Published in final edited form as:

J Mol Biol. 2011 September 30; 412(4): 710–722. doi:10.1016/j.jmb.2011.07.036.

The *Staphylococcus aureus* pathogenicity island 1 protein gp6 functions as an internal scaffold during capsid size determination

Altaira D. Dearborn¹, Michael S. Spilman^{1,2}, Priyadarshan K. Damle³, Jenny R. Chang^{1,4,*}, Eric B. Monroe¹, Jamil S. Saad¹, Gail E. Christie³, and Terje Dokland^{1,†}

¹Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294

²Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294

⁴Department of Biology, University of Alabama at Birmingham, Birmingham, AL 35294

³Department of Microbiology and Immunology, Virginia Commonwealth University, Richmond, VA 23298

Abstract

Staphylococcus aureus pathogenicity island 1 (SaPI1) is a mobile genetic element that carries genes for several superantigen toxins. SaPI1 is normally stably integrated into the host genome, but can become mobilized by “helper” bacteriophage 80α, leading to the packaging of SaPI1 genomes into phage-like transducing particles that are composed of structural proteins supplied by the helper phage, but having smaller capsids. We show that the SaPI1-encoded protein gp6 is necessary for efficient formation of small capsids. The NMR structure of gp6 reveals a dimeric protein with a helix-loop-helix motif similar to that of bacteriophage scaffolding proteins. The gp6 dimer matches internal densities that bridge capsid subunits in cryo-EM reconstructions of SaPI1 procapsids, suggesting that gp6 acts as an internal scaffolding protein in capsid size determination.

Keywords

bacteriophage; mobilization; virus assembly; NMR spectroscopy; cryo-electron microscopy

INTRODUCTION

Mobile genetic elements, such as bacteriophages, plasmids and pathogenicity islands, are commonly involved in horizontal gene transfer and genome evolution in bacteria ^{1,2}. *Staphylococcus aureus* is normally a harmless inhabitant of human skin and mucosal epithelia, but is also associated with serious systemic and local infections ³. A number of virulence determinants and antibiotic resistance genes have been linked to mobile genetic

© 2011 Elsevier Ltd. All rights reserved.

Contact information: Terje Dokland, Department of Microbiology, University of Alabama at Birmingham, 845 19th St South, BBRB 311, Birmingham, AL 35294, dokland@uab.edu, Tel: (205) 996 4502, Fax: (205) 996 2667.

*Current address: Dept. of Medicinal Chemistry, University of Washington, Seattle, WA 98195

†Communicating author

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.

elements traded within the staphylococcal population⁴, such as the staphylococcal cassette chromosome *mec* (SCC*mec*) elements that carry genes for antibiotic resistance⁵. Bacteriophage ϕ ETA2 carries the gene for exfoliative toxin A⁶ and bacteriophage ϕ Sa3USA carries genes (*lukGH*) for the leukocidin that aids in the selective lysis of neutrophils^{4, 7}.

S. aureus pathogenicity islands (SaPIs) are a family of 14–27 kb mobile genetic elements that harbor genes for superantigen toxins, antibiotic resistance genes and other virulence factors involved in host invasion^{8, 9}. Acquisition of up to three SaPIs in combination with other mobile genetic elements has led to the proliferation of aggressive bacterial strains that are resistant to multiple antibiotics^{10–13}. Certain SaPIs have also been shown to pass intergenerically between *Staphylococcus aureus* and the food pathogen *Listeria monocytogenes*¹⁴. A detailed description of the mechanisms by which such elements are mobilized is therefore critically important to understand the spread and establishment of pathogenicity in *S. aureus*.

SaPI mobilization is intimately associated with specific bacteriophages (helper phages)^{9, 15}. Though normally stably integrated, SaPIs are induced when a helper phage enters the lytic cycle, leading to the production of phage-like transducing particles that carry the SaPI genome, but are comprised of structural proteins supplied by the helper phage^{16–18}. One of the more well characterized SaPIs, SaPI1, has a 15.2 kb genome encoding 21 open reading frames (ORFs), including genes for the toxic shock syndrome toxin and enterotoxins K and Q^{9, 19}. SaPI1 can be mobilized by helper phage 80 α , a temperate, tailed double-stranded (ds) DNA bacteriophage in the *Siphoviridae* family (order *Caudovirales*)²⁰. 80 α can act as the helper phage for several additional SaPIs, including SaPI2, SaPIbov1 and SaPIbov2⁹. Bacteriophage 80 α capsids have *T*=7 icosahedral symmetry and are made of 415 copies of an HK97-like capsid protein, the product of gene 47 (gp47)²¹. Like other dsDNA phages, the 80 α capsid is assembled as an empty precursor, called the procapsid. Procapsid assembly requires a scaffolding protein (gp46) that acts as a chaperone and forms an internal core in the procapsids^{16, 21}. DNA is packaged into the procapsids by a headful mechanism through the portal vertex, leading to scaffolding removal and capsid expansion²¹. The DNA packaging is carried out by the terminase complex, which consists of oligomers of the small (TerS) and large (TerL) terminase subunits (gp40 and gp41, respectively).

Specificity between 80 α and SaPI1 is manifested at multiple levels. During mobilization, the 80 α protein Sri (gp22) acts as an antirepressor for the SaPI1 repressor Stl (SaPI1 gp22)^{22, 23}. Derepression is followed by excision and replication, which are independent of 80 α gene products²⁴, but SaPI1 DNA packaging requires a SaPI1-encoded TerS protein (gp3)²⁵. Strikingly, SaPI1 capsids are smaller (46 nm) than the regular 80 α capsids (63 nm), reflecting a corresponding difference in genome size^{16, 21, 26}. In the absence of SaPI1 TerS, only partial 80 α genomes are packaged²⁵. This process of capsid size determination resembles that of the *Escherichia coli* helper/satellite phage system P2/P4, wherein an external scaffolding protein, Sid, encoded by the P4 satellite, redirects the assembly of the P2 helper phage capsid protein, gpN, from a *T*=7 pathway to one resulting in a smaller *T*=4 capsid^{27–29}. By analogy, it is assumed that SaPI1 encodes one or more factors that shunt the 80 α structural proteins to an alternate assembly pathway.

