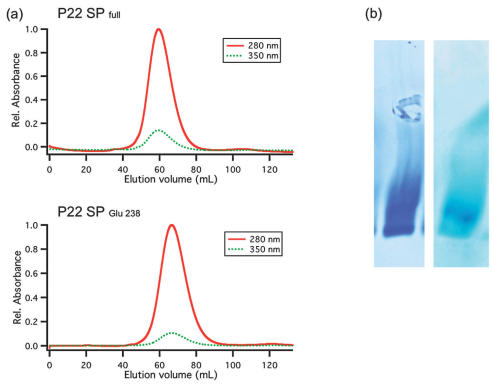



Figure 5.2: (a) SEC profiles of the P22 with full length SP (top) and P22with truncated SP-EA (bottom) after mineralization reaction. The profiles suggest that both P22 cages are still intact without measurable aggregation. (b) Agarose gel (0.8%) electrophoresis images of the mineralized P22 with truncated SP-EA. The gel was stained with Coomassie brilliant blue (left) and Prussian blue (right). The result shows co-migration of P22 protein and iron, suggesting encapsulationof iron oxide particle inside of P22 cage construct.

was present in the core–shell P22 structure. The encapsulated composites of iron oxide nanoparticles and P22 were analyzed after purification using size-exclusion chromatography (S-500 column). Elution from SEC was monitored for absorbance due to the protein (280 nm) and the iron oxide mineral (350 nm). The mineralized P22 samples had retention times identical to non-mineralized samples and showed co-elution of the protein and mineral components in all cases, which indicated that the mineral was associated with the viral capsid and the mineralization process did not significantly alter the overall structure of the viral capsid. This was confirmed by light scattering, which revealed that the size of the P22 capsid before (48 nm diam.) and after the mineralization

reaction (50 nm diam.) was unchanged. The samples were additionally analyzed using native gel electrophoresis on 0.8% agarose gels and stained for protein (coomassie) and iron (Prussian blue), which revealed colocalization of the mineral and protein (Figure 5.2). Two bands are visible on the Coomassie and Prussian blue stained gels possibly due o differences in the amount of SP or SP-fusion encapsulated within the P22 and lesser

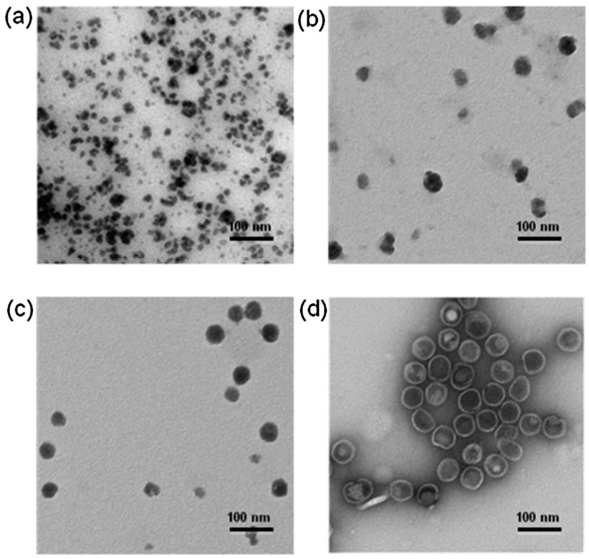



Figure 5.3: TEM images of various P22 cages, which are loaded with 20,000Fe/cage. (a) Unstained image of P22 with full length SP. (b) unstained image of P22 with truncated SP. (c) unstained image of P22 with truncated SP-EA peptide fusion. (d) uranyl acetate stained image of P22 with truncated SP-EA peptide fusion.

ability to direct $Fe_2O_3$ mineralization.

Mineralized samples were purified and subsequently imaged using transmission electron microscopy (TEM), and, addition- ally, samples that were negatively stained with uranyl acetate showed intact viral capsids (Figure 5.3). The iron oxide cores were observed in unstained TEM images and only subtle differences in the mineral cores were observed between the different SP constructs. The P22 $SP_{Glu238}$ consistently produced

more homogeneous particles with respect to shape and size than either the $SP_{238}$ or $SP_{full}$ samples (Figure 5.3), which is believed to be due to its anionic interior, which directs the controlled mineralization of $Fe_2O_3$. In the case of the mutant P22 $SP_{Glu238}$, which has an anionic interior, the iron oxide cores were spherical and had a monodisperse particle size distribution (41±5 nm diam.). TEM images of the other samples, P22 $SP_{full}$ (38±6 nm diam.) and P22 $SP_{238}$ (48±9 nm diam.) revealed irregularly shaped cores with a wide range of sizes. No electron diffraction could be detected on any of the mineralized samples, suggesting the formation of a highly disordered amorphous iron oxide.

We have demonstrated that the inner scaffold protein layer, in the protein core–shell architecture comprising the P22 procapsid, is competent to direct the nucleation of $Fe_2O_3$ via an oxidative hydrolysis reaction. Furthermore the outer, capsid, layer of the P22 assembly physically constrains the growth of the $Fe_2O_3$ particle limiting the particle size. Fusions of short polyanionic peptides to the scaffold protein increased the homogeneity of the $Fe_2O_3$ particles and removal of the scaffold protein destroyed the ability of the capsid to direct nucleation and growth. Synthetic conditions, aimed at controlling polymorph selection, are currently being explored using these core–shell capsid architectures. Synthetic conditions, aimed at con- trolling polymorph selection, are currently being explored using these core–shell capsid architectures.

# CHAPTER SIX

# ENCAPSULATION OF ACTIVE PEPTIDES IN BACTERIOPHAGE P22

## Introduction

It has been shown extensively that we can encapsulate well-ordered proteins through a P22 scaffold protein fusion strategy. However, many small peptides, that might not have well defined structures, have attractive activities for materials design. The previous chapter describes the encapsulation in P22 of a poly-glutamate peptide for sequestration of iron ions and their subsequent restricted mineralization (*87*). Peptides have been identified that bind to a variety of substrates. This includes gold, silver, platinum, palladium, zinc, iron oxide, cadmium, graphene, graphite, aluminum, silicon, etc. (*120*) Additionally, many synthetic peptides have been identified with immunogenic and therapeutic properties. (*121, 122*)

An interesting class of active peptides are the defensins and conotoxins. Defensins are cysteine rich peptides less than 100 amino acids in length, with antimicrobial properties (*123*). In animals, these peptides are responsible for the antimicrobial activity of granulocytes, the epithelium, and in the mucosal lining of the small intestine. They are particularly evident in humans and are ubiquitously found in inflamed and injured tissues. (*123*) Conotoxins are neurotoxic small peptides found in the venom of the marine cone snail (*124*). The small, less than 30 amino acids, peptides specifically target ion channels in their prey, which can lead to paralysis or death. (*124*)

The broad spectrum of available active peptides gives the peptide-scaffold fusion encapsulation strategy the possibility for many different functions. In this chapter, the encapsulation and activities of three additional peptide-scaffold protein fusions will be described.

## MVIIA Conotoxin Encapsulation*

*This work is in collaboration with Prof. Mandë Holford at Hunter College, City University of New York. As conotoxins are a select agent, protocols for preparation, use, and disposal of the resultant toxin protein were approved by the Montana State University Biosafety committee (IBC approval number 025-2011).

### Introduction

The overall goal of this project is to encapsulate the MVIIA-conotoxin (an approved analgesic) inside the P22 for delivery to the brain. Currently, there is no way to deliver the drug to the brain, besides direct injection, as it does not cross the blood brain barrier. The encapsulation of the drug inside P22 provides an exterior handle by which another functional peptide, TAT, can be linked to target the brain.

ω-MVIIA is a 25 amino acid long conotoxin purified out of the venom of *Conus magus* (*124*). This toxin specifically targets voltage-gated calcium-ion channels in the presynaptic endings. (*124*) The high specificity of this particular conotoxin has made it a promising target for treatment of severe chronic pain (*125*). MVIIA has been manufactured as a drug called Zinconotide and is used as an analgesic. As opposed to opiates, patients do not develop a tolerance to Zinconotide and thus it is well suited for

the treatment of chronic pain. (*125*) To be active, this peptide needs to make specific di-sulfide linkages between its cysteine residues. This aspect is of particular concern as the peptide, in our system, will be heterologously expressed in *E. coli*, and thus this important post-translational modification may not occur.

To traffic the P22 capsids across the blood brain barrier, the exterior of the capsid was decorated with a cell penetrating peptide. This peptide, known as TAT (sequence RRRQRRKKRGY), is a trans-activating transcriptional activator from HIV and has been shown to cause the uptake of nanoparticles in a broad spectrum of cell types (*126*). Additionally, the TAT peptide can be modified to incorporate a maleimide functionality on the N-terminus with no effect.

Results and Discussion

To construct the MVIIA- scaffold fusion, recursive PCR was used to add on the MVIIA sequence to the portion of the scaffold protein from amino acids 238-303. This version of scaffold protein was chosen as it has been shown that the smaller the scaffold, the more scaffold gets incorporated into the capsid (*80*). Thus, with a shorter scaffold, we could get a larger number of MVIIA into the capsid. An additional construct was made by inserting the gene for the red fluorescent protein, mCherry, in between the MVIIA and the scaffold protein. The rationale for the addition of the well-folded mCherry to the fusion was two fold. First, mCherry can act as a traceable marker for the fusion as it is fluorescent. Secondly, as shown in Chapter 2, the structured fluorescent proteins mCherry and GFP act as anchors, holding the scaffold and in this case another small peptide, inside the capsid. Maintaining the peptide inside the capsid is important

for its delivery and if the scaffold is released from the coat protein, it would leave the capsid and surely take the small, 25 amino acid, MVIIA peptide with it. Both constructs

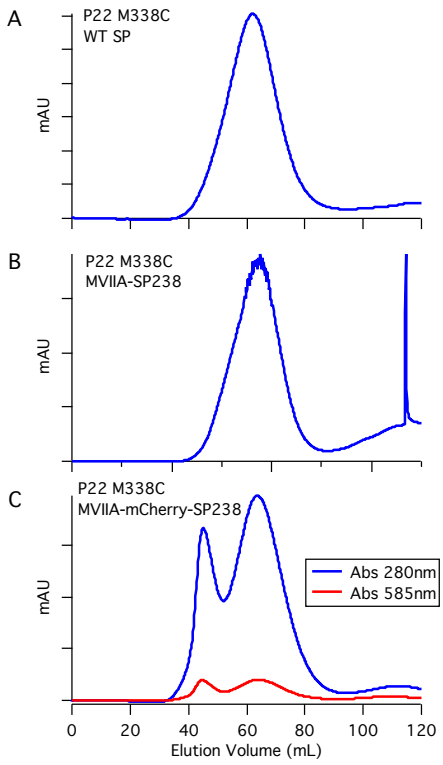



Figure 6.1: S500 size exclusion chromatography of the MVIIA constructs illustrating their proper assembly at ~60 mL elution volume. (A) the control capsid, P22 M338C with wild type scaffold protein, (B) P22 M338C with the MVIIA-SP238 scaffold fusion and (C) P22 M338C with the MVIIA-mCherry-SP238 fusion. In (C) the co-elution of the mCherry absorbance at 585nm is shown in red.

were made in the M338C coat protein background. This coat mutation exposes a cysteine to the outside of the capsid for downstream chemical modification. By size exclusion chromatography, both the MVIIA-SP238 and MVIIA-mCherry-SP238 result in the spontaneous *in vivo* assembly of capsids containing the fusion protein (Figure 6.1).

Additionally, the absorbance for mCherry (A585nm) co-elutes with the properly assembled capsids at ~60 mL.

The purified capsids (P22 M338C MVIIA-SP238) were successfully labeled with the maleimide-TAT peptide (TAT-FAM), which contains a fluorescent amino acid (fluorescein labeled lysine). By SDS-PAGE, there is a shift in the coat protein band to higher molecular weight indicating effective linking (Figure 6.2). Densitometry of the shifted band compared to unmodified coat shows that ~40% of the coat protein is being

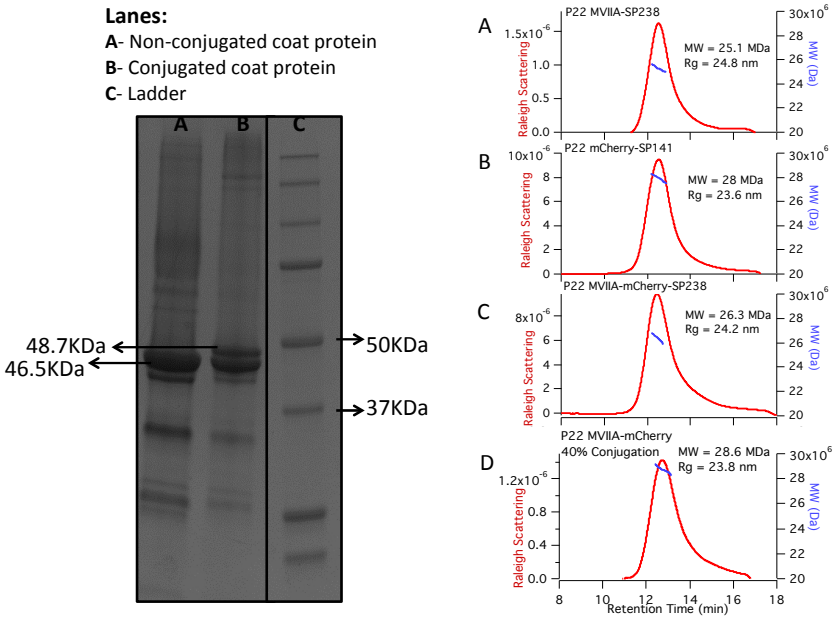



Figure 6.2: (Left) SDS-PAGE showing the conjugation of maleimidie-TAT-fluorescein to the M338C coat protein. M338C coat protein is 46.5 KDa while the modified coat is 48.7 KDa. The shift to the higher mass can be seen between lanes A and B. (Right) Stacked HPLC chromatograms showing that there is no change in retention time between the different scaffold constructs or upon conjugation. The MW is shown as a blue line across the elution peak. The average MW and radius of gyration for each elution peak is labeled on the graph.

conjugated with the TAT-fluorescein peptide. Additionally, by high-pressure liquid chromatography coupled to multi-angle light scattering (MALS), there was no change in capsid morphology (by the radius of gyration) and only slight variation in the average total particle molecular weight (Figure 6.2).

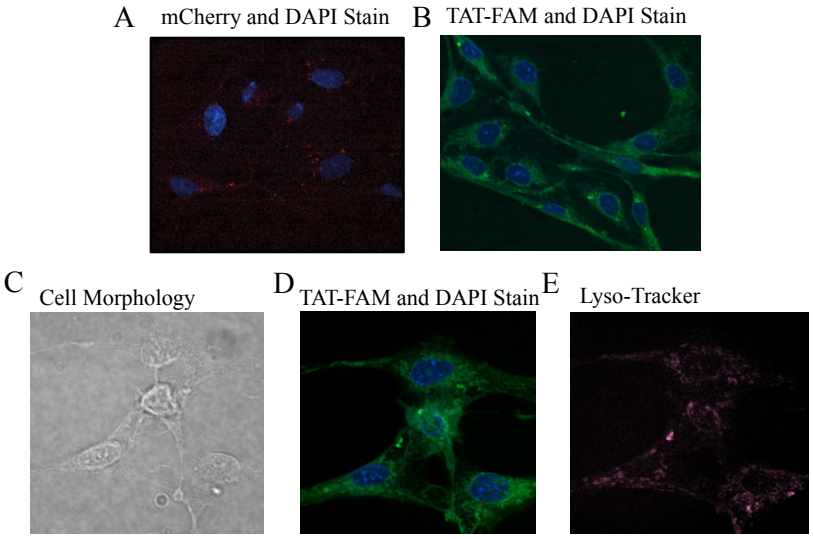



Figure 6.3: Fluorescent imaging results from incubation of the TAT-fluorescein (TAT-FAM) conjugated P22 MVIIA-mCherry-SP238 with Rat Brain Microvascular Endothelial Cells. (A) Red fluorescence is seen inside the cells from the mCherry, but not in the nuclei which are stained blue with DAPI. (B) The cells cytosol shows strong green fluorescence from the FAM conjugated on the exterior of the capsids. Again, the green capsids are not seen in the blue stained nuclei. (C) Showing the cell morphology, (D) the presence of the green capsids inside the cell, and (E) the Lyso-Tracker stain illuminating the lysosomes pink. By comparing D and E, it is concluded that the P22 capsids are in the lysosomes. Images courtesy of Dr. Prachi Anand, Hunter College.

Conjugated packaged capsids were successfully taken up by Rat Brain Microvascular Endothelial Cells (RBMVEC), a cell line that mimics the blood brain barrier. Incubation of the TAT-FAM P22 MVIIA-mCherry-SP238 conjugate, with rat brain cells for 30 minutes, followed by extensive washing and fluorescent imaging

showed the red fluorescence of the mCherry and the green fluorescence of the FAM inside the cytoplasm and not in the nuclei (Figure 6.3). The fact that the green signal from the exterior of the capsid and the red signal from the inside of the capsid are together in the cell, indicated that the capsids remain intact once engulfed. To see if the capsids were being engulfed and trafficked to the lysosome, a stain called Lyso-Tracker was employed that stains lysosomes pink. As seen in figure 6.3 D and E, the green fluorescence of the capsids correlates to the pink lysosomes. From this we can conclude that the capsids are entering the cells and are contained within the lysosome.

Conclusions and Future Directions

In this ongoing study, it has been shown that a conotoxin peptide, MVIIA, can be fused to the scaffold protein and packaged. Additionally, the exterior of the capsid can be chemically modified with a targeting peptide and the structure is not disrupted. The TAT targeting peptide allows for the conjugated and packaged capsids to be engulfed into the cells. Importantly, the capsids are not trafficked to the nucleus as the target of the MVIIA are the calcium ion channels, which can be accessed from the cytosol.

Now that it has been confirmed that the capsids can enter cells, and it appears that they remain intact upon entering the cell, and are not in the nucleus, experiments must be conducted to confirm that the MVIIA is active. One way to see if the disulfide bridges are present would be to analyze the mass of the fusion with and without a reductant present. However, mass spectrometry has remained elusive for studying the scaffold fusions. Solution NMR could also be used to elucidate the structure of the peptide. The activity could also be study *in vivo* be injecting it in mice and seeing if there is a numbing

effect. If the cargo proves to be active, a mode of release from the capsid also needs to be derived.

The preliminary results for the use of the P22 capsid as a vehicle for delivery of an analgesic across the blood brain barrier are very promising. However, the challenge of proving the toxins activity and devising a strategy for its controlled release are yet to be overcome.

Experimental

Using a recursive PCR strategy the gene for MVIIA was added on to the N-terminal end of the scaffold protein amino acids 238-303 (SP238). In this strategy, primers are designed such that the MVIIA gene is added in two PCR amplification steps. First, the C-terminal end of the MVIIA gene is appended to the N-terminal end of the SP238 using a forward primer containing the MVIIA gene section and a short section of the SP238 gene (Figure 6.4). The reverse primer anneals to the 3' end of the SP238 gene and adds a new BamHI restriction enzyme site. The product of this amplification is then used for the second amplification step to build the MVIIA gene. The second forward primer contains the N-terminal half of the MVIIA gene, adds a 5' XmaI restriction enzyme site and an overhang that anneals to the template (the product of first PCR) to join the MVIIA gene tog ether. The reverse primer stays the same. The product of this second PCR amplification yields the MVIIA-SP238 gene with a 5' XmaI site and a 3' BamHI site. The primers used for this are as follows with restriction enzyme sites in bold, the scaffold portion in italics, the overlapping "joining" section underlined, and the MVIIA start codon capitalized.

Fwd Primer 1: tatgattgctgtaccggtagctgccgcagcggcaagtgcccatgg*gtactcgactatccgaa*

Fwd Primer 2: aatcat**cccggg**ATGtgcaaaggcaaaggtgcgaagtgcagccgcctgatgtatgattgctgtacc

Rvs Primer: aaaa**gagctc***ttatcggattccttt*

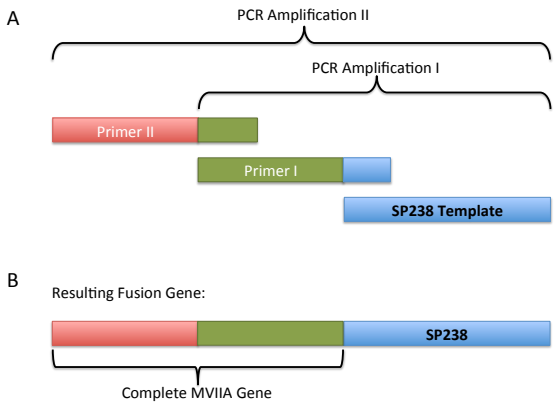



Figure 6.4: Diagram illustrating the recursive PCR strategy used to create the MVIIA-SP238 fusion gene. (A) The first round of PCR uses the SP238 gene as the template (blue) and a forward primer with the second half of the MVIIA gene (green) with an 3' end that anneals to the SP238 gene (blue). The result of the first PCR amplification is used as the template DNA for the second round of PCR. In the second round of PCR, the forward primer contains the first half (N-terminal) end of the MVIIA gene (red) and a region that overlaps with the first amplification (green). (B) The final fusion gene is a result of the sequence added from the two overlapping forward primers.

The resulting gene was digested with XmaI, SacI, and dpnI at 37°C for 3 hours and cleaned up using a Qiagen reaction cleanup kit. The digested insert was ligated into the similarly digested and gel purified pET11a vector. The ligation product was transformed into XL2-Blue ultracompetent cells (Agilent, CA) and the subsequent colonies screened for insert via DNA sequencing. Positive clones were transformed into BL21(DE3) cells and grown and purified as previously described herein.

To create the MVIIA-mCherry-SP238 construct, the same recursive PCR technique was used. The first forward primer was modified to anneal to the mCherry gene instead of the front half of the SP238. As a template, the pET11a mCherry Assembler plasmid was used (*102*). The second forward primer, which anneals to the MVIIA gene, and the reverse primer, which anneals to the end of scaffold protein, were kept the same. The same downstream steps were taken with this construct to ligate into an appropriate vector and transform into an expression cell line. The following primer was used with the mCherry annealing site in capitals and the MVIIA section underlined:
MVIIA-mCherry Fwd Primer:
tatgattgctgtaccggtagctgccgcagcggcaagtgcGTGAGCAAGGGCGAGGAGGAT

TAT-Peptide Synthesis and Conjugation. Prachi Anand preformed the peptide synthesis and conjugation at Hunter College, NY, NY. The TAT peptide was synthesized in house with the following sequence: RRRQRRKKRGY. During the synthesis, a fluorescently modified lysine was incorporated (underlined in the sequence). Additionally, the N-terminus was modified with a maleimide group. The fluorescent TAT peptide was conjugated to the M338C coat protein through maleimide-cysteine chemistry.

## Poly-Tyrosine Encapsulation

### Rationale

A commonly used X-ray contrast agent is iodine. An issue with this contrast agent is that when administered, it is filtered out of the body by the kidneys and thus the

iodine accumulates there. Accumulation of iodine in the kidneys can potentially cause renal failure and death. It has been shown that P22 capsids, when administered are filtered out by the liver. Therefore, the P22 capsid may be an effective vehicle for an iodine contrast agent as it would re-direct the contrast agent to the liver instead of the kidneys.

Tyrosine residues have been shown to be accessible sites for covalent iodination. The P22 scaffold fusion strategy has been shown effective in packaging poly-amino acid scaffold fusions (*87*). The goal of this research was to create and encapsulate a poly-tyrosine scaffold fusion providing many new sites for iodination inside the P22 capsid.

Preliminary Results

A poly-Tyrosine scaffold protein fusion was created adding 9 tyrosines, spaced by an alanine and glycine, to the N-terminus of scaffold protein 238-303 ($YAG_9$-SP238). The $YAG_9$-SP238 fusion was able to template P22 capsid formation *in vivo* and the S500 chromatogram shows little to no misassembled capsids (elution ~45 mL) (Figure 6.5A).

Preliminary results show that the poly-tyrosine construct holds more iodine than the wild-type capsid as evidenced by the greater yellow color post-iodination. Briefly, wild-type and $YAG_9$-SP238 were brought to the same total protein concentration, and treated with chloramine-T and NaI for five minutes at room temperature. The samples were then subjected to ultra centrifugation over a 35% w/v sucrose cushion. The resulting virus pellets were assayed for color (yellow for iodine). As seen in the pellets,

the $YAG_9$-SP238 looks much more yellow, indicating it was iodinated more than wild-type P22 which contains 8 natural tyrosines per coat protein (Figure 6.5).

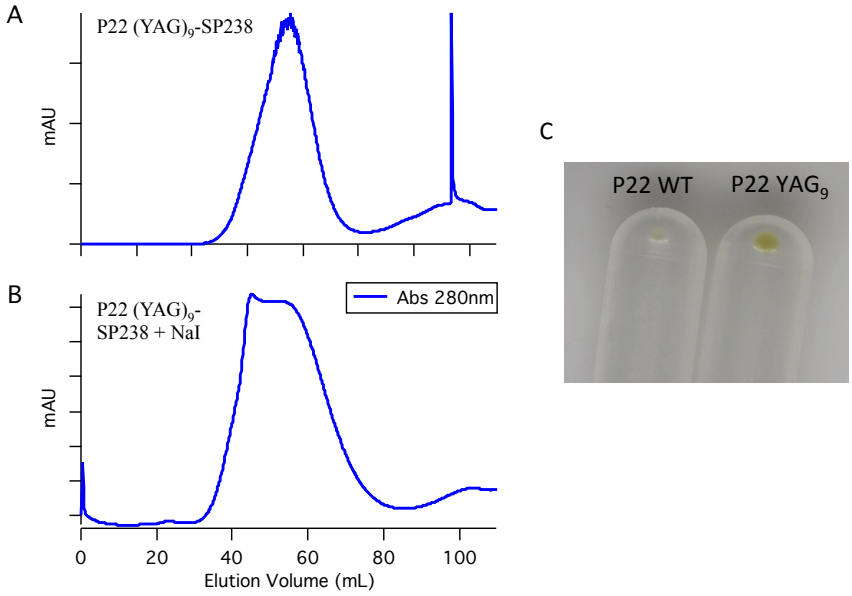



Figure 6.5: (A) S500 size exclusion chromatography of the P22 $YAG_9$-SP238 construct showing properly assembled capsids at the ~60 mL elution volume. (B) After iodination, the P22 $YAG_9$-SP238 retained its structure as there are particles eluting at 60 mL elution volume. (A) is a second size exclusion step, hence the lack of the mis-formed particle elution peak at ~45 mL. (B) Has been subjected to iodination before being SEC-purified. Therefore, the mis-formed peak at 45 mL could be a result of the assembly or from the iodination reaction. (C) Iodinated virus pellets showing the $YAG_9$-SP238 is much more yellow, indicating more iodine loading.

## Conclusions and Future Directions

A poly-tyrosine peptide can be successfully incorporated to the inside of the P22 capsid though fusion with the scaffold protein. Preliminary results indicate that the engineered tyrosine P22 can be loaded with more iodine than the wild-type capsids. The next step to characterize this material would be quantitatively determine the amount of

iodine in the P22 $YAG_9$-SP238 construct. However, iodine is difficult to quantitate. Another option is to forgo further characterization, put the iodinated capsids *in vivo* and see if they provide X-ray contrast.

Experimental

The tyrosine repetitive peptide was created as a tandem tyrosine-alanine-glycine (YAG) repeat to add more genetic diversity to the sequence such that recursive PCR could be used. The same method for recursive PCR was used as described in the previous section. The primers used are as follows with the new 5' restriction site (NcoI) in bold, the overlap section in italics, and the scaffold protein portion underlined.

YAG_SP238_Nco_FWD2:

5'-aaaaa**ccatgg**ggTATGCTGGGTACGCCGGCTATGCAGGATAC*GCGGGTTATG CTGGATA -3'*

YAG_SP238_Nco_FWD1:

*5'-GGGTTATGCTGGATA*CGCAGGGTATGCCGGCTATGCATACtgtactcgactat ccgaa-3'

Iodination was carried out in 50 mM phosphate pH 7.0 buffer. Fifty molar excess of NaI was added to the protein solution. Next,100 ug of Chloramine-T per 5 ug of P22, was added slowly while the protein/NaI mixture was stirring. The reaction was allowed to proceed, with mixing, for 10 minutes at room temperature. The reaction mixture was then subjected to dialysis overnight against 50 mM phosphate pH 7.0 buffer. The iodinated P22 was then concentrated using ultra centrifugation over a 35% w/v sucrose

cushion, at 45,000 rpm for 50 minutes. The resulting virus pellets were then imaged on a Multi-DocIt gel imager (UVP, CA).

