Published in final edited form as:
J Mol Biol. 2008 July 11; 380(3): 465–475. doi:10.1016/j.jmb.2008.04.065.

Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids

Anton Poliakov[1], **Jenny R. Chang**[2,3], **Michael S. Spilman**[3], **Priyadarshan K. Damle**[4], **Gail E. Christie**[4], **James A. Mobley**[1], and **Terje Dokland**[3,‡]

1 *Department of Surgery–Urology, University of Alabama at Birmingham, Birmingham, AL*

2 *Department of Biology, University of Alabama at Birmingham, Birmingham, AL*

3 *Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL*

4 *Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA*

Abstract

The *Staphylococcus aureus* pathogenicity island SaPI1 carries the gene for the toxic shock syndrome toxin TSST-1 and can be mobilized by infection with *S. aureus* helper phage 80α. SaPI1 depends on the helper phage for excision, replication and genome packaging. The SaPI1 transducing particles are comprised of proteins encoded by the helper phage, but have a smaller capsid commensurate with the smaller size of the SaPI1 genome. Previous studies identified only 80α-encoded proteins in mature SaPI1 virions, implying that the presumptive SaPI1 capsid size determination function(s) must act transiently during capsid assembly or maturation. In this study, 80α and SaPI1 procapsids were produced by induction of phage mutants lacking functional 80α or SaPI1 small terminase subunits. By cryo-electron microscopy, these procapsids have a rounded shape and an internal scaffolding core. Mass spectrometry (MS) was used to identify all 80α-encoded structural proteins in 80α and SaPI1 procapsids, including several that had not previously been found in the mature capsids. In addition, SaPI1 procapsids contained at least one SaPI1-encoded protein that has been implicated genetically in capsid size determination. MS on full-length phage proteins showed that the major capsid protein and the scaffolding protein are N-terminally processed in both 80α and SaPI1 procapsids.

Keywords

mass spectrometry; cryo; electron microscopy; bacteriophage; assembly; scaffolding protein

Introduction

Staphylococcus aureus pathogenicity islands (SaPIs) are a family of 15- to 27- kb genetic elements that carry genes encoding a variety of superantigen toxins [1,2,3]. SaPIs are stably integrated at specific chromosomal sites, but can be mobilized following infection by certain staphylococcal bacteriophages [4] or by SOS induction of endogenous prophages by environmental stress, including certain antibiotics [5,6]. The prototype member of the SaPI family, SaPI1, contains genes for the toxic shock syndrome toxin (TSST-1) as well as

‡Communicating author: Address: Terje Dokland, Department of Microbiology, University of Alabama aBirmingham, 845 19th St South, BBRB 311, Birmingham, AL 35294, Email: dokland@uab.edu, Tel: (205) 996 4502, Fax: (205) 996 2667.

enterotoxins Q and K [2,4]. SaPI1 can be mobilized by bacteriophage 80α [7], leading to the formation of SaPI1 transducing particles that have a morphology similar to that of 80α, but with capsids that are about one-third the volume of those of the helper phage, commensurate with the smaller size of the SaPI1 DNA [4,8]. Integration and excision of SaPIs is phage-like, occurring by site-specific recombination between a specific *att* sequence on the pathogenicity island and a corresponding chromosomal *att* site, which leads to the generation of short direct repeats flanking the integrated element. SaPI1 encodes an integrase that resembles phage integrases and is sufficient for the site-specific integration of the pathogenicity island, but cannot promote SaPI1 excision in the absence of the helper phage [4]. A second phage-like SaPI1 gene encodes a protein homologous to the small subunit of phage terminase, an enzyme involved in DNA encapsidation.

The major capsid protein of 80α is gp47, the gene product (gp) of open reading frame (ORF) 47 [8]. (See GenBank accession code NC_009526 for the most recent gene assignments.) ORF 46 encodes a scaffolding protein which is not present in the mature capsid [8]. This is typical of most dsDNA bacteriophages, where the major capsid protein is assembled into a precursor capsid, or procapsid, requiring the action of a scaffolding protein that acts catalytically during assembly [9,10]. Formation of viable capsids also requires the incorporation of a portal protein that acts as the entry and exit portal for the DNA genome and as a connector between head and tail. The 80α portal has been identified as gp42 [8]. 80αencodes typical small (gp40) and large (gp41) terminase subunits that are used for DNA packaging. SaPI1 normally utilizes the SaPI1-encoded small terminase subunit for efficient DNA packaging into small capsids (G.E.C. and N.P. Olivarez, unpublished data).

The SaPI1-dependent capsid size determination is reminiscent of the exploitation of helper phage P2 by the genetically unrelated satellite phage P4 in *Escherichia coli* [11,12], during which P4 causes the packaging of P4 genomes into smaller capsids constructed from P2-encoded structural proteins. The T=4 icosahedral P4 capsid is one-third the size of the normal T=7 P2 capsid [13], thereby excluding the three times larger P2 genome from being packaged. The P4 size determination is dependent on the P4-encoded protein Sid [14], which forms an external scaffold around the P4 precursor capsids, or procapsids, during P4 assembly [15,16]. By analogy, it is reasonable to hypothesize that SaPI1 encodes proteins that direct the formation of smaller capsids from 80α helper phage proteins. Indeed, genetic experiments have shown that homologues of SaPI1 ORFs 6 and 7 in the closely related SaPIbov1 are involved in the size determination, and deletion of either of these two genes led to failure to form small capsids [17]. In previous experiments, we showed that SaPI1 virions are composed of the same 80α-encoded proteins as 80α virions [8]. Likewise, SaPIbov particles are composed of proteins derived from its helper phage, φ11 [18]. These results imply that the presumptive SaPI1-encoded capsid size determination functions act transiently at an earlier step in the assembly process, during procapsid assembly or capsid maturation. This would be equivalent to the mode of action of the P4 Sid protein.

In this study, we analyze the structure and protein composition of 80α and SaPI1 procapsids produced by mutants lacking the respective small terminase subunit genes. In-gel digestion followed by mass spectrometry was used to identify the proteins in the capsids. The procapsids contain several proteins that were not previously found in 80α and SaPI1 virions. Most notably, SaPI1 procapsids contain at least one SaPI1-encoded gene product that has previously been implicated in size determination. We have also found that both the major capsid protein and the scaffolding protein are N-terminally processed and we have identified the cleavage sites. Such cleavage is likely an important step in the capsid assembly and maturation process.

Results

Production and characterization of procapsids

The 80α small terminase subunit gene (*terS*) was deleted in-frame from the *S. aureus* 80α lysogenic strain RN10616 by allelic exchange, resulting in the Δ*terS* strain ST24 (G.E.C. and N.P. Olivarez, unpublished). Likewise, the SaPI1 *terS* gene was deleted from the SaPI1-positive 80α lysogen RN10628, yielding strain ST37 (see Methods). The small terminase subunits of 80α and SaPI1 are required for DNA packaging of the respective genomes (G.E.C. and N.P. Olivarez, unpublished data); hence, these mutants were expected to be blocked at the procapsid stage, prior to DNA packaging. The Δ*terS* cells were grown as previously described [19]. Phage production was induced by the addition of 2 mg/L mitomycin C. Upon cell lysis the phage particles were harvested by precipitation with polyethylene glycol (PEG) 8,000 and purified on CsCl gradients. The protein-containing band from the gradients was observed by negative stain EM and separated further by velocity sedimentation on sucrose gradients. The fractions from the sucrose gradient were analyzed by SDS-PAGE.

