
All ETDs from UAB

UAB Theses & Dissertations

2007

Bacteriophage P22 Scaffold Ng Protein: Functions and Mechanisms in Procapsid Assembly

William R. Marion
University of Alabama at Birmingham

Follow this and additional works at: <https://digitalcommons.library.uab.edu/etd-collection>

Recommended Citation

Marion, William R., "Bacteriophage P22 Scaffold Ng Protein: Functions and Mechanisms in Procapsid Assembly" (2007). *All ETDs from UAB*. 6731.
<https://digitalcommons.library.uab.edu/etd-collection/6731>

This content has been accepted for inclusion by an authorized administrator of the UAB Digital Commons, and is provided as a free open access item. All inquiries regarding this item or the UAB Digital Commons should be directed to the [UAB Libraries Office of Scholarly Communication](#).

BACTERIOPHAGE P22 SCAFFOLDING PROTEIN: FUNCTIONS AND
MECHANISMS IN PROCAPSID ASSEMBLY

by

WILLIAM R. MARION

PETER E. PREVELIGE, JR., COMMITTEE CHAIR
CHRISTIE G. BROUILLETTE
HERBERT C. CHEUNG
JEFFREY A. ENGLER
N. PATRICK HIGGINS

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Science

BIRMINGHAM, ALABAMA

2007

BACTERIOPHAGE P22 SCAFFOLDING PROTEIN: FUNCTIONS AND
MECHANISMS IN PROCAPSID ASSEMBLY

WILLIAM R. MARION

BIOCHEMISTRY AND MOLECULAR GENETICS

ABSTRACT

Bacteriophage P22 scaffolding protein is responsible for controlling the assembly of 420 monomeric coat protein subunits into a spherical, T=7 procapsid lattice. The precise mechanism by which this occurs however, is unclear. *In vitro*, this process can be faithfully reproduced which allows for analysis of assembly kinetics. These analyses have shown that the assembly is nucleation limited, the nucleus is a pentameric complex of coat protein and that kinetics are dependent on the oligomeric character of the scaffolding protein. Scaffold exists in a monomer-dimer-tetramer equilibrium and previous studies have clearly demonstrated that dimers are actively involved during assembly. Cryo-EM reconstructions of P22 procapsids have suggested that the tetrameric form of the scaffolding protein may be involved as well. In this work we have further examined the role of these scaffold oligomers during assembly by engineering scaffolding protein to exist in homogeneous oligomeric populations.

ACKNOWLEDGEMENTS

I would like to thank Drs. Andrea Cochrane, Peter Kim and the late Don Wiley for their generous donation of the mutant GCN4 peptides which were used extensively in this work. I would also like to thank my committee for their critical analysis and insightful recommendations. Finally, I would like to recognize the contributions of the members of the Prevelige lab whose commentary was particularly helpful.

TABLE OF CONTENTS

| | <i>Page</i> |
|--|-------------|
| ABSTRACT..... | ii |
| ACKNOWLEDGEMENTS..... | iii |
| LIST OF TABLES..... | vi |
| LIST OF FIGURES..... | vii |
| LIST OF ABBREVIATIONS..... | ix |
| INTRODUCTION..... | 1 |
| RESULTS..... | 8 |
| Dimeric Scaffold Reduces the Concentration of Coat Protein Required for Assembly..... | 8 |
| Dimeric Scaffolding Protein Requires Both Binding Domains for Assembly Activity..... | 9 |
| Engineering Higher Scaffold Oligomers..... | 10 |
| GCN4 Fusions Recognize Pre-existing Binding Sites in Empty Shells..... | 12 |
| Re-entry Protects Wild Type Scaffold From Proteolysis..... | 13 |
| GCN4 Scaffold Fusions Initiate Coat Protein Polymerization..... | 16 |
| Stopping Assembly Allows Examination of Assembly Intermediates..... | 19 |
| DISCUSSION..... | 40 |
| The Role of Dimeric Scaffold..... | 40 |
| Contribution of Trimeric and Tetrameric Scaffold to Assembly..... | 41 |
| A Possible Role for the N-Terminus..... | 42 |
| Model of Scaffold-Assisted Procapsid Assembly..... | 43 |
| MATERIALS AND METHODS..... | 46 |
| Mutagenesis..... | 46 |
| Protein Expression and Purification..... | 47 |
| Bivalent Dimer Preparation..... | 48 |

| | |
|---------------------------------------|----|
| Empty Procapsid Shells | 48 |
| Monomeric Coat Protein..... | 48 |
| Shell Binding Reactions..... | 49 |
| Procapsid Assembly Reactions..... | 49 |
| Trypsin Digestions | 50 |
| Circular Dichroism Measurements | 50 |
| Analytical Ultracentrifugation | 50 |
| REFERENCES | 52 |

LIST OF TABLES

| <i>Table</i> | <i>Page</i> |
|--|-------------|
| 1) Deconvolution of CD spectra of GCN4 – 141-303 scaffold fusions..... | 24 |
| 2) Biophysical data collected on scaffold fusion proteins | 25 |

LIST OF FIGURES

| <i>Table</i> | <i>Page</i> |
|---|-------------|
| 1) P22 Assembly Pathway | 7 |
| 2) Scaffold reduces critical coat concentration | 20 |
| 3) Dimeric scaffold requires both binding domains to promote assembly (1) | 21 |
| 4) Dimeric scaffold requires both binding domains to promote assembly (2) | 22 |
| 5) CD spectra of GCN4 – 141-303 scaffold fusions..... | 23 |
| 6) Velocity sedimentation of scaffold fusion proteins..... | 26 |
| 7) Equilibrium sedimentation of scaffold fusion proteins | 27 |
| 8) Re-entry of scaffold fusion into empty shells..... | 28 |
| 9) Re-entry protects wild type scaffold from proteolysis | 29 |
| 10) R74C dimers are not protected from proteolysis by shell binding..... | 30 |
| 11) Scaffold fusion proteins are protected from proteolysis by re-entry (1) | 31 |
| 12) Scaffold fusion proteins are protected from proteolysis by re-entry (2) | 32 |
| 13) Scaffold fusion proteins promote coat polymerization (1)..... | 33 |
| 14) Scaffold fusion proteins promote coat polymerization (2)..... | 34 |
| 15) Scaffold fusions are stripped from some assembled products..... | 35 |
| 16) Morphology of products assembled by scaffold fusions | 36 |
| 17) Assembly reactions with saturating scaffold fusions | 37 |
| 18) Wt and trimeric scaffold reaction intermediates..... | 38 |

| | |
|--|----|
| 19) Scaffold fusion reaction intermediates | 39 |
| 20) Proposed model of scaffold assisted procapsid assembly | 45 |

LIST OF ABBREVIATIONS

| | |
|----------|---|
| Å | Angstrom (1×10^{-10} m) |
| CD | circular dichroism |
| CMV | Cytomegalovirus |
| D | light attenuation |
| dsDNA | double stranded deoxyribonucleic acid |
| EM | electron microscopy |
| HSV | Herpes simplex virus |
| kDA | kilodalton |
| K_D | dissociation constant |
| kRPM | thousands of revolutions per minute |
| MDa | megadalton |
| mRNA | messenger ribonucleic acid |
| NMR | nuclear magnetic resonance |
| NSEM | negative stain electron microscopy |
| OD | optical density |
| PLP | procapsid like particle |
| PMSF | phenylmethylsulfonylfluoride |
| SAXS | small angle X-ray scattering |
| SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |

T triangulation

WT wild type

INTRODUCTION

Icosahedral viral capsids are protein shells that house and protect the viral genome. When mature these capsids are angular in appearance, and some are built as spherical intermediates known as proheads, or procapsids. The least complex of the spherical structures requires 60 subunits arranged with icosahedral symmetry¹. The rules of quasi-equivalence tell us that these small structures are capable of self-assembling since each subunit makes identical contacts with its neighbors throughout the lattice. The predictions of quasi-equivalence tell us that in these structures each protein subunit adopts a similar fold and the icosahedral lattice represents the minimum energy structure they can adopt^{1,2}. Making larger structures from the same subunits requires that they are capable of adopting similar, 'quasi-equivalent' conformations throughout the icosahedral lattice. These structures can be classified by their triangulation, or T, number which is an index of the maximum number of quasi-equivalent conformations the coat subunits may occupy. The total number of coat protein subunits in these structures is $60 \cdot T^1$.

$T > 1$ viruses need a mechanism to control the balance between the alternate conformations of the coat subunits. This information can be encoded within the coat subunits. These viral coat proteins may be able to switch between conformational states as a function of their position in the lattice as they polymerize into a procapsid structure, usually the first detectable intermediate structure in the assembly pathway³. Viruses with

$T \geq 4$ sometimes require accessory molecules to direct the proper switching, association and polymerization of the monomeric coat subunits. These molecules, termed scaffolding proteins, are chaperones of the assembly process and not part of the mature virion⁴⁻⁷. Scaffolding proteins have been identified as necessary molecules in several viral morphogenetic pathways⁴⁻⁷. Although they are required for faithful procapsid assembly, the molecular mechanism scaffolding proteins use to direct this process is not well understood. Defining the events at the molecular level should lead to the discovery of methods of interfering with assembly⁸.

One of the best characterized scaffold-assisted assembly pathways is that of bacteriophage P22. P22 is a temperate dsDNA bacteriophage of *Salmonella typhimurium*. Over the last 40 years the laboratories of Botstein, Casjens, King and Prevelige have extensively studied P22 at the genetic and biochemical level and have characterized the assembly pathway in great detail^{3,6,9-27}. Figure 1 outlines the P22 assembly pathway which begins with the construction of an intermediate particle, the procapsid. It is composed of 415 copies of the 46.7 kDa coat protein arranged with $T=7$ icosahedral symmetry and an inner core of 200 – 300 molecules of a 33.6 kDa scaffolding protein⁶. A dodecameric ring of the 82.7 kDa portal protein occupies one of twelve vertices¹⁴. 12-20 copies each of three pilot proteins, gene products 7, 16, and 20 which are involved in DNA injection into the host are incorporated as well¹³. Once the procapsid is assembled, the viral dsDNA is actively translocated through the portal vertex in a headful manner^{18,19}. This packaging event occurs concomitantly with extensive, irreversible changes in the proteinaceous lattice²⁸. These changes include scaffold exit

from the structure³, a 30% volumetric expansion of the lattice and acquisition of the angular appearance of the mature viral capsid¹². When packaging is complete, the portal vertex is sealed by three minor proteins, (gps 4, 10, 26) which stabilize the packaged head¹⁵. The final modification occurs when up to six trimers of the 71.8 kDa tailspike protein are added at the portal vertex^{9,16}. Once the tailspikes attach, the phage is infectious⁹. This model for P22 assembly and maturation has been tested and refined for 30 years, and the intermediate steps are well characterized in the literature.

Mutagenic studies were essential in determining the steps of the P22 morphogenetic pathway. Establishing the roles of the three structural components of the procapsid - coat, scaffold and portal - required analyses of phage deficient in these proteins and characterizing the phenotypic defects^{3,17}. P22 bacteriophage that lack scaffolding protein must accumulate coat to high intracellular concentrations before it can polymerize¹⁷. In these infections the observed coat morphologies are procapsid-like-particles (PLPs) with T=7 (58%) and T=4 (25%) dimensions as well as large spirals of coat polymers (17%)^{3,6,17}. Although these spiral 'monsters' are low in total count, they account for nearly half of the polymerized coat protein in these lysates¹⁷. It is also worth noting that the T=7 particles, which appear normal by EM, lack the portal protein and cannot be packaged with the viral dsDNA. Infections with phage that are deficient in portal protein can assemble T=7 PLPs which contain all other procapsid-associated proteins^{3,17,20}. These phage are blocked at the DNA packaging step and accumulate scaffold-containing PLPs in the cell. In these cases, ~95% of the assembled particles have the proper T=7 symmetry. Phage that are deficient in coat protein are incapable of

assembling any intermediate structures³. These results indicate that the scaffolding protein is directly essential for infectivity by controlling assembly of the T=7 lattice^{3,6,17}, and indirectly via involvement with portal incorporation.

In vitro, purified P22 scaffold and coat proteins can be manipulated to generate procapsid-like-particles (PLPs)^{27,29}. This reaction allows for analysis of assembly kinetics and thermodynamics by monitoring light attenuation (D), or solution turbidity²⁷. The products of this reaction can be easily recovered for analysis by other techniques such as; equilibrium or velocity sedimentation, native or denatured gel electrophoresis, and negative stain-electron microscopy (NSEM). This *in vitro* system was used to identify scaffold's minimal C-terminal coat-protein binding domain^{24,30} and determine that scaffolding dimers are an active species during procapsid assembly^{26,31}. Analysis of concentration dependent kinetics suggests that the rate-limiting step in P22 assembly is the formation of coat pentamers³¹. This assay has proven an invaluable tool for studying scaffolding-assisted procapsid assembly.

Oligomerization of scaffolding proteins has been observed in several icosahedral viral systems (Φ 29, CMV, HSV, T4, λ , P22)^{5,26,32-35}. Bacteriophage T4 scaffolding protein, for example, forms a stable core onto which the coat protein assembles⁵. T4-like scaffolding cores have never been observed in Φ 29, λ , CMV or P22³⁴ but these proteins do self-associate. P22 scaffold exists in a reversible monomer / dimer / tetramer (1/2/4) equilibrium at physiologically relevant concentrations²⁶. Cryo-EM reconstructions of wild-type P22 procapsids reveal four scaffold C-terminal tails bound at the hexameric

coat clusters³⁶. Because of these observations, it seems likely that low order scaffold oligomers are necessary for regulating the energetic balance between alternate conformations of coat protein subunits during procapsid assembly.

With the exception of $\Phi 29$, high resolution structures of scaffolding proteins are scarce, due to their intrinsic flexibility and their propensity to oligomerize^{25,33}. Over the last 10 years the Prevelige lab has used a variety of techniques to generate a low-resolution structural model of P22 scaffolding protein²². Raman spectroscopy of scaffold indicates that it consists of several α -helical domains and lacks a defined hydrophobic core³⁷. Hydrodynamic modeling shows the monomer is a highly elongated molecule, 247 Å long and 22 Å in diameter^{26,29}. Small-Angle X-Ray Scattering (SAXS) collected on an assembly-active C-terminal fragment of wild-type scaffold (141-303) shows an unstructured linker region approximately 30 Å long which connects the C-terminal, coat-binding 'tail' and a dimerization interface (Tuma-unpublished results). The linker (~ res. 216-268) is believed to be critical for directing the spatial patterning of the bound coat subunits. The two properties of this domain that should be most crucial to procapsid assembly are therefore, flexibility and length. The NMR structure of a C-terminal fragment (res. 238-303) shows that the coat binding domain (res. 269-303) is an amphipathic helix-loop-helix dense with basic residues. Residues 238-268 were disordered under these conditions and are not represented in the structure²⁵. Interestingly, HSV and CMV scaffolding proteins bind coat proteins with an acidic C-terminal domain. These ~25 residue domains are predicted to have an amphipathic α -helical structure based on primary sequence analysis³⁸⁻³⁹.

The 141-303 fragment described above has been well characterized *in vitro*^{22,24,40}. Raman spectroscopy and SAXS studies have demonstrated its highly helical, elongated secondary structure²². Removal of the 140 N-terminal residues decreases the monomer - dimer association constant 8-fold, and eliminates tetramer formation²². Interestingly, this fragment increases assembly rates *in vitro*, however the fidelity of assembly is lost at high scaffold to coat ratios²⁴. This loss of fidelity has been attributed to the fragment's poor dimerization²⁴. We have used this fragment as a template to study the role of scaffolding oligomers during assembly by adding peptides to its N-terminus which spontaneously form homogeneous populations of dimers, trimers or tetramers with nanomolar affinities⁴¹⁻⁴³.