## Poly-Lysine Encapsulation

### Rationale

Reichhardt and co-workers demonstrated the incorporation of a large amount of negative charge on the inside of the P22 capsid without disrupting the capsid morphology (*87*). The re-engineering of the interior surface of the P22 capsid with a poly-glutamate peptide was then exploiting for the loading of iron oxide. (*87*) In conjunction with this study, a poly-lysine scaffold fusion was created to incorporate a large amount of positive charge to the interior of the P22 capsid. The highly positive cavity could then be used for the loading of negatively charged molecules. Of interest was the loading of RNA for use in RNA silencing.

### Preliminary Results

The poly-lysine peptide was designed as a lysine-alanine 5-repeat and added to the N-terminus of scaffold protein 238-303 through PCR ($(KA)_5$-SP238). The construct was able to direct assembly of capsids and was packaged *in vivo* (Figure 6.6). To assess if the capsids were able to sequester nucleic acid, small DNA (a random primer) and RNA (donated from the Copié Lab) was added to the purified $(KA)_5$-SP238 capsids. After incubation, the capsids were purified away from the nucleic acid by ultra centrifugation. The resulting virus pellet was then assayed by UV-Vis, and the

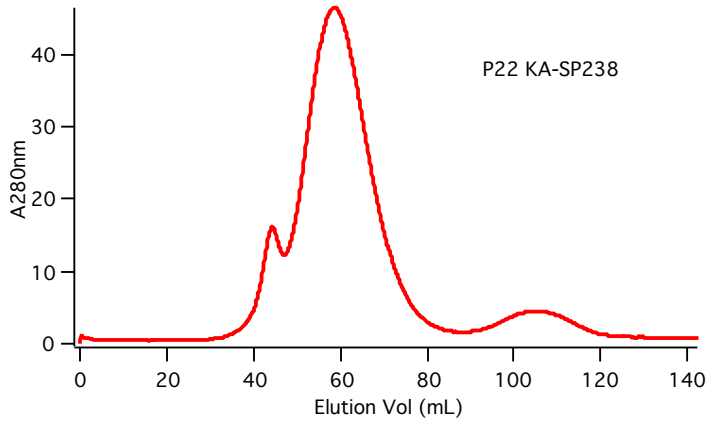



Figure 6.6: S500 chromatogoram of the P22-KA SP238 construct showing properly templated assembly. Assembled capsids elute at 60 mL elution volume.

A280nm:A260nm ratio was compared to the pre-incubation sample. Unfortunately, only small changes were observed in the ratios and all results were inconclusive.

## Conclusions and Future Directions

Here it is demonstrated that it is possible to incorporate a large number of positive residues on the interior of the P22 capsid. The addition of charge does not disrupt the capsid structure. Alas, experiments to load the highly-positive capsid with nucleic acid were inconclusive. For these experiments, a negatively charged dye molecule, which would be easily traceable, may be a better alternative to loading nucleic acid as proof of concept.

## Experimental

The lysine-alanine 5-repeat was added on to the N-terminus of the scaffold protein 238-303 through a one step PCR addition. The forward primer had a very long over-hang that contained the gene for the KA repeat. The following primers were used to

add the KA gene, add a 5' NdeI restriction enzyme site and a 3' SacI restriction enzyme site to the scaffold protein. Restriction enzyme sites are in bold, the scaffold portion is underlined, the KA repeat is in italics.

KA Fwd: 5'- *ggaatt**catatg**aaagcgaaggcaaaagctaaggccaaggcg*tgtactcgactatccga

SP Rvs: aa**gagctc**ttatcggattccttt

## Overall Conclusions and Future Directions

In this chapter it has been shown that it is possible to engineer active peptides as fusions to the P22 scaffold protein. These fusions are capable of templating the proper assembly of P22 capsids *in vivo.* Also described here is a robust method of recursive PCR to add on lengthy appendages by multiple PCR steps, dismissing the need for the purchase of expensive synthesized genes.

The main issue with these small active peptides is that they are hard to track. Our failure to be able to use mass spectrometry to detect the scaffold fusions severely limits our characterization. Mass spectrometry is needed to detect the post-translational modifications some peptides need for activity. MS would also confirm the presence of the peptides in the capsids as their small masses are difficult to image by SDS-PAGE.

However, despite these challenges, the plethora of active peptides described in the literature, and the seemingly general approach to incorporate peptides to the interior of the P22 capsid, leaves the door wide open to new material synthesis.

# CHAPTER SEVEN

# EXPLOITING THE NATURAL ASYMMETRY OF BACTERIOPHAGE P22 FOR NANOMATERIALS DESIGN

## Background

Protein containers are assembled from repeating monomer units, and as such, each symmetry related site of the container is chemically the same. If an addressable group is engineered into a protein container monomer, it will be displayed in an isotropic fashion over the entire container. This is a very useful feature for multivalent display of ligands on the exterior for enhanced response (*127*) and for creating many initiator sites for polymer growth or nucleation of an inorganic particle. However, isotropic display is not compatible with building larger assemblies as there is no directionality to the contacts and only large, uncontrolled aggregates can be made. Therefore, it is advantageous to create an asymmetric handle on protein container building blocks for increased utility and control over larger assemblages. Developing methods for creating asymmetric protein nanoparticles is non-trivial and only a few have been described. *(128-131, 91)* However, bacteriophage P22 has a natural asymmetric form that can be used for nanomaterial design.

### Masking Strategy

The *E. coli* glutamine synthetase enzyme has 24 solvent exposed histidines on two distinct sufaces that are utilized for a metal-dependent self-assembly pathway

resulting in tubes (*128*). Dabrowski and co-workers exploited this feature to create Janus-like monomers and hence affected the assembly outcomes. Monomers were attached to a Ni-NTA resin, where one histidine face binds, and the histidines on the exposed face were reacted with diethyl pyrocarbonate. When the Janus-monomers were mixed with the chelating metal, only dimers were formed and no native-like tubes. When the Janus-monomers were doped in with wild-type monomers, the Janus-monomers acted as capping agents; restricting the length of the assembled tubes. This masking technique yields facial distinction of the nanoparticles and creates a new class of assemblages. (*128*)

Using a similar masking technique, a *Listeria innocua* Dps (LiDps) Janus particle was made containing an anisotropic streptavidin (*129*). The authors used the genetically engineered KLFC LiDps, which has the KLFC peptide on each C-terminus of the monomers, to bind the cages to a thiol-reactive resin. This attachment masked one side of the LiDps by interacting with the terminal cysteines of the KLFC peptide. The exposed cysteine subunits were then selectively biotinylated. Next, a solution of streptavidin was added to the column, which selectively binds to the biotinylated subunits. The functionalized LiDps, is then released from the resin with DTT, exposing the released cysteines for further modification. The streptavidin, which has 4 biotin recognition sites, acts as an adapter protein through which any biotinylated molecule can be attached. The investigators show the utility of the system by fluorescently labeling the free cysteines via maleimide chemistry and attaching a biotinylated antibody to the streptavidin site. The antibody directed the cages to a microbial pathogen, while the fluorescent tag allowed for monitoring of the cages location. The dual-functionality of

this nanoparticle provided targeting and monitoring at the same time. (*129*) The described masking technique works well, but a whole surface is masked, not one monomer unit.

Dissociate, Mix, Reassemble Strategy

Taking a less directed approach, Kang and colleagues exploited the pH-dependent disassembly process of the LiDPS protein cage to create bi-functional chimeric nanoparticles (*91*). First, two populations of LiDps that were labeled on every monomer (as determined by mass spectrometry) were created and purified. The two populations had distinct functionalities; one with a tetrameric peptide KLFC (exposed to the exterior) and the other with an engineered cysteine S138C (on the interior). The terminal cysteine of the exposed KLFC peptide was labeled with maleimide-fluorescein while the interior cysteine was labeled with 5-iodoacetamide-1,10 phenanthroline (a strong metal chelator). (*91*) The two samples were then dissociated into monomers by lowering the pH to 2. The two types of monomers were then mixed in different ratios, and the pH was adjusted to 7, causing the spontaneous assembly of LiDps protein cages containing distinct ratios of the two monomers. Incorporation of the dual-functionalities were confirmed by the distinct absorbance spectrum of fluorescein and non-covalent mass spectrometry. (*91*)

Using CCMV, which also has a pH dependent disassembly, the same method was employed to make chimeric CCMV VLPs (*130*). Differentially labeled populations of CCMV capsids were dissociated into dimers at pH 7.5. The dimers were then mixed at differing ratios and the pH adjusted to 5 to spontaneously assemble the chimeric capsids. (*130*)

Through genetic engineering of the SV40 major coat protein (VP1) monomer, mono-functionalization of the SV40 VLP was accomplished (*131*). A mutant of VP1 was created to have an exposed histidine tag and cysteine. Mutant VP1 pentamers (the unit of assembly) were then mixed with excess wild-type VP1 pentamers to spontaneously form closed SV40 VLPs. The Janus-SV40 VLPs were then subjected to a Ni-NTA affinity column to pull out the VLPs containing a mutant VP1 pentamer (and hence a histidine tag) and allow the "all wild-type" VLPs to pass through. The resultant Janus-VLPs were then labeled with gold, which attached to the engineered cysteine in the mutant monomers. (*131*) This "label and mix" method is a simple and generic way to make chimeric protein nanoparticles, however it is restricted to protein cages that have a stimulus responsive disassembly-reassembly mechanism. This strategy also results in a mixed population with no spatial control.

The above examples for creating chimeric protein containers demonstrate synthetic strategies of adding multiple functionalities to one container to break the anisotropy. However, all of the current examples can be grouped in to two general methods: either masking and tagging or differentially label, dissociate, and mix to reassemble. These methods require reengineering of the protein container, extensive purification of components, and development of methods to combine individual components in a controlled manner to yield the intended outcome. In this chapter, a simple, generalized, *in vivo* assembly approach for incorporating an asymmetric unit to the P22 nanomaterial is discussed.

## Introduction

The bacteriophage P22 has a controlled, naturally occurring asymmetric unit: the portal complex. The portal is controllably added to the capsid assembly at only one of the 12 five-fold vertices. To incorporate this asymmetric unit into our bio-inspired nanomaterials approach, three strategies were investigated. First, a plasmid bearing the genes for portal, scaffold, and coat protein (gp 1,8,5) was re-engineered such that the exterior facing portion of the portal complex bears addressable cysteine residues. The second strategy takes a two-vector approach where the portal gene has been inserted on to a separate plasmid and hence can be co-transfected into any of the scaffold protein fusion vectors. This second approach yields P22 capsids packaged with the cargo fusion and bearing the asymmetric portal unit. An extension of the two-vector approach allows for differentially controlled expression of the portal-only plasmid and the scaffold-coat protein plasmid (scaffold and coat protein are on a single expression plasmid together) resulting in non-natural amino acid incorporation in to the portal only.

## Engineering New Cysteine Residues in the Portal Protein

Through the introduction of cysteine residues on proteins, free sulfhydryl groups are exposed for labeling reactions. The availability of the crystal structure of the P22 portal complex allowed for intelligent design of the location of the engineered new cysteine residues. The portal complex is composed of 12 monomers coming together to form a ring shaped core section that docks into the capsid with the N-terminal tails of each monomer wrapping together to form a barrel that extends into the interior of the

capsid (Figure 7.1a). Each monomer contains 4 naturally occurring cysteine residues (therefore 48 cysteines in the complex), however all of these cysteines are contained in the well-structured core of the protein (Figure 7.1b). The P22 capsid also contains 2

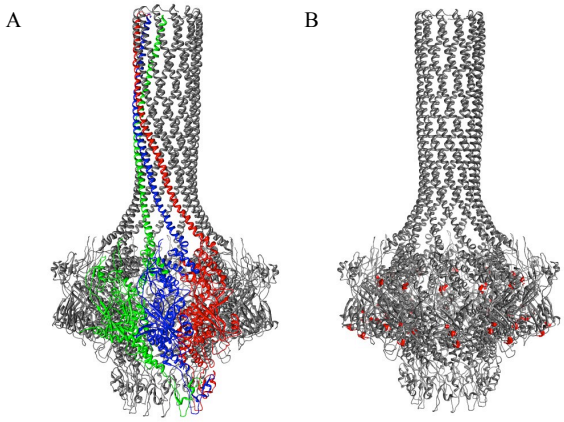



Figure 7.1: The crystal structure of the dodecameric portal complex. (A) The complex with three monomers colored to illustrate their position in the complex. (B) The portal complex has 4 natural cysteines per monomer, colored red, in the core of the structure. Images were generated using Chimera (UCSF) and the published coordinates (PDB 2LJ5).

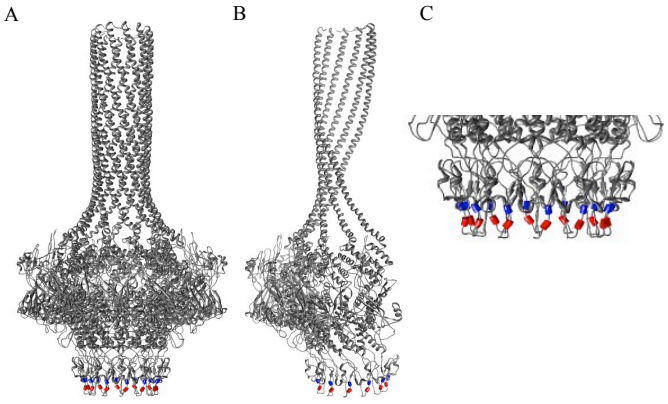



Figure 7.2: Engineering new cysteines into the exposed portal surface. (A) Location of the new cysteines. N377C in red T377C in blue. (B) The portal complex lacking 6 of its monomers to better show the new cysteins. (C) A close up view of the exposed region of the portal complex showing the location of the new cysteins. (PDB 2LJ5)

cysteines per subunit. However, the coat protein cysteines have been shown to be non-addressable via small molecules. By examining the structure of the portal protein, locations for single amino acid mutation were located that looked to be exposed to the exterior of the capsid and yet far enough apart in each monomer as to not be able to make di-sulfide linkage. From this analysis, asparagine 375 (N375C) and threonine 377 (T377C) were chosen for mutation (Figure 7.2).

Results and Discussion

The single amino acid changes from N375C and T377C were successfully made in the portal gene and did not affect the capsid assembly outcome. Using the gp 1,8,5 background vector, quick change PCR was used to mutate asparagine 375 or threonine 377 to a cysteine residue. Resulting clones were sequenced to confirm the base pair changes that result in the amino acid replacement. Upon expression of the mutated gp 1,8,5 plasmid in *E. coli*, the resultant capsids were indistinguishable from the wild-type gp 1,8,5 capsids via size exclusion chromatography, high pressure liquid chromatography coupled to multi-angle light scattering (MALS), and SDS-PAGE (Figure 7.3). Additionally, the ratio of coat protein to scaffold protein (415:12) on the SDS-PAGE gels looked similar to wild-type indicating the resulting capsids had the same amount of portal present in the capsids as the unmodified version (Figure 7.3).

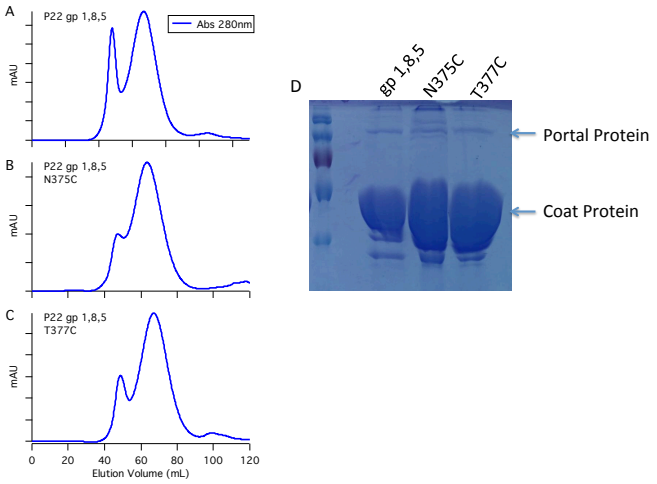



Figure 7.3: Characterization of the resulting N375C and T377C P22 capsids. (A-C) Size exclusion chromatograms of the control capsid, gp 1,8,5, and the new cysteine containing portal capsids. (D) SDS-PAGE of N375C and T377C compared to the wild-type control, gp 1,8,5, illustrating coat protein and portal protein in similar ratios.

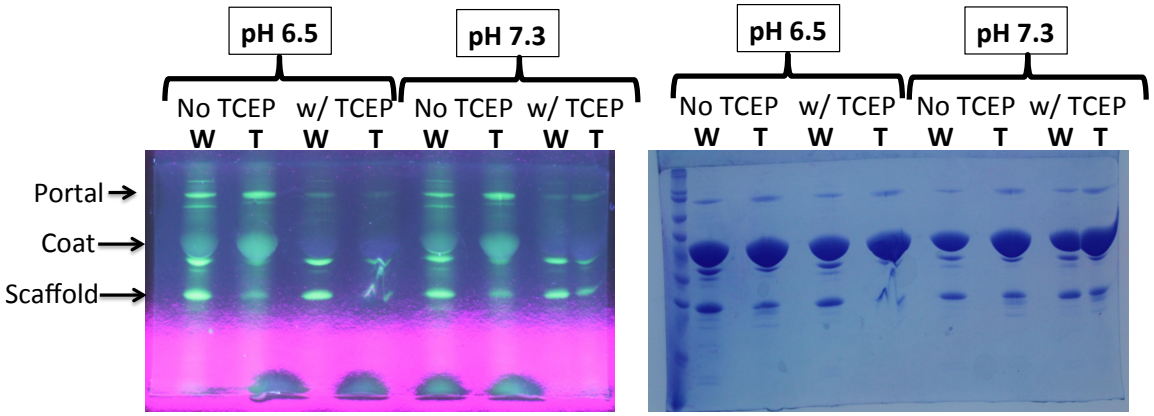



Figure 7.4: Fluorescein-maleimide labeling of T377C portal. On the left is the UV visualized SDS-PAGE and on the right is the coomassie stained gel. Four reaction conditions were explored, either pH 6.5 or pH 7.3 and with or without TCEP. The T337C (T) construct was compared to the wild-type construct (W) and the portal labeling efficiencies appear to be the same in the two samples.

To confirm the presence of the new cysteine functionality, the T377C capsids were labeled with fluorescein-maleimide. Maleimide groups react with free sulfhydryls

to form a covalent linkage while fluorescein is a green/yellow fluorescent molecule that is easily traceable due to its characteristic emission. Labeling reactions were carried out at two pH's, pH 6.5 and 7.5, and with or without TCEP (a disulfide reducing agent). The T377C mutant was compared to the wild-type gp1,8,5 capsids to assess for labeling efficiency. Unfortunately, these preliminary labeling reactions were inconclusive because the wild-type portal protein, which contains 48 natural cysteins compared to the mutants' 60 cysteins, labeled equally well as the mutants (Figure 7.4). Additioanlly, the fluorescein-maleimide appears to be labeling non-discriminantly as the coat protein, scaffold protein, and contaminants are all labeling. Both pH's worked equally as well, while TCEP was found to inhibit the labeling reaction. This means that the naturally occurring 4 cysteines in each monomer (one to all 4) are addressable by small molecule labeling. Because these naturally occurring cysteines seem to be a part of the structured core of the portal complex, it was hypothesized that a larger, bulkier labeling reagent could perhaps distinguish between the buried natural cysteines and the exposed, newly engineered cysteine.

To this end, an active protein was chosen to label the new cysteines. Bovine catalase is a large tetramer that catalyzes the disproportionation of hydrogen peroxide to oxygen and water. The rationale for linking catalase to the asymmetric portal vertex was two fold. First, it was large and hence should not be able to access the more buried, naturally occurring cysteines in the portal complex. Second, by attaching catalase to the asymmetric unit of the P22 capsid, we could create a "nano-rocket" and use hydrogen peroxide as a fuel to propel the capsid in a linear direction. The catalase would produce a

chemical gradient, acting as a nano-motor, to move the capsid in a non-Brownian motion trajectory.

Catalase was successfully linked to the portal protein via a bi-functional PEG linker (Figure 7.5a). On one end of the linker is an NHS-ester functionality that will label

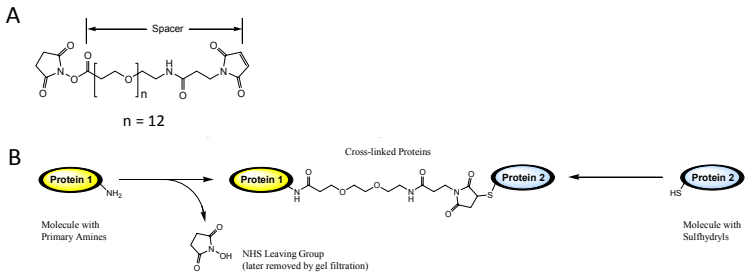



Figure 7.5: Portal-Catalase linking strategy. (A) The NHS-$PEG_{12}$-Maleimide linker. (B) Generalized scheme of the linked procedure where the lysines of catalase are first labeled (protein 1) and then this complex is added to the portal protein (protein 2). Image adapted from the Thermo-Fisher linking protocol.

free lysine residues. Spaced by 12 PEG units on the other end is a maleimide group that will label cysteines (labeling scheme Figure 7.5b). First, the linker was reacted with the catalase, labeling the exposed lysines on the enzyme. The reaction product was purified and subsequently reacted with the N375C and T377C mutants. This reaction mixture was purified via size exclusion chromatography to assure the capsid structure had not been disrupted (Figure 7.6). The first indication of successful linking was elucidated when hydrogen peroxide was added to the purified linked sample. Bubbles, presumably oxygen, indicating the catalase was present and active were seen in the linked sample but not the unreacted capsids (Figure 7.6c). From the "bubble" experiment, it was noticed that the T377C had a great deal more bubbles, thus indicating better labeling with the catalase.

To further analyze the linked samples, SDS-PAGE and anti-portal and anti-coat protein western blots were carried out. The SDS-PAGE gels showed the appearance of a higher molecular weight species in the linked samples versus the unreacted controls (Figure 7.7). This higher molecular weight species reacted with the anti-portal polyclonal

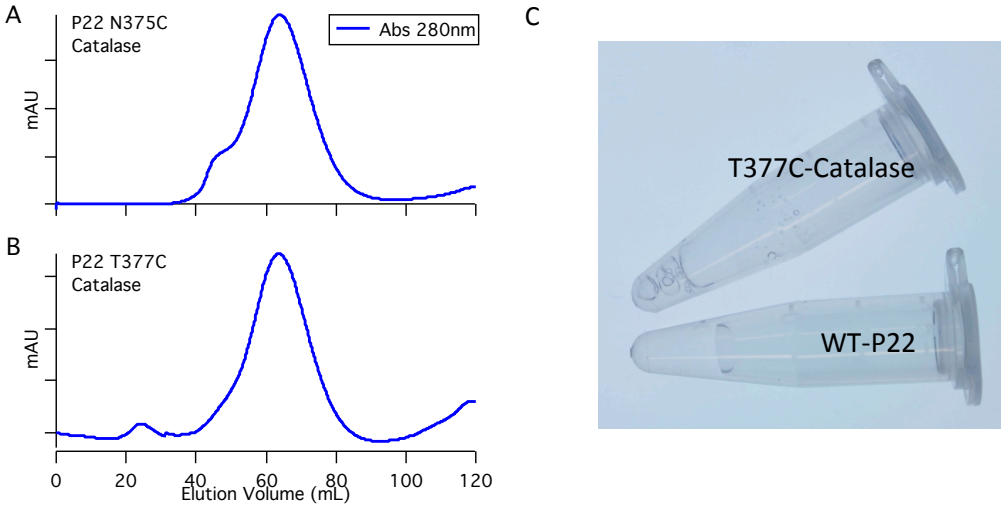



Figure 7.6: Characterization of the catalase linked samples. SEC of the catalase linked samples to either N375C (A) or T337C (B) showing the retention of the capsid assembly, as the complexes elute at the proper, ~60 mL elution volume. (C) When $H_2O_2$ is added to the T337C linked sample, bubbles appear, indicating the presence of the active catalase.

where as it did not react with the anti-coat protein polyclonal antibody. This indicates that the higher molecular weight species is composed of, in part, portal protein and linked catalase. The fact that the anti-coat protein polyclonal antibody did not react with the higher molecular weight species indicates that linking is specific to the portal protein.

To assess if the linked capsids were capable of motion upon the introduction of hydrogen peroxide fuel, samples were sent for testing at NanoSight Technologies (www.nanosight.com). The NanoSight instrument uses laser scattering to visualize nanoparticles and track them through 3 dimensional space. The nanoparticle tracking

analysis results in movies of nano-sized particles as they move due to Brownian motion. If the linked particles were propelled by the catalase reaction, they would not fit to a Brownian motion profile. Control tests showed that the P22 capsids could be seen by this technique. Unfortunately, when hydrogen peroxide was added to the samples, large entities appeared that drowned out any other signals. These large particles could be a contaminant from the hydrogen peroxide, bubbles of oxygen from the catalase reaction, or perhaps aggregates of protein, as hydrogen peroxide is a strong oxidant.

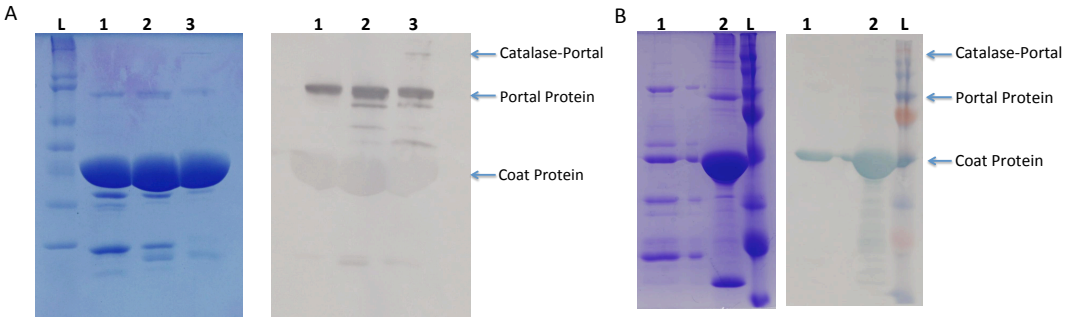



Figure 7.7: SDS-PAGE and western blots of the catalase-linked samples. (A) SDS-PAGE and anti-portal western of 1) wild-type unmodified gp 1,8,5, 2) T377C unmodified, 3) T377C-catalase linked. The appearance of a higher molecular weight band that cross-reacts with the anti-portal polyclonal, is hypothesized to be the catalase linked species. (B) SDS-PAGE and anti-coat western of 1) gp 1,8,5 positive control, and 2) the T377C linked sample. The higher molecular weight species does not cross react with the anti-coat polyclonal and as such, does not contain coat protein.

## Conclusions

In conclusion, it is possible to engineer new cysteine residues to the exposed surface of the portal protein while maintaining capsid assembly. Additionally, the naturally occurring cysteines are addressable with small molecule labeling reagents. As the wild-type cysteines label well with fluorescein-maleimide, it is difficult to determine if the new cysteine is also labeled. Unfortunately, it has not proved possible to see the

portal monomers by mass spectrometry, so the exact amount of labeling cannot be determined. The newly engineered cysteines can be selectively labeled by using larger reagents, such as enzymes. The new cysteines must be exposed on the exterior of the capsid because they are addressable to the large protein, catalase. By using simple maleimide chemistry, it should be possible to specifically label the asymmetric unit, as the wild-type coat protein does not label with maleimides.

Experimental

Molecular Biology. To create the cysteine mutants, quick-change PCR was employed. This PCR method amplifies the entire template vector with long, overlapping primers that match the gene of interest. The primers are designed such that flanking base pairs anneal before and after the codon that will be changed, with the new codon (that does not anneal to the vector) in the middle. The gp1,8,5 plasmid (from the Prevelige lab, University of Alabama, Birmingham) was used as the template vector. To make the N375C mutation the following primers were used (new codon in bold):

Fwd: 5'-ATACTACCTGCTC**TGC**CGCACTGACGAAAA-3'

Rvs: 5'-TTTTCGTCAGTGCG**GCA**GAGCAGGTAGTAT-3'

To make the T377C mutation, the following primers were used (new codon in bold):

Fwd: 5'-TGCTCAATCGC**TGC**GACGAAAATA-3'

Rvs: 5'-TATTTTCGTC**GCA**GCGATTGAGCA-3'

The reaction mixture contained 1 µl PfuUltra polymerase (Agilent Technologies, CA), 5 µl PfuUltra 10x buffer, 1 µl dNTP mix (25 mM each dNTP, New England Biolabs, MA),

50 ng of template DNA, 1 µl each of forward and reverse primer (at 100 nM), and ultra-pure water to a final volume of 50 µl. The reaction was placed in a thermo-cycler and heated to 95°C for 5 minutes, and then cycled 30 times at 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 5 minutes, with a final extension step at 72°C for 5 minutes. To each reaction 2 µl of DpnI was added and incubated overnight at 37°C (this step degrades the template DNA). The digested reaction was then cleaned and concentrated using a Qiagen PCR clean-up kit eluting the final product in 10 µl of provided elution buffer. Next, 3 µl of the resulting plasmid DNA was transformed in to XL-2 Blue ultra-competent cells (Agilent, CA). Colonies were then screened via single pass DNA sequencing for the N275C or N377C mutation. Positive clones were subsequently transformed in to BL21(DE3) *E. coli* for expression.