We previously showed that SaPI1 procapsids contained the SaPI1-encoded protein gp6 (8.3kDa) in addition to a full complement of 80 α -encoded structural proteins¹⁶. Gp6 was present in an estimated 48 copies per SaPI1 procapsid. Another SaPI1-encoded protein, gp7 (22.8kDa) was present as a minor component¹⁶. The genes corresponding to the equivalent two proteins were implicated in size determination in SaPIbov1²⁴. In SaPI1, these two proteins are thus likely candidates for the presumed capsid size determination factors. Both

gp6 and gp7 are adjacent and co-conserved among many members of the SaPI family⁹. Here, we show that the SaPI1 protein gp6 plays a critical role in the efficient formation of small, isometric capsids from 80 α -encoded structural proteins. The NMR structure of gp6 shows a dimer reminiscent of bacteriophage scaffolding proteins^{30, 31}. These dimers match internal densities in the cryo-electron microscopy (EM) reconstruction of SaPI1 procapsids, suggesting that gp6 acts as an internal scaffolding protein during capsid size determination.

RESULTS

SaPI1 gp6 is required for the efficient formation of small capsids

We previously identified gp6 as one of the prime candidates for a size determination factor for SaPI1¹⁶. Its C-terminal 10 residues are homologous to those of the 80 α scaffolding protein, gp46 (Fig. 1). To investigate the possible role of gp6 in scaffolding and capsid size determination, *orf6* was deleted from the SaPI1-harboring 80 α lysogen RN10628, generating the strain ST100 (Table S1). When this strain was induced with mitomycin C, it produced procapsids and virions, similar to a normal wildtype induction, that were separated on CsCl gradients. The protein composition of these particles was normal, and although the virions did not contain gp6, nor gp7, as expected for mature transducing particles¹⁷, gp7 was found associated with the procapsid fraction by mass spectrometry (MS) analysis.

Cryo-EM of the SaPI1 Δ *orf6* virion fraction revealed a mixture of empty and full particles in various sizes and shapes (Fig. 2a). The isometric shells clustered in two size classes centered on 45.3×47.8 nm (25%) and 59.7×63.8 nm (62%), consistent with $T=4$ and $T=7$ icosahedral symmetry, respectively (Fig. 2b). (Most of the difference in x and y for these nominally isometric particles is due to the angularity of the shell.) About 13% of the closed shells were clearly non-isometric or intermediate in size, implying a loss of icosahedral symmetry^{32, 33}. About 57% of the isometric shells were empty, with a somewhat larger proportion of empties (60%) among the large shells compared to 49% for the small shells. The empty shells presumably arise from spontaneous ejection of DNA during or after purification, suggesting that the DNA is less stable. This distribution of shapes and sizes contrasted sharply with a SaPI1 wildtype induction, in which virtually all the particles were small and only 2% were empty (Fig. 2c and d). This experiment shows that gp6 is essential to promote efficient formation of small isometric capsids, but is not solely responsible for size determination.

NMR structure of gp6

SaPI1 gp6 is a 72 amino acid protein predicted to have a predominantly α -helical structure (Fig. 3a). Full-length, His₆-tagged SaPI1 gp6 protein was characterized by sedimentation equilibrium analytical ultracentrifugation and found to exist as a stable dimer. The protein was also subjected to limited trypsin digestion, followed by MS analysis, which revealed preferential cleavage at several sites clustered near the C-terminus, indicating that this region is flexible (Fig. 3a).

The structure of the gp6 dimer was determined by NMR methods. Proton, nitrogen and carbon chemical shifts were assigned using a combination of 2D (¹H-¹⁵N HSQC and ¹H-¹³C HMQC), 3D and 4D (¹⁵N-, ¹³C-, ¹⁵N/¹³C-, ¹³C/¹³C- and ¹³C-edited/¹²C-double-half-filtered HMQC-NOESY) data (Fig. 3b). Numerous intermolecular NOE crosspeaks were observed between hydrophobic and aromatic residues located on the dimerization interface (Fig. 3c and d). A total of 635 NOE and 96 angle restraints were used to calculate an ensemble of 50 low-energy structures with an average pairwise r.m.s. deviation of 0.28 Å (Figs. 3e, 4a; Table 1). The gp6 dimer consists of a four-helix bundle (Fig. 4a), in which each monomer contributes two amphipathic α -helices, forming a hydrophobic core. The first helix (α 1)

includes residues 7–16, and the second helix ($\alpha 2$) includes residues 22–39 (Fig. 3a, 4b). The dimer is stabilized by extensive intermolecular contacts between the two helices, while intramolecular contacts are less prominent, consistent with the observed tight dimer in solution. The intervening residues (17–21) form a short loop (FGLNE) that is stabilized by close packing of the bulky hydrophobic residues F17 and L19 into the core (Fig. 4c). Although secondary structure prediction indicates an additional C-terminal α -helical region between residues 40 and 72 (Fig. 3a), this sequence is flexible and does not adopt a unique structure in solution. This is likely a result of the unusual right-handed twist of the bundle, which results in the $\alpha 2$ helices sprawling at a 60° angle, precluding further stabilizing intermolecular interactions between the two monomers. The C-termini, however, might adopt an ordered α -helical structure upon binding to the capsid (see below).