SDS-PAGE of the sucrose gradient-purified material produced in the 80α Δ*terS* mutant strain ST24 showed that tail proteins were mostly enriched in the more slowly sedimenting fractions (F3–F4; Fig 1A), while the major capsid and scaffolding proteins were found predominantly in the faster sedimenting fractions (F5–F8). Based on density measurements in the SDS-PAGE gel, the molar ratio of scaffolding to capsid protein (SP:CP) varies between 0.33 (lane 8) to 0.76 (lane 6). Very little soluble protein was seen at the top of the gradient, suggesting that structural proteins were effectively incorporated into stable procapsids.

The fractions containing predominantly capsid and scaffolding protein (F6–F8) were pooled, concentrated by pelleting and observed by cryo-EM (Fig. 2A). The 80α procapsids appear as a uniform population of 51 nm diameter particles with the thick-walled (≈4 nm thick), rounded shape that is characteristic of bacteriophage procapsids [20,21]. The procapsid particles contain a 37 nm inner core that is separated from the inside of the capsid shell by a ≈2 nm gap. The core seems to consist of elongated or punctate densities, assumed to correspond to the 80α scaffolding protein, gp46. A small number (<2 %) of thin-walled, larger capsids with a more angular outline were also present, presumably corresponding to spontaneously expanded shells and aberrant assemblies. By comparison, the capsids of DNA-filled 80α virions are larger (57–63 nm) and have a more angular shape and a thin capsid shell (Fig. 2B). The virions also have approximately 190 nm long, flexuous tails with large and elaborate baseplates.

The less highly purified 80α procapsid material from the CsCl density gradient included loose tails as well as procapsids (Fig. 2C). The tails are assembled via a separate pathway and co-purify with the procapsids on CsCl gradients. Surprisingly, a number of procapsids had tails attached. Normally, tail attachment should occur after DNA packaging, since premature tail attachment would lead to a failure to package DNA. As expected, no capsids filled with DNA were found in the absence of 80α TerS.

SaPI1 procapsids were made in the SaPI1 Δ*terS* mutant strain ST37 and purified on sucrose gradients in the same way (Fig. 1B). In this case, a large amount of capsid protein and scaffolding protein remained in the soluble fraction on top of the gradient, suggesting that assembly is less efficient than in 80α or that the SaPI1 procapsids are unstable. The rest of the protein separated into a fraction containing both tail and capsid protein (F2–F3) and a faster sedimenting fraction containing predominantly capsid and scaffolding protein (F4–F6; Fig. 1B). The ratio of scaffolding to capsid protein in these fractions was lower (around 0.2 for F4–F6) than in the 80α procapsid separation.

Cryo-EM of the sucrose-purified SaPI1 procapsids from the pooled capsid and scaffolding-containing fractions F4-F6 showed a number of 39 nm SaPI1 procapsid particles, smaller than, but similar in appearance to, the 80α procapsids (Fig. 2D). However, more than 50% of the particles were thin-walled shells with a diameter of 46 nm, lacking a scaffolding core (Fig 2D). This again suggests that the SaPI1 procapsids are inherently less stable than the 80α procapsids and more liable to spontaneous expansion and scaffolding loss, at least in the presence of sucrose, and is consistent with the lower measured SP:CP ratio.

The ST37 sample also yielded a higher density fraction in the CsCl gradient that was found to contain mature SaPI1 particles filled with DNA (not shown), showing that the 80αTerS is able to package DNA into SaPI1 size particles, albeit at lower efficiency than the SaPI1 TerS. This packaged DNA consists predominantly of 80α DNA fragments (G.E.C. and N. Olivarez, unpublished data). Thus, at least some of the empty, expanded shells seen in the sucrose purified sample could have resulted from packaged SaPI1 particles that had lost their DNA. Only a few larger 80α-size shells were seen (<5% of procapsids and expanded shells; Fig 2D, inset), consistent with the reported interference of SaPI1 with the lytic growth of 80α[7].

Identification of procapsid proteins

SDS-PAGE of the CsCl-purified 80α and SaPI1 proteins showed a set of bands mostly corresponding to those previously described for the CsCl-purified 80α and SaPI1 virions [8], including a full complement of head and tail proteins (Fig. 3). Consistent with previous results[8], the observed bands include minor tail proteins gp59, gp61, gp62 and gp68 (theoretical masses of 71.0, 73.7, 66.8 kDa and 43.8 kDa, respectively), portal protein (gp42; 59.5 kDa), the major capsid protein gp47 (36.8 kDa) and the major tail protein gp53 (21.5 kDa). In addition, both 80α and SaPI1 procapsid samples displayed a prominent band at an apparent mass of around 27 kDa, which corresponds to the gp46 scaffolding protein (calculated mass 23.4 kDa). SDS-PAGE of the sucrose-gradient purified 80α and SaPI1 procapsids showed a similar set of capsid proteins, but greatly reduced amounts of the tail proteins. In addition, the SaPI1 procapsids yielded a faint band at 8 kDa (Fig. 3), subsequently identified as SaPI1 gp6 (see below).

For a conclusive identification of the complete set of structural proteins in 80α and SaPI1 procapsids, mass spectrometry was employed. For this experiment, the CsCl-purified samples as well as the sucrose gradient-purified procapsids of both 80α and SaPI1 were separated by SDS-PAGE, in-gel digested with trypsin and analyzed by liquid chromatography and tandem mass spectroscopy. The resulting spectra were compared to a *S. aureus* protein database to identify all phage- and SaPI1-related proteins in the digest.

A complete list of all 80α and SaPI1 proteins detected in both samples is shown in Table 1. Although the MS experiment does not quantitate the amount of the different proteins, the percentage sequence coverage listed in the table gives a rough indication of the relative abundance, although the high sensitivity of the MS instrument ensures that even very low abundance proteins in the samples can be detected. Further details, including a list of the individual peptides identified, can be found in Supplementary Table 1 (80α proteins) and Supplementary Table 2 (SaPI1 proteins).

The proteins detected include all those previously identified in virions [8] and seen by SDS-PAGE. A number of additional proteins were also detected, as shown in Table 1. Several proteins were detected in the CsCl-purified material, but not in the sucrose-purified procapsids or were detected with a greatly reduced sequence coverage, suggesting a much lower abundance. Many of these correspond to tail-related proteins, while some may correspond to proteins that are only loosely attached to the procapsids. Others likely correspond to abundant non-structural phage proteins that co-purified on the CsCl gradients. We suspect this is the case

for proteins gp8, gp20, gp48 and gp71, which have described functions that are not of a structural nature. Nevertheless, it is possible that they have a dual role, similar to that of the Psu antiterminator protein of bacteriophage P4, which doubles as a capsid decoration protein [22,23]. A number of unrelated host proteins were also detected in the CsCl-purified material, including ribosome-related proteins and abundant enzymes (data not shown). Very few host proteins were found in the sucrose-purified sample. The 80α and SaPI1 proteins detected in the CsCl protein fraction include several ORFs listed as "hypothetical proteins" in the GenBank entry. The presence of these proteins demonstrates that these ORFs do correspond to real gene products.

Some of the most notable differences between the proteins found in the procapsids and those previously observed in the virions by Tallent *et al.* (2007)[8] are listed below:

Phage 80α protein gp10 corresponds to a small 6.7 kDa protein (YP_001285324.1) of unknown function that was found in SaPI1 procapsids, but not in 80α procapsids, and was not previously observed in virions. Similarly, 80α proteins gp16 (YP_001285330.1; 23.7 kDa) and gp36 (YP_001285350.1; 15.0 kDa) correspond to proteins of unknown function not previously seen in virions, but found in both 80α and SaPI1 procapsids. The abundance in the sucrose-purified procapsids appears to be low, except for gp36 in SaPI1, suggesting only a loose association with procapsids.

