Published in final edited form as:

J Mol Biol. 2007 March 2; 366(4): 1161–1173.

Molecular Dissection of Ø29 Scaffolding Protein Function in an *in vitro* Assembly System

Chi-yu Fu¹, Marc C. Morais², Anthony J. Battisti², Michael G. Rossmann², and Peter E. Prevelige Jr.³

¹Dept. of Biochemistry & Molecular Genetics, Univ. of Alabama at Birmingham, Alabama 35294, USA

²Dept. of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47907, USA

³Dept. of Microbiology, Univ. of Alabama at Birmingham, Alabama 35294, USA

Summary

An *in vitro* assembly system was developed to study prolate capsid assembly of phage ø29 biochemically, and to identify regions of scaffolding protein required for its functions. The crowding agent polyethylene glycol can induce bacteriophage ø29 monomeric capsid protein and dimeric scaffolding protein to co-assemble to form particles which have the same geometry as either prolate T=3 Q=5 procapsids formed *in vivo* or previously observed isometric particles. The formation of particles is a scaffolding-dependent reaction. The balance between the fidelity and efficiency of assembly is controlled by the concentration of crowding agent and temperature. The assembly process is salt sensitive, suggesting that the interactions between the scaffolding and coat proteins are electrostatic.

Three N-terminal ø29 scaffolding protein deletion mutants, Δ 1-9, Δ 1-15 and Δ 1-22, abolish the assembly activity. Circular dichroism spectra indicate that these N-terminal deletions are accompanied by a loss of helicity. The inability of these proteins to dimerize suggests that the N-terminal region of the scaffolding protein contributes to the dimer interface and maintains the structural integrity of the dimeric protein.

Two C-terminal scaffolding protein deletion mutants, Δ 79-97 and Δ 62-97, also fail to promote assembly. However, the secondary structure and the dimerization ability of these mutants are unchanged relative to wild type which suggests that the C-terminus is the likely site of interaction with the capsid protein.

Keywords

bacteriophage ø29; scaffolding proteins; *in vitro* capsid assembly; prolate procapsid

Correspondence to: Peter E. Prevelige, Jr..

Reprint requests to: Peter E. Prevelige Jr., Department of Microbiology, University of Alabama at Birmingham, BBRB 416 845 19th St South, Birmingham, AL 35294-2170, USA; e-mail: prevelig@uab.edu; phone: (205) 975-5327; fax: (205) 975-5479..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

The most economical way to build mega-dalton viral capsid shells that are large enough to accommodate and protect the viral genome is to assemble these from a small number of building block types, using repeated bonding interactions¹. For some viruses or bacteriophages that have icosahedrally symmetric capsids, proper form-determination requires the participation of scaffolding proteins. Scaffolding proteins which bind to the interior (internal) and exterior (external) of the capsid have been described. Internal scaffolding proteins have been identified in a number of dsDNA bacteriophages including P22, λ , T7, P2/P4, T4, and ϕ 29. They are also found in herpes viruses and the ssDNA bacteriophage ϕ X174. Based on biochemical studies, internal scaffolding proteins have been proposed to assist in nucleation of assembly and direct the assembly pathway by stabilizing the capsid protein interactions at particular quasi-equivalent positions^{2, 3; 4}.

Internal scaffolding proteins generally have extended structures. For example P22 scaffolding protein has dimensions of 247 Å by 22 Å and that of phage λ scaffolding protein has been estimated to be 90 Å by 6.4 Å^{5; 6}. Structures with high α helical content have also been suggested in phages such as external scaffolding proteins of ϕ 174 and P4 and internal scaffolding protein of ϕ 29, P22 and gp23 of T4⁷.

The crystal structure of the 11 kDa scaffolding protein of the *Bacillus subtilis* phage ϕ 29 (gp7) has been solved⁸. It consists of three α -helices with a disordered C-terminus. The N-terminal helix-loop-helix region is structurally similar to the capsid binding domain of P22 scaffolding protein^{8; 9; 10}. Approximately 180 copies of scaffolding protein assemble with 235 copies of gp8 capsid protein (50 kDa), and 12 copies of gp10 connector (portal) protein (35 kDa) to form 43 nm $\times$ 53 nm prolate procapsid particles with a triangulation number of T = 3 and an elongation factor of Q = 5^{11; 12}. In infections with mutant phage that do not produce functional scaffolding protein the capsid protein forms spiral or aberrant structures that fail to incorporate the connector protein^{13; 14}. In the absence of connector protein, scaffolding and capsid proteins form mainly T = 3 isometric particles rather than prolate capsids *in vivo*¹³. This suggests that scaffolding protein is required for assembling T=3 capsids, whereas the connector protein is involved in controlling the formation of prolate capsids.

Here we reported the development of an *in vitro* assembly system of ϕ 29 procapsids. Purified recombinant scaffolding and capsid proteins were induced to polymerize in the presence of the crowding agent polyethylene glycol (PEG). The resultant assembled particles had sizes and morphologies similar to the *in vivo* assembled particles. We had utilized this system to map the functional domains of the scaffolding protein.

Results

Capsid Protein is Primarily Monomeric in Solution

To determine the oligomeric state of the capsid protein in solution, sedimentation velocity measurements were performed on recombinant capsid protein at a concentration of 1 mg/ml (20 μ M) in the buffer of 100 mM Tris, pH8, 10 mM MgCl₂ and 38 mM NaCl at 13°C in a Beckman XL-A analytical ultracentrifuge. The distribution of s values was determined by direct fitting of the Lamm equation. The predominant species had a sedimentation coefficient of 2.7 s and a calculated molecular weight of 53 kDa (Fig 1). This value is 6 % higher than the theoretical molecular weight of 50 kDa calculated from the sequence. Trace amounts of the protein were broadly distributed at around 5.1 s. These results suggest that in the concentration range used for the *in vitro* assembly experiments described below the capsid protein was primarily monomeric before inducing assembly with polyethylene glycol.