The helix-loop-helix structure of gp6 is remarkably similar to that of *Bacillus* phage $\phi 29$ internal scaffolding protein gp7³⁰ (Fig. 4d). The 92-residue $\phi 29$ scaffolding protein is a dimer with much the same arrangement as gp6, except that the four-helix bundle has a left-handed twist and continues as a left-handed coiled coil before terminating in a 15-residue flexible tail. This flexible sequence is responsible for capsid binding in the $\phi 29$ procapsid³⁴. Like gp6, the capsid binding domain of the *Salmonella* phage P22 scaffolding protein gp8 also has a helix-loop-helix motif³¹. However, this domain is only part of a much larger scaffolding protein and existed as a monomer in solution.

Structure of the SaPI1 procapsid

SaPI1 procapsids were produced by induction of the SaPI1-containing 80 α lysogen ST63, in which the *terS* genes of both 80 α and SaPI1 had been deleted²⁵. The lack of functional terminase enzymes leads to an accumulation of SaPI1-sized procapsids that contain a full complement of scaffolding protein as well as gp6 and gp7, although MS analysis suggested that the amount of gp7 in the procapsids is very low¹⁶. The SaPI1 procapsids were subjected to cryo-EM analysis and three-dimensional reconstruction as previously described²¹.

By cryo-EM, the SaPI1 procapsids appear as 37 nm round particles having a ≈ 4 nm thick shell with a serrated outline (Fig. 5a). They are smaller, but otherwise similar in appearance to the 51 nm diameter 80 α procapsids²¹. A total of 3944 particles were selected for 3D reconstruction, using the program EMAN³⁵, resulting in a density map at a resolution of 9.9 Å by the 0.5 Fourier shell correlation criterion.

The SaPI1 procapsid has $T=4$ icosahedral symmetry, as expected for a capsid of this size (Fig. 5b), incorporating 240 copies of gp47, less five copies that would be replaced by the portal protein at one fivefold vertex. This asymmetry is averaged out by the icosahedral averaging applied in the reconstruction procedure. These 240 copies are found in four different environments, termed A, B, C and D. The capsid protein is thus clustered into twelve A₅ pentamers located at the fivefold vertices and thirty (BCD)₂ hexamers that sit on the icosahedral twofold axes (Fig. 5c). The twofold symmetric, elongated shape of the hexamers is similar to that in 80 α , even though the hexamers are located in positions of no symmetry in the $T=7$ –80 α procapsid²¹. The hexamers and pentamers interact via two types of trivalent connections, a type 1 (CCC) contact at the icosahedral threefold axis, and a type 2 (ABD) contact formed between one pentamer and two hexamers (Fig. 5c).

Unlike bacteriophage P4, in which the Sid protein forms an external scaffold, there were no additional features that might constitute a size determination factor on the outer surface of the SaPI1 procapsid (Fig. 5b). On the inside, however, there were 120 5-nm long finger-like densities that protruded towards the center (obcentrically) from the inner surface of the capsid (Fig 5d)²¹. The density for the finger-like protrusions was well defined, but at a

lower level than the shell itself, indicating a protein that is well ordered, but has low occupancy.

Subvolumes enclosing a threefold and a fivefold axis were extracted from both the SaPI1 and 80 α procapsid reconstructions and superimposed by cross correlation. The capsomers match well locally, and the finger-like protrusions in SaPI1 stand out clearly against the 80 α procapsid density, forming wishbone-like features whose legs straddle across and connect with the capsid protein in two adjacent capsomers (Fig. 5e, f). There are two types of protrusions: One type connects C and D subunits in a cluster around the icosahedral threefold axis. The other type connects A and B subunits around the pentamers and has weaker, less well defined density (Fig. 5c and e). There are no equivalent features in 80 α . Indeed, even when the density cutoff level was lowered until an internal scaffolding core appeared, separated from the shell by a 2.5 nm gap, there were no visible finger-like protrusions in the 80 α procapsid map (Fig. 5e).

As in the SaPI1 procapsid, protrusions on the inner surface of the shell were observed in P22³⁶. These were attributed to the P22 scaffolding protein, gp8. However, whereas the P22 scaffolding protein bridges the N-arms and A domains of two adjacent capsid protein monomers asymmetrically within a single capsomer, the densities in SaPI1 straddle a local twofold axis intercapsomerically and protrude further from the inner surface of the shell.

Modeling gp6 in the SaPI1 procapsid

We previously built a model for the gp47 capsid protein based on HK97 gp5²¹. This model was placed into the SaPI1 procapsid density at the A, B, C and D positions. Adjustments were made for the relative orientation of the A and P domains and P and E loops that were variable between the different non-equivalent subunits in 80 α ²¹. Connections were remodeled to accommodate the domain shifts; the models were otherwise unchanged. Ultimately, there were only small differences between gp47 in 80 α and SaPI1 procapsids.

From this model, a mask was generated and applied to the map. The resulting map showed only density not attributable to the modeled portion of gp47, leaving only the finger-like protrusions (Fig. 5f). The resolution of the map was sufficient to discern four ridges twisting around the surface of the protrusions, two of which were contiguous with the legs of the wishbones that connected with the capsid surface. Although 80 α gp46 is also present in the SaPI1 procapsids, a diffuse inner core similar to that observed in 80 α procapsids²¹ only appeared after lowering the cutoff level well beyond that used to visualize the protrusions.

