THE ROLE OF SCAFFOLDING PROTEIN IN
BACTERIOPHAGE P22 MORPHOGENESIS

by

Peter Richard Weigele

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

Doctor of Philosophy

Department of Oncological Sciences
The University of Utah
December 2003
This dissertation has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Date

Nov. 30, 2001

Nov. 17, 2001

Nov. 3, 2001

Nov. 30, 2001

Nov. 3, 2001

Christopher P. Hill

David M. Virshup
To the Graduate Council of the University of Utah:

I have read the dissertation of [Name Redacted] in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

[Signature]
Chair: Supervisory Committee

Approved for the Major Department

[Signature]
Barbara J. Graves
Chair

Approved for the Graduate Council

[Signature]
David S. Chapman
Dean of The Graduate School
ABSTRACT

The scaffolding protein of the dsDNA bacteriophage P22 was studied in order to better understand the molecular basis of procapsid assembly. The procapsid is a structural precursor to the capsid, a protein shell that encloses and protects a viral chromosome. The phage P22 procapsid is composed of icosahedrally arranged coat protein subunits encasing a core of scaffolding protein and contains a ring of 12 portal protein subunits at one of its 12 icosahedral vertices. The scaffolding protein activates the assembly of coat protein subunits, directs the positioning of those subunits such that a closed shell of the correct dimension is formed, and is required for the inclusion of portal and injection proteins into the procapsid. DNA enters the procapsid through the portal vertex and the scaffolding protein exits the procapsid intact before DNA packaging is completed. How scaffolding protein accomplishes these processes is the focus of the work presented in this dissertation.

The scaffolding protein was first examined by a functional domain analysis in vivo. Scaffolding protein truncation mutants were cloned and expressed in P22’s host, Salmonella enterica serovar Typhimurium, during infection with a P22 strain defective in scaffolding protein synthesis. Mutants were tested for their ability to complement multiple rounds of infection, exit from the procapsid, recruit portal and injection proteins, and form procapsid-like particles. Successive N-terminal deletions defined the boundaries of amino acids
necessary for these functions, including a 15 amino acid minimal coat protein binding domain.

Scaffolding protein’s ability to bind coat protein and promote assembly of procapsid-like particles both was also studied. Point mutations within the coat protein binding domain were generated and mutant scaffolding proteins were expressed \textit{in vivo} during scaffolding deficient P22 infection. Fifteen different single amino acid substitutions failed to abrogate coat protein binding \textit{in vivo}. Scaffolding protein mutants were subsequently overexpressed, purified, and assayed, \textit{in vitro}, for the ability bind coat protein as well as assemble procapsid-like particles from purified monomeric coat protein. Mutations that affected both processes were found.
CONTENTS

ABSTRACT ........................................................................................................... iv

LIST OF FIGURES .............................................................................................. viii

LIST OF TABLES .................................................................................................... x

ACKNOWLEDGMENTS ....................................................................................... xi

Chapter:

1. INTRODUCTION ............................................................................................ 1

   Bacteriophage and supramolecular assembly ............................................. 1

2. MODEL SYSTEM: BACTERIOPHAGE P22
   PROCAPSID MORPHOGENESIS ................................................................ 10

   Overview of the P22 assembly pathway ................................................... 10
   Procapsid assembly in vivo .................................................................. 12
   Procapsid assembly in vitro ................................................................ 13
   General principles of virus symmetry .................................................. 15
   P22 Procapsid anatomy ....................................................................... 19
      Procapsid and virion symmetry ...................................................... 19
      Procapsid and virion topographies ............................................... 20
      Procapsid scaffolding core and coat lattice contacts .................. 25
      Portal vertex .................................................................................. 27
   P22 procapsid subunits ....................................................................... 28
      Coat Protein .................................................................................. 28
      Scaffolding Protein .................................................................... 38
      Portal Protein .............................................................................. 43
   Scaffolding mediated assembly in other viruses .................................. 45
      Phage SPP1 .................................................................................. 45
      Phage T7 ...................................................................................... 46
      Phage T4 ...................................................................................... 47
      Phages P2/P4 .............................................................................. 48
      φX174 ......................................................................................... 48
      Herpesviridae ............................................................................. 49
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The bacteriophage P22 assembly pathway</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>&quot;T&quot; numbers</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>Seven quasi-equivalent postions</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>Cryo-electron micrographic reconstructions</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>Computationally isolated procapsid densities</td>
<td>24</td>
</tr>
<tr>
<td>2.6</td>
<td>The inner surface of the procapsid</td>
<td>26</td>
</tr>
<tr>
<td>3.1</td>
<td>Sucrose gradient analysis of <em>in vivo</em> assembly products</td>
<td>64</td>
</tr>
<tr>
<td>3.2</td>
<td>Electron micrograph of negatively stained procapsid-like particles</td>
<td>67</td>
</tr>
<tr>
<td>3.3</td>
<td>Sucrose sedimentation analysis of SCAFF141-294</td>
<td>68</td>
</tr>
<tr>
<td>3.4</td>
<td>Incorporation of the portal protein into procapsids</td>
<td>70</td>
</tr>
<tr>
<td>3.5</td>
<td>Minimal sequence required to support infection</td>
<td>72</td>
</tr>
<tr>
<td>3.6</td>
<td>Sucrose gradient analysis of assembly products produced during infections with scaffolding protein supplied in trans</td>
<td>74</td>
</tr>
<tr>
<td>3.7</td>
<td>Protein composition of purified virions</td>
<td>78</td>
</tr>
<tr>
<td>3.8</td>
<td>Electron microscopy of virions</td>
<td>79</td>
</tr>
<tr>
<td>3.9</td>
<td>DNA content of virions</td>
<td>81</td>
</tr>
<tr>
<td>3.10</td>
<td>Schematic summary of truncation and fusion mutants tested for coat protein binding</td>
<td>83</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.1</td>
<td>Purification of materials for <em>in vitro</em> shell binding and assembly</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>assays</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>Structural models of the P22 scaffolding protein C-terminal 40</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>amino acids</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Coat protein shell binding reactions</td>
<td>114</td>
</tr>
<tr>
<td>4.4</td>
<td>Electron microscopy of shell binding materials and reactions</td>
<td>115</td>
</tr>
<tr>
<td>4.5</td>
<td>Locations of point mutations within the structure of the scaffolding</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>protein coat binding domain</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>NaCl inhibition of scaffolding protein shell binding</td>
<td>122</td>
</tr>
<tr>
<td>4.7</td>
<td>Spectrophotometric analysis of <em>in vitro</em> assembly reactions</td>
<td>123</td>
</tr>
<tr>
<td>4.8</td>
<td>Coat and scaffolding protein assembly <em>in vitro</em></td>
<td>124</td>
</tr>
<tr>
<td>4.9</td>
<td>Sucrose sedimentation of <em>in vitro</em> assembly products</td>
<td>126</td>
</tr>
<tr>
<td>4.10</td>
<td>Agarose gel electrophoresis of <em>in vitro</em> assembly reactions</td>
<td>127</td>
</tr>
<tr>
<td>A.1</td>
<td>Sucrose gradient analysis of scaffolding:H6 filled procapidds</td>
<td>138</td>
</tr>
<tr>
<td>A.2</td>
<td>Immunoprecipitation (IP) of procapids</td>
<td>140</td>
</tr>
<tr>
<td>A.3</td>
<td>Electron microscopy of procapids and fusion protein filled procapids</td>
<td>141</td>
</tr>
<tr>
<td>B.1</td>
<td>Competitive filling</td>
<td>146</td>
</tr>
<tr>
<td>C.1</td>
<td>Southern blot of DNA from P22-like phage</td>
<td>151</td>
</tr>
<tr>
<td>C.2</td>
<td>Phage and procapid-like particles purified</td>
<td>153</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Primers used in PCR cloning of scaffolding protein truncation and fusion mutants</td>
<td>56</td>
</tr>
<tr>
<td>3.2 Summary of truncation and fusion mutants tested for complementation and coat binding in vivo</td>
<td>65</td>
</tr>
<tr>
<td>3.3 Infectivity of purified virions assembled from scaffolding protein truncation mutants</td>
<td>76</td>
</tr>
<tr>
<td>4.1 Coat protein binding domain point mutants and the oligonucleotide sequences used to generate them</td>
<td>101</td>
</tr>
<tr>
<td>4.2 Properties of P22 scaffolding protein with amino acid changes in the coat protein binding domain</td>
<td>112</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I am forever indebted to my advisor, Dr. Sherwood Casjens, for introducing me to a great experimental system, providing me with superior intellectual and technical support, transmitting history, and allowing me freedom in the direction of my projects. Excellent guidance and understanding were provided by my dissertation committee: Dr. Christopher Hill, Dr. John Roth, Dr. Wesley Sundquist, and Dr. David Virshup. Dr. Peter Prevelige communicated unpublished data and supplied essential instruction for the *in vitro* assembly experiments described in Chapter 4. Ms. Shirley Wise and Ms. Dee Dalponte gave timely and efficient administrative assistance. Three years of generous financial support came from the NIH Genetics Training Grant, and ultimately the U.S. taxpayer. Finally, I thank Rebecca Pye, my loving wife, for her profound strength, sustaining care, and unwavering support.
CHAPTER 1

INTRODUCTION

Bacteriophage and supramolecular assembly

Cellular life depends on a wonderfully intricate network of assembling and disassembling multi-subunit protein complexes. Crucial to this network is the genetic control of protein function through nucleic acid sequence, and the ability of this control to evolve through time. Our knowledge of gene sequence and regulation has increased dramatically during the last decade of the millennium through large scale sequencing projects, microarray expression analysis, and high throughput screening technologies; yet many questions remain about how gene products interact and are organized within the cell. Answers to some of these questions are being sought in the emerging field of proteomics. But even if high throughput and large scale screening technologies were able to reveal all the pairwise interactions between proteins within cells (Michnick, 2001; Uetz et al., 2000), our understanding of higher orders of protein organization/assemblies, their stoichiometries, temporal composition, and mechanisms of association, would be far from complete. Understanding the complex interactions of proteins within supramolecular assemblies that contain large numbers of subunits is one route to the goal of understanding how genes and their products cooperate to form the hierarchies of order and complexity that make up cellular ultrastructure and control signaling, metabolism and
macromolecular synthesis. It has been suggested that exploring this next level of organization is best addressed in studies on macromolecular machines (Alberts, 1998).

Macromolecular machines are large assemblies of proteins that utilize free energy (generally in the form of NTP hydrolysis, pH gradients, or light) coupled to conformational changes within the machine’s subunits to accomplish work (Nogales & Grigorieff, 2001). Understanding how they work requires integrating thermodynamic, kinetic, and structural information. In practical terms, this approach focuses on obtaining structures of the assembled machine and its components, determining the positions of components within the complex, as well as determining the structures of the machine’s reaction intermediates. This knowledge deepens our understanding of the relationship between structure and function not only by correlating structural information about the machine with the physical and/or chemical work it performs, but also by adding the extra dimension of time through the visualization/modeling of dynamic motions within the machine complex.

The success of this approach has been demonstrated in the studies of several large supramolecular assemblies, including the flagellar motor (Lloyd et al., 1999), F_oF_1 ATP synthase (Menz et al., 2001), rhodopsin (Menon et al., 2001), myosin (Irving et al., 2000), the GroES/EL chaperonin (Hunt et al., 1996), the phage Φ29 portal DNA packaging machine (Simpson et al., 2000), and the ribosome (Clemons et al., 1999). While these studies yield insights into how subunit conformational changes are coupled to work and catalysis, the information is limited to cyclic changes within an already assembled complex.
An important, but perhaps underappreciated question in the analysis of macromolecular machines is their assembly. Intracellularly, macromolecular machines need to be assembled and disassembled as well. Such questions can be addressed by studying virus structure and assembly.

The virus particle, or virion, is a protein machine whose job is to assemble itself, encapsidate a virus chromosome and deliver it into another cell. The virus capsid must protect its nucleic acid payload until delivery into a new host, which is accomplished either by “injection” as in the case of the dsDNA bacteriophages, or, as for many other viruses, by actually entering the host cell and uncoating or disassembling thereby releasing the viral chromosome. The process by which virus capsids and other supramolecular complexes assemble is commonly referred to as “self-assembly.” During self-assembly, the correct guidance, positioning and assembly of subunits within a macromolecular complex is controlled entirely by protein-protein and protein-nucleic acid interactions among the component subunits. These interactions are formed though non-covalent forces and occur in the absence of any additional input of energy. Thus, self-assembly is energetically favorable through primary sequence alone.

Early studies on the in vitro assembly of tobacco mosaic virus (TMV) from purified components not only established the phenomenon of self-assembly (Fraenkel-Conrat & Williams, 1955), but also demonstrated that, although energetically favorable, assembly of coat subunits is not spontaneous; it requires a specific viral RNA to nucleate assembly (Zimmern, 1977), implying some form of control of the assembly process. Implicit in self-assembly is the general problem that assembly of some subunits out of sequence might block or prevent the binding of subsequent subunits. Furthermore, for an extra-cellular structure
such as the flagellum, intracellular assembly could be lethal. Therefore, in addition to being energetically favorable, self-assembly must be a controlled process.

Combined genetic, biochemical and electron microscopy (EM) data on bacteriophage T4 baseplate/tail morphogenesis demonstrated that the assembly of this substructure, composed of 18 types of subunits, proceeds along an ordered pathway (Kikuchi & King, 1975a; Kikuchi & King, 1975b; Kikuchi & King, 1975c); certain subunits are added only subsequent to the assembly of other subunits. Thus, assembly progresses through the ordered accumulation of subassemblies pieced together to make the final structure. In attempting to explain this apparent control of T4 baseplate assembly, Kikuchi and King (1975a) suggested that the binding of subunits created active sites for subsequent addition of other subunits, but that unassembled subunits themselves did not contain active assembly sites. This idea was further elaborated by the concept of autostery, proposed by Caspar (Caspar, 1980), which suggests that for subunits capable of multimerizing, an energetic barrier exists between “inactive, unsociable” and “active, associable” subunit conformations and that interaction of free subunits with the bonding surfaces presented by polymerized subunits serves to overcome this barrier. Self-regulation during macromolecular-assembly has also been discussed elsewhere (Berget, 1985; King, 1980; Wood, 1980).

One consequence of autosterically controlled assembly is the appearance of cooperativity of assembly. In a transcription of a discussion session following Caspar’s paper (Caspar, 1980), Harold Erikson suggested cooperative assembly of polymerizing subunits could be sufficiently described by equilibria of bimolecular associations using physiologically plausible thermodynamic
parameters without invoking conformational control. However, subsequent high-resolution structural information as well as biochemical data on the assembly of TMV and the bacterial flagellum demonstrate a physical basis for autosteric control during self-assembly.

Autosteric control could be due to changes in folded protein domains or in the state of folding (e.g., unfolded regions become folded upon assembly). For example, unassembled coat protein subunits of TMV contain internal loops of disorder not seen in the fully assembled forms (Jardetzky et al., 1978; Namba & Stubbs, 1986). Assembly of the helical TMV rod begins with the multimerization of coat protein into a disk-like 20S intermediate composed of approximately 35 subunits comprising about two helical turns (Butler & Klug, 1971). Negatively charged, disordered loops from the coat subunits project into the central cavity of this disk. Charge repulsion between these loops and free subunits presents an energetic barrier to further assembly. The disk is able to recognize and bind an initiator sequence on the viral RNA (Zimmern & Butler, 1977). Binding induces folding of the exposed loops and the addition of subsequent coat protein subunits becomes energetically favorable. In this way, the disordered internal loop of the TMV coat subunit is used to prevent assembly until kinetically initiated by viral RNA, which ensures its incorporation into the completed structure.

The use of partially unfolded subunit conformations to control self-assembly is also used during flagellar assembly. The flagellum is essentially a bundle of 12 helical flagellar filaments. Flagellar filaments "grow" by addition of flagellin subunits to the tip of an elongating flagellar filament (Iino, 1969). Subunits reach the tip of the assembling filament by diffusing through the 30Å
diameter channel in the center of the filament (Samatey et al., 2001). Unassembled flagellin contains substantial unfolded sequences at both N- and C-termini (Vonderviszt et al., 1989). The disorder is believed to aid in transport across the membrane into and through the 30Å diameter inner filament channel as well as preventing the subunits from polymerizing intracellularly. Unassembled flagellin is prevented from diffusing away from the end of the filament by a pentameric cap structure composed of the HAP2 protein (Homma et al., 1984). Symmetry-mismatch between the fivefold cap structure, the 5.5 subunit per turn filament (Yonekura et al., 2000) as well as flexible arms in the HAP2 protein that interact with the filament end (Vonderviszt et al., 1998) are believed to allow rotation of the cap structure at the filament end and sequentially expose a space between the cap and filament where addition of a subunit could take place. The cap structure may therefore act either by giving flagellin subunits sufficient time to reach an assembly competent conformation and bind to the filament end, or by presenting binding surfaces that catalyze this change.

These assembly systems provide evidence for the use of partially unfolded conformations by proteins destined to participate in multi-subunit complexes to control initiation of the assembly process. However, coupling of folding with assembly might also be a general feature of proteins whose proper function requires either conformational flexibility or adaptability. Such flexibility might be important for the ability of flagellin to adopt nonidentical conformations within the flagellar filament as well as the ability of flagellar filaments to switch helical handedness in response to counterclockwise versus clockwise rotation of the flagellar motor. Many viral coat proteins must similarly have intrinsic conformational plasticity as demonstrated by being able to adopt similar but
distinct conformations within the assembled capsid (see "General principles of virus assembly" below for further discussion). The ability to adopt these multiple states is called conformational switching. Intrinsic flexibility of the subunits may be required for conformational switching. Indeed, the unassembled coat proteins of bacteriophages P22 (de Sousa et al., 1999) and PRD1 (Tuma et al., 1996b), and cowpea mosaic virus (CPMV) (Da Poian et al., 1994), are less stable or incompletely folded as compared to their polymerized counterparts. The demands of conformational plasticity are particularly salient for the øX174 external scaffolding protein, gpD, which adopts two types of nonidentical binding configurations on the surface of the øX174 procapsid (Dokland et al., 1997).

Conformational plasticity is also utilized by proteins in other kinds bonding configurations. Intrinsically unstructured domains within the cyclin dependent kinase inhibitor p21 have been shown to adopt structure upon binding to target proteins (Kriwacki et al., 1996). It has been suggested that these intrinsically unstructured domains permit p21 to bind several different proteins as well as mediate reversibility of binding. Similarly, the bacteriophage P22 scaffolding protein is at least partially unstructured in solution (Tuma & Thomas, 1997), interacts with multiple proteins (Greene & King, 1996), and binds its protein partners reversibly (Casjens & King, 1974). Folding changes are also seen in proteins whose DNA/RNA binding motifs fold upon interaction with their target nucleic acid (Spolar & Record, 1994), and it has been suggested that the lack of structure serves to increase the specificity of binding by offsetting energetically favorable binding interactions with energetically unfavorable folding transitions. These examples suggest that the role of partially unfolded
conformations may be utilized to accomplish not just timing of assembly, but reversibility and specificity as well.

Another way in which the conformation of free subunits can affect the manner in which they bind is revealed by the nature of their binding interface. Bonding interfaces between proteins are defined as the surface area between the subunits that is not accessible to solvent. For proteins that do not undergo large conformational changes upon binding, the interaction is generally thought to proceed through the docking of complementary surfaces and is held in place by favorable hydrogen bonding and ionic interactions (Xu et al., 1997a; Xu et al., 1997b). Calculating the interface surface areas for these types of protein-protein interactions has been used to reliably predict affinity. However, such calculations are difficult for proteins whose binding results in the interlacing of secondary and tertiary structural elements at the bonding interface, as is the case for proteins that multimerize by “domain swapping” (Bennett et al., 1995). Tsai et al. (1997) have used structural coordinates of protein-protein complexes in the protein data bank to calculate the degree of intimacy at interfaces by measuring, for each buried atom of the interacting proteins, what percentage of atoms in a surrounding radius are contributed by each protein partner (Tsai & Nussinov, 1997; Tsai et al., 1997; Tsai et al., 1996). These studies suggest that for some protein complexes, the binding interface is similar to the hydrophobic core of a globular protein, with tertiary-like packing of secondary elements provided by each binding partner. These folded interfaces imply a partially unfolded state for the unbound subunits. Interlacing contacts between viral coat subunits are observed in the structures of tomato bushy stunt virus (TBSV) (Olson et al., 1983) and bacteriophage HK97 (Wikoff et al., 2000) as well as SV-40 where the flexible
C-terminal arms mediate nonequivalent contacts between coat protein pentamers (Liddington et al., 1991). The interlacing of chains between subunits in these two very different viruses demands that their subunits must contain some partially unfolded domains, or at least undergo dramatic refolding upon assembly.

The examples discussed above suggest that partially unfolded conformations of proteins are used to control processes that are the basis of self-assembly: protein polymerization, plasticity of binding conformation, binding specificity/stringency, and the structure of the binding interface. Transition state theory offers a convenient way to illustrate what we might learn from studying self-assembly. For an energetically favorable process, the free energy is higher at the start of the reaction than the final product. However, for a complex assembly process the reaction path must traverse peaks of energetically unfavorable transitions and valleys of local minima where stable intermediates reside. Within this framework, studying intermediates in self-assembly processes may yield a more detailed topography of the “energy landscape” of assembly.

The bacteriophage assembly system offers an unparalleled model to examine these issues. Relative to many components of a cell, the bacteriophage is large, both in size and number of subunits, yet simple in that a phage capsid contains only a few kinds of subunits. Bacteriophages can be genetically manipulated, are biochemically tractable, and can be visualized by electron microscopy. Most importantly, the bacteriophage is not essential to the cell. Therefore, the intermediates of assembly, both naturally occurring and those trapped by genetic and biochemical means, are accessible to study.
CHAPTER 2

MODEL SYSTEM: BACTERIOPHAGE P22

PROCAPSID MORPHOGENESIS

The subject of this dissertation is the assembly of the bacteriophage P22 procapsid and the role of scaffolding protein in this process. Scaffolding protein acts in concert with two other viral proteins, the coat and portal proteins, to form a viral structural intermediate termed the procapsid. In the following sections, scaffolding protein and its properties are described in the context of procapsid assembly.

Overview of the P22 assembly pathway

The assembly of the *Salmonella enterica* (serovar Typhimurium) bacteriophage P22 epitomizes the assembly of dsDNA bacteriophages. Assembly proceeds via a well-defined pathway (see Figure 2.1). The genes and their products involved in phage morphogenesis have been identified and have been the subjects of research for three decades. Unlike the small animal and plant viruses that assemble their virions by condensing capsid subunits around the viral nucleic acid, P22 and all the large dsDNA viruses characterized thus far first assemble a capsid precursor called the procapsid into which the viral chromosome is subsequently packaged. DNA enters this structure through a portal protein complex at a unique vertex within the procapsid. For most of these
Figure 2.1 The bacteriophage P22 assembly pathway
viruses, the packaging of the viral chromosome is accompanied by a transformation of the procapsid shell that leads to shell expansion that increases the internal volume of the capsid (Zhang et al., 2000) and makes a physically more robust shell (de Sousa et al., 1999). Subsequent steps in P22 assembly include stabilization of the viral DNA within the capsid (Lenk et al., 1975; Strauss & King, 1984) and the assembly of the viral cell binding protein, the tailspike which binds to an digests O-antigen polysaccharide on the bacterial surface, allowing the phage access to the outer membrane for injection of DNA (Eriksson & Lindberg, 1976; Iwashita & Kanegasaki, 1976).

Procapsid assembly in vivo

For P22, only five gene products are required to make a procapsid and package DNA into it. The procapsid is composed of 420 molecules of coat protein (or 415 if the portal complex replaces a pentamer of coat protein at one vertex, see below), 250-300 molecules of scaffolding protein in the interior, and a dodecameric ring of portal protein located at a unique vertex (diagrammed in Figure 2.1). In addition, three kinds of “injection proteins” are incorporated into the procapsid that are not essential for its structural integrity or DNA packaging, but are required for successful delivery of DNA into the host (Bandyopadhyay et al., 1979).

Assembly of the procapsid requires copolymerization of coat and scaffolding proteins. Scaffolding protein directs the correct assembly of coat protein in vivo (Casjens & King, 1974; Earnshaw & King, 1978) as well as in vitro (Fuller & King, 1982), and several kinds of malformed arrays of coat polymerize in its absence. The dodecameric ring of portal protein is incorporated into the
procapsid only once per shell (Bazinet & King, 1988). How one portal ring precludes the incorporation of additional rings at the other vertices is unknown. Approximately when the P22 43kb dsDNA is translocated into the procapsid, the scaffolding molecules exit the procapsid (Casjens & King, 1974), presumably through the holes that exist in the coat lattice (Prasad et al., 1993), and are then free to catalyze more rounds of procapsid assembly (Casjens & King, 1974). The mature virion does not contain any scaffolding protein.

**Procapsid assembly in vitro**

Aspects of procapsid assembly have been successfully reproduced in vitro. Purified unassembled scaffolding and coat proteins, when mixed together under conditions where neither assembles alone, spontaneously associate to form closed shells that resemble procapsids built in vivo (Fuller & King, 1980; Prevelige et al., 1988). These particles do not contain portal protein, and portal protein is not incorporated when it is present in the assembly mixture. However, purified coat and scaffolding proteins, when added to a cellular extract from a coat deficient infection (these lysates contain a low amount of scaffolding protein due to autoregulation by unassembled scaffolding protein), are able to assemble some procapsids that are competent to package DNA and form infectious virions (Fuller & King, 1982). Thus, under these conditions, purified coat and scaffolding are able to incorporate at least some of the portal protein and injection proteins. Perhaps an interaction between scaffolding, portal and the minor proteins is responsible for portal incorporation under these more native conditions. Scaffolding protein alone does not form any large MW complexes in solution but is in equilibrium between monomeric, dimeric, and tetrameric forms (Parker et
al., 1997a). At the low salt concentrations generally used for procapsid assembly, coat remains monomeric at concentrations less than 1-2 mg/ml (Prevelige et al., 1988). At higher concentrations, coat protein self assembles into aberrant shells as well as insoluble aggregates (Prevelige et al., 1993). Thus, the presence of scaffolding lowers the concentration at which coat protein assembles in vitro and also increases the fidelity of assembly.