Polyethylene Glycol Induces *in vitro* Assembly of Capsid and Scaffolding Proteins

The first step to screen for *in vitro* conditions under which the capsid protein can assemble into procapsid-like particles was to express and purify recombinant capsid and scaffolding proteins separately. The capsid protein was then concentrated to 4 mg/ml (80 μ M) and diluted four fold into various assembly buffers in the presence of two-fold molar excess of scaffolding protein. Electrostatic interactions have previously been shown to be important in modulating capsid/scaffolding protein interactions in other phage assembly systems such as in phage P22, P4 and T7^{15 16 17}. Therefore, the NaCl concentration was maintained at 38 mM to minimize disruption due to charge/charge shielding. The kinetics of polymerization was followed by measuring the turbidity at 340 nm.

In buffer alone, neither capsid nor scaffolding proteins alone, nor the mixture of scaffolding and capsid protein formed structures large enough to detectably scatter light (Fig 2a). Various crowding agents, such as sucrose, dextran, and polyethylene glycol (PEG) have previously been used to facilitate protein polymerization by increasing the effective protein concentration¹⁸. Neither glycerol, sucrose, nor dextran, at a concentration of 10%, had any effect on assembly. While PEG did not promote scaffolding protein polymerization, PEG promoted capsid protein polymerization (Fig 2a). In the absence of scaffolding protein, PEG induced capsid self-association resulted in the formation of predominately spiral shaped aberrant structures as detected by electron microscopy (data not shown). To optimize the conditions that suppress the capsid self-polymerization and favor scaffolding promoted capsid assembly, the PEG-3350 concentration and temperature were adjusted. Lowering the temperature suppressed capsid protein self-polymerization and favored scaffolding promoted assembly. However, the yield was reduced. Higher PEG concentration increased the yield but also induced aberrant capsid self-polymerization. These two factors were adjusted to strike the optimal balance. The optimal conditions determined to be 20 μ M capsid protein, 40 μ M scaffolding protein at 7% PEG-3350 and 13°C. Under these conditions the increase of turbidity in the presence of scaffolding protein was about 2 fold higher than that of the no scaffolding background (Fig 2a)

Procapsid-Like Particles are Formed *in Vitro*

To analyze the products formed under the conditions described above, the samples were separated by centrifugation through a 10-40 % sucrose gradient and the protein composition in the fractions was analyzed by SDS-PAGE. In the absence of scaffolding protein, the majority of capsid protein did not polymerize and remained at the top of the gradient (Fig 2b). High molecular weight species that sedimented to the bottom of the gradient comprised approximately 1% of the total protein. However, when the reaction mixtures were rapidly pelleted to separate large structures, approximately 16 % of the capsid protein was pelleted (data not shown). This suggests that a significant fraction of the large structures were unstable and dissociated during the gradient separation, likely due to the decrease in protein concentration accompanying gradient centrifugation.

In the presence of scaffolding protein, approximately 45 % of the capsid protein assembled into structures that sedimented to a position about 1/3 from the top of the gradient (fractions 7-9) which are the same fractions where procapsid particles purified from a *Bacillus subtilis* infection are found (Fig 2b). These particles had a capsid to scaffolding protein ratio of 2.8/1 as determined by the relative intensity of bands on the SDS gel of separated fractions. This compares favorably to the 2.4/1 ratio observed for procapsids produced *in vivo* (Fig 2b). Approximately 33% of the capsid protein remained unassembled.

To examine the effect of scaffolding protein dosage on particle formation, assembly was performed at molar ratios of scaffolding to capsid protein ranging from 2/1 to 0.125/1 with the

capsid protein concentration fixed at 1 mg/ml. The assembly products were analyzed by 1% native agarose gel which separated the assembled particles from unassembled capsid and scaffolding protein subunits (Fig 2c). The yield of assembled particle increased with increasing scaffolding protein and was accompanied by a decrease in unassembled capsid protein. The broader distribution of the *in vitro* assembled particles as compared to those formed *in vivo* might reflect structural heterogeneity in the *in vitro* assembled particles.

***In vitro* Assembled Particles Have Correct Sizes and Morphologies**

Cryo-EM and image analysis were used to characterize the particles formed *in vitro*. Visual examination of cryo-EM micrographs indicated a morphologically diverse population of particles (Fig 3a). Although the majority of particles appeared to be prolate, there were a significant number of smaller spherical particles. Additionally, approximately 20% of the particles appeared to be irregularly shaped, or broken, and were excluded from subsequent structural analyses. One of the questions that can be addressed by cryo-EM image reconstruction analysis is whether *in vitro* assembled particles resemble any of the morphologies observed for *in vivo* assembled particles. In many dsDNA phages maturation from procapsid to capsid results in a clear difference in morphology, however this is not the case for $\phi 29$ where the procapsid and capsid forms appear indistinguishable^{19; 20; 21}. Three distinct morphologies of $\phi 29$ particles have been observed by cryo-EM image reconstruction analysis. Wild-type $\phi 29$ particles are prolate icosahedra which can be described by the triangulation numbers $T = 3$, $Q = 5$ ^{11; 12}. Amber mutations in the $\phi 29$ scaffolding protein (S65N) result in isometric icosahedra having predominantly $T = 3$ geometry, with a small minority of these particles having recently been shown to have $T = 4$ geometry^{11; 12; 22}. In order to sort particles based on overall morphology and orientation, individual particle images from the *in vitro* assembly reaction were compared to images of prolate, isometric $T = 3$, and isometric $T = 4$ particles projected in different orientations calculated to uniformly sample possible orientations within the asymmetric unit for each type of particle. Approximately 70 % of the ~2000 particles examined were classified as prolate, 20 % $T = 4$ isometric, and 10 % $T = 3$ isometric. Three different three-dimensional reconstructions were then calculated from the resulting class averages (see Methods, Fig 3a, 3b). Although the three reconstructions have the same approximate shape and size as their corresponding starting models, clearly defined capsomers are difficult to discern. In order to reduce noise and emphasize the most basic structural features of the maps, both the starting models and final reconstructions were low-pass filtered such that the amplitudes of terms at spatial frequencies greater than 45 Å⁻¹ in the Fourier synthesis were attenuated. If the filtered maps are then contoured so that only pixels whose density values are at least 3.5 standard deviations greater than the mean are visible, it is apparent that the prolate particles have $T = 3$, $Q = 5$ geometry and that the smaller of the two isometric particles has a triangulation number of $T = 3$ (Fig 3c). Other than being present in the starting models, the quasi-symmetries associated with particular T and Q numbers were not imposed at any stage of the reconstruction procedure. Thus, these local symmetries likely reflect the true geometry of the particles. The reconstruction of the larger of the isometric particles however, is especially featureless, and does not possess the $T = 4$ geometry present in the starting model for this class of particles. Thus, the projection database calculated from the $T = 4$ isometric $\phi 29$ particle likely served as a “trap” for aberrantly formed particles of intermediate size.