The shape and size of these protrusions matched those of the gp6 dimer structure, which was then fitted into the densities by aligning the twist of the four-helix bundle to the surface ridges of the protrusions. By analogy with ϕ 29, the dimer was oriented so that the C-termini pointed towards the shell, where the disordered portions could fit into the legs of density, although we could not conclusively make this assignment based on the structure of the protrusions alone. Nevertheless, the tubular shape of the legs is consistent with the presence of an α -helix and its length is sufficient to accommodate at least 18 amino acids of the 33-residue predicted α -helical region (Fig. 3a). The remaining 15 amino acids pack against the N-terminal arms of gp47 from two separate capsomers (Fig. 5f). The N-terminal arms were disordered in 80 α virions, but are ordered in both SaPI1 and 80 α procapsids, suggesting that both gp6 and gp47 become ordered upon binding²¹. This binding mode differs from that in P22, where the scaffolding protein bound to the capsid as a monomer³⁶. Previous studies had also suggested that the tip of the loop forms the capsid binding domain in P22³⁷. Hence, we conclude that the protrusions observed in the SaPI1 procapsids are attributable to the gp6 dimers.

DISCUSSION

The structural and functional evidence presented here strongly suggests that SaPII gp6 is an internal scaffolding protein utilized during the assembly of the smaller SaPII procapsids, and its presence is required to prevent procapsid shape promiscuity. Most dsDNA bacteriophages require a scaffolding protein for the assembly of a functional procapsid^{38, 39}. The majority of scaffolding proteins are internal; the exception being the external scaffolds of P4 and in the *Microviridae*^{40, 41}. All scaffolding proteins analyzed so far have a predominantly α -helical structure^{30, 31}. Scaffolding proteins may act as chaperones that prevent the formation of aberrant interactions between the capsid proteins, like the internal scaffolding proteins of P22 and P2 (gpO) and, by analogy, 80 α . These scaffolding proteins do not form permanent interactions with the capsid, but act transiently, presumably by restricting the conformational potential of the capsid proteins^{38, 39, 42}. In contrast, the P4 external scaffolding protein Sid forms a cage that stably associates with and directly limits the size of the underlying procapsid^{41, 43}.

Both gp6 and Sid work by restricting the conformational potential of their respective capsid proteins, thus resulting in the formation of small capsids. In the SaPII procapsid, as in 80 α , the gp47 capsid protein is clustered into hexamers and pentamers (Fig. 5b and c) that form prominent trivalent interactions²¹. The gp6 dimers bridge capsid protein subunits between two adjacent hexamers or between hexamers and pentamers (Fig. 5c). It is conceivable that this dimerization of capsomers leads to an increased curvature that is only compatible with the formation of a smaller $T=4$ shell. This binding mode differs from the intra-capsomeric binding mode of the P22 scaffolding protein³⁶.

In spite of an obvious resemblance between gp6 and the ϕ 29 scaffolding protein, gp7³⁰, the difference in hand between the two dimers suggests that the two proteins have evolved convergently rather than from a common ancestor. Most likely, their mode of action is also different; indeed, the ϕ 29 scaffolding protein acts as a nucleator of assembly from the portal protein⁴⁴, and it is unclear how and to what extent it also interacts with the capsid.

The observed homology between the presumed capsid-binding C-termini of gp6 and gp46 (Fig. 1) suggests that the two proteins may bind to the N-terminal arm of gp47 in the same way (Fig. 5f). However, although gp46 remains associated with both 80 α and SaPII procapsids¹⁶, it does not form stable interactions with gp47 and was not observed in association with gp47 in reconstructions of 80 α procapsids²¹.

Formation of isometric particles is strongly impaired, but not completely abolished, in the absence of gp6 (Fig. 2), suggesting the involvement of other SaPII-encoded factors in the assembly process. Without gp6, these factors have a detrimental effect on the formation of isometric particles, possibly by interfering with gp46. The most likely candidate for this activity is gp7, which might be involved in the incorporation of gp6 into the procapsids. The necessity for more than one size determination factor in SaPII differs from the P2/P4 system, where Sid alone was sufficient⁴³. Indeed, Sid does not even require the P2 scaffolding protein gpO to induce the formation of small ($T=4$) procapsids⁴⁵.

Proteins corresponding to gp6 and gp7 are found in many, but not all, SaPIs (Fig. 1), including ones associated with serious pathogenic conditions, such as toxic shock, gastric distress and impetigo^{8, 9}. The two genes always appear in tandem, suggesting that they function in concert to promote the formation of isometric $T=4$ particles. Capsid size switching is not a requirement for the mobilization of SaPIs^{9, 20}. However, the fact that gp6 and gp7 are co-conserved in several SaPIs suggests that the size determination confers some evolutionary advantage, presumably by preventing infectious helper phage production. This is in addition to the effects of other factors that interfere with helper phage replication and

packaging^{22, 24, 25}. Further studies will be necessary to better define the role of capsid size determination in helper phage interference and to understand the role of other SaPII proteins, such as gp7, in SaPII mobilization and size determination.

EXPERIMENTAL PROCEDURES

Production of SaPII particles

S. aureus strains used in this study (Table 2) were derivatives of 80 α lysogens RN10616 or RN10628, the latter carrying SaPII *tst::tetM*²⁵. In-frame deletion of SaPII or 80 α genes were made by allelic exchange using derivatives of pMAD⁴⁶, as described previously¹⁶, using primers shown in Table S1. SaPII “wildtype” virions were produced from *S. aureus* strain ST65, which carries a deletion of 80 α *orf44*. This deletion, which affects a gene product believed to be involved in phage DNA ejection²⁰, results in increased DNA stability in SaPII particles. SaPII procapsids were produced from ST63, in which both the 80 α *terS* and SaPII *terS* genes had been deleted²⁵. ST100 carries a deletion of SaPII *orf6*. All 80 α lysogens were grown in CY broth at 32°C and induced with mitomycin C as previously described⁴⁷. The particles were purified by PEG precipitation, chloroform extraction and banding on CsCl gradients²¹. SaPII procapsids were further separated on 10–40% sucrose gradients.