Co-assembly of coat and scaffolding proteins can be monitored spectrophotometrically by light scattering (measured at 90° to light path) or turbidity (Prevelige et al., 1988). The degree of light scattering is proportional to the mass of the particles produced. The increase of turbidity over time also gives a measurement of the rate of assembly. Spectrophotometric methods cannot distinguish procapsid-like particles from aberrant aggregates of coat protein, but the products of the reaction can be further assessed by sucrose gradient centrifugation, SDS-polyacrylamide gel electrophoresis (SDS-PAGE), agarose gel-electrophoresis, and electron microscopy. Rapid quenching of assembly by dilution and glutaraldehyde cross-linking shows few partially assembled shells and no smaller aggregates of protein suggesting that assembly of the procapsid is rapid once it starts and proceeds by the addition of monomers of coat protein and not by preformed oligomers (Prevelige et al., 1993).

Assembly experiments with limiting concentrations of coat and scaffolding demonstrate 0.8 mg/ml and 0.25 mg/ml minimum concentrations for assembly, respectively, (Prevelige et al., 1988). The rate of assembly has a fifth order dependence on coat protein concentration and a second order dependence on scaffolding protein concentration (Prevelige et al., 1993). Prevelige et al. (1993) have used these data to suggest that assembly is nucleated by a pentamer of coat
protein and dimers of scaffolding protein participate in assembly. Based on the *in vitro* assembly kinetics, procapsid assembly has been proposed to be nucleation limited, but a nucleation mechanism has not been proved. A subsequent simulation of virus assembly using purely thermodynamic parameters displayed features that mimic nucleation limited assembly (Zlotnick, 1994).

**General principles of virus symmetry**

In order to discuss the arrangement of coat protein in procapsids, an introduction to icosahedral symmetry and viral T-numbers is required. Not long after the elucidation of the genetic code, it was realized that the coding capacity of a viral genome was insufficient to encode a single protein large enough to surround the a viral nucleic acid. Therefore, the protein content of the virus must be made up of multiple copies of a single, or a few kinds of subunit. Early X-ray diffraction data supported this hypothesis. Based on that diffraction data, Crick and Watson (1956) proposed that only two kinds of virus coat protein symmetry would be utilized: helical for the rod shaped viruses and cubic point group for "spherical" viruses (Crick & Watson, 1956). Further experiments showed icosahedral (a subtype within "cubic") symmetry in all "spherical" viruses. However, in order to maintain identical contacts between an icosahedrally arranged aggregate of chiral objects such as coat protein subunits, only three subunits on each of the 20 icosahedral faces can be accommodated. Spherical viruses were known that contain many more than 60 subunits per virion, but the stoichiometries observed were always multiples of 60.

Quasi-equivalence is a geometric formalism developed by Caspar and Klug (1962) to account for viruses with more than 60 subunits per shell while
maintaining *nearly identical* contacts between those subunits (Caspar & Klug, 1962). The basic premise of this idea is that an icosahedral face can be divided into *triangular facets*, and that each of these facets can be populated (packed) with three subunits thereby constructing a larger structure in which similar, but not identical contacts are made between the various subunits.

Consider an infinite, two-dimensional triangulated net, bearing in mind each triangle corresponds to a facet (see Figure 2.2). A vector drawn between any two points of six-fold symmetry on the net can be described by two indices, \( h \) and \( k \), with the start of the vector being its origin \((0,0)\) drawn to a point \( h,k \) on the lattice. \( h \) is the number of integral units away from the origin in the horizontal direction and \( k \) the number in the direction \(60^\circ\) to horizontal. This vector then defines one edge of an icosahedral face, i.e., one edge of an equilateral triangle that can be superimposed on the triangulated net. It is important to note that the edge does not have to coincide with the lines of the triangulated net or the facet edges. If one were to cut out 20 such triangles as defined by the vector edge length, these could be assembled into an icosahedron. A simple algebraic equation based on the edge length describes all the possible icosahedral subtriangulations: the triangulation number, \( T = (h^2 + hk + k^2) \). \( T \) is an integer of the series \((1, 3, 4, 7, 9, 13, 16...,)\), and is the number of facets per face. Since there must be three subunits in each facet (related by three-fold symmetry) and 20 faces of an icosahedron, there are \( 60T \) subunits in an icosahedral virus shell.

In essence, the triangulated net represents a flat network of subunits arranged as hexamers where six facet points meet. Those subunits at the vertices of the triangular *face* are subunits that participate in pentamer interactions at each of the 12 five-fold symmetric vertices of an icosahedron. Therefore,
Figure 2.2 "T" numbers. A hexagonal lattice of chemically and structurally al coat protein subunits (top) can be subdivided to fit into the faces of an edron. The vectors define an edge length for a face of an icosahedron of T number. Only the T=7 face has all edge lengths shown. Examples of faces the context of complete icosahedrons are shown.
symmetric distribution of hexameric rings between the 12 pentameric vertices allows for building of capsids with greater than 60 subunits, while still preserving similar bonding interactions between those subunits.

One consequence of building shells with T-numbers larger than one is that coat subunits must participate in slightly different intersubunit interactions in order to be accommodated within the shell lattice. This difference is most obviously seen when noting that coat subunits display pentameric clustering at the icosahedral vertices, while the faces are packed with hexameric rings of coat subunits. However, the local differences between the subunits were presumed to be only slightly different, hence, the subunits were said to be in quasi-equivalent positions. In the absence of any detailed structural knowledge of viral subunits, the concept of "bond deformation" was invoked by Caspar and Klug in order to accommodate these differences; i.e., subunits utilized the same surfaces for intersubunit contacts but that changes in the bonding angles between subunits could account for the quasi-equivalent positions within the lattice. Since the publication of the theory of quasi-equivalence, high-resolution virus structures have demonstrated at least three ways for subunits to adopt nonsymmetry related positions. One is by mediating contacts between subunits using flexible loops as seen in interpentameric contacts in SV40 (Liddington et al., 1991) and between the E-loops in subunits of HK97 (Wikoff et al., 2000). Second, utilizing alternate packing of secondary structure at subunit interfaces in order to change dihedral angles as in the case of the T=3 RNA viruses (Olson et al., 1983; Rossmann et al., 1983), and finally, by changing angles between otherwise rigid domains by use of a flexible "hinge" connecting those domains as may be the case for P22 (Casjens, 1979; Lanman et al., 1999). We now understand that
conformational adaptability and utilization of different bonding surfaces is the underlying principle for quasi-equivalence.

**P22 Procapsid anatomy**

**Procapsid and virion symmetry**

The P22 procapsid shell is composed of coat protein molecules arranged in a T=7 *laevo* icosahedral lattice (*i.e.*, seven facets per icosahedral face) (Casjens, 1979) with 420 coat subunits that are clustered into 12 pentamers, one at each of the icosahedral vertices, and 60 hexamers packed three on each of the 20 faces. The corresponding features revealed by cryo-electronmicroscopy and computer aided image reconstruction are pentons and hexons, respectively (Prasad *et al.*, 1993). However, apparent hexon/penton boundaries are not necessarily indicative of subunit positions, whose precise boundaries await definition by a higher resolution structure. In the wild-type procapsid, one icosahedral vertex is occupied by the dodecameric portal complex which most likely replaces one coat pentamer, as is known to be the case for the φ29 procapsid (Ibarra *et al.*, 2000), thereby making the number of coat subunits 415 per P22 procapsid.

A T=7 lattice contains within each icosahedral face seven non-symmetrically related, nonidentical positions occupied by coat subunits (Zhang *et al.*, 2000). These seven subunits comprise the asymmetric unit which consists of one subunit from the pentamer and six subunits from a hexamer. Three asymmetric units together make an icosahedral face and are related to each other by perfect rotational symmetry around the three-fold axis of the face. However, each icosahedral face has three additional pseudo or local three-fold axes of symmetry. This local three-fold axis resides between two hexamers of a face and
a hexamer from an adjacent face across the face edge. These symmetries are schematically illustrated in Figure 2.3.

**Procapsid and virion topographies**

Cryo-EM reconstructions of the P22 procapsid reveal a “spherical” complex with an average diameter of 600Å and a bumpy outer surface featuring quasi-equivalently distributed penton and hexon rings of knob-like densities protruding from the procapsid shell floor (illustrated in Figure 2.4). The pentons consist of five club-shaped spokes symmetrically arranged around the five-fold axis. The hexons consist of six yoke or hairpin shaped spokes and have a pronounced skew. The hexons have a pronounced skew, also seen in HK97 (Conway et al., 2001; Lata et al., 2000), lambda (Dokland & Murialdo, 1993), and cauliflower mosaic virus (CMV) (Cheng et al., 1992) and others.

The 15Å resolution structure reveals subtle differences between subunits across the skew axis. Along the floor of the procapsid’s surface, two density features, a saddle and a smaller knob bridge the distances between the hexon-hexon and hexon-penton. A prominent feature of the procapsid shell is the presence of 35-40Å diameter holes in the centers of both pentons and hexons (Prasad et al., 1993). Short finger-like densities protrude from the inner rim of the hexon holes towards the procapsid interior (Thuman-Commike et al., 1999). The topography of the inner procapsid surface is different from the outer surface with the most prominent features being ridge-like densities found at the local and strict three-fold axes of symmetry as well as in those regions located between a penton and two adjacent hexons (Zhang et al., 2000).
Figure 2.3 Seven quasi-equivalent positions. This schematic of a $T=7$ face with some adjacent hexamers illustrates the seven unique subunit positions, A through G, within the lattice as well as strict and local symmetry axes. Figure adapted from Prasad et al. (1993).
Figure 2.4 Cryo-electron micrographic reconstructions. The procapsid (a) and the mature virion (b) are shown. A single icosahedral face is outlined with five-fold and a strict three-fold symmetry axis indicated. A single penton and hexon are each outlined. One subunit from a pentamer and six subunits from a hexamer make up the asymmetric unit of the icosahedral face. There are three asymmetric units per face, related to each other by 120° rotation about the three-fold symmetry axis. Figure adapted from Zhang et al. (2000).
There are striking changes in coat shell topography between the procapsid and the virion, and some of these changes are attributable to the specific movements of density domains described in the procapsid structure (Figure 2.4) (Zhang et al., 2000). The 650Å diameter virion is similarly textured although surface features are flattened and less pronounced as compared to the procapsid (shown in Figure 2.4) (Prasad et al., 1993; Zhang et al., 2000). In the virion, hexons display local six-fold symmetry, the skew having been lost during expansion. This loss of skew subsequent to expansion has also been observed in HK97 (Conway et al., 2001; Lata et al., 2000) and lambda capsids (Dokland & Murialdo, 1993) but not T4 (Aebi et al., 1974). The procapsid to capsid transition in T7 has not yet been observed by cryo-EM. The major changes around the P22 hexon occur in two of the yoke-shaped protrusions hairpin located at the narrow ends of the elliptically shaped hole (Zhang et al., 2000). In the procapsid, these have a closed hairpin or loop like appearance; in the virion, these densities have adopted an open or horseshoe shape (Zhang et al., 2000) (see Figure 2.5). Additionally, inter penton-hexon and hexon-hexon distances are greater than in the procapsid (Zhang et al., 2000). These are correlated with a flattening of the inner surface three-fold axis densities, resulting in a thinning of those areas of the procapsid shell and elongation of the saddle like densities present at all strict and local two-fold axes of symmetry (Zhang et al., 2000). Expansion also results in the contraction of lattice holes to 15Å diameter (Prasad et al., 1993; Zhang et al., 2000). This appears to be achieved by the upward hinging of the finger-like densities towards the center of the hole (Zhang et al., 2000).
Figure 2.5 Computationally isolated procapsid densities. The column on the left are isolated parts of the procapsid, mature shell on the left. A. and C. are outer and side views, respectively, of the strict three-fold axis of symmetry. Notice the flattening of the inner surface and radial displacement of densities in the expanded form seen in panels B. and D. E. and G. are outer and side views, respectively of computationally isolated local six-fold axes of symmetry, or hexons. Notice the the inwardly projecting finger like densities in the side view presented in G. These fingers are displaced upwards towards the center of the hole in F. and H. Notice also the loss of skew in F. Figure adapted from Zhang et al. (2000).
**Procapsid scaffolding core and coat lattice contacts**

Cryo-electron micrograph reconstructions of P22 procapsids and T=7 coat protein shells without scaffolding protein may partially reveal the organization of the scaffolding protein (Thuman-Commike et al., 2000; Thuman-Commike et al., 1999). Electron density differences between the scaffolding protein filled procapsids and empty coat protein shells map to the surfaces of four of six finger-like densities at the inner rim of the holes found at the centers of hexons (see Figure 2.6) (Thuman-Commike et al., 1999). These finger-like densities are suggested to be the site on coat protein where the scaffolding protein binds. Below these radii there is a region of very low protein density detected in the reconstruction, suggesting the volume of protein in this region is sparse, does not have icosahedral symmetry and/or is disordered (Thuman-Commike et al., 1996). A similar low density region just interior to the inner surface of the coat protein shell is observed in cryo-EM reconstructions of cytomegalovirus (Trus et al., 1999) and electron micrographs of phage T7 (Serwer, 1979; Steven et al., 1983). It has been suggested that this region represents a disordered link between the coat protein binding domain of the scaffolding protein and the more electron dense regions at smaller radii within the procapsid. This assertion is potentially supported by a region of disorder seen in the NMR solution structure of the scaffolding protein residues 244-303 just N-terminal to the coat protein binding domain (Sun et al., 2000) (see also Chapter 3), although the latter disorder could also be an artifact of truncation. Structural studies of herpesvirus procapsids (Newcomb et al., 2000), herpes B-capsids (Zhou et al., 1998) and the ssDNA phage oX174 (Dokland et al., 1997) also failed to reveal an icosahedrally organized scaffolding protein core.
Figure 2.6 The inner surface of the procapsid. The darker features are density differences between the scaffolding protein filled procapsid and the empty coat protein shell attributed to scaffolding protein. Note, only two of the four finger like coat protein densities are circled in the asymmetric unit. The other two circled densities are from an adjacent hexamer. Figure adapted from Thuman-Commike et al. (1999).
Portal vertex

P22, like all other dsDNA tailed-bacteriophages, has a portal protein complex located at a unique vertex. For P22, as in phages lambda (Kochan et al., 1984), T4 (Driedonks, 1981), T7 (Kocsis et al., 1995), T3 (Valpuesta et al., 1992), SPPI (Lurz et al., 2001), and ϕ29 (Carrascosa et al., 1985), the portal ring is composed of 12 subunits. For P22, it is not clear if the portal sits atop a coat pentamer or replaces it entirely, but in ϕ29 it appears that the connector is part of the shell, replacing a pentamer of capsid subunits (Ibarra et al., 2000). Computer assisted image reconstruction from electron microscopic images of negatively stained P22 portal rings shows arms projecting orthogonally to a radial distance of 90Å from the wide end of a slightly tapering cylinder 130Å long and 73Å wide with an inner channel of approximately 30Å diameter, which is sufficiently wide to permit the passage of DNA (Bazinet et al., 1988). One obvious feature of all portal complexes characterized to date is a symmetry mismatch between a 12-fold portal ring and the five-fold vertex at which it resides, an arrangement which has been proposed to allow for rotation of the portal complex within the coat lattice during DNA translocation (Hendrix, 1978). It is not known how the P22 portal ring’s dimensions are related to its orientation in the lattice; however, a C-terminal histidine epitope tag engineered onto the portal protein is accessible to antibodies only after disruption of the procapsid shell (Moore & Prevelige, 2000). Cryo-EM reconstructions of P22 procapsids and capsids do not reveal the portal complex in any detail since its density information is lost during icosahedral averaging (Prasad et al., 1993; Thuman-Commike et al., 1996). However, comparison of portal-less and portal containing procapsids do reveal
subtle density differences at the five-fold vertices (Thuman-Commike et al., 1996). Crystals of P22 portal rings have been grown, diffraction data gathered, and structural derivation is currently underway (S. Moore and P. Prevelige, personal communication).

**P22 procapsid subunits**

The coat, scaffolding, and portal proteins of the procapsid have been studied individually using biochemical and biophysical methods on homogeneous preparations of protein as well by genetic methods in vivo. It is useful to consider their individual properties when contemplating their contributions to the assembly and structure of the procapsid and virion.

**Coat protein**

Genetic, biochemical, and biophysical analyses of the coat protein indicate that coat protein folding and assembly is more complicated than the simple polymerization of prefolded subunits. Coat protein has a defined folding pathway populated by progressively structured intermediates, beginning with unfolded monomers proceeding in stages to coat protein assembled into the procapsid. Following the expansion of the capsid shell associated with DNA packaging, the conformation of coat protein in the mature infectious particle is its native, fully folded state. The following sections describe the properties of the coat protein folding and assembly intermediates.

*Genetic analysis of coat protein in vivo.* Gordon and King examined the genetic properties of 18 conditional-lethal alleles of the coat protein gene and characterized the defects in the assembly products made by the mutant alleles under restrictive conditions in vivo (Gordon & King, 1993). These alleles were all
shown to be temperature sensitive (ts) mutants and were shown to affect assembly at a stage prior to the assembly of the mature virion, since phage carrying these alleles were still infectious at the permissive temperature after incubation at a restrictive temperature. The ts alleles did not affect the levels of coat protein synthesis at the restrictive temperature; but they did appear to affect the synthesis of scaffolding protein, whose intracellular levels are proportionally correlated with the amount of procapsid assembly. At the nonpermissive temperature, the bulk of the coat protein was found in insoluble aggregates, and there was a reduction in the number of procapsids as seen in thin sections of infected cells by electron microscopy.

These mutations were all shown to be recessive to the wild-type and failed to show intragenic complementation or dominant negative properties (Gordon & King, 1994). Based on these results, it was proposed that the mutations destabilize a folding intermediate of the coat protein making it prone to aggregation at the restrictive temperature, and the alleles were classified as tsf, or temperature sensitive for folding. In contrast to the tsf alleles for coat, five cold sensitive (cs) alleles of coat support the formation of procapsids which fail to package DNA at the nonpermissive temperature (Gordon & King, 1994).

The tsf phenotype can be ameliorated by the overexpression of the groES/EL chaperonin (Gordon et al., 1994). Overexpression of the groES/EL chaperonin raised the minimum restrictive temperature of infection for the tsf mutants. GroES/EL overexpression lowered the amount of tsf coat protein found in aggregates at restrictive temperatures. Wild-type and tsf coat proteins co-immunoprecipitated with GroES/EL in P22 infected cells. A quantitative analysis of GroES/EL binding of wild-type and tsf coat proteins showed that the amount
of wild-type coat protein bound to GroES/EL was independent of GroES/EL concentration, but the amounts of tsf coat-chaperonin complexes were dependant on the concentration of GroES/EL (Nakonechny & Teschke, 1998). The duration of association between the chaperonin and the tsf mutants was higher than for wild-type, but the time of association independent chaperonin concentration (Nakonechny & Teschke, 1998). The authors concluded that the tsf coat proteins are a substrate of GroES/EL.

That increased GroES/EL levels do not affect the amount of wild-type coat bound to chaperonin suggests that interaction of tsf coat proteins with the chaperonin is a gain of function. This is supported by the observation that mutant groES/EL alleles that had been previously shown to block the morphogenesis of phages lambda, T4 and T5 do not affect wild-type P22 growth at any temperature, while infection of strains with these alleles by P22 strains bearing the tsf coat mutants showed a marked sensitivity to temperature (Nakonechny & Teschke, 1998). In vitro analysis of three tsf coat proteins demonstrated that these mutant proteins have altered fluorescence characteristics and increased susceptibility to proteolysis when bound to GroEL, and, that the mutant proteins were bound to GroEL with greater affinity than wild-type coat (Nakonechny & Teschke, 1998). Taken together, these results indicate that the tsf amino acid substitutions convert the coat protein to a substrate of GroEL rather than exacerbate a pre-existing requirement for GroEL, and the authors conclude that it is likely that P22 does not normally utilize the GroEL/ES chaperonin (Nakonechny & Teschke, 1998).

A second genetic strategy for studying the coat protein is intragenic second site suppression of tsf coat mutations. Some of these suppressors are only
able to suppress a single tsf allele; however, a subset of these suppressor mutants, termed “global suppressors,” are single amino acid substitutions that can suppress more than one tsf allele (Aramli & Teschke, 1999). Interestingly, these tsf strains bearing global suppressor mutations retained their requirement for GroEL at the nonpermissive temperature (Aramli & Teschke, 1999). Also, the three most promiscuous global suppressor mutations (D163G, T166I, and F170L) are located near a proposed hinge in the coat protein (Aramli & Teschke, 1999). This hinge, encompassing residues 180-205, was identified by susceptibility to proteolysis and is suspected to mediate domain motions between the N- and C-terminal halves of the coat subunit (Lanman et al., 1999).

In vitro folding and stability of wild type and mutant coat proteins. The folding of coat protein and some of its mutant forms have been characterized in vitro. Unassembled coat protein can be prepared by the dissociation and denaturation of coat subunits from purified procapsid shells (Fuller & King, 1981). The dissociated, unfolded subunits can be refolded by rapid dilution of denaturant, or by dialyzing away the denaturant. Refolded wild-type coat protein was shown, by centrifugation, to be monomeric. The monomer has exposed hydrophobic surfaces and some solvent exposed tryptophans as judged by binding of bis-ANS and fluorescence emission spectra, respectively (Teschke & King, 1993). Nonetheless, reduction of tryptophan fluorescence and bis-ANS binding of the folded monomer as compared to the fully unfolded subunit suggests that the folded monomer does have a hydrophobic core (Teschke & King, 1993).

Monitoring the folding reaction through tryptophan fluorescence and bis-ANS binding suggested that folding proceeds through a long lived kinetic
intermediate of ~300 sec half-life (Teschke & King, 1993). The kinetics of folding are not significantly altered for the tsf mutant alleles relative to wild-type (Teschke & King, 1995). In contrast to the tsf amino acid substitutions, cs mutant coat proteins showed significantly slower folding kinetics at the nonpermissive temperature as compared to the wild-type (Fong et al., 1997). These data suggest that the basis of tsf mutant protein aggregation is not due to a longer lived kinetic folding intermediate.

The productivity of in vitro refolding and the kinds of aggregates formed for both wild-type and mutant alleles are both strongly dependent on solvent conditions. By altering solvent conditions, nonproductive folding can be increased, and different forms of aggregation can occur including improperly dimensioned but soluble partial shells, soluble kinetically trapped dimers/trimers and insoluble aggregates/inclusion bodies (Earnshaw & King, 1978; Speed et al., 1996; Teschke, 1999). Despite this dependence on conditions, differences in folding and aggregation of wild-type versus tsf mutant subunits in vitro have been observed (Teschke, 1999). Early analyses of the folding of tsf coat alleles demonstrated that a significant percentage of the coat protein in a folding reaction ended up as dimers or trimers that were incompetent for assembly (Teschke & King, 1995). No such multimers were observed for the wild-type protein under the same conditions. However, under these conditions, both wild-type and mutant coat proteins showed an equal propensity towards forming insoluble aggregates at the higher temperatures. Using different buffer conditions, Teschke was able to show temperature dependent aggregation in vitro for three tsf mutant coat proteins, but under the same conditions, wild-type coat protein did not aggregate appreciably (Teschke, 1999). Furthermore, under
these conditions, tsf mutants were shown by circular dichroism and differential protease sensitivity to have non-wild-type secondary and tertiary structure (Teschke, 1999). These data demonstrated that the tsf mutant phenotypes could be replicated \textit{in vitro} and that these were possibly due to differences in the folded, unassembled coat proteins.

\textit{Wild type and mutant coat protein and in vitro procapsid assembly.} Assembly of procapsid-like particles can be achieved \textit{in vitro} using purified coat and scaffolding proteins (Fuller & King, 1982; Prevelige \textit{et al}., 1988; Prevelige \textit{et al}., 1993). \textit{In vitro} assembly has been used not only to examine the kinetics and concentration dependence of coat and scaffolding subunits on assembly, but also the effect of tsf mutations on the assembly process (Teschke & Fong, 1996). Early attempts to use tsf mutant coat subunits for assembly \textit{in vitro} failed since the coat proteins were trapped as assembly incompetent dimers and trimers after refolding. By altering the conditions of refolding, Teschke was able to refold assembly competent tsf mutant coat monomers that assembled with slower kinetics than the wild-type coat protein (Teschke, 1999). The altered kinetics were probably not due to competing aggregation since these proteins did not appreciably aggregate under the conditions used for assembly. Consequently, during assembly, folded but unassembled tsf monomer is present at higher concentrations than the more readily assembling wild-type. The assembly and aggregation of tsf mutant coat proteins bearing second site global suppressor mutations were similarly examined. Interestingly, coat monomers bearing these double mutations aggregated more severely under conditions which promote aggregation, yet they assembled more quickly than the single mutant tsf parent under those conditions that promoted assembly (Teschke, 1999). These
experiments may provide insight into the mechanism of tsf mutation induced aggregation by suggesting that: 1) tsf phenotype might be due to the increased concentration of folded, but unassembled coat protein and 2) the mechanism of tsf suppression by global suppressors may functions by increasing the rate of assembly of folded monomer thereby decreasing the concentration of folded monomer, which might be in rapid equilibrium with an aggregation prone folding intermediate.

