

Three-dimensional structure of the bacteriophage P22 tail machine

Liang Tang¹, William R Marion²,
Gino Cingolani^{1,3}, Peter E Prevelige Jr²
and John E Johnson^{1,*}

¹Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA, USA and ²Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL, USA

The tail of the bacteriophage P22 is composed of multiple protein components and integrates various biological functions that are crucial to the assembly and infection of the phage. The three-dimensional structure of the P22 tail machine determined by electron cryo-microscopy and image reconstruction reveals how the five types of polypeptides present as 51 subunits are organized into this molecular machine through twelve-, six- and three-fold symmetry, and provides insights into molecular events during host cell attachment and phage DNA translocation.

The EMBO Journal (2005) 24, 2087–2095. doi:10.1038/sj.emboj.7600695; Published online 2 June 2005

Subject Categories: structural biology; microbiology & pathogens

Keywords: bacteriophage; electron cryo-microscopy; molecular machine; symmetry; tail

Introduction

Most double-stranded DNA (dsDNA) bacteriophages evolve a highly specialized tail that is crucial to infection and is composed of numerous copies of proteins (Brussow and Hendrix, 2002). These tailed phages form the order of Caudovirales and are classified into three families based on tail morphology: Myoviridae, which have a long contractile tail; Siphoviridae, which have a long but noncontractile tail; and Podoviridae, which have a short tail (Ackermann, 1998). Bacteriophage P22, one of the prototypes in Podoviridae, is characterized by a short tail structure incorporated into a unique five-fold vertex of the icosahedral phage capsid (Poteete, 1994). This tail structure is a molecular machine that consists of multiple protein components with a total mass over 2.8 million Daltons and integrates various biological functions that are critical to the assembly and infection of the phage (Table I). The assembly of the tail machine is initiated by incorporation of the portal protein into one vertex of the procapsid during procapsid assembly together with several copies of each of three minor proteins (Bazin et al,

1988). The minor proteins and the portal complex are required for infectivity but not for procapsid assembly (Botstein et al, 1973). The portal forms a ring that provides a conduit for DNA entry during phage assembly and DNA exit during infection (Bazin et al, 1985), as in many other dsDNA phages and some eukaryotic viruses such as herpesvirus (Newcomb et al, 2001; Sheaffer et al, 2001; Dittmer and Bogner, 2005). Upon completion of DNA packaging, proteins gp4, gp10 and gp26 are attached to the portal to form a slender substructure, which stabilizes the packaged DNA by preventing it from leaking out of the capsid (King et al, 1973; Lenk et al, 1975; Strauss and King, 1984). In the last step of phage assembly, tailspikes are attached. The tailspike protein, a homotrimer of 72 kDa subunits, is anchored by its N-terminal head-binding domain in a noncovalent but irreversible manner (Berget and Poteete, 1980; Maurides et al, 1990). The C-terminal receptor-binding domain of the tailspike recognizes and cleaves the O-antigenic repeating units of lipopolysaccharide on the cell surface of *Salmonella* bacteria (Iwashita and Kanegasaki, 1976). This is followed by the injection of the phage DNA through the tail into the host cytoplasm.

The structure and assembly of the capsid of P22 as well as many other bacterial and eukaryotic dsDNA viruses have been studied extensively (Stewart et al, 1991; Wikoff et al, 2000; Zhou et al, 2000; Conway et al, 2001; Grunewald et al, 2003; Heymann et al, 2003; Jiang et al, 2003; Rixon and Chiu, 2003; Abrescia et al, 2004; Fokine et al, 2004). However, the current knowledge about the detailed architecture of the tail in dsDNA bacteriophage has been scarce, and is largely limited to the case of T4, a bacteriophage of Myoviridae family that infects *Escherichia coli*. Large-scale rearrangement of protein subunits was observed by comparison of the structures of the noncontracted and contracted tails of T4 (Kostyuchenko et al, 2003; Leiman et al, 2004). The electron cryo-microscopy (cryoEM) structure of *Bacillus* phage SPP1 portal in complex with head completion proteins gp15 and gp16, through which the tail is connected to the head, revealed local conformational change of the portal upon binding of gp15 and the function of gp16 as a valve (Orlova et al, 2003). Here, we present the three-dimensional structures of the tail machine and a C-terminally truncated form of the portal protein of bacteriophage P22 determined by cryoEM and image reconstruction. Combining data of biochemistry, X-ray crystallography, cryoEM and genomics sources, we have located the five protein components in this elegant molecular machine. The structures indicate novel functions of component proteins that were not recognized previously, and provided insights into the biological functions of the tail machine as a whole.

Results and discussion

Structure determination of the tail machine

The P22 tail machine was purified from the mature phage by disrupting the phage head. The C-terminally truncated portal

*Corresponding author. Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. Tel.: +1 858 784 9705; Fax: +1 858 784 8660; E-mail: jackj@scripps.edu

³Present address: Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13078, USA

Received: 8 March 2005; accepted: 3 May 2005; published online: 2 June 2005

Table 1 Protein components of the P22 tail machine

| | Function | Number of residues ^a | MW (Da) | pI | Copy number ^b |
|-----------------|---------------------------------|---------------------------------|---------|------|--------------------------|
| gp1 (portal) | DNA packaging | 724 | 82 611 | 4.75 | 12 |
| gp9 (tailspike) | Receptor binding and destroying | 667 | 71 857 | 5.34 | 6 × 3 |
| gp4 | Head completion | 166 | 18 025 | 4.91 | 12 |
| gp10 | Head completion | 472 | 52 457 | 5.90 | 6 |
| gp26 | Head completion | 232 | 24 603 | 5.14 | 3 |

^aThe number of residues, calculated molecular weight (MW) and isoelectric points (pI) are derived from SWISS-PROT (<http://us.expasy.org/spot/>).

^bThe copy numbers of each protein in the tail machine as implied in the cryoEM reconstruction.

protein was overexpressed in *E. coli*. CryoEM reconstructions were performed on the tail machine and the isolated portal assembly. The latter was fitted into the three-dimensional map of the tail machine to facilitate localization of components.

The electron micrographs of frozen-hydrated P22 tail machine showed excellent homogeneity (Figure 1A). The particle has six lobes surrounding a ring-like hub. The low density in the middle of the hub indicates the presence of a channel. There is a small piece of high density at the center or near the edge of the hub, which could be the axial projection of a rod-like substructure. This is more clearly visible in the class average of the images (Figure 1A). The particles have preferential orientation in the vitrified ice. That is, virtually all the particles are oriented in such a way that the central channel is perpendicular to the ice plane. To overcome the problem in three-dimensional reconstruction owing to the preferential orientation, we took series of micrographs by tilting the grids through various angles up to 60° (Figure 1B and C). An initial model was built manually by addition of six cylindrical density elements and a small spherical density to the reconstruction of the isolated truncated portal (see below). The final reconstruction was computed with 11 329 particles from tilted micrographs. The elongated needle-like density arising from the spherical one in the original map served as a validation for the reconstruction.