Cryo-electron microscopy

For cryo-EM, the sample was applied to C-flat holey film, flash frozen, transferred to a Gatan 626 cryo-specimen holder and observed in an FEI Tecnai F20 electron microscope operated at 200 kV at a magnification of 62,000x with typical defocus values from 0.8 to 2.4 μ m^{21, 48}. Images were collected on a Gatan Ultrascan 4000 CCD camera or Kodak SO-163 film.

Cloning, expression, and purification of gp6

SaPII *orf6* was amplified from SaPII DNA by PCR, incorporating an N-terminal His₆ tag, and inserted into the vector pET21a (EMD Biosciences). All cloning was done in *E. coli* strain DH5 α while expression was done in BL21 (DE3). Cells were grown in LB with 100 μ g/ml ampicillin at 37°C and gp6 expression was induced by the addition of 0.5 mM IPTG at OD₆₀₀ = 0.6. ¹⁵N-labeled and ¹⁵N-¹³C double-labeled samples were produced by the method of Marley et al.⁴⁹. Cultures were collected by centrifugation, resuspended in 100 mM Tris pH 7.4 with 200 mM NaCl, lysed in a high pressure cell disruptor (Avestin, Ottawa), clarified by centrifugation and filtered through a 0.45 μ m syringe filter. His₆-gp6 was purified over a 6 mL Ni-NTA gravity-flow column (Qiagen) using a 20–500mM imidazole step gradient in 100 mM Tris pH 7.4, 200 mM NaCl. His₆-gp6 eluted at 100 mM imidazole and was dialyzed against 10 mM sodium phosphate buffer (pH 7.0), 200mM NaCl, and concentrated to 10 mg/mL (1.02 mM) in a centrifugal concentrator, typically yielding 12 mg protein per liter of starting culture. Untagged gp6 protein was purified by ion exchange over a Hi-Q column (Bio-Rad) equilibrated in 100mM Tris pH 7.4, 50mM NaCl. The gp6-containing flow-through was further purified via size exclusion chromatography (Hi-prep 16/60, Amersham).

Trypsin digestion and mass spectrometry

Partial tryptic digests of His₆-gp6 were produced by incubating 200 μ g of protein (at 2.5 mg/mL) with 0.25–1.0 U of trypsin (Roche), overnight at 4 °C. Digestions were stopped with 1 mM EDTA, acidified with 1% formic acid, injected into a C4 trap and eluted by a 5–95% acetonitrile gradient into a Micromass LCT ESI-TOF mass spectrometer. Masses were calculated using MaxEnt. For proteomic identification of proteins, procapsids were

disrupted in 1% formic acid and digested for 2 min with 0.5 mg/ml pepsin, separated by reverse phase chromatography and analyzed by MS/MS in an LTQ-FT (Thermo Finnigan) spectrometer. Fragmentation spectra were annotated by a SEQUEST search against an 80 α /SaPI1 database.

NMR spectroscopy and structure calculation

NMR data were collected with a Bruker Avance II (700 MHz ^1H) spectrometer equipped with cryoprobes, processed with NMRPIPE⁵⁰ and analyzed with NMRVIEW⁵¹. A combination of two-, three- and four-dimensional NOESY data were obtained for natural abundance, ^{15}N - and ^{15}N -, ^{13}C -labeled protein samples at 30 °C. Protein backbone signals were assigned using standard triple resonance methods (HNCA, HNCOCACB, HNCACB, HNCOCACB), and side chain signals were assigned from 3D and 4D ^{15}N -, ^{13}C -, and $^{15}\text{N}/^{13}\text{C}$ -edited NOESY data^{52,53}. Intermolecular NOEs between the two halves of the dimer were assigned from 3D (^{13}C - and ^{13}C -edited/ ^{12}C -double-half-filtered and ^{15}N - and ^{15}N -edited/ ^{14}N -double-half-filtered) NOESY data collected on a 50:50 mixture of ^{15}N -, ^{13}C -labeled and unlabeled gp6.

Structures were calculated in torsion angle space with CYANA (<http://www.cyana.org>) starting from random initial angles. Upper interproton distance bounds of 2.7, 3.3 and 5.0 Å (with appropriate corrections for pseudoatoms) were employed for NOE cross peaks of strong, medium, and weak intensity respectively, which were qualitatively determined following intensity normalization of the different NOE data sets. Backbone hydrogen bond and chemical shift-based torsion angle restraints were employed. An ensemble of 20 low-energy structures was validated by PROCHECKNMR (Table 1). Structure figures were generated with Chimera⁵⁴.

3D reconstruction and modeling of SaPI1 procapsids

Film images were digitized with a Nikon CoolScan 9000ED film scanner at a raster of 4000 dpi and binned to 2000 dpi, corresponding to a pixel size of 2.048 Å, using a median value filter. Image processing, including particle picking, CTF correction and three-dimensional reconstruction were done primarily using the EMAN package^{21,35}. Map visualization and manipulation was done using UCSF Chimera^{54,55}.

The gp47 model from 80 α ²¹ was fitted into the SaPI1 procapsid map by a combination of manual and correlation-based fitting and geometric optimization, using the programs O⁵⁶ and Chimera⁵⁴. The NMR structure of the gp6 dimer was manually fitted into the map after subtracting the contribution of the capsid protein.

Accession numbers

The coordinates for the gp6 dimer has been deposited to the Protein Data Bank with PDB ID 2L8T and RCSB ID rscb102100. ^1H , ^{13}C and ^{15}N resonance assignments have been deposited to BioMagResBank with BMRB ID 17425. The electron density map for the SaPI1 procapsid reconstruction has been deposited to the Electron Microscopy Data Bank with accession code EMD-1869.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Drs. Peter Prevelige and Matthew Renfrow for technical support and access to MS instrumentation at UAB. This work was supported by The National Institutes of Health grants R21 AI071982 and R01 AI083255 to T.D.; R21 AI067654, R56 AI081837 and an A.D. Williams Trust and Baruch Foundation Trust Grant-In-Aid to G.E.C. E.B.M. was supported by a Ruth L. Kirschstein National Research Service award, F32 GM087994.