Gp40 and gp41 correspond to the 80α small and large terminase subunit, respectively. Both were present in the SaPI1 CsCl sample, but were not found in purified procapsids. The 80α procapsid mutant ST24 had the *terS* gene deleted, and consequently did not contain any TerS protein. The fact that TerL is only detected, even in the CsCl fraction, in the presence of TerS suggests that TerL binds to the procapsids via interaction with TerS. This is consistent with the proposed mechanism in bacteriophage P22, in which TerS forms a ring-like oligomer that docks with the procapsid during DNA packaging, while TerL binds to this TerS complex [24], but differs from other systems, like T4, in which it is TerL that binds to the procapsid [25,26]. The apparent association of a small amount of 80α terminase complex with SaPI1 procapsids is consistent with the demonstrated ability of the 80α TerS to package a limited amount of DNA into small (SaPI1-size) procapsids. The absence of terminase proteins in the purified procapsids indicates that the terminase complex is only loosely bound to capsids, probably only transiently during DNA packaging. It is possible that the SaPI1 TerS would bind the small procapsids more efficiently, thus providing a possible basis for the DNA packaging specificity.

ORFs 53–69 appear to constitute an operon that contains mainly tail-related genes. Some of the proteins detected that are encoded by this region can be directly related to analogous genes in other phages, including *E. coli* phage lambda [27] and *Lactococcus* phage Tuc2009 [28]. Thus, gp54 and gp55, which were present in the 80α CsCl-purified protein sample, appear to correspond to "tail assembly proteins" that would normally not be present in mature phage [29]. Gp58, gp64, gp66 and gp67, not previously detected, also most likely correspond to tail proteins. By comparison with Tuc2009, they are probably located in the baseplate, as are the previously described tail proteins gp59, gp61 and gp62, although they appear to have no direct parallels in Tuc2009. Deletion of the gp62 homolog (ORF54) in *S. aureus* phage φ11affects formation of the baseplate structure [18]. Curiously, while this protein is essential for growth of φ11 or 80α, deletion of this gene reduces, but does not eliminate, SaPI transduction by either phage. On the other hand, we see no evidence of several hypothetical proteins encoded by other small ORFs in the same operon, namely gp57, gp60, gp63 and gp65. It is likely that some or all of these ORFs do not actually code for any protein products or that they play only a catalytic role and are not stably associated with the tail structure. While some tail proteins are only seen in the CsCl-purified material, the more abundant ones are also found in the purified procapsid

fractions, since some tails and procapsids with attached tails would be expected to co-sediment with the procapsids in the sucrose gradients.

Five additional SaPI1-encoded proteins were found in the CsCl-purified SaPI1 protein sample, encoded by ORFs 5, 6, 7, 11 and 18, although only gp6 was seen at high abundance in the sucrose gradient-purified procapsids (Table 1). Gp5 and gp7 were not seen on SDS-PAGE of the CsCl material, due to overlap with the gp46 scaffolding protein band, but their presence there was confirmed by MALDI-MS. The 8.3 kDa gp6 protein corresponds to the 8 kDa band seen by SDS-PAGE (Fig. 3). Genetic evidence has implicated gp6 and gp7 in capsid size determination, while gp5 may play an accessory role [17,30]. The functions of gp11 and gp18 are unknown, but deletion of the genes encoding the homologous proteins from SaPIbov did not impair SaPI replication or transduction [30]. The low abundance of gp7 in the purified procapsids suggests that it either serves a non-structural function, acts at an early stage in assembly or is only loosely attached to the procapsids.

Characterization of full-length proteins

Plasmid pPD2 is a derivative of pET21 that overexpresses the 80α capsid and scaffolding proteins in *E. coli* (see Methods). When the proteins expressed following induction of pPD2 were analyzed by SDS-PAGE side-by-side with the purified 80α and SaPI procapsids, there was a discrepancy in the mobility of both the capsid protein (gp47) and scaffolding protein (gp46) (Fig. 4A). This suggested that both proteins were cleaved in the native procapsids, and hence also in the virions.

To determine the cleavage site of the two proteins, CsCl-purified 80α and SaPI1 procapsid-containing samples were disrupted in 6M urea and 0.1 % formic acid, and the proteins were separated by reverse phase chromatography and analyzed by ESI-MS as previously described [31](Fig. 5A and supplementary data). Only the most abundant proteins are detected in this experiment, and low molecular weight proteins are detected more efficiently than larger proteins. The measured masses of 35,062.20±0.18 Da in 80α and 35,061.86±0.53 Da in SaPI1 are within 1 Da of the calculated mass (35,062.9 Da) for residues 15–324 of the 80α major capsid protein (gp47). The measured mass of 21,700.5±0.07 Da (in 80α) and 21,701.05±0.34 Da (in SaPI1) are within 1 Da of the calculated mass (21,700.9 Da) for residues 14–206 of the 80α scaffolding protein gp46. Although the high mass accuracy of the measurement allowed the sequence to be assigned with high confidence, the N- and/or C-terminal sequences of the peptides were confirmed by tandem mass spectrometry (see supplementary data). This result shows that there is identical processing of capsid and scaffolding proteins during assembly of both 80α and SaPI1 procapsids.

In addition, masses of 21,394.86±0.30 Da and 21,394.47±0.61 Da were found in 80αand SaPI1, respectively. These correspond to gp53, the major tail protein, after removal of the N-terminal Met (theoretical mass 21,394.5 Da). In addition, the SaPI1 procapsids spectra showed measured masses of 8,271.57±0.19 Da and 22,804.03±0.41 Da, which correspond to the SaPI proteins gp6 (theoretical mass 8,272.2 Da) and gp7 (theoretical mass 22,805.3 Da), respectively, both proteins with their N-terminal Met residues retained. These are the two SaPI1 proteins implicated in size determination [17] that were also found in the tryptic digests (Table 1). The fact that SaPI1 gp6 and gp7 show up so clearly in the whole phage preparation suggests that these proteins are major constituents of SaPI1 procapsids. Failure to detect gp7 by SDS-PAGE is due to its comigration with the scaffolding protein.

Discussion

Procapsid and scaffold structure

Electron microscopy of 80α and SaPI1 procapsids confirmed the expected features of bacteriophage procapsids: smaller size than the mature capsids, a more rounded, thick-walled shape and an internal core. The 80α and SaPI1 procapsids appear to contain a complete core shell, consisting of radial segments of around 12–14 nm length (Fig. 1B), which would correspond to 80–100 residues of an α-helix. The mature, 193-residue gp46 protein is predicted to be predominantly α-helical, like other known scaffolding proteins, and has a strong propensity for coiled coil formation. Perhaps the observed radial segments correspond to these coiled coils. Density measurements on the SDS-PAGE gels indicate that scaffolding and capsid proteins are found in a molar ratio of between 0.33:1 and 0.76:1, corresponding to 137–315 copies of gp46 and 415 copies of gp47, assuming T=7 symmetry for the 80α capsid and one unique vertex. In SaPI1, the ratio is only 0.2:1 suggesting that other factors may substitute for some of the gp46. However, the SaPI1 measurements are complicated by the presence of expanded scaffold-less shells in this sample (Fig. 2D). Thus, the role of the scaffolding protein may be to switch the conformation of a specific subset of gp47 monomers.

The presence of procapsids with tails was surprising. It is possible that the high concentration of procapsids and tails that accumulate in the Δ*terS* mutant leads to spontaneous tail attachment. More likely, TerS normally blocks tail attachment by binding to the portal complex, so that in the 80α Δ*terS* mutant, tails would be free to attach to the heads. That there are fewer heads with tails attached in the SaPI1 sample supports this idea, since the MS results suggest that at least some SaPI1 procapsids have the 80α TerS attached during assembly.