Differences between Raman spectra of unassembled coat protein and purified empty procapsid shells reveals differences between the two states of the coat protein. Tuma et al. obtained a quantitative estimate of the secondary structure monomeric coat that was 22% alpha helix, 39% beta strand, and 38% undefined secondary structure (Tuma et al., 2001b). Raman signatures suggest that the majority of the tyrosines are hydrogen bonded, the tyrosines are solvent exposed, and that the lone cysteine sulphydryl group is hydrogen bonded. Raman monitored deuterium exchange shows 30% of monomer amide hydrogens are protected from exchange and the sulphydryl of Cys405 is exchange protected. These results are consistent with a monomer that has a folded core of defined tertiary structure and some exposed hydrophobic surfaces. Low angle solution X-ray scattering indicates the monomer has a larger radius of gyration than a globular protein of the same mass, and a computationally derived model based on the scattering data suggests the monomer as being elongated with a slightly bent shape rather than compact and spherical (Tuma et al., 2001b).

Coat subunits in the procapsid show significant changes in secondary but not tertiary structure. Quantitative estimates of a 4% increase in alpha-helicity between the monomer and procapsid suggest that approximately 17 of 428
peptide bonds are altered as a result of a folding transition upon assembly (Tuma & Thomas, 1997). The authors suggest that residues of the proposed hinge, which become both hydrogen exchange protected and more protease resistant (Lanman et al., 1999) upon procapsid assembly are the source of this folding transition. Additionally, amide protons are more exchange protected in the procapsid than in the monomer (Tuma et al., 2001a). Raman spectral markers specific to tryptophans, tyrosine phenoxyls, and the cysteine sulfhydryl show that the tertiary structure of the monomer is not significantly altered by assembly (Tuma & Thomas, 1997; Tuma et al., 1998a).

Polymerization of coat subunits into a procapsid shell also imparts significant stability to the coat protein. Differential scanning calorimetry (DSC) of monomeric and assembled coat protein measured significantly different melting temperatures and enthalpies for monomers and procapsids (Galisteo & King, 1993). The transition midpoint for monomer denaturation is 40°C and has an enthalpy of approximately 600 kJ/mol (Galisteo & King, 1993). In contrast, denaturation of procapsid subunits has a Tm of 87°C and an enthalpy of 1700 kJ/mol (Galisteo & King, 1993). These results are corroborated by pressure denaturation studies of the monomeric and procapsid forms of coat protein. The coat monomer is extremely labile under pressure, denaturing at 0.5 kBar (Prevelige et al., 1994). Coat subunits in the procapsid lattice, which must undergo dissociation before they denature, are labile at 2.2 kBar pressure in a time dependant manner (Prevelige et al., 1994). The rate of procapsid dissociation under pressure increases as the temperature is lowered indicating that subunits of the procapsid lattice are entropically stabilized (de Sousa et al., 1999; Foguel et al., 1995; Prevelige et al., 1994). These results suggest that the stability of
procapsids is derived from the exclusion of ordered water present at the subunit interfaces in the unpolymerized monomer rather than from significant changes in tertiary structure; a result that agrees with Raman spectroscopy findings.

Properties of wild type and mutant coat protein during in vitro shell expansion. The process which transforms the coat subunits in the procapsids to the mature form of the coat lattice is termed “expansion.” Expansion is marked by dramatic changes in the structure of the coat protein and is, in essence, the final folded state of coat protein. The physical characteristics of expansion have been monitored by time/temperature resolved Raman spectroscopy (Tuma & Thomas, 1997; Tuma et al., 1996a; Tuma et al., 1998a), differential scanning calorimetry (Galisteo & King, 1993), and spectroscopically monitored pressure induced denaturation (de Sousa et al., 1999). Shell expansion results in significant changes in tertiary structure as seen by Raman markers of side chain orientation and environment (Tuma & Thomas, 1997; Tuma et al., 1998a). Furthermore, expansion results in a two-fold increase in overall deuterium/hydrogen exchange protection. Changes in the Raman signature between the procapsid and expanded capsid suggest that expansion is driven by burial of hydrophobic residues that accompanies repacking of secondary structural elements (Tuma et al., 1998a). This model is supported by pressure denaturation studies of procapsids composed of tsf mutant subunits (G232D and W48Q) (Foguel et al., 1995). Each of these mutant procapsids is less stable under pressure, and G232D is also more cold sensitive to denaturation under pressure than wild-type procapsids. This increased sensitivity was attributed to packing defects at the procapsid subunit interfaces, which are proposed to destabilize an already
metastable structure. Accordingly, these procapsids expand more easily and the expanded shells are as stable as wild-type shells (Foguel et al., 1995).

_Implications of coat protein folding for assembly and expansion._ All the data discussed above can be interpreted a number of ways, but nonetheless suggest what condition the unassembled coat protein is in and how these qualities may contribute to assembly and the assembled state. Intuitively, one might expect that the mechanism of aggregation is a lengthening of the lifetime of an aggregation prone intermediate, but tsf coat protein mutants fold as monomers with similar kinetics to wild-type coat. Since tsf mutant coat proteins assemble slower, these mutations might increase the concentrations of unassembled but assembly competent coat protein monomers which might somehow bias the subunits towards aggregation. Instead, the tsf mutations seem to exacerbate an inherent quality of the coat protein, viz. aggregation. That the tsf subunits are substrates for GroES/EL might suggest that the mechanism of the tsf defect is to destabilize the folded state of the monomer, yet the stabilities of tsf mutants appear unaffected. However, their folded state does seem to be affected since the tsf mutants show differences in both secondary and tertiary structure as well as increased hydrophobicity of exposed surfaces. These changes appear to promote or contribute to improper interactions between subunits by aggravating intersubunit recognition. Speed et al. mixed P22 tailspike protein (tsp) with wild-type coat protein under conditions in which both proteins would form soluble multimeric aggregates. The aggregates were resolved by native gel electrophoresis and probed with coat and/or tailspike specific antibodies. The aggregates resolved by the native gel showed no mixtures of coat and tailspike proteins, suggesting that although aggregation is a defective state of the protein,
it nonetheless retains some elements of specificity and recognition (Speed et al., 1995).

**Scaffolding protein**

*In vivo properties of the scaffolding protein and its mutants.* The most salient feature of the scaffolding protein is that it is required for virus assembly, but is not present in the mature virion. Scaffolding catalyzes the correct formation of a closed icosahedral shell of the major head protein, the coat protein, and remains associated with the procapsid (Casjens & King, 1974; King & Casjens, 1974). It is released from the procapsid at the time of shell expansion and DNA packaging (King et al., 1976). Scaffolding protein is a true catalyst of assembly since it is recycled during infection (Casjens & King, 1974). The products of coat protein assembled *in vivo* during a scaffolding deficient infection are aberrantly shaped or partial shells that are often spiral in cross section, as if the growing edges of a polymerizing shell failed to meet properly and continued to polymerize instead (Earnshaw & King, 1978). Assembly of coat protein during a scaffolding amber infection yields, in addition to aberrant assemblies, some T=7 and T=4 shells, suggesting that scaffolding protein controls not only shell curvature as necessary for shell closure, but dimension as well (Thuman-Commike et al., 1998). Intracellular concentrations of scaffolding protein may have an affect on its role in morphogenesis, since scaffolding levels are correlated with coat synthesis, and in particular, the formation of procapsids (Casjens et al., 1985; King et al., 1978). Scaffolding protein not incorporated into procapsids depresses its own synthesis and this autoregulation has been shown to occur posttranscriptionally (Wyckoff & Casjens, 1985).
Mutants of scaffolding protein reveal additional aspects of its function. Aberrant assemblies of coat protein produced \textit{in vivo} during a scaffolding defective infection do not contain the portal protein, nor do the small amounts of T=7 empty shells produced (Earnshaw & King, 1978). Likewise, at restrictive temperatures, the scaffolding protein \textit{ts} mutants S242F and Y214W fail to incorporate portal protein into their procapsids (Greene & King, 1996). Interestingly, procapsids assembled with the S242F mutant also fail to incorporate gp16 and gp20, two proteins important for infectivity but not required for virion assembly (Greene & King, 1996). Y214W procapsids do, however, contain these minor proteins (Greene & King, 1996). These findings suggest that scaffolding protein actively recruits portal protein to the procapsid complex rather than allowing portal to bind by controlling shell curvature.

Three additional scaffolding mutants with apparently altered affinity for coat protein, also have defects in the recruitment of gp16 (Greene & King, 1996). At restrictive temperatures \textit{in vivo}, the temperature sensitive scaffolding protein mutant L177I fails to package DNA and accumulates procapsids that lack gp16 (Greene & King, 1996). \textit{In vitro} this mutant scaffolding was more resistant to extraction by GuHCl from purified procapsids than wild-type scaffolding protein, suggesting that at the restrictive temperature, L177I scaffolding protein might bind coat protein more tightly (Greene & King, 1996). Procapsids assembled at restrictive temperatures by the cold sensitive mutants Q149W and Q149Y are mostly empty procapsids that have reduced amounts of gp16 and gp20 (Greene & King, 1996). Yet, paradoxically, Q149W procapsids assembled at permissive temperatures are more resistant to extraction by GuHCl at room temperature \textit{in vitro} than wild-type procapsids (Greene & King, 1996). The \textit{in vivo}
and *in vitro* phenotypes of these scaffolding protein mutants suggest that scaffolding recruits portal and the minor proteins gp16 and gp20 to the procapsid and that there is also a subtle but complex relationship between this recruiting function and the scaffolding protein’s affinity for the procapsid.

*In vitro properties of unassembled wild type and mutant scaffolding protein.* The solution properties of purified scaffolding protein have been extensively characterized. Thermal and GuHCl induced denaturation of scaffolding protein was measured both by fluorescence of the scaffolding protein’s lone tryptophan and by circular dichroism (CD) (Greene & King, 1999a). A lack of coincidence between the two unfolding curves suggested that scaffolding denaturation is not a two-state process (Greene & King, 1999a). Unfolding of scaffolding protein began around 15°C with a loss of apparent α-helicity (Greene & King, 1999a). Unfolding in GuHCl was completely reversible and thermal denaturation was at least partially reversible (Greene & King, 1999a). Protease treatment of partially denatured scaffolding protein demonstrated that unfolding reveals protease sites in the C-terminal half of the protein suggesting that a domain in the C-terminal half of the protein unfolded first (Greene & King, 1999b). This was consistent with the tryptophan being located in the N-terminal half of scaffolding protein and the fluorescence data, which gave a higher estimated Tm than that calculated from the CD data. A less stable C-terminal domain is also consistent with the thermal denaturation profile of the scaffolding Y214W mutant, which displays a sharp unfolding transition when monitored by tryptophan fluorescence. This transition is believed not to be the result of destabilized mutant protein, since the CD denaturation profile is nearly identical to the wild-type, but rather due to the additional tryptophan revealing the unfolding of the
C-terminal half not seen in the wild-type fluorescence profile (Greene & King, 1999b).

Scaffolding protein forms oligomers in solution, being in rapid equilibrium between monomer and dimer with a small amount of tetramer at physiological concentrations (Parker et al., 1997a). Truncation of the first 140 amino acids of scaffolding results in the loss of tetramer formation and a Kd of dimerization two orders of magnitude higher than wild-type (Parker et al., 1998). Hydrodynamic modeling of the scaffolding protein based on its sedimentation profile suggests the protein is an elongated rod about 21 by 242 Å (Parker et al., 1997a). Interestingly, the truncation mutant lacking the N-terminal 140 amino acids (hereafter referred to as the 141-303 fragment) has a similar sedimentation profile suggesting that the N-terminal half of the molecule is folded back on itself (Parker et al., 1998). These results are further supported by the observation that residues in the C-terminal half are more H/D exchange protected in the full-length protein than in the deletion mutant (Tuma et al., 1998b). Although the wild-type and 141-303 fragment show differences in exchange protection, the full-length scaffolding protein is unusually labile to deuterium exchange with all amide protons being exchanged by 30 minutes at 20°C (Tuma et al., 1998b). The amide exchanges can be grouped into fast and slow exchanging populations. Raman spectra of an 11 amino acid C-terminal truncation mutant of scaffolding protein consisting of residues 141-292 indicates to a loss of alpha helicity to an additional 5-10 residues. This truncation mutant is assembly incompetent in vitro suggesting the C-terminal 30 amino acids may participate in coat protein binding (Parker et al., 1998).
Scaffolding protein and in vitro assembly of procapsids. In vitro assembly of procapsids with wild-type and mutant scaffolding proteins reveal aspects of its function. Partial proteolytic digests of scaffolding protein when mixed with purified unassembled coat protein leads to assembly of the coat protein into procapsid-like structures (Greene & King, 1999b). Purification of the assembled products and separation of their proteins by SDS-PAGE reveal two proteolytic fragments with the same N-terminal sequence, but with 30 amino acid difference at the C-termini, the longer of the two being assembly competent (Greene & King, 1999b). This finding suggested the C-terminal 30 amino acids are part of an assembly domain of scaffolding protein.

Prevelige et al. saw an apparent second order dependence of scaffolding protein concentration on the rate of procapsid assembly in vitro (Prevelige et al., 1993). The authors concluded that dimers of scaffolding protein were active in assembly. This hypothesis is further supported by the observation that covalently linked scaffolding protein dimers are more efficient at promoting assembly than unmodified scaffolding protein (Parker et al., 1998). The N- and C-terminal truncation mutant of scaffolding protein consisting of residues 141-292 cannot promote assembly, in vivo or in vitro (Parker et al., 1998). However, when mixed with wild-type scaffolding protein, assembly is inhibited, suggesting that the wild-type and mutant proteins heterodimerize and cannot participate in assembly (Parker et al., 1998). It has been suggested that the favorable energy of scaffolding protein dimerization is used to prolong association between coat protein monomers increasing the probability of binding (Sun et al., 2000).
Portal protein

Portal protein in vivo. The portal proteins of bacteriophages form multimeric rings through which DNA is packaged into the capsid. The P22 protein gp1 was shown to be a portal protein when it was demonstrated that 1) portal-less procapsids fail to package DNA, 2) that the portal protein was an integral part of the procapsid/capsid structure, 3) that it was present at only one vertex in the virion, and 4) that, in vitro, it formed dodecameric rings with a 25-30Å channel through the center. The presence and absence of the portal protein has been shown to have no effect on the rate of coat and scaffolding assembly in vivo, suggesting it is likely not part of a kinetic initiator of assembly (Bazinet & King, 1988; Bazinet et al., 1988; Botstein et al., 1972; King et al., 1976). The mechanism of portal incorporation remains unknown. Although the absence of portal protein does not affect the fidelity of coat shell assembly, overexpression of portal protein during P22 morphogenesis leads to defects in assembly including the formation of T=7 shells containing twice the normal complement of portal protein that are also incompetent for packaging of DNA as well as the formation of smaller T=4 shells that are able to package ~70% of a genome and mature into stable, tailed mini-capsids. T=4 shells of coat protein are observed during scaffolding deficient infections, although these particles do not contain portal protein. These results argue for an indirect effect of portal protein levels on procapsid form determination. Excess portal protein may be titrating away scaffolding protein needed for correct subunit-subunit interactions during the growth phase of coat polymerization. Indeed, genetic escape mutants able to overcome the toxicity of portal protein overexpression all map to a single amino acid substitution, G287E, in the scaffolding gene, a residue in the loop of the
helix-loop-helix motif of the scaffolding coat protein binding domain (Moore, 2001).

Genetically, portal has been shown to have additional interactions with other phage morphogenetic proteins. Suppressors of a cold sensitive portal mutant were isolated in the gene for tailspike protein (Bazinet et al., 1990). Tailspike protein addition occurs late in P22 assembly, after procapsid formation and DNA packaging. It has been shown, in vitro, that tailspike and the plug proteins (gp 4, gp10, and gp26) as well as terminase (gp2/3) can interact with portal rings, but only tailspike and the plug proteins interacted with portal protein monomers (Moore, 2001). Suppression might therefore involve competition between tailspike and terminase for prematurely assembled rings that might otherwise have inappropriate interactions with phage DNA.

Portal protein in vitro. Portal protein has been purified from lysates of cells infected with coat protein deficient phage P22. It appears to be monomeric and only assembled into rings after prolonged storage at cold temperatures suggesting a kinetic barrier to ring assembly (Bazinet et al., 1988). Purification of a recombinant histidine-tagged portal showed similar properties (Moore & Prevelige, 2000). At concentrations greater than 40 mg/ml, portal protein spontaneously assembles into rings. Assembled and unassembled forms can be quantified at different times during ring assembly, allowing for the determination of an apparent rate constant. Arrhenius analysis of ring assembly at different temperatures yielded an activation energy of 12 kcal/mol for assembly. The basis of a kinetic barrier to portal oligomerization could be due to partial disorder of an N-terminal domain (Moore & Prevelige, 2000). Differences in the folded states for the assembled vs. unassembled state are
increased $\alpha$-helicity in the ring form as measured by CD, decreased susceptibility to proteolysis after ring assembly, and higher binding of the hydrophobic dye, bis-ANS, to the monomeric form (Moore & Prevelige, 2000). As previously mentioned, both monomeric and dodecameric forms of portal protein have been used to find binding partners \textit{in vitro} during affinity chromatography. These experiments did not find an interaction between portal and scaffolding proteins.

\textbf{Scaffolding mediated assembly}

\textbf{in other viruses}

A scaffolding protein-assisted capsid assembly strategy is employed by essentially all dsDNA bacteriophages as well as the medically relevant Herpesviridae. It is not known if this strategy is derived from a single ancestor, or may have been arrived at multiple times during evolution. The following summarizes and discusses some of the well characterized phage assembly systems.

\textbf{Phage SPP1}

The bacteriophage SPP1 is a T=7 isometric dsDNA tailed virus of Bacillus subtilis (Becker \textit{et al.}, 1997). SPP1 procapsids consist of major capsid, scaffolding, and portal proteins as well as an accessory protein, gp7 (Becker \textit{et al.}, 1997; Droge \textit{et al.}, 2000). Co-expression of SPP1 scaffolding protein, gp11, and the major capsid protein, gp13, is sufficient to form closed particles of procapsid-like morphology (Droge \textit{et al.}, 2000). Gp13 expression by itself forms heterogeneous aberrant structures. gp11 expression is reduced in the absence of gp13 suggesting that gp11 expression, like the P22 scaffolding protein, is autoregulated. Co-immunoprecipitation of SPP1 procapsid proteins demonstrated an interaction
between scaffolding and capsid proteins and between the portal protein and the
accessory protein (Droge et al., 2000). Interestingly, although the presence of
scaffolding protein increased the interaction between coat and portal proteins, no
direct scaffolding portal protein interaction was observed. The absence of the
portal protein during procapsid assembly results in a subpopulation of smaller
particles, probably of T=4 symmetry. Like P22, the expression of portal protein
had no effect on the rate of capsid protein/scaffolding assembly in vivo (Droge et
al., 2000). The physical and structural properties of the SPP1 scaffolding protein
in solution have so far not been characterized.

Bacteriophage T7

T7 also has genes that encode proteins with coat, scaffolding, and portal
functions (Dunn & Studier, 1983; Steven et al., 1983). Coexpression of the head
protein, gp10 and the scaffolding protein, gp9, leads to the formation of small
amounts of closed capsid like structures consisting of the two proteins alone
(Cerritelli & Studier, 1996). Similar structures could be formed in vitro by
incubation of purified head and scaffolding proteins in the presence of 20% polyethylene glycol 200 (Cerritelli & Studier, 1996). Like P22 assembly, this
assembly was inhibited by high ionic strength. Lowering the ratio of scaffolding
to coat protein reduced the amounts of assembly as well, however, the head
protein assembled into polycapsids suggesting that once assembly was initiated
by scaffolding, stoichiometric amounts were needed to maintain the fidelity of
assembly. Because of its purification properties by size exclusion chromatography, as well as a high degree of protease sensitivity, the T7 scaffolding protein is suggested, like the P22 scaffolding protein, to be elongated
and flexible (Cerritelli & Studier, 1996). Also like P22, the T7 scaffolding protein contains a high amount of negatively charged amino acids, as deduced from the amino acid sequence (Dunn & Studier, 1983).

**Bacteriophage T4**

Bacteriophage T4 utilizes more than one scaffolding protein (Caldentey *et al.*., 1987; Keller *et al.*., 1988). The major scaffolding protein, gp22, forms structures in the absence of the capsid protein (Caldentey *et al.*., 1987). Like other phages, T4 capsid protein forms aberrant structures in the absence of the scaffolding protein, the most dramatic of these being polyheads which are large tubes of polymerized coat protein (DeRosier & Klug, 1972). The coat proteins assembled into these tubes are able to undergo conformational changes during expansion that are similar to those seen in normal heads (Steven *et al.*., 1976). That the T4 scaffolding protein plays a role in size and morphology of the head is also seen in mutants of the scaffolding protein that result in the formation of isometric heads (Keller *et al.*., 1988). T4 gp22 scaffolding protein leaves the capsid during DNA packaging after approximately ~25% of a headful has been packaged (Jardine & Coombs, 1998). Proteolytic clipping of the T4 scaffolding protein appears to be the signal for exit, and a small fragment of the protein is retained in the capsid (Wagner & Laemmli, 1976). T4 also has minor scaffolding proteins and proteins that are part of the scaffolding core. Among these proteins are gp68 and IPIII. These proteins are targeted to the head by a 10 amino acid capsid targeting sequence (Mullaney *et al.*., 2000). This sequence fused to a heterologous protein can mediate the incorporation of that protein into the T4 capsid (Mullaney *et al.*., 2000).
**Bacteriophages P2/P4**

The *E. coli* bacteriophage P2 is an isometric, T=7 icosahedral bacteriophage with a contractile tail. Plasmid driven expression of the P2 major capsid protein gpN results in the formation of both aberrantly polymerized gpN as well as a small fraction of closed T=4 and T=7 shells (Wang *et al.*, 2000). Coexpression of the major capsid protein gpN with the P2 scaffolding protein, gpO, increases the amount of T=7 shells, although gpO is not found associated with these structures (Rishovd *et al.*, 1994). It is not known if the gpN-gpO association is more transient in the assembly pathway or is more labile so that it is lost during purification. The P2 coat protein can also interact with an external scaffolding protein, gpSid, encoded by the smaller parasitic phage P4, whose genome encodes no major capsid protein of its own. Assembly of gpN with gpSid forces gpN to form T=4 capsids which package the smaller P4 genome. Structures of the T=7 gpN shell as well as T=4 gpN-gpSid shells have been determined by cryo-electron microscopy (Dokland *et al.*, 2002). gpSid forms an external lattice of trimeric complexes and appears to act as “belts” that constrain the gpN to smaller symmetry (Dokland *et al.*, 1992; Marvik *et al.*, 1995).

**Bacteriophage φX174 (Microviridae)**

The small, icosahedral ssDNA bacteriophage φX174 is like the dsDNA bacteriophages in that it utilizes scaffolding proteins in the assembly of a procapsid into which viral DNA is packaged. φX174 uses internal and external scaffolding proteins (Dokland *et al.*, 1999). The internal scaffolding protein, gpB, is used to assemble a pentamer of capsid protein, gpF, and a pentamer of the vertex spike protein, gpG into preformed pentons (Ekechukwu & Fane, 1995). An
external scaffolding protein, gpD, is used to assemble twelve of these preformed pentamers into a T=1 icosahedral shell (Ekechukwu & Fane, 1995). Why two scaffolding proteins are required to mediate the assembly of 60 subunits of coat protein that make identical contacts is unknown, but it has been speculated to involve the incorporation of the spike protein at each vertex (Dokland et al., 1997).

The crystal structure of the øX174 procapsid is known (Dokland et al., 1997). The structure reveals not only the structure of the capsid protein, but that of the external scaffolding protein as well. The structure of the external scaffolding, gpD, displays a remarkable conformational plasticity; four subunits bind each icosahedral face using two nonidentical conformations. Roughly twenty amino acids of the internal scaffolding protein’s C-terminus are ordered in the crystal structure and are in an α-helical conformation. Like bacteriophage P22, there appears to be a region of disorder N-terminal to the coat protein binding domain of gpB, and the remainder of the molecule is probably not icosahedrally ordered.

**Herpesviridae**

The virion morphogenesis of the Herpesviridae rather closely resembles that of the dsDNA bacteriophages. For those Herpesviridae whose assembly pathways have been characterized, formation of the T=16 capsid (Newcomb et al., 1996) begins with the scaffolding protein assisted assembly of a procapsid into which DNA is packaged and from which the scaffolding protein must exit (Homa & Brown, 1997). Cleavage of the scaffolding protein by a protease packaged in the procapsid triggers the exit of the scaffolding protein (Newcomb
et al., 2000). The protease is translated from an open reading frame that overlaps the scaffolding protein such that the C-terminus of the protease and the scaffolding protein are identical, which probably serves to ensure the localization of the protease to the procapsid (Newcomb et al., 1996).

Other than the P22 scaffolding protein, the Herpes virus scaffolding proteins are the best characterized scaffolding proteins. Binding to the major capsid protein is mediated by a 12 amino acid domain at the extreme C-terminus of the scaffolding protein (Hong et al., 1996). This binding domain is dominated by hydrophobic residues. Replacement of these residues with charged residues, as well as the presence of chaotropic detergents abolished the interaction between the scaffolding and capsid proteins in an in vitro protein interaction assay, suggesting that scaffolding binding is largely a hydrophobic interaction (Hong et al., 1996), unlike P22 where electrostatic interactions have been suggested to play a significant role.