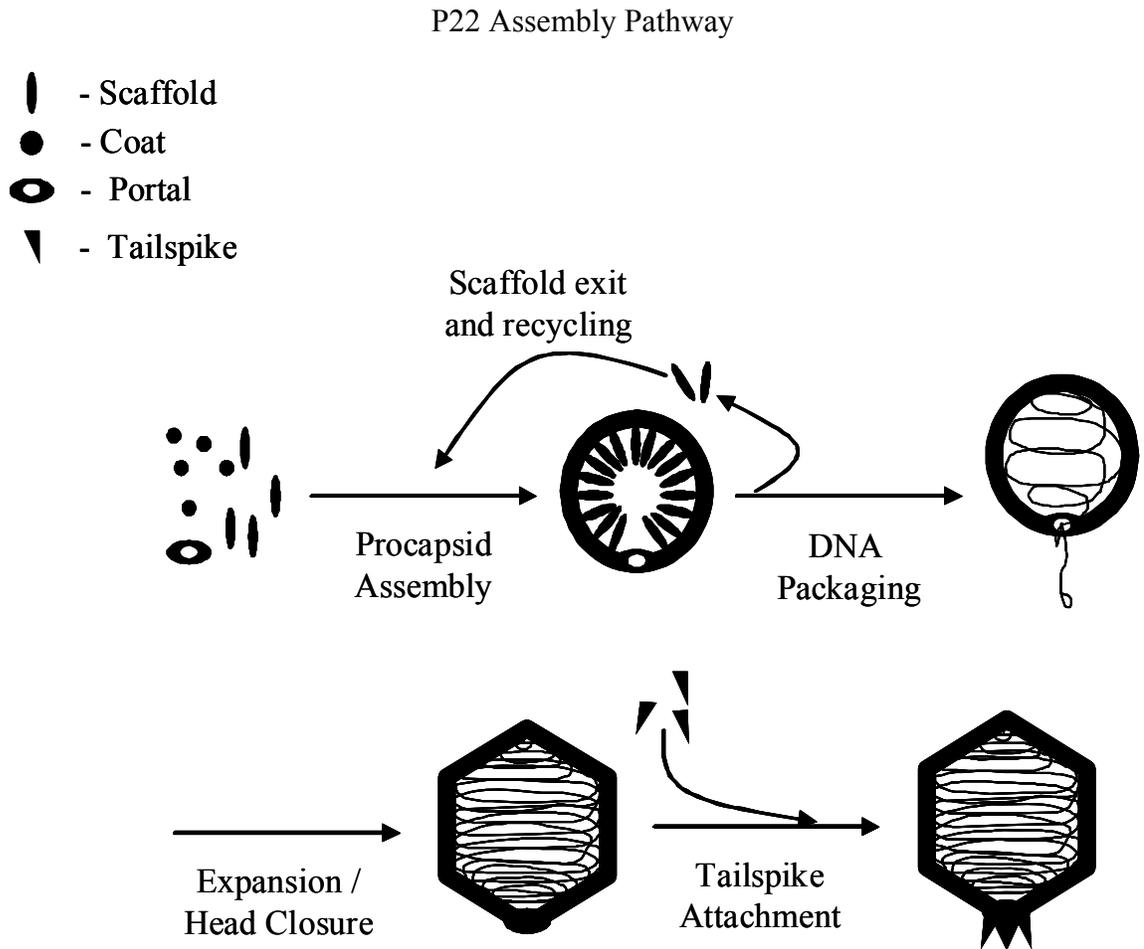


Figure 1 – Assembly pathway of bacteriophage P22. Portal, scaffold and coat subunits are translated from late mRNA and self-assemble into the first intermediate structure, the procapsid. Viral dsDNA is actively translocated through the unique portal vertex in a headful manner. Concomitant with this scaffold exits the structure through holes in the procapsid hexons and is used for further rounds of procapsid assembly. Once packaging is complete, the portal vertex is plugged by gene products 4, 10 and 26 which stabilizes the head. Tailspike adds to the portal vertex, completing the replication cycle.

RESULTS

Dimeric Scaffold Reduces the Concentration of Coat Protein Required for Assembly

A previously characterized scaffold point mutation, R74C, was shown to enhance the kinetics of in vitro PLP assembly 1.5 – 2 fold as an oxidized dimer (R74C_{ox})²⁶. We noticed that these full-length covalent scaffold dimers also increased the final amount of light scattered in these experiments. This could be explained by the formation of large, aberrant particles which scatter more photons than PLPs, increasing solution turbidity. This did not seem likely, since there was no evidence of aberrant structures in sucrose gradients or by negative-stain EM²⁶. Another possibility was that the increase in scattering reflected an increase in PLP assembly. Since assembly is nucleation limited^{27,31}, this could result from a lowered critical concentration⁴⁴. Prevelige, et. al. demonstrated that wild type scaffold, even when supplied in excess, was unable to initiate assembly below 0.3 mg/ml (~6.5 μ M) monomeric coat protein³¹. We repeated this experiment with our covalent dimers to determine if this value would be lowered in their presence. Figure 2 shows the concentration of soluble coat protein remaining after a 24 hour incubation with wild-type and R74C_{ox}, as determined by integrating the intensity of coomassie-stained coat protein bands in 10% SDS-PAGE gels. The data shows that the covalent dimers reduce the amount of unpolymerized coat protein remaining in this assay by 50-75% relative to wild type. This data is evidence that scaffold dimers lower the

energetic barrier of PLP nucleation, most likely by eliminating the energetic cost of forming scaffold dimers before binding coat proteins.

Dimeric Scaffolding Protein Requires Both Binding Domains for Assembly Activity

Parker et.al., demonstrated that a C-terminal scaffold fragment, 141-303, is capable of assembling PLP's *in vitro* and is non-competitively inhibited by a similar fragment (141-292) with an 11-residue C-terminal deletion (Δ tail)²⁴. The proposed mechanism of this inhibition is the transient formation of heterodimeric, monovalent scaffolding molecules, which would be incapable of promoting PLP assembly. A critical element of this model is the need for two coat protein binding domains at the extreme C-terminus of scaffold dimers. By introducing the Δ tail mutation (Δ 293-303) into the R74C scaffold and creating a disulfide crosslinked scaffold dimer with only one coat binding domain, i.e. monovalent, we were able to test this hypothesis directly. As a negative control, a nonvalent (no binding domains) covalent dimer was constructed and tested as well. Figure 3 shows the kinetics of *in vitro* PLP assembly initiated by each of these dimeric constructs, as well as wild type. The scaffold dimers which lack a coat binding domain show no time-dependent increase in solution turbidity, indicating they are assembly incompetent. Figure 4 shows the sedimentation character of particles assembled after 4 hours by these constructs. In these assays, the co-migration of coat and scaffolding proteins through the sucrose gradients is indicative of particle formation. The coat and scaffold have migrated approximately 2/3 of the length of the gradient in the reactions catalyzed by wild-type and the bivalent dimer. Covalent dimers missing one or both coat binding domains at the scaffold C-terminus generate nothing that is capable of

sedimenting beyond the top of the gradient. Together, these data demonstrate the efficacy of dimeric scaffold during assembly, and suggest a role for dimeric scaffold during nucleation.

Engineering Higher Scaffold Oligomers

Cryo-EM difference mapping of scaffold-lacking and scaffold-containing procapsids showed the presence of the scaffolding protein's coat-binding domain at defined positions in the T=7 lattice^{36,46}. This positioning led to a presumptive role in form determination for the tetrameric scaffold species. According to the model of assembly proposed by Thumann-Commike et. al., dimeric scaffold would be involved at early steps during assembly, and tetramers may be important during later stages. They proposed that a trimeric cluster of scaffold could result in the formation of T=4 lattice as well. We decided to test this model by engineering scaffolding proteins to exist as homogeneous populations of a specific oligomer. This was accomplished by fusing peptides that spontaneously form homogeneous populations of parallel, coiled-coil dimers, trimers and tetramers^{43,47,48} to the N-terminus of the previously characterized 141-303 scaffold fragment, which is mostly monomeric at low concentration ($K_{D12} \sim 640 \mu\text{M}$)²². The fused peptides were developed and described by Harbury, et. al. and are point mutations based on the structure of the GCN4 leucine zipper dimerization domain⁴³. They are small (34-36 residues), and form homogeneous populations of specific oligomers as parallel coiled-coils with nanomolar affinities. Based on our model of wild type scaffold, these peptides fit critical elements of our hypothesis of scaffold structure and function during procapsid assembly.

To determine the structural effects of forcing oligomerization on the 141-303 fragment, the CD spectra of the fragment and each fusion protein were collected and deconvoluted using the method of Chang et. al.⁴⁹. The spectra are shown in figure 5 and the corresponding deconvolutions are shown in table 1. It is clear from the spectra that there is a significant increase in helical signature for each scaffold, which results from the interaction of the parallel N-terminal coiled-coils. Because the fused peptides are all the same length, the magnitude of this increase should be similar for all three proteins, yet it is clear from the figure that they are not. Attempts to determine the source of this discrepancy using CD were inconclusive, yet the fact that the scaffold fusions are capable of recognizing monomeric and assembled coat protein binding sites (results) indicates that these structural perturbations were not sufficient to inhibit coat scaffold interactions. We interpret this as evidence that the binding domains are natively folded and the structural perturbations likely represent structural changes elsewhere in the protein.

We also wanted to characterize the oligomeric homogeneity of these engineered proteins in solution. This was examined by static light scattering analysis and analytical ultracentrifugation. The proteins were eluted through an S-200 size exclusion column (Pharmacia) and the elution peaks characterized by 90° light scattering and change in solution refractive index. This analysis provides a direct determination of solute molecular weight at the point of elution. The collected peak fractions were sedimented in a Beckman XLA-80 analytical ultracentrifuge in order to determine their composition as a function of sedimentation values, where increases would presumably reflect the

oligomeric state of the proteins in solution. The data from these combined analyses are shown in table 2, and indicate that the solutions of the fusion peptides are nearly, but not completely, homogeneous. Figure 6 shows the size distribution as resolved by velocity sedimentation of 5 μM solutions of each fusion protein. Light scattering analysis of molecular weights indicates that the most prominent species in each experiment represents the predicted oligomeric state. The appearance of at least two species in each experiment demonstrates that the peptides exist in an equilibrium of monomer-oligomer.

To determine the monomer-oligomer association constants, equilibrium sedimentation experiments were conducted over an 8-fold range of concentrations at three speeds. Figure 7 shows a representative residual trace from each of these experiments. Global analysis of these data shows that the dimeric fusion has a K_D of 1.2 μM . This represents an increase in dimeric association of nearly 600-fold for the 141-303 scaffold fragment. The trimeric fusion was nearly homogeneous as trimer, and gave values beyond the lower limit of detection for the experiment, and the tetrameric construct was resolved with a K_D of $1.4 \times 10^{-17} \text{ M}^3$. These data show that over the concentration range used for assembly and binding experiments described in this work, the fusion peptides are occupying their predicted oligomeric states at levels of 50-95%.

GCN4 Fusions Recognize Pre-existing Binding Sites in Empty Shells

A unique behavior of wild type scaffolding protein is that of shell re-entry^{30,50}. Wild type scaffold can be titrated into solutions of intact, empty shells and it will recognize the internal binding sites, and completely re-enter the shell. Presumably this

occurs through the 30 Å x 45 Å holes in the center of the hexons on the T=7 lattice^{23,30,50}. As an initial control for activity, the fusion proteins were assayed for their ability to recognize pre-existing binding sites. The peptides were titrated in excess into solutions of empty shells and the resulting particles were sedimented through sucrose gradients. The results, shown in figure 8, demonstrate that the fusions do indeed comigrate with the coat protein shells.

Parker, et.al. showed that the 141-303 peptide could bind empty shells in numbers nearly double that of wild type scaffold when supplied in excess²⁴. Since the binding and assembly characteristics of the 141-303 scaffolding fragment had been previously described, we were provided a background to help understand how forced oligomerization would alter its activity. To determine the scaffold : coat stoichiometry of the fusions in the filled shells, the peak fractions were harvested and analyzed by densitometry. We found that the 141-303 fragment and its oligomeric fusions bound empty P22 shells with double the stoichiometry of wild type. Parker et.al. made note of this re-entry phenomenon with the 141-303 fragment and suggested that this behavior was driven by mass, rather than binding affinity or stoichiometry. The 2-fold increase in binding stoichiometry observable when one-half the scaffold mass is removed would tend to support this model.

Re-entry Protects Wild Type Scaffold From Proteolysis

Scaffolding protein binding sites are on the inner surface of the procapsid lattice³⁶. The P22 procapsid however, has only 420 coat protein molecules, i.e. possible

binding sites. The stoichiometry of the 141-303 fragment and the fusion proteins, which are presumably inside the capsid, is significantly greater than this value. We investigated the possibility that some, or all of the scaffold molecules were not stably internalized. To test this hypothesis, scaffolding protein constructs which had been incubated with empty shells for one hour were exposed to trypsin. Since empty procapsids are stable under these conditions and free scaffold is rapidly digested by protease³⁰, we chose these assays to digest those scaffold molecules which were not stably maintained within the intact procapsid. Figure 9 shows that wild type scaffolding protein is protected from digestion under these conditions, while purified scaffold is rapidly consumed. The 141-292 fragment, which does not bind the empty shell, is completely unprotected from trypsin digestion in the presence of empty shells, demonstrating the inability of P22 procapsids to inhibit trypsin activity. Presumably, the protection is due to the inability of trypsin to penetrate the interior of the intact procapsid. Greene and King demonstrated that there are high and low affinity binding sites in the P22 lattice³⁰. Parker et. al. later showed that the dimeric R74C_{ox} construct is capable of binding only high affinity sites, which amount to approximately 60 molecules per procapsid⁴⁰. In our hands, the binding stoichiometry of this construct was not consistent between sucrose gradient sedimentation experiments. To determine if this was due to any unstably bound scaffold, we digested shells which were preincubated with this protein in its oxidized and reduced forms. We wanted to determine whether these molecules, although bound to high-affinity sites, were stably incorporated within the procapsid. The results in figure 10 show that there is a significant amount of protection of this scaffold under reducing conditions. A control experiment with wild type scaffold and shells under both conditions revealed that β -ME had no effect

on trypsin digestion (data not shown). This implies that while the high affinity sites are accessible to the dimeric scaffold, the cysteine crosslink prevents complete re-entry. Presumably this is due to steric hindrance in the N-terminal domain, since the reduced scaffold is protected by the procapsid shell. These experiments again suggest that this phenomenon occurs via small holes in the lattice hexons. Indeed, experiments using empty shells derived from portal-deficient phage indicate that re-entry does not involve this unique vertex (data not shown). We interpret these results to indicate that this process involves a conformational change at the N-terminus which is forbidden by the cysteine crosslink at residue 74.