Structure of the C-terminally truncated version of P22 portal

The preliminary crystallographic analysis of full-length P22 portal protein to 7 Å resolution was previously reported (Cingolani *et al*, 2002). Further X-ray structural studies have been hampered by difficulties in improvement of crystal quality, which is probably due to the high structural flexibility of the portal. We have conducted cryoEM and image reconstruction of the truncated portal assembly encompassing residues 1–602. The cryoEM structure shows a ring-like morphology shared by portals of phi29, SPP1, T3 and herpes simplex virus (Simpson *et al*, 2000; Guasch *et al*, 2002; Orlova *et al*, 2003; Trus *et al*, 2004) (Figure 2A and B). The crown, wing and stalk domains are discernible (Figure 2B). The crown and wing domains are connected by thin density, implying plausible dynamics in terms of relative orientation/position between these two domains, which might be required for the regulatory role of the crown domain during DNA packaging and/or ejection (Simpson *et al*, 2000; Guasch *et al*, 2002; Orlova *et al*, 2003). The diameter of the portal channel varies with the narrowest 35 Å at the bottom opening and 40 Å at the top opening. The maximal diameter of the channel is ~70 Å below the crown domains, creating a

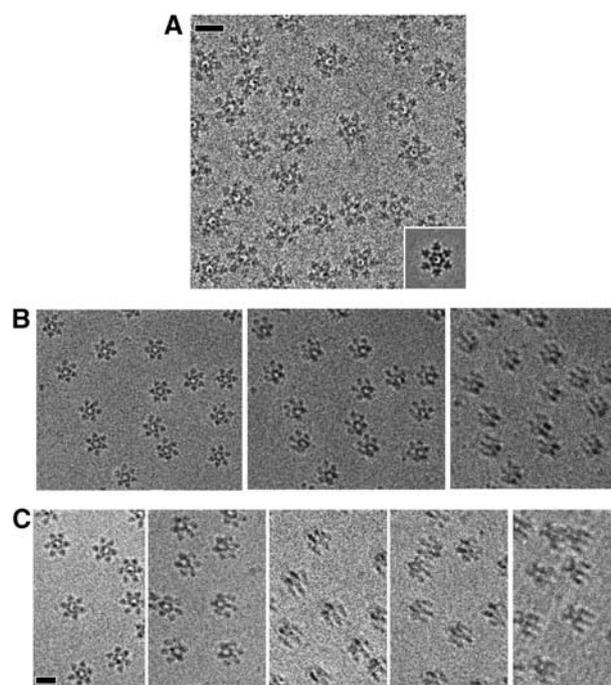


Figure 1 Electron micrographs of the frozen-hydrated P22 tail machine. (A) Representative view of the electron micrograph. Inset: a class average. (B) CCD images over the same field taken at tilt angles of 0, 20 and 40°. (C) Tilt series by 0, 20, 30, 40 and 60° (from left to right). Bars: 200 Å.

hollow space inside the portal complex. The X-ray structure of phi29 portal shows a channel of 36 Å at the bottom opening, which widens up to ~60 Å at the top opening. However, residues 231–244 that were not visible in the structure of phi29 portal presumably point inward from the channel wall, and could occupy a position similar to that of the P22 portal crown domain, which would limit the actual inner diameter of the channel. In contrast, the channel of the SPP1 portal has a relatively constant diameter as was shown in a cryoEM reconstruction (Orlova *et al*, 2003). Hence, assuming that a B-form DNA passes through this channel, two contact sites would be present in P22 portal involving the crown domain and the distal end of the stalk domain, respectively, whereas the DNA–portal interactions in SPP1 portal would be more evenly distributed over the whole channel. The truncated C-terminal 122 residues are most likely situated above the crown domain on the outer surface of the portal assembly, which is consistent with the fact that the C-terminal His tag is accessible to the Ni column during purification of the tail machine (see Materials and methods).

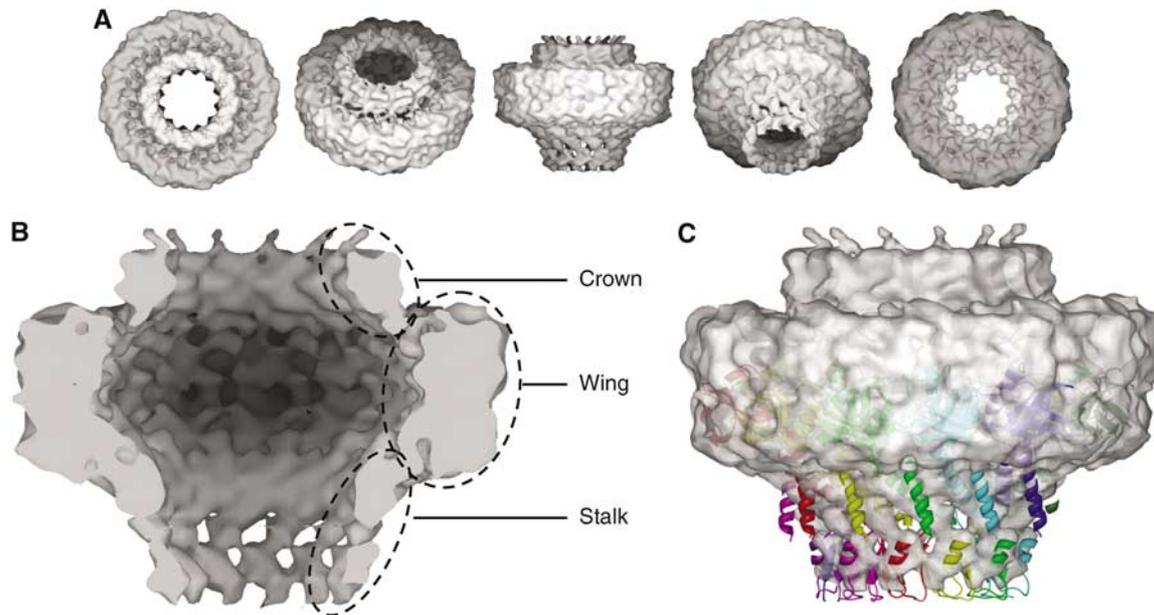


Figure 2 Structure of the C-terminally truncated P22 portal. (A) The map was rotated by 45° about the horizontal axis sequentially from the left to the right. (B) Cutaway view. The three domains are outlined by dashed ellipses. (C) Fitting of the X-ray structure of phi29 portal (the ribbon diagram) into the cryoEM map of the P22 portal (semitransparent isosurface). Each subunit of the phi29 portal is in a unique color. Only the front half is shown for clarity.