Model bias is always a concern in reconstruction procedures utilizing model-based reference projections. However, the absence of $T = 4$ geometry in the reconstruction calculated from the larger of the isometric particles indicates that the reconstructions resulting from the procedure outlined above are not overly model-biased. Furthermore, the absence of pRNA in the prolate reconstruction also suggests that the reconstruction procedure utilized here is able to overcome model bias as the pRNA is present in the starting model but not the final model, consistent

with the absence of pRNA from the *in vitro* assembly reaction. Although distorted or defective particles were likely included in the calculation of the reconstructions, introducing errors in the resulting electron density maps and a subsequent loss of resolution, the data suggest that the capsid and scaffolding protein assembled *in vitro* predominantly form particles with either T=3, Q=5 prolate or T=3 isometric geometry.

Assembly is Controlled in a Narrow Range of PEG Concentration

In the $\phi 29$ *in vitro* assembly system, the crowding agent PEG is required to promote interactions between components and induce assembly. To examine the effect of PEG on the kinetics of assembly and the distribution of products, assembly reactions at 7%, 8%, and 9% PEG-3350 were compared using fixed (20 μ M) capsid protein concentration and various ratios of scaffolding protein. The initial rates of assembly as measured by the increase in turbidity were 3-6 fold faster at 9% PEG-3350 compared to 7% PEG-3350 (Fig 4a), with 8% being intermediate. The effect of PEG-3350 concentration on the fraction of capsid protein partitioning into uncontrolled assemblies was monitored by determining the fraction of large aberrant structures that pelleted upon low speed centrifugation using SDS-PAGE analysis (Fig 4b). At 9% PEG-3350, capsid self-association was significantly enhanced relative to 7% PEG-3350 as evidenced by the fact that approximately 70% of the capsid protein formed pelletable aggregates in 9% PEG-3350 solutions versus approximately 15% in 7% PEG-3350. The presence of scaffolding protein, even at a two-fold molar excess, had little effect on the amount of aggregates produced. When these assembly reactions were analyzed on a sucrose gradient, the aggregates were found to be polydisperse and had scaffolding protein associated with them (data not shown). Thus, while scaffolding protein can interact with capsid protein under these conditions, it is not able to redirect assembly towards particle formation. The data suggest that the aggregates became dominant at the faster assembly rates induced by high PEG concentration. Similar results were observed in parallel studies using PEG-8000 although the concentrations required to promote controlled and uncontrolled assembly were lowered (data not shown). The enhanced capsid self-association and resulting off-pathway assembly are likely due to the incorrect incorporation of capsid subunits during the lattice growth.

Sodium Chloride Reduces Particle Formation

To examine the chemical nature of interactions participating in particle assembly, the NaCl concentration in the reactions was varied from 18 mM to 160 mM and the yield of procapsid like particles was analyzed by native agarose gel electrophoresis and densitometry. The effect of NaCl on the relative particle yield is shown in Fig 5. At 160 mM NaCl, the yield of particles decreased to 16% of the low salt value. The suppression of particle formation can be due to the disruption of capsid /scaffolding, capsid/capsid or scaffolding/scaffolding interactions. However at 160 mM NaCl, capsid self-association remained at 80% of the low salt value as measured by turbidity (Fig 5). This suggests that the reduction of particle formation at high salt more likely results from the disruption of electrostatic interactions between scaffolding-capsid rather than capsid-capsid interactions.

Scaffolding Deletion Mutants Abolish Activity

The N-terminal helix-loop-helix of $\phi 29$ scaffolding protein which is structurally similar to the capsid binding domain of phage P22 scaffolding protein might be involved in capsid protein binding interactions^{8; 9; 10}. Hydrogen-deuterium exchange has shown that the scaffolding protein residues 20-31 have faster exchange kinetics in the procapsid than in the free form suggesting that it may interact with the capsid.²³

To determine the regions of scaffolding protein that are important in assembly, three deletion mutants Δ 1-9, Δ 1-15 and Δ 1-22 were constructed. For all three N-terminal deletions the rate of assembly was similar to the background rate of assembly seen in the capsid only assembly

reactions. Very few assembled particles were detectable by SDS-PAGE after sucrose gradient fractionation (Fig 6). Densitometry of the stained gels suggested that the amount of assembled particles was decreased approximately 25-fold. Consequently, closed shells were rarely seen in the negatively stained electron micrographs.