Figure 11 shows the results of a 20 minute tryptic digestion of the 141-303 fragment and the fusion peptides pre-incubated with empty shells. In order to determine the stoichiometry of binding after tryptic digestion, the reactions were quenched by the addition of PMSF after twenty minutes. The shells were pelleted, resuspended in Tris-Tricine sample buffer and the stoichiometry determined by densitometry of 12.5% Tris-Tricine gels. Figure 12 shows the results of this assay, which were surprising. While wild-type scaffold and the dimeric fusion remain protected in similar amount by mass, the 141-303 fragment, trimeric and tetrameric scaffolds are significantly less protected. The 141-303 fragment, due to its poor self-association, does not form scaffold-scaffold interactions as well as wild type, which may significantly reducing its binding strength. This situation would create two classes of bound scaffold, one that remains tightly bound at high affinity sites, and one that binds either low-affinity sites,⁴⁰ or is capable of freely entering and exiting the interior of the procapsid, due to decreased self-association. In

these assays, the second class would be rapidly digested by external proteases, such as trypsin, leaving only the tightly bound scaffold protected. The wild type and dimeric fragment may not be able to exit as freely as the 141-303 fragment due to stronger self-association. Within this framework, the trimeric and tetrameric scaffolds would bind at least as well as the dimeric fusion. Indeed, stripping these oligomeric scaffolds from refilled shells shows very little stoichiometric difference between the dimer, trimer and tetramer (data not shown). However, in order to be completely protected from external protease, the proteins must be fully internalized. The data suggests that there are two populations of these shell-associated oligomeric scaffolds. One that is able to bind to the internal surface and fully re-enter, and another that is bound to the shells yet is hindered at the re-entry step. Since the binding sites straddle the hexon holes, it is possible that the internal (protected) trimeric and tetrameric scaffolds are doing the same. This would present a significant energetic barrier to a second scaffold oligomer trying to enter the empty shell at this site. This behavior is similar to that observed for the R74C_{ox} dimer, which is completely unprotected in these assays.

GCN4 Scaffold Fusions Initiate Coat Protein Polymerization

The fusion peptides were able to recognize pre-existing binding sites during re-entry experiments, suggesting that they presented the coat binding domain in a manner similar to wild type scaffold. It was then necessary to determine how these oligomeric scaffolds would initiate *in vitro* PLP assembly from monomeric coat proteins. To assess this, assembly reactions were done with 5 scaffolding proteins for every 7 coat monomers. This represents the maximum scaffold to coat ratio observed in wild type

procapsids^{3,6,40}. Figure 13 shows the time-dependent increase in solution turbidity measured at 250 nm with 10 μ M scaffold and 14 μ M coat. The kinetics and the final amount of light scattered (which reflects the amount and size of the polymerized material) are clearly dependent upon the engineered valency of the scaffolds. While this dependence is noteworthy, it does not reveal any information about the products of the reactions. Figure 14 shows the sedimentation profiles of the particles assembled by the 141-303 fragment and the oligomeric scaffolds. The 141-303 fragment and its dimeric fusion generated gradient peaks in the position expected for a properly assembled PLP. The tetrameric fusion generated a smear of scaffold and coat which co-sedimented. In this case it was clear to us that assembly had gone off-pathway. The trimeric scaffold generated a detectable peak of coat and scaffold protein which sedimented slower than wild type particles. As a check for particle stability / scaffold binding, the products of these reactions were incubated with 0.7 M Gdn•HCl for 1 hour. This treatment removes 85-90% of scaffolding protein from *in vivo* assembled procapsids without disrupting the lattice. The results in figure 15 show that scaffold was indeed stripped from wt, 141-303, R74C_{ox} and the dimeric fusion assembly products, yet it had no effect on the products of trimeric and tetrameric scaffold assembly. Figure 16 shows NSEMs of the reaction products and reveal what appear to be properly dimensioned PLP's in all reactions, except those assembled by the trimeric and tetrameric scaffolds.

Wild type scaffold can be titrated at >50-fold excess into solutions of monomeric coat protein without compromising the fidelity of the assembly process. Parker et. al. observed that this is not the case with the 141-303 fragment²⁴. This property was

attributed the decreased self-association observed for this fragment. We tested this hypothesis by performing assembly reactions over a range of fusion concentrations while holding the coat concentration constant. The reactions were monitored at 250 nm and deemed to be complete when no further increase in turbidity was noted with added scaffold (data not shown). When these scaffold saturated reaction products were analyzed by sucrose gradient sedimentation, they fell into two distinct classes and are shown in figure 17. The first class was assembled by wt, the 141-303 fragment and its dimeric fusion. At low scaffold to coat ratios particles assembled via the truncated peptides consisted mainly of PLP's. However, as Parker observed for the 141-303 fragment, the amount of aberrant product was increased concomitant with the concentration of added scaffold. This pattern was unchanged by the dimeric scaffold, indicating that the lack of assembly control by this fragment is not completely due to its decreased dimeric association. Another feature of these gradients is the presence of unassembled coat protein. The second class of products was mainly aberrant and generated by the trimeric and tetrameric fusions. This second group presents soluble scaffold, but not coat protein after assembly is complete. This data shows that under our experimental conditions the products assembled by trimeric and tetrameric scaffolds, as we have constructed them, are not PLP's. The lack of soluble coat protein in these sucrose gradients indicates that these oligomers are generating off-pathway structures due to increased, or improper nucleus formation.

Stopping Assembly Allows Examination of Assembly Intermediates

By adding 200 mM NaCl at the reaction half-times, we were able to view intermediates in these reactions and better clarify what was being initially assembled. This method was applied to assembly reactions initiated by the trimeric fusion in order to determine why assembly appeared to be normal at low scaffold to coat ratios yet went off-pathway at saturating levels. Figure 18 shows the results from this reaction at equimolar scaffold / coat levels. The trimeric fusions, as was seen at 5:7 scaffold to coat ratios, assemble what appear to be complete PLP-like structures. However these structures continue to grow larger and more massive, strongly indicating that the structures are not closed.

Figure 19 shows sucrose gradients of the assembled particles at their respective reaction half times with equimolar scaffold and coat for all 5 scaffolds discussed in this work. Again, two types of behavior are evident. The first is the wild type case, which is closely mimicked by the 141-303 fragment and the dimeric scaffold. These scaffolds generate discernable comigrating peaks of coat and scaffolding protein, and very few large, aberrant structures. The second, which eventually leads to an aberrant particle, is seen with the trimeric and tetrameric scaffold. These preformed scaffold oligomers not only initiate off-pathway assembly, but are capable of over-riding any control of PLP assembly that may be encoded within the coat protein subunit.

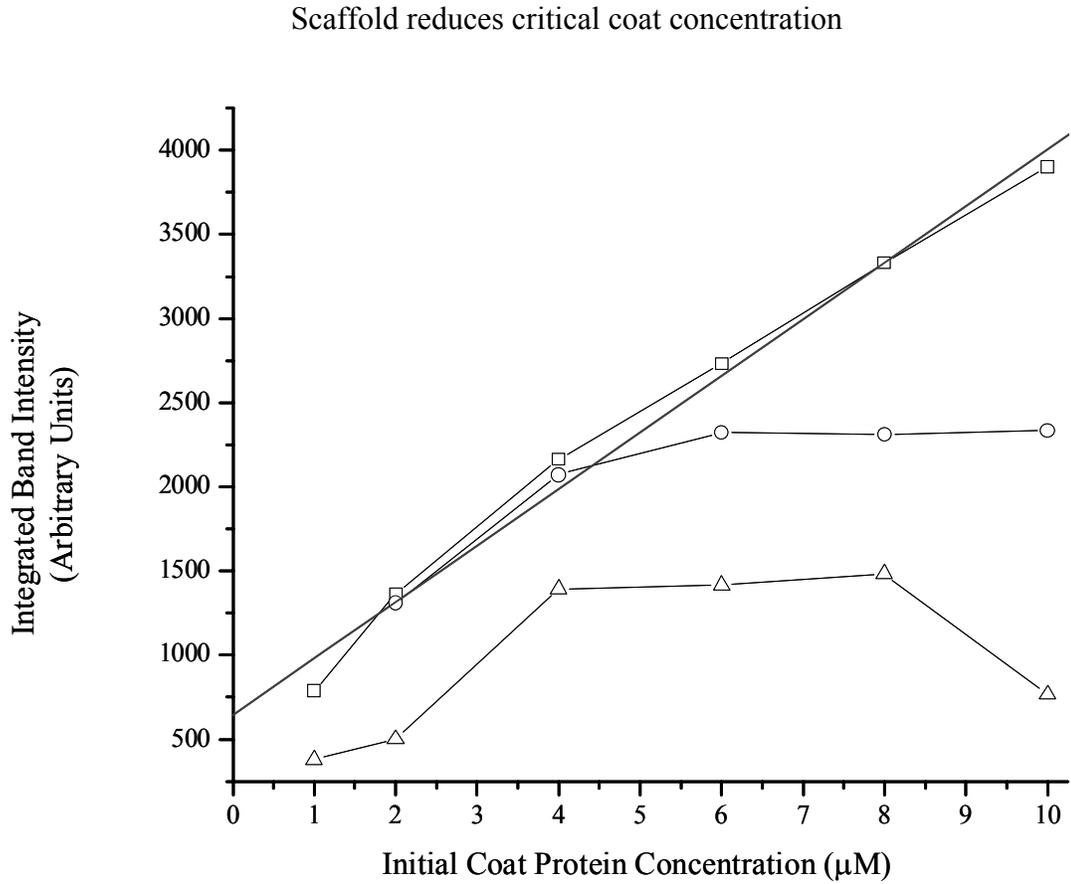


Figure 2 – Full length dimeric scaffold reduces the concentration of coat protein required for in vitro PLP assembly 75 - 50%. Soluble coat protein remaining after 12 hours of incubation with 30 μM wild type scaffold or 30 μM R74C_{ox} scaffold was determined by pelleting the samples to 100s. The graph shows the integrated coat protein band intensity from coomassie-stained SDS-PAGE gels vs. the starting coat concentration. Legend: Squares – No Added Scaffold; Circles – 30 μM WT scaffold; Triangles – 30 μM R74C_{ox}; The solid line represents a linear fit of the average intensity of the staining of the scaffold-lacking controls in both experiments.

Dimeric scaffold requires both binding domains to promote assembly (1)

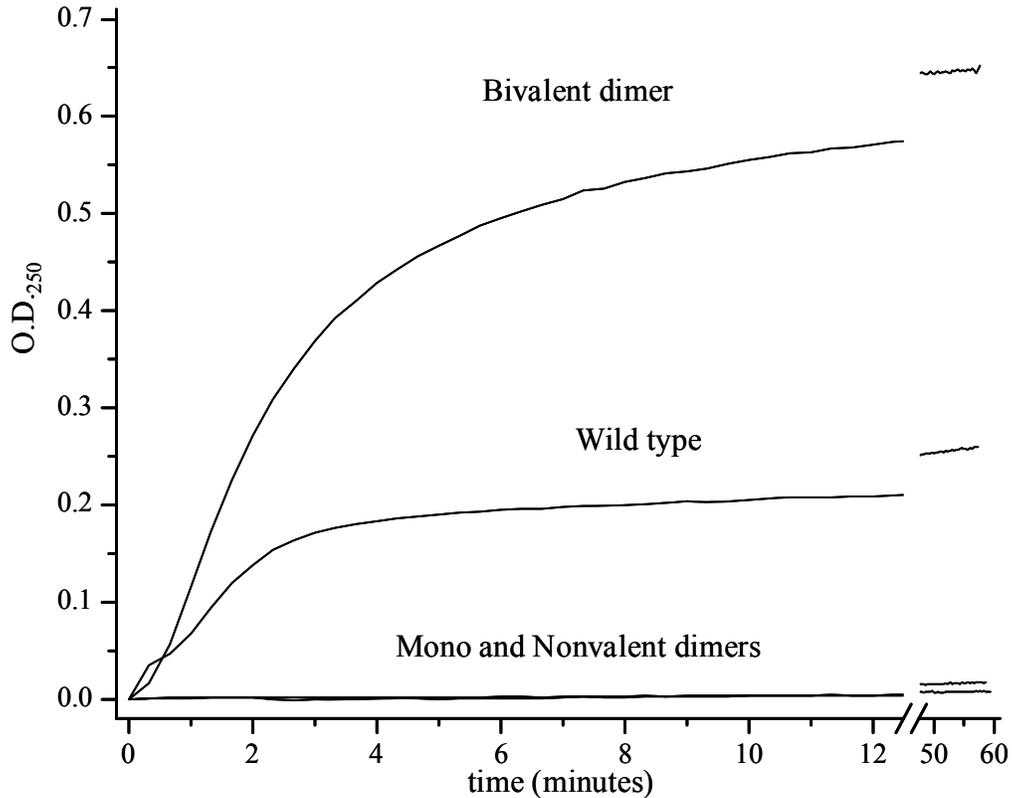


Figure 3 - Kinetics of *in vitro* PLP assembly from monomeric coat proteins by wild type and dimeric scaffold species. Assembly is followed by monitoring the change in solution turbidity at 250 nm. Equimolar solutions of scaffolding proteins were titrated into thermally equilibrated cuvettes containing monomeric coat protein. The increase in initial rate and final amount of turbidity by preassembled scaffold dimers over wild type of was shown by Parker et. al. and is indicative of the critical role of dimeric scaffold during assembly. The lack of significant change in turbidity for the mono and nonvalent dimeric scaffolds indicates that these proteins are unable to promote PLP assembly.

Dimeric scaffold requires both binding domains to promote assembly (2)

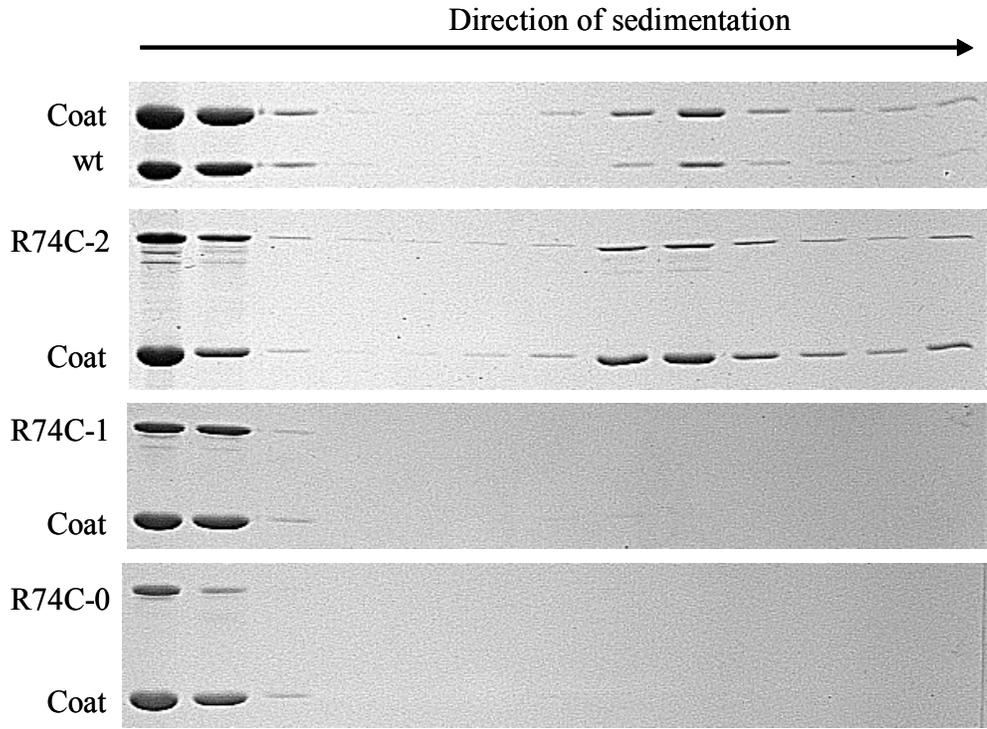


Figure 4 - Sedimentation of particles assembled after four hours by wt and multivalent dimers. Gradient conditions were chosen so that PLPs assembled by wild type scaffold would migrate approximately 2/3 of the length of the gradient. Bivalent dimers assemble particles which sediment to the same position as wt . Note that there is no detectable co-sedimentation of coat and scaffold in reactions where dimers lack a coat binding domain. This is consistent with the absence of spectroscopically detectable assembly in these reactions. Gradients from top to bottom: wild type, R74C_{ox} – bivalent (2), monovalent (1) and nonvalent (0) dimers.