In the P22 portal cryoEM map, the molecular boundary in the stalk domain is readily recognizable (Figure 2). The stalk domain has an elongated shape, and the long axis is skewed by $\sim 25^\circ$ with respect to the 12-fold rotational axis. The X-ray structure of the phi29 portal was docked into the cryoEM map of the P22 portal (Figure 2C) (Simpson *et al*, 2000; Guasch *et al*, 2002). The elongated density of the stalk domain in the P22 portal fits well with the coordinates of the characteristic long α helices (residues 130–155 and 206–225) in phi29 portal (Guasch *et al*, 2002). This similarity in stalk domain orientation implies a common mechanical role in the DNA packaging molecular motor (Simpson *et al*, 2000; Guasch *et al*, 2002; Serwer, 2003). On the other hand, differences in the size and local structure of the other two domains correspond to more diverse regulatory functions, such as sensing and signaling during DNA packaging and ejection, and interactions with other proteins in the phage (Casjens *et al*, 1992; Molineux, 2001; Isidro *et al*, 2004).

Overall structure of the tail machine

The reconstruction of the tail machine reveals a structure with a height of 416 Å and a maximal diameter of 260 Å, which is made up of three major parts: a central tube, six appendages and a needle (Figures 3 and 4A). The atomic structures of the C-terminal receptor-binding domain and the N-terminal head-binding domain of the tailspike gp9 were determined at high resolution (Steinbacher *et al*, 1996, 1997), whereas X-ray structures of other protein components are lacking. Combination with genetic and biochemical data as well as the cryoEM reconstruction of the isolated P22 portal has allowed us to localize the five-component proteins in the three-dimensional map of the tail machine. The upper major part of the tail machine is a tube with a channel of 43 Å diameter in the middle (Figures 3 and 4A), which is presumably the conduit for viral DNA entry during packaging

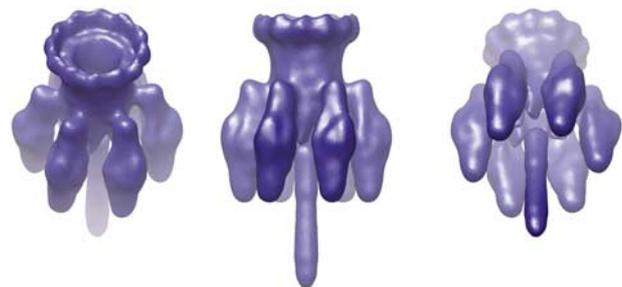


Figure 3 CryoEM reconstruction of the P22 tail machine. The map was rotated by 45° along the horizontal axis stepwise from left to right. The map was contoured at 3.7σ above background and the resultant volume corresponds to 100% of the mass.

and exit during infection. The diameter of the channel varies from 113 Å at the top opening, to the narrowest 43 Å, and 52 Å at the distal end (Figure 4A). The outer diameter of the tube ranges from 80 Å for the bottom end to 164 Å for the top wider end. The upper portion of this tube is assigned as the portal and the lower portion as gp4 and gp10. Attached to the lower half of the tube are six elongated appendages, each of which is a trimer of the tailspike protein gp9. The bottom opening of the tube is plugged by a 218 Å long needle-like density, which is coaxial with the tube. This needle has been interpreted as gp26.

The tailspike

The tailspike polypeptide consists of two domains, a C-terminal domain that binds the lipopolysaccharide receptor and an N-terminal head-binding domain that is responsible for attachment to the phage, with the domain boundary at residues 100–120. The structures of both domains were solved by X-ray crystallography independently, which provided atomic models for residues 5–108 and 113–666

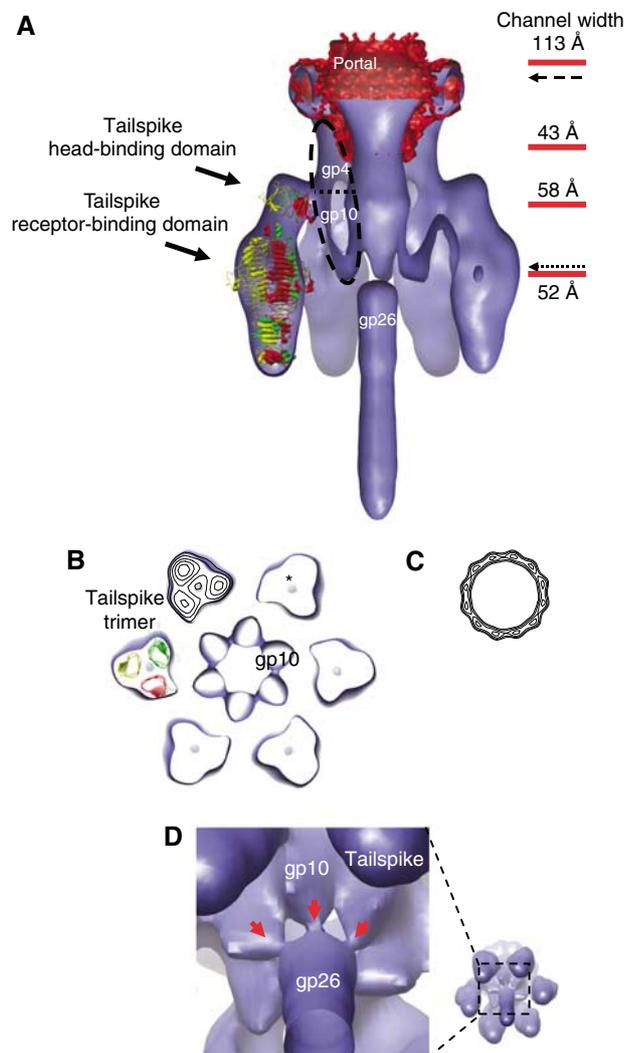


Figure 4 Structure of the P22 tail machine. (A) Cutaway view of the P22 tail machine (blue isosurface). The cryoEM map of C-terminally truncated portal (red isosurface) is superimposed. The X-ray structures of the receptor-binding domains and head-binding domains of the homotrimeric tailspike (ribbon diagrams with each polypeptide in red, green and yellow) are docked into the map. The density for gp4 and gp10 is indicated by a dashed curve, and the putative boundary between them by a dotted line. (B) Cross-section of the tail machine map at the position indicated by a dotted arrow in panel A. The atomic model of the receptor-binding domains (ribbon diagrams with each polypeptide in red, green and yellow) is superimposed with one of six tailspike trimers. A contour map (density value from 4.35σ to maximum with an interval of 1.74σ) is superimposed to show the trimeric conformation of the tailspike. The tiny feature (asterisk) in the middle of each tailspike represents lower density. (C) A contour (density value from 4.35σ to maximum with an interval of 0.58σ) of a cross-section in the tail machine map corresponding to the portal. The position of the cross-section is indicated by a dashed arrow in panel A. (D) The bottom opening of the tail tube viewed from the gp26 needle. The density that connects gp26 and gp10 is indicated by red arrows.