References

1. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol.* 2005; 3:722–732. [PubMed: 16138100]
2. Ziebuhr W, Ohlsen K, Karch H, Korhonen T, Hacker J. Evolution of bacterial pathogenesis. *Cell Mol Life Sci.* 1999; 56:719–728. [PubMed: 11212331]
3. Gordon RJ, Lowy FD. Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clin Infect Dis.* 2008; 46:S350–S359. [PubMed: 18462090]
4. Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet.* 2006; 367:731–739. [PubMed: 16517273]
5. Katayama Y, Ito T, Hiramatsu K. A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother.* 2000; 44:1549–1555. [PubMed: 10817707]
6. Yamaguchi T, Hayashi T, Takami H, Nakasone K, Ohnishi M, Nakayama K, Yamada S, Komatsuzawa H, Sugai M. Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. *Mol Microbiol.* 2000; 38:694–705. [PubMed: 11115106]
7. Ventura CL, Malachowa N, Hammer CH, Nardone GA, Robinson MA, Kobayashi SD, DeLeo FR. Identification of a novel *Staphylococcus aureus* two-component leukotoxin using cell surface proteomics. *PLoS ONE.* 2010; 5:e11634. [PubMed: 20661294]
8. Novick RP. Mobile genetic elements and bacterial toxins: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. *Plasmid.* 2003; 49:93–105. [PubMed: 12726763]
9. Novick RP, Christie GE, Penades JR. The phage-related chromosomal islands of Gram-positive bacteria. *Nat Rev Microbiol.* 2010; 8:541–551. [PubMed: 20634809]
10. Diep BA, Carleton HA, Chang RF, Sensabaugh GF, Perdreau-Remington F. Roles of 34 virulence genes in the evolution of hospital- and community-associated strains of methicillin-resistant *Staphylococcus aureus*. *J Infect Dis.* 2006; 193:1495–1503. [PubMed: 16652276]
11. Diep BA, Otto M. The role of virulence determinants in community-associated MRSA pathogenesis. *Trends Microbiol.* 2008; 16:361–369. [PubMed: 18585915]
12. Hu DL, Omoe K, Inoue F, Kasai T, Yasujima M, Shinagawa K, Nakane A. Comparative prevalence of superantigenic toxin genes in methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* isolates. *J Med Microbiol.* 2008; 57:1106–1112. [PubMed: 18719180]
13. O'Neill AJ, Larsen AR, Skov R, Henriksen AS, Chopra I. Characterization of the epidemic European fusidic acid-resistant impetigo clone of *Staphylococcus aureus*. *J Clin Microbiol.* 2007; 45:1505–1510. [PubMed: 17344365]
14. Chen J, Novick RP. Phage-mediated intergeneric transfer of toxin genes. *Science.* 2009; 323:139–141. [PubMed: 19119236]
15. Novick RP, Subedi A. The SaPIs: mobile pathogenicity islands of *Staphylococcus*. *Chem Immunol Allergy.* 2007; 93:42–57. [PubMed: 17369699]
16. Poliakov A, Chang JR, Spilman MS, Damle PK, Christie GE, Mobley JA, Dokland T. Capsid size determination by *Staphylococcus aureus* pathogenicity island SaPII involves specific incorporation of SaPII proteins into procapsids. *J Mol Biol.* 2008; 380:465–475. [PubMed: 18565341]
17. Tallent SM, Langston TB, Moran RG, Christie GE. Transducing particles of *Staphylococcus aureus* pathogenicity island SaPII are comprised of helper phage-encoded proteins. *J Bacteriol.* 2007; 189:7520–7524. [PubMed: 17693489]