Protein Production and Purification. To express the mutants, 10 mL of an overnight culture was used in inoculate 1 L of LB broth. The cultures were incubated at 37°C were rigorous shaking until the optical density at 600 nm was approximately 0.6. The cultures were then induced with 1 mM IPTG, and allowed to grow for 4 additional hours. The culture was then centrifuged at 4,500 rpm for 15 minutes to pellet the cells. The cell pellets were resuspended in 50 mM phosphate, 100 mM NaCl, pH 7.6 lysis buffer and frozen at -20°C. The lysate was then thawed and incubated with DNAse (20 mg/ml), RNAse (30 mg/ml), and lysozyme (15 mg/ml) (Sigma-Aldrich) for 30 minutes at room temperature. Cells were lysed further by sonication (40% duty cycle, 5 minutes three times with a 2 minute cool down between) on ice. The cell debris was removed via centrifugation at 14,500 rpm for 30 minutes.

Clarified lysate (25 mL) was then layered on top of a 35% (w/v) sucrose cushion (5 mL) and centrifuged at 48,000 rpm for 50 minutes in an ultra centrifuge (Sorvall WX Ultra). The resulting P22 capsid pellet was resuspended in 50mM phosphate, 25mM NaCl, pH 7.0 buffer and spun at 17,000 x g for 20 minutes to remove any particulates and lipid. The resulting capsid containing supernatant was further purified by size-exclusion chromatography over a preparative S-500 Sephadex (GE Healthcare) 80 mL column using an AKTA Pharmacia FPLC. The flow rate was 1 mL/minute using 50mM phosphate buffer (25mM NaCl pH 7.0) and the elution volume was 150 mL. Mis-assembled capsids elute at ~40 mL while the properly assembled capsids elute at ~60 mL.

Fluorescein-Maleimide Labeling. SEC-purified P22 samples were brought to an $A_{280nm}$ of 3 in 50 mM phosphate 25 mM NaCl at either pH 6.5 or 7.5, and with or without 10 mM TCEP. The samples were then labeled with 50 molar excess of fluorescein-5-maleimide (F5M, Sigma-Aldrich). The fluorescein stock solution was 25 mg/mL in DMSO and was added drop-wise to a vigorously shaking protein sample. The reaction mixtures were wrapped in foil and incubated at room temperature with mixing for 3 hours. Reacted samples were then subjected to 100k MWCO spin filters to remove unreacted F5M and TCEP. The samples were spun and washed with 50 mM phosphate 25 mM NaCl pH 7, on the filter 3 times.

SDS-PAGE. 15 µl of each sample was mixed with 5 µl of DTT-containing loading dye and boiled for 5 minutes. The samples were subjected to a quick centrifugation to gather condensate from the tube. 15 µl of the boiled samples were

loaded on to a poured 12% acrylamide gel and ran at 40 mAmps until the dye front exited the bottom of the gel. The gels were then illuminated with wide-UV and imaged on a Multi-DocIt imaging system (UVP, CA). The gels were then stained with coomassie brilliant blue stain and subsequently destained with methanol/acetic acid. Protein stained gels were also imaged on the Multi-DocIt imaging system (UVP, CA).

Catalase Linking. Lyophilized catalase from bovine liver was purchased from Sigma-Aldrich, CA and the bi-functional linker was purchased from Thermo Scientific, CA. Catalase was rehydrated in 50 mM phosphate, 25 mM NaCl, pH 7.2 at a final concentration of 5 mg/mL. The NHS-$PEG_{12}$-Maleimide linker (succinimidyl-[(*N*-maleimidopropionamido)-dodecaethyleneglycol] ester, SMPEG) was resuspended in DMSO to a final concentration of 250 mM. SMPEG was added to the catalase (5 mg in 1 mL buffer) to a final concentration of 1 mM (4 µl of stock) and allowed to react for 30 minutes at room temperature with end-over-end mixing. Excess linker was removed from the catalase via a 30k MWCO spin filter, spinning and washing the catalase at least 4 times. The catalase was subjected to a final spin in the filter resulting in the volume being reduced to ~100 µl. The entirety of the labeled catalase was added to the P22 N375C portal or P22 N377C portal (3 mg in 1 mL pH 7.2 phosphate buffer), and allowed to react for 30 minutes at room temperature with end-over-end mixing. The reaction mixture was then subjected to S-500 size exclusion chromatography as described above.

Anti-Portal and Anti-Coat Protein Western Blots. Samples at an approximate $A_{280nm}$ of 3 were run on 12% acrylamide gels as described above. The gels were then

sandwiched in the following order: black side of the transfer cassette, sponge, filter paper, gel, membrane, filter paper, sponge, red side of the transfer cassette. The cassette was then placed in the electrophoretic transfer apparatus and transferred for 2 hours at 200 mAmps (transfer buffer = 3.03 g Tris HCl, 14.4 g glycine, 200 mL methanol, per liter). The membrane was then blocked in 5% milk (in TTBS, 20 mM Tris 1 mM tween, 100 mM NaCL, pH 7.6) for 1 hour at room temperature. Anti-portal polyconal was added at 1:10,000 dilution while the anti-coat protein polyclonal was added at 1:25,000 dilution. The primary anti-bodies were allowed to react with the membrane for 2 hours at room temperature with gentle mixing. The membranes were then washed vigorously, 5 times, with TTBS. Secondary anti-body (HRP-conjugated) was added at 1:25,000 diltion in 5% milk, TTBS and allowed to incubate for 1 hour at room temperature. The membranes were then washed vigorously, 5 times, with TTBS. The membranes were developed using an HRP-kit (Bio-Rad, CA) and imaged on the Multi-DocIt imaging system (UVP, CA).

## Two Vector Approach for Portal Incorporation

As described in the earlier chapters of this work and by Patterson, et al. (*7, 36*), the Douglas lab has in hand a library of scaffold-fusion P22 assembler plasmids. To expand the utility of these constructs for nanomaterial synthesis, it would be advantageous to be able to integrate the portal complex asymmetric unit into these capsids. It has been shown that a minimal amount of scaffold protein will assemble closed capsids as well as the Douglas lab showing various fusions to the scaffold protein

also allow for proper P22 capsid assembly. However, is it possible to assemble coat protein, scaffold fusions or truncations, and the portal into one closed capsid? Weigele et al. showed that the scaffold protein amino acid residues 229-303 are needed for portal recruitment (*17*). Thus, the Douglas lab fusions containing the scaffold protein 141-303 should be able to incorporate the portal protein.

To create a generalized method to incorporate the portal complex into any of the Douglas lab scaffold protein fusion constructs, a 2-vector approach was taken. By inserting the portal gene on to a vector that is compatible with our fusion vectors (which are either pET11a or pET-Duet), it is possible to co-transfect the expressing *E. coli* resulting in all three proteins (coat, scaffold-fusion, and portal) being expressed. Driving portal expression from a separate plasmid from the coat and scaffold proteins, in a heterologously expressed system, has never been shown. Herein it is shown that this method does lead to portal incorporation in to the assembled capsids and that the SP141 and GFP-SP141 fusion allow for portal incorporation. Interestingly, the wild-type scaffold construct did not incorporate portal protein into the resulting capsids. Therefore, the 2-vector co-transfection method should allow for the asymmetric portal handle to be incorporated into any of our scaffold (141-303)-fusion constructs.

Results and Discussion

The portal protein gene was successfully amplified from the gp 1,8,5 plasmid and inserted in to the pRSFDuet plasmid. pRSFDuet is a kanamycin resistant vector with a RSF1030 origin of replication. This vector is therefore compatible with all pET derived vectors (which have a ColE1 origin of replication) that have a different antibiotic

selection from kanamycin. Different origins of replication are needed as it is thought that with the same origin of replication, vectors are effectively competing for replication factors. This competition leads to the plasmids being expressed differentially, or lost all together. (*132*) It is also important that the two vectors to be used have different antibiotic resistance genes, such that the resulting transformed bacterial colonies can be selected based on the dual-resistance. Thus, as all of the Douglas scaffold-fusion constructs are on pET derived vectors, the pRSFDuet vector was chosen.

The correct insertion of the portal gene into the first multiple cloning site of the pRSFDuet vector was confirmed by single-pass DNA sequencing. This vector was then co-transformed with the ampicillin resistant pET11a vector containing: GFP-SP141 with K118C coat (described in Chapter 2), SP141 with K118C coat (described in Chapter 2), or the wild-type scaffold with wild-type coat. The pRSFDuet and pET11a vectors are both inducible with IPTG. Thus, the co-transformed BL21(DE3) *E. coli* are singly induced to express all three proteins.

The co-transformed cultures all showed P22 capsid formation, however only the GFP-SP141 and SP141 constructs showed portal incorporation (Figure 7.9). All samples assembled well as seen by the SEC chromatogram showing the majority of the protein is in the proper assembled ~60 mL elution peak (Figure 7.8). However, when these purified samples were run on an SDS-PAGE gel, only the GFP-SP141 and SP141 constructs showed portal incorporation (Figure 7.9). The incorporation of portal was further confirmed in these two constructs by an anti-portal western blot (Figure 7.9).

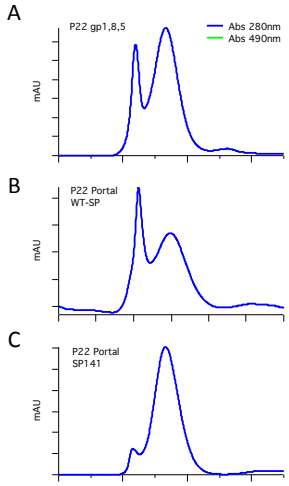



Figure 7.8: SEC of the capsids resulting from co-transformation of the portal and coat/scaffold plasmids showing proper assembly (60 mL elution peak). (A) Positive control, gp 1,8,5 (B) WT scaffold (C) SP141

It is very interesting that the wild-type scaffold, which is what is seen in nature, did not incorporate the portal protein in this approach. The one difference between this vector and the other two (besides the type of scaffold) was that the coat protein was wild type while the others had the K118C coat mutation. It is known to us that this K118C cysteine is available for modification on the inside of the assembled capsid. It is unlikely, but the 118 cysteine could be having an effect in the SP141 and GFP-SP141 constructs. Another, more likely possibility, is that during the co-transforming procedure, a contaminating empty pRSFDuet vector was transformed instead of the portal containing pRSFDuet. This colony would therefore have kanamycin (from pRSFDuet) and ampicillin (from pET11a) resistance but would not express the portal protein. The expression of portal protein alone was not assayed in these experiments, only the incorporation of portal in to the purified capsids. To truly be able to say the portal is not

incorporated into the wild-type scaffold construct, it must be verified that the portal protein is, in fact, being expressed in the co-transformed cells.

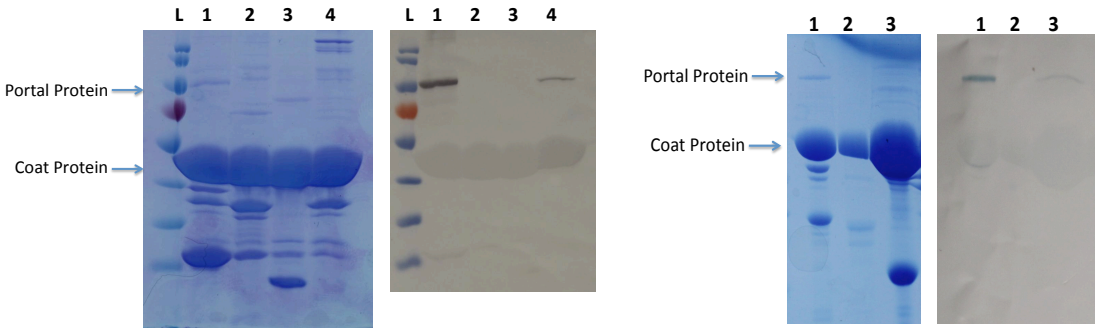



Figure 7.9: SDS-PAGE and anti-portal western blots. On the left: 1)gp 1,8,5 positive control 2) WT SP 3) SP141 4) GFP-SP141. This western shows the incorporation in the GFP-SP141 construct. On the right: 1) gp 1,8,5 positive control 2) negative control 3) SP141. This western shows the portal incorporation into the SP141 construct. For an unexplained reason, the SP141 portal band did not show up on the left western.

## Conclusions

The described general method of co-transformation with the portal protein on a compatible vector, allows for the incorporation of the asymmetric portal unit into hopefully any of the Douglas lab scaffold-fusion constructs. This will broaden the applications for these constructs as they can now have an asymmetric handle for further material synthesis. The asymmetric portal unit will allow for mono-layer coating of surfaces and the controlled assembly of larger architectures with our protein packaged P22 capsids.

## Experimental

The SDS-PAGE, western blot, and SEC protocol can be found described in the first section of this chapter.

Molecular Biology. First, an interior, naturally occurring NcoI restriction enzyme site (CCATGG) had to be knocked out of the portal gene. This was achieved using quick change PCR, as described in the preceding section of this chapter, to change the A in the recognition site to a C. The change was confirmed via DNA sequencing and by the lack of digestion with NcoI. The following primers were used, with the change in bold:

Fwd: 5'- AACGATTGGGTATTTCC**C**TGGCTGACGCAGGACA -3'

Rvs: 5'- TGTCCTGCGTCAGCCA**G**GGAAATACCCAATCGTT -3'

The mutated gene was then amplified using standard PCR to add a 5' NcoI site and a 3' EcoRI site. OneTaq 2x Master Mix was used from New England Biolabs, with 2 µl of each primer (stock at 50 pmol), and 50 ng of the template DNA (gp 1,8,5 plasmid) added to each 50 µl reaction. Cycling parameters were as follows: 95°C for 2 minutes, 25 cycles of 95°C for 20 seconds, 55°C for 20 seconds, 68°C for 2 minutes and then a final extension step at 68°C for 5 minutes. The following primers were used for these additions, with the new restriction site underlined:

Fwd: 5'- AAAA<u>CCATGG</u>GGGCCGACAATGAAAA -3'

Rvs: 5'- AAAA<u>GAATTC</u>TTATTGAGGTGT -3'

The resulting portal gene was cleaned up with a Qiagen PCR cleanup kit, digested with excess NcoI, EcoRI, and dpnI for 2 hours at 37°C, then cleaned up with a Qiagen reaction cleanup kit. The pRSFDuet vector was digested with excess NcoI, EcoRI, and Antarctic phosphotase for 2 hours at 37°C. The linear vector was then gel purified on a 1% agarose slab gel. The band corresponding to linear vector was eluted out of the gel with a Qiagen gel extraction kit. The linearized vector was then ligated to the digested

portal gene using T4 ligase (New England Biolabs, MA). The ligation reaction (2 µl) was then transformed into XL2-Blue ultracompetent cells and the colonies were screened for gene insertion.

Co-Transformation. First, assembler construct containing *E. coli* had to be rendered competent for electroporation. Overnight cultures of BL21(DE3) *E. coli* of wild-type scaffold/coat, GFP-SP141/K118C coat, and SP141/K118C coat were grown, separately, at 37°C. These overnights were used to inoculate 100 mL cultures of fresh LB broth. The cultures were incubated at 37°C with rigorous shaking, until the optical density at 600 nm was approximately 0.6. The cultures were then centrifuged at 4500 rpm for 10 minutes. The resulting cell pellets were then repeatedly resuspended and washed with cold, sterile 10% glycerol. After 4 washing steps, the cell pellets were resuspended in 1 mL of cold 10% glycerol, aliquoted into 40 µl aliquots, and frozen at -80°C. Each cell line is now electro-competent.

One aliquot of each competent cell line, was then subjected to electroporation with 1 µl of pRSFDuet-portal. The resulting cells were then plated on LB-agar plates containing both ampicillin and kanamycin, and incubated at 37°C overnight. The resulting colonies contain both the pET11a coat/scaffold and pRSF portal plasmids.

Protein Production and Purification. The co-transformed cells were grown, harvested, and purified in the same manner as described in the previous section. Both the pET11a and pRSF plasmids are induced with IPTG, so there is no change to the protocol

other than ampicillin (100 μg/mL final concentration) and kanamycin (30 μg/mL final concentration) being added to the culture medium.

## Non-Natural Amino Acid Incorporation into the Portal Protein

Through the incorporation of non-natural amino acids (NNAAs), the chemical and biological diversity of proteins can be expanded. NNAAs have been incorporated *in vivo* to the growing polypeptide chain through the use of a 4 base pair codon, the use of a rare codon, or through reassignment of a sense codon. (*133*) Herein, the incorporation of a methionine analog, azidohomoalanine, into the portal protein is described using the reassignment of the sense codon for methionine. Kiick and coworkers have shown that azidohomoalanine (AHA) is a very good analog to methionine in *E. coli*, activating the methionyl-tRNA synthetase and being incorporated into the protein where a methionine was coded for 95% of the time (*134*). This incorporation strategy is highly effective and simple to employ as it does not require a special tRNA synthetase and the only requirement is to express the protein in a methionine-depleted culture medium. (*134*)

The incorporation of AHA into proteins confers a reactive azide group where a methionine once was. Azides can be specifically addressed through the efficient copper catalyzed cyloaddition of terminal alkyne molecules to form a covalent linkage between the azide-protein and the alkyne-molecule (called “click chemistry”). With the crystal structure of the portal protein published, the incorporation site of an AHA can be intelligently designed wherever desired.

As described in this chapter, the portal asymmetric unit can be incorporated into the vast array of packaged P22 nanocontainers. It has also been shown that the portal protein's cysteines are readily modifiable. To expand the chemical utility of the portal protein, the selective incorporation of AHA into an engineered location on the portal complex is described providing a specific, asymmetric, chemical handle. The advantages of an addressable azide, as opposed to a cysteine, include that azides are not susceptible to any biological molecules while cysteines can form disulfide bridges *in vivo,* rendering them inert for further modification. Additionally, click-chemistry is highly efficient and specific while maleimide-sulfhydryl chemistry is pH dependent and maleimides can label aberrantly. Click chemistry is also very general, as any molecule with a terminal alkyne can be covalently attached to the azide group.

The two vector approach for generally incorporating portal protein into the P22 scaffold-fusion capsids (described in the previous section) is not suitable for selectively incorporating AHA into the portal *only.* In this method, IPTG induces the transcription of both plasmids at the same time and therefore all three proteins are translated at the same time. If this incorporation method were to be used, AHA would be in all three proteins wherever a methionine is coded for and the unique asymmetric reactive site would be lost. Thus, a new genetic strategy was developed for the controlled differential expression of the portal protein allowing for AHA incorporation with the subsequent expression of the coat and scaffold protein in the presence of methionine to initiate the assembly of P22 capsids.

Results and Discussion

The portal protein monomer contains 21 naturally occurring methionine residues (labeled blue in Figure 7.10). It would be incredibly laborious to knock out all 21 natural methionines, and as all that is needed is an asymmetric chemical handle, perhaps the more azide functionalities available in the portal, the better. However, it does appear that there are no severely exposed methionines to the exterior of the capsid. Therefore, a new methionine N380M was engineered into the flexible loops extending outward from the portal complex (labeled red in Figure 7.10). The N380M mutation was made with quick change PCR into the gp 1,8,5 assembler plasmid. The change did not disrupt capsid formation or the incorporation of portal into the plasmid (Figure 7.10).

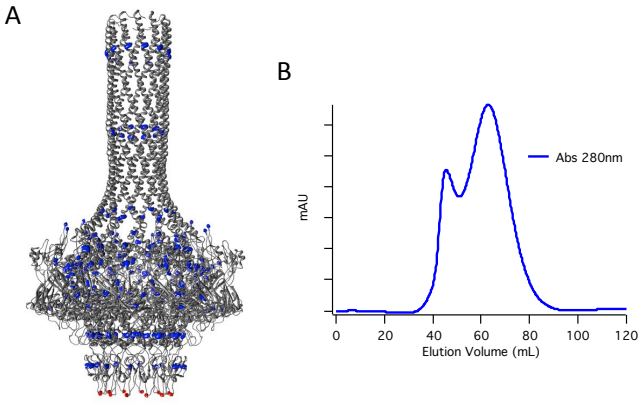



Figure 7.10: (A) Location of natural methionines on the portal complex, colored bue. The installment of a new cysteine (red in A) does not disrupt the capsid structure as seen by SEC (B).

The new two-vector approach results in capsid assembly and portal incorporation. First, the N380M portal gene was inserted into the pBAD vector. This vector has ampicillin resistance, is induced with L-arabinose, and has a ColE1 origin of replication.

Next, the genes for the residues of scaffold protein 141-303 (SP141) and the wild-type coat protein were inserted into the pRSFDuet vector into multiple cloning sites 1 and 2, respectively. The pRSFDuet vector confers kanamycin resistance, is induced with IPTG, and has a RSF1030 origin of replication. The different antibacterial resistances, origin of replications, and chemical inducers make the pBAD and pRSF vectors compatible for this study. Additionally, IPTG has been shown to inhibit expression off of the pBAD vector, adding another layer of expressional control over the system (*135*). To assess if pre-loading cells with expressed portal and then expressing coat/scaffold to assemble capsids would result in portal incorporated capsids, the first experiments were carried out under non-limiting conditions.

Briefly, the two vectors were co-transformed into BL21(DE3) *E. coli* and expressed in LB broth. First, L-arabinose was added to express the portal protein off the pBad plasmid. After several hours, the cells were spun down, washed to remove excess L-arabinose, and resuspended in fresh LB. Next, IPTG at 1 mM (a concentration that inhibits pBad expression) was added inducing the production of coat and scaffold protein and the culture was allowed to grow further. Capsids were then purified out of the culture using a 35% sucrose cushion ultra-centrifugation and assayed for assembly and portal incorporation. The SEC chromatography showed a majority of the sample was assembled capsids (Figure 7.10). Additionally, SDS-PAGE shows portal protein in the assembled capsid SEC fraction. The ratio of portal to coat protein in the gel looks similar to the ratio seen in the gp 1,8,5 control and hence, it is believed that the differential expression yields capsids containing one portal complex per capsid.

The incorporation of AHA does not inhibit capsid assembly or portal incorporation. To incorporate the AHA into the portal only, the two plasmids were co-transformed into pLyseS *E. coli.* This strain has additional controls over its production of T7 polymerase and thus is supposed to have less leaky expression of proteins controlled by the T7 promoter (which controls the expression of coat and scaffold protein in this system). The co-transformed pLyseS cells were rendered methionine-depleted and then used in a similar fashion as described above for AHA incorporation. The culture was grown in a minimal media lacking methionine and containing AHA only when the portal was induced. When the coat and scaffold were expressed, the media was changed to fresh LB (which contains all 20 amino acids). The resulting capsids were well assembled y SEC and showed portal incorporation at the same ratio as the gp 1,8,5 control (Figure 7.11).

AHA is incorporated into the portal selectively and is addressable via click-chemistry. To assess if AHA was incorporated into the purified capsids, they were reacted with an alkyne-fluorescein molecule. The reaction was then run on an SDS-PAGE gel and illuminated with UV light to see where the fluorescein labeled. As seen on the gel, the band corresponding to the portal protein illuminates brightly under UV (Figure 7.11). However, there is some background emission from the coat protein band. However, these proteins are in a ratio of 12 portal monomers to 415 coat protein monomers. Thus, the emission seen for coat, especially in the high concentration on the gel, is very minor as compared to the AHA-incorporated portal fraction. Alternatively, the background emission of the coat protein could be the result of a small fraction of coat

protein being expressed during the AHA-incorporation step. To avoid this, an expression strain with a more tightly controlled T7 polymerase could be used.

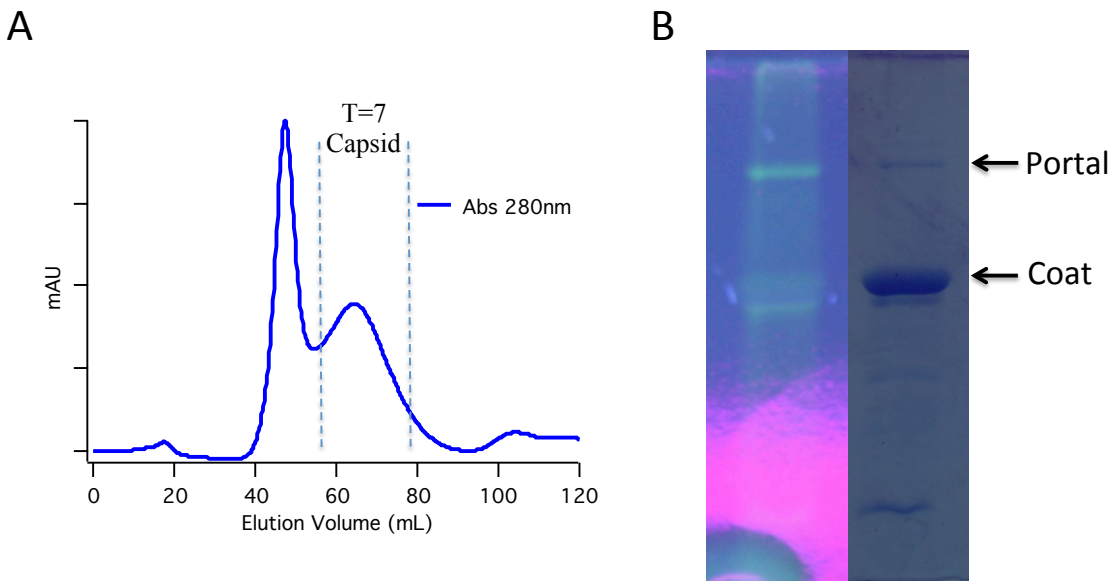



Figure 7.11: Characterization of the azidohomoalanine (AHA) containing capsids. (A) AHA incorporation does not disrupt the capsid morphology by SEC. (B) The AHA is addressable by an alkyne-fluorescein. On the left, UV illumination of the gel and right coomassie stain. The higher, portal band is brightly illuminated showing specific labeling by the alkyne-fluorescein.

To add an asymmetric functionality to the P22-Portal AHA capsids, a gold binding peptide was successfully linked to the portal protein via click chemistry. A gold binding peptide that was identified through phage display (*136*) was synthesized on a soild support to contain an N-terminal alkyne (see experimental). The peptide was subsequently covalently attached to the portal-AHA through click-chemistry. The purified P22-portal-AHA-peptide, when mixed with gold nanoparticles, induces a color change from red to blue indicating binding. This color change was not seen in the P22 no

AHA control reaction. These results are very preliminary, but encouraging, and more experiments need to be carried out to determine if the gold is binding asymmetrically to the portal or decorating the entire capsid.

## Conclusions

The use of non-natural amino acids greatly diversifies the addressable chemical moieties contained in a protein. Azidohomoalanine has been shown to incorporate well as a replacement for methionine when expressed in *E. coli*. Using this system, AHA has been successfully incorporated into the portal protein of the P22 capsid. Using an intelligently designed two-vector approach, it was possible to direct the AHA to only the portal protein unit of the P22 capsid assembly. The incorporated AHA molecules are readily addressable by a small molecule alkyne dye and a functional peptide. By covalently linking a gold binding peptide to the asymmetric unit of the P22 capsid, it is possible to build monolayer films of P22 on to gold particles in solution. If this same peptide were to decorate the whole capsid, gold particle-P22 aggregates would form as one P22 could interact with two or more gold particles. In conclusion, the incorporation of a specific chemical handle on the asymmetric unit of the P22 capsid will greatly increase the utility of the P22 as a nanomaterial.

## Experimental

Standard PCR reaction conditions and SDS-PAGE protocol can be found described in the first section of this chapter.