(Steinbacher *et al*, 1996; Steinbacher *et al*, 1997). However, the connection between the two domains has not been established owing to lack of structural information for the residues in the domain boundary. The six appendages in the cryoEM map of the tail machine clearly correlate with the two atomic models. The trimeric organization of each tailspike is evident both in the cryoEM micrograph (Figure 1A) and in a

cross-section contour of the cryoEM map (Figure 4B). The X-ray structure of the receptor-binding domain of the tailspike protein was unambiguously docked into the appendage in a global search over the whole cryoEM map with the program colores (Wriggers *et al*, 1999) (Figure 4). The domain is oriented in such a way that one monomer points outward away from the global six-fold axis while the other two monomers face to the tail tube.

The atomic model of the trimeric head-binding domains was then manually fit into the rest of the appendage density on the top. The orientation of the head-binding domain trimer about its internal three-fold axis is somewhat arbitrary. However, the intratrimeric rotational three-fold axes of the head-binding domains and of the receptor-binding domains intersect at 20° . In addition, the three-fold axis of the head-binding domains points outward by 18° away from the global six-fold axis of the tail machine, and that of the receptor-binding domains inclines by 8° when viewed from tailspike toward the global six-fold axis. This misorientation between the two domains is allowed by the flexible linkage near residues 100–120 between the two domains (Steinbacher *et al*, 1997). It is not clear what force holds these two domains into the current conformation. Nevertheless, the current structure may represent a closed conformation. Upon host cell attachment, the receptor-binding domains bind to the lipopolysaccharide and may spread out like a blooming flower so that the bottom opening of the channel in the tail tube is closer, and therefore has easier access, to the host cell membrane.

The interface between the tailspike and the tail tube appears to encompass more area than the lateral surface of the whole head-binding domain of one tailspike subunit (Figure 4). This suggests that binding of tailspike on the tail tube involves the interface formed by two head-binding domains. In fact, an Asp100Asn mutation that disrupted the interface between head-binding domains decreased the affinity of the tailspike to the tail tube to 1% of the wild-type tailspike (Schwarz and Berget, 1989; Steinbacher *et al*, 1997). This large contact area and plausible large conformational change in the local structure of the head-binding domain polypeptides could explain the noncovalent but irreversible binding of the tailspike (Berget and Poteete, 1980; Maurides *et al*, 1990). We propose that the tailspike-binding site on the tail tube is located on the interface of the two tail tube proteins gp4 and gp10, as gp10 is required for attachment of the tailspike (Strauss and King, 1984), and the mass of gp4 and gp10 implies that the molecular boundary between them crosses the tailspike-binding site (see below).

The portal in the tail machine

The upper end of the tail tube shows a wider opening. Inspection of the morphology indicates that it accounts for the wing domain in the portal. Indeed, the cryoEM structure of the C-terminally truncated portal fits well with the upper part of the tail tube in the map of the tail machine (Figure 4). Nevertheless, the crown domain was not visible in the tail machine reconstruction, putatively resulting from the dynamic nature of this domain upon tail assembly.

The 12-fold symmetry of the wing domains of the portal is clearly visible in the tail machine map, although only six-fold symmetry was imposed during reconstruction (Figures 3 and 4C).

The tail accessory proteins gp4 and gp10

Early data on the assembly pathway of P22 suggested that, after DNA packaging and terminase release, the portal provides docking sites for gp4, and subsequently gp10 is attached (Strauss and King, 1984). Binding of gp10 requires binding of gp4, whereas gp4 is able to bind to the phage head without the presence of gp10 (Strauss and King, 1984). Therefore, the middle and the lower parts of the tail tube are accounted for by gp4 and gp10, respectively (Figure 4). Genomics data of P22-like phages revealed as high as 93% amino-acid sequence identity for gp10 (Casjens *et al*, 2004). Interestingly, high identity was detected for the N-terminal 48, 77 and 115 residues of gp4, gp26 and gp9, respectively, with their putative homologs, whereas the other portions of these proteins show considerably lower identity (Casjens *et al*, 2004). These data support the role of gp10 as a molecular junction in the tail assembly of P22-like phages. That is, gp10 serves as a common platform for binding of the N-terminal domains from gp4, gp26 and tailspike. The variable C-terminal domain of gp4, gp26 and tailspike either binds to the portal or performs respective biological functions that are related to phage/host specificity.

We have extracted the putative density for gp4 and gp10 by computationally subtracting the density for the tailspike, the needle and the portal. The resultant density predicts a molecular mass of 521 kDa, when the map is contoured at a level that is used for the whole tail machine map to yield 100% mass. This value is in good agreement with the mass of a gp10 hexamer and a gp4 12-mer, which is 531 kDa. Therefore, we infer that gp4 and gp10 are present in the tail assembly as a 12-mer and hexamer, respectively. In fact, the cryoEM map clearly shows six protrusions at the bottom end of the tail tube, and the density distribution in horizontal sections shows six peaks (Figures 3, 4B and D). In contrast, gp4 has more contiguous density. In phage T4, gp13, gp14, gp15 and gp3 form a neck region that connects the tail to the portal, and both gp13 and gp14 are dodecameric, whereas gp15 and gp3 appears to be hexameric (Leiman *et al*, 2004). The cryoEM structure of SPP1 connector with the head completion proteins gp15 and gp16 showed that gp15 and gp16 are both dodecameric (Orlova *et al*, 2003). Gp4 appears as a nut that fits around a bolt formed by the distal end of the portal (Figures 4A and 6), in sharp contrast to the SPP1 connector where the portal slightly encompasses gp15 (Orlova *et al*, 2003). Taking into consideration the molecular mass of the two proteins (Table I), a putative line of the boundary between gp4 and gp10 can be drawn, and this line turns out to be within the tailspike-binding site (Figure 4). Thus, the interface between gp4 and gp10 provides the binding site for the tailspike. This ensures that gp4 and gp10 together form only six sites for tailspike binding, although gp4 is a 12-mer, as it would create irregular binding patterns other than the normal six-fold binding of tailspike if 12 binding sites were present. This is consistent with the highly homogeneous hexavalent binding of the tailspike in the tail assembly in the experimental preparation.