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. 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]
20. Christie GE, Matthews AM, King DG, Lane KD, Olivarez NP, Tallent SM, Gill SR, Novick RP. The complete genomes of *Staphylococcus aureus* bacteriophages 80 and 80 alpha - implications for the specificity of SaPI mobilization. *Virology.* 2010; 407:381–390. [PubMed: 20869739]
21. Spilman MS, Dearborn AD, Chang JR, Damle PK, Christie GE, Dokland T. A conformational switch involved in maturation of *Staphylococcus aureus* bacteriophage 80alpha capsids. *J Mol Biol.* 2011; 405:863–876. [PubMed: 21129380]
22. Tormo-Mas MA, Mir I, Shrestha A, Tallent SM, Campoy S, Lasa I, Barbe J, Novick RP, Christie GE, Penades JR. Moonlighting bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature.* 2010; 465:779–782. [PubMed: 20473284]
23. Harwich, MD, Jr. PhD thesis. Dept. of Microbiology and Immunology, Virginia Commonwealth University; Richmond, VA: 2009. Transcriptional profiling of staphylococcal bacteriophage 80α and regulatory interactions with pathogenicity island SaPI1.
24. 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]
25. Ubeda C, Olivarez NP, Barry P, Wang H, Kong X, Matthews A, Tallent SM, Christie GE, Novick RP. Specificity of staphylococcal phage and SaPI DNA packaging as revealed by integrase and terminase mutations. *Mol Microbiol.* 2009; 72:98–108. [PubMed: 19347993]
26. 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]
27. Christie GE, Calendar R. Interactions between satellite bacteriophage P4 and its helpers. *Annu Rev Genet.* 1990; 24:465–490. [PubMed: 2088176]
28. Dokland T, Lindqvist BH, Fuller SD. Image reconstruction from cryo-electron micrographs reveals the morphopoietic mechanism in the P2–P4 bacteriophage system. *EMBO J The European Molecular Biology Organization Journal.* 1992; 11:839–846.
29. 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]
30. Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BAL, McMurray CT, Anderson DL, Rossmann MG. Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly. *Nat Struct Biol.* 2003; 10:572–576. [PubMed: 12778115]
31. Sun Y, Parker MH, Weigele P, Casjens S, Prevelige PE, Krishna NR. Structure of the coat protein-binding domain of the scaffolding protein from a double-stranded DNA virus. *J Mol Biol.* 2000; 297:1195–1202. [PubMed: 10764583]
32. Kellenberger E. Form determination of the heads of bacteriophages. *Eur J Biochem.* 1990; 190:233–248. [PubMed: 2194799]
33. Caspar DLD, Klug A. Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol.* 1962; 27:1–24. [PubMed: 14019094]
34. Fu CY, Morais MC, Battisti AJ, Rossmann MG, Prevelige PEJ. Molecular dissection of phi29 scaffolding protein function in an in vitro assembly system. *J Mol Biol.* 2007; 366:1161–1173. [PubMed: 17198713]
35. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol.* 1999; 128:82–97. [PubMed: 10600563]
36. Chen DH, Baker ML, Hryc CF, Dimaio F, Jakana J, Wu W, Dougherty M, Haase-Pettingell C, Schmid MF, Jiang W, Baker D, King JA, Chiu W. Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci U S A.* 2011
37. Weigele PR, Sampson L, Winn-Stapley D, Casjens SR. Molecular genetics of bacteriophage P22 scaffolding protein's functional domains. *J Mol Biol.* 2005; 348:831–844. [PubMed: 15843016]

38. Dokland T. Scaffolding proteins and their role in viral assembly. *Cell Mol Life Sci.* 1999; 56:580–603. [PubMed: 11212308]
39. Fane BA, Prevelige PE. Mechanism of scaffolding-assisted viral assembly. *Adv Prot Chem.* 2003; 64:259–299.
40. Dokland T, McKenna R, Ilag LL, Bowman BR, Incardona NL, Fane BA, Rossmann MG. Structure of a viral procapsid with molecular scaffolding. *Nature (London).* 1997; 389:308–313. [PubMed: 9305849]
41. Marvik OJ, Dokland T, 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]
42. Chang JR, Spilman MS, Rodenburg CM, Dokland T. Functional domains of the bacteriophage P2 scaffolding protein: identification of residues involved in assembly and protease activity. *Virology.* 2009; 384:144–150. [PubMed: 19064277]
43. 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]
44. Fu CY, Uetrecht C, Kang S, Morais MC, Heck AJ, Walter MR, Prevelige PEJ. A docking model based on mass spectrometric and biochemical data describes phage packaging motor incorporation. *Mol Cell Proteomics.* 2010; 9:1764–1773. [PubMed: 20124351]
45. Wang S, Chang JR, Dokland T. Assembly of bacteriophage P2 and P4 procapsids with internal scaffolding protein. *Virology.* 2006; 348:133–140. [PubMed: 16457867]
46. 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]
47. Novick RP. Genetic systems in Staphylococci. *Methods Enzymol.* 1991; 204:587–636. [PubMed: 1658572]
48. Dokland, T.; Ng, ML. World Scientific Press. Dokland, T.; Huttmacher, DW.; Ng, ML.; Schantz, JT., editors. Singapore: 2006.
49. Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR.* 2001; 20:71–75. [PubMed: 11430757]
50. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995; 6:277–293. [PubMed: 8520220]
51. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol.* 2004; 278:313–352. [PubMed: 15318002]
52. Kay LE, Clore GM, Bax A, Gronenborn AM. Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 beta in solution. *Science.* 1990; 249:411–414. [PubMed: 2377896]
53. Wüthrich K. Determination of three-dimensional protein structures in solution by nuclear magnetic resonance: An overview. *Methods Enzymol.* 1989; 177:125–131. [PubMed: 2691844]
54. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. *J Struct Biol.* 2007; 157:281–287. [PubMed: 16963278]
55. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25:1605–1612. [PubMed: 15264254]
56. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density and the location of errors in these models. *Acta Cryst A.* 1991; 47:110–119. [PubMed: 2025413]
57. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics.* 1999; 15:305–308. [PubMed: 10320398]

Highlights

- SaPI1 gp6 is required for assembly of T=4 capsids from phage 80 α capsid protein
- SaPI1 gp6 forms a dimer that resembles the scaffolding protein of phage ϕ 29
- gp6 dimers bridge capsid protein capsomers on the inside of SaPI1 procapsids
- gp6 functions as an internal scaffolding protein during SaPI1 procapsid assembly

Figure 1.

Gp6 sequence homology. Sequence alignment of SaPI1 gp6, its homologs from several other pathogenicity and resistance islands in *S. aureus*, *S. hominis* (ShoPI1) and *S. haemolyticus* (ShPI2) and the C-terminal 136 residues of 80a gp46. Black boxes indicate identity; open boxes, sequence homology, based on the Risler matrix ⁵⁷. (GenBank accession numbers: SaPI1, U93688; SaPI2, EF010993; SaPI3, CP000046; ShoPI1, NZ_ACLP01000007; ShPI2, AP006716; SaRIfusB, AM292600.)

Figure 2.