Molecular Biology. The N380M mutation in the portal protein was engineered into the gp 1,8,5 plasmid using quick change PCR as described in the previous sections. The primers used were as follows, with the change in bold:

Fwd: 5'- AATCGCACTGACGAA**ATG**AGTGGAGACCTTCCGA -3'

Rvs: 5'- TCGGAAGGTCTCCACT**CAT**TTCGTCAGTGCGATT -3'

The N380M portal gene was amplified using portal gene amplification primers rom the previous section to add an N-terminal NcoI site and a C-terminal EcoRI site. The digested gene was ligated into the NcoI/EcoRI linearized pBad vector. The wild-type coat protein was amplified out of the gp 1,8,5 plasmid adding an N-terminal NdeI and a C-terminal XhoI site. This product was ligated into MCS2 of the linearized pRSF vector. The SP141 gene was amplified off of the pET11a GFP-SP141 plasmid adding an N-terminal EcoRI and a C-terminal SacI site. The primers used for these amplifications are as follows, with the new restriction enzyme sites underlined:

CP Fwd: 5'- AAAACATATGGCTTTGAACGAA

CP Rvs: 5'- AAAACTCGAGTTACGCAGTCTGACC

SP141 Fwd: 5'- AAAAGAATTCCTGGTGCCGCGCGG

SP141 Rvs: 5'- AAAAGAGCTCTTATCGGATTCCTT

Differential Expression, Non-limiting Conditions. BL21(DE3) were co-transformed with 1 µl each of the pBAD-N380M portal plasmid and the pRSFDuet SP141-WT Coat plasmid. The resulting transformation was spread on a LB agar plate containing both ampicillin and kanamycin. The resulting colonies had both plasmids, as they are both needed for resistance to the antibiotics. A colony was chosen to inoculate a

10 mL overnight culture. This overnight culture was used to inoculate a 1 L culture of LB broth. The culture was grown at 37°C with shaking to an optical density at 600 nm of 0.6 and then induced with 10 mL of a 20% L-arabinose solution. The culture was allowed to grow for 3 more hours. After this time, the cells were centrifuged at 4,500 rpm for 15 minutes. The cell pellet was resuspended in fresh LB, pelleted, and washed a second time to assure the removal of the L-arabinose. Only half of the cells were then resuspended in 1 L of fresh LB broth, to reduce the cell density, and allowed to recover in the shaking 37°C incubator for 30 minutes. Next, IPTG was added to a final concentration of 1 mM to induce the production of coat and scaffold protein. The culture was allowed to grow for 3 more hours, and then the cells were pelleted at 4,500 rpm for 15 minutes, resuspended in lysis buffer, and frozen at -20°C. The capsids were then harvested as previously described in this chapter.

Differential Expression, AHA Incorporation. First, pLyseS cells (Agilent, CA) were co-transformed with 1 µl each of the pBAD-N380M portal plasmid and the pRSFDuet SP141-WT Coat plasmid. The resulting transformation was spread on a LB agar plate containing both ampicillin and kanamycin. The resulting colonies had both plasmids, as they are both needed for resistance to the antibiotics. A colony was chosen to inoculate a 10 mL overnight culture in LB broth. This overnight culture was used to inoculate a 1 L culture of M9 minimal media supplemented with all 19 amino acids except methionine. The culture was grown at 37°C with shaking over night. In the morning, a 10 mL aliquot was used to inoculate a fresh 1 L of M9 minimal media, and this was also allowed to grow over night. The two passages in minimal media lacking

methionine effectively render the pLyseS cells methionine deficient. Once depleted, the cells were used to inoculate a third, fresh 1 L M9 minimal culture. Once in mid-log, the cells were induced with L-arabinose to express portal and AHA was added to a final concentration of 100 μg/mL. After 4 hours, the cells were spun down, and washed and resuspended four times to remove any residual AHA and arabinose. The cells were finally resuspended in LB broth in twice their original volume to dilute the cells back to a mid-log phase. The cells were allowed to recover for 30 minutes, and then were induced with 1 mM IPTG to turn on coat and scaffold protein production and allowed to grow overnight. The resultant cells were then harvested, purified, and assayed for capsid assembly, portal incorporation, and AHA incorporation to the portal.

Gold-Peptide Synthesis. The peptide VSGSSPDS was chosen as it has been shown to bind gold tightly (*136*). The entirety of the synthesis was carried out in a 5 mL fritted syringe on an amide resin (2-chlorotrityl resin, Novabiochem). 200 mg of dry amine resin (0.75 mM amines per gram of resin loading capacity) was first swollen and deprotected with 5 mL of 20% piperidine/80% dimethylformamide (DMF) (solution 1) for 40 minutes, with rocking, at room temperature. After deprotection, the resin was washed with 5 mL of DMF then 5 mL dichloromethane (DCM) two times, then the wash is expelled from the syringe. Next, the first amino acid was added (amino acids are added from C to N-terminus). 4 equivalents of Fmoc-Serine (for 200 mg resin, this is 0.6 mM Fmoc-Serine), 4 equivalents of O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium Hexafluorophosphate (HBTU), and 6 equivalents of N,N-Diisopropylethylamine

(DIEA) in 3 mL of DMF, were quickly mixed, briefly sonicated, and loaded on to the resin. The resin was allowed to react with the first amino acid for 1 hour at 35°C with shaking. The amino acid solution was then expelled, the resin washed with 5 mL DMF then 5 mL DCM two times, and then the deprotecting solution 1 was added to the resin, and allowed to react for 40 minutes at room temperature with shaking. The resin was then washed twice with 5 mL each DMF/DCM, and the next amino acid, Fmoc-asparagine was loaded in the same manner as the first. This method was cycled loading all of the Fmoc-protected amino acids from the C to the N-terminus of the peptide. After the last amino acid was added, Fmoc-valine, it was deprotected as described and then an alkyne was added to the N-terminus of the peptide. To do this, 850 µl of 1.2 M bromoacetic acid and 200 µl of N,N'-Diisopropylcarbodiimide (DIC) were added to the resin and reacted for 30 minutes at room temperature with shaking. The resin was then washed with DMF/DCM four times. A solution of 1 M propargyl amine in DMF was then added to the resin and allowed to react for 40 minutes at room temperature with shaking. The resin was then washed a last time with DCM/DMF three times and the resin was brought to complete dryness in a lyophilizer.

Next, the alkyne-peptide was cleaved off the resin using 1 mL of 95:4:1 trifluoroacetic acid (TFA):Water:triisopropylsilane (TIPS) for 2 hours at room temperature with shaking. The resultant solution was eluted off the resin, which contains the peptide. The peptide solution was diluted 1:10 in cold ether and stored at -20°C over night to precipitate it out of solution. The precipitate was then filtered out of solution, washed with cold ether, dried, and then resuspended in water.

The alkyne-peptide was then purified via HPLC on a Zorbax 300SB-C18, 5μm, 9.4 x 250 mm, reverse phase column. 3 mL of sample were loaded at a time, and eluted by a 0-60%B gradient for 15 minutes (flow rate 1 mL/min, solution A= water, 0.01% TFA, B = 95:5 ACN:water, 0.01% TFA. The major peaks were identified using electrospray mass spectrometry and the peaks corresponding to 829 m/z (peptide calculated mass 868 m/z) from all runs were pooled. The sequence of this fraction was confirmed using MS/MS and the N-terminal sequence, SSGS was confirmed with certainty. The fractions were then lyophilized to complete dryness and the resulting white powder resuspended in 500 μL of 50 mM phosphate, 25 mM NaCl, pH 7.2 and stored at 4°C.

Click-Chemistry Reactions: The protocol used is from Hong, 2009. (*137*) Stock solutions were made fresh as follows: $CuSO_4$ (20 mM in water), TBTA ligand (50 mM in DMSO), sodium ascobate (100 mM in water), aminoguanidine hydrochloride (100 mM in water), and fluorescein alkyne (20 mM in DMSO). The P22-portal AHA was brought to an $A_{280nm}$ of 3. The alkyne gold peptide, due to its lack of aromatic residues leading to no $A_{280nm}$ signal, was used as purified in PBS pH 7.2. For the reaction mixture, the following was mixed in this order: 50 μl of P22-portal AHA, 382.5 μl phosphate buffer at pH 7.2, 10 μl of the alkyne molecule (either the fluorescein or the peptide), 7.5 μl of premixed $CuSO_4$ and ligand (2.5 μl $CuSO_4$ + 5 μl ligand), 25 μl aminoguanidine, and 25 μl sodium ascorbate. The reaction mixtures were vortexed and allowed to react for one hour at room temperature with end-over-end mixing. The fluorescein reaction was protected from light by wrapping the reaction tube in foil. The reaction was then

subjected to a 100k MWCO spin filter and the sample was concentrated and washed with phosphate buffer four times. Lastly, the resulting washed reaction was reduced to ~50 µl in volume.

## Overall Conclusions and Future Directions

Exploitation of the naturally occurring portal complex of the P22 capsid is a simple, *in vivo,* method to incorporate an asymmetric unit to the P22 nanomaterial building block. This biomimetic strategy is by far the easiest method described, as the asymmetric unit is inherent to the assembly. By using a compatible plasmid carrying the portal gene and co-transformation, the incorporation of portal can be generally used in any of our scaffold-fusion assemblers (i.e. fluorescent protein or enzyme fusions). Although this approach appears to be general, it should be tested with more of our constructs, especially the concatenated fusions, as their larger size may inhibit portal incorporation.

In this chapter, it has been shown that the cysteines in the portal (naturally occurring or engineered) are addressable through maleimide chemistry for modification with small molecules or for linkage to large proteins. Additionally, a molecular strategy to selectively incorporate a non-natural amino acid into the portal complex has been described. This method adds an azide functionality specifically to the portal that can be used for site-specific modification via click chemistry.

The addition of the portal complex to our biomimetic approach to materials synthesis, greatly expands the "tool-box" for building P22-based materials. By giving

our P22 building block directionality, we can have greater control over the larger assemblages we create. The advantage of a single reactive surface on an other wise anisotropic material, makes it possible to coat nanoparticles in solution without causing inter-particle interactions (i.e. make aggregates). The directionality of the P22 building block will allow for building materials in a controlled way.

An idea not discussed in this chapter, is utilizing the portal complex as a linking point between two P22 capsids, creating a barbell shaped "Gemini-particle". A method for incorporating an alkyne containing non-natural amino acid has been described by reassignment of the methionine codon (*133*). Using the non-natural amino acid described herein, one could make an alkyne-portal construct and an azide-portal construct and covalently click the two capsids together creating yet another class of biomaterials. As the portal genes are on separate plasmids from the coat and scaffold protein, this can be used as a "plug-and-play" approach; allowing for expansion of any capsid material. One could link capsids containing different enzymes in a metabolic pathway, bring a metal catalyst containing capsid to an enzyme containing capsid, have two differentially packaged small molecules the need to be delivered together for a desired effect, or even package a pro-drug in one capsid and the pro-drug cleaving enzyme in the other.

Overall, the portal incorporation strategies described in this chapter are potentially powerful tools for expanding the types of nanomaterials that can be designed. The strategies are general and can be incorporated into any P22 platform being devised.

# CHAPTER 8

# CONCLUDING REMARKS

Protein containers have been utilized for the encapsulation of active proteins, inorganic particles, and polymers all with advantageous effects. The defined structure of protein containers leads to the controlled incorporation of cargo molecules either by spatial control of the docking agent or through restriction of the cargos synthesis. Through encapsulation the cargos, functionality is increased as the container lends a handle for further modification. The encapsulated cargos can thus be targeted, shielded from the immune response, or assembled in to larger architectures. Encapsulation also confers protection of the cargo from the exterior environment. For encapsulated enzymatic cargos, catalytic rates can be modulated and in some cases enhanced due to the sequestration of high concentrations of substrates.

The bacteriophage P22 is a remarkable protein container. The assembly process is well defined, leading to precise T=7 capsids that are easily modified genetically or chemically. The capsids are malleable and three additional morphologies are accessible for nanomaterial design. The ability to access the expanded morphology, allows for the investigation of the same encapsulated cargo at two distinct concentrations when compared to the procapsid morphology. Additionally, the wiffle-ball structure, with its large pores allows for release of the cargo or uninhibited diffusion in and out of the capsid. Another remarkable feature of the P22 capsid, is the ability to truncate or fuse poly-peptides to the scaffold protein and still template proper assembly. The naturally occurring asymmetric unit of the P22 capsid, the portal complex, confers another level of control over our P22

nanomaterial. With all of these advantages to the P22 system, the P22 capsid is an amazing, versatile building block for nanomaterial design.

In chapter 2 of this work, the initial strategy for directed protein encapsulation inside the P22 capsid is described. It is shown that the N-terminal fusion of fluorescent proteins GFP and mCherry to the truncated scaffold protein results in templated assembly and directed encapsulation. Using this method, it is possible to package up to 300 copies of the fluorescent fusion with no effect to the capsid stability of structure. The cargo occupies more than 24% of the available volume. By accessing the alternative forms of the capsid, the cargo can be released from the wiffle ball through its large 10 nm pores. Additionally, when taken to the expanded form, it is shown that the scaffold proteins access the exterior environment and are sensitive to thrombin cleavage. This treatment leaves the cleaved cargo molecules encapsulated in the capsid. This chapter illustrates the robustness of the P22 encapsulation system and the ability to modulate the capsid after assembly for advantageous affects.

Chapter 3 is an extension of chapter 2 in that it illustrates the fusion of GFP and mCherry with the scaffold protein as a concatamer. This triple fusion is also able to template assembly and package cargo in high copy number. By encapsulating GFP and mCherry together, the FRET communication of the pair can be probed inside the capsid as compared to the fusion free in solution. This is the first example of measuring protein communication in a protein container. The concatamer was spaced with glycine repeats of varying lengths between the donor and acceptor in influence the FRET behavior. The fusions were investigated in the densely packaged procapsid form and the expanded form which has twice the volume as the procapsid. In the expanded form, the fusions are effectively diluted in half but remain encapsulated.

The FRET results show that in the procapsid form, the FRET is greatly enhanced and hence, the fluorescent fusions are forced next to their neighbors and exhibit inter-molecular FRET as well as intra. In the expanded capsid, where the fusions are diluted in half, some of the FRET signal is lost however it is still three times higher than seen in the unencapsulated samples. The linker length between the donor and acceptor showed a trend of less FRET in the longest linker and more FRET in the shortest linker. However, this trend was only seen in the encapsulated samples. Thus, the flexibility of the linkers could only be fully realized when the fusions were free in solution and not constrained in the capsid. This study demonstrates that the scaffold fusion is tolerant of concatemers and through encapsulation in a protein container, it is possibly to force communication between proteins.

In chapter 4, the encapsulation of a phosphotriesterase enzyme in the P22 capsid and its subsequent increased stability. Phosphotriesterase (PTE) is a mesophilic enzyme that readily degrades nerve agents and pesticides. The fusion of PTE to the scaffold protein lead to the assembly of non-perfect capsids. It was hypothesized, and suggested by data, that because this fusion is expressed in such low quantities, the templating of the capsid assembly was affected. However, through encapsulation, the thermal stability of the enzyme was greatly increased. Additionally, the capsid conferred protection of the enzyme to proteolytic degradation and the process of lyophilization and rehydration. This is the first example of a mesophilic enzyme being packaged in P22 and thus illustrates the stability conferring capabilities the P22 capsid can have. This is also the first report investigating the role of the expression level of the scaffold-fusion on the assembly outcomes.

Investigating the types of fusions that can be encapsulated further, chapter 5 demonstrates the encapsulation of a poly-glutamate peptide in the P22 capsid. The poly-glutamate adds a great deal of negative charge to the interior surface of the capsid. This re-engineering allows for the sequestration and subsequent mineralization of iron oxide particles inside the capsid. The capsid sequesters, templates, and restricts the size of the iron oxide particles. This research establishes the use of repetitive extensions on the scaffold protein for re-engineering of the interior of the P22 capsid. It also highlights the utility of the P22 capsid as a restrictive reaction vessel for nanoparticle growth.

The promise of the poly-glutamate fusion lead to the development of a poly-lysine and a poly-tyrosine scaffold fusion discussed in chapter 6. Both poly-peptides templated the assembly of homogeneous capsids and are presumably packaged. In the case of the poly-lysine, a positively charged fusion, the subsequent loading of nucleic acid was inconclusive. However, the poly-tyrosine fusion was successfully iodinated for use as an X-ray contrast agent. In this chapter a third peptide fusion was discussed that codes for a triple fusion of the the neurotoxic conotoxin, MVIIA, the fluorescent protein mCherry, and the scaffold protein. This triple fusion templated assembly and was sequestered on the interior of the capsid. Through encapsulation of the MVIIA peptide, the capsid was used as a molecular handle to add a targeting moiety. The targeting peptide TAT was successfully appended to the exterior cysteines of the capsid and resulted in the engulfment of the capsids by a blood-brain mimicking cell line. The green fluorescent tag on the TAT combined with the red fluorescence of the mCherry inside the capsid, allowed for both the capsid and cargo to be tracked simultaneously. The two signals remained co-localized thus suggesting

the capsid retains its assembly after entering the cells. It was also shown that the capsids are trafficked to the lysosomes inside the cell. This chapter illustrates the wide possibilities of activities that can be accessed through genetic fusion with the scaffold protein. Additionally, the exterior of the capsid can be modified with targeting peptides resulting in the capsids entry into cells.

Lastly, methods for the exploitation of the naturally occuring asymmetric unit of the P22 capsid was described in chapter 7. Three incorporation strategies were devised for different applications. The first modified the in-hand plasmid (gp 1,8,5) that results in capsids with new cysteines in their portal unit. It was shown that the naturally occurring cysteines readily label with small molecule-maleimides. However the newly engineered portal cysteines on the exterior surface could be selectively labeled by a large protein, catalase. The linked samples maintained the capsid structure and catalase activity. The second method describes a general method of co-transformation of a compatible plasmid carrying the portal gene with that of potentially any of the Douglas lab pET based fusion assembly vectors. Herein, the expression of the SP141 truncation and the GFP-SP141 plasmids allowed for the incorporation of portal, whose expression is driven off a separate plasmid. A third incorporation strategy was described that allowed for the selective integration of a non-natural amino acid, azidohomoalanine, in to the portal as well as the portals incorporation into capids. This two-vector approach exploits the use of different chemical inducers to drive first the expression of portal, and the incorporation of AHA, and then subsequently to drive the expression of coat and scaffold protein without AHA, to assemble capsids. The AHA functionality allowed for the selective chemical modification of he portal through click chemistry. Fluorescein-alkyne and a

gold binding peptide-alkyne were both successfully linked to the portal. This chapter highlights the general ability to incorporate an asymmetric handle to any of he P22 fusion capsids. Additionally, through the use of non-natural amino acid incorporation, a specific chemical handle can be added *in vivo* to the asymmetric unit. The addition of the asymmetric portal to the capsid greatly increases the control of the P22 nanomaterial for the construction of larger assemblages.

In conclusion, the beauty of the bacteriophage P22 has inspired a wide variety of nanomaterials. The ability of all of our building blocks to be made *in vivo* sets this system apart, and way ahead, of the competition. The precise architecture of the P22 capsid has leant itself to the creation homogeneous nano-building blocks encapsulating active cargos. The scaffold protein-fusion genetic design, described herein, has proved to be an incredibly versatile encapsulation strategy. The scaffold fusions accurately template capsid assembly and provide a signal for encapsulation. The availability to add in the asymmetric portal complex unit to any of our P22 nanomaterial designs, further enhances our ability to engineer new materials.

Now that the P22 nanomaterial toolbox has been filled with numerous strategies, the synthesis of new P22 materials is only limited by the imagination. Our approach is extremely robust and versatile and will certainly result in the advancement of the bio-inspired nanomaterials field.

# REFERENCES CITED

1. de la Escosura, A., Janssen, P. G. A., Schenning, A. P. H. J., Nolte, R. J. M., and Cornelissen, J. J. L. M. (2010) Encapsulation of DNA-Templated Chromophore Assemblies within Virus Protein Nanotubes, *Angewandte Chemie International Edition 122*, 5463-5466.

2. Comellas-Aragonès, M., de la Escosura, A., Dirks, A. J., van der Ham, A., Fusté-Cuñé, A., Cornelissen, J. J. L. M., and Nolte, R. J. M. (2009) Controlled Integration of Polymers into Viral Capsids, *Biomacromolecules 10*, 3141-3147.

3. Lucon, J., Qazi, S., Uchida, M., Bedwell, G. J., LaFrance, B., Prevelige, P. E., and Douglas, T. (2012) Use of the interior cavity of the P22 capsid for site-specific initiation of atom-transfer radical polymerization with high-density cargo loading, *Nature Chemistry 4*, 781-788.

4. Tang, J., Lander, G. C., Olia, A., Li, R., Casjens, S., Prevelige, P., Cingolani, G., Baker, T. S., and Johnson, J. E. (2011) Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22, *Structure (London, England : 1993) 19*, 496-502.

5. Douglas, T., and Young, M. (2006) Viruses: Making Friends with Old Foes, *Science 312*, 873-875.

6. Worsdorfer, B., Woycechowsky, K. J., and Hilvert, D. (2011) Directed Evolution of a Protein Container, *Science 331*, 589-592.

7. Patterson, D. P., Prevelige, P. E., and Douglas, T. (2012) Nanoreactors by Programmed Enzyme Encapsulation Inside the Capsid of the Bacteriophage P22, *ACS Nano 6*, 5000-5009.

8. Comellas-Aragones, 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., and Nolte, R. J. M. (2007) A virus-based single-enzyme nanoreactor, *Nat Nano 2*, 635-639.

9. Chen, C., Daniel, M.-C., Quinkert, Z. T., De, M., Stein, B., Bowman, V. D., Chipman, P. R., Rotello, V. M., Kao, C. C., and Dragnea, B. (2006) Nanoparticle-Templated Assembly of Viral Protein Cages, *Nano Letters 6*, 611-615.

10. Kostiainen, M. A., Ceci, P., Fornara, M., Hiekkataipale, P., Kasyutich, O., Nolte, R. J. M., Cornelissen, J. J. L. M., Desautels, R. D., and van Lierop, J. (2011) Hierarchical Self-Assembly and Optical Disassembly for Controlled Switching of Magnetoferritin Nanoparticle Magnetism, *ACS Nano 5*, 6394-6402.

11. Inoue, T., Kawano, M.-a., Takahashi, R.-u., Tsukamoto, H., Enomoto, T., Imai, T., Kataoka, K., and Handa, H. (2008) Engineering of SV40-based nano-

capsules for delivery of heterologous proteins as fusions with the minor capsid proteins VP2/3, *Journal of Biotechnology 134*, 181-192.

12. Lander, G. C., Tang, L., Casjens, S. R., Gilcrease, E. B., Prevelige, P., Poliakov, A., Potter, C. S., Carragher, B., and Johnson, J. E. (2006) The Structure of an Infectious P22 Virion Shows the Signal for Headful DNA Packaging, *Science 312*, 1791-1795.

13. van Eldijk, M. B., Wang, J. C. Y., Minten, I. J., Li, C., Zlotnick, A., Nolte, R. J. M., Cornelissen, J. J. L. M., and van Hest, J. C. M. (2012) Designing Two Self-Assembly Mechanisms into One Viral Capsid Protein, *Journal of the American Chemical Society 134*, 18506-18509.

14. Uchida, M., Flenniken, M. L., Allen, M., Willits, D. A., Crowley, B. E., Brumfield, S., Willis, A. F., Jackiw, L., Jutila, M., Young, M. J., and Douglas, T. (2006) Targeting of Cancer Cells with Ferrimagnetic Ferritin Cage Nanoparticles, *Journal of the American Chemical Society 128*, 16626-16633.

15. Olia, A. S., Prevelige, P. E., Johnson, J. E., and Cingolani, G. (2011) Three-dimensional structure of a viral genome-delivery portal vertex, *Nature Structural & Molecular Biology 18*, 597-603.

16. Minten, I. J., Hendriks, L. J. A., Nolte, R. J. M., and Cornelissen, J. J. L. M. (2009) Controlled Encapsulation of Multiple Proteins in Virus Capsids, *Journal of the American Chemical Society 131*, 17771-17773.

17. Weigele, P. R., Sampson, L., Winn-Stapley, D., and Casjens, S. R. (2005) Molecular Genetics of Bacteriophage P22 Scaffolding Protein's Functional Domains, *Journal of Molecular Biology 348*, 831-844.

18. Wu, W., Hsiao, S. C., Carrico, Z. M., and Francis, M. B. (2009) Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery, *Angewandte Chemie International Edition 48*, 9493-9497.

19. Padilla-Meier, G. P., and Teschke, C. M. (2011) Conformational Changes in Bacteriophage P22 Scaffolding Protein Induced by Interaction with Coat Protein, *Journal of Molecular Biology 410*, 226-240.

20. Tanaka, S., Sawaya, M. R., and Yeates, T. O. (2010) Structure and Mechanisms of a Protein-Based Organelle in Escherichia coli, *Science 327*, 81-84.

21. Fiedler, J. D., Brown, S. D., Lau, J. L., and Finn, M. G. (2010) RNA-Directed Packaging of Enzymes within Virus-like Particles, *Angewandte Chemie International Edition 49*, 9648-9651.

22. Padilla, J. E., Colovos, C., and Yeates, T. O. (2001) Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments, *Proceedings of the National Academy of Sciences of the United States of America 98*, 2217-2221.

23. Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C., and Yeates, T. O. (2008) Atomic-Level Models of the Bacterial Carboxysome Shell, *Science 319*, 1083-1086.

24. Ponka, P., Beaumont, C., and Richardson, D. R. (1998) Function and regulation of transferrin and ferritin, *Seminars in Hematology 35*, 35-54.

25. Zhou, Z. H., McCarthy, D. B., O'Connor, C. M., Reed, L. J., and Stoops, J. K. (2001) The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes, *Proceedings of the National Academy of Sciences of the United States of America 98*, 14802-14807.

26. Seebeck, F. P., Woycechowsky, K. J., Wei, Z., Rabe, J. P., and Hilvert, D. (2006) A Simple Tagging System for Protein Encapsulation, *Journal of the American Chemical Society 128*, 4516-4517.

27. Worsdorfer, B., Pianowski, Z., and Hilvert, D. (2011) Efficient in Vitro Encapsulation of Protein Cargo by an Engineered Protein Container, *Journal of the American Chemical Society 134*, 909-911.

28. Fan, C., Cheng, S., Liu, Y., Escobar, C. M., Crowley, C. S., Jefferson, R. E., Yeates, T. O., and Bobik, T. A. (2010) Short N-terminal sequences package proteins into bacterial microcompartments, *Proceedings of the National Academy of Sciences 107*, 7509-7514.

29. Rhee, J.-K., Hovlid, M., Fiedler, J. D., Brown, S. D., Manzenrieder, F., Kitagishi, H., Nycholat, C., Paulson, J. C., and Finn, M. G. (2011) Colorful Virus-like Particles: Fluorescent Protein Packaging by the Qβ Capsid, *Biomacromolecules 12*, 3977-3981.

30. O'Neil, A., Prevelige, P. E., Basu, G., and Douglas, T. (2012) Coconfinement of Fluorescent Proteins: Spatially Enforced Communication of GFP and mCherry Encapsulated within the P22 Capsid, *Biomacromolecules 13*, 3902-3907.

31. Miklos, A. C., Sarkar, M., Wang, Y., and Pielak, G. J. (2011) Protein Crowding Tunes Protein Stability, *Journal of the American Chemical Society 133*, 7116-7120.

32. Minton, A. P., and Wilf, J. (1981) Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase, *Biochemistry 20*, 4821-4826.

33. Wang, Y., Sarkar, M., Smith, A. E., Krois, A. S., and Pielak, G. J. (2012) Macromolecular Crowding and Protein Stability, *Journal of the American Chemical Society 134*, 16614-16618.

34. Zhou, H.-X., Rivas, G., and Minton, A. P. (2008) Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences, *Annual Review of Biophysics 37*, 375-397.

35. Minten, I. J., Claessen, V. I., Blank, K., Rowan, A. E., Nolte, R. J. M., and Cornelissen, J. J. L. M. (2010) Catalytic capsids: the art of confinement, *Chemical Science*, 358-362.

36. Patterson, D. P., Schwarz, B., El-Boubbou, K., van der Oost, J., Prevelige, P. E., and Douglas, T. (2012) Virus-like particle nanoreactors: programmed encapsulation of the thermostable CelB glycosidase inside the P22 capsid, *Soft Matter 8*, 10158-10166.

37. Glasgow, J. E., Capehart, S. L., Francis, M. B., and Tullman-Ercek, D. (2012) Osmolyte-Mediated Encapsulation of Proteins inside MS2 Viral Capsids, *ACS Nano 6*, 8658-8664.

38. Minten, I. J., Claessen, V. I., Blank, K., Rowan, A. E., Nolte, R. J. M., and Cornelissen, J. J. L. M. (2011) Catalytic capsids: the art of confinement, *Chemical Science 2*, 358-362.

39. Brasch, M., and Cornelissen, J. J. L. M. (2012) Relative size selection of a conjugated polyelectrolyte in virus-like protein structures, *Chemical Communications 48*, 1446-1448.

40. Brasch, M., de la Escosura, A., Ma, Y., Uetrecht, C., Heck, A. J. R., Torres, T., and Cornelissen, J. J. L. M. (2011) Encapsulation of Phthalocyanine Supramolecular Stacks into Virus-like Particles, *Journal of the American Chemical Society 133*, 6878-6881.

41. Kwak, M., Minten, I. J., Diana-Milena, A., Musser, A. J., Brasch, M., Nolte, R. J. M., Müllen, K., Cornelissen, J. J. L. M., and Herrmann, A. (2010) Virus-like Particles Templated by DNA Micelles: A General Method for Loading Virus Nanocarriers, *Journal of the American Chemical Society 132*, 7834-7835.