Like P22, the herpes scaffolding protein forms homomeric interactions that are mediated by sequences N-terminal to the capsid binding domain (Oien et al., 1997). These self interactions are proposed to be coiled-coil interactions between parallel scaffolding proteins (Pelletier et al., 1997). Interfering with the scaffolding-scaffolding interactions reduced binding to the capsid protein in vitro, but heterologous replacement of the scaffolding N-terminal sequences with the GCN4 leucine-zipper, which itself forms coiled-coil interactions, was sufficient to sustain coat protein binding (Pelletier et al., 1997). These results suggest that stable self-association of the herpes scaffolding protein is required for efficient procapsid assembly.
CHAPTER 3

FUNCTIONAL DOMAINS OF
THE BACTERIOPHAGE P22
SCAFFOLDING PROTEIN

Introduction

The bacteriophage P22 scaffolding protein mediates assembly of the procapsid, a 27 megadalton protein complex that is a preformed container into which the viral genome is packaged. All of the scaffolding protein exits from the structure during virus maturation. Scaffolding protein-mediated procapsid assembly is a strategy employed by most dsDNA phages (Dokland, 1999), the ssDNA phage øX174 and related Microviridae (Burch et al., 1999), and the Herpesviridae viruses herpes simplex virus type 1 (HSV-1) and human cytomegalovirus (CMV) (Gibson & Roizman, 1972).

The role of scaffolding protein in virus morphogenesis is best understood for the development of the Salmonella bacteriophage P22. The procapsid is the first structure detected in the phage assembly pathway. It is formed by the co-assembly coat scaffolding protein (located in the interior of the procapsid but not found in the mature virion) into an T=7 icosahedral shell composed of ~415 coat subunits, ~250-300 molecules of scaffolding and a dodecameric ring of portal protein located at a unique vertex. Although coat protein subunits can
polymerize at high concentrations \textit{in vivo} and \textit{in vitro}; the presence of scaffolding protein lowers the concentration at which coat subunits assemble (Prevelige \textit{et al.}, 1988). The assembly-active coat protein species is monomeric (Prevelige \& King, 1993) and scaffolding protein is in equilibrium between monomers, dimers and tetramers in solution (Parker \textit{et al.}, 1997a). There is evidence that dimers and monomers of scaffolding protein are the active species in assembly (Parker \textit{et al.}, 1998; Prevelige \textit{et al.}, 1993).

The P22 scaffolding protein also acts as an assembly chaperone that ensures the correct size and curvature of the procapsid shell (Earnshaw \& King, 1978). When scaffolding protein is defective \textit{in vivo}, coat protein assembles a heterogeneous mixture of structures ranging in defect from insoluble inclusion bodies to soluble aggregates such as, polyheads, spiral shells and a few $T=4$ and $T=7$ shells that lack portal (Thuman-Commike \textit{et al.}, 1998). This polymorphism of assembled structures presumably reflects the conformationally flexible nature of the coat protein, which is required for the protein’s ability to occupy seven nonidentical positions within polymerized shell.

The mechanism by which scaffolding protein promotes assembly of the coat protein is unknown. Since high concentrations of coat protein self-assemble, unassembled coat protein may be in equilibrium between assembly competent and incompetent states. Scaffolding protein may recognize, bind and stabilize an assembly-competent state of coat protein, thereby increasing the concentration of assembly-competent monomers of coat protein to levels favorable for polymerization. An alternate possibility is that binding of scaffolding protein induces a conformational change in coat protein converting it to an assembly competent state. Both aggregation and assembly of coat protein are irreversible,
therefore it is difficult to hypothesize a thermodynamic model for scaffolding protein's chaperone function.

When assembled into a procapsid, the coat protein occupies seven distinct quasi-equivalent positions (Casjens, 1979; Prasad et al., 1993). Furthermore, as DNA is packaged, the procapsid shell undergoes an exothermic expansion resulting in seven new “final states” of coat protein (Galisteo & King, 1993; Prasad et al., 1993). Genetic evidence suggests coat protein has folding intermediates. Of 18 temperature sensitive mutations in coat, 15 are temperature sensitive folding (tsf) mutations that affect the folding of the coat protein resulting in aggregation and formation of inclusion bodies at the nonpermissive temperature (Gordon & King, 1993; Gordon & King, 1994). However, capsids assembled at the permissive temperature by tsf coat mutant proteins show no defects in stability (Gordon & King, 1993). Scaffolding protein’s chaperone action might be a kinetic result of its ability to lower the concentration required for initiation of coat protein assembly or through altering the conformation of coat protein in such a way as to prevent incorrect or inaccurate assembly. In this way the scaffolding protein acts as a kinetic trap or partition by diverting coat protein from aggregation pathways during folding and assembly towards the correct assembly pathway resulting in closed, correctly sized shells.

Other functions of the scaffolding protein have been correlated to regions of the primary sequence of the scaffolding protein. Scaffolding protein is required for the incorporation of the portal protein, as well as minor proteins (gp7, gp16, and gp20) required for infectivity, into the procapsid (Greene & King, 1996). The scaffolding gene missense mutations Y214 W and S242F are conditional lethal at restrictive temperatures (Greene & King, 1996). Procapsid-
like structures that contain the mutant scaffolds are assembled \textit{in vivo} at the nonpermissive temperature but they lack portal. Scaffolding protein autoregulates its own expression, through the action of an N-terminal domain (Casjens & Adams, 1985). Scaffolding protein controls the assembly of the coat protein into the procapsid lattice, therefore, autoregulation of scaffolding protein adds additional control to the assembly process (Casjens \textit{et al.}, 1985). Limiting concentrations of scaffolding protein \textit{in vivo} might ensure fidelity in the assembly process and maximize the production of infectious particles. Scaffolding protein in solution at physiological concentrations \textit{in vitro} is in equilibrium between monomeric, dimeric and tetrameric forms and the Kd of dimerization has been determined to be approximately 50 \(\mu\)M. However, the amino acid 141-303 fragment is severely reduced in its ability to form dimers having a Kd of 1500 \(\mu\)M (Parker \textit{et al.}, 1997a). Therefore, this truncation lacks a functional domain necessary for tight dimerization. Finally, during the morphogenesis of the phage virion, scaffold exit is required before DNA packaging can be completed. Greene and King have isolated mutants of scaffolding protein that fail to exit at the nonpermissive temperature (Greene & King, 1996). These mutants are blocked in phage maturation and accumulate procapsids (Greene & King, 1996). These exit mutant procapsids are also resistant to having their scaffolding protein extracted by treatment with GuHCl relative to wild-type (Greene & King, 1996).

In spite of having no observable hydrophobic core (Tuma & Thomas, 1997), this highly flexible, elongated molecule nevertheless contains functional domains that enable it to perform multiple, complex functions. In the present study, we attempt to increase the resolution at which some of these functional domains are defined within the primary amino acid sequence of the scaffolding
protein by testing the ability of N- and C- terminal truncation mutants of scaffolding protein to complement some of these functions during scaffolding defective infections.

**Materials and methods**

**Plasmid construction**

Plasmids encoding scaffolding protein truncation mutants for use in the *in vivo* assembly assay were made by cloning fragments of P22 gene 8 into the expression vector pET3a (Novagen, Madison, WI) between the NdeI and BamHI restriction sites (Rosenberg et al., 1987; Studier & Moffatt, 1986; Studier et al., 1990). Gene fragments were generated by polymerase chain reaction (PCR) using tailed primers containing appropriate restriction sites, sequences complementary to the desired regions of the gene and, where necessary, either a start (ATG) or a stop codon. These primers are listed in Table 3.1. Restriction sites are underlined, P22 complementary sequences are boldfaced, and start and stop codons are italicized. PCR reactions were performed in 50 μl reactions, using the Elongase® enzyme mixture (Invitrogen, Carlsbad, CA), consisting of Taq polymerase and a proofreading enzyme for greater processivity and accuracy, in the supplied buffer with 1.5 mM MgCl₂, 1 μM of each primer, and 200 μM each dNTP. The cycling parameters included an initial 2-minute 94°C denaturation step followed by 30 cycles consisting of a 30-second, 94°C denaturation step; a 30-second, 55°C annealing step; and a 1-minute, 68°C extension step. Enzymes as well as unincorporated nucleotides and primers were removed by purification using a Qiaquick PCR Cleanup Kit (Qiagen, Valencia, CA).
### Table 3.1

Primers used in PCR cloning of scaffolding protein truncation and fusion mutants

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
<th>1st Codon</th>
</tr>
</thead>
</table>
| 8AJ  | GACGATGCA
CATATGGCTGCAACAAACAAAGACCGG | Ala52 |
| 8AK  | GACGATGCA
CATATGGCTCGGAATAAGTCGCCGCG | Ala196 |
| 8AL  | GACGATGCA
CATATGGCTGAACCCGAGAGCCCGCC | Ala218 |
| 8AM  | GACGATGCA
CATATGGCTGCGTACCTATTACCGG | Gly266 |
| 8W   | GACGATGCA
CATATGGCTGCGAATCGTACAGC | Ala141 |
| 8AH  | GACGATGCA
CATATGGCTGAACCCGAGAGCCCGCC | Ala218 |
| 8AS  | GACGATGCA
CATATGGCTGCGAATCGTACAGC | Ala218 |
| 8AT  | GACGATGCA
CATATGGCTCTCACTCGACTATCGGAACGC | Leu238 |
| 8AU  | GACGATGCA
CATATGGCTCGCTTAACTCTCAAGCCTCGC | Arg254 |
| 8AV  | GACGATGCA
CATATGGCTGCTACCCATGCTGACCC | Ala257 |
| 8BG  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Ala270 |
| 8BH  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Ala277 |
| 8BI  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Ala277 |
| 8BJ  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Ala275 |
| 8BL  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Ala279 |
| 8BM  | GACGATGCA
CATATGGCTCAGCAAAATAAGGAGGCCATTCG | Met280 |
| 8KSD  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Met1 |
| 821A  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Ala21 |
| 836A  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Ala36 |
| 858A  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Ala58 |
| 864A  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Ala64 |
| 8L41A  | AAACGGGG
GATTACC(AAGGAGATATA)ATGGAAACAAACCACCGAAATTCAGGC | Ala141 |

**Upstream Primers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
<th>1st Codon</th>
</tr>
</thead>
</table>
| 8XX  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Pro25 in gp5 |
| 8CA  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Lys296 |
| 8CC  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Lys294 |
| 8AS  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Tyr292 |
| 8XX  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Pro25 in gp5 |
| X8XR | TAAGGGCT
CTGAGTAGTGGAGGTGATGCGGACCGAATTC | Pro25 in gp5 |

**Downstream Primers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence*</th>
<th>Last Codon</th>
</tr>
</thead>
</table>
| 8X   | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Pro25 in gp5 |
| 8CA  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Lys296 |
| 8CC  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Lys294 |
| 8AS  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Tyr292 |
| 8XX  | GTAGAGAGG
CATCTGGAGGTGATGCGGACCGAATTC | Pro25 in gp5 |

*Sequences are given 5' to 3'. Non-coding tails are italicized, restriction sites boldfaced, and complementary coding sequences follow in nonitalicized text. Special sequences are in parentheses.

The first or last amino acid codon of scaffolding gene is underlined.

*bEncodes NdeI recognition site.

*cEncodes BamHI recognition site.

*dEncodes KpnI recognition site.

*eEncodes XhoI recognition site.

*fSine-Dalgarno sequence in parentheses.

*iEndogenous stop codon for gp8 is located within the overlapping gp5 reading frame.
bovine serum albumin (BSA). After one hour at 37°C, NaCl was added to a final concentration of 100 mM, and 20 units of *Bam*HI were added. After incubation at 37°C for an additional hour, restriction enzymes were removed either by phenol/chloroform (1:1) extraction followed by ethanol precipitation or by using a Wizard® DNA cleanup kit (Promega, Madison, WI). Digested DNA was incubated overnight at 16°C with five units of T4 DNA ligase in a reaction volume of 20 µl. Ten µl of the ligation mixture were used to transform CaCl$_2$ competent (Hanahan, 1983) *E. coli* strain NF1829 (Shultz et al., 1982) to ampicillin resistance on L-plates with 100 µg/ml ampicillin. Plasmids were recovered from transformants by alkaline-lysis mini-prep method using a Qiaprep spin miniprep kit (Qiagen, Valencia, CA) and the presence of cloned inserts was determined by visualization of *Nde*I and *Bam*HI DNA fragments agarose electrophoresis gels stained with ethidium bromide.

Plasmids encoding scaffolding protein fusion mutants were made by cloning fragments of the scaffolding protein gene into the glutathione-S-transferase (GST) fusion protein expression vector pGEX4T-1 (Smith & Johnson, 1988) (Genbank accession: U13853) between its *Bam*HI and *Xho*I restriction sites using the same methodology described above. Cloning scaffolding gene fragments into, and expressing the fusion protein from pGEX4T-1 results in a scaffolding protein fragment with an N-terminal fusion to 26 kDa GST protein.

Plasmids used to test viability of scaffolding truncation mutants were made by cloning gene fragments into pNFPW. pNFPW was constructed by replacing the *P$_{lac}$* controlled operon of pTAT13 (Franklin, 1989) with a similar operon from p2702 (both plasmids were a gift from N. Franklin), resulting in plasmid pULT1. The “tester operon” of p2702 is similar to that of pTAT13, but
also contains a P22 N-protein utilization site (nut) allowing, in the presence of the P22 N protein, readthrough of transcriptionally terminated genes (N. Franklin, personal communication). The EcoRI/Scal fragment of pULT1 was replaced with the EcoRI/Scal fragment from pGEX4T-1 containing the lacI⁰ repressor. The resultant vector was named pNFPW. The promoter and adjacent nut site are followed by an open reading frame encoding the bacterial alkaline phosphatase, phoA. Expression of phoA in the E. coli strain NF1829, which has low endogenous alkaline phosphatase activity, in the presence of 0.1 mM 5-bromo-4-chloro-3-indolylphosphate-p-toluidene (XP), a chromogenic substrate, yields blue colonies on agar plates. Expression of phoA from pNFPW can be induced by isopropyl-1-thio-β-D-galactopyranoside (IPTG) in solid medium at a concentration of 0.1 mM. The phoA gene is flanked by a unique 5’ KpnI site and a unique 3’ XbaI site. Truncations of the scaffolding gene were generated by PCR using primers listed in Table 3.1. Sine-Dalgarno sequences were provided by the upstream primers used to amplify the scaffolding gene fragments. These PCR-generated gene fragments and vector were digested and ligated as described above using KpnI, XbaI and T4 DNA ligase (New England Biolabs, Beverly MA). Blue/white colony screening was used to detect the insertion of new DNA between these cloning sites. White colonies were picked, plasmid DNA was isolated, and plasmids with inserts were confirmed by restriction digest with KpnI and XbaI followed by agarose electrophoresis.
**In vivo coat protein binding assay**

Truncation and fusion mutants were expressed in *Salmonella typhimurium*. Truncation mutants cloned into pET3a (Novagen, Madison, WI) (Studier *et al.*, 1990) were transformed into TR6578 cells already carrying the T7 RNA polymerase encoding plasmid pGPl-2 (Tabor & Richardson, 1985). TR6578 is strain LB5000 of Bullas and Ryu (1983). T7 RNAP expression is repressed by the temperature labile (cl ts857) phage lambda repressor, which is expressed constitutively from the plasmid. Expression of truncated scaffolding protein mutants was induced by shifting these cells from 30°C to 37°C. After fifteen minutes, cultures were infected at a multiplicity of infection (MOI) of 5 with P22 strain 8-amN123, 2-amH202, 13-amH101, cl-7. This strain was deficient in DNA packaging (2-') to assure that products of phage assembly could not progress beyond the procapsid stage and was defective in lysis of the host cells (13-'), which allowed the products of phage assembly to be concentrated by pelleting the infected cells. Establishment of lysogeny was averted by the inclusion of a clear mutation (cl-). After 120 minutes at 37°C, the infected cells were pelleted, resuspended in TM (10 mM Tris pH 7.5, 1 mM MgCl₂) and lysed by vortexing with 10% CHCl₃. Excess free DNA was digested by addition of DNaseI (Roche Applied Science, Indianapolis, IN) at 0.5 µg/ml to lower the viscosity, and cellular debris was removed by centrifugation at 10,000 RCF for 20 minutes at 4°C. The supernatant was layered onto 20% sucrose dissolved in TE7.5 in an ultracentrifuge tube. TE7.5 is 10 mM Tris-hydroxymethyl-aminomethane (Tris) chloride at pH 7.5 and 1 mM ethylenediaminetetraacetic acid (EDTA). The tube was centrifuged at ~100K RCFs for 90 minutes. The pellet containing the large molecular weight (MW) complexes was carefully drained and resuspended with
gentle agitation overnight at 4°C in TE7.5. The resuspended pellet was applied to tubes containing 10-40% sucrose gradient and separated by centrifugation in a SW41 swinging bucket rotor (Beckman-Coulter, Fullerton, CA) at 25K rpm for 120 minutes. Gradients were formed using a Gradient Master (BioComp Instruments, New Brunswick, Canada) and were harvested by fractionation on a Fractionator (BioComp Instruments). Fractions were analyzed by SDS-PAGE, native electrophoresis through 1% agarose, and negative stain electron microscopy. Co-migration of coat and scaffolding through the gradient and the presence of intact shells, as indicated by native gel electrophoresis and/or electron microscopy, demonstrated interaction of coat and scaffolding protein.

**Plaque assay**

To determine the minimal domain of scaffolding protein required for virion production, truncation mutants were cloned into pNFPW under the control of the IPTG inducible Ptac promoter as described above. The resulting plasmid was used to transform TR6578 by electroporation under conditions described elsewhere (Casjens et al., 1991). An 0.2 ml aliquout of an overnight culture of each of the resulting transformants was plated by soft LB agar overlay on LB agar plates with ampicillin selection (100 µg/ml) and IPTG at 0.2 mM in the bottom agar. Serial dilutions of scaffolding deficient phage (8" amN123) were spotted on to the cells which were constitutively expressing full-length or truncated scaffolding protein from the pNFPW derivative plasmids, and plates were incubated overnight at 37°C. The resulting lawns examined for the presence of plaques.
Western blotting

Unincorporated portal protein is insoluble when lysates are made by lysis with CHCl₃ (P. Weigele, unpublished observation); therefore whole cell lysates were prepared by boiling cells that have been resuspended in 1X SDS-PAGE sample buffer (Ausubel, 1987; Ausubel, 1999). Whole cell lysates from infections and purified procapsids containing truncated mutants of the scaffolding protein were resolved by SDS-PAGE (Ausubel, 1987; Ausubel, 1999). Proteins were electroblotted to PVDF membrane (Bio-Rad Laboratories, Hercules, CA). The presence of portal protein was determined by probing with mouse anti-portal polyclonal antibody (gift of S. Moore and P. Prevelige). Antibody bound to immobilized portal protein was visualized by an alkaline phosphatase (AP) conjugated anti-mouse secondary antibody and chemiluminescent detection using a Phototope Star® detection kit (New England Biolabs, Beverly, MA).

Negative staining and electron microscopy

Suspensions of phage or procapsids at concentrations between $10^4$-10$^{10}$ particles per milliliter were spotted on 400-mesh, formvar/carbon coated copper grids (Ted Pella, Redding, CA) for 1 minute. Grids were rinsed with three drops of buffer TM (Tris, pH 7.6, 1mM MgCl₂) and stained for 1 minute with 2% uranyl acetate. Stain was wicked away with bibulous paper and grids were dried under vacuum with dessicant. Virus particles were visualized by electron microscopy using a Hitachi H7100 transmission electron microscope (Hitatchi Instruments, San Jose, CA) with an accelerating voltage of 75 kV and a magnification between 50,000X and 60,000X.
Results

Scaffolding protein’s N-terminal 279 amino acids and 9 C-terminal amino acids are dispensable for binding coat protein

To define the boundaries of the coat protein binding domain of scaffolding protein, N- and C-terminal truncations of the scaffolding protein were constructed. These truncation mutants were tested for their ability to assemble into procapsids using an \textit{in vivo} assembly assay. Scaffolding fragments were expressed from a pET3a derived plasmid from which the mutant scaffolding proteins were expressed from a T7 RNA polymerase promoter. These cells were infected with P22 phage deficient in scaffolding protein (gp8); gp2, a terminase subunit that is essential for DNA packaging; and the lysis protein, gp13. The $8^-$ nonsense mutation causes synthesis of only a 7 amino acid N-terminal fragment of scaffolding protein. Because this phage cannot package DNA, the products of infection cannot progress beyond the procapsid stage and, due to the lysis defect, those products accumulate intracellularly.

120 minutes after infection, any large structures produced during these infections were first purified by pelleting them through 20\% sucrose in an ultracentrifuge. These procapsid-like structures, mainly assembled coat protein and proteins bound to them, were further purified by sedimentation into a 10-40\% sucrose gradient. Gradients were fractionated and their protein content resolved by SDS-PAGE. Co-migration of coat protein and the scaffolding protein mutant through the gradient was considered as evidence for binding of the mutant scaffolding protein to coat protein. Representitive gels are shown in
Figure 3.1 and the overall results of these experiments are summarized in Table 3.2.

Using this assay for the ability of scaffolding protein to interact with coat protein, we found that truncated scaffolding proteins missing the N-terminal 20, 51, 140, 188, 195, 217, 219, 228, 237, and 243 acids are able to bind to coat protein structures. The sedimentation of procapsids containing endogenous full-length scaffolding protein (303 amino acids) from a P22 2', 8' infection is shown in Figure 3.1, panel a. Figure 3.1, panel c, shows data for the smallest stably expressed scaffolding protein fragment composed of residues 244-303, SCAFF244-303. The portion of the gradient that contained this scaffolding fragment corresponds exactly to the location that wild-type procapsids sediment as seen in panel a of Figure 3.1. Negative stain electron microscopy and electrophoresis in agarose gels confirmed that the fragments are largely in procapsid-like particles that appear to be filled by electron microscopy (data not shown). A few aberrant particles were present in these fractions, but it is not known if they contained bound scaffolding fragment.

Truncation mutants shorter than SCAFF244-303 were difficult to both express stably and visualize by standard SDS-PAGE. Therefore scaffolding protein’s N-terminus was replaced with glutathione-S-transferase (GST), a 26 kD globular protein from *Schistosoma japonicum* (see Table 3.2). GST alone does not associate with procapsids or coat protein aggregates (see Figure 3.1, panel d). The small amount of GST that co-sediments at the procapsid position in gradient in panel d most likely represents non-specific trapping within topologically closed coat protein shells, since it is very highly expressed and is the most abundant protein in these cells (data not shown).
Figure 3.1 Sucrose gradient analysis of in vivo assembly products. Material harvested after an infection was pelleted through 20% sucrose, resuspended (L) and resolved on a 10-40% sucrose velocity sedimentation gradient. The protein composition of gradient fractions was resolved by SDS-PAGE through 12% gels and visualized by silver staining. Infection with a packaging and lysis deficient phage (P22 2-, 13-) results in the production and accumulation of "wild-type" procapsids (a). A packaging, lysis, and scaffolding deficient phag (P22 3-, 13-, 8-) produces large MW aggregates of coat that migrate through the gradient (b). The same phage as in b. when supplied with protein in trans during infection can incorporate it into larger structures as seen when a scaffolding fragment encompassing amino acids 244-303 is used (c). Overexpression of GST alone during infection (d) does not result in significant incorporation of protein into the coat material, but fusion of scaffolding protein amino acids 280-303 to the C-terminus of GST results in incorporation of this fusion protein into large MW structures of coat (e). A similar fusion containing scaffolding protein amino acids 284-303 does not sustain the ability to bind to coat and be incorporated into coat structures (f).
Table 3.2
Summary of truncation and fusion mutants tested for complementation and coat binding in vivo

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Primers(^{a})</th>
<th>gp8 fragment</th>
<th>Binding</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSCAFF21-303</td>
<td>AG, X(^b)</td>
<td>21-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF52-303</td>
<td>AJ, X(^b)</td>
<td>52-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF141-303</td>
<td>W, X(^b)</td>
<td>141-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF189-303</td>
<td>Z, X(^b)</td>
<td>189-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF196-303</td>
<td>AK, X(^b)</td>
<td>196-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF218-303</td>
<td>AL, X(^b)</td>
<td>218-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF220-303</td>
<td>AS, X(^b)</td>
<td>220-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF229-303</td>
<td>AH, X(^b)</td>
<td>229-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pSCAFF238-303</td>
<td>AL, X(^b)</td>
<td>238-303</td>
<td>+ used in NMR analysis</td>
<td></td>
</tr>
<tr>
<td>pSCAFF244-303</td>
<td>AU, X(^b)</td>
<td>244-303</td>
<td>+ shortest frag. expressed that assembles</td>
<td></td>
</tr>
<tr>
<td>pSCAFF257-303</td>
<td>AV, X(^b)</td>
<td>257-303</td>
<td>nd</td>
<td>didn’t express</td>
</tr>
<tr>
<td>pSCAFF266-303</td>
<td>AM, X(^b)</td>
<td>266-303</td>
<td>nd</td>
<td>didn’t express</td>
</tr>
<tr>
<td>pSCAFF141-294</td>
<td>W, CC(^b)</td>
<td>141-294</td>
<td>+</td>
<td>N and C-terminal deletions</td>
</tr>
<tr>
<td>pSCAFF141-292</td>
<td>W, AB(^b)</td>
<td>141-292</td>
<td>-</td>
<td>N and C-terminal deletions</td>
</tr>
<tr>
<td>pGST:244-303</td>
<td>BA, XX(^c)</td>
<td>GST:244-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:247-303</td>
<td>BC, XX(^c)</td>
<td>GST:247-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:250-303</td>
<td>BB, XX(^c)</td>
<td>GST:250-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:257-303</td>
<td>BD, XX(^c)</td>
<td>GST:257-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:270-303</td>
<td>BG, XX(^c)</td>
<td>GST:270-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:275-303</td>
<td>BE, XX(^c)</td>
<td>GST:275-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:277-303</td>
<td>BL, XX(^c)</td>
<td>GST:277-303</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>pGST:280-303</td>
<td>BM, XX(^c)</td>
<td>GST:280-303</td>
<td>+ shortest assembling N-terminal fusion</td>
<td></td>
</tr>
<tr>
<td>pGST:284-303</td>
<td>BJ, XX(^c)</td>
<td>GST:284-303</td>
<td>- longest non-assembling N-term fusion</td>
<td></td>
</tr>
<tr>
<td>pGST:293-303</td>
<td>BK, XX(^c)</td>
<td>GST:293-303</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>141-303:H(_{a})</td>
<td>DC, DD(^d)</td>
<td>141-303:mycH(_{a})</td>
<td>+ 23 aa C-terminal fusion that assembles</td>
<td></td>
</tr>
<tr>
<td>141-303:α</td>
<td>W, α(^e)</td>
<td>141-303:α</td>
<td>+ 74 aa C-terminal fusion that assembles</td>
<td></td>
</tr>
<tr>
<td>pNFPWS</td>
<td>8KSD, XX8R(^f)</td>
<td>1-303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSCAFFA1-20</td>
<td>821A, XX8R(^f)</td>
<td>21-303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSCAFFA1-35</td>
<td>836A, XX8R(^f)</td>
<td>36-303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSCAFFA1-57</td>
<td>858A, XX8R(^f)</td>
<td>58-303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSCAFFA1-63</td>
<td>864A, XX8R(^f)</td>
<td>64-303</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSCAFFA1-140</td>
<td>8141A, XX8R(^f)</td>
<td>141-303</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Primers given in pairs: upstream, downstream
\(^{b}\) PCR product cloned into and protein expressed from pET3a
\(^{c}\) PCR product cloned into and protein expressed from pGEX4T-1
\(^{d}\) PCR product cloned into and protein expressed from pTrcHis2B.
See Appendix A for details on construction.
\(^{e}\) PCR product cloned into and protein expressed from pUHE1
\(^{f}\) PCR product cloned into and protein expressed from pNFPW
Fusion to GST did allow stable, high level expression of hybrid proteins containing short C-terminal fragments, and fusions missing 243, 246, 249, 256, 269, 274, and 279 N-terminal amino acids were able to bind coat protein. Again, the presence of procapsid-like structures was confirmed by agarose gel electrophoresis (data not shown) and negative stain electron microscopy (see Figure 3.2 for example). The association of the GST fusion to the C-terminal 34 AA of scaffolding (GST:SCAFF280-303) with coat protein is demonstrated in Figure 3.1, panel e. A similar fusion containing the scaffolding protein residues 284-303 (GST:SCAFF284-303) did not associate with coat protein (see Figure 3.1, panel f) although it was made in similar amounts. Therefore, the N-terminal boundary of the coat protein binding domain resides between amino acids (AA) 280 and 284.