To elucidate the mechanism of how the deletion mutations in the scaffolding protein abolish its function, the ability of the mutant proteins to dimerize was examined by equilibrium ultracentrifugation. The data for wild type scaffolding protein obtained under assembly conditions was well fit as a single species with a molecular weight of 21.8 kDa. This value represents a 2% deviation from the molecular weight expected for a scaffolding protein dimer based on the amino acid sequence. (Fig 7a). This data suggests that under assembly conditions the scaffolding protein functions as a dimer. Surprisingly, the mutant in which the nine N-terminal residues were deleted did not form dimers as the data were best fit as a single species with the molecular weight of 10.0 kDa, which represents a 2% deviation from the expected molecular weight of the monomer (10.2 kDa) (Fig 7b). Deletion of the N-terminal nine residues also resulted in unfolding of the protein as indicated by the nearly complete loss of helical structure as monitored by circular dichroism spectroscopy (Fig 8).

To evaluate the role of the C-terminus of the scaffolding protein in assembly, two different truncations were constructed. In the first of these mutants the disordered C-terminal, residues 79-97 were deleted (Δ 79-97 mutant), and in the second mutant residues 62-97, which comprises the last three helical turns of helix 3, were deleted. Both C-terminal deletions abolished the assembly activity (Fig 6). In contrast to the N-terminal deletions, these proteins maintained a high percentage of α -helix content (Fig 8). Additionally, these C-terminal deletion mutants still form tight dimers in solution as the sedimentation equilibrium data was best fit as a single species with a molecular weight within 3% of that predicted for a dimer (Fig 7c).

Discussion

The ϕ 29 scaffolding protein is a relatively small 11 kDa molecule compared to other scaffolding proteins, such as 33 kDa of P22 (gp8), 34 kDa of HSV-1 (VP22*), and 32 kDa of P2/P4 (gpO). No core of scaffolding protein was detected that might form a template surface for capsid protein polymerization. Therefore, as is the case for P22, the assembly active species of ϕ 29 scaffolding protein is probably dimers or small oligomers. Electrostatic interactions play a key role in the interaction between scaffolding and capsid proteins of ϕ 29 as shown by the salt sensitivity of assembly - a property seen in other phage scaffolding proteins including P22, P4 and T7^{15 16 17}. These interactions seem to be weak as no binding between the capsid and scaffolding proteins was detected using an affinity column or in fluorescence polarization experiments (data not shown). Weak interactions are thought to be important to allow transient association and dissociation during the assembly reaction and the subsequent scaffolding release²⁴.

The two C-terminal deletion mutants, Δ 79-97 and Δ 62-97, have lost assembly activity, although their secondary structure and the dimerization capacity remained unchanged relative to wild type scaffolding protein suggesting that the effects of deletion were local. It therefore appears likely that the C-terminal region interacts with capsid protein directly. However, it remains possible that this region is involved in scaffolding/scaffolding interactions to form oligomers larger than dimers. While higher oligomers have not been observed in solution, dimers of dimers interacting through the C-terminal region have been observed in the crystal structure of the scaffolding protein. These tetramers have been fitted into the cryo-EM density reconstruction of ϕ 29 procapsids⁸.

The deletion of the first nine residues perturbs the secondary structure as shown by the loss of helicity in the CD spectra and results in dimer dissociation. The complete loss of helicity indicates that not just the helix-loop-helix region has been perturbed but also the extended helix 3 which forms a long coiled-coil motif and appears to stabilize the dimer in the crystal structure. Thus the N-terminal region of the scaffolding molecule appears to be important in maintaining the structural integrity even though previous hydrogen/deuterium exchange studies have demonstrated that it is highly dynamic and undergoes continual unfolding and refolding in solution²³. The $\phi 29$ scaffolding protein exits intact from the procapsid during DNA packaging²⁵. However, the diameter of the hole in the center of the hexamers and pentamers of the procapsid is approximately 14 Å¹². The folded scaffolding dimer has a maximum width of 30 Å in the N-terminal region and 23 Å in the C-terminal coil-coil region⁸. A trigger that shifts N-terminal region toward the unfolded state and thus leads to unfolding of the whole molecule could be a possible mechanism for scaffolding protein release and escape during DNA packaging.

Based on the above observations of the N and C terminal properties of the scaffolding protein, it appears that the entire scaffolding molecule is required for procapsid assembly.

The thermodynamics and kinetics aspects of assembling icosahedral structures have been studied in a number of *in vitro* assembly systems and also by computer simulations^{26; 27; 28; 29}. It has been demonstrated that the assembly of icosahedral capsids, like the assembly of helical capsids, consists of a slow nucleation phase and a rapid growth phase³⁰. To initiate assembly, the protein concentration needs to exceed a critical concentration at which stable nuclei can form. Once formed, the nuclei then proceed to the lattice growth phase. In the $\phi 29$ system, the crowding agent PEG is required to increase the effective concentration of the components and induce productive collisions and drive the reaction. In the case of $\phi 29$, increasing the PEG concentration and the rate of assembly leads to the formation of aberrant particles indicating the fidelity of assembly varies with the rate. A link between assembly rate and malformation has been modeled in computer simulations of assembly^{26; 31; 32}. In these simulations the underlying cause of malformation is either the formation of an excess of nuclei leading to kinetic trapping, or subunit addition before proper conformational switching can be achieved leading to the incorrect geometrical incorporation of capsid subunits and the formation of aberrant particles. In the case of $\phi 29$, it was possible to empirically vary the assembly conditions and strike a balance between fidelity and efficiency of assembly

In vitro the monomeric capsid and dimeric scaffolding proteins (in the absence of connector protein) form 70% T=3, Q=5 prolate particles with 10% T=3 isometric particles. A small fraction of the particles have an extended long axis, but no grossly elongated particles such as phage T4 tubular polyheads were seen³³. The tendency to form particles with high polymorphism seems to be the nature of capsid protein, which is designed to be flexible enough to accommodate different subunit bonding in the lattice. In various *in vitro* assembly systems of phages and viruses, capsid proteins can assemble to form either spherical particles with various triangulation numbers or spiral or tubular structures as a function of chemical conditions such as temperature, pH, ionic strength, or stabilizing divalent metal ion etc³⁴. The data presented here shows that $\phi 29$ capsid protein itself forms mainly aberrant structures on its own and therefore doesn't encode sufficient information for form determination. The scaffolding protein promotes capsid assembly to form closed shells with defined morphologies. The connector is likely not required to build T=3, Q=5 particles but rather it shifts the thermodynamic trajectory toward prolate particle formation and enforces the stringent control of assembly fidelity *in vivo*.