Effect of deletion of gp6. **(a)** Cryo-electron micrograph of SaPI1 Δ orf6 particles. Examples of small (S), large (L) and non-isometric (N) capsids are indicated. **(b)** Size distribution plot of SaPI1 Δ orf6 particles. The short and long axes of each particle (nm) were plotted on the x and y axes, respectively. Gray circles represent empty particles, while black circles represent filled particles. Clusters corresponding to small (S) and large (L) isometric shells are demarcated (dashed lines). **(c)** Cryo-EM and **(d)** size distribution plot of SaPI1 wildtype particles. Scale bar, 100 nm.

Figure 3.

Characterization of gp6. **(a)** The predicted and observed secondary structure of gp6 are shown, with black boxes representing α -helices. Pertinent residue numbers are indicated. The wavy line represents the disordered C-terminal sequence in the NMR structure. The open triangles indicate predicted tryptic cleavage sites, while filled triangles show sites that were both predicted and found to be cleaved by MS analysis. The numbers in circles indicate the order in which cleavage was detected. **(b)** ^1H - ^{15}N HSQC spectrum of gp6, with peaks assigned to residues as labeled. Inset, magnified view of the crowded central region. **(c)** Three examples of ^{13}C -edited/ ^{12}C -double half-filtered HMQC-NOESY spectra, showing intermolecular contacts formed by the β , $\delta 1$ and $\gamma 2$ methyl groups of residues A30, L19 and

V13, respectively. **(d)** Schematic diagram of the intermolecular interactions shown in (c). The circles represent the α helices, viewed from the connecting loop (gray line). **(e)** Schematic diagram of the intra- (left) and intermolecular (right) interactions used as restraints in the structure calculation. The rectangles represent side views of the $\alpha 1$ and $\alpha 1$ helices. See also Fig. S1.

Figure 4.

Structure of gp6. **(a)** Stereo view of the superposition of 20 top solutions (gray) with the lowest energy structure colored red and blue, by chain. **(b)** Ribbon representation of the lowest energy solution of the gp6 dimer. Secondary structure elements are labeled, and the 33-residue disordered C-termini are indicated by dotted lines. **(c)** Stereo view of a detail of the dimer tip, viewed down the axis of the helical bundle. The loop residues (FGLNE) are depicted in ball-and-stick representation, colored by atom type. **(d)** Ribbon diagram of the bacteriophage φ29 gp7 scaffolding protein (PDB ID 1NO4), colored yellow and green by chain.

Figure 5.

The SaPI1 procapsid. **(a)** Cryo-electron micrograph of SaPI1 procapsids (ST63), showing a typical field of particles used for reconstruction. Scale bar, 100 nm. **(b)** Outside view of the SaPI1 procapsid in isosurface representation, colored radially from the center (red to blue). The large triangle represents one icosahedral face, with the fivefold, threefold and twofold symmetry axes indicated (pentagons, triangle and ovals, respectively). Subunits A, B, C and D of one asymmetric unit are labeled. **(c)** Schematic diagram showing the triangular face from (b). The four subunits (A, B, C and D) in one asymmetric unit are colored red, blue, green and yellow, respectively. Type 1 and 2 trimers are indicated by the filled and open triangles, respectively. The large ovals represent the threefold-associated (pink) and fivefold-associated (purple) gp6 dimers on the inside of the shell. **(d)** Inner surface of procapsid, colored radially from the center (red to blue). A magnified view of a portion of the map is shown on the right. **(e)** Superposition of a local region of the SaPI1 (solid surface, colored radially from red to blue) and 80α (mesh) procapsid reconstructions. The internal core in 80α is visible (arrow). A fivefold and a threefold symmetry axis are indicated. **(f)** Fit of the gp6 dimer in the SaPI1 procapsid map (mesh). A difference map between the gp47 model and the map is shown as a transparent solid at a higher cutoff level and clearly shows the wishbone-like structure. The NMR structure of the gp6 dimer is shown in red and blue; the predicted α-helical C-terminal extensions of gp6 are shown in turquoise and brown; Two gp47 N-terminal helices are shown in purple and orange. See also supplementary video.

Table 1

NMR and refinement statistics for gp6 protein

NMR distance and dihedral constraints	
Distance constraints	
Total NOE	635
Intra-residue	52
Inter-residue:	
Sequential ($ i - j = 1$)	153
Medium-range ($ i - j < 4$)	221
Long-range ($ i - j > 5$)	112
Intermolecular	80
Hydrogen bonds	17
	96
Total dihedral angle restraints	
ϕ	48
ψ	48
Structure statistics	
CYANA target function	0.26 ± 0.08
Maximum violations (mean and s.d.)	
Distance constraints ($\text{\AA}$)	0.12 ± 0.04
Dihedral angle constraints ($^{\circ}$)	1.76 ± 0.70
Max. dihedral angle violation ($^{\circ}$)	4.01
Max. distance constraint violation ($\text{\AA}$)	0.19
Average pairwise r.m.s. deviation* ($\text{\AA}$)	
Heavy	0.89 ± 0.11
Backbone	0.28 ± 0.10

* Pairwise r.m.s. deviation was calculated among 50 refined structures.

Table 2

Strains and plasmids used in this study.

Strain	Description	Reference
RN10616	RN4220 (80 α)	Ubeda et al. 2009 ²⁵
RN10628	RN4220 SaPI1 _{1st::tetM} (80 α)	Ubeda et al. 2009 ²⁵
ST63 *	RN10628 Δ terS _{SaPI1} Δ terS _{80α}	Ubeda et al. 2009 ²⁵
ST65	RN10628 Δ 80 α orf44	this study
ST100	RN10628 Δ SaPI1 orf6	this study
Plasmid	Description	Reference
pMAD	allelic exchange vector	Arnaud et al. 2004 ⁴⁶
pPD17	pMAD Δ 80 α orf44	this study
pPD35	pMAD Δ SaPI1 orf6	this study

* this strain is erroneously referred to as ST27 in Ubeda et al. (2009)²⁵