42. Liepold, L. O., Abedin, M. J., Buckhouse, E. D., Frank, J. A., Young, M. J., and Douglas, T. (2009) Supramolecular Protein Cage Composite MR Contrast Agents with Extremely Efficient Relaxivity Properties, *Nano Letters 9*, 4520-4526.

43. Lucon, J., Abedin, M. J., Uchida, M., Liepold, L., Jolley, C. C., Young, M., and Douglas, T. (2010) A click chemistry based coordination polymer inside small heat shock protein, *Chemical Communications 46*, 264-266.

44. Pokorski, J. K., Breitenkamp, K., Liepold, L. O., Qazi, S., and Finn, M. G. (2011) Functional Virus-Based Polymer: Protein Nanoparticles by Atom Transfer Radical Polymerization, *Journal of the American Chemical Society 133*, 9242-9245.

45. Kovacs, E. W., Hooker, J. M., Romanini, D. W., Holder, P. G., Berry, K. E., and Francis, M. B. (2007) Dual-Surface-Modified Bacteriophage MS2 as an

Ideal Scaffold for a Viral Capsid-Based Drug Delivery System, *Bioconjugate Chemistry 18*, 1140-1147.

46. Raja, K. S., Wang, Q., Gonzalez, M. J., Manchester, M., Johnson, J. E., and Finn, M. G. (2003) Hybrid Virus,àíPolymer Materials. 1. Synthesis and Properties of PEG-Decorated Cowpea Mosaic Virus, *Biomacromolecules 4*, 472-476.

47. Huang, X., Bronstein, L. M., Retrum, J., Dufort, C., Tsvetkova, I., Aniagyei, S., Stein, B., Stucky, G., McKenna, B., Remmes, N., Baxter, D., Kao, C. C., and Dragnea, B. (2007) Self-Assembled Virus-like Particles with Magnetic Cores, *Nano Letters 7*, 2407-2416.

48. Huang, X., Stein, B. D., Cheng, H., Malyutin, A., Tsvetkova, I. B., Baxter, D. V., Remmes, N. B., Verchot, J., Kao, C., Bronstein, L. M., and Dragnea, B. (2011) Magnetic Virus-like Nanoparticles in N. benthamiana Plants: A New Paradigm for Environmental and Agronomic Biotechnological Research, *ACS Nano 5*, 4037-4045.

49. Sun, J., DuFort, C., Daniel, M.-C., Murali, A., Chen, C., Gopinath, K., Stein, B., De, M., Rotello, V. M., Holzenburg, A., Kao, C. C., and Dragnea, B. (2007) Core-controlled polymorphism in virus-like particles, *Proceedings of the National Academy of Sciences 104*, 1354-1359.

50. Daniel, M.-C., Tsvetkova, I. B., Quinkert, Z. T., Murali, A., De, M., Rotello, V. M., Kao, C. C., and Dragnea, B. (2010) Role of Surface Charge Density in Nanoparticle-Templated Assembly of Bromovirus Protein Cages, *ACS Nano 4*, 3853-3860.

51. Aniagyei, S. E., Kennedy, C. J., Stein, B., Willits, D. A., Douglas, T., Young, M. J., De, M., Rotello, V. M., Srisathiyanarayanan, D., Kao, C. C., and Dragnea, B. (2009) Synergistic Effects of Mutations and Nanoparticle Templating in the Self-Assembly of Cowpea Chlorotic Mottle Virus Capsids, *Nano Letters 9*, 393-398.

52. Loo, L., Guenther, R. H., Basnayake, V. R., Lommel, S. A., and Franzen, S. (2006) Controlled Encapsidation of Gold Nanoparticles by a Viral Protein Shell, *Journal of the American Chemical Society 128*, 4502-4503.

53. Loo, L., Guenther, R. H., Lommel, S. A., and Franzen, S. (2007) Encapsidation of Nanoparticles by Red Clover Necrotic Mosaic Virus, *Journal of the American Chemical Society 129*, 11111-11117.

54. de la Escosura, A., Verwegen, M., Sikkema, F. D., Comellas-Aragonès, M., Kirilyuk, A., Rasing, T., Nolte, R. J. M., and Cornelissen, J. J. L. M. (2008) Viral capsids as templates for the production of monodisperse Prussian blue nanoparticles, *Chemical Communications 0*, 1542-1544.

55. Uchida, M., Terashima, M., Cunningham, C. H., Suzuki, Y., Willits, D. A., Willis, A. F., Yang, P. C., Tsao, P. S., McConnell, M. V., Young, M. J., and Douglas, T. (2008) A human ferritin iron oxide nano-composite magnetic resonance contrast agent, *Magnetic Resonance in Medicine 60*, 1073-1081.

56. Terashima, M., Uchida, M., Kosuge, H., Tsao, P. S., Young, M. J., Conolly, S. M., Douglas, T., and McConnell, M. V. (2011) Human ferritin cages for imaging vascular macrophages, *Biomaterials 32*, 1430-1437.

57. Flenniken, M. L., Liepold, L. O., Crowley, B. E., Willits, D. A., Young, M. J., and Douglas, T. (2005) Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture, *Chemical Communications 0*, 447-449.

58. Flenniken, M. L., Willits, D. A., Harmsen, A. L., Liepold, L. O., Harmsen, A. G., Young, M. J., and Douglas, T. (2006) Melanoma and Lymphocyte Cell-Specific Targeting Incorporated into a Heat Shock Protein Cage Architecture, *Chemistry & Biology 13*, 161-170.

59. Patterson, D. P., Desai, A. M., Holl, M. M. B., and Marsh, E. N. G. (2011) Evaluation of a symmetry-based strategy for assembling protein complexes, *RSC Advances 1*, 1004-1012.

60. Ni, T. W., and Tezcan, F. A. (2010) Structural Characterization of a Microperoxidase Inside a Metal-Directed Protein Cage, *Angewandte Chemie International Edition 49*, 7014-7018.

61. Prevelige, P. E. (2005) Bacteriophage P22, In *The Bacteriophages* (Calendar, R., Ed.) 2 ed., pp 457-468, Oxford University Press, New York.

62. Zinder, N. D., and Lederberg, J. (1952) Genetic exchange in Salmonella, *Journal of Bacteriology 64*, 679-699.

63. Byl, C. V., and Kropinski, A. M. (2000) Sequence of the Genome of SalmonellaBacteriophage P22, *Journal of Bacteriology 182*, 6472-6481.

64. Wu, H., Sampson, L., Parr, R., and Casjens, S. (2002) The DNA site utilized by bacteriophage P22 for initiation of DNA packaging, *Molecular Microbiology 45*, 1631-1646.

65. Roy, A., and Cingolani, G. (2012) Structure of P22 Headful Packaging Nuclease, *Journal of Biological Chemistry 287*, 28196-28205.

66. Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E., and Chiu, W. (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions, *Nature Structural & Molecular Biology 10*, 131-135.

67. Parent, K. N., Khayat, R., Tu, L. H., Suhanovsky, M. M., Cortines, J. R., Teschke, C. M., Johnson, J. E., and Baker, T. S. (2010) P22 Coat Protein

Structures Reveal a Novel Mechanism for Capsid Maturation: Stability without Auxiliary Proteins or Chemical Crosslinks, *Structure 18*, 390-401.

68. Moore, S. D., and Prevelige Jr, P. E. (2002) Bacteriophage P22 portal vertex formation in vivo, *Journal of Molecular Biology 315*, 975-994.

69. Galisteo, M. L., and King, J. (1993) Conformational transformations in the protein lattice of phage P22 procapsids, *Biophysical Journal 65*, 227-235.

70. Greene, B., and King, J. (1994) Binding of Scaffolding Subunits within the P22 Procapsid Lattice, *Virology 205*, 188-197.

71. Teschke, C. M., McGough, A., and Thuman-Commike, P. A. (2003) Penton Release from P22 Heat-Expanded Capsids Suggests Importance of Stabilizing Penton-Hexon Interactions during Capsid Maturation, *Biophysical Journal 84*, 2585-2592.

72. Wyckoff, E., and Casjens, S. (1985) Autoregulation of the bacteriophage P22 scaffolding protein gene, *Journal of Virology 53*, 192-197.

73. Earnshaw, W., and King, J. (1978) Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein, *Journal of Molecular Biology 126*, 721-747.

74. Parker, M. H., Stafford Iii, W. F., and Prevelige Jr, P. E. (1997) Bacteriophage P22 scaffolding protein forms oligomers in solution, *Journal of Molecular Biology 268*, 655-665.

75. Parker, M. H., Brouillette, C. G., and Prevelige, P. E. (2001) Kinetic and Calorimetric Evidence for Two Distinct Scaffolding Protein Binding Populations within the Bacteriophage P22 Procapsid, *Biochemistry 40*, 8962-8970.

76. Chen, D.-H., Baker, M. L., Hryc, C. F., DiMaio, F., Jakana, J., Wu, W., Dougherty, M., Haase-Pettingell, C., Schmid, M. F., Jiang, W., Baker, D., King, J. A., and Chiu, W. (2011) Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus, *Proceedings of the National Academy of Sciences*.

77. Parker, M. H., Prevelige, P. E., Jablonsky, M., Casens, S., Sampson, L., and Krishna, N. R. (1997) Cloning, purification, and preliminary characterization by circular dichroism and NMR of a carboxyl-terminal domain of the bacteriophage P22 scaffolding protein, *Protein Science 6*, 1583-1586.

78. Sun, Y., Parker, M. H., Weigele, P., Casjens, S., Prevelige Jr, P. E., and Krishna, N. R. (2000) Structure of the coat protein-binding domain of the scaffolding protein from a double-stranded DNA virus, *Journal of Molecular Biology 297*, 1195-1202.

79. Parker, M. H., and Prevelige Jr, P. E. (1998) Electrostatic Interactions Drive Scaffolding/Coat Protein Binding and Procapsid Maturation in Bacteriophage P22, *Virology 250*, 337-349.

80. Parker, M. H., Casjens, S., and Prevelige Jr, P. E. (1998) Functional domains of bacteriophage P22 scaffolding protein, *Journal of Molecular Biology 281*, 69-79.

81. Tuma, R., Parker, M. H., Weigele, P., Sampson, L., Sun, Y., Krishna, N. R., Casjens, S., Thomas Jr, G. J., and Prevelige Jr, P. E. (1998) A helical coat protein recognition domain of the bacteriophage P22 scaffolding protein, *Journal of Molecular Biology 281*, 81-94.

82. Casjens, S., Wyckoff, E., Hayden, M., Sampson, L., Eppler, K., Randall, S., Moreno, E. T., and Serwer, P. (1992) Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA, *Journal of Molecular Biology 224*, 1055-1074.

83. Moore, S. D., and Prevelige, J., Peter E. (2002) A P22 Scaffold Protein Mutation Increases the Robustness of Head Assembly in the Presence of Excess Portal Protein, *Journal of Virology 76*, 10245-10255.

84. Moore, S. D., and Prevelige, P. E. (2001) Structural Transformations Accompanying the Assembly of Bacteriophage P22 Portal Protein Rings in Vitro, *Journal of Biological Chemistry 276*, 6779-6788.

85. Rodriguez-Casado, A., Moore, S. D., Prevelige, P. E., and Thomas, G. J. (2001) Structure of Bacteriophage P22 Portal Protein in Relation to Assembly: Investigation by Raman Spectroscopy, *Biochemistry 40*, 13583-13591.

86. Herzog, A. B., Pandey, A. K., Reyes-Gastelum, D., Gerba, C. P., Rose, J. B., and Hashsham, S. A. (2012) Evaluation of Sample Recovery Efficiency for Bacteriophage P22 on Fomites, *Applied and Environmental Microbiology 78*, 7915-7922.

87. Reichhardt, C., Uchida, M., O'Neil, A., Li, R., Prevelige, P. E., and Douglas, T. (2011) Templated assembly of organic-inorganic materials using the core shell structure of the P22 bacteriophage, *Chemical Communications 47*, 6326-6328.

88. Carbonell, X., and Villaverde, A. (1996) Peptide display on functional tailspike protein of bacteriophage P22, *Gene 176*, 225-229.

89. Carbonell, X., and Villaverde, A. (1998) Insertional Mutagenesis in the Tailspike Protein of Bacteriophage P22, *Biochemical and Biophysical Research Communications 244*, 428-433.

90. Kang, S., Uchida, M., O'Neil, A., Li, R., Prevelige, P. E., and Douglas, T. (2010) Implementation of P22 Viral Capsids as Nanoplatforms, *Biomacromolecules 11*, 2804-2809.

91. Kang, S., Oitrogge, L. M., BroomeIi, C. C., LiepoId, L. O., Prevelige, P. E., Young, M., and DougIas, T. (2008) Controlled Assembly of Bifunctional Chimeric Protein Cages and Composition Analysis Using Noncovalent Mass Spectrometry, *Journal of the American Chemical Society 130*, 16527-16529.

92. Suzuki, M., Abe, M., Ueno, T., Abe, S., Goto, T., Toda, Y., Akita, T., Yamadae, Y., and Watanabe, Y. (2009) Preparation and catalytic reaction of Au/Pd bimetallic nanoparticles in Apo-ferritin, *Chemical Communications*, 4871-4873.

93. Jolley, C., Klem, M., Harrington, R., Parise, J., and Douglas, T. (2011) Structure and photoelectrochemistry of a virus capsid-TiO2 nanocomposite, *Nanoscale 3*, 1004-1007.

94. 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., and Rome, L. H. (2005) Engineering of vault nanocapsules with enzymatic and fluorescent properties, *Proceedings of the National Academy of Sciences, USA 102*, 4348-4352.

95. Chen, Q., Schönherr, H., and Vancso, G. J. (2010) Encapsulation and Release of Molecular Cargos via Temperature-Induced Vesicle-To-Micelle Transitions, *Small 6*, 2762-2768.

96. Prevelige Jr, P. E., Thomas, D., and King, J. (1988) Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro, *Journal of Molecular Biology 202*, 743-757.

97. Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins, *Nature Methods 2*, 905-909.

98. Stephanopoulos, N., Tong, G. J., Hsiao, S. C., and Francis, M. B. (2010) Dual-Surface Modified Virus Capsids for Targeted Delivery of Photodynamic Agents to Cancer Cells, *ACS Nano 4*, 6014-6020.

99. Douglas, T., and Young, M. (1998) Host-guest encapsulation of materials by assembled virus protein cages, *Nature 393*, 152-155.

100. Wallrabe, H., and Periasamy, A. (2005) Imaging protein molecules using FRET and FLIM microscopy, *Current Opinion in Biotechnology 16*, 19-27.

101. Miller, R. A., Stephanopoulos, N., McFarland, J. M., Rosko, A. S., Geissler, P. L., and Francis, M. B. (2010) Impact of Assembly State on the Defect Tolerance of TMV-Based Light Harvesting Arrays, *Journal of the American Chemical Society 132*, 6068-6074.

102. O'Neil, A., Reichhardt, C., Johnson, B., Prevelige, P. E., and Douglas, T. (2011) Genetically Programmed In Vivo Packaging of Protein Cargo and Its Controlled Release from Bacteriophage P22, *Angewandte Chemie International Edition 50*, 7425-7428.

103. Albertazzi, L., Arosio, D., Marchetti, L., Ricci, F., and Beltram, F. (2009) Quantitative FRET Analysis With the E0GFP-mCherry Fluorescent Protein Pair, *Photochemistry and Photobiology 85*, 287-297.

104. Morillo, S. G., and Timenetsky, M. d. C. S. T. (2011) Norovirus: An Overview, *Rev. Assoc. Med. Bras. 57*, 462-467.

105. Dumas, D. P., Caldwell, S. R., Wild, J. R., and Raushel, F. M. (1989) Purification and properties of the phosphotriesterase from Pseudomonas diminuta, *Journal of Biological Chemistry 264*, 19659-19665.

106. Dumas, D. P., Durst, H. D., Landis, W. G., Raushel, F. M., and Wild, J. R. (1990) Inactivation of organophosphorus nerve agents by the phosphotriesterase from Pseudomonas diminuta, *Archives of Biochemistry and Biophysics 277*, 155-159.

107. Ghanem, E., and Raushel, F. M. (2005) Detoxification of organophosphate nerve agents by bacterial phosphotriesterase, *Toxicology and Applied Pharmacology 207*, 459-470.

108. Roodveldt, C., and Tawfik, D. S. (2005) Directed evolution of phosphotriesterase from Pseudomonas diminuta for heterologous expression in Escherichia coli results in stabilization of the metal-free state, *Protein Engineering Design and Selection 18*, 51-58.

109. Caldwell, S. R., and Raushel, F. M. (1991) Detoxification of organophosphate pesticides using an immobilized phosphotriesterase from Pseudomonas diminuta, *Biotechnology and Bioengineering 37*, 103-109.

110. Allen, M. A., Prissel, M. M., Young, M. J., and Douglas, T. (2008) Constrained Metal Oxide Mineralization: Lessons from Ferritin Applied to other Protein Cage Architectures, In *Handbook of Biomineralization*, pp 312-333, Wiley-VCH Verlag GmbH.

111. Klem, M. T., Young, M., and Douglas, T. (2008) Biomimetic synthesis of β-TiO2 inside a viral capsid, *Journal of Materials Chemistry 18*, 3821-3823.

112. Jolley, C. C., Uchida, M., Reichhardt, C., Harrington, R., Kang, S., Klem, M. T., Parise, J. B., and Douglas, T. (2010) Size and Crystallinity in Protein-Templated Inorganic Nanoparticles, *Chemistry of Materials 22*, 4612-4618.

113. Okuda, M., Suzumoto, Y., Iwahori, K., Kang, S., Uchida, M., Douglas, T., and Yamashita, I. (2010) Bio-templated CdSe nanoparticle synthesis in a cage shaped protein, Listeria-Dps, and their two dimensional ordered array self-assembly, *Chemical Communications 46*, 8797-8799.

114. Yamashita, I., Iwahori, K., and Kumagai, S. (2010) Ferritin in the field of nanodevices, *Biochimica et Biophysica Acta (BBA) - General Subjects 1800*, 846-857.

115. Uchida, M., Kang, S., Reichhardt, C., Harlen, K., and Douglas, T. (2010) The ferritin superfamily: Supramolecular templates for materials synthesis, *Biochimica et Biophysica Acta (BBA) - General Subjects 1800*, 834-845.

116. Douglas, T., and Ripoll, D. R. (1998) Calculated electrostatic gradients in recombinant human H-chain ferritin, *Protein Science 7*, 1083-1091.

117. Kang, S., Lander, G. C., Johnson, J. E., and Prevelige, P. E. (2008) Development of Bacteriophage P22 as a Platform for Molecular Display: Genetic and Chemical Modifications of the Procapsid Exterior Surface, *ChemBioChem 9*, 514-518.

118. Shen, L., Bao, N., Prevelige, P. E., and Gupta, A. (2010) Fabrication of Ordered Nanostructures of Sulfide Nanocrystal Assemblies over Self-Assembled Genetically Engineered P22 Coat Protein, *Journal of the American Chemical Society 132*, 17354-17357.

119. Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, V. B. V., King, J., Prevelige Jr, P. E., and Chiu, W. (1996) Three-dimensional Structure of Scaffolding-containing Phage P22 Procapsids by Electron Cryo-microscopy, *Journal of Molecular Biology 260*, 85-98.

120. Seker, U. O. S., and Demir, H. V. (2011) Material Binding Peptides for Nanotechnology, *Molecules 16*, 1426-1451.

121. Van Regenmortel, M. H. V. (2001) Antigenicity and Immunogenicity of Synthetic Peptides, *Biologicals 29*, 209-213.

122. Shadidi, M., and Sioud, M. (2003) Selective targeting of cancer cells using synthetic peptides, *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy 6*, 363-371.

123. Ganz, T. (2003) Defensins: antimicrobial peptides of innate immunity, *Nature Reviews Immunology 3*, 710-720.

124. Terlau, H., and Olivera, B. M. (2004) Conus Venoms: A Rich Source of Novel Ion Channel-Targeted Peptides, *Physiological Reviews 84*, 41-68.

125. Miljanich, G. P. (2004) Ziconotide: Neuronal Calcium Channel Blocker for Treating Severe Chronic Pain, *Current Medicinal Chemistry 11*, 3029-3040.

126. Brooks, H., Lebleu, B., and Vivès, E. (2005) Tat peptide-mediated cellular delivery: back to basics, *Advanced Drug Delivery Reviews 57*, 559-577.

127. Yang, Z.-Q., Puffer, E. B., Pontrello, J. K., and Kiessling, L. L. (2002) Synthesis of a multivalent display of a CD22-binding trisaccharide, *Carbohydrate Research 337*, 1605-1613.

128. Dabrowski, M. J., Chen, J. P., Shi, H., Chin, W.-C., and Atkins, W. M. (1998) Strategies for protein-based nanofabrication: Ni2+-NTA as a chemical mask to control biologically imposed symmetry, *Chemistry & Biology 5*, 689-697.

129. Suci, P. A., Kang, S., Young, M., and Douglas, T. (2009) A Streptavidin-Protein Cage Janus Particle for Polarized Targeting and Modular Functionalization, *Journal of the American Chemical Society 131*, 9164-9165.

130. Gillitzer, E., Suci, P., Young, M., and Douglas, T. (2006) Controlled Ligand Display on a Symmetrical Protein-Cage Architecture Through Mixed Assembly, *Small 2*, 962-966.

131. Li, F., Chen, Y., Chen, H., He, W., Zhang, Z.-P., Zhang, X.-E., and Wang, Q. (2011) Monofunctionalization of Protein Nanocages, *Journal of the American Chemical Society 133*, 20040-20043.

132. Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., and Bradbury, A. R. M. (2007) Plasmid incompatibility: more compatible than previously thought?, *Protein Engineering Design and Selection 20*, 309-313.

133. Strable, E., Prasuhn, D. E., Udit, A. K., Brown, S., Link, A. J., Ngo, J. T., Lander, G., Quispe, J., Potter, C. S., Carragher, B., Tirrell, D. A., and Finn, M. G. (2008) Unnatural Amino Acid Incorporation into Virus-Like Particles, *Bioconjugate Chemistry 19*, 866-875.

134. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proceedings of the National Academy of Sciences 99*, 19-24.

135. Lee, S. K., Chou, H. H., Pfleger, B. F., Newman, J. D., Yoshikuni, Y., and Keasling, J. D. (2007) Directed Evolution of AraC for Improved Compatibility of Arabinose- and Lactose-Inducible Promoters, *Applied and Environmental Microbiology 73*, 5711-5715.

136. Huang, Y., Chiang, C.-Y., Lee, S. K., Gao, Y., Hu, E. L., Yoreo, J. D., and Belcher, A. M. (2005) Programmable Assembly of Nanoarchitectures Using Genetically Engineered Viruses, *Nano Letters 5*, 1429-1434.

137. Hong, V., Presolski, S. I., Ma, C., and Finn, M. G. (2009) Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation, *Angewandte Chemie International Edition 48*, 9879-9883.

# APPENDICES

# APPENDIX A

## SUPPORTING INFORMATION FOR CHAPTER 2

## Experimental Section

### Plasmid Design

The background plasmid was pET11a (Novagen) with the WT P22 scaffold protein and coat protein genes inserted between the NdeI and HindIII sites. A BamH1 restriction site was introduced between the scaffold protein and coat genes using PCR based mutagenesis. This plasmid is known as the assembler plasmid and was used as the background plasmid for all subsequent experiments.[1]

To create the fusion assembler plasmid, standard PCR was used to amplify the gene for the truncated scaffold protein (AA 141-303) and add the 5' NdeI and 3' BamHI recognition sequences along with the new multiple cloning site (MCS) and a thrombin recognition site (product organization: 5'-**NdeI**-SacI-NcoI-*thrombin site*-SP141 gene-STOP-**BamHI**-3'). This product was inserted to replace the full length scaffold protein in the aforementioned assembler plasmid. Next, the gene for either EGFP or mCherry was cloned with the stop codons omitted and a 5' SacI and 3' **NcoI** recognition site was added using standard PCR methods. This product was ligated in to the engineered MCS upstream of the SP141 gene resulting in a genetic fusion. The ligated plasmid was transformed into XL-2Blue ultra competent cells (Agilent Technologies, Santa Clara, CA) and colonies were screened for insert. The fusion plasmid sequence was confirmed by single-pass sequencing (SeqWright, Houston, TX). The confirmed sequence fusion assembler plasmid was subsequently transformed into BL21 *E. coli* (EMD Chemicals) for protein expression. All plasmid maps were generated using pDraw32 (acaclone).

Primers Used:

| | |
|---|---|
| SP141-FWD | 5’ggaatt**catatg**gagctcccatgg*ctggtgccgcgcggcagc*tgtcgcagcaatgccgt3’ |
| SP141-RVS | 5’ aa**ggatcc**ttatcggattccttaaagttttgcctttagcttgcgg 3’ |
| EGFP-FWD | 5’ aagagctcaaggggtgaaggaagtaatgaagatc 3’ |
| EGFP-RVS | 5’ cag**ccatgg**cttgttcttctcaggcagcagaga 3’ |
| mCherry-FWD | 5’ aagagctcatggtgagcaagggcga 3’ |
| mCherry-RVS | 5’ caga**ccatgg**cttgtacagctcgtccat 3’ |

Protein Purification

Transformed BL21 (DE3) *E. coli* were grown to an $OD_{600nm}$ = 0.6 and then induced with 1mM IPTG. The culture was allowed to grow an additional 4 hours at 37°C with vigorous shaking. The bacteria were separated from the media by centrifugation at 3700 x g for 15 minutes. The cell pellets were resuspended in 50mM phosphate buffer (100mM NaCl, pH 7.2) and frozen at -20°C overnight. The next day, the cell pellets were thawed and incubated with DNAse ( 20 mg/ml), RNAse (30 mg/ml), and lysozyme (15 mg/ml) (Sigma-Aldrich) for 30 minutes at room temperature. Cells were lysed further by sonication (40% duty cycle, 5 minutes, three times with a 2 minute cool down between) on ice. The cell debris was removed via centrifugation at 12,000 x g for 45 minutes. The supernatant (25 mL) was floated on top of a 35% (w/v) sucrose cushion (5 mL) and centrifuged at 237,404 x g for 50 minutes in an ultra centrifuge (Sorval). The resulting virus pellet was resuspended in 50mM phosphate buffer (25mM NaCl pH 7.6) and spun at 17,000 x g for 20 minutes to remove any particulates and lipid. The virus was then dialyzed against the aforementioned buffer overnight.

Cryo-Electron Microscopy

Cryo-EM was done by standard methods[2] 3 μL of sample was applied to non-glow-discharged C-flat holey film (Electron Microscopy Sciences, Hatfield, PA), blotted briefly before plunging into liquid ethane, and transferred to a Gatan 626 cryo-sample holder. All samples were observed in an FEI Tecnai F20 electron microscope operated at 200 kV, and the images were captured on a Gatan UltraScan 4000 CCD camera at a magnification of 38,000.

Size Exclusion Chromatography -
Multi Angle Light Scattering

P22 capsid samples were separated by size exclusion chromatography column WTC-0200S (Wyatt Technologies). Samples were injected and the mobile phase was applied using an Agilent 1200 HPLC at a flow rate of 0.7 mL/min. The mobile phase was 50mM phosphate buffer (100mM NaCl, 200ppm sodium azide, pH 7.2). Sample (25μL) was injected on the column and run for 30 minutes. Samples were monitored using both a UV-vis detector (Agilent) and a refractive index detector (Wyatt) to determine concentration. Multi angle light scattering was detected with a Dawn 8 (Wyatt). Average Mw and Rg were determined from a fit to the data using a Zimm plot in the Astra software from Wyatt.

Dynamic Light Scattering

Dynamic light scattering was measured at 90º with a 661nm diode laser (Brookhaven 90Plus particle size analyzer) and the correlation function was fit using a

non negatively constrained least squares analysis to yield a hydrodynamic radius (*29*) for the particles.

Size Exclusion Chromatography

Samples were centrifuged at 17,000 rpm for 10 minutes and the supernatant was loaded on to a preparative S-500 Sephadex (GE Healthcare) 80 mL column using an AKTA Pharmacia FPLC. The flow rate was 1 mL/minute using 50mM phosphate buffer (25mM NaCl pH 7.6) and the elution time was 150 mL. Protein (280 nm) and the chromophore (490 nm for EGFP and 585 nm for mCherry) wavelengths were monitored simultaneously.

Agarose Gel

Samples were mixed with loading dye (40% glycerol, bromophenol blue) and loaded into the wells of a 0.8% agarose slab gel. Gels were run in 40mM tris 5mM acetate 1mM EDTA pH 8.2 at 65 V for three hours at room temperature. Gels were stained with coomassie blue and destained using acetic acid and methanol. Gels were imaged using a UVP MultiDoc-It Digital Imaging System.