CD spectra of GCN4 – 141-303 scaffold fusions

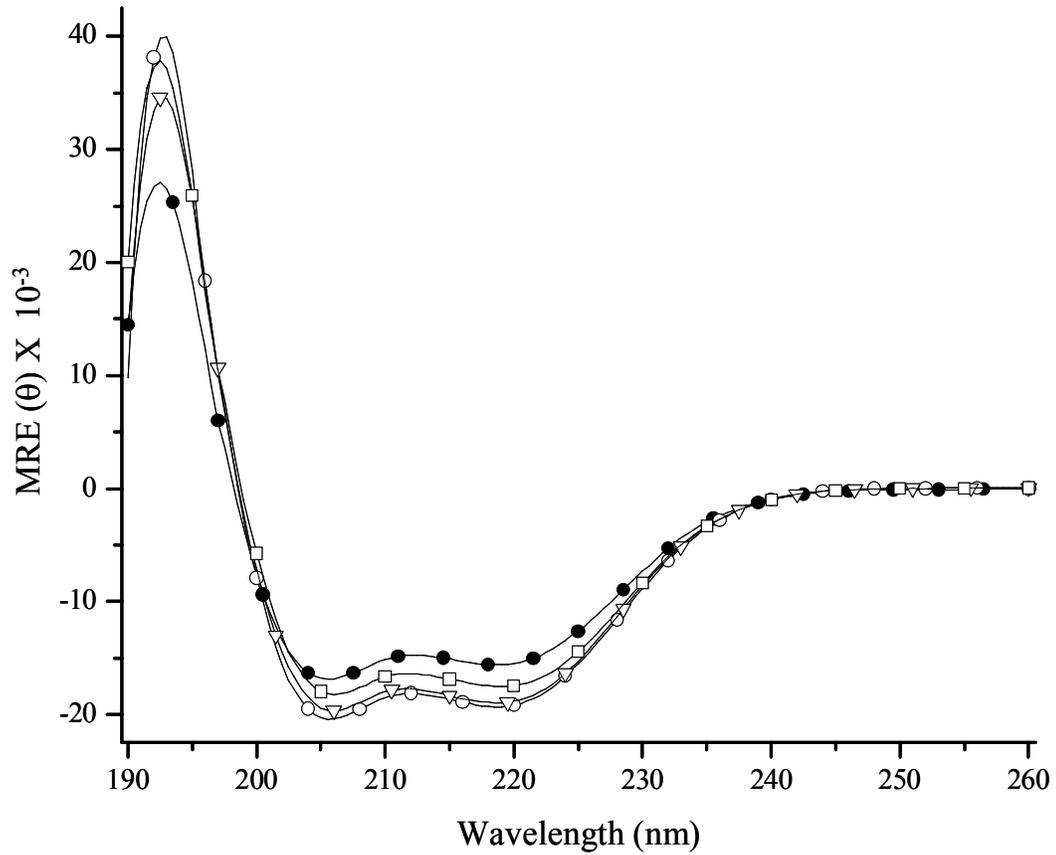


Figure 5 – CD spectra of the GCN4 fusion proteins and the 141-303 fragment. All fusion proteins show an increase in helical signature. Deconvolutions of the spectra are shown in table 1. Legend; Closed circles– 141-303 fragment, Open circles - Dimeric scaffold, Triangles - Trimeric scaffold, Squares - Tetrameric scaffold.

Deconvolution of CD spectra of GCN4 – 141-303 scaffold fusions

| Scaffold | % | % | % | % Random |
|-----------|-----------------|----------------|---------------|----------|
| Construct | α -helix | β -sheet | β -turn | Coil |
| 141-303 | 31 | 14 | 17 | 38 |
| Dimer | 48 | 8 | 14 | 30 |
| Trimer | 46 | 12 | 12 | 30 |
| Tetramer | 42 | 10 | 14 | 34 |

Table 1 – Secondary structure content of scaffold constructs as determined by the deconvolution method of Chang et. al.. The fused GCN4 peptide domains represent 17% of the amino acids in the fused proteins. The increase in helical content for each of the oligomerized scaffolds ranges from 11% (tetramer) to 17% (dimer).

Biophysical data collected on scaffold fusion proteins

| Scaffold | MW (kDa) | MW (kDa) | |
|-----------|----------|----------|------|
| Construct | Monomer | 90° LS | S |
| Dimer | 21932 | 47860 | 2.51 |
| Trimer | 21973 | 85110 | 3.28 |
| Tetramer | 22214 | 95500 | 3.91 |

Table 2 - Summary of biophysical data collected for the fusion peptides collected by 90° light scattering and analytical ultracentrifugation. The molecular weights determined by light scatter are slightly higher than predicted for each construct indicating that some higher association may be occurring in solution. This is visible in the velocity sedimentation analysis for each peptide as well. This provided a source of non-ideality during the equilibrium centrifugation which was reduced by the addition of 500 mM NaCl for those experiments.

Velocity sedimentation of scaffold fusion proteins

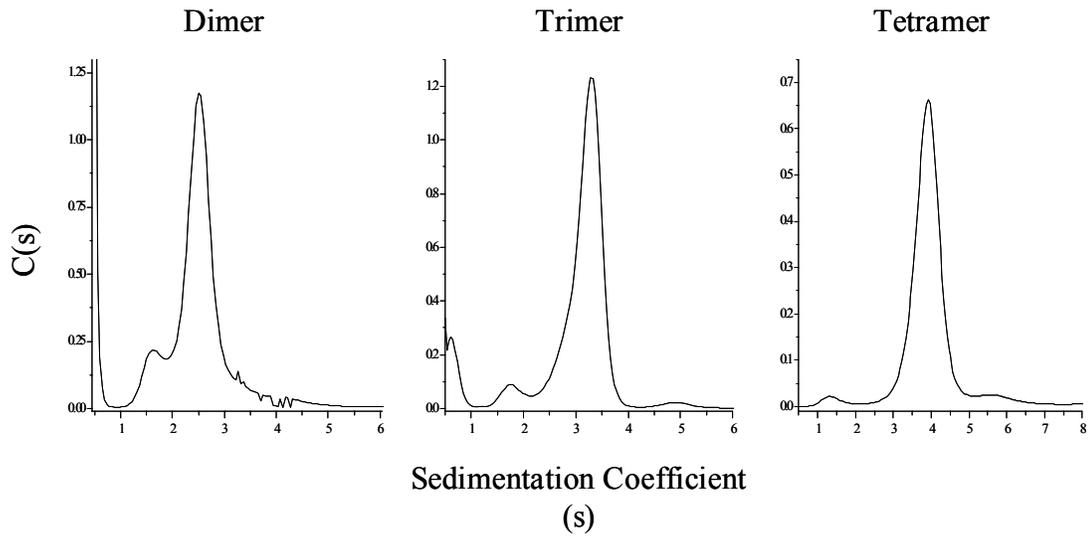


Figure 6 - Velocity sedimentation of engineered oligomers. 5 μ M samples of each peptide were sedimented at 59 kRPM and sedimentation coefficients were determined using sedfit 8.4. None of the samples are homogeneous, yet the s-value of the major peak increases as the engineered oligomeric state of the fusion peptides increases.

Equilibrium sedimentation of scaffold fusion proteins

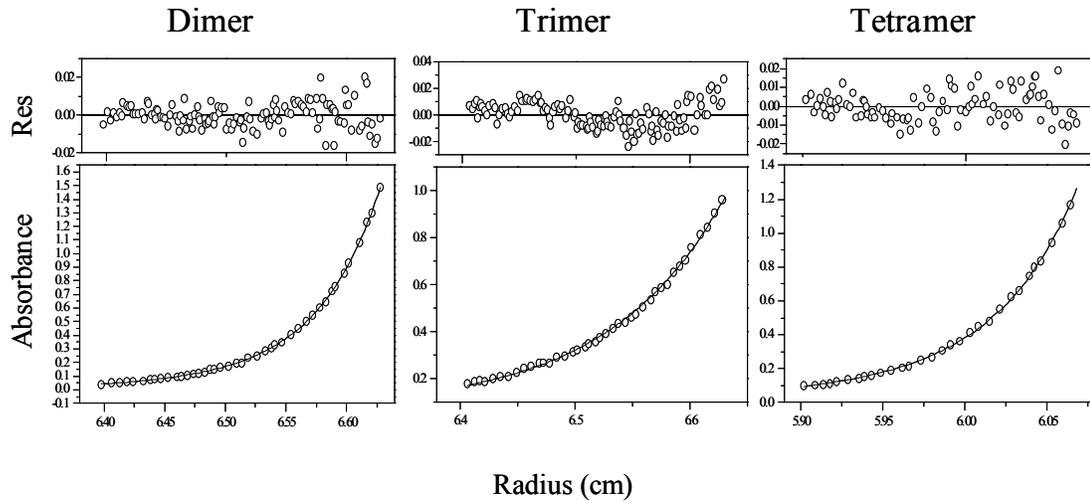


Figure 7 – Representative equilibrium sedimentation scans for each of the three engineered oligomers and corresponding residuals to global fit. The two species fits are shown as solid lines, for clarity only every third data point is shown on the fitted curve.

Re-entry of scaffold fusion into empty shells

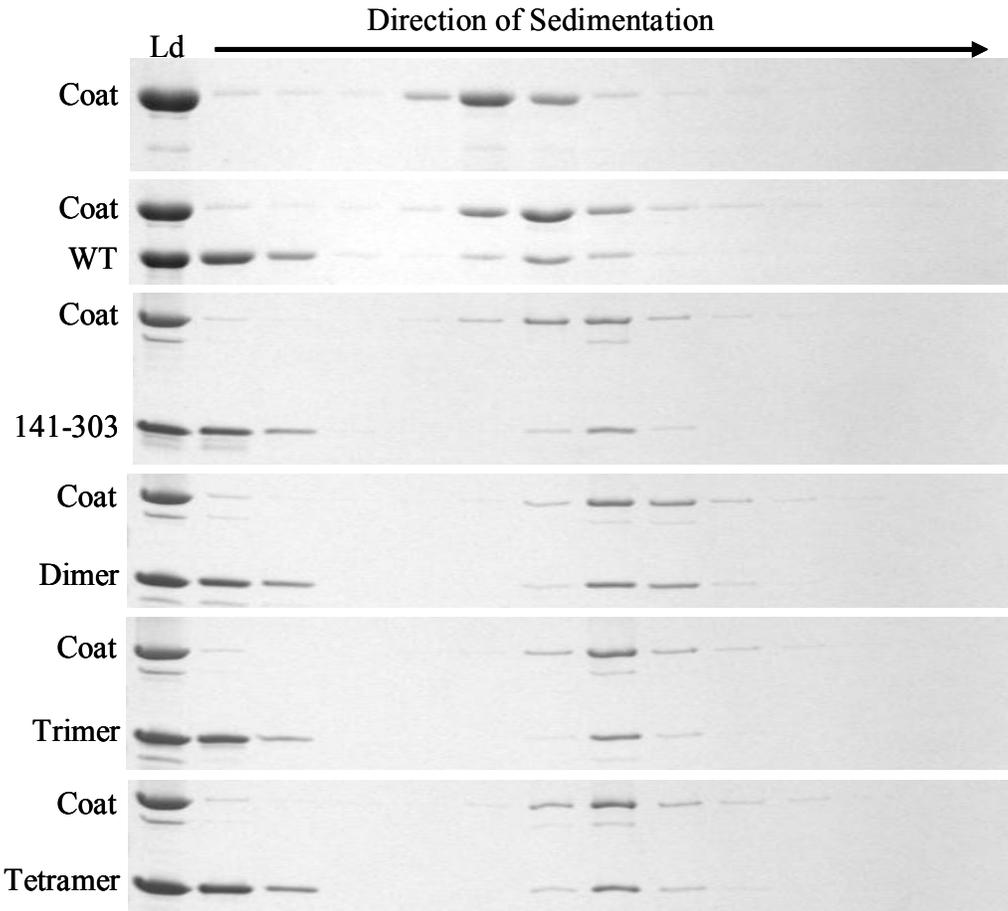


Figure 8 - Re-entry of fusion peptides into empty shells. The increased sedimentation rates of the fragment and fusion filled particles versus those filled by wild type scaffold is due to the mass increase of the refilled particles. At 6-700 peptides bound per shell, this represents a 3.5 - 6.0 MDa increase in total particle mass. Scaffold fusions were added to solutions of empty shells in 3-fold molar excess. Gradients from top to bottom: Wild type, 141-303, dimer, trimer, tetramer. Gradient loads are shown in the first fraction on the left.

Re-entry protects wild type scaffold from proteolysis

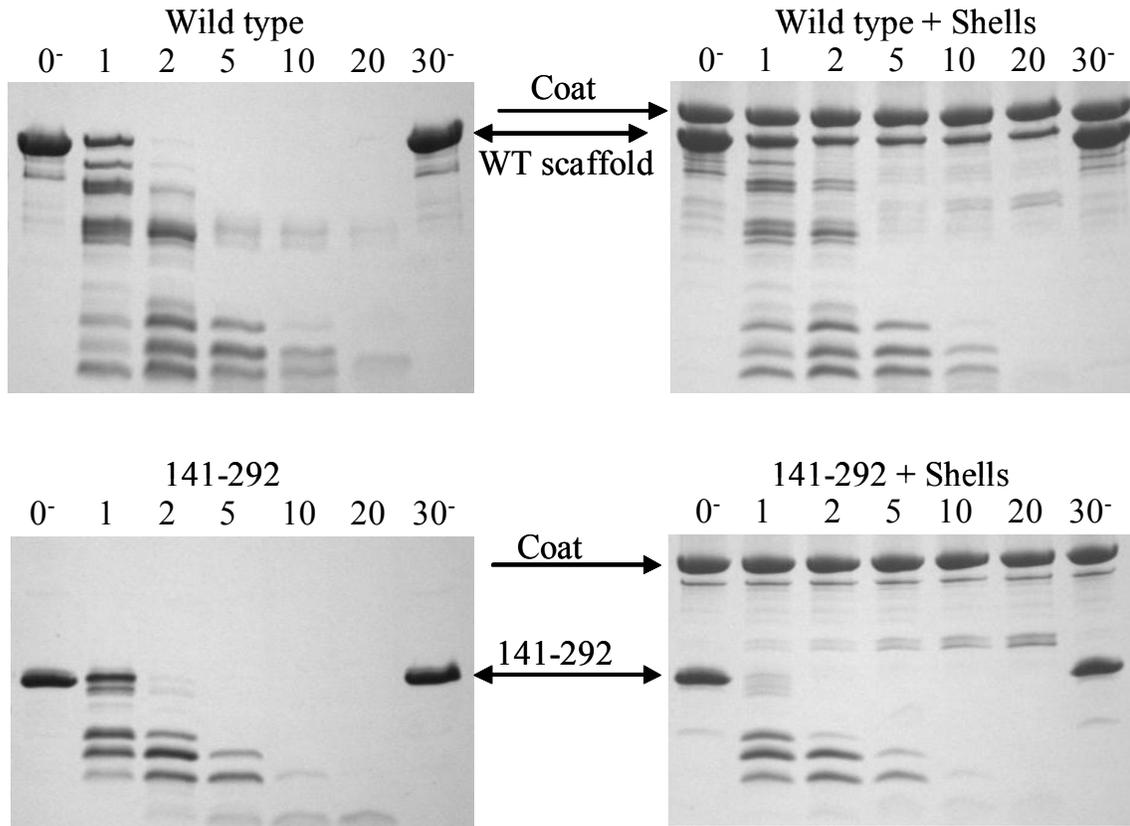


Figure 9 – Protection of wild type scaffold from trypsin digestion due to shell re-entry. The top left panel shows the rapid and complete digestion of purified wild type scaffold. Top right shows the protection of wt scaffold from digestion due to shell re-entry. The bottom panels show the results of similar digestions of a scaffold fragment which lacks the coat binding domain and does not re-enter the shells. The pattern and kinetics of digestion for this fragment are nearly identical in both gels, indicating that protection is due to complete re-entry. For each image, the time course of trypsin digestion is labeled at the top of the gel. Shown also are the non-digested (- trypsin) reaction mixtures at 0 and 30 minutes at the experimental temperature (37°C).