To determine the C-terminal boundary of the coat binding site, we performed a similar analysis with scaffolding protein fragments that begin at AA 141 and terminate at different sites. We find that SCAFF141-294 does assemble (see Figure 3.3) but SCAFF141-292 does not (data not shown). Therefore the C-terminal boundary of the scaffolding protein’s coat protein binding domain lies between AA 292 and 294.

**Scaffolding protein residues 229 to 238 affect portal protein recruitment to procapsids**

Western blot analysis was used to determine the presence or absence of portal protein in preparations of purified procapsids that have incorporated
Figure 3.2 Electron micrograph of negatively stained procapsid-like particles. Particles assembled *in vivo* are filled with GST:257-303 (60,000X magnification).
Figure 3.3 Sucrose sedimentation analysis of SCAFF141-294. A Coomassie-blue stained, 12% SDS-polyacrylamide gel was used to determine protein composition of samples. Lane order: MW standards (StdS), crude lysate (CL) from infected cells expressing scaffolding protein truncation mutant 141-294 during an $\delta^-$ infection, material loaded (L) and fractions 1-12 encompassing the entire gradient. Co-sedimentation of coat and mutant scaffolding protein indicates that the high MW products contain both proteins.
truncated scaffolding protein (see Figure 3.4). Procapsids containing scaffolding protein fragments were prepared as described above and peak fractions from sucrose gradients were concentrated by pelleting. To demonstrate that portal protein was expressed in each infection, whole cell lysates of infected cells were prepared before purification of procapsids. The protein content of whole cell lysates and sucrose gradient purified procapsids were resolved by SDS-PAGE. Lanes from infections producing wild-type procapsids, portal deficient procapsid-like particles, and scaffolding-deficient coat assemblies were included as controls. An identical gel was run, transferred to PVDF and probed with anti-portal protein polyclonal antibody. Using this approach, portal protein was found associated with procapsid-like particles containing scaffolding protein fragments consisting of amino acids 141-303, 189-303, and 218-303 (data not shown). Procapsids assembled by a GST:SCAFF275-303 fusion protein did not contain portal protein (data not shown). The N-terminal limit of portal recruiting functional domain of scaffolding protein is demonstrated by the gel and blot in Figure 3.4. SCAFF229-303 procapsids contain portal protein, but SCAFF238-303 procapsids do not. Therefore, at least some amino acids between 229 and 238 are required for portal protein recruitment. The downstream boundary for portal protein recruitment was not determined.

**Scaffolding protein residues 21-303 are sufficient to complement a scaffolding amber mutation, residues 36-303 are not**

Since the extreme C-terminus of scaffolding protein binds to coat, and the bulk of the N-terminus is apparently dispensable for coat protein binding,
Figure 3.4 Incorporation of the portal protein into procapsids. A. Whole cell lysates (WCL) and sucrose gradient purified particles (SGP) from P22 infected *Salmonella typhimurium* LT2 are resolved on 12% SDS-PAGE. Truncation mutants of scaffolding protein were supplied *in trans* during an 8' infection from an pET3a derived expression plasmids as described in Table 3.2. B. Identical gel to that in part A, transferred to PVDF, probed with anti-portal and coat protein polyclonal antibody.
truncation mutants were tested for their ability to complement, \textit{in trans}, a scaffolding deficient mutant phage for full infectivity. An \textit{8amN123} mutant strain of P22 was plated on a host constitutively expressing a plasmid borne, full-length scaffolding protein or scaffolding protein truncation mutants. These plasmids are listed in Table 3.2. As shown in Figure 3.5, full-length scaffolding protein was able to complement the defective phage, and plaques were seen at normal efficiency. This effect was not the formation of a wild-type 8 gene on the phage chromosome through recombination between the wild-type allele on the plasmid and the phage genome since virions assembled in this host are not infectious on a non-suppressing host (data not shown). Expression of scaffolding protein residues 21-303 from pSCAFFΔ1-20 also sustained infection, although the plaques are smaller than those produced in the presence of the full-length protein. Plaques formed by complementation with pSCAFFΔ1-20 are also not due to recombination repair of the phage to 8\textsuperscript{+} since the amber mutation in the phage lies in codon seven of the scaffolding protein gene and recombinational repair of the P22 amber mutation is impossible. In contrast to the full-length protein and the scaffolding mutant missing the first 20 amino acids, removal of coding sequence for the N-terminal 35 amino acids in pSCAFFΔ1-35 resulted in loss of plaque formation (data not shown).

**Scaffolding protein residues 57-303 are sufficient to sustain capsid assembly and DNA packaging, but virions are not infectious.**

In order to determine the nature of the lack of complementation by the N-terminal truncation mutants of scaffolding protein, similar complementation
Figure 3.5 Minimal sequence required to support infection. Serial dilutions of P22 (8amN123, c1-7) containing the amounts of phage indicated are spotted on a host lawn of the indicated genotype. In the nonsuppressing hosts, scaffolding protein (the product of gene 8) is supplied in trans from an expression plasmid. No plaques are seen on: sup alone; sup\textsuperscript{o}, pSCAFF\textDelta1-20 without IPTG; sup\textsuperscript{5}, pSCAFF\textDelta1-35; sup\textsuperscript{o}, pSCAFF\textDelta1-57; sup\textsuperscript{o}, pSCAFF\textDelta1-63; and sup\textsuperscript{o}, pSCAFF\textDelta1-140.
experiments were performed in liquid and the products of assembly analyzed by sucrose density gradient centrifugation. Briefly, host strains carrying plasmids expressing N-terminal truncation mutants of scaffolding protein (listed in Table 3.2) were infected with an $\delta^-$ strain also defective in lysis (13), allowing the products of assembly to accumulate intracellularly. Sucrose gradients were run on lysates prepared from these infections to determine if scaffolding truncation mutants were capable of forming DNA filled, virion-sized particles or if they were blocked at the stage of scaffolding exit/DNA packaging, which would be revealed as an accumulation of procapsids during the infection.

The material for the gradients was prepared in essentially the same manner as that for the minimal coat protein binding experiments described previously. The infecting strain used was P22 8amN123, 13amH101, c1-7. Procapsids and DNA filled virions migrate at different positions within a sucrose gradient (see Figure 3.6, panel a). During a lysis defective, but otherwise wild-type infection, most of the material made is virion particles (Figure 3.6, panel a). In the absence of scaffolding protein, no virion sized material is made (Figure 3.6, panel b) and most material sediments slightly above the procapsid position, where empty procapsids would normally run. Below this band, the gradients have a whitish opalescence due to a broad distribution of heterogeneous coat protein aggregates. Additionally, some minor products of assembly were seen in the gradients. The scaffolding deficient infection yielded material that ran as a discreet band near the top of the gradient. This material is likely off pathway $T=4$ shells that coat protein can form (Thuman-Commike et al., 1998). An additional band of unkown origin is seen running slightly above where procapsids and empty shells run.
Figure 3.6. Sucrose gradient analysis of assembly products produced during infections with scaffolding protein supplied in trans. Infection with a lysis defective (13\textsuperscript{−}) strain results in the production of infectious virions and trace amounts of procapsids (a; see Table 3.3 for infectivity data). A lysis defective, scaffolding deficient (13\textsuperscript{−}, 8\textsuperscript{−}) strain assembles coat aggregates and closed empty shells that migrate similarly to procapsids (b). Infection with the same strain as in b, but with the full-length scaffolding protein over-expressed from a plasmid results in the production of infectious virions (c). A scaffolding protein truncation mutant with the 35 N-terminal amino acids deleted (SCAFF\textDelta1-35) is able to assemble virions, but with reduced infectivity (d). A scaffolding mutant with an N-terminal 57 amino acid deletion (SCAFF\textDelta1-57) is still able to assemble virions but this material is not infectious (e).
When full-length scaffolding protein is supplied \textit{in trans} from the “SCAFF\(\Delta X\)” plasmids during the scaffolding protein deficient infection, the assembly of virions is restored (Figure 3.6, panel c). Virion-sized material is also produced in the presence of gpSCAFF\(\Delta 1\)-35 (Figure 3.6, panel d) and gpSCAFF\(\Delta 1\)-57 (Figure 3.6, panel e). The trans-complemented infected cell lysates were also sedimented to equilibrium in a CsCl step gradient consisting of \(\rho=1.4 \text{ gm/cc} \) and \(\rho=1.6 \text{ gm/cc} \) density steps. The material banded between these layers at approximately \(\rho=1.5 \text{ gm/cc} \), the banding density of DNA-filled P22. This material was then resolved in a 10-40\% sucrose velocity gradient. The slower sedimenting procapsid bands were not present (as expected), but the bands at the virion position were present. These findings indicated that these particles contained DNA. Interestingly, the band running slightly slower than the empty shell position was not removed by the CsCl step, suggesting that these might be mini-virions with \(T=4\) coat protein shells.

Although SCAFF\(\Delta 1\)-35 failed to support plaque formation in the complementation assay, it was capable of forming virion-sized material (as determined by sucrose gradient centrifugation) that has virion density in CsCl gradients. The particles present in the gradients were apparently virion-like particles assembled during an infection in liquid culture. However, complementation by SCAFF\(\Delta 1\)-35 did not yield plaque forming units on solid medium. In order to understand this difference, the bands at the virion position in each of the sucrose gradients were harvested and the amount of infectious phage was determined by serial dilution and plating on a sup\(^*\) host. The titers of phage purified from these infections are given in Table 3.3. All solutions used contained buffer TM (10 mM Tris, pH 7.5; 1 mM MgCl\(_2\)) to stabilize DNA filled
Table 3.3

Infectivity of purified virions assembled from scaffolding protein truncation mutants

<table>
<thead>
<tr>
<th>P22 strain and scaffolding plasmid</th>
<th>Plaque forming units (pfu)/mL ( ^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 13^- )</td>
<td>( 1.28 \times 10^{11} )</td>
</tr>
<tr>
<td>( 13^-, 8^- )</td>
<td>( 2.45 \times 10^9 )</td>
</tr>
<tr>
<td>( 13^-, 8^-, pNFPW8 )</td>
<td>( 6.4 \times 10^{10} )</td>
</tr>
<tr>
<td>( 13^-, 8^-, pSCAFFΔ1-35 )</td>
<td>( 4.1 \times 10^8 )</td>
</tr>
<tr>
<td>( 13^-, 8^-, pSCAFFΔ1-57 )</td>
<td>( 1.06 \times 10^6 )</td>
</tr>
</tbody>
</table>

\( ^a \) 150 µg of material isolated using a mechanical gradient fractionator (BioComp Instruments, New Brunswick, Canada) from each infection was loaded onto a 10-40% sucrose gradient and centrifugated in a Beckman SW41 rotor at 25 K, 8°C, for 1h 45'. Equal volumes of sucrose gradient were harvested from identical postion within each gradient and titered on suppressing host.
particles. Precise harvesting of equal volumes at identical positions among the gradients was achieved by use of a Gradient Fractionator (Biocomp Instruments, New Brunswick, Canada).

Although no infectious virions are expected to be made during an $\delta^-$ infection, and no virion band was visible in the gradient at the virion position (see Figure 3.6, panel b), material harvested from the position in the gradient where virions were expected to run contained $2.45 \times 10^6$ plaque forming units. These are likely the phage used in the infection, which was done at an MOI of 7, that carried through the purification process and/or revertants since the $\delta^-$ defect is a single amber nonsense mutation.

The titers demonstrate that trans-complementation with the full-length scaffolding protein restored the titers to about half of the wild-type infection. Virion-like particles made by SCAFFΔ1-35 complementation (“transvirions”) are “weakly infectious” since its titer is 100-fold less than particles made by full-length trans-complementation, but also 100-fold greater than background. I conclude that SCAFFΔ1-57 transvirions, although they contain DNA, are not infectious on the sup$^+$ host.

SCAFFΔ1-57 transvirions have altered protein and DNA composition

In order to determine why the SCAFFΔ1-57 transvirions are non-infectious, the protein components in purified wild-type P22 and SCAFFΔ1-57 DNA-containing particles were resolved by SDS-PAGE and are shown in Figure 3.7. Electron micrographs of these particles are shown in Figure 3.8. The gel reveals two differences in protein composition between the two samples.
Figure 3.7 Protein composition of purified virions. 12% SDS-PAGE followed by staining with Coomassie blue was used to resolve and visualize the protein composition of virion-like particles purified from a P22 8amN123 infection complemented in trans by plasmid based expression of either the full-length scaffolding protein, expressed from pNFPW8, or the N-terminal truncation mutant, SCAFF\textsubscript{Δ1-57}. Particles from both infections sedimented through CsCl at \( \rho = 1.4 \) gm/cc and banded atop a CsCl layer between \( \rho = 1.6 \) gm/cc. The band of protein in the Δ1-57 lane running at approximately 34 kDa, not seen in the 1-303 lane, was eluted and microsequenced, revealing it to be the truncated scaffolding protein.
Figure 3.8 Electron microscopy of virions. A) wild-type P22 virions B) SCAFFD1-57 transvirions. Magnification in both panels is approximately 60,000X.
The SCAFFΔ1-57 transvirion material has a greatly reduced amount of gp16 relative to wild-type and also has an additional major protein band migrating at approximately 34 kDa. In order to determine the identity of the extra protein, the proteins were electroblotted to PVDF for protein microsequencing. Twelve out of 12 residues identified were identical to scaffolding protein residues 58 through 69, showing that this band is SCAFFΔ1-57.

The presence of the truncated scaffolding protein in the SCAFFΔ1-57 transvirions indicates that it was still complexed with the DNA-containing virion. Electron micrographs of SCAFFΔ1-57 transvirions were indistinguishable from wild-type virions (see Figure 3.8), suggesting that the truncated scaffolding protein was inside the capsid. To test for altered DNA composition, DNA was extracted from SCAFFΔ1-57 and wild-type virions and subjected to contour-clamped homogenous electric field (CHEF) agarose gel electrophoresis (Figure 3.9). Whole lambda DNA and lambda DNA digested with HindIII were included in the gel as size standards. The DNA from wild-type P22 virions resolved as a slightly more diffuse band than the whole lambda DNA. Normal P22 DNA is heterogeneous in length (±850 bp) due to the headful mechanism of DNA packaging (Casjens & Hayden, 1988). The SCAFFΔ1-57 resolved as a more diffuse band relative to wild-type P22 DNA indicating a much greater heterogeneity in DNA length. Calibration of the gel with the known size standards and densitometric analysis of the gel (data not shown) indicated that the center of the SCAFFΔ1-57 transvirion DNA band migrated at a distance corresponding to 38.5 kb with a range of ±4 kb.
**Figure 3.9.** Wild-type and transvirion phage DNA. DNA from wild-type P22 and the D57 transvirion was resolved by CHEF gel electrophoresis and compared to DNAs of known length.
Discussion

The coat protein binding functional domain of scaffolding protein

By analyzing truncation and truncation/fusion mutants of scaffolding protein, we identified a small region of the scaffolding protein that is sufficient to mediate coat protein binding. Selected scaffolding protein mutants are schematically illustrated in Figure 3.10. This minimal coat protein binding domain, functionally defined by assembly \textit{in vivo}, is only 15 AA long and spans residues 280 through 294. This region has a high proportion of charged amino acids and is predicted to be composed of \(\alpha\)-helices. Using this information and our expression plasmid pSCAFF238-303, Sun \textit{et al.} (2000) were able to determine an NMR solution structure for our truncated scaffolding protein consisting of residues 238-303. This structure shows that the coat protein binding functional domain of the scaffolding protein is contained within, but does not encompass all of a structural domain consisting of a helix-loop-helix motif. The N-terminal, or proximal helix is formed by 15 AA (269-283) followed by a loop of 5 AA (284-288). The C-terminal, or distal helix is composed of the 15 AA residues 289-303. It is interesting to note that the minimal coat-binding functional domain spans the loop and only the adjacent portion of each helix, and that portions of this structural domain are dispensable for binding function. It is not known how the truncations affect the folding and structure of the helix-loop-helix motif. Additionally, contributions that sequences from GST might make to either the folding of the coat protein binding domain and or coat binding have not been determined.

The longest N-terminal truncation of scaffolding protein that still binds coat protein removes 11 of the 15 amino acids of the proximal helix. Similarly,
<table>
<thead>
<tr>
<th>scaffolding mutant:</th>
<th>coat binding?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-303</td>
<td>Yes</td>
</tr>
<tr>
<td>244-303</td>
<td>Yes</td>
</tr>
<tr>
<td>GST:280-303</td>
<td>Yes</td>
</tr>
<tr>
<td>GST:284-303</td>
<td>No</td>
</tr>
<tr>
<td>141-303</td>
<td>Yes</td>
</tr>
<tr>
<td>141-294</td>
<td>Yes</td>
</tr>
<tr>
<td>141-292</td>
<td>No</td>
</tr>
<tr>
<td>141-303:myc/H₆</td>
<td>Yes</td>
</tr>
<tr>
<td>141-303:α-peptide</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 3.10 Schematic summary of truncation and fusion mutants tested for coat protein binding.
the longest C-terminal truncation of scaffolding protein with a functional coat-binding domain removes 9 of the 15 amino acids of the distal helix. These data suggest the possibility that the complete folded helix-loop-helix motif might not be necessary for binding function and further suggests the possibility that the conformation of scaffolding protein bound to coat might be different from scaffolding protein in solution. Structural differences between free scaffolding and procapsid-bound scaffolding have been observed by Raman spectroscopy (Tuma et al., 1996a) and inferred by differential scanning calorimetry (Galisteo & King, 1993). Comparisons between a synthetic Raman spectrum consisting of the summed spectra for free scaffolding and empty procapsid shell versus scaffolding filled procapsids suggest an increased \( \alpha \)-helical content for scaffolding protein within procapsids (Tuma et al., 1998b). Free scaffolding in solution displays a broad noncooperative denaturation profile by differential scanning calorimetry (DSC) but procapsid bound scaffolding is thermostabilized. Raman spectra demonstrate exit of scaffolding protein from the procapsid is strongly correlated with both main chain changes in the scaffolding protein, which may be indicative of an unfolding transition, and a reorientation of a small number of side chains that could reflect changes in interactions between the procapsid lattice and scaffolding protein. Therefore, the folded state of free scaffolding is different from scaffolding bound to the procapsid. However, there is no evidence to suggest that a reordering or change in folding occurs in the coat protein binding domain.

Scaffolding binds to coat protein via it’s C-terminus, in a manner highly analogous to the Herpesviridae and Microviridae scaffolding proteins. Binding of the herpes simplex virus 1 (HSV-1) scaffolding protein, preVP22a
(alternatively called ICP35) to the B-capsid lattice (the Herpes equivalent of a procapsid) is mediated by 12 amino acids very near the C-terminus of preVP22a (Hong et al., 1996). However, unlike the P22 scaffolding protein, whose binding has been proposed to be electrostatic, the interaction between the HSV-1 scaffolding protein and the coat protein is likely hydrophobic (Hong et al., 1996). Mutational analysis of the HSV-1 minimal coat binding domain shows that a phenylalanine residue is absolutely essential for binding and that hydrophobic residues clustered around this phenylalanine are important as well (Beaudet-Miller et al., 1996). Furthermore, the binding interaction is abrogated by the presence of Nonidet P-40, further evidence supporting a hydrophobic interaction (Hong et al., 1996). The representative Microviridae member, øX174, a ssDNA bacteriophage, also utilizes a scaffolding protein, gpB, in the assembly of its procapsid, and like herpes preVP22a and P22 gp8, binds to the capsid via its C-terminus (Dokland et al., 1999). Contact between gpB and the capsid gpF occurs through a small amphipathic helix that fits into a pocket on gpF (Dokland et al., 1997). Interestingly, these three apparently evolutionarily unrelated scaffolding proteins all bind their partners through their C-termini. It has been proposed that encoding a coat binding domain at the C-terminus of a scaffolding protein might be a way to guard against poisoning of assembly by prematurely stopped translation products (Eppler et al., 1991).

The portal protein recruiting domain of scaffolding protein

The N-terminal boundary of the domain necessary for recruiting portal to the procapsid complex resides between residues 229 and 238. These residues likely represent the region of the scaffolding protein that either directly binds to
the portal protein or completes a domain that binds to portal. The possibility that
another protein serves as a bridge between scaffolding protein and portal protein
has not been formally ruled out, but is unlikely since coexpression of coat,
scaffolding, and portal proteins from a plasmid is sufficient to generate
procapsid-like particles that contain portal protein (P. Weigele & S. Casjens,
unpublished observation) and there is no evidence for participation of any host
proteins. There are data suggesting gp16 might be involved in portal protein
recruitment. The ts scaffolding protein conditional mutant, S242W, is deficient in
portal incorporation under nonpermissive conditions and also fails to
incorporate gp16 into the procapsid (Greene & King, 1996). However, at
restrictive temperatures, the ts mutant L1771 assembles procapsids that contain
portal protein but lack gp16 (Greene & King, 1996). Similarly, transvirions
assembled using the SCAFFΔ1-57 lack gp16. However, 16− infections yield
similar amounts of virion-like particles as wild-type infections, suggesting that
gp16’s role in portal recruitment is not essential.

Surprisingly the longest truncation of scaffolding that recruited portal
protein to the procapsid was missing the region (Y214W) where a scaffolding-
portal interaction had previously been mapped genetically (Greene & King,
1996). How are truncations of scaffolding that exclude this region able to recruit
portal? One possibility is that Y214 is within a region of scaffolding that
regulates the binding of scaffolding to portal, perhaps through allosteric
interactions between scaffolding protein domains, either within a scaffolding
protein subunit or between scaffolding molecules.

These data constitute further evidence supporting a direct physical
interaction between the scaffolding and portal proteins. Portal incorporation can
occur in the absence of gp16 and all other phage proteins except, of course, the scaffolding and coat proteins. Portal has been shown, in vivo, not to be a kinetic initiator of assembly for both P22 (Bazinet & King, 1988) and the phage SPP1 (Droge et al., 2000). So the question of how only one portal protein complex is incorporated in the procapsid remains unanswered.

**Scaffolding protein release and DNA packaging**

*S. typhimurium* that carries a plasmid that constitutively expresses the truncation mutant SCAFFΔ1-20 scaffolding protein can trans-complement a scaffolding deficient phage infection as seen by plaque formation on this strain; however, SCAFFΔ1-35 and SCAFFΔ1-57 cannot efficiently trans-complement to form plaques. Analysis of particles assembled during trans-complemented infections performed in liquid demonstrate that truncation induced loss of plaque formation is not due to a failure of DNA packaging since the SCAFFΔ1-35 and SCAFFΔ1-57 mutants can mediate the assembly of virus particles that contain DNA. The yield of infectious virions produced by trans-complementation with Δ1-36 was three orders of magnitude less than the 1-303 trans-complemented infection suggesting that the burst size during Δ1-35 trans-complemented infection on a host lawn was probably too small to make visible plaques. The SCAFFΔ1-57 transvirions were completely non-infectious.