Materials and Methods

Protein Preparation

Recombinant capsid protein was subcloned from plasmid pARgp 7-8-8.5-10 with NdeI and BamHI cloning sites to pET 3c vector³⁵. The forward and reverse primers were (5'-gggaattccatatgcgaattacatttaacgacgtgaaaacg-3') and (5'-cgcggtatccatcattacgcctgagcacc-3') respectively. The plasmid was transformed to *Escherichia coli* BL21 (DE3) pLysS for protein expression. The capsid protein were purified from inclusion bodies and refolded at 0.5 mg/ml in 0.7 M Arginine at pH 8. The refolded proteins were dialyzed to 100 mM Tris, pH 8, 10 mM MgCl₂ and 50 mM NaCl. Capsid proteins were bound and eluted from HiTrap SP HP column at around 150 mM NaCl in the same buffer. The purified capsid protein was dialyzed and stored in 100 mM Tris, pH 8, 10 mM MgCl₂ and 150 mM NaCl.

Recombinant scaffolding proteins were expressed in *Escherichia coli* BL21 (DE3) pLysS cells harboring plasmid pARgp7³⁶. The proteins were purified with HiTrap Q HP column using a NaCl gradient in 50 mM Tris-HCl pH 8, 1 mM EDTA. The scaffolding proteins eluted at 210 mM NaCl and were subsequently dialyzed against the same buffer at 50 mM NaCl. Further purification was carried out with Superdex G-75 size exclusion chromatography. The purified scaffolding proteins were stored in 50 mM Tris-HCl pH 8, 50 mM NaCl and 1 mM EDTA.

The C-terminal truncated mutants Δ 62-97 and Δ 79-97 were constructed by introducing a stop codon to replace residue 62 and 79 respectively by QuikChange Site-Directed Mutagenesis kit with primers (5'-gccgctgaaaagatgatctgatcgtgtgtaaatagtaagc-3') for Δ 62-97 and (5'-gacagacaaacaggaataagatcacaagaagctgatattagtg-3') for Δ 79-97. The proteins were expressed and purified as described. The N-terminal truncated mutants, Δ 1-9, Δ 1-15 and Δ 1-22, were constructed by subcloning desired region framed by NdeI and NheI restriction sites and inserting back to pET 3c vector. The NdeI site provides the start codon for translation. The tailed primers with the NdeI site right in front of the 1st nucleotide encoded for residue 10, 16 and 23 are (5'-gagatatacatatggatatttgaataagcttttagacc-3'), (5'-gagatatacatatgttagaccctgagttggctcaatc-3'), and (5'-gagatatacatatgtcagaagaacggaagctcttc-3') respectively. The reverse primer carried NheI site (5'-ccaccagtcagctagcggcacataac-3'). The proteins were expressed and purified as described. The mass of purified proteins were confirmed using ESI-TOF mass spectrometry (Micromass LCT).

Procapsid particles assembled *in vivo* were produced from infecting non-suppressor *Bacillus subtilis* strain RD2 with the ϕ 29 mutant phage *sus* 8.5(900) *sus* 16(300) *sus* 14(1241) and purified on a 10-40 % (w/v) sucrose gradient in TMS buffer (50 mM Tris-HCl pH 7.8, 10 mM MgCl₂, and 100 mM NaCl) at 4°C using a Beckman SW55 at the rotor speed of 45,000 rpm for 55 minutes. The fractions containing procapsids were pelleted and resuspended in TMS buffer as previously described¹¹.

Analytical Ultracentrifugation

Sedimentation velocity experiments on capsid protein were performed at 13°C with the protein concentration of 1 mg/ml (20 μ M) in the buffer of 100 mM Tris, pH8, 10 mM MgCl₂ and 38 mM NaCl. The sedimentation at 50,000 rpm with An-60 Ti rotor was followed by the absorbance at 280 nm over time with Beckman Optima XL-A analytical ultracentrifuge. The partial specific volume was calculated by amino acid sequence and the solvent density was calculated according to the solvent composition using the Sednterp program. The sedimentation coefficient distributions were calculated with a continuous c(s) distribution model of Lamm equation solution implemented in SEDFIT program³⁷. The diffusion coefficients were estimated by a weight-average frictional ratio. Svedberg equation relates

these two factors to molecular weight. The molecular weight distributions were derived with a continuous $c(M)$ distribution model in SEDFIT program³⁸.

Sedimentation equilibrium experiments on wild type and truncated scaffolding proteins were done at 50 μ M and 25 μ M loading concentration in the buffer of 100 mM Tris, pH8, 10 mM $MgCl_2$ and 38 mM NaCl at 13°C. The data were collected with Beckman Optima XL-A analytical ultracentrifuge at OD 238 nm at rotor speeds of 33,000 rpm, 39,000 rpm and 47,000 rpm with An-60 Ti rotor. The SedAnal 4.02 program³⁹ was used to fit the data globally to obtain the molecular weight and dissociation constant.

***In vitro* Assembly Reactions**

Assembly reactions were carried out at 20 μ M (1 mg/ml) capsid protein and 10 μ M scaffolding protein. 360 μ l of 10.2% PEG-3350 in 100 mM Tris, pH 8, 10 mM $MgCl_2$ was added to the cuvette and allowed to temperature equilibrate to 13°C. 45 μ l of scaffolding protein and 120 μ l of capsid protein were mixed subsequently to give a final buffer composition of 7% PEG-3350, 100 mM Tris, pH 8, 10 mM $MgCl_2$ and 38 mM NaCl. The assembly kinetics was monitored by recording the turbidity at 340 nm at 40 second intervals for 2 hours using the Beckman DU640 spectrometer. The dead-time between mixing and the start of data acquisition was 40 seconds.