The needle is formed by gp26

The lower end of the channel is plugged by a needle with a length of 214 Å and a diameter of about 38 Å (Figures 3 and 4). We assigned this needle density to gp26 based on several lines of evidence. First, gp26 is required for preventing

packaged DNA from leaking out of the capsid, and was proposed to be the last protein in the sequential assembly of the tail (Israel, 1977; Strauss and King, 1984). Lack or mutation of gp4 and gp10 blocks attachment of gp26 (Strauss and King, 1984). Second, sequence analysis and secondary structure prediction show two long helices characterized by multiple heptad repeats, which implies that this protein folds as a trimeric coiled coil (Figure 5). This structural motif was observed in the C-terminal 119 residues of fibrin that forms the whiskers of phage T4 (Tao *et al*, 1997), and type I membrane fusion proteins in many enveloped viruses such as hemagglutinin of influenza virus (Gamblin *et al*, 2004; Stevens *et al*, 2004) and spike proteins in coronavirus (Xu *et al*, 2004). The length of the predicted long helix composed of ~100 residues would be 150 Å, which fits well with the overall length of the needle taking into consideration that the remaining residues form domain structures near the N- and C-termini. The trimerization of heterologously expressed and purified gp26 in solution was observed by sedimentation equilibrium analysis (Andrews *et al*, 2005). Third, genetic mosaicism studies suggested that gp10 serve as a common platform for binding of the N-terminal domains from gp4, gp26 and gp9 with homology among P22-like phages (see the previous section). We postulate that the N-terminal proximity of gp26 is involved in binding with gp10, and most likely forms a subdomain to interact with the bottom tip of gp10. The cryoEM map shows slender but clear density between the top of gp26 and the bottom tip of gp10, and this interaction holds gp26 in the present conformation (Figure 4D).

Early work on P22 indicated that gp26 was the last in the three accessory proteins to be attached to the tail (Strauss and King, 1984). The location of gp26 appears to support a long-standing thought that it acted as a plug to prevent the packaged DNA from leaking. In *Bacillus* bacteriophage SPP1, there is no counterpart of gp26 and the role of a plug is played by a portion of accessory protein gp16, which forms a seal near the distal end of the connector and blocks the channel (Orlova *et al*, 2003). Interestingly, the length of 214 Å places gp26 at the forefront when the phage interacts with the host cell. Considering the position of lipopolysaccharide on the cell surface where the tailspikes are bound, the bottom tip of the gp26 needle could be deeply inserted into the outer membrane, and reach the peptidoglycan and even the inner membrane. Given the rigid nature of the triple-helical coiled-coil motif, a natural thought is that gp26 could serve as a mechanism for perturbation of the host cell outer membrane, if not as a cell-puncturing device such as that in T4 (Kanamaru *et al*, 2002), which in turn allows DNA translocation across the cell membrane. The contact of gp26 with the host cell outer membrane and/or peptidoglycan may physically strip it off the gp10. Indeed, gp26 was among the ejection proteins that were released from the phage virion after host cell attachment (Israel, 1977). The release of gp26 may produce a signal to be transduced through gp10, gp4 and the portal into the interior of the phage head, and trigger the phage DNA translocation.

Implications for P22 assembly and infection

The molecular anatomy described above allows us to structurally understand the assembly of the P22 tail machine (Figure 6). Upon DNA packaging and terminase dissociation, gp4, gp10, gp26 and gp9 are attached in a precisely controlled

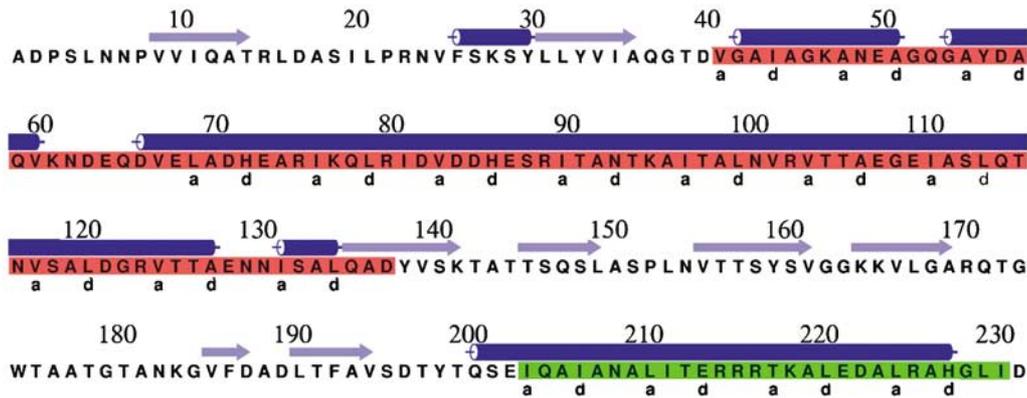


Figure 5 Secondary structure prediction of gp26 by PHD (<http://www.predictprotein.org/>) (Rost *et al*, 2004). The two putative heptad repeat regions are indicated by red and green background, and the characteristic first and fourth hydrophobic residues in each repeat unit are indicated beneath the sequence by letters a and d, respectively. The secondary structure is labeled by blue cylinders for α helix and light blue arrows for β strand.

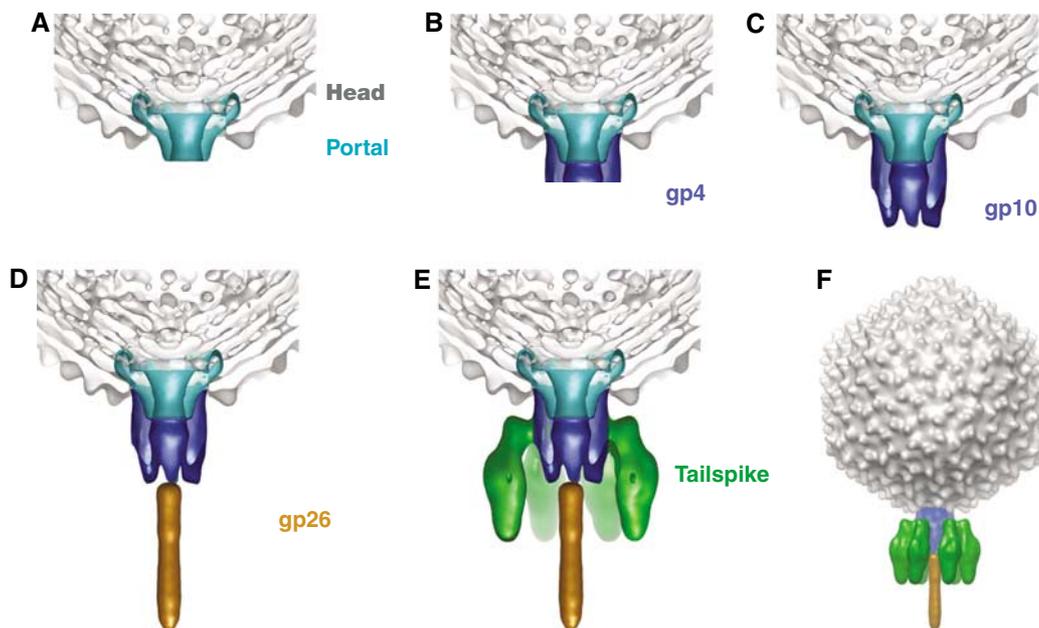


Figure 6 The assembly cascade of the P22 tail machine, starting from the phage capsid with the portal incorporated in a unique five-fold vertex (A), followed by sequential attachment of gp4 (B), gp10 (C), gp26 (D) and tailspikes (E). The tail machine is putatively superimposed with a part of a cryoEM map (gray isosurface) of the wild-type P22 (L Tang, JE Johnson and SR Casjens, unpublished data). The front half of the structures in (A–E) was computationally removed to show the internal features. A front view of the putative P22 structure is shown in (F).