R74C dimers are not protected from proteolysis by shell binding

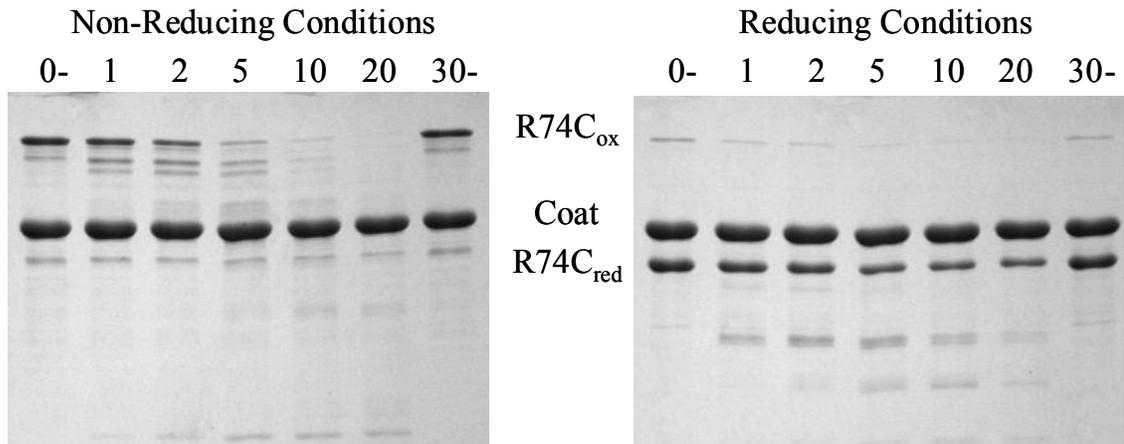


Figure 10 – Digestion of R74C oxidized dimer due to incomplete re-entry. The left panel shows the rapid digestion of the oxidized dimer after a 1-hour pre-incubation with empty shells. Right panel: Incubation of the same mixture with β -Mercaptoethanol prior to digestion reduces the dimer and leads to protection from trypsin digestion, suggesting that the cystine crosslink blocks re-entry.

Scaffold fusion proteins are protected from proteolysis by re-entry (1)

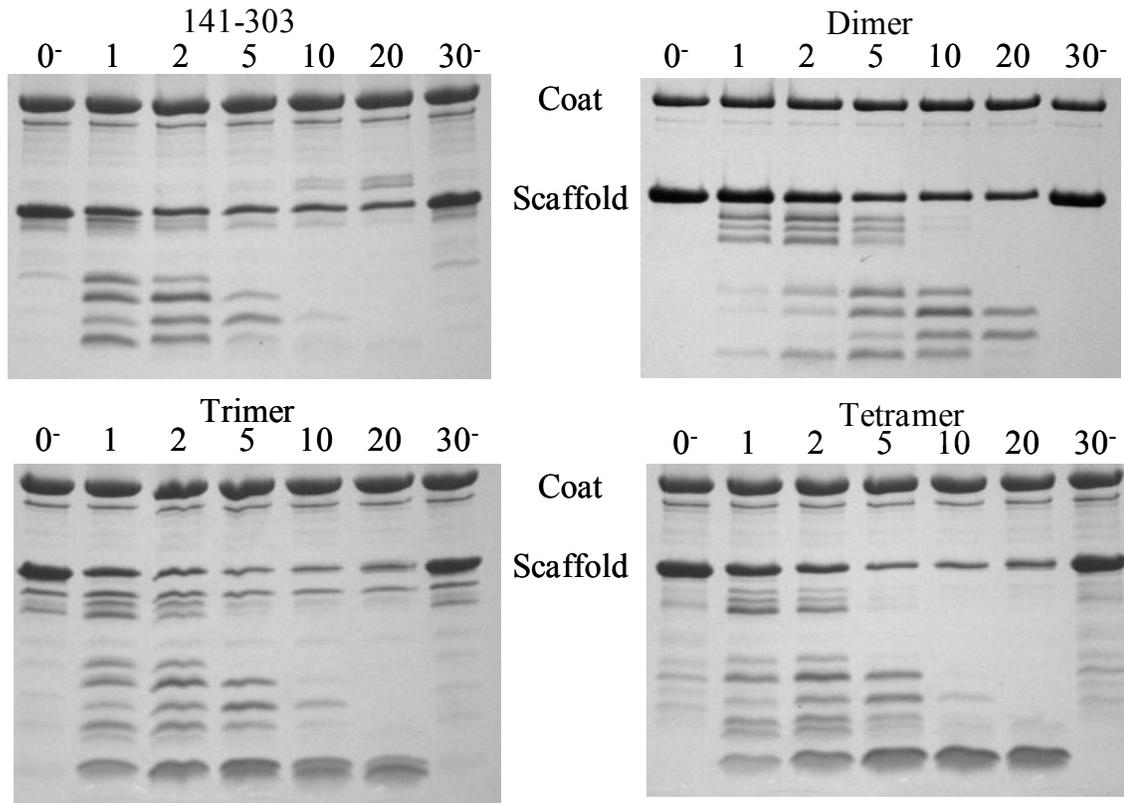


Figure 11 – Protection of the 141-303 fragment and corresponding fusion peptides from trypsin digestion. For each image, the time course of trypsin digestion is labeled at the top of the gel. Shown also are the non-digested (- trypsin) reaction mixtures at 0 and 30 minutes at the experimental temperature (37°C). In the absence of empty shells, the fusion peptides are completely digested after 10 minutes of incubation (data not shown).

Scaffold fusion proteins are protected from proteolysis by re-entry (2)

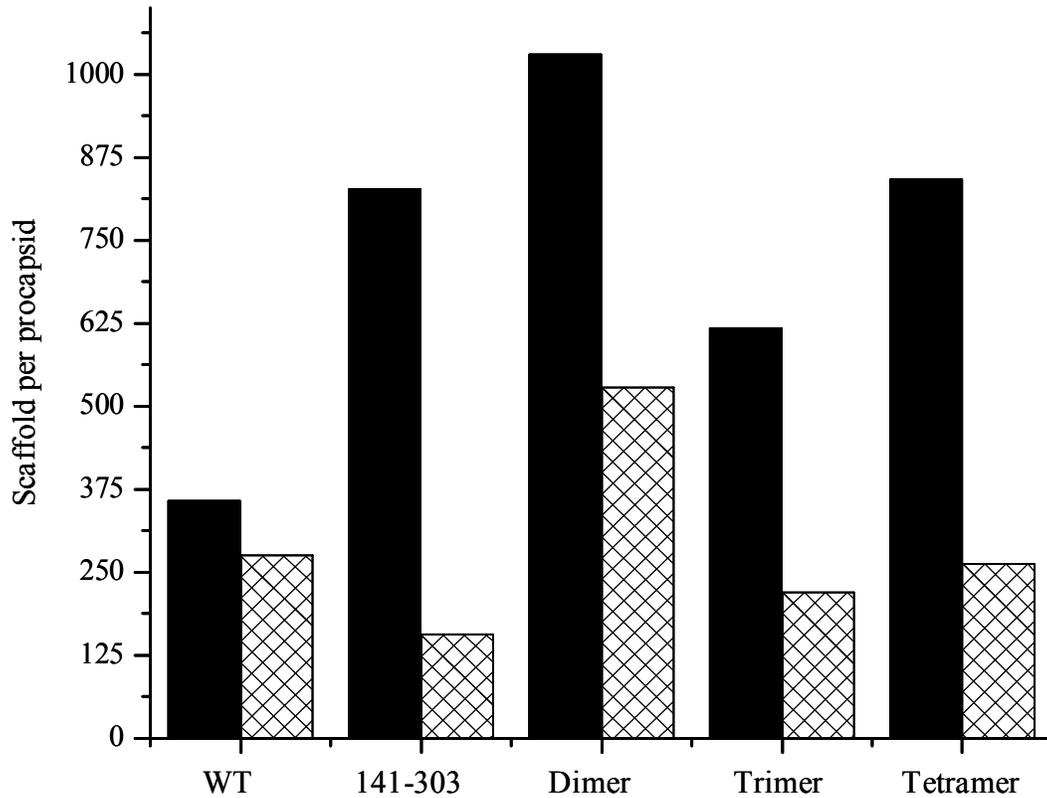


Figure 12 – Protection of scaffold from tryptic digestion in refilled shells. Empty P22 shells were incubated with excess scaffolding constructs. The refilled shells were harvested by pelleting and subjected to a 30 minute tryptic digestion as described in materials and methods. The digestions were spun to 100s and the pellets resuspended in 2X Tris-Tricine sample buffer. Stoichiometry was determined by scanning the intensity of the coomassie stained gel bands. The darkened columns represent scaffold molecules per procapsid before digestion. The crosshatched bars represent scaffold molecules per procapsid after digestion.

Scaffold fusion proteins promote coat polymerization (1)

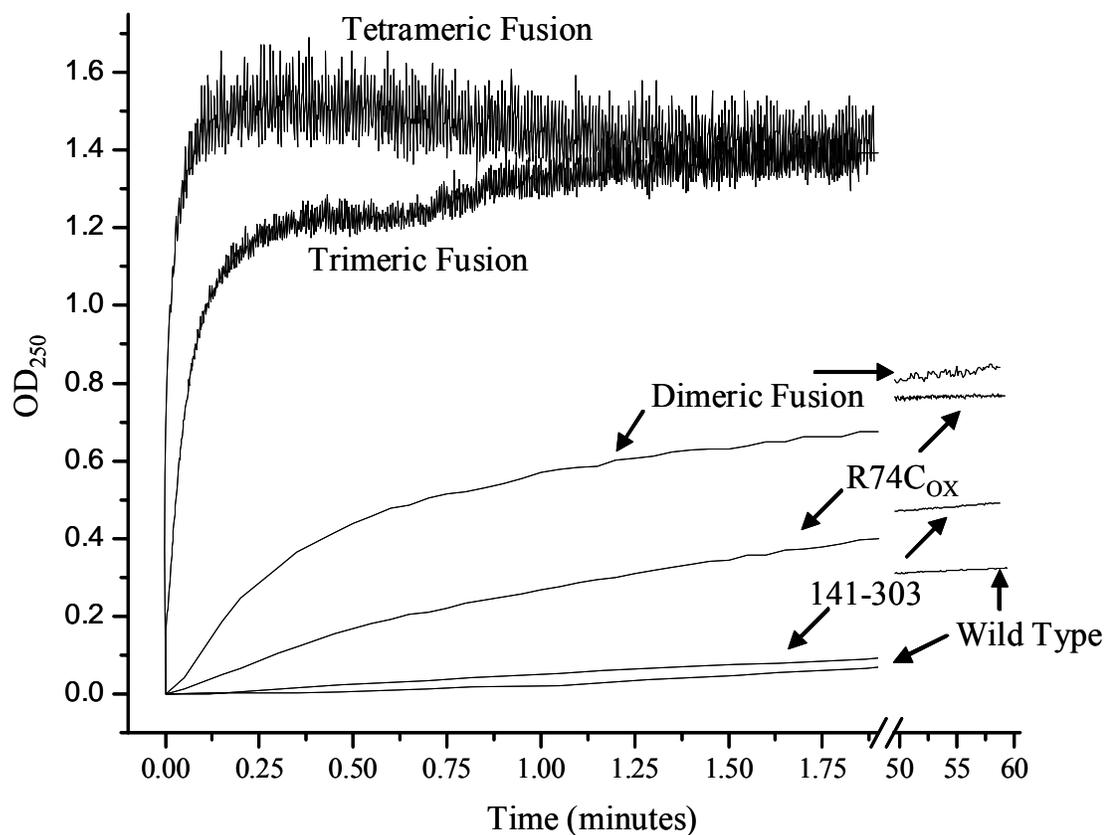


Figure 13 - Dependence of initial assembly rates and final amount of turbidity as a function of engineered valency of the scaffolds. Fusion proteins were titrated into thermostatted cuvettes containing 14 μ M monomeric coat protein. At these concentrations, reactions catalyzed by the tetrameric and trimeric fusions are complete after 2 minutes, while the dimeric fusions require 6 minutes. The 141-303 fragment, R74C_{ox} dimers and wild type are >90% complete after 1 hour. Kinetics of assembly catalyzed by all 6 scaffolds at 10 μ M are shown.

Scaffold fusion proteins promote coat polymerization (2)

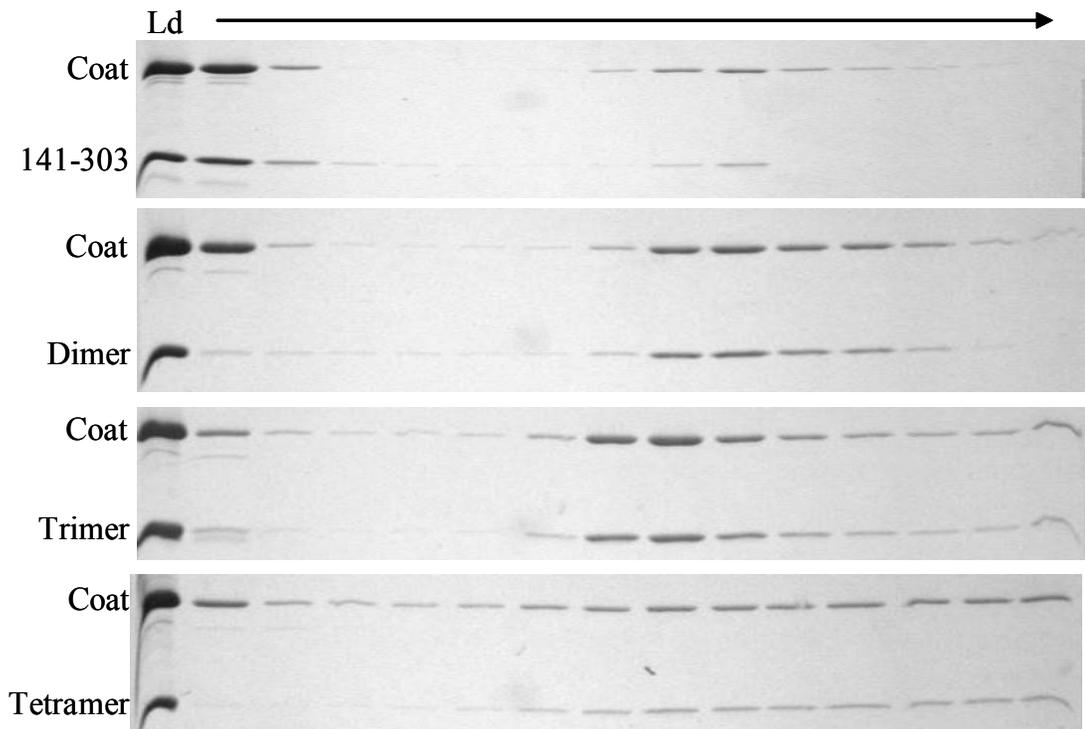


Figure 14 - Partices assembled by fusion peptides at wild type stoichiometric (5 Scaffold : 7 Coat) ratios of GCN4 fusions. The peak fractions of the 141-303 peptide and the dimeric fusion are coincidental, while the trimeric fusion generates a peak which migrates slightly less rapidly. The tetrameric fusion generates a smear of coat and scaffold which comigrates through the entire gradient and is indicative of large, aberrant structures.