Role of protein cleavage in capsid assembly and maturation

Comparison of the full-length 80α major capsid and scaffolding proteins expressed in *E. coli* to those found in procapsids revealed that both proteins are cleaved in 80α and SaPI1 procapsids and hence also in the virions (Fig. 4). Analysis of sequences around the truncation sites of the major capsid and scaffolding proteins revealed significant similarities, with a consensus sequence of KLKxNLQxF*A, where * denotes the cleavage site (Fig. 4). This similarity strongly suggests that the proteins are cleaved by the same protease. There is no difference in the processing of the two proteins between 80α and SaPI1, ruling out an involvement of differential protein processing in the capsid size determination.

Cleavage of the structural proteins is a common, albeit not universal feature of dsDNA bacteriophages. Bacteriophages P2, T4 and HK97 all exhibit cleavage of the major capsid protein. Such cleavage occurs after procapsid assembly and is considered an important control point in the assembly pathway that is essential for capsid expansion to occur [21,32]. The effect of the cleavage may be to cause capsid expansion by destabilizing procapsid protein–protein interactions or to induce release of the scaffolding proteins. In bacteriophage P4 procapsids, experimental cleavage of the capsid protein with trypsin led to capsid expansion and scaffold removal [33]. Cleavage and even complete degradation of the scaffolding protein is also commonly found and provides a mechanism for scaffolding removal and escape from the procapsid in some systems [9,10].

The maturation cleavage in these systems is carried out by a phage protease that is usually located immediately upstream of the scaffolding protein in the same operon and sometimes the protease function is embedded in the scaffolding protein itself. Thus, the most likely candidate for a protease function in 80α is gp44. This protein was found in both 80αand SaPI1 procapsids, and previously in 80α virions [8], and appears to be a minor component of the core

itself. If gp44 is indeed the protease, deletion of ORF 44 would most likely lead to the accumulation of uncleaved, immature procapsids.

Capsid size determination

The most significant finding in our study in terms of understanding capsid size determination was the identification of several SaPI1-encoded proteins in the SaPI1 procapsids. Two of these proteins corresponded to the SaPI1 proteins gp6 and gp7, which had previously been implicated in size determination [17]. Although both proteins were abundant in the CsCl sample and easily detected by ESI-MS of full-length proteins (Fig. 5), only gp6 was found in abundance in sucrose-purified SaPI1 procapsids, suggesting that gp7 may act earlier in assembly or is only loosely attached to the procapsids. The identification of gp6, and possibly gp7, as structural components of the SaPI1 procapsid, but not of the mature capsid, suggests that they may function as alternative size-determining scaffolding proteins, similar to the Sid protein of bacteriophage P4 [15]. On SDS-PAGE gels (Fig. 3) gp6 appears to exist in a molar ratio of 0.156:1 with capsid protein, or a total of 37 copies, assuming that the SaPI1 capsid has T=4 symmetry (235 copies of gp47). This is most likely an underestimate, since Coomassie staining of such small proteins is not quantitative. The amount of gp7 in the CsCl-purified SaPI1 material cannot be estimated directly from the SDS-PAGE since it overlaps with the band that includes both gp5 and the gp46 scaffolding protein.

What are the roles of gp6 and gp7 in the size determination? In the P2/P4 system, capsid size is determined by the 244 residue Sid protein, which forms an external scaffold on P4 procapsids [15,16]. Secondary structure prediction (PredictProtein) shows that the 192 amino acid gp7 protein is almost 100% α-helical. This is a typical characteristic of scaffolding proteins [9,10], including Sid. If gp7 forms a similar, external scaffold on SaPI1 procapsids, however, it appears to be only transiently or loosely associated with the procapsids. The 72 amino acid, predominantly α-helical gp6 protein, on the other hand, is stably associated with procapsids. In cryo-electron micrographs of P4 procapsids, the external scaffold is readily visible [16]. However, no such external scaffold is visible in SaPI1 procapsids, which have a similar outline to the 80α procapsids (Fig. 2). Furthermore, density measurements suggest that the ratio of scaffolding to capsid protein is less in SaPI1 than in 80α procapsids. Together, these data suggest that gp6 may act as an *internal* scaffolding protein that substitutes for some of the gp46 in the SaPI1 procapsids.

The function of gp6 and gp7 is in addition to, rather than instead of, the scaffolding protein of the helper phage, as deletion of 80α ORF46 in a SaPI1-containing 80α lysogen eliminated the ability to produce any viable phage or SaPI1 transducing particles (data not shown). However, it is possible that gp6 and gp7 are sufficient for small procapsid assembly *per se*, and that gp46 is required for some other activity, such as incorporation of the portal or capsid maturation. This would be equivalent to the situation in the P2/P4 system, where morphologically correct small capsids can be made in the presence of the P4 Sid protein without the gpO internal scaffolding protein even though gpO is essential for the formation of viable capsids that can package DNA [34,35].

SaPI1 protein gp5 was reported to be non-essential for the formation of small procapsids, but has an effect on the amount of SaPI-sized DNA packaged into capsids [17,30]. In our assay, it is only seen in the CsCl-purified material and not in purified procapsids. One possible role for this protein could be to modulate the interaction of procapsids with terminase.

Ultimately, a more careful structural analysis of 80α and SaPI1 procapsids will be needed to determine which of these models is correct. The tailed dsDNA bacteriophages are evolutionarily related [36], so structural and functional similarities between 80α and P2 would be expected. However, the satellite phage P4 and SaPI1 belong to different lineages (not phage-

related) and are not likely to be evolutionarily related. It should therefore be very interesting to examine how these diverse replicons have acquired functionally analogous helper phage exploitation and capsid determination mechanisms.

Materials and methods

Generation of mutants

Allelic exchange was used to construct ST37, a strain of *S. aureus* carrying an 80α prophage and a copy of SaPI1 with an in-frame deletion removing most of the coding sequence of the SaPI1 small terminase gene. Flanking DNA fragments of ~1 kb each were amplified by PCR and ligated together. A region from ORF6 to the beginning of SaPI1 *terS* was amplified from SaPI1 *tst::tetM* with primers SMT53 and SMT54 (Table 2), and a region from *tetM* to the end of SaPI1 *terS* was amplified with primers SMT51 and SMT52. The PCR products were digested with *Hin*dIII, ligated together, and used as template for amplification with SMT51 and SMT54. After purification, the resulting PCR product was digested with *Nco*I and ligated with *Nco*I-digested pMAD [37]. The resulting plasmid, pPD11, was introduced into *E. coli* DH5α by electroporation. Following verification of the construct by sequencing, the plasmid was introduced by electroporation into the SaPI1 *tst::tetM* containing 80α lysogen RN10628 (R.P. Novick, unpublished) for allelic exchange. Cointegrates were selected by incubation at 44 °C in tryptic soy broth (TSB, Gifco) containing 5 μg/ml erythromycin in the presence of 200 μg/ml X-gal. Individual Ery^R blue colonies were purified, grown at 30 °C in TSB to select for cointegrate resolution, and then plated onto prewarmed TSB plates and incubated at 44 °C to cure the cells of the plasmid. Individual white colonies were screened by PCR to confirm the desired deletion and verified by sequencing. Construction of ST24, an 80α lysogen carrying an in-frame deletion of the 80α small terminase subunit gene, was accomplished using a similar strategy (G.E.C. and N.P. Olivarez, unpublished).