order (Strauss and King, 1984). The structure of the tail machine illustrates an organization of multiple proteins governed through a variety of symmetry relationships (Figure 6). The ring-like dodecameric portal is located at one of the 12 five-fold vertices of the phage head, which is the first symmetry transition in the phage assembly. The accessory proteins gp4 and gp10 are attached to the portal to form a cylindrical tube with a central channel. This channel provides the conduit for phage DNA entry and exit. gp4 is dodecameric and gp10 is hexameric, as indicated in the cryoEM map. Thus, the second symmetry transition occurs between gp4 and gp10. The third symmetry transition takes place where the gp26 trimer is attached to the hexameric gp10. It forms a plug at the distal opening of the tail tube to prevent the packaged DNA from leaking. The six copies of

trimeric tailspikes bind to the interface of gp4 and gp10 through its N-terminal domain. The C-terminal receptor-binding domain of the tailspike recognizes the O-antigenic repeat units of lipopolysaccharide on the surface of the host cell. The needle formed by gp26 most likely penetrates the outer membrane to reach the peptidoglycan layer, and then undergoes major conformational changes such as being dislodged from gp10 driven by the force created by tailspike-receptor binding. This opens the channel, and a signal is generated and relayed through gp10, gp4 and the portal into the internal head, and thus triggers a series molecular events leading to phage DNA translocation and/or delivery of minor phage proteins in some phages such as P22 and T7 (Israel, 1977; Strauss and King, 1984; Molineux, 2001).

Materials and methods

Production of a C-terminally truncated P22 portal protein $\Delta C(1-602)$

Full-length P22 portal protein was expressed in *E. coli* strain BL21 (DE3) and purified as described previously (Cingolani *et al*, 2002). Concentrated portal rings displayed a characteristic pattern of degradation after ~3 months storage at 4°C. Using a combination of mass spectrometry and N-terminal sequencing, we determined that full-length protein was cleaved around residue 602, which yields a large C-terminally truncated fragment $\Delta C(1-602)$ of ~70 kDa. The portal protein DNA region encoding this fragment was amplified by PCR and ligated in a pET-21b vector. The $\Delta C(1-602)$ portal protein was expressed in *E. coli* strain BL21 (DE3). Cells were grown in LB media and expression of the recombinant protein induced by addition of 1 mM final concentration of IPTG. Expression was performed at 30°C for 3 h. Cells were harvested by centrifugation, resuspended in the lysis buffer (20 mM Hepes, pH 7.5, 250 mM sodium chloride, 5 mM β -mercaptoethanol, plus various protease inhibitors) and disrupted by sonication. Cellular debris was removed by centrifugation and the soluble fraction incubated with 5 ml of NTA-Ni-agarose beads, pre-equilibrated in the same buffer. After extensive washing with the wash buffer (20 mM Hepes, pH 7.5, 600 mM sodium chloride, 30 mM imidazole), bound P22 $\Delta C(1-602)$ portal protein was eluted with four bead-volumes of the elution buffer (20 mM Hepes, pH 7.5, 150 mM sodium chloride, 500 mM imidazole). The elution was then dialyzed against low-salt buffer (20 mM Hepes, pH 7.5, 50 mM sodium chloride, 5 mM β -mercaptoethanol) and concentrated to ~50 mg/ml using a Millipore concentrator (molecular weight cutoff 50 kDa) to promote ring formation. Typically, a liter of *E. coli* yielded ~50 mg of pure protein.

Purification of the P22 tail machine

Lysis-defective (13⁻am H101) phage carrying the gene of a C-terminal 6xHis-tagged portal protein was generated as described previously (Moore and Prevelige, 2001). A booster lysate of viable phage was used to infect 12.01 (37°C) of *Salmonella enterica* Serovar *Typhimurium* strain DB 7000 at an MOI of 10.0. The cells were harvested after 4 h and resuspended in a total volume of 240 ml of 10 mM MgCl₂ and 25 mM Tris-HCl (pH 7.6, 20 ml/l) and immediately frozen at -80°C. Cells were lysed by 3 × freeze-thaw and addition of lysozyme to 0.1 mg/ml. The lysate was centrifuged at 6000 g for 10 min (4°C) and the resultant supernatant centrifuged through a 20% (w/v) sucrose cushion at 195 000 g for 2 h at 4°C. The phage pellets were resuspended in a total volume of 5 ml 25 mM Tris-HCl, pH 7.6, by gentle shaking overnight at 4°C.

The resuspended phage was diluted 1:100 in a solution of 0.1 M EDTA, 0.1% (v/v) Triton X-100, 2.0 M urea and 50 mM Na₂HPO₄, pH 8.0. This mixture was incubated in a 60°C water bath for 45 min to burst the phage heads and liberate the tail machines. This is a modification of the procedure originally described (Hartwig *et al*, 1986). The suspension was cooled on ice, clarified by centrifugation (15 000 g, 30 min, 4°C) and dialyzed 3 × versus 1.0 l of 15 mM imidazole, 150 mM NaCl and 50 mM Na₂HPO₄, pH 8.0. This was loaded onto a 5 ml Pharmacia Hi-Trap Ni-NTA[®] column equilibrated with dialysate. Tail machines were specifically eluted using an imidazole gradient. The collected fractions were simultaneously concentrated and buffer-exchanged (50 mM Na₂HPO₄, pH 8.0) by centrifugation through a Pall-Filtron[®] 30 kDa cutoff membrane.

Electron cryo-microscopy and three-dimensional reconstruction

The sample was frozen-hydrated in vitrified ice on holey EM grids by a standard procedure (Adrian *et al*, 1984; van Heel *et al*, 2000). Briefly, 3–5 μ l of sample was applied for 1 min onto a previously glow-discharged copper grid coated with a holey carbon film. The grid was blotted with Whatman #2 filter paper, plunged into a slush of liquid ethane and stored in liquid nitrogen. The grid was transferred with a Gatan 626 cryostage (Gatan Inc., Pleasanton, CA) into a Phillips CM200FEG transmission electron microscope (Phillips/FEI, Eindhoven, The Netherlands) operated at an accelerating voltage of 120 kV. Electron micrographs were recorded on Kodak SO156 films under low-dose conditions (5–20 electrons/Å²). Micrographs were digitized using a Zeiss SCAI scanner (Z/I Imaging Corporation, Madison, AL) with a step size of 7 μ m. Particles were

boxed with the program EMAN (Ludtke *et al*, 1999), and image reconstruction was carried out with the program SPIDER (Frank *et al*, 1996). The resolution was estimated with the Fourier shell correlation method (van Heel *et al*, 2000).