200 μ l aliquots were overlaid on a 10-40 % sucrose gradient with 7% PEG-3350. A 50 μ l cushion of 60 % CsCl in 50 % sucrose were laid at the bottom. The samples were centrifuged at 4°C at 45,000 rpm for 1.5 hours in a Beckman SW-55Ti rotor. The gradient were fractionated from the bottom and analyzed by SDS-PAGE. The intensity of the coomassie blue stained protein bands was quantified using a Bio-Rad gel documentation system. The stoichiometry of the capsid to scaffolding protein in the procapsids or particles assembled *in vitro* was quantified by the relative protein intensity taking into consideration the relative molecular weight of capsid protein (50 kDa) and scaffolding protein (11 kDa) to calculate the molar ratio.

The native gel electrophoresis was performed at 1% agarose in TBM buffer (45 mM Tris, 45 mM borate, 1 mM $MgCl_2$) at 4 V/cm for three to four hours at 4°C.

Electron Microscopy

In vitro assembled particles were purified by sucrose gradient centrifugation and dialyzed out of PEG and sucrose. Purified particles were flash frozen on holey grids in liquid ethane. Images were recorded at 38,000x magnification using a Philips CM200 FEG microscope, with electron dose levels of approximately $15 \text{ e}^-/\text{\AA}^2$. Suitable micrographs of assembled particles were digitized at 14 μ m intervals ($3.68 \text{ \AA pixel}^{-1}$) with a Zeiss SCAI scanner.

Three-dimensional image reconstruction

Individual particle images were boxed, floated, and preprocessed to normalize mean intensities and variances and to remove linear background gradients. Contrast transfer function parameters were determined with the program EMAN⁴⁰. Visual inspection of several particles indicated morphological heterogeneity (Fig 3a). Although the majority of particles appeared similar in size and shape to prolate wild-type $\phi 29$ particles, many particles appeared to be approximately spherical with varying radii. Several particles also appeared to be irregularly shaped, or broken, and were subsequently excluded from structural analysis. In order to classify particles based on size and morphology, the following procedure was used.

1. Projection data for prolate procapsid, $T = 3$ isometric and $T = 4$ isometric $\phi 29$ particles were calculated in 3.2° angular increments to uniformly sample possible orientations in the asymmetric unit of each particle. Each calculated projection was the seeding

image for that class. The cryo-EM reconstruction of these three models has been previously reported^{11; 12; 22}.

2. Each *in vitro* assembled particle as observed in cryo-EM was compared to the calculated projections of each data base. Individual particles were assigned a class based on how well they match a projection of a particular model in a particular orientation using optimized real-space variance as a similarity measure⁴⁰.
3. Members of each class were averaged to create class averages. Class averages corresponding to each of the three models were then used to calculate three different three-dimensional reconstructions. Icosahedral symmetry was imposed for the two isometric reconstructions, and 5-2 symmetry imposed on the prolate reconstruction. Both phases and amplitudes were corrected according to the parameters defining the contrast transfer function. The program EMAN⁴⁰ was used to implement each step of the procedure outlined above.
4. The resulting reconstructions were then used to begin a new cycle of particle classification and orientation (steps 1-3 above). A total of eight cycles were carried out, though convergence, as indicated by Fourier shell correlation between reconstructions from adjacent cycles⁴⁰, occurred after only five cycles.

Circular Dichroism Spectroscopy

Circular Dichroism (CD) spectra of wild type and truncated scaffolding proteins were obtained on an AVIV model 62DS spectropolarimeter (Lakewood, NJ) equipped with a single cell thermoelectric device. Data were collected on 1 mg/ml protein solutions at 13°C, from 250 nm to 190 nm in 1 nm increments, with a 16 sec averaging time per point. The spectrum for baseline correction was identical to the sample solutions less protein. Raw data, in millidegrees, were baseline corrected, smoothed, and multiplied by a scaling factor to obtain spectra in units of mean residue ellipticity. Secondary structure calculations were performed with PROSEC.

Acknowledgements

This work was supported by National Institutes of Health, GM47980 (P.E.P) and NSF, MCB0443899 (M.G.R). We thank Dr Paul Jardine for ø29 clones and helpful discussions.

Abbreviations

cryo-EM, cryo-electro microscopy; CD spectroscopy, circular dichroism spectroscopy; PEG, polyethylene glycol.

References

1. Caspar DL, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol 1962;27:1–24. [PubMed: 14019094]
2. Prevelige PE Jr, Thomas D, King J. Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro. J Mol Biol 1988;202:743–57. [PubMed: 3262767]
3. Prevelige PE Jr, Thomas D, King J. Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. Biophys J 1993;64:824–35. [PubMed: 8471727]
4. Zhou ZH, Macnab SJ, Jakana J, Scott LR, Chiu W, Rixon FJ. Identification of the sites of interaction between the scaffold and outer shell in herpes simplex virus-1 capsids by difference electron imaging. Proc Natl Acad Sci U S A 1998;95:2778–83. [PubMed: 9501166]
5. Fane BA, Prevelige PE Jr. Mechanism of scaffolding-assisted viral assembly. Adv Protein Chem 2003;64:259–99. [PubMed: 13677050]