The SCAFFΔ1-57 transvirions have multiple defects. They do not contain the full complement of gp16, a protein required for virion infectivity. Surprisingly, the truncated scaffolding protein is retained in the virion, demonstrating that, for this mutant, the initiation of DNA packaging and scaffolding protein release are uncoupled. No differences are apparent between
the SCAFFΔ1-57 transvirions and wild-type by negative stain electron microscopy, suggesting that the retained scaffolding fragments are inside the capsid. The average length of DNA packaged in the SCAFFΔ1-57 transvirion is 5 kb shorter than for wild-type virions. The wild-type P22 virion packages a linear 43.5 kb DNA, or ~800 bp more than the 41,725 bp genome (Casjens & Hayden, 1988). The redundant termini of P22 DNA are substrates for the host homologous recombination machinery and are required for circularization of P22 DNA after injection. P22 normally packages DNA according to capsid volume, since P22 strains containing chromosomal deletions package the same length DNA as wild-type (Tye et al., 1974) and overexpression of portal protein during infection results in the off pathway assembly of smaller T=4 capsids that package a shorter length DNA (Moore & Prevelige, 2002). In the case of the SCAFFΔ1-57 transvirion, the unreleased scaffolding protein almost certainly reduces the volume in the capsid available to packaged DNA resulting in the packaging of shorter length chromosomes. These chromosomes, being shorter than wild-type, do not contain the terminal redundancy that is required to support homologous recombination mediated DNA circularization.

The first 57 amino acids of scaffolding protein may be important for sensing the entry of DNA into the procapsid and altering the conformation of scaffolding protein thus allowing it to exit the procapsid before expansion and closure of the holes in the coat protein shell. But how is the SCAFFΔ1-57 scaffolding protein “stuck” inside the capsid? One possibility is that the truncation prevents the scaffolding protein from releasing from its coat protein binding partner. This seems somewhat unlikely since the binding sites for scaffolding protein inside the capsid are probably lost during the shell expansion
that marks the maturation transition between procapsid and capsid. The
topography of the virion’s inner surface as revealed by cryo-electron microscopy
is much smoother than that of the procapsid. Four of six finger-like coat protein
densities lining the interiors of holes observed at each of the quasi-six-fold axes
have been identified as potential scaffolding protein binding sites. In the virion,
these fingerlike densities are displaced towards the center of the hole, apparently
closing the hole and altering the accessibility of the possible scaffolding binding
sites. A second possibility is that truncation of scaffolding protein’s first 57 amino
acids might alter the conformation of the scaffolding protein such that it is
sterically preventing exit from the procapsid. A third possibility is that the
scaffolding protein is held in the procapsid by scaffolding-scaffolding protein
interactions. If SCAFFΔ1-57 is indeed trapped in the virion by scaffolding-
scaffolding interactions, one specific prediction would be that co-expression of
SCAFFΔ1-57 with the wild-type scaffolding might trap the otherwise exit
competent wild-type protein in the virion as well. A fourth possibility is that the
deletion mutant simply fails to respond to the presence of DNA and is simply
trapped in the capsid by hole closure during expansion.

One general consideration in trying to understand scaffolding protein exit
from the procapsid is the question of which protein actively releases the other
during maturation. Previously published data are consistent with protein exit
being due to conformational changes in the scaffolding protein and not the coat
protein. The binding of scaffolding protein to the procapsid is perturbed by low
concentrations of denaturant, 0.5M GuHCl being sufficient to extract most of the
scaffolding protein from the procapsid. Scaffolding protein extraction is
reversible; removal of denaturant by dialysis or dilution is sufficient to allow re-
entry into and binding of the procapsid (Fuller & King, 1981). Temperature induced exit of scaffolding protein is a cooperative process and is correlated with structural changes in the scaffolding protein and not the coat protein (Tuma et al., 1996a). Furthermore, the conformation of scaffolding protein within the procapsid is more resistant to thermal denaturation than free scaffolding (Tuma et al., 1996a). These observations suggest that the conformation of scaffolding protein in solution is different to that of scaffolding protein bound to the procapsid. Comparisons of cryo-electron micrographic reconstructions of scaffolding protein filled procapsids versus empty coat protein shells show only slight differences in the conformations of the coat proteins. If control and cooperativity of scaffolding protein release were mediated by coat protein, one might expect large changes in coat protein structure associated with scaffolding protein exit. It would be interesting to see if biophysical characterization of the SCAFFΔ1-57 scaffolding protein mutant might reveal differences in conformation or stability that can be correlated with its failure to exit during DNA packaging.

Summary and Model: Allosteric control of scaffolding protein function

The data presented in this chapter demonstrate that the coat protein binding, portal recruitment, and scaffold exit functions can be localized to specific amino sequences of the scaffolding protein. It is interesting to note that binding of scaffolding protein to coat protein is mediated by amino acids very close to the C-terminus, whereas the control of scaffolding protein exit is mediated by amino acids near the N-terminus. This suggests that conformational changes in the scaffolding protein which control its function might be relayed
between functional domains along the length of the molecule, or that the N- and C-terminal domains contact each other directly. Previously characterized scaffolding protein mutants provide tentative support for a model of allosteric control of scaffolding protein function.

*Scaffolding protein L177I in light of SCAFFΔ1-57.* L177I and SCAFFΔ1-57 mutants of scaffolding protein have paradoxical phenotypes since both are apparently unable to exit the procapsid, but SCAFFΔ1-57 packages DNA and L177I does not. The temperature sensitive mutant L177I fails to exit the procapsid at nonpermissive temperatures (Greene & King, 1996). Unlike SCAFFΔ1-57, L177I mutants do not package DNA and instead accumulate at the procapsid stage. *In vitro,* the L177I scaffolding protein was shown to have a higher affinity for the procapsid, since it was more resistant to GuHCl extraction (Greene & King, 1996). Procapsids containing the L177I mutant protein have a higher energy of activation for scaffolding protein release plus expansion. Cryo-electron micrographic reconstruction of L177I procapsids suggest that the mutant protein binds to all six of the finger-like densities in the hexon holes rather than just four as the wild-type scaffolding protein does (Thuman-Commike et al., 2000). If the increased affinity for the capsid lattice is the source of the temperature sensitive phenotype, the L177I data suggest that scaffolding protein exit is a necessary precondition to expansion, since this mutant accumulates unexpanded, unpackaged capsids.

How are procapsids containing SCAFFΔ1-57 able to expand and package DNA when L177I cannot? One possibility is that SCAFFΔ1-57 might have a normal affinity for coat protein. SCAFFΔ1-57 might be able to receive the exit signal, release from coat protein, but is retained in the procapsid by an unknown
mechanism. In the case of L177I, the phenotype might be due to gain of function; binding to the additional sites within the procapsid might prevent the movement of the finger-like densities into the hexon holes thereby sterically blocking shell expansion. How the higher affinity of L177I protein for the procapsid, or its ability to bind additional sites within the procapsid avoids interference with shell expansion and DNA packaging at permissive temperatures is unknown. L177I is not dominant since the mutation was mapped by plasmid rescue using a wild-type copy of gene 8. Despite these differences, the behavior of both mutants suggests that sequences N-terminal to the coat protein binding domain can affect its function. For L177I, sequences in the region of this mutation might modulate the affinity of scaffolding protein for the coat protein. For SCAFFΔ1-57, the truncation may directly remove controlling sequences or affect the conformation of the protein C-terminal to the truncation. In either case, the affected sequences of SCAFFΔ1-57 appear to control scaffolding-scaffolding interactions or the exit conformation of the scaffolding protein.

Other scaffolding truncation mutants. A scaffolding truncation mutant consisting of amino acids 141-303 has previously been characterized and apparently has a higher affinity for coat protein. The scaffolding protein fragment, SCAFF141-303, stimulates coat protein to assemble with higher kinetics in vitro than does wild-type scaffolding protein (Parker et al., 1998). Compared to wild-type scaffolding, SCAFF141-303 binds with higher stoichiometry within the procapsid and is more difficult to remove by GuHCl extraction (P. Weigele & S. Casjens, unpublished observations). In a shell binding assay, SCAFF141-303 is better able to fill “empty procapsids” than wild-type (see Appendix B). Biophysical characterization of the 141-303 scaffolding protein
revealed a destabilized domain N-terminal to the coat protein binding domain (Tuma et al., 1998b). Interestingly, this region was found by circular dichroism and tryptophan fluorescence to be similarly destabilized in the L177I mutant (Greene & King, 1999a). These two mutants suggest that removal or destabilization of regions N-terminal to the coat protein binding domain can affect its affinity for coat protein.

The increased coat affinity by the truncated mutants might also be explained by decreased steric hindrance. A truncated protein, being smaller, might have greater access to binding sites within the procapsid and also diffuse more quickly during assembly. Indeed, a scaffolding truncation mutant composed of residues 238-303, SCAFF238-303 promoted polymerization of unassembled coat protein subunits with even faster kinetics in vitro than SCAFF141-303 (P. Prevelige, personal communication). It seems unlikely that this additional truncation of the scaffolding protein would further increase its affinity for coat protein, rather the truncation might increase the rate of assembly by increasing the rate of diffusion and rotation of the scaffolding protein’s coat protein binding domain. Increasing diffusion through solvent might increase scaffolding-coat protein collisional frequency, and a shorter scaffolding molecule might "tumble" more quickly and could therefore more rapidly find a suitable docking orientation to bind to coat protein.

Truncation of the scaffolding protein also leads to a reduction in the fidelity of assembly. Procapsid assembly reactions performed in vitro with either SCAFF141-303 or the SCAFF238-303 yield procapsid-like particles as well as a large proportion of malformed shells and coat protein aggregates (P. Prevelige, personal communication). This lowered fidelity may be due either to loss of an
N-terminal domain required to mediate scaffolding-scaffolding contacts that are required for accuracy of assembly, or, since these fragments assemble with faster kinetics, an N-terminal domain of scaffolding protein might ensure the fidelity of assembly by lowering the rate of assembly, thereby allowing coat proteins to find the best conformations.

The fact that truncation of the scaffolding protein causes reduced fidelity of assembly in vitro could be the result of a lower rate of self-association by these mutant proteins. Free scaffolding protein in solution is in equilibrium between monomers, dimers, and tetramers (Parker et al., 1997a). The 141-303 scaffolding protein does not form tetramers and the Kd for dimer to monomer dissociation is 10 fold higher (Tuma et al., 1998b). It had been observed that a covalently linked scaffolding protein dimer was more active in assembly than unmodified scaffolding protein (Parker et al., 1998) and that the rate of assembly in vitro displayed a second order dependence on scaffolding protein concentration (Prevelige et al., 1988). This led to the suggestion that scaffolding dimerization acted as an “entropy sink” with the favorable energy of scaffolding protein dimerization being used to tether coat protein subunits together thereby promoting productive binding interactions (Parker & Prevelige, 1998). However, by NMR spectroscopy, the 238-303 scaffolding protein is entirely monomeric (Parker et al., 1997b) and yet it promotes rapid polymerization of coat protein in vitro, contradicting the previous suggestion that dimerization of scaffolding protein is important for initiation of assembly. These data suggest that scaffolding protein’s ability to activate coat protein assembly is due to binding to coat protein alone and that scaffolding protein dimerization may be only needed for accurate shell assembly.
CHAPTER 4

MUTATIONAL ANALYSIS OF THE COAT PROTEIN-BINDING DOMAIN OF THE P22 SCAFFOLDING PROTEIN

Introduction

The central process in virus morphogenesis is the assembly of a protein shell or capsid that protects the viral chromosome when it is between hosts. In the case of the dsDNA bacteriophages, and some eukaryotic icosahedral viruses, the polymerization of coat protein subunits during capsid assembly is controlled by scaffolding proteins (Dokland, 1999). Scaffolding proteins stimulate and regulate coat protein polymerization by binding coat protein subunits and lowering the concentration at which they assemble (Prevelige et al., 1988). Furthermore, scaffolding proteins ensure the fidelity of assembly by controlling the size and curvature of the shell and by assisting in the incorporation of the portal and other minor nonstructural proteins into assembling capsids. These functions are necessary for the correct assembly of the procapsid, a viral capsid intermediate composed, minimally, of icosahedrally arranged coat protein subunits, an inner core of scaffolding protein, and a portal protein at a single vertex.

The procapsid acts as a preformed container from which the scaffolding protein must exit and into which the viral chromosome is packaged. DNA
packaging is accompanied by conformational changes in the coat protein subunits that result in an expanded and physically more robust shell. As such, the procapsid is metastable: 1) physically, in that it is more susceptible to dissociation by denaturants or pressure as compared to the virion, 2) conformationally, since the coat subunits are not in their terminal conformation until expansion is complete, and 3) temporally, as a transient structural intermediate whose composition changes en route to the mature virion.

Its multiple roles of binding, recruitment, and release, indicate that scaffolding proteins do not function merely as a passive switch for coat protein polymerization during procapsid assembly. Regions of primary amino acid sequence necessary to these roles have previously been defined, in vivo, for the bacteriophage P22 scaffolding protein (see Chapter 2). Furthermore, a growing body of complementary genetic, biochemical, and structural data on P22, and other viruses that use scaffolding proteins, provides a framework in which to fit a more refined dissection of P22 scaffolding protein functions.

Unassembled scaffolding protein does not have the characteristics of a “normal” globular protein (Fuller & King, 1981). Its thermal denaturation profile is noncooperative, and all amide hydrogens are exchangeable in deuterium suggesting it has no hydrophobic or exchange protected core (Tuma et al., 1996a). Raman spectroscopy and circular dichroism indicate, however, that it is rich in secondary structure, with α-helices predominating (Greene & King, 1999a; Tuma et al., 1998b). Scaffolding protein is also highly sensitive to proteases (Greene & King, 1999b). The hydrodynamic profiles of scaffolding protein during velocity sedimentation and gel chromatography suggest an axial ratio of approximately 10:1 (Parker et al., 1997a). These observations lead to a model of scaffolding as an
elongated molecule composed of mostly of α-helices connected by flexible random coil. Some similar characteristics have been ascribed to the scaffolding proteins of lambda, T7, T4 and herpes virus (Dokland, 1999). Scaffolding protein bound in the procapsid interior, however, displays a cooperative melting transition as observed both by differential scanning calorimetry (DSC) and Raman spectroscopy (Galisteo & King, 1993; Tuma & Thomas, 1997). Furthermore, the cooperativity of exit is correlated with structural transitions in the scaffolding protein but not the coat protein (Tuma & Thomas, 1997). The computed Raman difference spectrum between free and bound scaffolding proteins indicates not only changes in α-helicity but side chain orientations as well, suggestive of protein-protein interactions present within the procapsid that are not present in the free subunits (Tuma & Thomas, 1997). These results imply a conformational control of scaffolding protein binding to and exit from the procapsid, although the relative contributions from scaffolding-scaffolding versus scaffolding-coat interactions to these processes have not been fully resolved.

Scaffolding protein remains stably associated with coat protein inside the procapsid until the time of DNA packaging, whereupon it exits, intact, from the procapsid (Casjens & King, 1974). Cryo-electron microscopic reconstruction of scaffolding protein filled procapsids reveals patches of density not seen in empty coat protein shells (Thuman-Commike et al., 1999). These densities, tentatively attributed to the C-terminal region of scaffolding protein, are localized to four of six finger-like projections lining the inner rim of holes at the centers of coat hexons within the procapsid lattice. Scaffold ing subunits at these locations
would be well poised to exit through those holes during DNA packaging. Scaffolding protein exit can be mimicked, \textit{in vitro}, by treatment of procapsids with mild concentrations of denaturant or heat (Casjens \& King, 1974; Fuller \& King, 1981). Heat treated procapsids not only lose their scaffolding proteins, but undergo expansion as well, an irreversible process resulting in the closure of the hexon holes (Galisteo \& King, 1993). Scaffolding exit, therefore, appears to precede expansion, a requirement supported by a conditional mutant of scaffolding protein L177I, which accumulates procapsids under nonpermissive conditions (Greene \& King, 1996). This mutant has an apparent increased affinity for the procapsid lattice and procapsids containing the L177I scaffolding mutant expand at higher temperatures. These data suggest that \textit{in vivo}, a failure of scaffolding protein exit blocks DNA packaging. However, an N-terminal scaffolding truncation mutant, SCAFFΔ1-57, remains associated within the capsid after DNA packaging and shell expansion have occurred, contradicting the exit requirement for expansion (see Chapter 2). The smoothness of the virion’s inner surface (Zhang \textit{et al.}, 2000) likely indicates that scaffolding protein binding sites are lost during expansion implying that SCAFFΔ1-57 may be trapped within the procapsid by scaffolding-scaffolding interactions. The paradoxical phenotypes of these two mutants suggest a subtle relationship between scaffolding-coat and scaffolding-scaffolding interactions and exit of the scaffolding protein from the procapsid.

The P22 scaffolding protein binds to coat protein by a 15 amino acid functional domain spanning residues 280 to 294 of the 303 AA long protein (see Chapter 3). This coat protein binding domain is within a highly charged helix-loop-helix structural motif whose solution structure has been determined by
NMR (Sun et al., 2000). Capsid binding domains have been characterized for Herpes Simplex Virus (HSV-1) (Beaudet-Miller et al., 1996) and Microviridae scaffolding proteins (Burch & Fane, 2000), which, like the P22 coat binding domain, are located very near the proteins’ C-termini and are small with respect to the entire length of the protein. P22 scaffolding protein’s ability to bind coat protein can be examined by in vitro assembly of purified scaffolding and coat subunits (Prevelige et al., 1988), as well as by monitoring binding of scaffolding protein to preassembled shells of coat protein derived from procapsids assembled in vivo from which scaffolding protein has been removed by mild treatment with denaturant (Greene & King, 1994). Unlike the scaffolding-coat interactions in herpesviruses that are thought to be hydrophobic (Hong et al., 1996), electrostatic interactions have been proposed as the basis for the P22 scaffolding-coat protein interaction since high concentrations of NaCl inhibit both assembly and shell binding in vitro (Parker & Prevelige, 1998). The kinetics and thermodynamics of the scaffolding protein-shell binding interaction have also been examined, and both approaches suggest two classes of binding interactions between scaffolding protein and the procapsid lattice (Greene & King, 1994; Parker et al., 2001). In this study, the solution structure of the scaffolding coat protein binding domain (Protein Data Bank identification code: 1GP8) is used to facilitate a “rational mutagenesis” approach. We examined the effects amino acid substitutions within the coat protein binding domain of the P22 scaffolding protein, as well as truncation of the N-terminus, have on shell binding and assembly in vitro.
Materials and methods

Cloning of mutant scaffolding proteins

Wild-type and mutant scaffolding protein genes were generated by PCR amplification of the scaffolding gene using restriction site and mutation containing primers and cloning these fragments into the expression vector pET15b (Novagen, Madison, WI) which adds histidine codons to the 5’-end of the cloned reading frame. Primers used for PCR are listed in Table 4.1. Briefly, PCR reactions were performed in 50 µl reactions, using the Elongase® enzyme mixture (Invitrogen, Carlsbad, CA), consisting of Taq polymerase and a proofreading enzyme for greater processivity and accuracy, in the supplied buffer with 1.5 mM MgCl₂, 1 µM each primer, and 200 µM each dNTP. The cycling parameters included an initial 2-minute 94°C denaturation step followed by 30 cycles of consisting of a 30-second, 94°C denaturation step; a 30-second, 55°C annealing step; and a 1-minute, 68°C extension step. Enzymes as well as unincorporated nucleotides and primers were removed by purification using a Qiaquick PCR Cleanup Kit (Qiagen, Valencia, CA).

Five µl of purified PCR reaction and approximately 500 ng of pET15b vector DNA, in a total volume of 50 µl, were digested together first with 10 units of NdeI in NEB Buffer 2 (New England Biolabs, Beverly, MA) supplemented with bovine serum albumin (BSA). After 1 hour at 37°C, NaCl was added to a final concentration of 100 mM, and 20 units of BamHI were added. After incubation at 37°C for an additional hour, restriction enzymes were removed either by phenol/chloroform extraction followed by ethanol precipitaton or by using a Wizard® DNA cleanup kit (Promega, Madison, WI). Digested DNA was
Table 4.1

Coat protein binding domain point mutants and the oligonucleotide sequences used to generate them

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Primer Sequence b</th>
</tr>
</thead>
<tbody>
<tr>
<td>D288A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>K286A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>E290A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>G287A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>V289A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>D281A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>R293A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>K294A</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>D281K</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>K286E</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>D288R</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>E290K</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>R293E</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>K294E</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>R293E</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>K294E</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>V289D</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>G287P</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>A284D</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>Y292D</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>T291D</td>
<td>CGGAGGATCC</td>
</tr>
<tr>
<td>8W</td>
<td>GAAGATCCATATG</td>
</tr>
<tr>
<td>8LNde</td>
<td>CAGAGATCCATATGA</td>
</tr>
</tbody>
</table>

*aUpstream primers encode a NdeI restriction site; downstream primers encode a BamHI restriction site.

*bNoncoding tails are italicized, restriction sites are boldfaced and mutations are underlined. Sequences are given 5' to 3'.

*Upstream primers: 8W amplifies from codon 141 of gene 8, 8LNde amplifies from the start codon
incubated overnight at 16°C with five units of T4 DNA ligase in a reaction volume of 20 μl.

Ten μl of the ligation mixture were used to transform CaCl₂ competent *E. coli* strain NF1829 to ampicillin resistance. Plasmids were recovered from transformants by alkaline-lysis mini-prep method and the presence of cloned inserts was determined by visualization of *Nde*I and *Bam*HI and DNA fragments in agarose electrophoresis gels stained with ethidium bromide. The presence of the engineered mutation was confirmed by sequencing (Core Facility, University of Utah). The appropriate plasmids were used to transform CaCl₂ competent BL21(DE3) (Stratagene, La Jolla, CA), an *E. coli* B strain that contains the λDE3 prophage expressing T7 RNA polymerase under control of an IPTG inducible promoter.

**Expression and purification of N-terminally histidine-tagged scaffolding protein**

An overnight culture of BL21(DE3) carrying the appropriate plasmid was diluted 1:100 into 500 ml L-broth (LB) containing ampicillin at 100μg/ml, and cultures were incubated at 37°C with shaking. When cultures reached an OD₅₉₅ of 0.5-0.6, expression of his-tagged scaffolding protein was induced by addition of IPTG to a final concentration of 0.4 mM. Cultures were incubated at 37°C with shaking for an additional 2 hours. Cells were harvested by centrifugation at 7,000 RCF for 10 minutes. Cell pellets were either lysed immediately for purification or stored frozen at −80°C until use.

Cells to be lysed were transferred to 50 ml conical polypropylene tubes (Sarstedt, Newton, NC) and resuspended in 5 ml lysis buffer per gram of wet
weight cells. Lysis buffer is 50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole brought to pH 8.0 by addition of NaOH. Lysozyme was added to 1 mg/ml and the slurry incubated on ice for 30’. Cells were sonicated for 3 minutes using a Misonix XL2015 (Misonix, Farmingdale, NY) with a microtip on 50% pulse cycle at a power setting ≤3. The lysate was treated with DNaseI (Roche Applied Science, Indianapolis, IN) at 5 μg/ml and RNaseA (Sigma, St. Louis, MO) at 10 μg/ml and incubated for 15 minutes at room temperature, or until no longer viscous. Phenylmethylsulphonyl flouride (PMSF) was added to 1 mM immediately after addition of nucleases. The lysate was cleared by centrifugation at 4°C at 10,000 RCF for 30 minutes.

One ml of resuspended Ni-NTA conjugated agarose (Qiagen, Valencia, CA) slurry was added per 4 ml of clarified lysate. The Ni-NTA-agarose lysate mixture was incubated with gentle agitation on a Labquake® rotator (Barnstead/Thermolyne, Dubuque, IA) for 2 hours at 4°C. After incubation, the lysate/Ni-agarose mixture was poured into a Poly-Prep (Bio-Rad Laboratories, Hercules, CA) disposable chromatography column connected by luer fitting to a stopcock and tubing. The Ni-NTA matrix washed with 4 ml wash buffer per ml of slurry. Wash buffer is lysis buffer, but with 20 mM imidazole. The his-tagged scaffolding protein was then eluted using 5 ml of elution buffer. Elution buffer is wash buffer, but with 250 mM imidazole. The eluate was collected in 500 μl fractions. The presence and purity of his-tagged scaffolding protein in each fraction was determined by separating and visualizing the protein composition of each fraction by SDS-PAGE and staining with Coomassie Brilliant blue. A typical gel is shown in Figure 4.1, panel a. The appropriate fractions were pooled
Figure 4.1. Purification of materials for *in vitro* shell binding and assembly assays. A. Coomassie stained SDS-PAGE gel showing components of the his-tagged scaffolding protein purification. B. Coomassie stained SDS-PAGE gel showing that coat protein shells have lost their complement of scaffolding protein after being sedimented twice through 0.5 M GuHCl.
and dialyzed against buffer B (50 mM Tris, pH 7.6, 25 mM NaCl, 2 mM EDTA). The purification generally yielded 4 ml of scaffolding protein solution at approximately 7 mg/ml. The protein was used without further purification and was found to be stable, when frozen in aliquots at −20°C, for at least 3 months.

**Preparation of procapsids and coat protein shells**

*Salmonella enterica* (serovar Typhimurium) strain DB7000 (Botstein, 1968) was grown at 37°C with gentle agitation at 37°C to OD595 = 0.5-0.6 in LB made with 5 g/L NaCl and infected with P22 strain 2amH200, 13amH101, c1-7 at a multiplicity of infection (MOI) of seven. The 2− defect in DNA packaging prevents morphogenesis beyond the procapsid stage, 13− prevents lysis and allows procapsids to accumulate intracellularly, and c1− blocks lysogeny to ensure the infection proceeds via the lytic pathway. Incubation of the infected cultures was continued for 90 minutes and cells were harvested by centrifugation at 7,000 RCF for 10 minutes. The cell pellet was resuspended in 1/125th volume buffer TM (10 mM Tris pH7.5, 1 mM MgCl2) and lysed by the addition of 1/100th of the resuspended pellet volume of chloroform followed by vortexing. The chloroform lysate was treated with DNasel at 50 μg/mL for 10 minutes at room temperature. Immediately after the addition of nuclease, PMSF was added to the lysate to a final concentration of 1 mM. After 10 minutes, or when the lysate was no longer viscous, cellular debris was removed by centrifugation at 15,000 RCF for 15 minutes at 4°C. This clarified lysate was layered atop 20% sucrose buffered with TE (10 mM Tris pH7.5, 1 mM EDTA) at a lysate volume to centrifuge tube volume ratio between 0.2 and 0.4, and spun in the ultracentrifuge at 100,000-
140,000 RCF for 90 minutes. After the supernatant was decanted, the resulting glassy pellet consisting of mostly procapsids, was resuspended by gentle agitation overnight in 100-500 µl buffer TE at 4°C.