Scaffold fusions are stripped from some assembled products

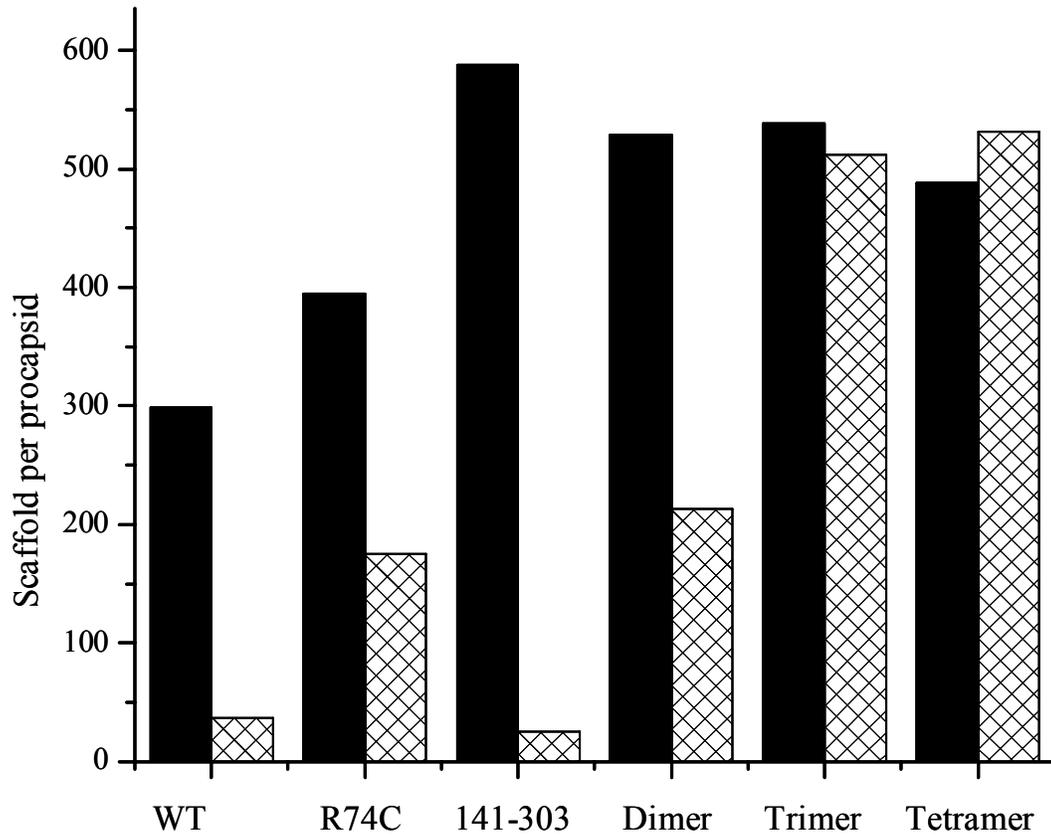


Figure 15 - Stoichiometry of assembled products after ‘stripping’. Assembled products were incubated for 1 hour in 700 mM Gdn•HCl, conditions which normally remove >85-90% of wt scaffold from procapsids assembled *in vivo*. Stoichiometry is expressed as the number of scaffolding molecules per 420 coat molecules (per procapsid). Note the increase in the number of scaffolding proteins as the scaffold ‘valency’ is increased for full length and the truncated scaffolds.

Morphology of products assembled by scaffold fusions

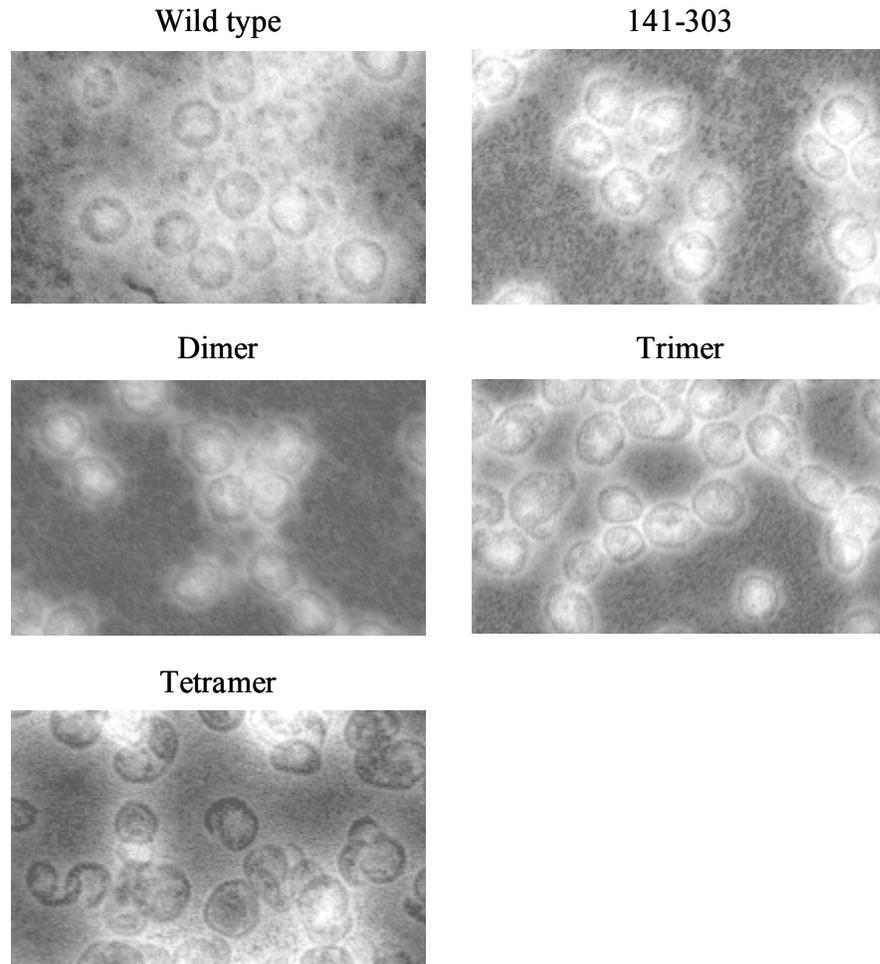


Figure 16 - Morphology of particles assembled at wild type scaffold : coat ratios. Shown above are representative regions of uranyl acetate-stained reaction products. Note the increase in aberrant morphologies in the case of the trimeric and tetrameric reaction products.

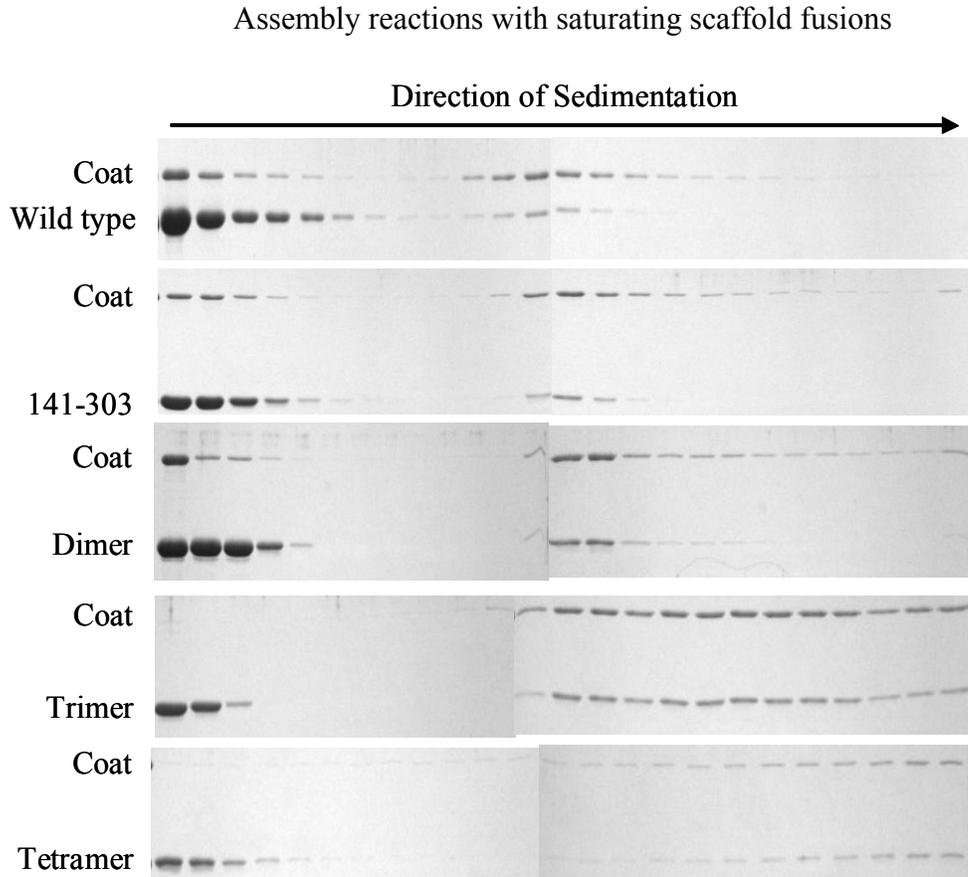


Figure 17 - Particles assembled at saturating scaffold concentrations by all constructs.

The smearing of the coat protein towards the bottom of the gradient indicates large, aberrant structures. The smear is present in the reactions initiated by the unfused fragment and the dimeric peptide, but represents less than half of the total product. The trimeric fusion generates this pattern, indicating that most of the reaction product is aberrant when this construct is supplied saturating amounts. The products assembled by the tetrameric scaffold have nearly completely migrated through the entire gradient. Note also the absence of unassembled coat protein in the reactions initiated by the trimeric and tetrameric scaffolds

Wt and trimeric scaffold reaction intermediates

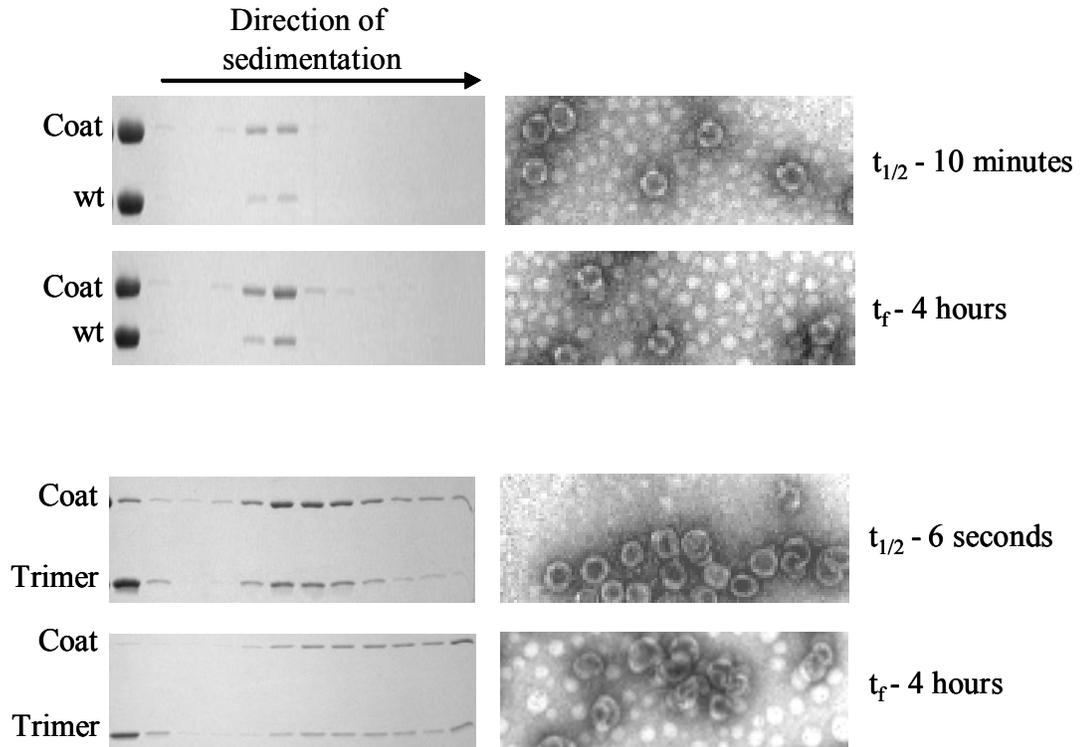


Figure 18 – Assembly intermediates of reactions initiated *in vitro* by wt and trimeric scaffolds. Above: Sucrose gradients and negative stain of products of *in vitro* assembly by wt scaffold at 10 minutes and 4 hrs. Below: Sucrose gradients and negative stain of products of *in vitro* assembly by trimeric scaffold at 6 seconds and 4 hrs.