Mass Spectrometry

MS analyses were performed on a ESI-Q-TOF mass spectrometer (Q-TOF Premier, Waters). Purified P22 capsids (0.1-2 $\mu$L, 1.0 mg/mL) were injected onto a BioBasic SEC-300 (Thermo Scientific) column and eluted with 40% IPA, 0.1% formic acid. Deconvoluted spectra were generated with the software MaxEnt1 provided by Waters.

In-gel trypsin digestions were performed as described by ref. 3. Digested protein samples were analyzed by LC/MS/MS (liquid chromatography mass spectrometry) methods for protein identification. Following protein digestion, samples were injected onto a reverse phase C18 column for separation of the injected peptides. Eluted LC peaks underwent MS/MS detection on a Qtof (or ion trap) mass spectrometer. A Mascot search of the MS/MS data was performed using the NCBI protein database and this returned high confidence protein identifications.

References Cited

1. P. E. Prevelige, D. Thomas, J. King, *Journal of Molecular Biology* **1988**, *202*, 743.

2. T. Dokland and M.L. Ng, Transmission electron microscopy of biological specimens. In: T. Dokland, D.W. Hutmacher, M.L. Ng and J.T. Schantz, Editors, Techniques in Microscopy for Biomedical Applications, World Scientific Press, Singapore (2006).

3. Shevchenko, A., et al. *Nature Protocols*, 1:6, pg 2856 (2006).

Supplementary Figures

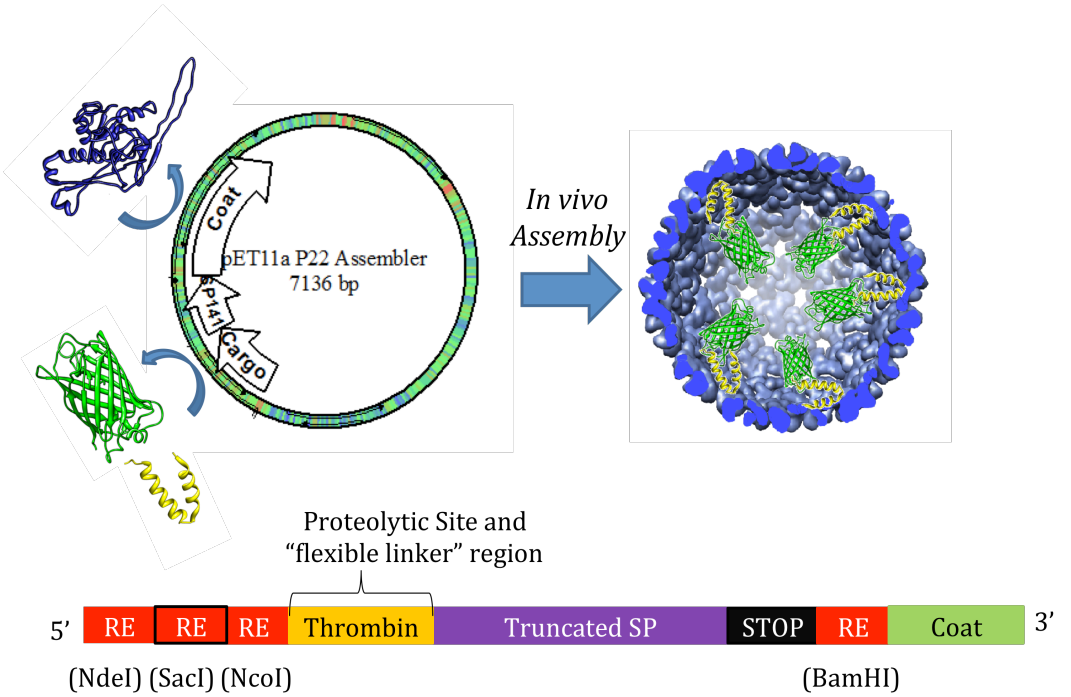



Figure A.1: Plasmid map illustrating the two gene products (coat protein and scaffold protein fusion) that lead to the spontaneous *in vivo* assembly of cargo packaged P22 procapsids. The organization of the 5' engineered multiple cloning sites (NdeI, SacI, and NcoI restriction enzymes (RE) sites), the flexible linker (and thrombin cleavage site) to the truncated scaffold protein and the P22 coat protein is shown schematically in a linearized form.

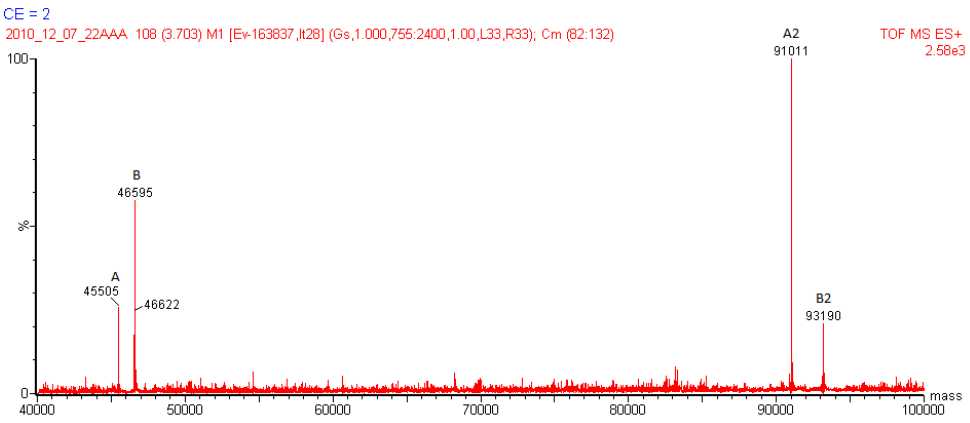
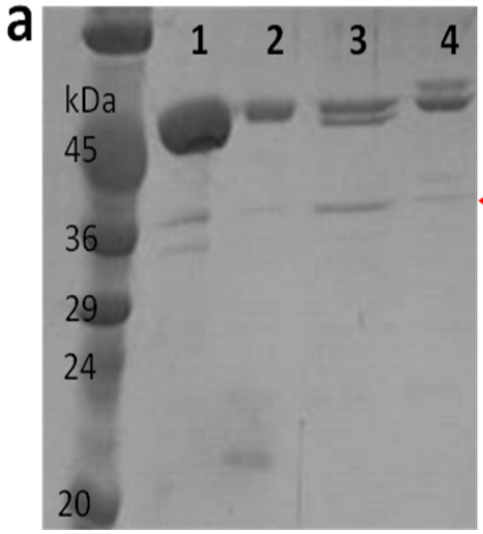



Figure A.2: Max ent mass spectrometry spectra (see methods for details) of the purified P22 GFP-SP141 capsids. Peak A refers to the EGFP-SP141 fusion and A2 its dimer (calc. mass = 45515 Da) while peak B and B2 refer to the P22 coat protein and coat dimer respectively (calc. mass = 46596, no Met1).




| Lane | Sample |
|---|---|
| 1 | WT P22 |
| 2 | P22 $SP_{141}$ |
| 3 | P22 EGFP- $SP_{141}$ |
| 4 | P22 mCherry- $SP_{141}$ |

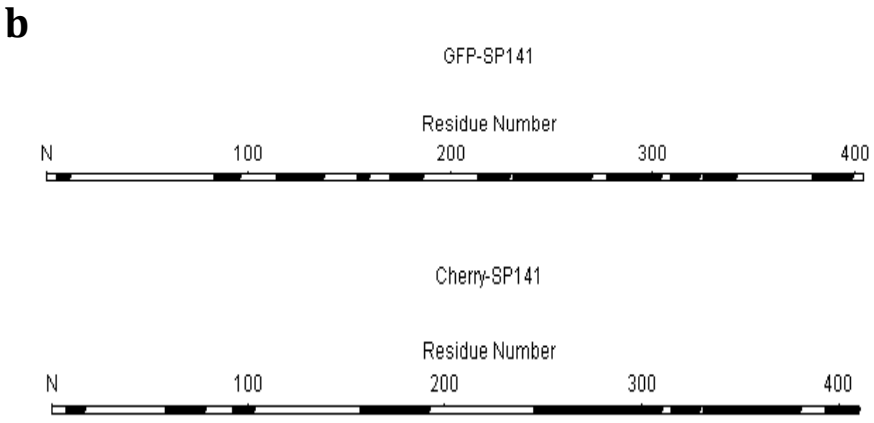
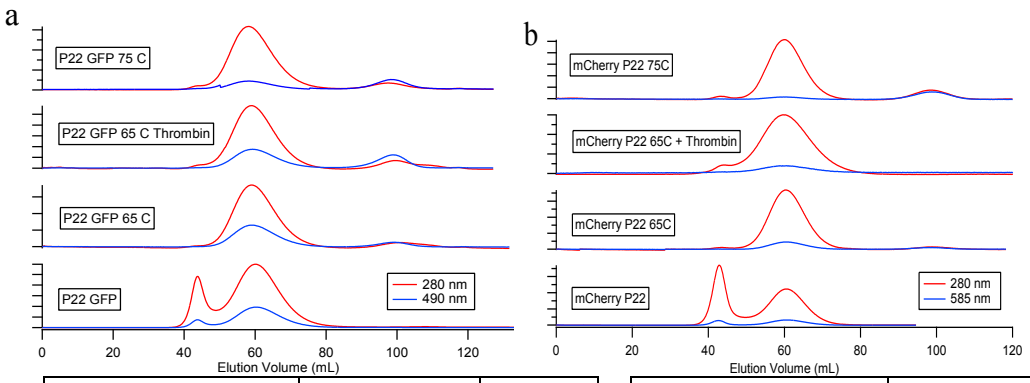



Figure A.3: a) SDS-PAGE (12%) denaturing gel of the purified cargo packaged P22 procapsids. Wild type (WT) P22 coat protein runs at approximately 46 kDa (calc. mass 46.6 kDa) across all samples. Full length scaffold protein (SP) runs at approximately 35 kDa (slightly higher than its calculated mass of 33.5 kDa) in the WT P22 lane. A co-purifying *E. coli* contaminant (identified as OmpF by in-gel trypsin digestion and subsequent mass spectrometry) at 37 kDa is seen across all P22 forms. The truncated SP (AA 141-303) runs at approximately 21.5 kDa and can only be seen in the SP141 alone P22. The SP mass is greatly increased in both fusions and in the case of the GFP-SP141 fusion runs directly beneath the coat protein while the mCherry-SP141 runs directly above the coat protein. b) Both fusion bands identity have been confirmed by in-gel trypsin digestion and MS/MS and the peptide coverage maps are shown. Black bars

indicate a region where peptides were identified whereas white bars had no experimental findings.




| Sample | A490:A280 | % Loss |
|---|---|---|
| P22 GFP | 0.32 | 0 |
| + 65°C | 0.35 | 0 |
| + 65°C + Thrombin | 0.30 | 6.25 |
| + 75°C | 0.14 | 56.25 |

| Sample | A585:A280 | % Loss |
|---|---|---|
| P22 mCherry | 0.14 | 0 |
| + 65°C | 0.12 | 14.29 |
| + 65°C + Thrombin | 0.13 | 7.14 |
| + 75°C | 0.03 | 78.57 |

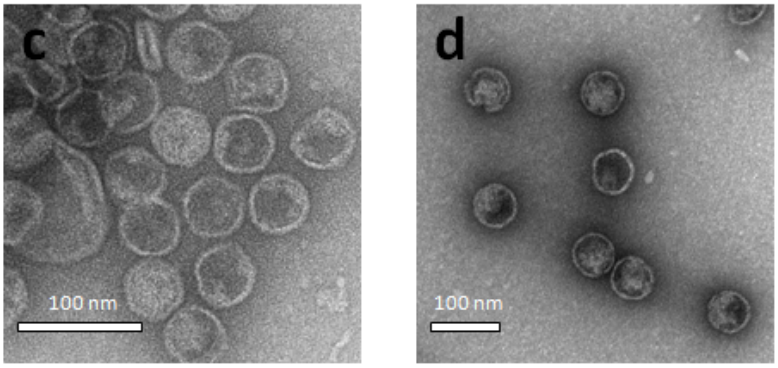



Figure A.4: Sephacryl-500 (GE Healthcare) size exclusion chromatographs of fusion-cargo packaged P22 capsids under various conditions; (a) EGFP-SP141 P22, (b) mCherry-SP141 P22. The characteristic absorbance for EGFP is 490nm while mCherry absorbs at 585nm. The bottom chromatogram shows the characteristic soluble aggregate peak at 43 mL (which was discarded) and the soluble capsid fraction at 60 mL. The capsid fraction was collected and pooled and used for all further experiments. This fraction was then heated to the indicated temperature and treated with excess thrombin overnight where denoted. The heating process causes a small amount of aggregation, which is easily removed from the soluble fraction by brief centrifugation at 17,000 g. The small amount of capsid breakdown results in the appearance of the fusion protein at 100 mL. The amount of fusion increases at 75°C due to its ability to leave the capsid through the large holes in the capsid in the WB morphology lacking pentamer units. Additionally, the visible absorbance is decreased in these samples further indicating the ability of the fusion cargo to diffuse out of the capsid. This is quantified in the tables below each set of chromatograms. The ratio between the protein absorbtion (A280nm) and the fluorescent fusion are listed along with the percentage of fluorescent absorbtion lost for each sample. (c) TEM micrgraph of mCherry-SP141 P22 heated to 65°C or (d) 75°C showing further that intact capsids are present after heating.

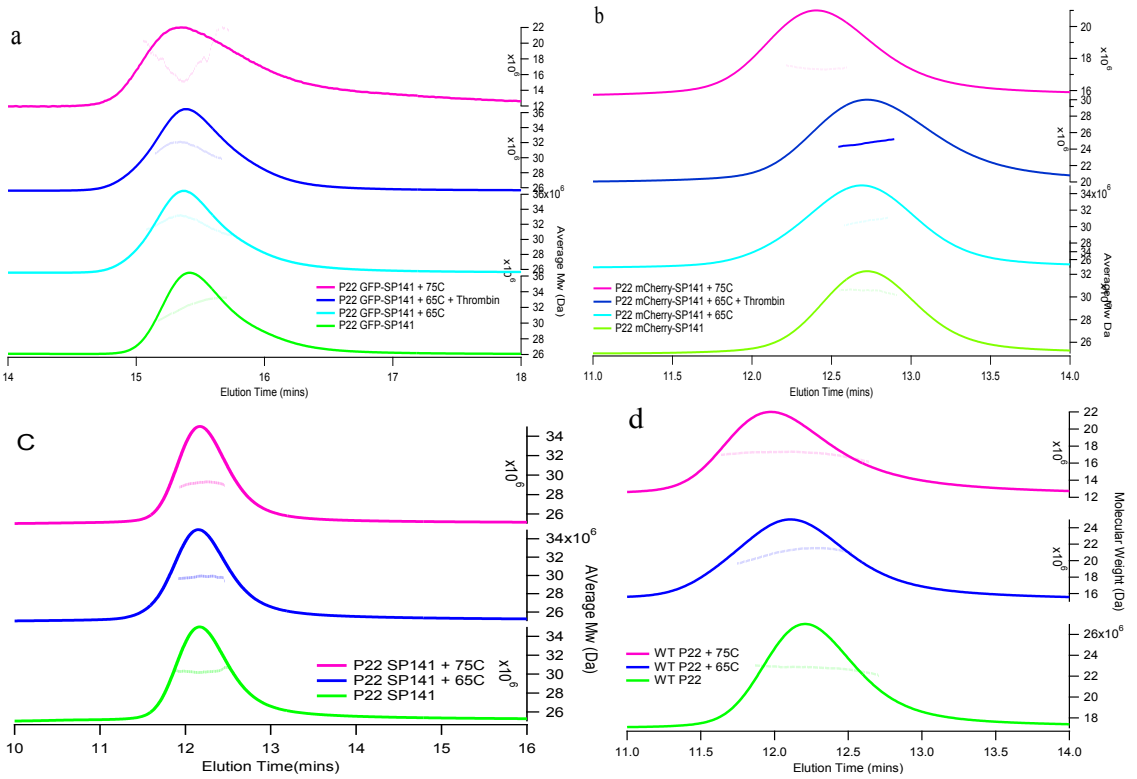



Figure A.5: HPLC-MALS data of the (a) EGFP-SP141 fusion and (b) mCherry-SP141 fusion capsids, (c) SP141 only capsids, and (d) WT P22 under various conditions. Shown is the HPLC chromatogram with the molecular weight distribution across the elution peak overlayed. The elution times remain constant across conditions for each capsid while the molecular weight changes (right y-axis) depending on the treatment. (Note: Elution times shift between experiments due to column to column variation; this shift is confirmed by running a standard. All stacked chromatograms were run on the same column and same day.)

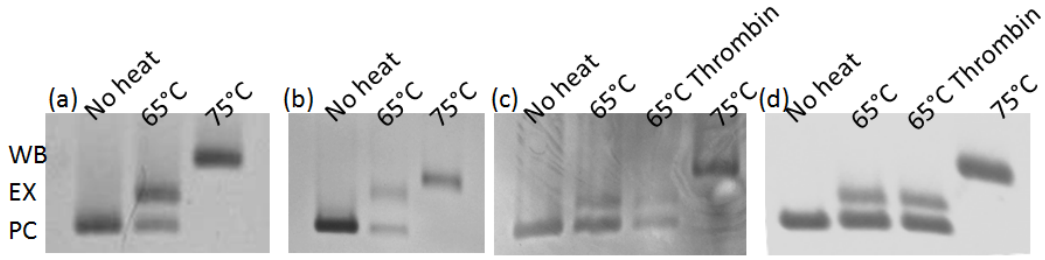
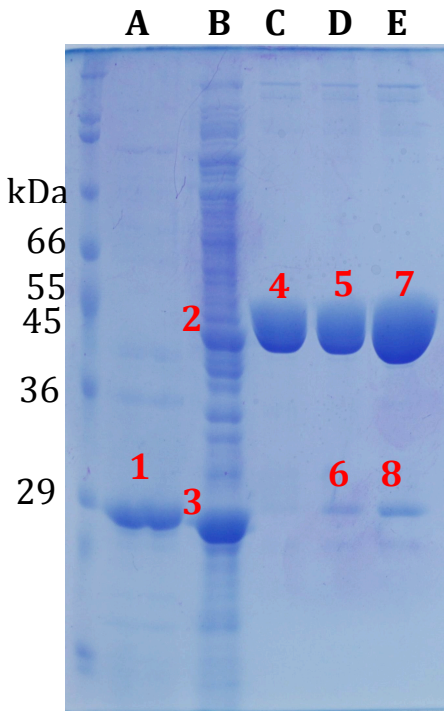



Figure A.6: Non-denaturing 0.8% agarose gels illustrating the various morphology changes available to (a) WT P22 (b) P22 SP141 (c) P22 GFP-SP141 and (d) mCherry-SP141. Samples were heated at the indicated temperatures for 20 minutes, centrifuged at 17,000 g for 5 minutes, and then the supernatant loaded onto the gel. The furthest running bands indicates procapsid (PC), the middle bands indicate an expanded shell (EX) morphology, while the highest bands represent the expanded shell with the additional loss of the pentamer units (wiffle ball, WB).




| Lane | Sample |
|---|---|
| A | EGFP-6XHis |
| B | P22 EGFP-SP141 Ultra Centrifuge Supt.* |
| C | P22 EGFP-SP141 (SEC Pure) + 65°C min |
| D | C + Thrombin (@ same concentration C) |
| E | D undiluted |

| Band | Identity | Calc. Mw (Da) |
|---|---|---|
| 1 | EGFP-6xHis | 26823 |
| 2 | $SP_{141}$-EGFP | 45515 |
| 3 | Free EGFP (auto lysis product) | unknown |
| 4 + 5 + 7 | P22 Coat + EGFP-$SP_{141}$ (over loaded and over lapping) | Coat = 46596 Fusion = 45515 |
| 6 + 8 | Cleaved EGFP (from $SP_{141}$) | 27422 |

Figure A.7: Reducing and denaturing 12% acrylamide SDS-PAGE showing the appearance of free EGFP protein after capsid expansion (via heat treatment to 65°C) and thrombin digestion. The tables on the right describe the samples in lanes A through E and the identitiy of the band 1-8. *The ultra centrifuge supernatant results from the purification method and is usually thrown away as no assembled virus is contained in it. Please see the experimental section for more details.

| Sample | Mw (MDa) | Radius ($R_g$) (nm) | Sample Size* |
|---|---|---|---|
| WT P22 | 21.7 | 25.1 | 6 |
| WT P22 65°C | 19.8 | 28.0 | 6 |
| WT P22 75°C | 18.3 | 27.4 | 6 |
| P22 $SP_{141}$ | 30.4 | 23.3 | 15 |
| P22 $SP_{141}$ 65°C | 27.6 | 24.2 | 3 |
| P22 $SP_{141}$ 75°C | 23.8 | 25.0 | 3 |
| P22 EGFP-$SP_{141}$ | 32.4 | 22.5 | 12 |
| P22 EGFP-$SP_{141}$ 65°C | 31.8 | 25.2 | 12 |
| P22 EGFP-$SP_{141}$ thrombin | 33.1 | 22.6 | 6 |
| P22 EGFP-$SP_{141}$ 65°C thrombin | 29.6 | 23.3 | 6 |
| P22 EGFP-$SP_{141}$ 75°C | 20.8 | 26.2 | 12 |
| P22 mCherry-$SP_{141}$ | 30.5 | 22.5 | 12 |
| P22 mCherry-$SP_{141}$ 65°C | 29.5 | 24.4 | 12 |
| P22 mCherry-$SP_{141}$ thrombin | 31.5 | 23.3 | 6 |
| P22 mCherry-$SP_{141}$ 65°C thrombin | 24.7 | 23.5 | 9 |
| P22 mCherry-$SP_{141}$ 75°C | 20.9 | 25.2 | 9 |

Table A.1: Summary of average Mw and capsid diameter from all SEC-HPLC-MALS data. *Sample size denotes number of individual runs averaged together to get the Mw and radius.

# APPENDIX B

## SUPPLEMENTARY INFORMATION FOR CHAPTER 3

## Percent Occupied Calculation

EGFP(PDB: 2Y0G) h= 4.9nm r= 1.3nm  v= 28.1 $nm^3$

mCherry (PDB: 2H5Q) h= 3.77nm r= 1.38nm  v= 22.56 $nm^3$

Total fusion volume = 50.66 $nm^3$ x 100 fusions = 5066 $nm^3$

P22 Procapsid = 40nm interior diameter (PDB: 2XYY), volume = 33510 $nm^3$

Assuming random packing of spheres, the total volume that can be occupied is 64% of the total volume of the capsid.  (Jaeger, H. M. and Nagel, S. R. "Physics of Granular States." *Science* **255**, 1524, 1992.)

Total available capsid volume = 21446 $nm^3$

5066 $nm^3$ / 21446 $nm^3$ = 24% occupied*

*does not take into account the volume contributed by the scaffold protein therefore the fusions take up more than 24% of the capsid volume.

Supplementary Figures

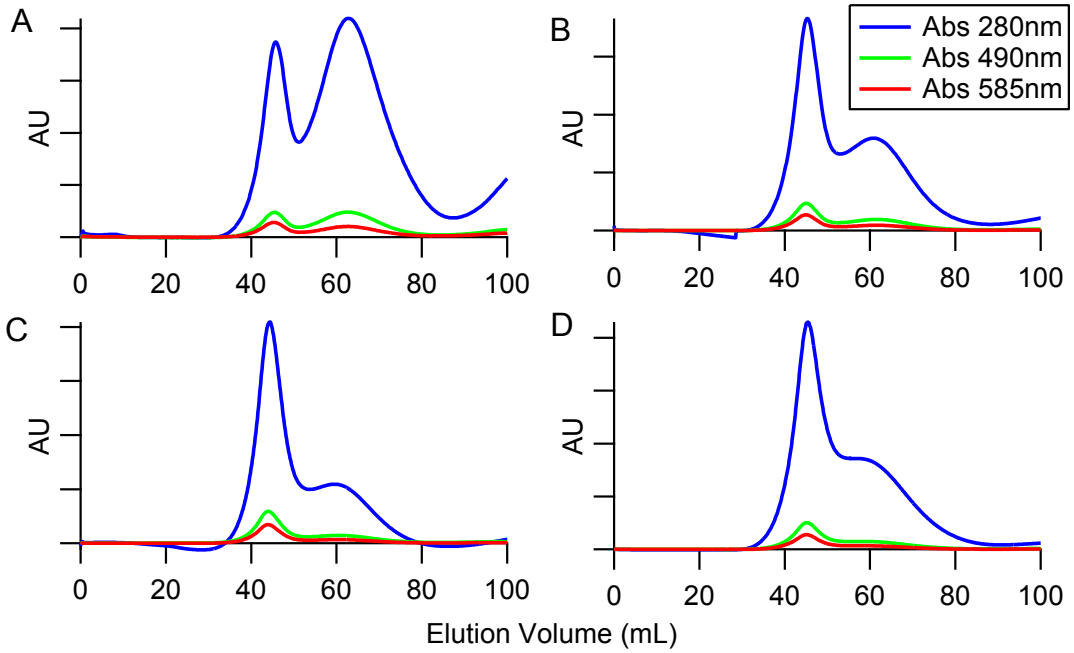



Figure B.1: P22 Concatenated samples are purified over a preparative S-500 size exclusion column. This column has the capacity to separate aggregate/mis-formed particles (40 mL elution volume) from the correctly assembled capsids (60 mL elution volume). Fractions from the later half of the 60 mL peak are used for downstream applications. Protein absorbance (280nm) is monitored simultaneously with the GFP absorbance (490nm) and mCherry absorbance (585nm) illustrating that the two fluorescent proteins co-elute with intact capsids. (A) P22 Cat 0gly (B) P22 Cat 6gly (C) P22 Cat 12gly (D) P22 Cat 18gly

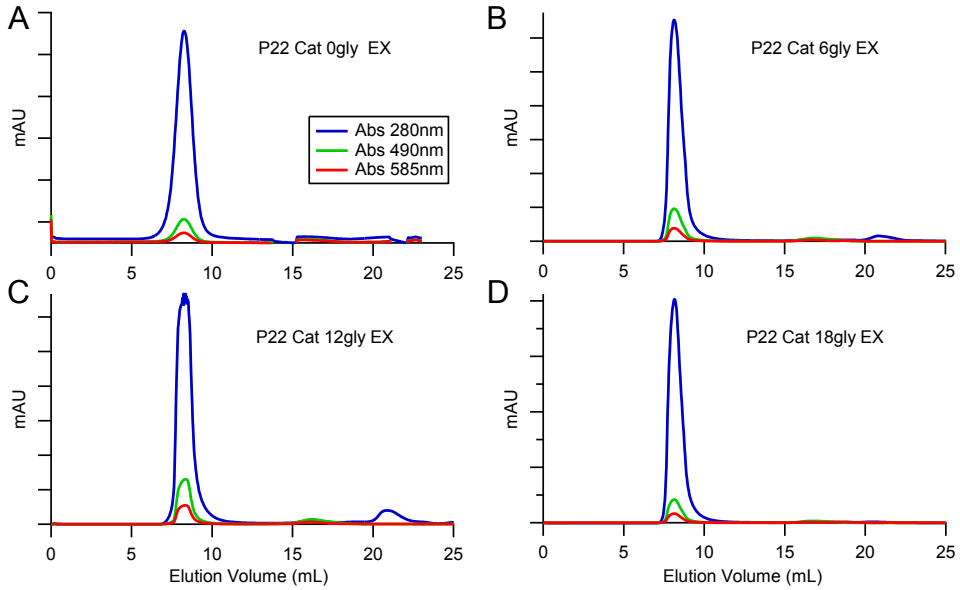



Figure B.2: Size exclusion (S500) purified capsids were further purified after heating to 65°C by Sup6 size exclusion chromatography. The capsid elutes in the void volume of ~8 mL. The GFP-mCherry fusion is still encapsulated within the capsid as illustrated by the co-elution of the GFP absorbance (490nm, green line) and the mCherry absorbance (585nm, red line) with the capsid peak (~8mL elution volume).