Cloning and expression of capsid and scaffolding protein

The 80α capsid gene was amplified from 80α DNA using primers SMT34 and SMT35, cleaved with *Nde*I and *Bam*HI, and ligated into the pET21a (Novagen) vector. The resulting plasmid, pPD1, was cleaved with *Bam*HI and *Sal*I and ligated with a fragment containing the 80α scaffolding protein gene and a synthetic ribosome binding site, generated by PCR amplification with primers SMT61 and SMT62. Expression of these two proteins from pPD2 in *E. coli* strain BLR (DE3) (Invitrogen) was accomplished by induction with 0.5 mM IPTG at 37°C. The cells were harvested 2 hr post induction, lysed by passage through an EmulsiFlex C3 high pressure homogenizer (Avestin Inc., Ottawa) and prepared for SDS-PAGE.

Production and purification of procapsids

The 80α ΔterS lysogen ST24 was used as the source of 80a procapsids, while SaPI1 procapsids were isolated from the SaPI1ΔterS-containing 80α lysogen ST37. Strains were grown at 32 °C in CY broth as previously described [19], and the 80α prophages were induced by the addition of 2 mg/L mitomycin C (Sigma) at OD_{540}=0.4–0.5 with lysis occurring 3 hr post induction. The lysates were clarified by centrifugation at 5,400 × *g* for 20 min. Procapsids were precipitated with 10% (w/v) polyethylene glycol 6000 and 0.5M NaCl overnight at 4°C followed by centrifugation at 5,400 × *g* for 20 min. The pellets were resuspended in phage buffer (50 mM Tris, pH 7.8, 100 mM NaCl, 1 mM $MgSO_4$, 4 mM $CaCl_2$) and 0.30–0.50 g CsCl was added per ml of suspension. In some cases, half a volume of chloroform was added to the PEG pellet to facilitate resuspension, resulting in a cleaner sample. No difference was observed in the structure or protein composition between treated and untreated samples. The resulting solution was centrifuged at 339,000×*g* for 20 h in a Beckman NVT90 rotor. The procapsid-containing bands from the CsCl gradients were collected and dialyzed against

dialysis buffer (20 mM Tris, pH 7.8, 50 mM NaCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$) for further analysis.

To separate the procapsids from tails, the dialyzed CsCl bands were loaded onto 10–40% sucrose gradients in phage buffer and centrifuged for 2 h at 110,000×*g* in a Beckman SW41 rotor. Twelve fractions were collected manually from the sucrose gradients and analyzed by SDS-PAGE. Fractions containing predominantly procapsids and procapsids with attached tails were pelleted by centrifugation at 110,000×*g* for 1 hr. The pellet was resuspended in dialysis buffer and used for EM and MS experiments.

Electron microscopy

Samples for negative stain were prepared by applying 3 μl procapsid suspension to glow-discharged carbon-only grids (Electron Microscopy Sciences), washed 2× with dialysis buffer and stained with 1% uranyl acetate. Cryo-EM was done by standard methods [38]: 3μl of sample was applied to C-flat holey film (Electron Microscopy Sciences), blotted briefly before plunging into liquid ethane and transferred to a Gatan cryo-sample holder. All samples were observed in an FEI Tecnai F20 electron microscope operated at 200kV and images were captured on a 4k × 4k Gatan Ultrascan CCD camera or on Kodak SO-163 film at magnifications from 38,000× to 81,200×.

Protein identification

Coomassie-stained SDS-PAGE gels of 80α and SaPI1 procapsids were cut into 10–12 strips. The strips were destained in 60% methanol, 0.1% trifluoroacetic acid (TFA) and dried with pure acetonitrile. The acetonitrile was removed by evaporation in a Speedvac centrifugal evaporator and protein digestion was performed by addition of 10 μg/ml Trypsin Gold (Promega) solution in 100 mM ammonium bicarbonate and incubated for 8 hours at 37°C. After extraction with 10% acetonitrile, the peptides were loaded on a 100 nm ×10 cm capillary column in-house packed with C18 Monitor 100 A-spherical silica beads and eluted by a one hour gradient of 10–100% acetonitrile, 0.1% TFA. Mass spectrometric analysis was performed on an LTQ XL (Thermo Finnigan) spectrometer. The search for matching peptide sequences was performed using the SEQUEST search engine with the UniProt database including Staphylococcal and phage entries. Only peptides with a probability of >0.99 were considered [39].

Measurement of full-length protein masses

Measurement of the whole masses of most abundant proteins composing 80α and SaPI1 procapsids was performed as previously described [31]. Briefly, purified procapsids were disrupted in 6 M urea and 0.1% TFA, loaded on a C4 microtrap reverse phase column and eluted with a gradient of 0–50% isopropanol, 0–50% acetonitrile, in 0.1% formic acid. Spectra were acquired on a time-of-flight electrospray mass spectrometer (LCT, Micromass).

Confirmation of cleavage sites of scaffolding and major capsid proteins was achieved by tandem mass spectrometry of N- and C-terminal peptides on MALDI-TOF/TOF tandem mass spectrometer (Ultraflex III, Bruker Daltonics). In-gel digests of 80α scaffolding and major capsid proteins were spotted on a Bruker 384-spot plate and allowed to dry. Equal volumes of 5 mg/ml of α-cyano-hydroxycinnamic acid in 60% acetonitrile and 0.1% TFA were spotted on top of the digests and allowed to dry. Peptides matching expected C- or N-terminal peptides of truncated proteins to less than 0.02 Da were fragmented by laser-induced dissociation. Fragment assignment was performed with BioTools software (Bruker Daltonics) using 0.5 Da fragment mass accuracy cutoff.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank Dr. Peter E. Prevelige for access to his ESI mass spectrometer. This work was funded by NIH grants AI071982 to T.D. and AI067654 to G.E.C.

References

1. Foster TJ. Immune evasion by staphylococci. Nat Rev Microbiol 2005;3:948–958. [PubMed: 16322743]
2. Novick RP. Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. Plasmid 2003;49:93–105. [PubMed: 12726763]
3. Alouf JE, Muller-Alouf H. Staphylococcal and streptococcal superantigens: molecular, biological and clinical aspects. Int J Med Microbiol 2003;292:429–440. [PubMed: 12635926]
4. Ruzin A, Lindsay JA, Novick RP. Molecular genetics of SaPI1 - a mobile pathogenicity island in *Staphylococcus aureus*. Mol Microbiol 2001;41:365–377. [PubMed: 11489124]
5. Ubeda C, Maiques E, Knecht E, Lasa I, Novick RP, Penades JR. Antibiotic-induced SOS response promotes horizontal dissemination of pathogenecity island-encoded virulence factors in staphylococci. Mol Microbiol 2005;56:836–844. [PubMed: 15819636]
6. Maiques E, Ubeda C, Campoy S, Salvador N, Lasa I, Novick RP, Barbe J, Penades JR. beta-lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in *Staphylococcus aureus*. J Bacteriol 2006;188:2726–2729. [PubMed: 16547063]
7. Lindsay JA, Ruzin A, Ross HF, Kurepina N, Novick RP. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus.* Mol Microbiol 1998;29:527–543. [PubMed: 9720870]
8. Tallent SM, Langston TB, Moran RG, Christie GE. Transducing particles of *Staphylococcus aureus* pathogencity island SaPI1 are comprised of helper phage-encoded proteins. J Bacteriol 2007;189:7520–7524. [PubMed: 17693489]
9. Dokland T. Scaffolding proteins and their role in viral assembly. Cell Mol Life Sci 1999;56:580–603. [PubMed: 11212308]
10. Fane BA, Prevelige PE. Mechanism of scaffolding-assisted viral assembly. Adv Prot Chem 2003;64:259–299.
11. Christie GE, Calendar R. Interactions between satellite bacteriophage P4 and its helpers. Annu Rev Genet 1990;24:465–490. [PubMed: 2088176]
12. Lindqvist BH, Deho G, Calendar R. Mechanisms of genome propagation and helper exploitation by satellite phage P4. Microbiol Rev 1993;57:683–702. [PubMed: 8246844]
13. Dokland T, Lindqvist BH, Fuller SD. Image reconstruction from cryo-electron micrographs reveals the morphopoietic mechanism in the P2-P4 bacteriophage system. EMBO J 1992;11:839–846. [PubMed: 1547786]
14. Shore D, Deho G, Tsipis J, Goldstein R. Determination of capsid size by satellite bacteriophage P4. Proc Natl Acad Sci USA 1978;75:400–404. [PubMed: 272656]
15. Marvik OJ, Dokland TE, Nøkling RH, Jacobsen E, Larsen T, Lindqvist BH. The capsid size-determining protein Sid forms an external scaffold on phage P4 procapsids. J Mol Biol 1995;251:59–75. [PubMed: 7643390]
16. Dokland T, Wang S, Lindqvist BH. The structure of P4 procapsids produced by coexpression of capsid and external scaffolding proteins. Virology 2002;298:224–231. [PubMed: 12127785]
17. Ubeda C, Maiques E, Tormo S, Campoy S, Lasa I, Barbe J, Novick RP, Penades JR. SaPI operon I is required for SaPI packaging and is controlled by LexA. Mol Microbiol 2007;65:41–50. [PubMed: 17581119]