6. Ziegelhoffer T, Yau P, Chandrasekhar GN, Kochan J, Georgopoulos C, Murialdo H. The purification and properties of the scaffolding protein of bacteriophage lambda. *J Biol Chem* 1992;267:455–61. [PubMed: 1530932]
7. Dokland T. Scaffolding proteins and their role in viral assembly. *Cell Mol Life Sci* 1999;56:580–603. [PubMed: 11212308]
8. Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BA, McMurray CT, Anderson DL, Rossmann MG. Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly. *Nat Struct Biol* 2003;10:572–6. [PubMed: 12778115]
9. Parker MH, Casjens S, Prevelige PE Jr. Functional domains of bacteriophage P22 scaffolding protein. *J Mol Biol* 1998;281:69–79. [PubMed: 9680476]
10. Sun Y, Parker MH, Weigele P, Casjens S, Prevelige PE Jr, Krishna NR. Structure of the coat protein-binding domain of the scaffolding protein from a double-stranded DNA virus. *J Mol Biol* 2000;297:1195–202. [PubMed: 10764583]
11. Tao Y, Olson NH, Xu W, Anderson DL, Rossmann MG, Baker TS. Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* 1998;95:431–7. [PubMed: 9814712]
12. Morais MC, Choi KH, Koti JS, Chipman PR, Anderson DL, Rossmann MG. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol Cell* 2005;18:149–59. [PubMed: 15837419]
13. Hagen EW, Reilly BE, Tosi ME, Anderson DL. Analysis of gene function of bacteriophage phi 29 of *Bacillus subtilis*: identification of cistrons essential for viral assembly. *J Virol* 1976;19:501–17. [PubMed: 822175]
14. Camacho A, Jimenez F, De La Torre J, Carrascosa JL, Mellado RP, Vasquez C, Vinuela E, Salas M. Assembly of *Bacillus subtilis* phage phi29. 1. Mutants in the cistrons coding for the structural proteins. *Eur J Biochem* 1977;73:39–55. [PubMed: 402269]
15. Parker MH, Prevelige PE Jr. Electrostatic interactions drive scaffolding/coat protein binding and procapsid maturation in bacteriophage P22. *Virology* 1998;250:337–49. [PubMed: 9792844]
16. Cerritelli ME, Studier FW. Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. *J Mol Biol* 1996;258:286–98. [PubMed: 8627626]
17. Wang S, Palasingam P, Nokling RH, Lindqvist BH, Dokland T. In vitro assembly of bacteriophage P4 procapsids from purified capsid and scaffolding proteins. *Virology* 2000;275:133–44. [PubMed: 11017795]
18. Ellis RJ. Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* 2001;26:597–604. [PubMed: 11590012]
19. Lata R, Conway JF, Cheng N, Duda RL, Hendrix RW, Wikoff WR, Johnson JE, Tsuruta H, Steven AC. Maturation dynamics of a viral capsid: visualization of transitional intermediate states. *Cell* 2000;100:253–63. [PubMed: 10660048]
20. Zhang Z, Greene B, Thuman-Commike PA, Jakana J, Prevelige PE Jr, King J, Chiu W. Visualization of the maturation transition in bacteriophage P22 by electron cryomicroscopy. *J Mol Biol* 2000;297:615–26. [PubMed: 10731416]
21. Xiang Y, Morais MC, Battisti AJ, Grimes S, Jardine PJ, Anderson DL, Rossmann MG. Structural changes of bacteriophage phi29 upon DNA packaging and release. *Embo J* 2006;25:5229–39. [PubMed: 17053784]
22. Choi KH, Morais MC, Anderson DL, Rossmann MG. Determinants of Bacteriophage varphi29 Head Morphology. *Structure* 2006;14:1723–7. [PubMed: 17098197]
23. Fu CY, Prevelige PE Jr. Dynamic motions of free and bound O29 scaffolding protein identified by hydrogen deuterium exchange mass spectrometry. *Protein Sci* 2006;15:731–43. [PubMed: 16522798]
24. Zlotnick A. Are weak protein-protein interactions the general rule in capsid assembly? *Virology* 2003;315:269–74. [PubMed: 14585329]
25. Nelson RA, Reilly BE, Anderson DL. Morphogenesis of bacteriophage phi 29 of *Bacillus subtilis*: preliminary isolation and characterization of intermediate particles of the assembly pathway. *J Virol* 1976;19:518–32. [PubMed: 822176]
26. Schwartz R, Shor PW, Prevelige PE Jr, Berger B. Local rules simulation of the kinetics of virus capsid self-assembly. *Biophys J* 1998;75:2626–36. [PubMed: 9826587]

27. Reddy VS, Giesing HA, Morton RT, Kumar A, Post CB, Brooks CL 3rd, Johnson JE. Energetics of quasiequivalence: computational analysis of protein-protein interactions in icosahedral viruses. *Biophys J* 1998;74:546–58. [PubMed: 9449355]
28. Zlotnick A. Theoretical aspects of virus capsid assembly. *J Mol Recognit* 2005;18:479–90. [PubMed: 16193532]
29. Endres D, Zlotnick A. Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys J* 2002;83:1217–30. [PubMed: 12124301]
30. Prevelige PE Jr, King J. Assembly of bacteriophage P22: a model for ds-DNA virus assembly. *Prog Med Virol* 1993;40:206–21. [PubMed: 8438077]
31. Zlotnick A. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J Mol Biol* 1994;241:59–67. [PubMed: 8051707]
32. Zhang T, Schwartz R. Simulation study of the contribution of oligomer/oligomer binding to capsid assembly kinetics. *Biophys J* 2006;90:57–64. [PubMed: 16214864]
33. Cummings DJ, Bolin RW. Head length control in T4 bacteriophage morphogenesis: effect of canavanine on assembly. *Bacteriol Rev* 1976;40:314–59. [PubMed: 786254]
34. Kanesashi SN, Ishizu K, Kawano MA, Han SI, Tomita S, Watanabe H, Kataoka K, Handa H. Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro. *J Gen Virol* 2003;84:1899–905. [PubMed: 12810885]
35. Guo PX, Erickson S, Xu W, Olson N, Baker TS, Anderson D. Regulation of the phage phi 29 prohead shape and size by the portal vertex. *Virology* 1991;183:366–73. [PubMed: 1905079]
36. Lee CS, Guo P. Sequential interactions of structural proteins in phage phi 29 procapsid assembly. *J Virol* 1995;69:5024–32. [PubMed: 7609072]
37. Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 2002;11:2067–79. [PubMed: 12192063]
38. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 2000;78:1606–19. [PubMed: 10692345]
39. Stafford WF, Sherwood PJ. Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. *Biophys Chem* 2004;108:231–43. [PubMed: 15043932]
40. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 1999;128:82–97. [PubMed: 10600563]

Fig 1.