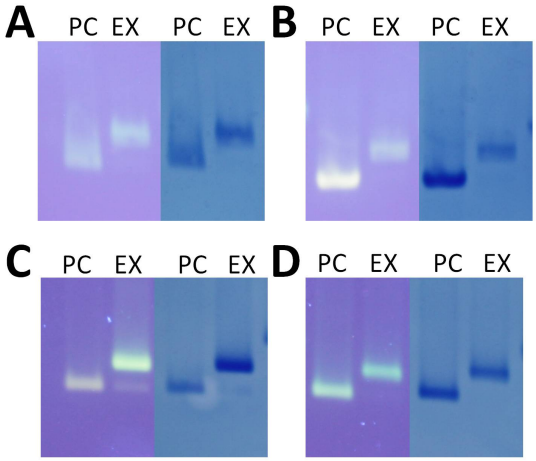



Figure B.3: Non-denaturing agarose shift gels showing the procapsid (PC) and the shift to expanded capsid (EX) in all samples. The GFP-mCherry fusion is retained in both the PC and EX capsid, in all samples, as seen by the co-migration of the fluorescence signal, seen by UV illumination (left panel), and the coomassie protein stain (right panel). (A) P22 Cat 0gly (B) P22 Cat 6gly (C) P22 Cat 12gly (D) P22 Cat 18gly

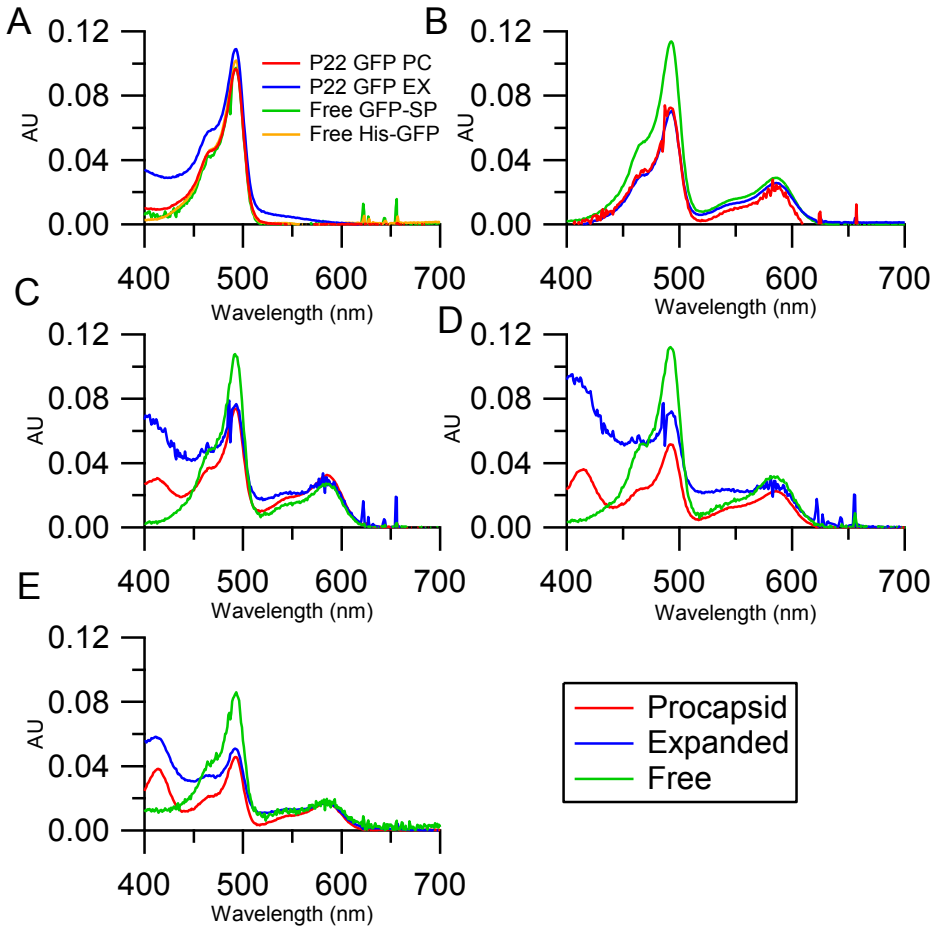



Figure B.4: Scattering corrected UV-Vis spectra of all encapsulated and free in solution fusions. Procapsid traces are in red, expanded capsid traces in blue, and free fusions are in green. (A) P22 GFP only (B) P22 Cat 0gly (C) P22 Cat 6gly, (D) P22 Cat 12gly, (E) P22 Cat 18gly.

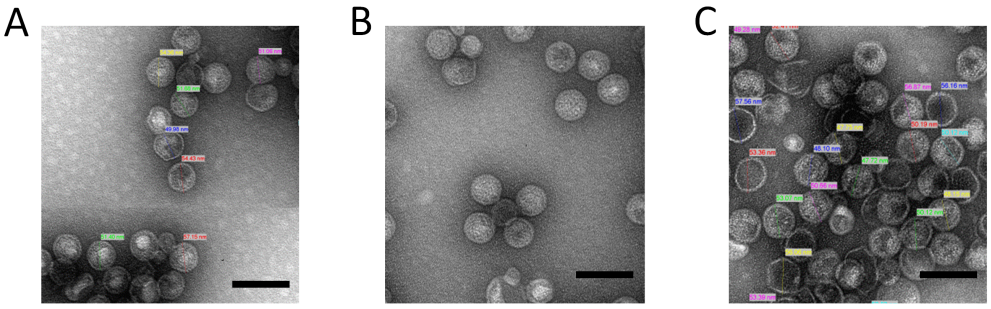



Figure B.5: Transmission electron micrograph of all the P22-Cat constructs. (A) P22 Cat6gly (B) P22 Cat12gly (C) P22 Cat18gly. Black bar represents 100nm. Colored text in panels A and C are measured values of the particles.

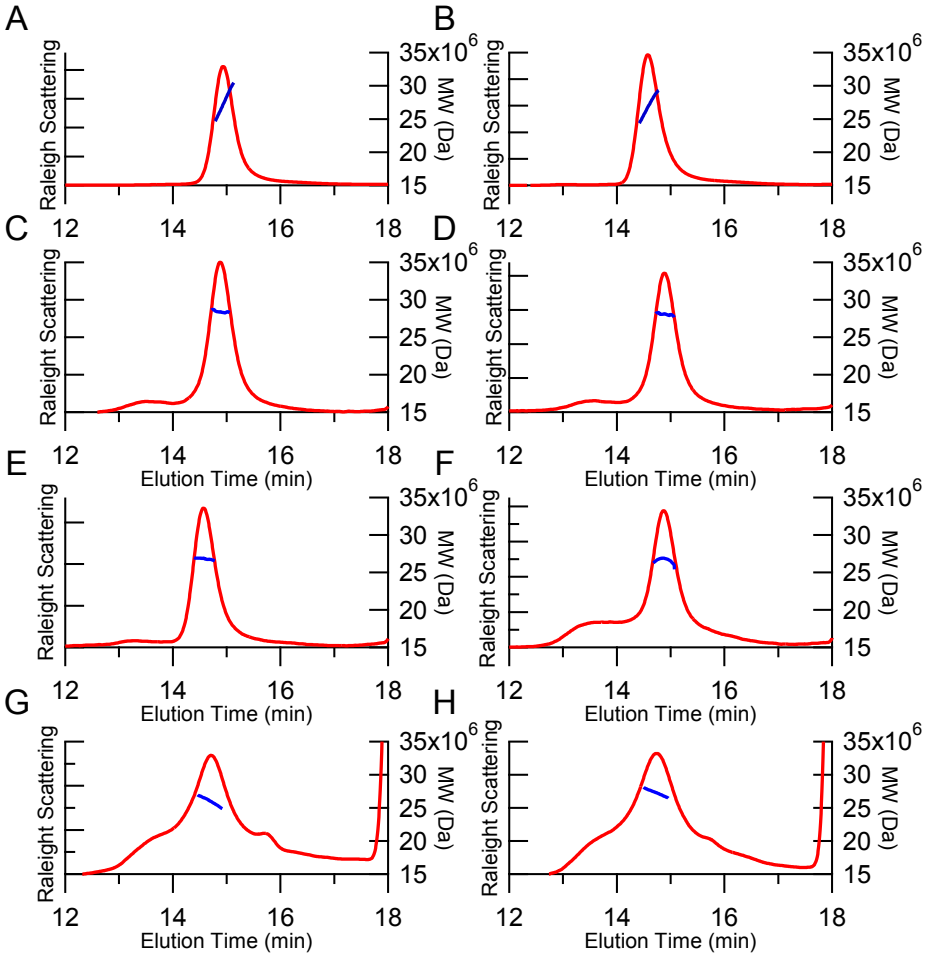



Figure B.6: High pressure liquid chromatography coupled to multi-angle light scattering data for the P22 Cat glycine linker series in the procapsid and expanded capsid morphology. Raleigh scattering versus elution time is plotted in red with the total molecular weight across the capsid peak in blue. A) P22 Cat 0gly, procapsid B) P22 Cat 0gly, expanded C) P22 Cat 6gly, procapsid D) P22 Cat 6gly, expanded E) P22 Cat 12gly, procapsid F) P22 Cat 12gly, expanded G) P22 Cat 18gly, procapsid H) P22 Cat 18gly, expanded

# APPENDIX C

## ALTERNATIVE VERSION OF MANUSCRIPT IN CHAPTER 3

## Forward

This manuscript was submitted, and rejected, from Biomacromolecules before the version in Chapter 3 was published. From the reviewer's comments, we discovered that the story we were trying to tell was too large for one paper. Hence, we made the decision to only publish one set of data, that for the glycine spaced FRET pair. In this original manuscript, an additional set of data is included for the "amino-acid-stiff" spaced FRET pair. Additionally, the data was analyzed differently in the published manuscript, where many calculations and normalizations were done to extract the FRET trends. In the original manuscript, we examined the raw data closely and tried to extrapolate the overall trends. The experimental section herein has been truncated, as most sections were identical to the experimental in chapter 3.

## Original Manuscript Version for Chapter 3

### Co-Confinement of Fluorescent Proteins: Spatially Enforced Interactions of GFP and mCherry Encapsulated Within the P22 Capsid

Authors: Alison O'Neil, Gautam Basu, Peter E. Prevelige, Trevor Douglas

### Abstract

The precise architectures of viruses and virus-like particles are proving to be highly advantageous in synthetic materials applications. Not only can these nano-containers be harnessed as active materials, but they can be exploited for examining the effects of in vivo "cell-like" crowding and confinement on the properties of the

encapsulated cargo. Here we report the first example of inter-molecular communication between two proteins encapsulated within the capsid of the bacteriophage P22. Using a genetically engineered fusion between the P22 scaffold protein with GFP and a red fluorescent protein (mCherry) in tandem as a concatamer, we were able to spontaneously assemble P22 capsids and direct the programmed encapsulation of the genetic fusion when co-expressed with P22 coat protein. These capsids are densely packaged, with more than 15% of the volume being occupied, and yet the fused GFP and mCherry retain their fluorescence and are FRET active. To examine the distance dependence of FRET inside the capsid, linkers of different lengths were engineered between the GFP and mCherry fusions. The P22 system is unique in that the capsid morphology can be altered, without losing the encapsulated cargo, resulting in a doubling of the capsid volume. Thus we have examined the encapsulated fusions at two different internal concentrations. Our results indicate that FRET is sensitive to the expansion of the capsid. The capsid, in both morphological states, still enforces significant inter-molecular communication and the dependence on linker length is overwhelmed. This P22 co-encapsulation system is a promising platform for studying crowding, enforced proximity, and confinement effects on active proteins.

## Introduction

There is a growing interest in investigations of viruses, not as vectors of disease, but to harness their naturally evolved structures as nano-containers in synthetic materials approaches.[1] Virus capsids, and other protein cage architectures, have been exploited to encapsulate a wide range of catalytic metal centers, addressable polymers, drugs, and

active proteins such as enzymes and fluorescent proteins. [2] These versatile bio-compatible vessels have been shown to have applications in medicine and catalysis but also provide novel environments for studying the effects of molecular crowding in the confined interior of the viral capsid.[3]

The concentrations of macromolecules inside the cell can be in the range of 300 mg/mL, however biochemical reaction studies are more usually carried out at concentrations of less than 1mg/ml.[4] The high concentration of many different types of mutually volume excluding macromolecules in the cell causes it to be a crowded and confining environment in which to carry out reactions. The crowding effect is important because all forms of life have evolved complex systems for regulating cellular volume (and in-turn local macromolecular concentrations). [4] Additionally, it has been shown that small changes in the cell volume can cause large over-arching changes in intracellular processes. [5] Therefore, non-specific background interactions due to confinement and crowding are unavoidable and to gain a true understanding of cellular and molecular processes, these background interactions must be taken into account and understood. Biochemists have realized the importance of understanding the effects of macromolecular crowding and confinement on biological reactions and these insights could provide new impetus for new materials design and implementation. [6]

To probe the effects of crowding and confinement on the interaction and communication between two macromolecules, we report the use of the bacteriophage P22 capsid as a container for the co-encapsulation of two fluorescent proteins. We have developed a system by which a genetically fused pair of fluorescent proteins, that act as

Förster resonance energy transfer (FRET) partners, can be co-encapsulated within the P22 bacteriophage capsid in high copy number and high local concentration (Figure 1a). By comparing the FRET of the packaged materials to the proteins free in solution, we have probed the molecular interactions occurring in this engineered crowded nano-environment. To explore the effects of a constrained environment on the distance dependence of the FRET interaction of the encapsulated cargo more closely, the fluorescent protein fusions were engineered with several different lengths of tethers between the fluorescent proteins.

The P22 capsid, derived from the *Salmonella typhimurium* bacteriophage, is a virus like particle (VLP) composed of 420 copies of a 46.6 kDa coat protein that assembles into a T=7 icosahedral capsid with the aid of approximately 100-330 copies of a 33.6 kDa scaffold protein. [7] Heterologous expression, in *E. coli*, of P22 coat protein together with scaffold protein results in self-assembly of the compact procapsid structure, which has an external diameter of 58 nm. The scaffold proteins are incorporated on the interior of the capsid through electrostatic interactions localized at the C-teminus of the scaffold protein. [8] The P22 VLP undergoes a series of characterized structural transitions upon heating, which yields large changes in internal volume and external accessibility to the hollow interior. [9] Gentle heating at 65º C for 20 minutes causes irreversible expansion of the procapsid structure yielding a 64 nm P22 structure referred to as the expanded form together with loss of the scaffold protein. [9] The resulting change in diameter from procapsid to expanded morphology results in a near doubling of the effective volume of the internal cavity (~58,000 $nm^3$ to ~113,000 $nm^3$). From structural studies both the

procapsid and expanded form are seen to have small 2 nm pores in the VLP shell between the exterior and interior. [10]

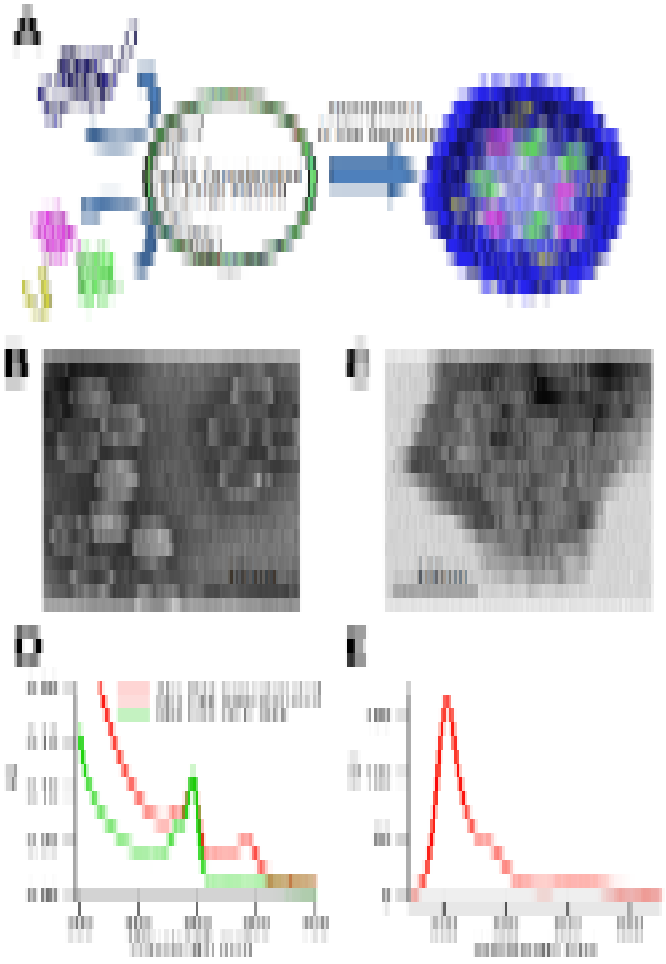


Figure C.1: The P22 scaffold protein- mCherry-GFP catenated fusion retains its function while accurately templating VLP assembly. (A) Cartoon illustrating the genetic layout of the catenated fusion and P22 coat protein and their spontaneous *in vivo* assembly into P22 VLPs. (B) TEM of WT P22 and (C) the P22 encapsulating the catenated fusion (P22 Cat). (D) UV-Vis of a P22 GFP only sample versus the P22 Cat which show absorbtion at both GFP (490nm) and mCherry (585nm) wavelengths. (E) Emission spectra of P22Cat, Ex 450nm, showing GFP emission (507nm) and the mCherry FRET signal at 610nm.

Previous studies have shown that the scaffold protein of P22 interacts with the coat protein through the last few residues of the C-terminal region and the full length

scaffold protein can be highly modified at the N-terminus while still maintaining the ability to guide the assembly of coat protein into the native-like P22 capsid structure. [8, 11] These modifications can include the deletion of a large number of amino acids from the N-terminus and/or fusion of another gene product to the truncated scaffold protein. Truncated scaffold proteins have been exploited as engines for the programmed *in vivo* incorporation of proteins and peptides to the interior of the P22 VLP. [12, 13] We have shown that cargo proteins, specifically fluorescent proteins, expressed as scaffold protein fusions are incorporated into the interior of P22 VLPs when expressed together with coat protein. Using the scaffold protein fusion strategy, fluorescent proteins were efficiently incorporated with up to 300 copies per VLP, corresponding to a local concentration of fluorescent protein of 8.6 mM within the capsid.[13, 14] Packing of the three protein fusion system described here corresponds to >15% occupancy of the capsid interior (see supplemental information). In addition, incorporation of fluorescent protein-scaffold protein fusion to the interior did not alter the structural transitions of the VLP from procapsid to the expanded form. [13] Unlike the few VLP encapsulation systems described to date, the availability of the structural transition by the P22 VLP makes it an intriguing system because the volume can be systematically modulated for investigating crowding and confinement effects on a specific sample of encapsulated cargo through the procapsid to expanded morphological transformation.

FRET is a well established method for studying protein-protein interactions *in vivo* and *in vitro*. [15] Excitation of the fluorescent donor can result in energy transfer if a suitable acceptor molecule is within a well-defined distance (50% energy transfer at

FRET radius $R_0$), and if there is significant overlap between the donor fluorescence and acceptor absorption spectra. FRET results in reduction in donor fluorescence and if the acceptor is also fluorescent, there is concomitant appearance of acceptor fluorescence. This phenomenon has a strong distance dependence ($1/r^6$), which makes it an extremely sensitive tool for measuring the effects of spatial confinement of the FRET pair interactions.[16] In the present work we have used the enhanced green fluorescent protein (EGFP) as the donor and mCherry protein as the fluorescent acceptor.

Experimental

Fluorescence Emission Scans**.** All samples were brought to the same EGFP 490nm absorbance on a UV spectrometer. 200ul of these samples were loaded into a quartz cuvette with a 1cm pathlength into a Cary Eclipse (Varian) Fluorescence Spectrometer. Samples were excited at 450nm, with a 5nm excitation slit width, and a fluorescence emission scan from 450nm to 700nm was taken. Fluorescence measurements were taken at 90° with a 5nm emission slit width. The PMT detector voltage was set to 600v and the scan rate set to 600nm/min.

Emission Scan Correction and FRET Efficiency Calculation. To correct the emission spectra for small differences in absorbtion, the following calculations were made. First, the absorbtion spectra was corrected for scattering by graphing the log of the absorbtion wavelength versus the log of the absorbtion for each sample. A line ($y = mx + b$) was fit to this graph in the region where there is no absorbtion, only scattering (i.e. at the higher wavelengths, 700-800nm) to yield a slope and intercept. The absorbtion

spectra where then corrected as: corrected abs = abs – 10^(slope * log(wavelength) + intercept)

Next, the emission spectra were corrected using a factor created from the absorbtion of the control (P22 GFP procapsid for all procapsid samples, P22 GFP expanded for all expanded samples, and free GFP-scaffold protein for all the free in solution samples) and the corrected absorbtion of the samples at the GFP peak (493nm). The equation used was:

corrected em = em * (abs of control / corrected abs)

To calculate FRET efficiency, the following equation was used:

E = 1 – (corrected emission/emission of control) As the emission control, P22 GFP procapsid was used for all procapsid samples, P22 GFP expanded was used for all expanded samples, and free GFP-scaffold protein was used for all the free in solution samples at the GFP emission peak (501nm).

Results and Discussion

Here we describe a robust self-assembling protein cage system for the programmed encapsulation of a triple protein fusion, consisting of the truncated scaffold protein, GFP, and mCherry catenated together and packaged within the P22 capsid (P22Cat). Building on previous work demonstrating the fusion of truncated scaffold protein and mCherry[13], we incorporated the additional fusion of GFP by cloning in the gene using engineered restriction sites at the 5' end of the mCherry-scaffold protein gene (Figure 1a). The resulting construct was sequenced and expressed in *E. coli* and the particles purified as previously described. [13] When characterized by transmission electron

microscopy (TEM), the P22 capsids were homogenous and morphologically indistinguishable from the capsids with either full length scaffold protein or capsids incorporating a single scaffold protein-fluorescent protein fusion.[13] By TEM, particles were 47± 4 nm in diameter (Figure 1b,c) and measurement of the particles by multiangle light scattering revealed diameters of 45.4 nm with an average total particle Mw of 27.54±0.16 MDa (Table 1). This corresponds to approximately 111 scaffold protein-mCherry-GFP fusions encapsulated within the P22 (Table 1). The characteristic absorption (495nm for GFP, 585nm for mCherry[17]) and emission spectra (507nm for GFP, 610nm for mCherry[17]) are seen for both GFP and mCherry in purified virus samples (Figure 1d,e) illustrating that the two fluorescent proteins are functional in the properly assembled capsid. The emission spectrum of the construct P22Cat, obtained by exciting only GFP (450 nm) exhibits mCherry fluorescence (610 nm) in addition to GFP fluorescence (507 nm), indicating that P22Cat is FRET-active. (Figure 1e)

To probe the distance dependence of FRET inside our viral cages, we created a series of variable length spacers between the donor GFP and acceptor mCherry proteins within the fusion. To do this, a segment of DNA encoding 6 glycine residues was ligated into our vector system in tandem, between the two proteins. The glycine linkers are as follows: 6Gly: (EL)$(G)_6$(EL); 12Gly: (EL)$(G)_6$(EL)$(G)_6$(EL); 18Gly: (EL)$(G)_6$(EL)$(G)_6$(EL)$(G)_6$(EL). The glutamic acid and leucine (EL) residues are encoded by the inherent restriction enzyme site of the DNA insert. However, an artifact of our cloning method is that the DNA can ligate in forwards, encoding the 6 glycine amino acids (as shown above), or in reverse encoding for serine-threonine-proline-serine-

alanine-threonine (STPSAT). In screening for tandem linkers of all glycines, we found linkers with hybrids of glycine segments and the STPSAT. Instead of disregarding these "backwards" linkers, we decided to investigate how they affect interactions in the FRET pair, especially because this linker has a proline in the center of it, creating a kink instead of a fully extended glycine linker. To differentiate between the two linker types, we named this backwards, kinked, spacer simply the amino acid spacer (aa). Thus, P22 Cat6aa has a linker of (EL)(STPSAT)(EL), P22 Cat12aa has a linker of $(EL)(G)_6(EL)(STPSAT)(EL)$, and P22 Cat18aa has a linker of $(EL)(G)_6(EL)(G)_6(EL)(STPSAT)(EL)$.

The catenated fusions with variable length spacers spontaneously assembled into capsids indistinguishable from P22Cat by TEM. (Figure S1) Analysis by TEM, DLS, and MALS all yielded similar diameters to P22Cat as well and showed the same relative amount of packaging as the no spacer fusion capsids (103-123 fusions per capsid, Table 1). In addition, the expanded capsid morphology was accessible in these spacer fusion capsids and non-denaturing agarose gels of these materials were identical to those of the no spacer fusion capsids (Figure S2). Regardless of the linker length or type, both the characteristic GFP and mCherry absorbances at 495nm and 585nm, respectively, were seen illustrating the proper folding of the fluorescent proteins in all spacer fusion capsids (Figure S3). As expected, FRET communication was also observed in all of the spacer fusion encapsulated samples (Figure C.2).

Table **C1.** All P22 encapsulated systems were measured for homogeneity, size, and total particle molecular weight using high-pressure liquid chromatography coupled to multi-angle light scattering (MALS), dynamic light scattering (DLS), and transmission electron microscopy (TEM). All samples were run in triplicate and reported values are averages of the three runs.

| Sample | Fusion Mw (Da) | $R_g$ (nm, MALS) | Diameter (nm, DLS) | Diameter (nm, TEM) | Total Mw (MDa) | # cargo |
|---|---|---|---|---|---|---|
| P22 GFP PC | 45515 | 22.5 | 50±4 | 45 ±1 | 32.4 | 282 |
| P22 mCherry PC | 45994 | 22.5 | 49±5 | 49±1 | 30.5 | 238 |
| P22 Cat PC | 71582 | 22.7±0 | 60.7 ±13 | 47±4 | 27.54±0.16 | 111 |
| P22 Cat EX | | 25.0 | | | 26.87±0.12 | |
| P22 Cat6aa PC | 72369 | 22.4±0.1 | 58.1±7 | 53±3 | 27.23±0.71 | 106 |
| P22 Cat12aa PC | 72954 | 22.9±0.1 | 57.8±7 | 55±3 | 27.84±1.44 | 113 |
| P22 Cat18aa PC | 73538 | 22.8±0.1 | 64.5±8 | | 27.03±0.89 | 101 |
| P22 Cat6Gly PC | 72167 | 22.3±0.07 | 69.8 | 44±4 | 28.42±0.01 | 123 |
| P22 Cat6Gly EX | | 24.0±1.4 | | | 27.34±0.62 | |
| P22 Cat12Gly PC | 72751 | 23.4±1.2 | 70.5 | 53±2 | 27.03±0.23 | 103 |
| P22 Cat12Gly EX | | 24.4±1.4 | | | 25.65±0.94 | |
| P22 Cat18Gly PC | 73336 | 24.9±0.12 | 69.5 | 60±4 | 26.42±0.74 | 93 |
| P22 Cat18Gly EX | | 24.8±0.1 | | | 26.4±0.77 | |

When the P22 capsid is heated to 65ºC it undergoes a morphological change to the expanded form that breaks the interactions between the coat proteins and scaffold proteins. The P22 capsids with the catenated cargo can still access the expanded morphology after heating to 65°C for 20 minutes at pH 7.0 as seen in the non-denaturing agarose shift gels (Figure S2). These gels also show that in the expanded form, the capsid still retains the fluorescent cargo, as seen by the co-migration of fluorescence and coomassie protein stain and also the by the co-elution of GFP and mCherry absorbance with the capsid in the expanded form when analyzed by size exclusion chromatography (Figure S2, S4). In the unmodified wild type system, the conformational change in the capsid and release of scaffold allows it to diffuse out of the capsid entirely. However in

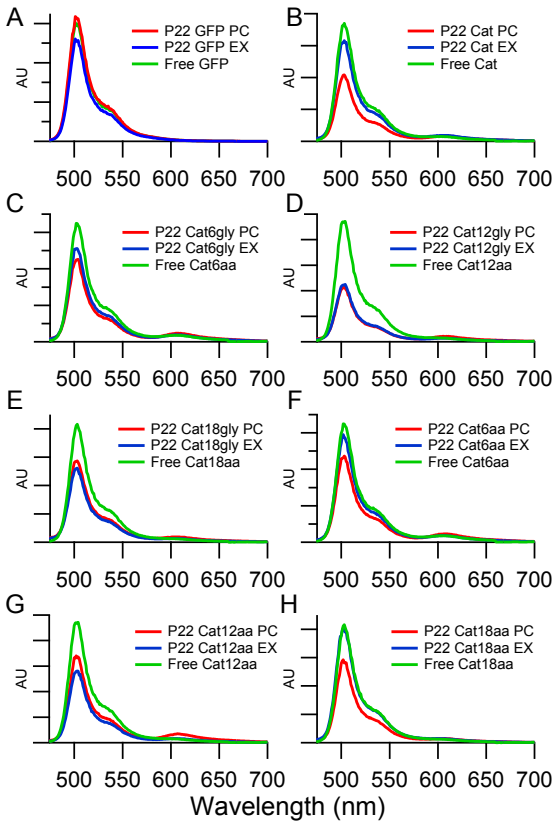



Figure C.2: All P22 concatenated fusions show characteristic GFP emission at 507nm with a FRET signal from mCherry appearing at 610nm. Each panel compares the fluorescent fusion in the free state to the encapsulated state in the procapsid (PC) or the expanded (EX) capsid morphology. Shown are the absorbance corrected emission spectra, Ex 450nm, showing that the catenated procapsids and expanded capsids are all quenched in their GFP signal compared to their free in solution counterparts. (A) P22 GFP only (B) P22 Cat (C) P22 Cat6gly, (D) P22 Cat12gly, (E) P22 Cat18gly, (F) P22 Cat6aa, (G) P22 Cat12aa, (H) P22 Cat18aa

our scaffold-fusion system, because of the size and ordered structure of the fusion (either GFP, mCherry, or the catenated GFP-mCherry) it cannot diffuse out of the expanded capsid shell. Thus, because the scaffold protein is no longer anchored to the coat, the catenated fusion proteins have substantially more rotational, and translational, freedom in the expanded form of the capsid. Also, the internal volume of the expanded form is roughly double that of the procapsid form (and so the concentration of the encapsulated

fusion is therefore halved) suggesting even more internal flexibility for the fusion proteins in this form of the P22.