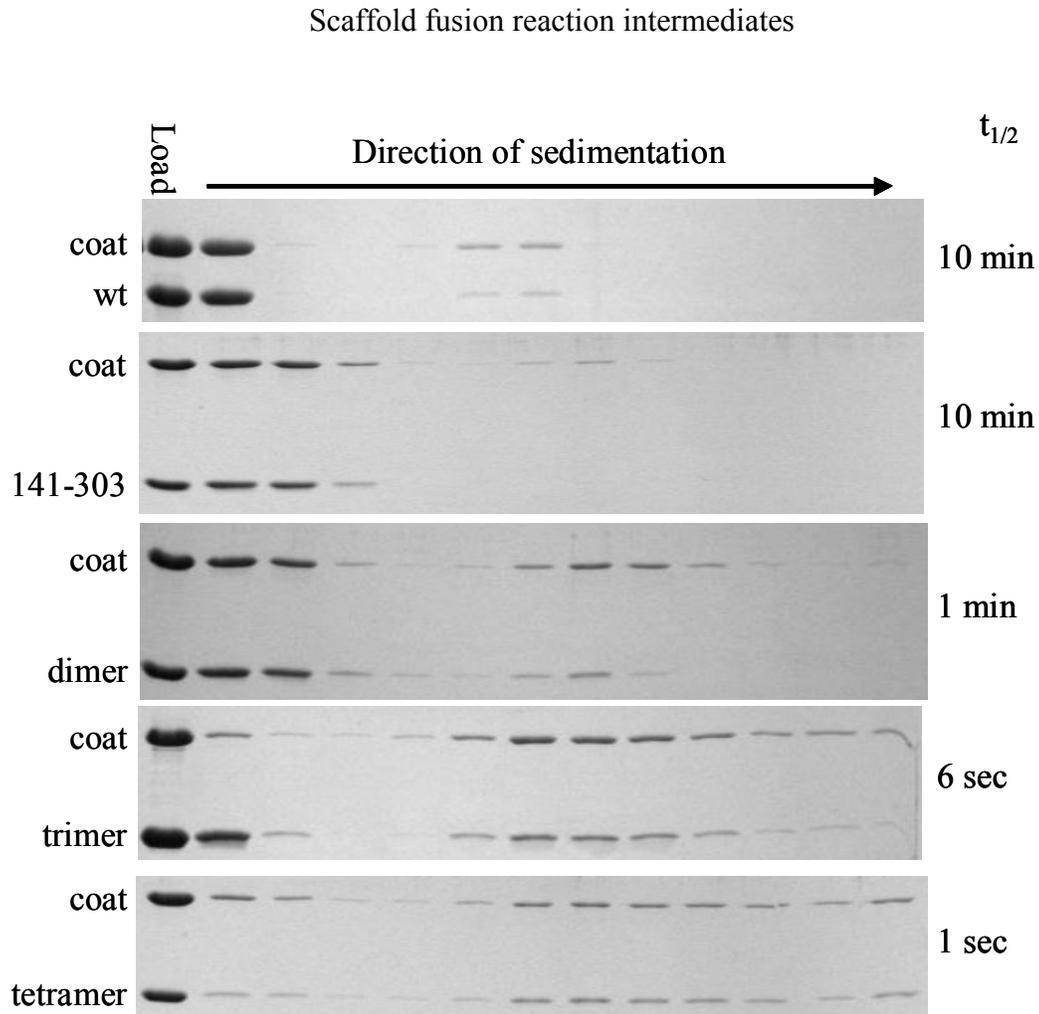


Figure 19 - Assembly intermediates for each scaffold construct. Scaffold was titrated into solutions of monomeric coat protein at equimolar final concentrations. Reactions were quenched by the addition of NaCl to 200 mM final at the time points indicated to the right of the gel. Note the nearly complete disappearance of the coat protein from the top of the gradients in the trimeric and tetrameric reactions. Note also the appearance of the rapidly sedimenting structures in the gradients.

DISCUSSION

The Role of Dimeric Scaffold

Monomeric P22 coat protein cannot be concentrated beyond 1 mg/ml (21.5 μ M) without generating large structures similar to the aberrants observed *in vivo* by King et al.^{3,17}. Presumably, below this concentration interactions between coat protein monomers are unfavorable for nucleation. Addition of wild type scaffold lowers this concentration of coat protein to approximately 4 μ M and directs polymerization toward the correct T=7 morphology. The experiments in this work demonstrate that full length dimeric scaffold further reduces the critical concentration 50 - 75% without comprising product morphology. In addition, we were able to demonstrate that these dimers must be bivalent. This suggests that scaffold dimers are binding two coat monomers and stabilizing a coat-scaffold heterotetramer. Covalent dimers eliminate the scaffold dimerization equilibrium, which likely permits the formation of more heterotetramers, and as a result, critical concentration is lowered. The heterotetramers may then provide a platform for proper procapsid nucleation. This behavior can be explained by the 'entropy sink' model proposed by Parker, Casjens and Prevelige²⁴. They suggest that the coat-scaffold binding event stabilizes coat oligomers by the lowering the entropic energy barrier associated with coat polymerization via scaffold dimerization.

Our results from biophysical experiments performed on the fusion proteins showed them to be 50 – 95% oligomeric in the concentration ranges utilized in this work. Our results from the CD experiments demonstrated however, that the trimeric and tetrameric scaffold were not natively folded near the C-terminus. This was not completely unexpected and we believe that this result is informative in its own right. Primarily we interpret this to indicate that the wild type scaffold tetramer is likely not a dimer of dimers with two-fold symmetry. If it were, the secondary structure alterations we noted at the C-terminus of the tetrameric fusion protein should not have occurred, as the coat binding domains would have presumably been in native environments. Thumann-Commike et. al. proposed that the tetrameric species, if it did not possess two-fold symmetry, would impose the correct local interactions to force a T=7 icosahedral lattice during growth³⁶. They proposed that a scaffold tetramer which possessed this type of two-fold symmetry would force lattice growth toward T=9. We were unable to confirm this hypothesis due to rapid nucleation and the formation of strictly aberrant particles by the preformed scaffold tetramers. It is also possible that the coat protein's internal binding angles and interface positions forbid the formation of lattices larger than T=7, since this is the largest closed structure observed with P22 coat protein.

Contribution of Trimeric and Tetrameric Scaffold to Assembly

The experiments in this work with engineered scaffold oligomers showed that increasing the local concentration of coat protein leads to increased nucleation and enhancements in polymerization kinetics. By increasing the number of coat protein subunits bound to scaffold in a small space, the energetic cost of coat nucleation is

notably reduced. Indeed, reactions driven by trimeric and tetrameric scaffolds lacked any observable critical concentration of coat protein, an indication that nucleation was favored over lattice growth. In the context of viral capsid assembly this was presented by Berger et. al. as the local rules hypothesis⁵¹. In this model, subtle conformational changes of coat protein subunits which occur via local interactions during nucleation have meaningful consequences on the morphology of the final product. *In silico*, Schwartz et. al. demonstrated that slight reductions in energetic barriers toward nucleation led to enhanced kinetics with a concomitant loss of fidelity⁵². Our data would tend to support this model by demonstrating that pre-formed trimeric and tetrameric scaffolds create areas of high local concentration which perilously reduce the energetic penalty of nucleation. This could lead to over-nucleation and according to this model, would be detrimental to fidelity.

A Possible Role for the N-Terminus

As Parker, et. al. observed, the products of *in vitro* assembly generated by the 141-303 fragment become increasingly aberrant as its concentration is increased²⁴. This behavior was attributed to this fragment's poor dimerization. By fusing the GCN4 dimerization domain to the N-terminus of this peptide, we increased its dimeric association constant 600 fold. Assembly reactions with this protein showed a marked increase in kinetics however, the expected 'rescue' of particle morphology was not observed. This result indicates that the N-terminus may have an active role controlling the morphology of P22 procapsids, although there is no evidence that this domain of the scaffold interacts directly with the coat protein. However, there is evidence that the N-terminus of the protein is responsible for oligomer formation²². The 141-303 fragment

used in this work forms dimers approximately 10 fold less efficiently than wild type and does not form tetramers²². Alternatively, the N-terminus may act to alter the disposition of the coat binding domains in the dimeric state. It is impossible, with our results, to eliminate contributions by the tetrameric form of P22 scaffold during lattice growth. Specifically, we cannot rule out the possibility that non-symmetric scaffold tetramers form during critical steps during growth which force the capsid toward T=7.

Model of Scaffold-Assisted Procapsid Assembly

The observation that covalent scaffold dimers lower the critical concentration of coat protein required for assembly of procapsids *in vitro*, suggests that scaffold dimers are involved in the nucleation step^{31,44}. Thumann-Commike et. al., presented a model of assembly that suggested trimeric clusters of coat proteins provided critical interactions for procapsid growth around the pentameric vertices³⁶. They proposed the nucleus to be a pentameric coat cluster, to which scaffold-coat heterotetramers would bind^{31,36}. The data in this work indicates that dimeric scaffolding protein may encourage nucleation of these vertices by creating a platform for a third subunit to bind. This third subunit would be one of the five which makes up the pentameric vertex. This would make the rate limiting factor in procapsid assembly the formation of nuclei composed of 15 coat subunits and a maximum of 10 scaffolding subunits assembled as a pentameric cluster of coat protein trimers, stabilized by scaffold dimers. A model of this assembly pathway is represented in figure 20.

A second aspect of their model was the scaffold interaction with the growing lattice and the ‘help’ scaffold provided during ‘difficult’ steps in assembly by providing

the energy necessary to promote stable coat-coat interactions via scaffold tetramer formation³⁶. Our evidence also suggests that dimeric scaffold interactions are necessary, but not sufficient for proper control of P22 procapsid assembly. This more subtle aspect of scaffold function almost certainly resides in the N-terminus of the protein. Figure 20 depicts how Thumann-Commike et. al. proposed that tetrameric scaffold could stabilize coat : coat interactions and direct growth toward T=7 lattice formation³⁶. The pathway is described using the nomenclature of the quasi-equivalent conformations as they were assigned in their model. The dimeric scaffold : scaffold interactions stabilize the a, b, g coat trimer which forms the first stable building block of the pentameric procapsid nucleus³¹. Subsequent binding events by the c and f subunits are unstable until scaffold binds, it at this step that tetramer formation would be critical. Without the stabilizing influence of the asymmetric scaffold tetramer these subunits could adopt conformations which lead to the formation of a T=4 lattice. If this happens, growth would proceed towards T=7 in one direction and T=4 in another and an aberrant spiral would result. The tetrameric scaffold stabilized subunits would be energetically compelled to complete the hexon, thus ensuring the growth of the T=7 lattice.

Proposed model of scaffold assisted procapsid assembly

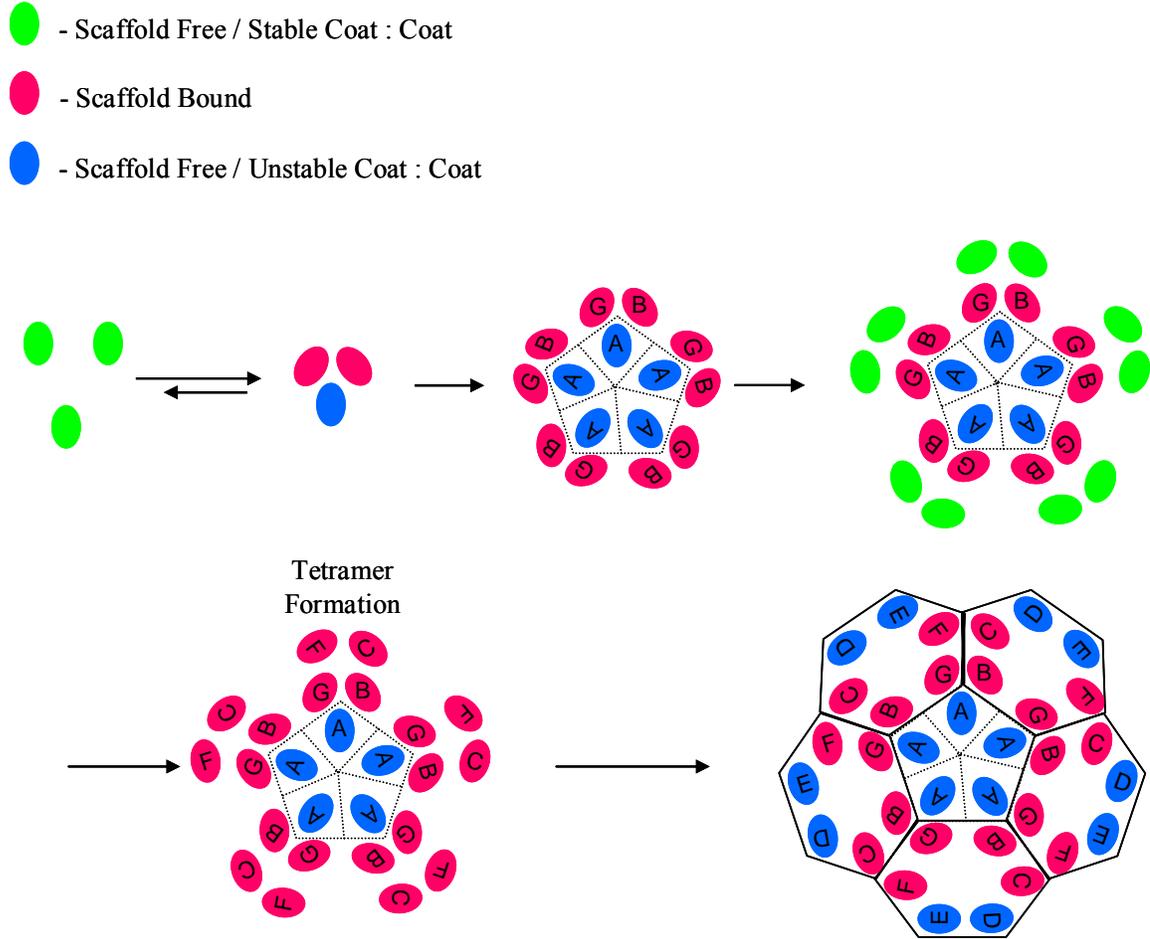


Figure 20 – Proposed assembly pathway of the P22 procapsid. Free monomeric coat protein and scaffold form a heteropentamer which forms the pentameric vertex. The addition of the subunits adjacent to the B and G subunits is unstable until scaffold binds and forms a stable tetramer as a dimer of dimers. Once the tetramer is formed, the subunits are locked into the C and F conformations and provide a stable platform for the addition of the D and E subunits, which complete the hexons.

MATERIALS AND METHODS

Mutagenesis

Plasmids encoding wt scaffold (res. 1-303), R74C and R74C292 were generously supplied by Peter Weigele of the Casjens lab. The oligomeric fusion peptides were created by introducing a Kpn1 site into a pET 3a vector containing the 141-303 fragment²⁴. The Kpn1 site was introduced in frame immediately downstream of residue 141. The coding region for the wt GCN4 dimerization domain was donated by Don Wiley. The coding sequence was cloned from this background and while introducing an Nde1 site immediately upstream of the coding region and a Kpn1 site in frame at the end of the coding region. The insert was digested and ligated with a pre-digested pET 3a (Kpn1 / Nde1) 141-303 background. The ligated DNA was transformed into competent DH5 α and BL21/pLysS cells as described previously. This created a cassette for the insertion of the inserts coding for the trimeric and tetrameric fusion peptides. These were donated by Peter Kim and Andrea Cochrane respectively, and were cloned from the supplied parent vectors similarly. For the trimeric peptide, the primers were 5'-GGGAGACCACAACGGTTTC-3' and 5'-GAGTGGGGTACCGCGTGCTTCAC-CGATCAGT-3', for the tetrameric fusion the primers were 5'-GAGACCACAACGGTTTCCC-3' and 5'-ATATTAGGTACCACGTTACCAAC-TAGTTTTTTTA-3'. The amplified products were digested, ligated and transformed as described above.

Protein Expression and Purification

All scaffold constructs were overexpressed in *e. coli* BL21/pLysS cells and induced with 0.5 mM IPTG. Cell pellets were resuspended in 50 mM Tris-Cl, pH7.6, 25 mM NaCl, 2 mM EDTA (buffer B) (30 ml / L culture) and frozen at -80°C. Cells were lysed by thawing at room temperature and adding lysozyme to 0.1 mg/ml and PMSF to 1.0 mM. Lysates of the wt, R74C, R74C Δ tail, G266C, G251C, Δ 216-251 and Δ 216-234 were spun to 2500s in a Sorvall ss-34 rotor at 4°C, the supernatants were then spun to 75s in a Beckman Ti-42 rotor at 4°C. The clarified lysates were loaded onto a 5 ml Pharmacia Hi-Trap Q-sepharose (Anion exchange) column equilibrated in buffer B and eluted with linear NaCl gradients between 0 and 350 mM NaCl. The pooled fractions were dialyzed versus 50 mM NaOac, pH 5.5, 2 mM EDTA (buffer A) and loaded onto a 5 ml Pharmacia Hi-Trap SP-sepharose (cation exchange) column equilibrated in buffer A. These were eluted with linear NaCl gradients between 0 and 350 mM NaCl.