The opalescent blue solution of resuspended procapsids was layered atop a 10-40% sucrose gradient at a protein mass to gradient volume ratio of 10-40 µg/mL. Gradients were formed in Ultraclear® SW41 ultracentrifuge tubes (Beckman-Coulter, Fullerton, CA) using a Gradient Master (BioComp Instruments, New Brunswick, Canada). Sample volume was 500 µl. Gradients were centrifuged at 25,000 rpm in an SW41 swinging bucket rotor (Beckman-Coulter, Fullerton, CA) 120 minutes at 8°C. The visible procapsid band, slightly overloaded at the given conditions, migrated to the middle quarter of the gradient, and was harvested by syringe and needle through the wall of the centrifuge tube. Insoluble aggregates were presumed to have pelletted to the bottom of the tube. This solution of purified procapsids was subsequently pelleted by ultracentrifugation at 100,000-140,000 RCFs for 90 minutes and resuspended in an appropriate small volume of buffer B (50 mM Tris, pH 7.6, 25 mM NaCl, and 2 mM EDTA).

Coat protein shells were prepared by first incubating procapsids in buffer B with 0.5 M GuHCl for 20 minutes followed by pelleting procapsids through sucrose and buffer B containing 20% sucrose and 0.5 M GuHCl. The resulting pellet was resuspended in a small volume of buffer B and pelleted again through the same mildly denaturing conditions. This pellet was thoroughly drained and resuspended in an appropriate small volume of buffer B. Loss of scaffolding protein was checked by comparing material from pellets with the original procapsid stock by SDS-PAGE/Coomassie staining (see Figure 4.1, panel b).
this manner, coat protein shells with >95% scaffolding protein removed were prepared. Coat protein shells were stable for at least 3 months when stored at 4°C in buffer B.

Shell binding assay and salt sensitivity assay

Sixty-five µg of scaffolding protein (wild-type or mutant) was added to 50 µg of coat protein shells (hereafter referred to as “shells”) at molar ratio two scaffolding/coat subunit in a final volume of 200 µl with buffer B and incubated for 2 hours at room temperature. Protein concentrations were previously determined by measuring their absorbance at 280 nm and calculating extinction coefficients from their amino acid sequence (von Hippel & Gill, 1989). After incubation, the mixture was loaded atop a layer of 20% sucrose in buffer B and spun in an ultracentrifuge for 90 minutes at 40,000 rpm at 20°C. The supernatant was decanted and the resulting pellet resuspended in 50 µl of TE overnight at 4°C with gentle agitation. Resuspended pellets were mixed with 20 µl 6X SDS-sample buffer (0.28 M Tris, pH 6.8, 30% glycerol, 1% SDS, 6% 2-mercaptoethanol, 0.0012% bromophenol blue) and boiled for 3 minutes prior to loading. Samples were resolved by 12% SDS-PAGE and proteins were visualized by Coomassie staining. To assay for salt sensitivity, shell-binding reactions were performed under similar conditions but in the presence of 25, 50, 75, 100, 250 and 500 mM NaCl. The binding reactions were then centrifuged through 20% sucrose-buffer B solutions containing the same concentrations of NaCl.
Coat monomer preparation and in vitro assembly reactions

Preparation of coat protein monomers and in vitro assembly reactions were performed as previously described. Briefly, concentrated shells prepared as described above were diluted into 6 M GuHCl/buffer B to a final concentration of 1.2 mg/ml such that the final GuHCl concentration was not below 3 M. The dissociation and denaturation reactions were performed at room temperature. The dissociated coat protein was transferred to Spectrapor dialysis tubing, MWCO 10,000-12,000 Daltons (Spectrum Laboratories, Rancho Dominguez, CA), and dialyzed against pre-chilled buffer B at 4°C with three buffer changes to remove denaturant. Each dialysis step was at least 90 minutes. Coat protein concentration after dialysis was usually about 0.8-1.0 mg/ml due to dilution by osmosis. Coat monomers were used on the same day they were prepared.

In vitro assembly reactions were performed in buffer B and consisted of mixing coat monomer and his-tagged scaffolding protein in a quartz cuvette thermostatted at 20°C and monitoring turbidity at 250 nm at 30 second intervals for 90 minutes in a spectrophotometer. Mixing of the assembly reaction was accomplished by adding the components to the cuvette and pipetting up and down three times using a Pipetman P200 (Rainin, Oakland, CA) set to 200 μl. After assembly, material from the cuvette was resolved by centrifugation through a 5-20% sucrose gradient and fractionated. An aliquot of the assembly mixture after 90 minutes was also electrophoresed through 1% SeaKem LE agarose (FMC BioProducts, Rockland, ME) in 1X TBM (40 mM Tris, 20 mM acetic acid, and 1 mM MgSO₄). Gels were fixed by three 20-minute washes in 95% ethanol and dried thin on a slab gel dryer. The gel was stained in 25%
isopropanol, 10% acetic acid, and 0.4% Coomassie Brilliant Blue R-250 and destained first by washing in 25% isopropanol, 10% acetic acid for 1 to 5 minutes, and finally by incubation in 10% acetic acid until the background was sufficiently clear to allow visualization of protein bands. Spectrophotometric data from the assembly reactions were saved in spreadsheet format and plotted using Kaleidagraph™ (Synergy Software, Reading, PA) graphing software.

Results

Point mutational analysis of scaffolding protein in vivo

As described in the previous chapter, an in vivo assembly approach was used to define a minimal coat protein binding domain consisting of the scaffolding protein residues 280-294. In the course of those experiments, SCAFF238-303 was shown to assemble by the in vivo methodology described in Chapter 3. We supplied our plasmid construct that expresses this truncated protein to Sun et al. who used it to produce protein for NMR spectroscopic determination of the solution structure of the coat protein binding domain (Sun et al., 2000). This structure is shown in Figure 4.2. Based on this structure, point mutations were generated within the 15 amino acid coat protein binding domain that were expected to perturb residues that might participate in specific amino acid interactions that support binding as well as residues that contribute to the stabilization of the helix-loop-helix motif. These point mutations were generated by PCR with mutant primers, the mutant genes were cloned into the pET3a expression vector as described in the Materials and Methods section of Chapters 3 and 4.
Figure 4.2. Structural models of the P22 scaffolding protein C-terminal 40 amino acids. A. Ribbon diagram of the scaffolding protein's C-terminus. B. An α-carbon trace with labels shows the positions of amino acids within the structure. Amino acids at the boundary of the minimal coat protein binding domain are underlined. C. A ball and stick rendering shows backbone carbon, nitrogen, and oxygen as well as amino acid side chains. Each representation of the scaffolding protein's C-terminus depicted was modeled first in Deep-View (Swiss PDB-viewer) and rendered in the Persistence of Vision ray-tracing software. All renderings are from the same view of the molecule based on structural coordinates found in the PDB file 2GP8 (www.pdb.org).
The point mutant scaffolding protein genes were expressed from these plasmids and assayed for *in vivo* function during a scaffolding protein deficient P22 infection as described for the truncation mutants in Chapter 3. After infection, the products of assembly were separated by sucrose gradient centrifugation. Fractions were collected and resolved by SDS-PAGE in order to determine the protein content of material large enough to sediment at a position similar to procapsids.

Seventeen single point mutations and one double mutation were generated, and these are listed in Table 4.2. The mutations were constructed in the context of the SCAFF141-303 fragment. This truncated scaffolding protein was initially chosen because of concerns about autoregulation by the N-terminal half of the scaffolding protein affecting expression levels. It was not known at the time that truncation of the N-terminal half of the scaffolding protein could affect its affinity for the coat protein. Since it had been previously suggested that charged residues were important for the binding of scaffolding protein to coat protein (Parker & Prevelige, 1998), six charged residues were substituted individually by alanine. All six proteins were able to bind to coat protein in the *in vivo* assembly assay (data not shown). In order to make the amino acid substitutions at these sites more severe, the charged amino acids were substituted by oppositely charged residues. These six charge switching mutants all bound coat protein in the *in vivo* assembly assay as did a double point mutant that switched two residues to opposite charge (data not shown). Similarly, a point mutant substituting a solvent exposed hydrophobic valine at position 289 to aspartate, as well as five other amino acid substitutions placing charged residues at hydrophobic and polar sites at the interface between helices in the
Table 4.2

Properties of P22 scaffolding protein with amino acid changes in the coat protein binding domain

<table>
<thead>
<tr>
<th>Mutation</th>
<th>SCAFF1-303 shell binding</th>
<th>SCAFF141-303 shell binding</th>
<th>SCAFF1-303 in vitro assembly</th>
<th>SCAFF141-303 in vivo binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild-type</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>D281A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>D281K</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>A284D</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>K286A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>K286E</td>
<td>Y</td>
<td>+/-</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>G287A</td>
<td>+/-</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>G287P</td>
<td>N</td>
<td>+/-</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>D288A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>D288R</td>
<td>Y</td>
<td>+/+</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>V289A</td>
<td>+/+</td>
<td>+/+</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>V289D</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>E290A</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>E290K</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>T291D</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Y292D</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>R293A</td>
<td>+/-</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>R293E</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>K294A</td>
<td>Y</td>
<td>+/—</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>K294E</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>R295E</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

aY indicates activity similar to wild-type protein, N indicates that no activity was observed, and +/- indicates a partial effect was observed. Gray boxes indicate conditions not tested. *In vitro* coat shell binding and *in vitro* assembly reactions described in the text.

Proteins were expressed from pET15b, purified and tested in N-terminally his-tagged form.

Mutation is in context of SCAFF141-303 scaffolding protein truncation mutant expressed from pET3a, not histidine tagged.

Wild-type in this instance is SCAFF141-303 fragment without amino acid substitutions.

Sensitivity to NaCl was not increased relative to wild-type.

None of the mutants tested assembled as well as wild-type.
coat protein binding domain, all bound to coat protein in this shell binding assay (data not shown). The results of these experiments are summarized in the last column of Table 4.2.

**Point mutational analysis using an in vitro shell binding assay**

Since attempts at finding simple non-functional coat protein binding domain mutants using an *in vivo* approach failed, an *in vitro* shell binding assay was developed to screen mutant scaffolding proteins for modified binding ability. The ability of scaffolding protein to enter and bind procapsid shells whose normal complement of scaffolding protein had been extracted by treatment with GuHCl has previously been characterized (Fuller & King, 1981; Greene & King, 1996). The use of preassembled coat protein to assay scaffolding function is arguably biologically not relevant since entry into and binding of empty coat protein shells may not be a normal function of scaffolding protein. However, on the other hand, there is no data to suggest that binding of scaffolding protein to already assembled procapsids does not occur *in vivo*. The assay, described in Materials and methods and Figure 4.3 provides a practical means to measure binding of scaffolding protein mutants to coat protein.

SDS-PAGE display of proteins in purified preparations of coat protein shells can be seen in Figure 4.1. The coat protein shells are relatively intact as can be seen in the electron micrographs in Figure 4.4. Purified recombinant wild-type and mutant scaffolding proteins, each bearing an N-terminal polyhistidine tag, were tested for their ability to co-sediment with shells through a buffered solution containing 20% sucrose. In these experiments, the influence of
Figure 4.3 Coat protein shell binding reactions. A) Schematic outline of the *in vitro* procapsid binding assay. B) The shell binding assay. Representative Coomassie blue stained, SDS-PAGE gels of some of the mutant scaffolding proteins tested in the shell binding assay are shown. Above, the lanes contain the total reaction components; below, lanes contain the re-isolated shells with any bound scaffolding protein.
Figure 4.4. Electron microscopy of shell binding materials and reactions. 
A) Wild-type procapsids containing scaffolding protein. B) Empty coat protein shells after GuHCl induced release of scaffolding protein. C) Shells reacted \textit{in vitro} with full-length scaffolding protein. D) Shells reacted \textit{in vitro} with SCAFF141-303. Samples were prepared by negative staining as described in Materials and Methods and viewed at a magnification of 40,000X.
truncation-induced changes in scaffolding protein function was avoided by engineering the mutations into the full-length protein using mutagenic PCR primers and cloning the resultant product into the pET15b vector as described in Materials and Methods. A typical preparation of purified recombinant his-tagged scaffolding protein is shown in Figure 4.1.

When shells were used at 5 μM coat protein concentration, purified scaffolding was able to reenter at 2.5, 5, 10 and 20 μM concentrations, but the saturating ratio of scaffolding protein bound to coat protein shells after in vitro re-entry was about half of that seen for procapsids assembled in vivo (data not shown). This presumably reflects either differences in accessibility to sites available during de novo assembly versus re-entry of scaffolding protein into an already polymerized coat protein shell or the intactness of the coat protein shells.

The particular amino acid substitutions chosen to be tested for coat shell binding were predicted to be disruptive to the helix-loop-helix motif of the coat protein binding domain, or to salt bridges predicted to form between charged residues of the scaffolding and coat proteins. The mutations expected to disrupt the structure of the motif were substitutions of hydrophobic and hydrogen bonding polar residues located at the interface of the helices by charged residues. The glycine at position 287 could function as a flexible hinge to allow for the correct structure of the loop and positioning of the helices with respect to each other. Substitution of glycine by proline at position 287 was expected to result in peptide backbone angles that would be different from the wild-type structure. In the case of surface accessible charged residues, the charge substitutions that were chosen are opposite in charge. The locations of the mutations are illustrated in Figure 4.5.
Figure 4.5. Locations of point mutations within the structure of the scaffolding protein coat binding domain. The five mutations that abrogated coat protein shell re-entry are shown in panel A. The positions that tolerated substitution in this assay are shown in panel B. Those mutations in parentheses showed partial binding.
Shell binding assay demonstrates loss of binding ability for scaffolding proteins with certain non-conservative amino acid substitutions

The ability of the scaffolding protein mutants to reenter the procapsid are shown in Table 4.2, panel b. Of the 18 changes made, only five mutants, A284D, G287P, V289D, Y292D, and R293E, show no ability to bind to coat protein shells. The loss of binding displayed by these mutants may be the result of a complete abrogation of coat protein binding, or a drastically reduced affinity of scaffolding protein for coat protein. Weakly binding scaffolding protein mutants could diffuse out of procapsids during sedimentation. The large volume of the sucrose medium through which the proteins must sediment (5 ml) relative to the small volume of the binding reaction (0.5 ml) may also contribute to the loss of scaffolding protein from shells through a 10-fold dilution of the reaction components during sedimentation. Nevertheless, the results shown in Figure 4.4 demonstrate a binding or loss of binding phenotype for the wild-type and mutant scaffolding proteins in this assay.

Shell binding ability retained in scaffolding protein mutants with alanine substitutions

Because the first set of mutations tested could be considered harsh in terms of protein function (i.e., they are not conservative substitutions), we also tested a number of alanine substitutions at positions where the nonconservative changes abolished coat binding. The alanine substitutions are predicted to be more benign and therefore might serve as a better indicator of an absolute requirement for a specific amino acid class at a given site within the primary
sequence of the protein (Baase et al., 1992; Matthews, 1996). The three substitutions, G287A, V289A, and R293A were tested for ability to bind coat protein as described above. All four of the alanine substitutions tested were able to bind, although the stoichiometry appeared reduced in each case: about half as much protein incorporated compared to wild-type scaffolding protein (data not shown). The results from these experiments are summarized in Table 4.2, and their location within the structure of the scaffolding protein C-terminal domain is illustrated in Figure 4.5.

**Scaffolding protein mutations K286E and E290K do not alter of shell binding at high ionic strength**

Scaffolding protein has been proposed to bind to coat protein via electrostatic interactions. High concentrations of NaCl have been shown to inhibit both in vitro procapsid assembly and shell binding by scaffolding protein, suggesting that high ionic strength perturbs coat-scaffolding and/or scaffolding-scaffolding interactions rather than coat-coat interactions. To test whether or not charge substitutions within the coat protein binding domain would change the sensitivity of shell binding to salt, we performed the shell binding assay in the presence of increasing concentrations of NaCl. Although the two mutations tested did not themselves have a discernable effect on re-entry under the low salt conditions used in our assay, we expected that a single charged amino acid substitution that did not abolish re-entry might still be able to weaken the affinity
Figure 4.6 NaCl inhibition of scaffolding protein shell binding. Shell binding reactions were set up as described in Materials and Methods in the presence of the indicated concentration of NaCl. Reactions were pelleted through sucrose containing the same salt concentration as the filling reaction. Wild-type and mutant scaffolding protein were run in pairs for comparison. Load not shown for E290K, but it was indistinguishable from the K286E samples. Neither mutant shows increased sensitivity to salt inhibition.
of scaffolding for the coat protein and that the strength of this affinity might be revealed by an increased sensitivity to inhibition of shell binding by high ionic strength.

Wild-type or mutant scaffolding proteins were mixed with coat protein shells in the presence of 25, 50, 75, 100, 250, and 500 mM NaCl buffered with Tris pH 7.6 and 2 mM EDTA. After 2 hours, the mixture was pelleted through 10% sucrose in assembly reaction buffer. The amount of scaffolding protein bound to shells was visualized by SDS-PAGE of the resuspended pellets. As shown in Figure 4.6, neither the K286E, nor the E290K mutant showed any change in sensitivity to salt by qualitative assessment of the gels. Densitometric analysis of the gels further supported this result (data not shown).

**Scaffolding protein point mutant G287P in SCAFF141-303 background is able to bind shells**

The mutants that we tested above for assembly in vivo were in the background of a scaffolding protein truncation mutant composed of residues 141-303. After we discovered that truncation of the N-terminus could have effects on the scaffolding protein’s affinity for the procapsid, either by affecting the affinity of scaffolding protein for coat protein or by affecting scaffolding-scaffolding interactions within the procapsid, we decided to test, by in vitro shell binding assay, four nonbinding mutants within the context of the truncated protein. SCAFF141-303/A284D, SCAFF141-303/Y292D, SCAFF141-303/R293E
all failed to bind shells (data not shown). However, SCAFF141-303/G287P was able to bind (data not shown) although the full-length G287P mutant did not bind.

**In vitro assembly of wild type and mutant scaffolding proteins reveals phenotypes not congruent with shell binding phenotypes**

Prevelige and co-workers have characterized the *in vitro* association of unassembled wild-type coat and scaffolding proteins into procapsid-like particles. In order to further characterize the effects scaffolding protein single amino acid substitutions have on protein function, selected mutant scaffolding proteins were tested for their ability to mediate *de novo* assembly of coat protein into procapsid-like particles *in vitro*. Unassembled coat protein subunits were prepared as described in Materials and Methods. Coat protein (between 15 and 20 μM) and scaffolding protein (between 13 and 27 μM) were mixed in a cuvette thermostatted at 20° and the kinetics of assembly were spectrophotometrically monitored by measuring turbidity at 250 nm over time. The change in turbidity over time is plotted in Figures 4.7 and 4.8. The concentrations of coat and scaffolding proteins used in each assembly experiment are given in the insets of Figures 4.7 and 4.8.

Coat protein alone increases in turbidity over the course of the experiment by less than 0.05 OD units showing that at these concentrations it self-associates by itself very poorly. However, assembly reactions containing coat protein mixed with the his-tagged wild-type scaffolding protein had OD changes ranging from
Figure 4.7. Spectrophotometric analysis of in vitro assembly reactions. Assembly reactions of purified coat and scaffolding protein as described were monitored spectrophotometrically by absorbance at 250 nm over time. A) Assembly curves are shown. The numbers in parentheses are the micromolar concentrations of coat and scaffolding proteins respectively. The products of assembly in this top panel were analyzed by sucrose gradient centrifugation and agarose gel electrophoresis and shown in Figures 4.9 and 4.10. B) The bottom panel shows assembly of G287 mutants at different concentrations. The early time points have been expanded to emphasize lag in G287P stimulated assembly.
Figure 4.8 Coat and scaffolding protein assembly *in vitro*. *In vitro* assembly of purified coat and scaffolding protein was monitored spectrophotometrically by absorbance at 250 nm over time. A) The assembly curves of two mutant proteins are shown. The numbers in parentheses are the micromolar concentrations of coat and scaffolding proteins respectively. B) Assembly of additional mutants altered at V289.
1.0 to 1.5 units. None of the mutant assembly reactions tested gave OD changes as high as wild-type. At completion, assembly reactions containing T291D and K286E mutant scaffolding protein were approximately half as turbid as the wild-type reaction. G287P and G287A reactions were approximately one third the turbidity of the wild-type assembly reaction. The assembly reaction containing V289D did not increase in turbidity, and V289A showed a weak increase turbidity. Neither the R293E nor the R293A scaffolding protein mutants supported an increase in turbidity of the assembly reaction.

The assembly of coat protein into procapsid-like particles was confirmed both by sedimentation of assembly products through a sucrose gradient as well as resolving the assembly products by agarose gel electrophoresis. These data are shown in Figures 4.9 and 4.10. Approximately half of the coat protein in the assembly reaction containing wild-type his-tagged scaffolding protein assembled into structures that sedimented at the position expected for procapsids made during a 2^+ infection. Coomassie stained SDS-PAGE gels of assembly reaction fractions sedimented in a sucrose gradient for mutant scaffolding proteins altered at positions 287 and 293 are shown in Figure 4.9. An aliquot of the total assembly mixture was also resolved by nondenaturing electrophoresis through 1% agarose, shown in Figure 4.10. The gels show unassembled reaction components staining as a diffuse band near the top of the gel. The larger, procapsid-like material ran further into the gel that was visualized with Coomassie staining as a tight band. These particles ran at positions where authentic procapsids made during a 2^+ infection had previously been shown to migrate in other gels (not shown).
Figure 4.9 Sucrose sedimentation of in vitro assembly products. Coomassie stained, SDS-PAGE gels of fractions of in vitro assembly reactions sedimented into 5-20% sucrose gradient are shown. An aliquot of the total mixture loaded onto the gradient was resolved in the first lane of each gel, indicated by L below. The arrow shows the direction of sedimentation with respect to the position of the fractions within the gradient. Of the two prominent bands, the higher MW band in each lane is coat protein, the lower MW band is the his-tagged scaffolding protein. Note: in some lanes, the coat protein appears to be less concentrated than the scaffolding protein. This difference was later determined to be due to binding of coat protein to polystyrene during sample preparation (S. Moore, personal observation.)
<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>coat alone</td>
<td></td>
</tr>
<tr>
<td>coat + scaffolding wild type</td>
<td></td>
</tr>
<tr>
<td>coat + scaffolding G287P</td>
<td></td>
</tr>
<tr>
<td>coat + scaffolding G287A</td>
<td></td>
</tr>
<tr>
<td>coat + scaffolding R293E</td>
<td></td>
</tr>
<tr>
<td>coat + scaffolding R293A</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.10. Agarose gel electrophoresis of *in vitro* assembly reactions.** PC indicates migration position of authentic procapsids made in P22 2^- infection. Unassembled proteins migrate near top of gel.
The ability of mutant scaffolding proteins to support assembly in vitro is summarized in Table 4.2. Interestingly, the G287P mutant, which was unable to bind shells in vitro, was able to promote assembly. By comparison, the G287A mutant protein, which was able to bind shells, was only moderately better than the G287P protein at promoting assembly. In contrast to the mutations at position 287, the R293A mutation which caused only a moderate defect in ability to bind shells, did not promote assembly at all.

G287P mutant scaffolding protein displays lag relative to wild type in the early part of assembly in vitro

Expansion of the early time points in the assembly curve shows a marked lag phase not seen in the wild-type protein or in the G287A mutant (see Figure 4.7, panel b). The slopes of the assembly curves approximate the rate of assembly since the spectrophotometric signal contains the average of all structures formed in the cuvette. By this measure, once G287P assembly gets started, the two mutants display similar rates of assembly. However, the final yield of assembly is apparently slightly reduced for G287P relative to G287A.

Discussion

Scaffolding protein mutant phenotypes in vivo

A total of 21 different amino acid substitutions were made and tested for coat protein binding in vivo. Each mutant protein was found to bind coat protein. Several explanations can rationalize the persistence of the scaffolding-coat protein interaction in vivo in the face of so many mutational changes in scaffolding protein. The amount of coat and scaffolding proteins expressed
during the above infections constitute a significant percentage of the total protein content. Therefore, even a severely weakened scaffolding-coat protein interaction might be driven towards association at the high concentrations of present in the cell. If, once assembled in the procapsid, scaffolding protein binding to coat protein is further stabilized by scaffolding-scaffolding interactions, then weakened binding might be further masked. Finally, truncation of the N-terminal half of scaffolding protein has been shown to apparently increase the affinity of scaffolding protein for coat protein (see also Appendix B), although the exact basis for this apparently increased affinity is not yet known. Nonetheless, the in vivo assembly data also suggest that binding of scaffolding protein to coat protein is genetically very robust, since most single mutations within a domain essential for binding to coat protein show no phenotype in our assay. Either the interaction involves only a few amino acids or it involves more amino acids but is flexible and can tolerate many different changes.