Sedimentation Velocity Analysis of Ø29

Capsid Protein Ø29 capsid protein at an initial concentration of 0.5 mg/ml was sedimented at 13°C and 50,000 rpm in a Beckman Optima XL-A analytical ultracentrifuge. The sedimentation coefficient distribution was calculated with a continuous $c(s)$ distribution model with the program SEDFIT.

Capsid

M 1 2 3 4 5 6 7 8 9 10 11 12 13

Capsid plus Scaffolding

M 1 2 3 4 5 6 7 8 9 10 11 12 13

Bottom

Top

Scaffolding to Capsid Molar Ratio

Fig 2. Scaffolding-Promoted Capsid Assembly *in vitro*

(a) The kinetics of assembly were followed by turbidity at 340 nm. The composition of the assembly reactions were as follows: 40 μ M scaffolding protein, 20 μ M capsid protein in the absence of PEG (crosses); 40 μ M scaffolding protein, 7% PEG (closed triangles); 20 μ M capsid protein, 7% PEG (open squares); 40 μ M scaffolding protein, 20 μ M capsid protein, 7% PEG (closed squares). The arrow indicates the time of mixing the components.

(b) Sucrose gradient fractionation of assembly reactions.

The assembly reactions of capsid protein (top) and of capsid plus scaffolding proteins (bottom) in 7% PEG were centrifuged through a 10-40% sucrose gradient containing 7% PEG, fractionated from the bottom (fraction 1) and analyzed by SDS-PAGE. Procapsids purified *in vivo* were run as a marker (M) in the left most lane.

(c) Native agarose gel analysis of assembly reactions.

10 μ l aliquots of assembly reactions containing 20 μ M capsid protein and scaffolding protein at the indicated molar ratios were electrophoresed towards the cathode, and visualized by Coomassie blue staining. Procapsids purified *in vivo* were run as a marker (M) in the left most lane.

Fig 3.
Class Sorting and 3-D Cryo-EM Reconstructions of *in vitro* Assembled Particles

- a. Cryo-EM micrograph of a typical field of *in vitro* assembled particles.
- b. Three-dimensional starting models used for initial particle classification are shown in purple on the left. These models come from previously published reconstructions of $\phi 29$ particles observed *in vivo* (see text). From top to bottom, these particles are T=3, Q=5 prolate proheads, T=3 isometric particles, and T=4 isometric particles. To the right of each starting model, two projections of the particle are shown (purple). These projections are examples of references used for the first cycle of classification and alignment. Examples of seven *in vitro* assembled particles assigned to that particular class (grey) follow each projection.

- c. Cryo-EM reconstructions of *in vitro* assembled particles (bottom, purple) after eight cycles of the reconstruction procedure are compared to initial starting models (top, green) used to begin the classification procedure. Triangulation nets for the prolate particles are drawn as black lines, with the positions of hexamers and pentamers shown in red. Icosahedral asymmetric units are shown in black for the isometric particles, with the position of 5-, 3-, and 2-fold axes shown as pentagons, triangles, and ovals, respectively. All maps were low-pass filtered at 45 Å, and contoured at 3.5 σ . Note that the contrast of the particles in (b) is reversed relative to those in (a) in order to follow the convention adopted by the image processing software.

Fig 4.
The Effect of PEG Concentration on Assembly
(a) The initial rate of turbidity increase in the presence of 7% (open squares) and 9% PEG (close squares) at the indicated scaffolding to capsid protein ratios. (b) The fraction of aggregates in the reactions with 7% and 9% PEG (open and slashed bars respectively)

Fig 5.

The Effect of NaCl on Assembly

The relative amount of particles formed at 18-160 mM NaCl normalized to that at 18 mM NaCl (open bar). The final turbidity level of assembly reactions with capsid protein alone at the corresponding NaCl concentrations normalized to that at 18 mM NaCl (slashed bar).

Fig 6.**The Effect of Truncated Scaffolding Mutants on Assembly**

Assembly reactions containing only capsid protein, capsid protein in the presence of wild type scaffolding protein, and capsid protein in the presence of the indicated mutant scaffolding proteins were analyzed by centrifugation through a 10-40% sucrose gradient, fractionated from the bottom, and analyzed by SDS-PAGE.

Fig 7.
Sedimentation Equilibrium Analysis of Scaffolding Protein

- a.** Equilibrium sedimentation of wild type scaffolding protein at 50 μM (left panel) and 25 μM (right panel) loading concentration were recorded at rotor speeds of 33,000 rpm (circle) and 39,000 rpm (square). The fit to a single species model is shown as a solid line with the residual plots shown in the bottom panels.
- b.** Equilibrium sedimentation of Δ 1-9 scaffolding protein at 50 μM (left panel) and 25 μM (right panel) loading concentration were recorded at rotor speeds of 39,000 rpm (square) and 47,000 rpm (triangle). The fit to a single species model is shown as a solid line with the residual plots shown in the bottom panels.
- c.** Equilibrium sedimentation of Δ 62-97 scaffolding protein at 50 μM (left panel) and 25 μM (right panel) loading concentration were recorded at rotor speeds of 33,000 rpm (circle), 39,000 rpm (square) and 47,000 rpm (triangle). The fit to a single species model is shown as a solid line with the residual plots shown in the bottom panels.

Fig 8.

The CD Spectra of Wild Type and Truncated Scaffolding Proteins

The mean residual ellipticity recorded from 190 nm to 250 nm of wild type scaffolding protein (close squares), Δ 79-97 (open triangles), Δ 62-97 (inverted triangles) Δ 1-9(open squares), Δ 1-15 (crosses) and Δ 1-22 mutants (open stars).