In the P22 portal reconstruction, particles were boxed into a 160 × 160 dimension with a pixel size of 1.84 Å. A total of 20271 particles were used in the three-dimensional reconstruction. Image analysis by classification and averaging of images of negatively stained P22 portal showed 12-fold symmetry (data not shown), which was in agreement with the rotation function results from the X-ray crystallographic studies (Cingolani *et al*, 2002). Therefore, the 12-fold rotational symmetry was imposed during the reconstruction. The defoci of these micrographs ranged from 1.4 to 3.7 μ m, and the contrast transfer function was corrected by phase flipping. The resolution of the final reconstruction was estimated to be 8 Å, where the Fourier shell correlation between two independent reconstructions calculated from half data sets dropped to 0.5 (Figure 7A). The map revealed the crown, wing and stalk domains with a central channel. We were able to distinguish the subunit boundary at the stalk region, and the connection between the crown and wing domains was clearly defined in the map. The handedness of the map was determined assuming that the orientation of the stalk domain of the portal monomer is similar to that in the portal complex of bacteriophage phi29 whose atomic structures are available (Simpson *et al*, 2000; Guasch *et al*, 2002).

As the P22 tail machine shows preferential orientation in ice, we took micrographs by tilting the grids by angles of 0, 20, 30, 40 and

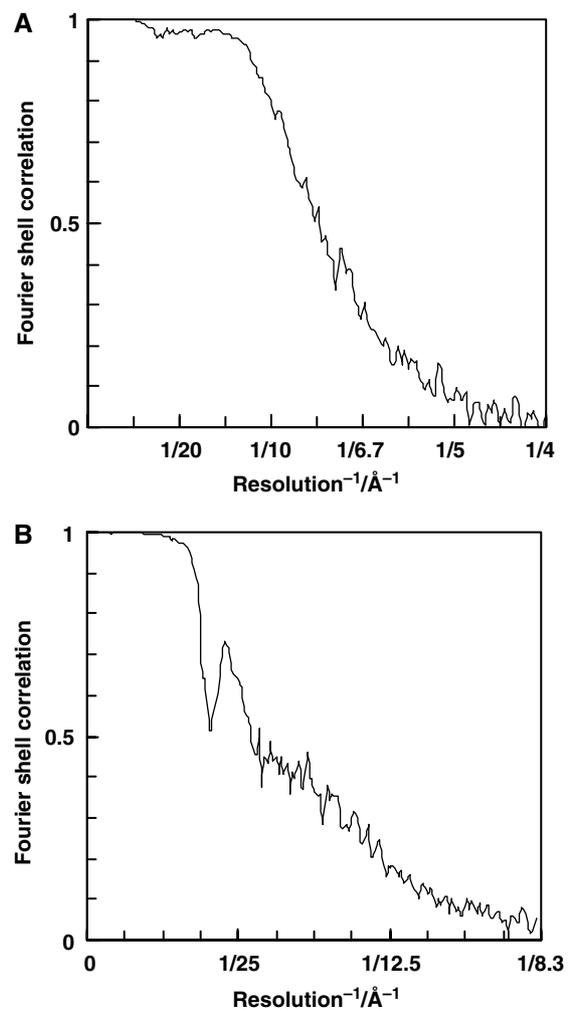


Figure 7 Fourier shell correlation plots of reconstructions of (A) the truncated P22 portal and (B) the P22 tail machine. In (B), the Fourier shell correlation drops to 0.5 at 1/23 Å⁻¹, 0.4 at 1/22 Å⁻¹, and 0.3 at 1/16 Å⁻¹.

60°. Inspection of the tilt series assisted manual building of an initial model by addition of six cylindrical density and a small spherical density to the portal reconstruction. In this initial model, one end of the cylindrical density was in close proximity to the lower part of the wing domain of the portal and the other end was slightly pointed outward so that the long axes of the cylinders and the portal channel form an angle of 15°. The small spherical density was placed next to the bottom opening of the portal. An initial data set of 3087 particles was used to calculate a preliminary reconstruction, which gave rise to an elongated needle-like density for the originally spherical one and therefore served as a validation for the reconstruction. A total of 17 517 particles were boxed from tilted micrographs into a 160 × 160 dimension with a pixel size of 4.2 Å. A reconstruction was calculated and refined using all the data and imposing C6 symmetry. In order to avoid possible inaccuracy in the length of the gp26 needle arising from saturated density in top or bottom views, the particles from untilted micrographs were

excluded from the reconstruction. The gp26 needle in the reconstruction computed from all data was removed. With such an initial model and 11 329 particles from tilted micrographs, a new reconstruction was calculated and refined. The contrast transfer functions were not corrected. The resolution of the final reconstruction was estimated to be 23 Å, where the Fourier shell correlation for two independent reconstructions calculated from half data sets started to drop to 0.5 (Figure 7B) (van Heel *et al*, 2000). The map has been deposited with EMDB database at the European Bioinformatics Institute with an accession code of EMD-1119.

Acknowledgements

We are grateful to Dr SR Casjens for helpful discussion. This work was supported by grants from the National Institutes of Health to JEJ (AI40101) and PEP (GM47980).