18. Tormo MA, Ferrer MD, Maiques E, Ubeda C, Selva L, Lasa I, Calvete JJ, Novick RP, Penades JR. SaPI DNA is packaged in particles composed of phage proteins. J Bacteriol 2008;190:2434–2440. [PubMed: 18223072]
19. Novick RP. Genetic systems in Staphylococci. Methods Enzymol 1991;204:587–636. [PubMed: 1658572]
20. Dokland T, Murialdo H. Structural transitions during maturation of bacteriophage lambda capsids. J Mol Biol 1993;233:682–694. [PubMed: 8411174]
21. Conway JF, Duda RL, Cheng N, Hendrix RW, Steven AC. Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. J Mol Biol 1995;253:86–99. [PubMed: 7473720]
22. Isaksen ML, Dokland T, Lindqvist BH. Characterization of the capsid associating activity of bacteriophage P4's Psu protein. Virology 1993;194:647–681. [PubMed: 8389077]
23. Dokland T, Isaksen ML, Fuller SD, Lindqvist BH. Capsid localization of the bacteriophage P4 Psu protein. Virology 1993;194:682–687. [PubMed: 8503181]
24. Nemecek D, Gilcrease EB, Kang S, Prevelige PE, Casjens S, Thomas GJ. Subunit conformations and assembly states of a DNA translocating motor: the terminase of bacteriophage P22. J Mol Biol 2007;374:817–836. [PubMed: 17945256]
25. Maluf NK, Gaussier H, Bogner E, Feiss M, Catalano CE. Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine. Biochemistry 2006;45:15259–15268. [PubMed: 17176048]
26. Kanamaru S, Kondabagil K, Rossmann MG, Rao VB. The functional domains of bacteriophage T4 terminase. J Biol Chem 2004;279:40795–40801. [PubMed: 15265872]
27. Hendrix, RW.; Casjens, S. Bacteriophage lambda and its genetic neighborhood. In: Calendar, R., editor. The Bacteriophages. 2. Oxford University Press; New York: 2006. p. 409-447.
28. McGrath S, Neve H, Seegers JFML, Eijlander R, Vegge CS, Brøndsted L, Heller KJ, Fitzgerald GF, Vogensen FK, van Sinderen D. Anatomy of a Lactococcal phage tail. J Bacteriol 2006;188:3972–3982. [PubMed: 16707689]
29. Xu J, Hendrix RW, Duda RL. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell 2004;16:11–21. [PubMed: 15469818]
30. Ubeda C, Maiques E, Barry P, Matthews A, Tormo MA, Lasa I, Novick RP, Penades JR. SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. Mol Microbiol 2008;67:493–503. [PubMed: 18086210]
31. Chang JR, Poliakov A, Prevelige PE, Mobley JA, Dokland T. Incorporation of scaffolding protein gpO in bacteriophages P2 and P4. Virology 2008;370:352–361. [PubMed: 17931675]
32. Marvik OJ, Jacobsen E, Dokland T, Lindqvist BH. Bacteriophage P2 and P4 morphogenesis: assembly precedes proteolytic processing of the capsid proteins. Virology 1994;205:51–65. [PubMed: 7975237]
33. Wang S, Chandramouli P, Butcher S, Dokland T. Cleavage leads to expansion of bacteriophage P4 procapsids in vitro. Virology 2003;314:1–8. [PubMed: 14517054]
34. Lengyel JA, Goldstein RN, Marsh M, Sunshine MG, Calendar R. Bacteriophage P2 head morphogenesis: cleavage of the major capsid protein. Virology 1973;53:1–23. [PubMed: 4574872]
35. Marvik OJ, Sharma P, Dokland T, Lindqvist BH. Bacteriophage P2 and P4 assembly: alternative scaffolding proteins regulate capsid size. Virology 1994;200:702–714. [PubMed: 8178454]
36. Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci USA 1999;96:2192–2197. [PubMed: 10051617]
37. Arnaud M, Chastanet A, Debarbouille M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. Appl Environ Microbiol 2004;70:6887–6891. [PubMed: 15528558]
38. Dokland, T.; Ng, ML. Electron microscopy of biological samples. In: Dokland, T.; Hutmacher, DW.; Ng, ML.; Schantz, JT., editors. Techniques in microscopy for biomedical applications. World Scientific Press; Singapore: 2006.
39. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003;75:4646–4658. [PubMed: 14632076]

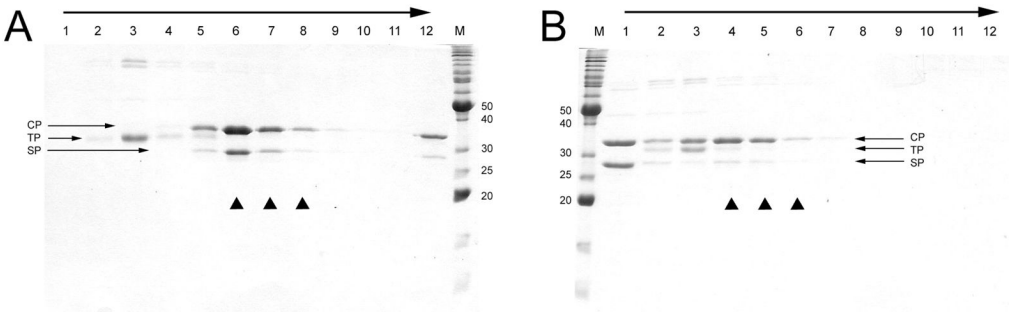


Figure 1.

SDS-PAGE of sucrose gradient-separated 80α(A) and SaPI1 (B) procapsids. Fraction numbers (1 ml fractions) and the direction of sedimentation (arrow) are indicated. M, marker, MW as indicated (kDa). The positions of the capsid protein (CP), major tail protein (TP) and scaffolding protein (SP) are indicated by arrows, while the triangles indicate fractions that were pooled and used for cryo-EM and MS analysis.