**Scaffolding mutant phenotypes in vitro**

In the procapsid binding assay, 5 out of 14 single amino acid substitutions (in the context of the full-length scaffolding protein) failed to stably interact with the preassembled coat protein lattice. Two obvious models to account for the loss of binding are 1) disruption of specific scaffolding-coat amino acid interactions and 2) unfolding or destabilization of coat protein binding motif. A third and more subtle possibility is that mutations interfere with conformational changes necessary to effect binding. The types and distribution of mutations that abolished binding do not appear to be especially revealing about the mode of binding.
Electrostatic interactions have previously been suggested as the basis for scaffolding protein binding interaction. Only one out of six charge switching substitutions, R293E, abolished shell binding, and this change did not block in vivo assembly of SCAFF141-303/R293E. Two of those substitutions which did permit shell binding, K286E and E290K, did not display weakened shell binding when challenged with salt. In contrast, two hydrophobic to charge substitutions, A284D and V289D, abolished binding. One neutral polar to charge substitution, Y292D, abolished binding, the other, T291D, did not. These findings might suggest the importance of hydrophobic and noncharged polar interactions in binding. However, only V289 is solvent exposed. The remaining hydrophobic residues are found at the interface between the helices and are therefore less likely able to interact with coat protein. Based on the structure, A284 and Y292 appear to be involved in stabilizing helix-helix interactions and loss of binding might therefore be due to disruption of the helix-loop-helix motif.

Out of four mutations expected to perturb the two amphipathic helices, either by altering the loop conformation (G287P), or by interfering with residues expected to mediate interactions between helices (A284D, T291D, Y292D), only the T291D mutant retained the ability to bind. These data suggest that the folded helix-loop-helix is required for binding of scaffolding protein to the coat lattice. The G287P mutant seems to best support a folded motif requirement since this mutant almost certainly disrupts the conformation of the coat protein binding domain. The proline substitution constrains the peptide bond at position 287 to a φ angle of \(-65^\circ\), which does not fit with the configuration of the native loop revealed in the NMR structure and prevents the packing of the two helices against each other. However, requirement for a stable helix-loop-helix motif
seems to be contradicted by the ability of the SCAFF141-303/G287P mutant to bind shells. Furthermore, the full-length G287P mutant which cannot bind preassembled shells is nonetheless active in assembly. These two mutants suggest the possibility that the conformation of scaffolding bound to coat protein might be different from that revealed by the solution structure of the coat protein binding domain.

What could the nature of such a conformational change be and how could it be achieved? A conformational change that mimics the G287P mutant would involve separation of the two helices. The helix-loop-helix motif is a thermodynamically favorable structure and therefore energy would be required to separate the helices. This energy could be provided by favorable interactions between charged scaffolding residues with partner amino acids in the coat protein. However, if there were a conformational change in binding, one would expect it to involve more than a simple separation of helices, since mutations expected to disrupt helix packing do not increase binding. Recently, a screen for phage resistant to toxic effects of overproduction of heterologously expressed portal protein in the bacterial host generated survivors that all carry the same mutation in the scaffolding gene, G287E (S. Moore, personal communication). Modeling of this mutation within the coat protein binding domain predicts conversion of the loop portion of the motif to an alpha helix connecting the proximal and distal helices into one long helix.
Differences in shell binding and coat assembly phenotypes suggest functional differences in how scaffolding protein binds coat

Procapsid-like particles assembled in vitro in the presence of excess scaffolding protein contain appear to contain more scaffolding protein than coat protein shells that have been re-filled by incubation in a similar excess of scaffolding protein. For shell binding, the entry and binding of a certain number scaffolding protein molecules to the inner surface may preclude the binding of additional molecules of scaffolding protein. Or, a subset of scaffolding protein binding sites on coat protein subunits may only be available to during assembly. The latter possibility seems to be suggested by the G287P and the R293A mutations. G287P is able to stimulate assembly of coat protein in vitro, yet is unable to bind shells in vitro. Conversely R293A binds shells, but is unable to stimulate assembly. Therefore, G287P may be able to bind a conformation of coat protein only available during assembly, but not to coat protein in its assembled form. R293A may be unable to bind a preassembled conformation of coat protein.

The inability of G287P to bind shells may, however, reflect the specific properties of this assay. If G287P alters the conformation of the C-terminus of the scaffolding protein such that it is “splayed,” this mutant scaffolding protein might be sterically blocked from entering through the hexon holes even if it were able to bind coat should it get inside. Assembly with free coat protein subunits would not present this steric hurdle. The SCAFF141-303/G287P mutant is able to bind empty shells. Perhaps this mutant is smaller and can fit through the hexon holes. Hydrodynamic modeling of full-length scaffolding and SCAFF141-303
proteins suggest these proteins have the same length; however, actual measurement of the mutant’s hydrodynamic properties might be informative.

Another possibility to explain the different assembly and shell binding phenotypes of the G287P and R293A mutant proteins is that for these mutants, binding to coat and activation of assembly have been uncoupled. This possibility may not easily accommodate the G287P phenotype, but it is possible that the conformation of G287P is the conformation required to activate assembly of the coat protein. R293A binds to coat protein in the assembled form, but in the unassembled state, this mutant might not supply the correct surface to coat protein to stimulate assembly. If scaffolding protein has separable binding and assembly activation functions, this might explain the lag phase observed during G287P mediated assembly of coat protein. The G287P mutant might be slower in forming associations with coat protein, but after sufficient time, enough scaffolding protein-coat protein complexes might form to nucleate assembly.

**Models of binding and assembly**

Fundamental questions about how scaffolding protein is able to promote assembly of the coat protein remain. P22 procapsid assembly has been proposed, but not proved, to be nucleation limited. However, a simple simulation of virus assembly using only thermodynamic parameters behaves similarly to nucleation limited assembly processes (Zlotnick, 1994). The simplest model for how scaffolding protein is able to promote assembly is that the scaffolding protein is able to recognize, bind and stabilize an assembly competent conformation of coat protein and thereby increase the concentration of assembly competent coat protein subunits to the point where assembly becomes favorable. A wealth of
data from assembly of other macromolecular complexes such as the flagellum and TMV demonstrate the importance of assembly-induced conformational changes during the self-assembly process. The *in vitro* phenotypes of the G287P and R293A scaffolding protein mutants discussed above suggest that assembly is more complex than simple binding of scaffolding protein to coat protein.

Further speculation as to the mechanism of scaffolding protein binding might not be fruitful, but based on the available data, we propose a model to offer testable predictions for scaffolding protein function. Binding of the scaffolding protein requires both hydrophilic and hydrophobic amino acids in the coat protein binding domain. The primary purpose of the hydrophilic contribution might be to provide orienting docking interactions but might not contribute to the overall strength of the scaffolding-coat interaction. If this were the case, one might expect that the binding of certain charge substitutions might have an effect on the rate of association but not dissociation. A rate of dissociation might be estimated if one were able, for example, to chase labeled scaffolding protein out of procapsids with unlabelled scaffolding protein. Assuming one could have wild-type and mutant scaffolding filled procapsids with the same starting stoichiometry, this might be achieved.
APPENDIX A

CONFORMATION OF SCAFFOLDING
PROTEIN C-TERMINUS
WITHIN PROCAPSIDS

Introduction

During phage assembly in vivo, the scaffolding protein exits the procapsid at around the time of DNA packaging. Exactly how the scaffolding protein is able to do this is unknown. DNA packaging may alter the charge properties of the procapsid interior such that the scaffolding protein’s affinity for the coat protein lattice is diminished. Alternatively, the DNA packaging process may alter the conformation of the coat protein lattice thereby reducing its affinity for scaffolding protein. There may also be conformational changes to the scaffolding protein itself that facilitate its release from the procapsid. Whatever the mechanism, scaffolding protein faces the topological challenge of getting out of the procapsid. Phage P22 scaffolding protein presumably does so through 25-35 Å diameter holes that have been observed at the centers of hexameric clusters (Prasad et al., 1993). Density difference mapping of cryo-electron micrographic reconstructions of scaffolding protein filled procapsids versus empty, unexpanded coat protein shells suggests that a small portion of the scaffolding protein is localized to finger-like projections that protrude into the procapsid interior from the inner rim of the hexon centers, while the bulk of the scaffolding
protein is not icosahedrally ordered (Thuman-Commike et al., 2000; Thuman-Commike et al., 1999). At these locations, scaffolding protein would be well poised to exit from the procapsid. However, the protein is believed to be entirely within the procapsid interior, suggesting there must be some mechanism to guide it out. One hypothesis is that part of the scaffolding protein protrudes slightly from the procapsid shell so that the rest of the molecule is simply extruded out through the holes during DNA packaging thereby eliminating any element of diffusion to exit. In order to test this hypothesis, an epitope was fused to the C-terminus of scaffolding protein and accessibility of the epitope to antibody when the scaffolding fusion-protein was inside the procapsid was examined.

Materials and methods

Codons 141 to 303 of the scaffolding protein gene were amplified by PCR using tailed primers: 5′-AGCTAGCCATGGCCCGCAGCAATGCCGTAGC-3′ (upstream) and 5′-TAAGCGTCTAGAGCTCGGATTCCTTTAAGTTTTGCC-3′ (downstream). Sequences complementary to P22 are given in plain text, the restriction sites bold faced, the first and last codons of the reading frame are underlined and the primer tail is italicized. The PCR product was cloned into the vector pTrcHis2B (Invitrogen, Carlsbad, CA) between the NcoI and XbaI sites. The resulting plasmid adds a 23 amino acid extension (FLQNKLISQQDLNSAVDHHHHH) to the C-terminus of the scaffolding protein fragment and includes six penultimate histidine residues recognized by anti-His polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, CA). Procapsids incorporating this scaffolding-fusion protein were made by expressing the protein in Salmonella during infection with P22 3amN6,
8amN123, 13amH101, c1-7) and purified under as described in Chapter 2. The gel in Figure A.1 demonstrates the incorporation of the fusion protein into procapsid-like material.

Immunoprecipitations using whole procapsids and anti-histidine antibodies (Ab) were performed using 10 µl of protein-G protein conjugated agarose beads (Roche Diagnostics Corporation, Indianapolis, IN), 10 µg of procapsids, and 1 µl of anti-histidine polyclonal antibody. Briefly, procapsids were incubated, with rotation, for 4 hours at 4°C with anti-His Ab in 250 µl phosphate buffered saline (PBS). 10 µl of agarose beads were added to the mixture and incubation was continued for an additional 2 hours. Beads were concentrated by centrifugation at 3,000 RCFs for 1 minute and resuspended in 1 ml PBS. This rinse step was repeated and the beads were concentrated a third time and resuspended in 100 µl 1X SDS-sample buffer. The beads were boiled for 3' and 20 µl of sample were resolved by SDS-PAGE through a 12% gel. The gel was electrophoretically transferred to PVDF membrane (Bio-Rad Laboratories, Hercules, CA) and subsequently probed with anti-portal polyclonal antiserum (generously provided by Sean Moore). One µl of purified procapsids was included as a control to show the presence of portal and coat proteins.

Procapsids were prepared for electron microscopy by spotting 1µl of concentrated phage material (~10 mg/ml) containing $10^9$ to $10^{10}$ particles on a Formvar/carbon coated 400-mesh copper grid (Ted Pella, Redding, CA) for 30 seconds, rinsed once with three drops of 100 mM KCl, and stained with 2% uranyl acetate. Procapsids treated with antibody were similarly prepared, except that 1 µl of a 1:10 dilution of antibody was mixed with 1 µl of procapsid material and incubated for 10 minutes at room temperature prior to spotting on a grid.
Figure A.1. Sucrose gradient analysis of scaffolding:H6 filled procapsids. Material assembled in cells expressing scaffolding fusion protein during infection with a scaffolding deficient phage was separated by sucrose gradient centrifugation as described in Chapter 2 and the protein composition of gradient fractions analyzed by Coomassie stained SDS-gels. Lane designations: CL, crude lysate; L, load; 5 through 11, fractions from sucrose gradient.
Results and discussion

The results in Figure A.2 demonstrate that an anti-His antibody/fusion protein filled procapsid complex bound to the protein-G conjugated agarose beads (lane 3), but no such complex was precipitated with wild-type procapsids (lane 2). This result suggests that the hexavalent histidine epitope was available to the antibody. The protein from the immunoprecipitation was resolved by SDS-PAGE and probed with anti-portal protein polyclonal antibody during Western blot analysis. The presence of portal protein in the immunoprecipitate indicates that a complex consisting of at least scaffolding and portal protein was immunoprecipitated (model "1" in Figure A.2b). The polyclonal anti-portal antibody cross reacts with coat protein, which was also present in the immunoprecipitation. The presence of coat protein suggests that the complex is in fact a procapsid, although the complex could be a partially disrupted shell (models "2" and "3" in Figure A.2B). Partial disruption of the shell might provide access of the epitope to the antibody, allowing the complex to bind to protein-G conjugated beads.

Electron microscopy was performed to see if an antibody/procapsid complex could be visualized directly. As seen in Figure A.3, the scaffolding-fusion protein filled procapsids incubated with anti-his antibodies appear to have thicker shells than either the wild-type or fusion filled procapsids alone as well as wild-type procapsids incubated with antibody. Initially, this additional electron density was interpreted as a layer of antibody bound to the exterior of the shell. However, measurement of the particle diameters in the micrographs suggests that the diameters of procapsids among all the samples may be the same; the additional density is due to thicker shells. One hypothesis to explain
Figure A.2. Immunoprecipitation (IP) of procapsids. A. Western blot probed by polyclonal antibody reacting with portal and coat proteins demonstrates that whole procapsids containing C-terminally histidine-tagged scaffolding protein can be immunoprecipitated using anti-His antibodies. B. Cartoons of complexes capable, in theory, of interacting with the protein G conjugated beads.
Figure A.3 Electron microscopy of procapsids and fusion protein filled procapsids. Negatively stained samples of whole procapsids containing C-terminally histidine-tagged scaffolding protein incubated with anti-His antibodies display additional electron density at the procapsid shell. A) procapsids, wild-type; B) procapsids, fusion filled; C) wild-type procapsids + anti-His Ab; D) fusion protein filled procapsids + anti-His Ab. Samples were viewed with 40,000X magnification.
this thickening of the shell in the presence of antibody is that antibody is interacting with scaffolding-fusion protein histidine epitopes, or a subset thereof, protruding from the procapsid (either from the lattice or perhaps near the portal complex) and that this interaction promotes a change to the procapsid lattice resulting in a thick shell. Procapsids with thick shells have been seen before. Thomas and Prevelige (1991) showed that procapsid-like particles assembled in vitro in the absence of gp16 had thicker, or fuzzy, shells. The particles assembled under their conditions look similar to the fusion-protein filled procapsids incubated with antibody. Later reports of in vitro assembled procapsid-like particles lacking gp16 do not show thick shells for reasons unknown (Prevelige et al., 1993). It is interesting to note that there is a genetic relationship between the scaffolding protein and incorporation of gp16 into the procapsid (Greene & King, 1996).
APPENDIX B

COMPETITIVE PROCAPSID FILLING:
FULL LENGTH VS. SCAFF141-303

Introduction

How is scaffolding protein’s affinity for coat protein regulated? Several lines of evidence demonstrate that scaffolding protein’s N-terminus affects the binding of scaffolding protein to coat protein. A truncation mutant of scaffolding protein consisting of residues 141-303 can promote assembly of coat protein in vitro and does so at a rate of about twice that of full-length scaffolding protein. Gel densitometric analysis of procapsid-like particles assembled in vitro using the SCAFF141-303 reveal a scaffolding protein about twice that of procapsid-like particles similarly assembled using the full-length scaffolding protein. A similar result is observed for procapsids assembled in vivo using full-length versus SCAFF141-303. Mass measurements of these particles using STEM demonstrate that SCAFF141-303 filled procapsids contain twice the normal complement of scaffolding protein (M. Simon and S. Casjens, unpublished data). These data argue for some form of conformational control of the scaffolding protein’s coat protein binding domain. Two ways in which this conformational control might be achieved would be either the N-terminus of the molecule directly interacting with the coat protein binding domain or somehow a conformational change is
relayed along the molecule from the N- to C-terminus. In order to better define how truncation affects scaffolding protein binding to coat protein, binding of scaffolding protein to preassembled coat protein was examined. By using the shell binding assay, the element of shell assembly is removed.

**Materials and methods**

Full-length scaffolding protein and SCAFF141-303 were cloned, expressed and purified as described in Chapter 3. The primers used to amplify codons 141-303 of the scaffolding gene were:

5′-GACGATGCATATGGCTCGCAGCAATGCCGTAGCAG-3′ (upstream) and

5′-GTAGAGAGGATCCTTGGAGTGATTGCGGAGATG-3′ (downstream).

The primers used to amplify the full-length of the scaffolding protein gene were:

5′-CACGATGCATATGGAAACCAACCACCGAAATTCAGG 3′ (upstream) and

5′-GTAGAGAGGATCCTTGGAGTGATTGCGGAGATG-3′ (downstream).

The restriction sites within the primer sequences are indicated in bold type. See Tables 3.1 and 4.1 for more details about oligonucleotides. Coat protein shells were purified from GuHCl treated procapsid preparations as described in Chapter 3. Filling reactions were performed using buffer conditions described in Chapter 3.

Competitive filling reactions were performed as follows. Fifty µg of shells were mixed with full-length scaffolding or SCAFF141-303 protein in a 1:1 coat protein/scaffolding protein molar ratio in a total volume of 200 µl and incubated at room temperature for two hours. Two hours had previously been determined as sufficient time for filling to reach completion. After 2 hours, an additional
molar equivalent of competitor scaffolding protein was added to the filling reaction; i.e., SCAFF141-303 was added to the mixture containing full-length scaffolding, and full-length protein was added to the mixture containing SCAFF141-303. Control reactions included adding the same form of scaffolding protein to the mixture. After an additional 2 hours, the mixtures were layered atop 20% sucrose and procapsids were pelleted by spinning at 40K rpm in an SW50.1 rotor at 16° C for 90 minutes. Pellets were resuspended in 100 µl of buffer B and 25 µl of 6X SDS-sample buffer and boiled for 3 minutes followed by rapid cooling in ice. Ten µl of this mixture were resolved by electrophoresis through a 12% SDS-polyacrylamide gel and visualized by staining with Coomassie blue.

Results and Discussion

Figure B.1 shows an SDS-PAGE of samples from a competitive filling experiment. The first four lanes show a sample of the total protein in each of the filling reactions. The last four lanes are protein recovered from the pellet after centrifugation. As expected, both full-length and 141-303 scaffolding proteins bound to the coat protein shells (lanes 5 and 6). In the filling reaction where full-length was followed with 141-303, the 141-303 was able to bind (lane 7), suggesting that the truncated scaffolding protein a) was able to bind additional sites not accessible to the full-length protein and/or b) was able to displace some of the full-length protein. In the filling reaction where coat protein shells were incubated with the 141-303 scaffolding protein fragment first, almost no full-length protein was able to bind (lane 8), suggesting that all available binding sites within the lattice were occupied by SCAFF141-303 and that wild-type scaffolding protein was not able to displace SCAFF141-303.
Figure B.1 Competitive filling. Coat protein shell filling reactions were carried out with either full- or half-length scaffolding protein. After 2 hours, competitor scaffolding was added and the reaction was allowed to incubate for another 2 hours. Aliquots of the filling reaction were removed and the remainder of the filling reaction was pelleted and both load and pellet samples were resolved by SDS-PAGE.
The observed results might be explained by steric and/or thermodynamic arguments. If the scaffolding protein were in equilibrium between free and bound states in the filling reaction, and both full-length and truncated scaffolding protein had equal affinity for the coat protein lattice, then one might expect that, over time, equal amounts of each form of the scaffolding protein might be found in the pellet at the end of the experiment. If the truncated form of scaffolding had a higher affinity for the coat protein, then it might be able to replace the full-length scaffolding protein in shells that had been preincubated with the full-length scaffolding protein, and exclude full-length scaffolding protein from shells that had been first incubated with the truncated scaffolding protein. Alternatively, the ability of truncated scaffolding protein to better fill coat protein shells may be simply due to the size of the protein. More sites may be available to the smaller scaffolding protein fragment. A truncated scaffolding protein might also be able to more easily fit through the holes in the coat protein lattice than the full-length scaffolding protein. Clearly, the presence or absence of the N-terminal half of the scaffolding protein is affecting how the protein binds to coat protein. How binding is affected remains to be determined.
APPENDIX C

PUTATIVE P22 HOMOLOGS

Introduction

There is a reasonable probability that for every species or serotype of *Salmonella*, there exists at least one and probably more P22-like phages. Generalized transducing phages with P22-like morphology and sequence identity have been isolated from natural isolates of *S. enterica* serotypes (Schicklmaier & Schmieger, 1995). A recent examination of the *Salmonella paratyphi* B genome sequence revealed the presence of a P22 like prophage (data not shown). Many of the phages whose host ranges are used as one criterion in classifying *Salmonella* species (e.g., the Anderson typing phages) have P22-like morphology (Schmieger, 1999). Two phages, epsilon 15 (ε₁⁵) and epsilon 34 (ε³⁴), studied for their ability to alter the O-antigen composition of their host bacterium (a phenomenon called lysogenic conversion) are indistinguishable from P22 in electron micrographs (data not shown). With the goal of finding homologous scaffolding proteins from a P22-like phage, Southern blot analysis was used to probe for P22 homologous prophage within genomic DNAs from 10 different serotypes of *Salmonella* natural isolates (kind gift from John Roth). Similarly, Southern blot analysis was used to test for homology in phages ε₁⁵ and ε³⁴.
Materials and methods

Phages \( \epsilon^{15}, \epsilon^{31} \) (both gifts from Andrew Wright) and P22, were grown by infecting exponentially growing hosts, \( S.\ newington \) and \( S.\ ado \) respectively, at an MOI of 0.1 in 500 ml LB at 37°C with shaking until lysis was achieved. Phage were concentrated by centrifugation and resuspended in buffer TM (10 mM Tris, 1mM MgCl\(_2\), pH 7.6). Phage were further purified by centrifugation through 20% sucrose buffered with TM onto a layer of TM buffered CsCl (\( \rho=1.6 \)). The concentrated phage band was harvested by a hypodermic needle through the side of the tube and then dialyzed against buffer TM. Phage genomic DNA was isolated as described previously. One µg of phage DNA was digested with either \( \text{EcoRI}, \text{BamHI}, \text{or HindIII} \) (New England Biolabs, Beverly, MA). Restriction fragments were resolved by electrophoresis through 1% agarose. The DNA was transferred to membrane by Southern blotting, probed with a PCR product encoding the P22 coat protein gene, and visualized using Phototope-Star Chemiluminescent Detection Kit (New England Biolabs, Beverly, MA) and exposure to x-ray film. The sequences of the primers used to amplify gene 5 (the coat protein gene) of bacteriophage P22 were:

\[
5' - \text{AAACGGGGTACCGCAAGCTAAAGGC} - 3' \quad \text{(upstream)} \quad \text{and} \\
5' - \text{TAAGCGTCTAGACTATTATTACGCAGTCTGACCAGGAGGC} - 3' \quad \text{(downstream)}.
\]

Genomic DNA was isolated from overnight cultures of \( S.\ salmonella \) serovars using a DNeasy kit (Qiagen, Valencia, CA). Approximately 100 ng of genomic DNA from each strain was spotted onto a Hybond-N\(^*\) nylon membrane (Amersham-Biosciences, Piscataway, NJ) using a vacuum-manifold dot-blotter. The DNA was fixed to the membrane by alkali treatment followed by drying. The membranes were probed with whole P22 genomic DNA labeled and
visualized using the Phototope-Star Chemiluminescent Detection Kit (New England Biolabs, Beverly, MA) and exposure to x-ray film.

Phage induction from putative lysogens was achieved by growing cells in LB at 37°C, with shaking, to OD_{595}=0.5 and adding mitomycin-C (Sigma, St. Louis, MO) to a final concentration of 2 μg/ml and continuing incubation for 3 hours. Phage and phage-like particles were purified as described above. Concentrated material was applied to formvar/carbon-coated grids and negatively stained with 2% uranyl acetate.

**Results and discussion**

The digested phage genomic DNA is shown in Figure C.1, panel A. This DNA was transferred in the same configuration to a membrane subsequently probed with the P22 coat protein gene. As seen in panel B of Figure C.1, the probe hybridized, as expected, to the P22 DNA. The probe also hybridized, albeit less strongly, to the ε^{34} DNA. No signal was detected in the ε^{15} lanes, even after prolonged exposure to x-ray film. This result indicates that nucleotide identity between P22 and ε^{34} is sufficiently high enough to support hybridization and suggests their structural genes are homologous and quite similar.

Genomic DNA from various *Salmonella* strains was probed with P22 genomic DNA. As expected, probe DNA hybridized to DNA from two P22 lysogens, but not to DNA from *S. typhimurium* (LT2). The probe DNA also hybridized to genomic DNA from *S. heidelberg*, and, to a lesser degree, to *S. dublin* and possibly *S. abortus-ovus*. To test whether these strains harbored, as lysogens, P22-like phages, liquid cultures were treated with mitomycin-C to induce phage production. Purified material from *S. heidelberg* and *S. abortus-ovus*
Figure C.1 Southern blot of DNA from P22-like phage. A. Genomic DNA of phages P22 epsilon34 and epsilon15 were each digested singly with EcoRI, HindIII, and BamHI. B. DNAs were blotted to a membrane and probed with P22 gene 5 (coat protein). C. Dot blot of genomic DNA from wild-type isolates of Salmonella enterica was probed with P22 genomic DNA.
cultures are shown in panels A and B, respectively, of Figure C.2. The particles in both micrographs appear to lack obvious tailspikes. The lack of tailspikes may be due to the gene being defective in the prophage. The tailspikes of these phages may look altogether different to P22 or they may have simply failed to fold properly under the temperature at which the virion-like particles were produced. Many of the particles purified from the mitomycin-C treated S. abortus-ovus cultures appear to be defective, having the appearance of empty shells. These data suggest the presence of P22-related prophage in the genomes of these bacterial strains.
Figure C.2 Phage and procapsid-like particles purified from mitomycin-C induced cultures. Panel A are particles from *S. heidelberg*, Panel B are particles from *S. abortus-ovis*. 
REFERENCES


Spolar, R. S. & Record, M. T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. Science 263, 777-784.