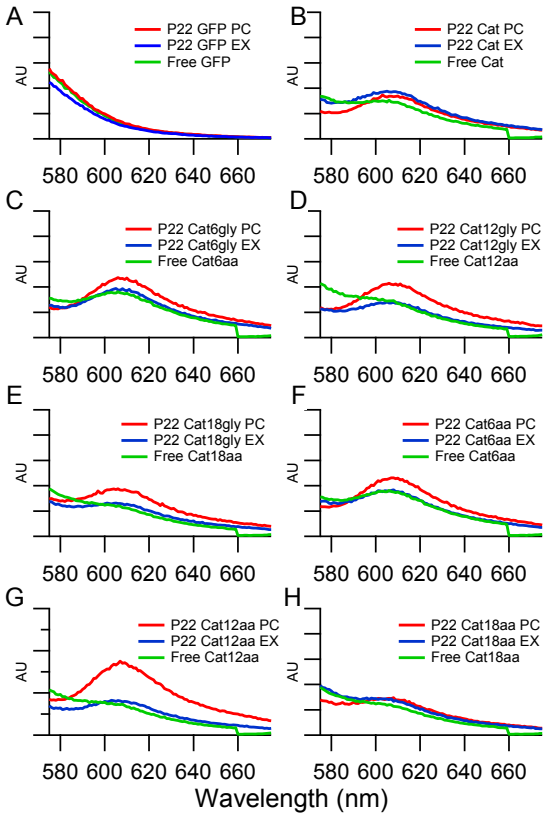



Figure C.3: The trend of the FRET signal (mCherry 610nm) in the P22 Concatenated fusions shows the most FRET in the free samples, with less FRET in the procapsid form, and even less FRET in the expanded (EX) capsid form. Emission scan blow-up, Ex 450nm, of the FRET signal region of P22 GFP PC, EX, and free showing no FRET signal as expected (A). Emission scan blow-up, Ex 450nm, of the FRET signal region of P22 Cat (B), P22 Cat6gly (C), P22 Cat12gly (D), P22 Cat18gly (E), P22 Cat 6aa (F), P22 Cat12aa (G), and P22 Cat18aa (H).

A comparison of FRET-induced mCherry fluorescence (excitation at 450nm) shows that the overall trend in all samples is towards more mCherry fluorescence (and therefore more FRET) in the procapsid form compared to the expanded and free in solution form, with the expanded and free form having similar mCherry emission (Figure

3). The only exceptions to this are the no spacer sample and the 18aa spacer sample where all three forms in both samples show essentially the same amount of FRET emission. The lack of change in the no spacer sample between the three forms may be due to the covalent contact between the two proteins and therefore the FRET signal is constant in all environments whereas differences in the GFP quenching may be caused by crowding conditions and inter-molecular interactions (Figure 2). The 18aa spacer, with its long kinked spacer, may fill enough of the volume of the capsid to maintain the same amount of inter-molecular FRET regardless of capsid form. The FRET signal in the expanded capsids, regardless of linker type or length, are similar to their counterparts free in solution. From this, it appears that as we go to the larger capsid volume, the FRET induced mCherry fluorescence, and therefore the fusion's environment, is more "bulk-like".

Interestingly, across all procapsids, regardless of linker length or type, the level of FRET-induced mCherry fluorescence signal is relatively the constant (Figure 3). This correlates well with the low level of change in the encapsulated FRET efficiencies calculated from decrease in GFP fluorescence (Figure S5). All procapsids have about the same amount of fluorescent proteins within them (93-123 per capsid) so this observation indicates that the FRET signal may be related to the packaging density of the particles. Another noteworthy observation is that the FRET signal is affected by the volume of the capsid interior whereas the level of GFP quenching seems to have little dependence on volume (Figure 3 and 2). FRET is therefore a more sensitive measurement of the effects of confinement within a virus like particle than GFP quenching alone.

When comparing all of the procapsid samples together, the 12aa sample shows the most mCherry fluorescence detected FRET however it does not show the most GFP quenching (12gly shows the most GFP quenching) (Figure 3 and 2). The FRET-induced mCherry fluorescence is expected to correlate with GFP emission only if the quantum yield of mCherry is identical in molecules being compared. The observed disconnect, between the FRET and the GFP quenching, may indicate that other pathways for energy transfer from the donor GFP exist beside FRET to the mCherry acceptors or that the mCherry fluorescence is affected by crowding in the capsid. A comparison of mCherry fluorescence, at the same concentration free in solution and P22 encapsulated, reveals that the encapsulated mCherry shows markedly less fluorescence than the free mCherry (Figure 4). This means that the mCherry fluorescence is also affected by confinement and crowding and that in these experiments GFP quenching may not proportionately lead to mCherry fluorescence from energy transfer.

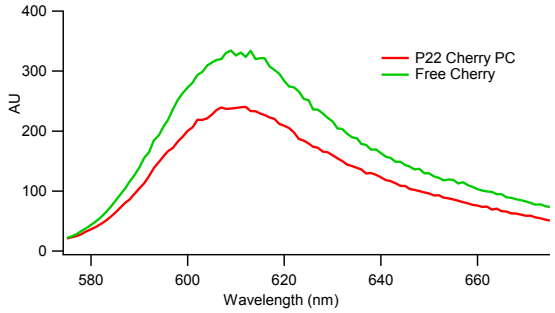



Figure C.4: mCherry shows less fluorescence when encapsulated in the P22 than when free in solution. Emission scan, Ex 550nm, of the P22 mCherry and free mCherry at the same mCherry absorbance.

FRET has a strong dependence on the distance between two fluorophores ($1/r^6$) and this can be seen in the non-encapsulated fusion FRET efficiency (Figure S5). As the

linker gets longer, the efficiency is expected to go down. However this dependence is not seen in our encapsulated system. Thus, the crowding effects experienced by the fusions in the both capsid forms, mainly inter-molecular interactions, dominate over the FRET distance dependence behavior.

Packaged fusions experience inter-molecular as well as intra-molecular interactions when packaged at high density inside the capsid. In the form having the most densely packaged material, the un-heated procapsid, quenching of the GFP was observed in all samples, regardless of linker length or linker type as compared to a P22 with only GFP encapsulated or a sample of free, un-encapsulated, GFP (Figure 2). Additionally, in the procapsid morphology the GFP emission for the encapsulated fusions in all samples (regardless of linker type or length), was quenched in comparison to their free un-encapsulated counterparts (Figure 2). This demonstrates that even though all samples have the same GFP concentration, the enforced packaging has a significant effect on the GFP fluorescence. Due to the crowded environment in the procapsid form, this effect is most likely due to inter-molecular energy transfer. If the level of quenching was due entirely to intra-molecular FRET, one would expect the level of GFP quenching of the encapsulated fusions to be similar, if not the same, as the free un-encapsulated fusions. To rule out the possibility that the observed quenching of GFP in packaged GFP-mCherry is not due to any other capsid associated non-radiative deactivation pathways, GFP fluorescence from all encapsulated catenated fusion samples were compared to the P22 encapsulated GFP-only sample. This revealed that all catenated samples were quenched by about 50% (Figure 2). Additional evidence that the capsid and crowding do

not affect the GFP fluorescence is that the absorbance corrected GFP emission is essentially the same in the procapsid form as compared to free in solution (Figure 2). The observed quenching is thus largely due to FRET between GFP and mCherry.

A comparison of the quenching of the GFP fluorescence in the expanded form compared to the procapsid from, in general, does not change (Figure 2). This is somewhat surprising in that going from the procapsid to the expanded form, the internal volume is doubled and the fusion concentration is thus halved. One might expect this to have a relieving effect on the intra-molecular interactions of the fusions. However this is not what is observed. Therefore the expanded capsid is still a very crowded environment and the enforced proximity of the fusions causes a significant amount of intra-molecular interactions leading to the quenching of GFP. When the emission from the expanded capsid fusions are compared to the free in solution counterparts, in general, there is quenching in the expanded encapsulated samples. This is additional evidence for enforced intra-molecular communication in the expanded forms. However, there are some examples (no linker, 6aa, and 18aa) where the expanded capsid samples show essentially the same GFP fluorescence as their free in solution counterparts (figure 2a). In these samples, the intra-molecular interactions appear to be relieved.

## Conclusions

We have successfully demonstrated the first example of co-encapsulation of two proteins capable of molecular communication within the confined environment of a virus-like particle. Not only does our P22 scaffold protein-mCherry-GFP fusion fold correctly and maintain its fluorescence, but the scaffold protein accurately templates the

assembly of P22 capsids around the triple catenated protein fusion. The resulting capsids are homogeneous in size with dense packaging on the order of ~100 fusions per capsid, occupying >15% of the capsid interior. Our multi-protein encapsulation system demonstrates the possibility for "hand-shake" reactions inside a viral capsid and is poised for work with multimeric functional complexes.

The fluorescent fusion pairs experience intra as well as inter-molecular FRET due to their confinement and crowding inside the capsid. The fusion fluorescence spectra show a clear trend when encapsulated versus when free in solution with the encapsulated samples showing more quenching than their free in solution partners. Additionally, the FRET signal is much higher in the procapsid encapsulated state and becomes more bulk-like in the expanded capsid. Unlike the capsid-free state, when encapsulated there is minimal dependence of FRET on the linker length or type between the two chromophores. These findings highlight the fact that the effects of local crowding can overwhelm the linker length dependence of intra-molecular FRET in bulk solution and further illustrates that confinement and crowding challenges the established picture of intra- and inter-molecular interaction in bulk solution.

Crowding and confinement have a profound effect on cellular processes when compared to the same reactions in dilute solution. Our P22 encapsulation system is a promising platform to study these effects as crowding is inherent and the confinement volume is highly definable, predictable, and most importantly, modular- as the capsid can be expanded without loss of encapsulated cargo.

Acknowledgements

The authors would like to thank Candace Goodman for many fluorescence conversations. This research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Science and Engineering DE-FG02-07ER46477.

References Cited

(1). Bode, S. A.*, et al. Nanoscale.* **2011,** *3*, 2376-2389; Destito, G.*, et al.* In *Viruses and Nanotechnology;* M. Manchester and N. Steinmetz, Eds.; Springer Berlin Heidelberg, 2009; Vol. 327, pp 95-122; Douglas, T.*, et al. Advanced Materials.* **2002,** *14*, 415-418; Huang, X.*, et al. Nano Letters.* **2007,** *7*, 2407-2416; de la Escosura, A.*, et al. Angewandte Chemie.* **2010,** *122*, 5463-5466; Fiedler, J. D.*, et al. Angewandte Chemie International Edition.* **2010,** *49*, 9648-9651; Seebeck, F. P.*, et al. Journal of the American Chemical Society.* **2006,** *128*, 4516-4517; Tissot, A. C.*, et al. PLoS ONE.* **2010,** *5*, e9809.

(2). Rhee, J.-K.*, et al. Biomacromolecules.* **2011,** *12*, 3977-3981; Lucon, J.*, et al. Chemical Communications.* **2010,** *46*, 264-266; Brasch, M.*, et al. Journal of the American Chemical Society.* **2011,** *133*, 6878-6881; Tsai, S. J. *Progress in molecular biology and translational science.* **2011,** *103*, 1-20; Huang, X.*, et al. ACS Nano.* **2011,** *5*, 4037-4045; Worsdorfer, B.*, et al. Science.* **2011,** *331*, 589-592; Chen, Q.*, et al. Small.* **2009,** *5*, 1436-1445; Comellas-Aragones, M.*, et al. Nat Nano.* **2007,** *2*, 635-639; Jolley, C.*, et al. Nanoscale.* **2010,** Advance Article; Kickhoefer, V. A.*, et al. Proceedings of the National Academy of Sciences of the United States of America.* **2005,** *102*, 4348-4352.

(3). Kitagawa, T.*, et al. Molecular Imaging and Biology.* **2011,** 1-10; Varpness, Z.*, et al. Nano Letters.* **2005,** *5*, 2306-2309; Bolis, D.*, et al. Journal of Molecular Biology.* **2004,** *336*, 203-212.

(4). Minton, A. P. *Journal of Biological Chemistry.* **2001,** *276*, 10577-10580.

(5). Lang, F.*, et al. Physiological Reviews.* **1998,** *78*, 247-306; Minton, A. P. *Journal of Cell Science.* **2006,** *119*, 2863-2869.

(6). Zhou, H.-X.*, et al. Annual Review of Biophysics.* **2008,** *37*, 375-397; Dhar, A.*, et al. Proceedings of the National Academy of Sciences.* **2010,** *107*, 17586-17591; Cheung, M.*, et al. Proceedings of the National Academy of Sciences.* **2005,** *102*, 4753-4758; Adrian H, E. *Current Opinion in Structural Biology.* **2010,** *20*, 196-206.

(7). Earnshaw, W.*, et al. Journal of Molecular Biology.* **1976,** *104*, 387-410.

(8). Parker, M. H.*, et al. Journal of Molecular Biology.* **1998,** *281*, 69-79.

(9). Teschke, C. M.*, et al. Biophysical Journal.* **2003,** *84*, 2585-2592.

(10). Prasad, B. V. V.*, et al. Journal of Molecular Biology.* **1993,** *231*, 65-74.

(11). Weigele, P. R.*, et al. Journal of Molecular Biology.* **2005,** *348*, 831-844.

(12). Reichhardt, C.*, et al. Chemical Communications.* **2011,** *47*, 6326-6328.

(13). O'Neil, A.*, et al. Angewandte Chemie International Edition.* **2011,** *50*, 7425-7428.

(14). Minten, I. J.*, et al. Chemical Science.* **2010,** *2*, 358-362.

(15). Wang, Y.*, et al. Integrative Biology.* **2009,** *1*, 565-573.

(16). Wallrabe, H.*, et al. Current Opinion in Biotechnology.* **2005,** *16*, 19-27.

(17). Shaner, N. C.*, et al. Nat Meth.* **2005,** *2*, 905-909.

# APPENDIX D

## SUPPLEMENTAL INFORMATION FOR CHAPTER 4

## Calculation of Number of Enzymes Per Capsid

P22 Coat Protein = 46596 Da x 420 coat protein per capsid = 19570320 Da
PTE-SP141 fusion = 56756 Da
P22-PTE average MW = 21840560 ± 570820 Da (*n=3*)
P22-PTE (21840560 Da) – P22 Capsid (19570320 Da) = 2270240 Da
2270240 Da / PTE-Fusion (56756 Da) = 40±10 PTE per capsid

## Supplemental Figures

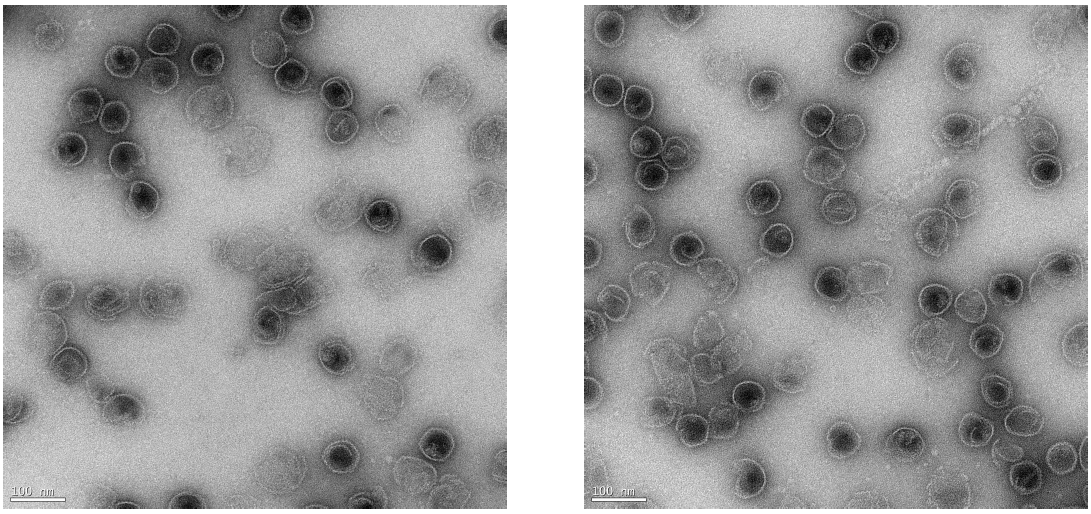



Figure D.1: Representative TEM images of the purified P22-PTE grown under optimized conditions. Particles are not perfectly spherical and some seem to have multiple layers of coat protein. Scale bar in bottom left represents 100nm.

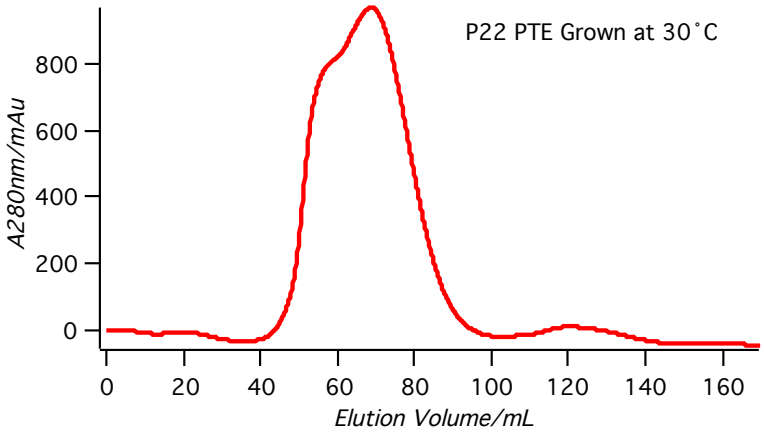



Figure D.2: The P22-PTE virus pellet was resuspended in HEPES buffer (50 mmol, pH 8.5) and subjected to size-exclusion purification. A preparative sephacryl-S500 (Amersham) was used at a flow rate of 1 mL/min using the aforementioned buffer. Fractions from the later half of the 65 mL elution peak were kept as the purified P22-PTE sample.

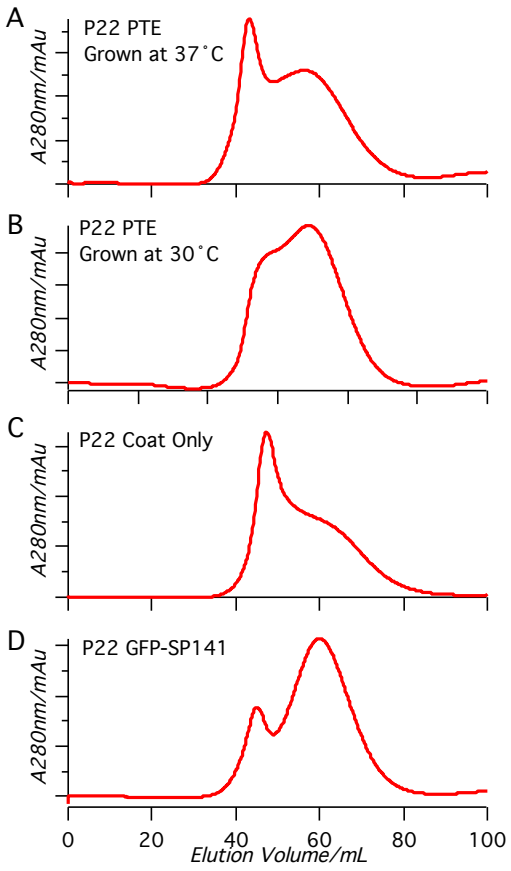



Figure D.3: Sephacryl-S500 size exclusion chromatograms illustrating the differing ratio of mis-assembled P22 capsids (~45mL) and properly assembled P22 capsids (60mL) depending on the growth conditions and abundance of P22 scaffold protein. (A) P22-PTE grown at 37°C shows a higher ratio of mis-assembled capsids then the same P22-PTE grown at 30°C (B). The chromatographic profile of the mostly mis-assembled P22-PTE (A, grown at 37°C) looks very similar to the profile of (C) where the P22 coat protein is expressed and purified without scaffold protein present. This suggests that the abundance of scaffold protein in the expression, greatly affects the capsid assembly outcome. (D) An example of a SEC chromatogram resulting from a P22 growth with a very well folded and expressed scaffold protein fusion (GFP-SP141). In this well expressing construct, the two peaks are more resolved and the properly assembled peak (60mL) is the majority of the sample.

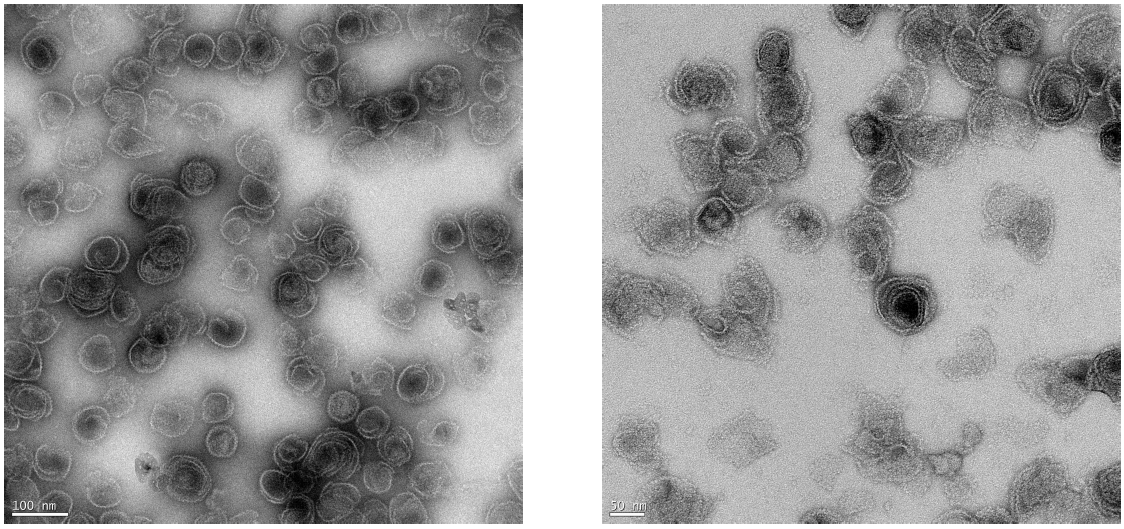



Figure D.4: Representative TEM image of the P22-PTE grown at 37°C illustrating the poor assembly of capsids. Capsids are vaired in size and completeness with many capsids having multiple shells (see center of right panal for excellent example).

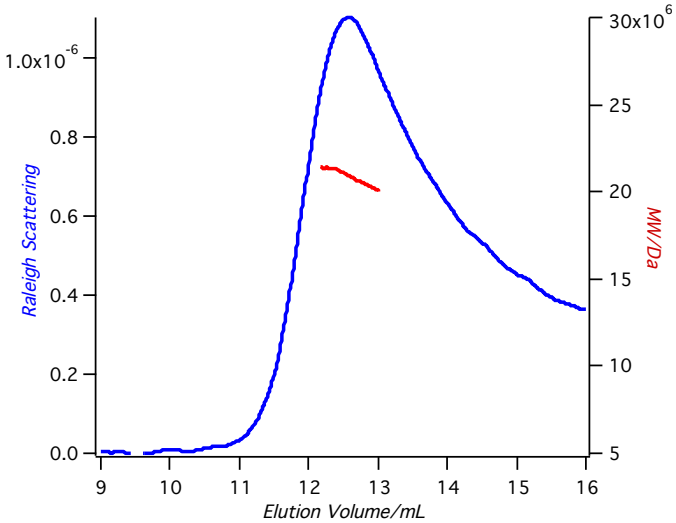



Figure D.5: HPLC-SEC coupled to multi-angle light scattering chromatogram of the P22-PTE construct. The left axis follows the Raleigh scattering while the right axis shows the average particle molecular weight (in Da) across the elution peak.

# APPENDIX E

## SUPPLEMENTAL INFORMATION FOR CHAPTER 5

## Experimental Section

### Site-Directed Mutagenesis of the P22 Scaffolding protein using Polymerase Chain Reaction (PCR)

The P22 assembler plasmid (pET11a (Novagen)) contains the gene for scaffolding protein (SP) upstream from the gene for coat protein. The SP gene is flanked by the upstream restriction enzyme site NdeI and downstream site BamHI. In order to make P22 virus cages with truncated scaffolding protein ($SP_{238}$) we used PCR to create a gene for truncated SP flanked by the restriction enzyme sites NdeI and BamHI. We then digested both the PCR product containing the gene for $SP_{238}$ and the assembler plasmid with NdeI and BamHI and ligated the $SP_{238}$ gene into the P22 assembler plasmid. The oligonucleotide primers for the mutagenesis are (Forward: 5'-GGAATTCCATATGGAGCTCCCATGGATGTGTACTCGACTATCCG-3' and Reverse: 5'-AAGGATCCTTATCGGATTCCTTTAAGTTTTGCCTTTAGCTTGCGG-3') along with the template gene $SP_{238}$ (in a background of pET3a (Novegen). These were used to create a $SP_{238}$ gene with the upstream restriction enzyme site NdeI and downstream site BamHI. The $SP_{238}$ gene was successfully cloned into the P22 assembler plasmid, resulting in a plasmid containing both the truncated SP and coat protein, and was confirmed by DNA sequencing. To add the polyanionic peptide (ELEAE) to the N-terminus of the truncated scaffolding protein the mutagenic oligonucleotide primer (5'-CATATGGAGCTCGAAGCTGAGCCATGGCTGGTGCCG-3') and the P22 $SP_{238}$ DNA clone as a template were used. This mutation, called P22 $SP_{Glu\ 238}$, was

confirmed by DNA sequencing.

Protein Expression

The P22 viral coat protein and scaffolding protein were heterologously expressed in BL21(DE3) cells. High levels of protein expression were induced with IPTG and yielded assembled viral cages devoid of nucleic acid. The viral cages were purified to homogeneity by lysis (50 mM phosphate, 100 mM NaCl, pH 7.6) followed by centrifugation (12,000xg, 60 minutes) to remove cell debris and ultra- centrifugation (48,000 rpm, 50 minutes) to pellet the virus particles. The resulting virus containing pellet was resuspended in buffer (50 mM phosphate, 25 mM NaCl, pH 7.6) and purified over a Sephacryl S-500 High Resolution column (GE Healthcare).

Iron Oxide Mineralization

The various P22 constructs were suspended in buffer (100 mM MOPS, 37.5 mM $Na_2SO_4$, pH 6.8) at a concentration of 6 mg/mL and treated with $Fe^{2+}$ with air as the only oxidant. $Fe^{2+}$ was added to the P22 solution at a rate of 166 Fe per viral capsid per minute to a final amount of 20,000 Fe atoms per viral capsid. Mineralized virus samples were purified using a Sephacryl S-500 High Resolution column (GE Healthcare) and the elution was monitored simultaneously at 280 nm (protein) and 350 nm (mineral).
*Transmission Electron Microscopy (TEM)*: TEM data were collected on a Leo 912, with Ω filter, operating at 80 keV. Samples were concentrated using microcon ultrafilters (Microcon YM-100) with 100 kDa Mw cutoff and transferred to carbon coated copper grids. Samples were imaged both stained with uranyl acetate and unstained.

Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) measurements were made on a Brookhaven Instruments ZetaPals (phase analysis light scattering) particle size analyzer. DLS was measured at 90° using a 661 nm diode laser, and the correlation functions were fit using a non-negatively constrained least-squares analysis.

## Supplemental Figures

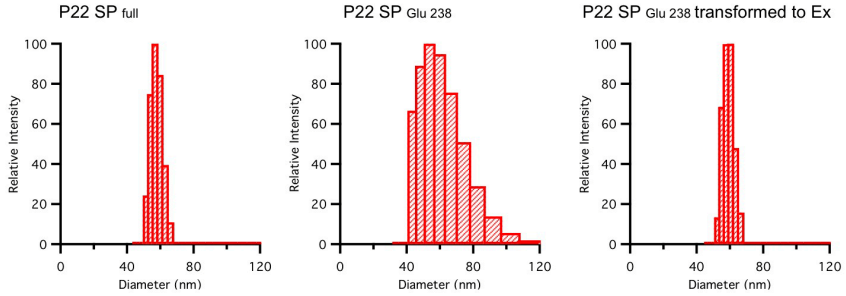



Figure E.1: DLS analysis of various P22 cage constructs utilized in this study.