References

- Abrescia NG, Cockburn JJ, Grimes JM, Sutton GC, Diprose JM, Butcher SJ, Fuller SD, San Martin C, Burnett RM, Stuart DI, Bamford DH, Bamford JK (2004) Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* **432**: 68–74
- Ackermann HW (1998) Tailed bacteriophages: the order caudovirales. *Adv Virus Res* **51**: 135–201
- Adrian M, Dubochet J, Lepault J, McDowell AW (1984) Cryo-electron microscopy of viruses. *Nature* **308**: 32–36
- Andrews D, Butler JS, Al-Bassam J, Joss L, Winn-Stapley DA, Casjens S, Cingolani G (2005) Bacteriophage P22 tail accessory factor gp26 is a long triple-stranded coiled-coil. *J Biol Chem* **280**: 5929–5933
- Bazinet C, Benbasat J, King J, Carazo JM, Carrascosa JL (1988) Purification and organization of the gene 1 portal protein required for phage P22 DNA packaging. *Biochemistry* **27**: 1849–1856
- Bazinet C, King J (1985) The DNA translocating vertex of dsDNA bacteriophage. *Annu Rev Microbiol* **39**: 109–129
- Berget PB, Poteete AR (1980) Structure and functions of the bacteriophage P22 tail protein. *J Virol* **34**: 234–243
- Botstein D, Waddell CH, King J (1973) Mechanism of head assembly and DNA encapsulation in *Salmonella* phage p22. I. Genes, proteins, structures and DNA maturation. *J Mol Biol* **80**: 669–695
- Brussow H, Hendrix RW (2002) Phage genomics: small is beautiful. *Cell* **108**: 13–16
- Casjens S, Winn-Stapley DA, Gilcrease EB, Morona R, Kuhlewein C, Chua JE, Manning PA, Inwood W, Clark AJ (2004) The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. *J Mol Biol* **339**: 379–394
- Casjens S, Wyckoff E, Hayden M, Sampson L, Eppler K, Randall S, Moreno ET, Serwer P (1992) Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. *J Mol Biol* **224**: 1055–1074
- Cingolani G, Moore SD, Prevelige Jr PE, Johnson JE (2002) Preliminary crystallographic analysis of the bacteriophage P22 portal protein. *J Struct Biol* **139**: 46–54
- Conway JF, Wikoff WR, Cheng N, Duda RL, Hendrix RW, Johnson JE, Steven AC (2001) Virus maturation involving large subunit rotations and local refolding. *Science* **292**: 744–748
- Dittmer A, Bogner E (2005) Analysis of the quaternary structure of the putative HCMV portal protein PUL104. *Biochemistry* **44**: 759–765
- Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG (2004) Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci USA* **101**: 6003–6008
- Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**: 190–199
- Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasishth N, Steinhauer DA, Daniels RS, Elliot A, Wiley DC, Skehel JJ (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* **303**: 1838–1842
- Grunewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* **302**: 1396–1398
- Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, Sousa N, Carrascosa JL, Coll M (2002) Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle. *J Mol Biol* **315**: 663–676
- Hartwig E, Bazinet C, King J (1986) DNA injection apparatus of phage P22. *Biophys J* **49**: 24–26
- Heymann JB, Cheng N, Newcomb WW, Trus BL, Brown JC, Steven AC (2003) Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat Struct Biol* **10**: 334–341
- Isidro A, Santos MA, Henriques AO, Tavares P (2004) The high-resolution functional map of bacteriophage SPP1 portal protein. *Mol Microbiol* **51**: 949–962
- Israel V (1977) E proteins of bacteriophage P22. I. Identification and ejection from wild-type and defective particles. *J Virol* **23**: 91–97
- Iwashita S, Kanegasaki S (1976) Enzymic and molecular properties of base-plate parts of bacteriophage P22. *Eur J Biochem* **65**: 87–94
- Jiang W, Li Z, Zhang Z, Baker ML, Prevelige Jr PE, Chiu W (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**: 131–135
- Kanamaru S, Leiman PG, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Arisaka F, Rossmann MG (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**: 553–557
- King J, Lenk EV, Botstein D (1973) Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22. II. Morphogenetic pathway. *J Mol Biol* **80**: 697–731
- Kostyuchenko VA, Leiman PG, Chipman PR, Kanamaru S, van Raaij MJ, Arisaka F, Mesyanzhinov VV, Rossmann MG (2003) Three-dimensional structure of bacteriophage T4 baseplate. *Nat Struct Biol* **10**: 688–693
- Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118**: 419–429
- Lenk E, Casjens S, Weeks J, King J (1975) Intracellular visualization of precursor capsids in phage P22 mutant infected cells. *Virology* **68**: 182–199
- Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**: 82–97
- Maurides PA, Schwarz JJ, Berget PB (1990) Intragenic suppression of a capsid assembly-defective P22 tailspike mutation. *Genetics* **125**: 673–681
- Molineux IJ (2001) No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Mol Microbiol* **40**: 1–8
- Moore SD, Prevelige Jr PE (2001) Structural transformations accompanying the assembly of bacteriophage P22 portal protein rings *in vitro*. *J Biol Chem* **276**: 6779–6788
- Newcomb WW, Juhász RM, Thomsen DR, Homa FL, Burch AD, Weller SK, Brown JC (2001) The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J Virol* **75**: 10923–10932

- Orlova EV, Gowen B, Droge A, Stiege A, Weise F, Lurz R, van Heel M, Tavares P (2003) Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. *EMBO J* **22**: 1255–1262
- Poteete AR (1994) P22 bacteriophage. In *Encyclopedia of Virology*, Webster RG, Granoff A (eds) pp 1009–1013. London: Academic Press
- Rixon FJ, Chiu W (2003) Studying large viruses. *Adv Protein Chem* **64**: 379–408
- Rost B, Yachdav G, Liu J (2004) The PredictProtein server. *Nucleic Acids Res* **32**: W321–W326
- Schwarz JJ, Berget PB (1989) Characterization of bacteriophage P22 tailspike mutant proteins with altered endorhamnosidase and capsid assembly activities. *J Biol Chem* **264**: 20112–20119
- Serwer P (2003) Models of bacteriophage DNA packaging motors. *J Struct Biol* **141**: 179–188
- Sheaffer AK, Newcomb WW, Gao M, Yu D, Weller SK, Brown JC, Tenney DJ (2001) Herpes simplex virus DNA cleavage and packaging proteins associate with the procapsid prior to its maturation. *J Virol* **75**: 687–698
- Simpson AA, Tao Y, Leiman PG, Badasso MO, He Y, Jardine PJ, Olson NH, Morais MC, Grimes S, Anderson DL, Baker TS, Rossmann MG (2000) Structure of the bacteriophage phi29 DNA packaging motor. *Nature* **408**: 745–750
- Steinbacher S, Baxa U, Miller S, Weintraub A, Seckler R, Huber R (1996) Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc Natl Acad Sci USA* **93**: 10584–10588
- Steinbacher S, Miller S, Baxa U, Budisa N, Weintraub A, Seckler R, Huber R (1997) Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *J Mol Biol* **267**: 865–880
- Stevens J, Corper AL, Basler CF, Taubenberger JK, Palese P, Wilson IA (2004) Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* **303**: 1866–1870
- Stewart PL, Burnett RM, Cyrklaff M, Fuller SD (1991) Image reconstruction reveals the complex molecular organization of adenovirus. *Cell* **67**: 145–154
- Strauss H, King J (1984) Steps in the stabilization of newly packaged DNA during phage P22 morphogenesis. *J Mol Biol* **172**: 523–543
- Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG (1997) Structure of bacteriophage T4 fibrin: a segmented coiled coil and the role of the C-terminal domain. *Structure* **5**: 789–798
- Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, Steven AC (2004) Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J Virol* **78**: 12668–12671
- van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, Patwardhan A (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* **33**: 307–369
- Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**: 2129–2133
- Wriggers W, Milligan RA, McCammon JA (1999) Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* **125**: 185–195
- Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao GF, Rao Z (2004) Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. *J Biol Chem* **279**: 49414–49419
- Zhou ZH, Dougherty M, Jakana J, He J, Rixon FJ, Chiu W (2000) Seeing the herpesvirus capsid at 8.5 Å. *Science* **288**: 877–880