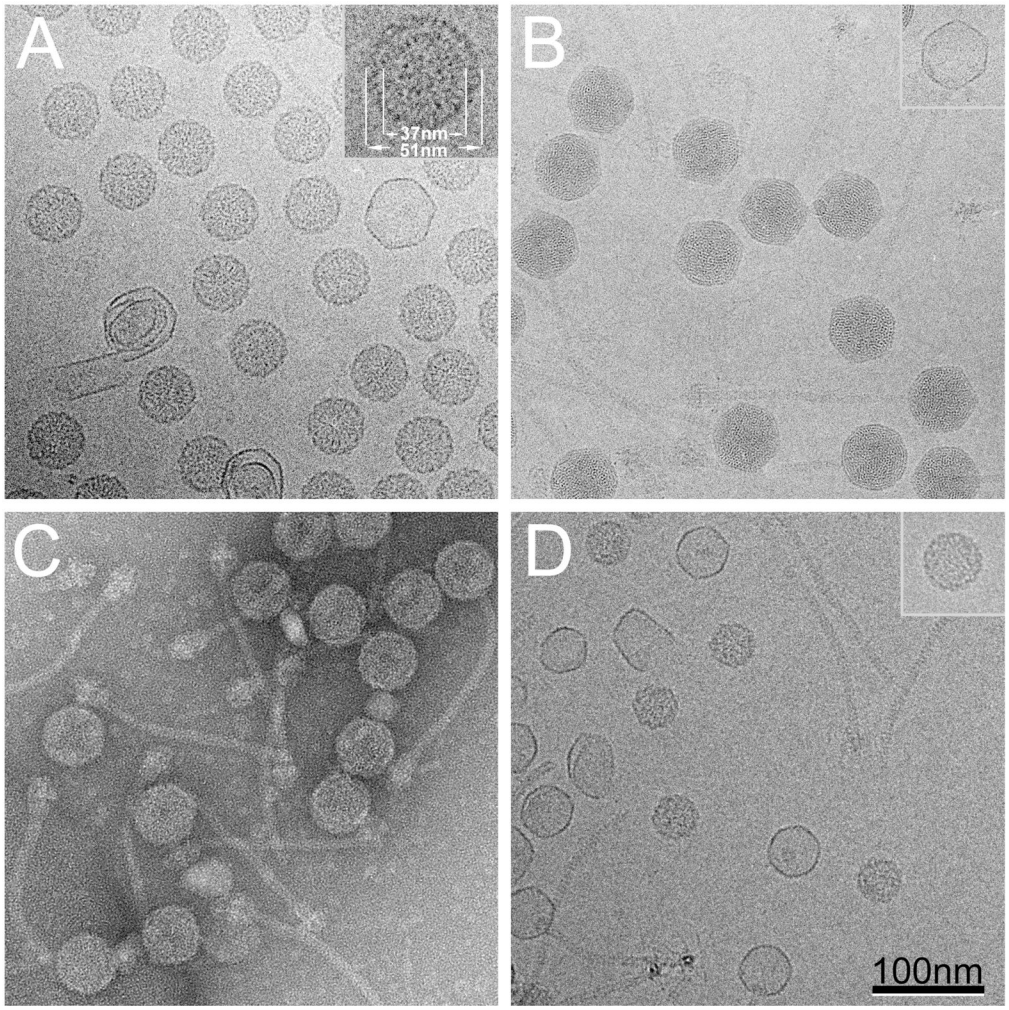


Figure 2.
(A) Cryo-EM of sucrose gradient–purified 80α procapsids. Inset, 2× magnified view of one procapsid with dimensions of the shell and the inner core indicated. (B) Cryo-EM of CsCl-purified 80α virions. The 2 nm spacing of the internal DNA is clearly visible. One thin-walled, empty capsid is also shown (inset). (C) Negatively stained CsCl-purified 80α procapsid fraction, containing a mixture of procapsids, tails and procapsids with attached tails. (D) Cryo-EM of sucrose gradient-purified SaPI1 procapsids. Inset, one of the about 5% 80α-size procapsids found in the SaPI1 procapsid sample. Scale bar, 100 nm.

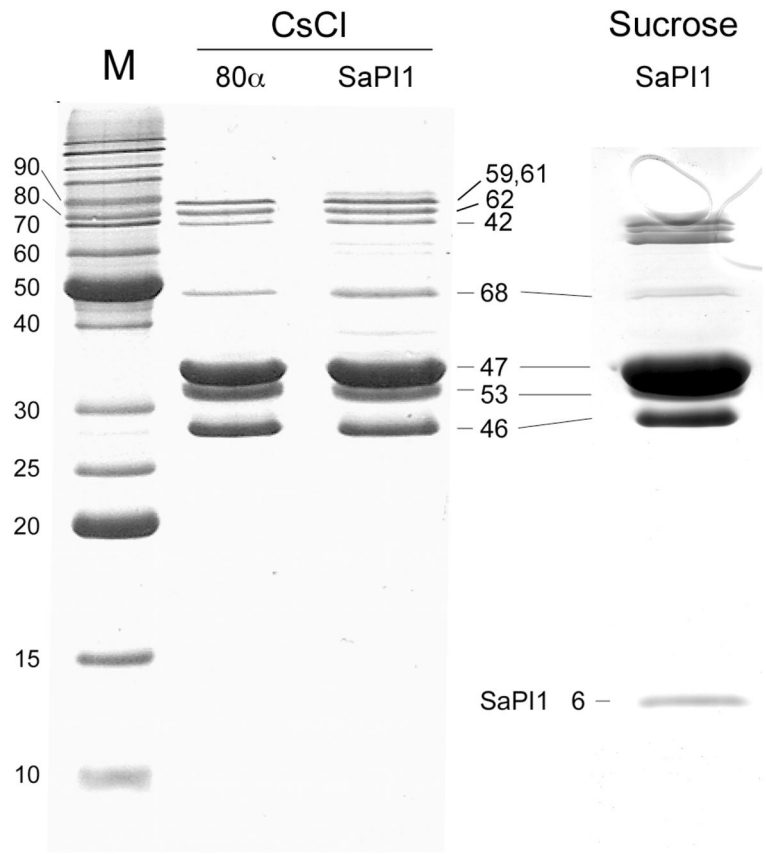
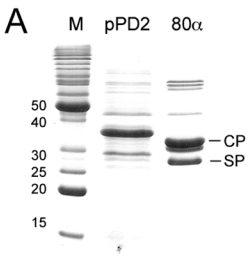


Figure 3.
SDS-PAGE of the 80α and SaPI1 CsCl-purified protein fraction as well as SaPI1 procapsids purified on sucrose gradients. The identities of major bands are indicated by the gene product number.



B

	Measured mass (Da)		Theoretical	Sequence
	80a	SaPI1	mass (Da)	(* = cleavage site)
SP, gp46	21,700.5	21,701.1	21,700.9	MEENKLKFNLQFF*ADQSDDPDEPGGDGKKGNPDKKEND...
CP, gp47	35,062.2	35,061.9	35,062.9	MEQTQKLKLNLQHF*ASNNVKPQVFNPDNVMMHEKKDGT...

Figure 4.

(A) SDS-PAGE of capsid and scaffolding proteins expressed from the clone pPD2 in *E. coli* compared to the same proteins from 80α procapsids. The bands corresponding to gp46 (SP) and gp47 (CP) are indicated on the 80α lane. (B) Table of measured masses of gp46 and gp47 in 80α and SaPI1 procapsids, compared to the theoretical mass for the cleaved protein, as well as the N-terminal sequence of the two proteins. The cleavage site is indicated by an asterisk (*).