6.0 M Urea was added to the lysates of the oligomeric fusions which were centrifuged as described above. The clarified lysates were loaded onto a 5 ml Pharmacia SP-sepharose column equilibrated in buffer A and 6.0M Urea and eluted with a linear gradients from 0 to 1.0 M NaCl. Pooled fractions were refolded by dialysis against buffer B then loaded onto a 5 ml Pharmacia Hi-Trap SP-sepharose column and eluted with linear gradients from 0 to 1.0 M NaCl.

Bivalent Dimer Preparation

The heterodimeric R74C / R74C Δ tail was created by batch oxidizing purified R74C and R74C Δ tail overnight at room temperature in buffer B (20-22°C). The mixture was separated by elution from a 5 ml Pharmacia Hi-Trap Q column. This separated monomeric (reduced) scaffold from the dimeric (oxidized) scaffold. The dimeric fraction was then dialysed versus buffer A and eluted from a Pharmacia Hi-Trap SP-sepharose column equilibrated in buffer A with a linear gradient from 50 to 150 mM NaCl.

Empty Procapsid Shells

Empty procapsid shells were obtained from procapsids generated by packaging deficient phage in *s. typhimurium*. Cell lysates were clarified by centrifugation to 2500s, the supernatant was centrifuged to 75s over a 35% sucrose cushion. The procapsids were resuspended in buffer B and diluted with an equal volume of 1.0 M GdnHCl and centrifuged to 75s. This procedure was performed twice. The shells were resuspended in buffer B and stored at 4°C.

Monomeric Coat Protein

Monomeric coat protein was harvested from a plasmid encoding the portal scaffold and coat genes cloned from wild type phage and bearing the R74C mutation in the scaffolding protein. These 'procapsids' were harvested and stripped as wt procapsids with the addition of 1.0 mM 2-Mercaptoethanol to reduce scaffold dimers. The resultant shells were denatured in 6.0 M GdnHCl and dialyzed vs. buffer B. The monomeric coat

protein was further purified by elution from a 5-ml Pharmacia Hi-Trap Q-sepharose column from 100 -300 mM NaCl.

Shell Binding Reactions

Empty shells were titrated into cuvettes containing buffer B thermally equilibrated to 20°C so that the final concentration in a 400 μ L reaction would be 10 μ M. Scaffolding protein constructs which had been warmed to room temperature were titrated to concentrations between 20 and 50 μ M final. Re-entry kinetics were monitored at 250 nm at intervals between 5 and 20s. Reaction mixtures were allowed to incubate for 4 hours prior to being centrifuged through 5%-20% sucrose gradients at 30,000 RPM for 30 m in a Beckman SW-55 rotor ($s_{app} = 325$).

Procapsid Assembly Reactions

Monomeric coat protein obtained as described above was dialysed to remove NaCl at 4°C overnight and titrated into thermally equilibrated cuvettes containing buffer B at concentrations indicated in the text (typically 10 or 14 μ M final). Scaffolding protein constructs, warmed to room temperature, were titrated to final concentrations indicated in the text. Kinetics of assembly were monitored at 250 nm over various lengths of time (2 min to 60 min). Readings were taken over a range of 100 ms at 2 min reaction times up to every 20 s for 60 minute reactions. Reactions were allowed to incubate for 4 hours prior to sucrose gradient centrifugation.

Trypsin Digestions

Trypsin digestions were performed at 37°C at a scaffold : trypsin molar ratio of 250 : 1. The reaction mixtures were incubated for > 1 hr. at room temperature prior to digestion. Samples were removed from the reactions at the indicated time points and immediately quenched by boiling in 2X Tris-Tricine sample buffer, or by the addition of PMSF to 1.0 mM as indicated in the text.

Circular Dichroism Measurements

Scaffold constructs were exhaustively dialysed vs 50 mM NaH₂PO₄, pH 7.0, 25 mM NaCl. The samples were diluted to 10, 5, and 2.5 μM for spectra collection. CD spectra were collected at 20°C from 260 to 185 nm with a 0.5 nm step on an Aviv 62DS CD spectrometer using a pathlength of 1.0 mM, bandwidth of 2.0 nm, and averaging time of 10 s. Deconvolution of the spectra was done according to the method of Chang et. al.⁴⁹

Analytical Ultracentrifugation

Velocity sedimentation experiments were done in a Beckman Optical XLA Analytical ultracentrifuge at 59,000 RPM in a 4-hole rotor. Due to a lack of tryptophan residues, sample absorbances were monitored at 230 nm. The extinction coefficient at this wavelength was determined to be $1.00 \times 10^5 \text{ L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$. Data was analyzed using sedfit 8.4. The fusion protein samples were diluted from the S-200 elution peaks to a concentration of 5.0 μM.

Equilibrium sedimentation was performed in the same centrifuge using a four hole rotor and 6-channel centerpieces. The proteins were sedimented in 50 mM NaH_2PO_4 , pH 7.0 and 500 mM NaCl. The salt was added to reduce non-ideality in previous centrifuge runs observed for the trimeric and tetrameric species. Addition of salt had no effect on the results obtained with the dimeric scaffold. Again, the absorbance was monitored at 230 nm in order to more accurately determine the oligomeric nature of the proteins over the concentration range used in our experiments. The proteins were sedimented to apparent equilibrium at concentrations of 9, 4.5 and 1.5 μM at speeds of 10, 16 and 19 kRPM. Data was analysed with Origin 4.0 (Microcal).

REFERENCES

1. Caspar, D.L.D., & Klug, A. (1962). Principles in the construction of regular viruses. *Cold Spring Harbor Symposia on Quantitative Biology*, 27, 1-24.
2. Johnson, J.E. (1996). Functional implications of protein-protein interactions in icosahedral viruses. *Proc Natl Acad Sci U S A*, 93(1), 27-33.
3. King, J., Lenk, E.V., & Botstein, D. (1973). Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. II. Morphogenetic pathway. *J Mol Biol*, 80(4), 697-731.
4. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(259), 680-5.
5. Showe, M.K. & Black, L.W. (1973). Assembly core of bacteriophage T4: An intermediate in head formation. *Nat. New Biol.*, 242, 70-75.
6. Casjens, S., & King, J. (1974). P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly. *J Supramol Struct*, 2(2-4), 202-24.
7. Ray, P., & Murialdo, H. (1975). The role of gene Nu3 in bacteriophage lambda head morphogenesis. *Virology*, 64(1), 247-63.
8. Teschke, C.M., King, J., & Prevelige, P.E., Jr. (1993). Inhibition of viral capsid assembly by 1,1'-bis(4-anilinonaphthalene-5- sulfonic acid). *Biochemistry*, 32(40), 10658-65.
9. Israel, J.V., Anderson, T.F., & Levine M. (1967). In vitro morphogenesis of phage P22 from heads and base-plate parts. *Proc Natl Acad Sci U S A*, 57(2), 284-91.
10. Botstein, D., Chan, R.K., & Waddell, C.H. (1972). Genetics of bacteriophage P22. II. Gene order and gene function. *Virology*, 49(1), 268-82.
11. Botstein, D., Waddell, C.H., & King, J. (1973) Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. I. Genes, proteins, structures and DNA maturation. *J Mol Biol*, 80(4), 669-95.

12. Earnshaw, W., Casjens, S., & Harrison, S.C. (1976). Assembly of the head of bacteriophage P22: X-ray diffraction from heads, proheads and related structures. *J Mol Biol*, 104(2), 387-410.
13. Israel, V. (1977). E proteins of bacteriophage P22. I. Identification and ejection from wild-type and defective particles. *J Virol*, 23(1), 91-7.
14. Bazinet, C., & King, J. (1985). The DNA translocating vertex of dsDNA bacteriophage. *Annu Rev Microbiol*, 39, 109-29.
15. Strauss, H., & King, J. (1984). Steps in the stabilization of newly packaged DNA during phage P22 morphogenesis. *J Mol Biol*, 172(4), 523-43.
16. Sauer R.T., Krovatin W., Poteete A.R., & Berget P.B. (1982). Phage P22 tail protein: gene and amino acid sequence. *Biochemistry*, 21(23), 5811-5.
17. Earnshaw, W., & King, J. (1978) Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein. *J Mol Biol*, 126(4), 721-47.
18. Bazinet C., Benbasat J., King J., Carazo J.M., & Carrascosa J.L. (1988). Purification and organization of the gene 1 portal protein required for phage P22 DNA packaging. *Biochemistry*, 27(6), 1849-56.
19. Casjens, S., & Huang, W.M. (1982). Initiation of sequential packaging of bacteriophage P22 DNA. *J Mol Biol*, 157(2), 287-98.
20. Moore, S.D., & Prevelige, P.E., Jr. (2002). Bacteriophage P22 portal vertex formation in vivo. *J Mol Biol*, 315(5), 975-94.
21. Tuma, R., Prevelige, P.E., Jr., & Thomas, G.J., Jr. (1998). Mechanism of capsid maturation in a double-stranded DNA virus. *Proc Natl Acad Sci U S A*, 95(17), 9885-90.
22. Tuma, R., Parker, M.H., Weigele, P., Sampson, L., Sun, Y., Krishna, N.R., et al. (1998). A helical coat protein recognition domain of the bacteriophage P22 scaffolding protein. *J Mol Biol*, 281(1), 81-94.
23. Parker, M.H., & Prevelige, P.E., Jr. (1998). Electrostatic interactions drive scaffolding/coat protein binding and procapsid maturation in bacteriophage P22. *Virology*, 250(2), 337-49.
24. Parker, M.H., Casjens, S., & Prevelige, P.E., Jr. (1998). Functional domains of bacteriophage P22 scaffolding protein. *J Mol Biol*, 281(1), 69-79.

25. Sun, Y., Parker, M.H., Weigele, P., Casjens, S., Prevelige, P.E., Jr., & Krishna, N.R. (2000). Structure of the coat protein-binding domain of the scaffolding protein from a double-stranded DNA virus. *J Mol Biol*, 297(5), 1195-202.
26. Parker, M.H., Stafford, W.F., & Prevelige, P.E., Jr. (1997) Bacteriophage P22 scaffolding protein forms oligomers in solution. *J Mol Biol*, 268(3), 655-65.
27. Prevelige, P.E., Jr., Thomas, D., & King, J. (1988). Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro. *J Mol Biol*, 202(4), 743-57.
28. Casjens, S. (1979). Molecular organization of the bacteriophage P22 coat protein shell. *J Mol Biol*, 131(1), 1-14.
29. Fuller, M.T., & King, J. (1982). Assembly in vitro of bacteriophage P22 procapsids from purified coat and scaffolding subunits. *J Mol Biol*, 156(3), 633-65.
30. Greene, B., & King, J. (1994). Binding of scaffolding subunits within the P22 procapsid lattice. *Virology*, 205(1), 188-97.
31. Prevelige, P.E., Jr., Thomas, D., & King, J. (1993). Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophys J*, 64(3), 824-35.
32. Darke, P.L., Cole, J.L., Waxman, L., Hall, D.L., Sardana, M.K., & Kuo, L.C. (1996). Active human cytomegalovirus protease is a dimer. *J Biol Chem*, 271(13), 7445-9.
33. Morais, M.C., Kanamaru, S., Badasso, M.O., Koti, J.S., Owen, B.A., McMurray, C.T., et. al. (2003). Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly. *Nat Struct Biol*, 10(7), 572-6.
34. Ziegelhoffer, T., Yau, P., Chandrasekhar, G.N., Kochan, J., Georgopoulos, & C., Murialdo, H. (1992). The purification and properties of the scaffolding protein of bacteriophage lambda. *J Biol Chem*, 267(1), 455-61.
35. Preston, V.G., al-Kobaisi, M.F., McDougall, I.M., & Rixon, F.J. (1994). The herpes simplex virus gene UL26 proteinase in the presence of the UL26.5 gene product promotes the formation of scaffold-like structures. *J Gen Virol*, 75(Pt 9), 2355-66.
36. Thuman-Commike, P.A., Greene, B., Malinski, J.A., Burbea, M., McGough, A., Chiu, W., et.al. (1999). Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids. *Biophys J*, 76(6), 3267-77.

37. Tuma, R., Prevelige, P.E., Jr., & Thomas, G.J., Jr. (1996). Structural transitions in the scaffolding and coat proteins of P22 virus during assembly and disassembly. *Biochemistry*, 35(14), 4619-27.
38. Beaudet-Miller, M., Zhang, R., Durkin, J., Gibson, W., Kwong, A.D., & Hong, Z. (1996). Virus-specific interaction between the human cytomegalovirus major capsid protein and the C terminus of the assembly protein precursor. *J Virol*, 70(11), 8081-8.
39. Hong, Z., Beaudet-Miller, M., Durkin, J., Zhang, R., & Kwong, A.D. (1996). Identification of a minimal hydrophobic domain in the herpes simplex virus type 1 scaffolding protein which is required for interaction with the major capsid protein. *J Virol*, 70(1), 533-40.
40. Parker, M.H., Brouillette, C.G., & Prevelige, P.E., Jr. (2001). Kinetic and calorimetric evidence for two distinct scaffolding protein binding populations within the bacteriophage P22 procapsid. *Biochemistry*, 40(30), 8962-70.
41. Sevrioukov, E.A., Walenta, J.H., Sunio, A., Phistry, M., & Kramer, H. (1998). Oligomerization of the extracellular domain of Boss enhances its binding to the Sevenless receptor and its antagonistic effect on R7 induction. *J Cell Sci*, 111(Pt6), 737-47.
42. Yang, X., Farzan, M., Wyatt, R., & Sodroski, J. (2000). Characterization of stable, soluble trimers containing complete ectodomains of human immunodeficiency virus type 1 envelope glycoproteins. *J Virol*, 74(12), 5716-25.
43. Harbury, P.B., Zhang, T., Kim, P.S., & Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, 262(5138), 1401-7.
44. Oosawa, F., & Kasai, M. (1962). A theory of linear and helical aggregations of macromolecules. *J Mol Biol*, 4, 10-21.
45. Hughson, A.G., & Hazelbauer, G.L. (1996). Detecting the conformational change of transmembrane signaling in a bacterial chemoreceptor by measuring effects on disulfide cross-linking in vivo. *Proc Natl Acad Sci U S A*, 93(21), 11546-51.
46. Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, B. V. V., King, J., Prevelige, P. E., Jr., et.al. (1996). Three-dimensional structure of scaffolding-containing phage P22 procapsids by electron cryo-microscopy. *J. Mol. Biol.*, 260, 85-98.
47. Hu, J.C., O'Shea, E.K., Kim, P.S., & Sauer, R.T. (1990). Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science*, 250(4986), 1400-3.

48. Weissenhorn, W., Calder, L.J., Dessen, A., Laue, T., Skehel, J.J., & Wiley, D.C. (1997). Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 gp41 ectodomain expressed in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 94(12), 6065-9.
49. Chang, C.T., Wu, C.S., & Yang, J.T. (1978). Circular dichroic analysis of protein conformation: Inclusion of the beta-turns. *Anal Biochem*, 91(1), 13-31.
50. Galisteo, M.L., & King, J. (1993). Conformational transformations in the protein lattice of phage P22 procapsids. *Biophys J*, 65(1), 227-35.
51. Berger, B., Shor, P.W., Tucker-Kellogg, L., & King, J. (1994). Local rule-based theory of virus shell assembly. *Proc Natl Acad Sci U S A*, 91(16), 7732-6.
52. Schwartz, R., Shor, P.W., Prevelige, P.E., Jr., Berger, B. (1998). Local rules simulation of the kinetics of virus capsid self-assembly. *Biophys J*, 75(6), 2626-36.