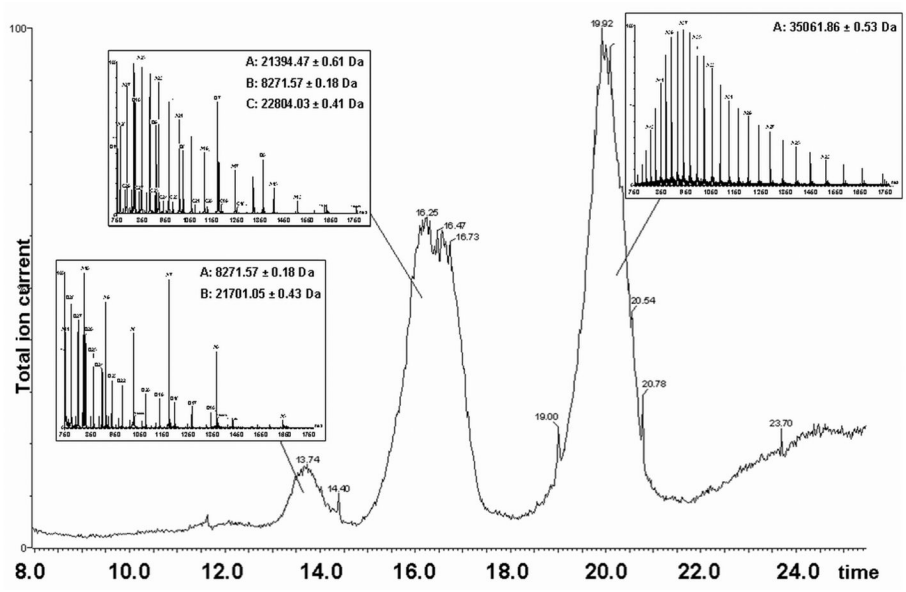


Figure 5.

Reverse phase chromatography and ESI-MS of SaPI1 procapsids. The total ion current of the MS detector is plotted against elution time in minutes. The ion current does not accurately reflect the abundance of the measured ionic species. Peaks and their corresponding spectra are shown with the measured masses listed. Peak one (approximately 14 min) includes SaPI1 gp6 (theoretical mass 8,271.5 Da) and N-terminally truncated scaffolding protein gp46 (21,700.9 Da); peak two also contains gp6 as well as SaPI1 gp7 (22,805.3 Da) and major tail protein gp53 (21,394.5 Da). The third peak in the chromatogram contains N-terminally truncated major capsid protein gp47 (35,062.9 Da).

Table 1

Listing of all proteins detected by ESI-MS of trypsin digests.

ORF (gp)	Accession #	#aa	MW (kDa)	Function/Location	% sequence coverage			
					CsCl fraction		Procapsids	
					80α	SaPI1	80α	SaPI1
80α proteins								
03	YP_001285317.1	301	35.6	Unknown	6.0	20.6	—	—
08	YP_001285322.1	262	30.1	Putative antirepressor	40.1	39.7	—	—
10	YP_001285324.1	58	6.7	Unknown	—	22.4	—	29.3
15	YP_001285329.1	73	8.6	Unknown	19.2	—	—	—
16	YP_001285330.1	207	23.7	Unknown	27.5	30.9	5.3	9.7
17	YP_001285331.1	142	16.0	ssDNA-binding protein	8.5	35.2	—	—
20	YP_001285334.1	256	29.7	Putative replisome organizer	12.5	—	—	—
32	YP_001285346.1	170	19.0	dUTPase	—	37.1	17.1	—
36	YP_001285350.1	128	15.0	Unknown	26.6	46.9	8.6	33.6
39	YP_001285353.1	140	16.5	Transcription activator (rinA)	—	9.3	—	—
40	YP_001285354.1	146	16.8	Small terminase subunit	—	29.5	—	—
41	YP_001285355.1	447	52.5	Large terminase subunit	—	13.0	—	—
42	YP_001285356.1	511	59.5	Portal protein	47.0	59.7	49.5	48.5
44	YP_001285358.1	331	38.5	Minor capsid protein, putative protease	37.1	38.1	30.0	28.2
46	YP_001285360.1	206	23.4	Scaffolding protein	47.6	74.8	67.0	54.4
47	YP_001285361.1	324	36.8	Major capsid protein	71.0	70.4	71.3	73.5
48	YP_001285362.1	95	10.9	Unknown	23.2	61.1	—	—
49	YP_001285363.1	110	12.8	Putative DNA packaging protein	33.6	28.2	40.0	37.3
50	YP_001285364.1	100	11.8	Unknown	20.0	41.0	44.0	—
52	YP_001285366.1	127	14.7	Unknown	28.3	31.5	30.7	—
53	YP_001285367.1	193	21.5	Major tail protein	75.6	54.9	42.5	62.7
54	YP_001285368.1	121	13.1	Putative tail assembly protein	19.0	—	12.4	—
55	YP_001285369.1	114	13.6	Putative tail assembly protein	42.1	—	—	—
56	YP_001285370.1	1154	125.8	Tail tape measure protein	19.2	11.6	—	2.1
58	YP_001285372.1	315	37.1	Likely minor tail protein	61.3	51.7	5.7	35.2
59	YP_001285373.1	633	71.0	Lipase, likely tail tip	53.2	47.6	25.9	30.3
61	YP_001285375.1	636	73.7	Minor tail protein	24.7	67.8	28.9	54.9
62	YP_001285376.1	607	66.8	Minor tail protein, baseplate	37.9	67.2	35.0	56.2
64	YP_001285378.1	125	14.1	Likely minor tail protein	33.6	54.4	—	—
66	YP_001285380.1	99	11.8	Likely minor tail protein	—	17.3	—	—
67	YP_001285381.1	632	72.0	Cell wall hydrolase, likely tail tip	22.3	29.9	5.1	—
68	YP_001285382.1	390	43.8	Putative tail fiber protein	70.8	76.2	36.9	52.6
69	YP_001285383.1	131	14.4	Likely minor tail protein	67.9	58.8	42.7	60.3
71	YP_001285385.1	484	53.8	Lysis protein	13.6	—		—
SaPI1 proteins								
5	AAC28956.2	175	20.6	Putative accessory capsid protein	—	32.6	—	—
6	AAC28957.2	72	8.3	Putative capsid size determinator	—	91.7	—	91.7
7	AAC28958.2	192	22.8	Putative capsid size determinator	—	79.7	—	9.9
11	AAL67613.1	131	15.2	Unknown	—	16.8	—	—
18	AAL67617.1	106	12.6	Unknown	—	36.8	—	—

Table 1 Legend: The ORF numbering follows that in the 80α and SaPI1 GenBank entries, NC_009526 and U93688, respectively. Accession numbers for individual loci follow. Known or putative functions and locations are listed, based on the information listed in the GenBank entry, homology detected by BLAST search, inferred by location in the genome (tail genes) and experimental evidence

from the present work and others [8,17]. The percentage coverage of peptides detected by MS is given for each protein found in 80α and SaPI1 CsCl purified protein fraction and sucrose-purified procapsids. (−), not detected.

Table 2

Oligonucleotide Primers

Primer	Sequence (5′-3′) restriction sites indicated by italics
SMT34	GACT*CATATG*GAACAAACACAAAAATTAAAAT
SMT35	GACT*GGATCC*TTATTAAACTTCTCCTGGAACT
SMT51	ACA*CCATGG*GCATACAGATATTCTCTGGA
SMT52	ACA*AAGCTT*GTGGATGATATACCGTTAGAG
SMT53	ACA*AAGCTT*CGCTTGTTTTGCCGTTAA
SMT54	ACA*CCATGG*CAATATGCAGGAGATTTCAAG
SMT61	TCAG*GGATCC*GAAGGAGATATCTCATGGAAGAAAATAAACTTAAG
SMT62	CTTC*GTCGAC*TTAAATGCCTCCGTTAATTTTTAA