Cell, Vol. 118, 419–429, August 20, 2004, Copyright ©2004 by Cell Press

Three-Dimensional Rearrangement of Proteins in the Tail of Bacteriophage T4 on Infection of Its Host

Petr G. Leiman,[1] Paul R. Chipman,[1] Victor A. Kostyuchenko,[1,2] Vadim V. Mesyanzhinov,[2] and Michael G. Rossmann[1,*]

[1]Department of Biological Sciences
Purdue University
915 West State Street
West Lafayette, Indiana 47907

[2]Laboratory of Molecular Bioengineering
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry
16/10 Miklukho-Maklaya Street
117997 Moscow
Russia

Summary

The contractile tail of bacteriophage T4 undergoes major structural transitions when the virus attaches to the host cell surface. The baseplate at the distal end of the tail changes from a hexagonal to a star shape. This causes the sheath around the tail tube to contract and the tail tube to protrude from the baseplate and pierce the outer cell membrane and the cell wall before reaching the inner cell membrane for subsequent viral DNA injection. Analogously, the T4 tail can be contracted by treatment with 3 M urea. The structure of the T4 contracted tail, including the head-tail joining region, has been determined by cryo-electron microscopy to 17 Å resolution. This 1200 Å-long, 20 MDa structure has been interpreted in terms of multiple copies of its approximately 20 component proteins. A comparison with the metastable hexagonal baseplate of the mature virus shows that the baseplate proteins move as rigid bodies relative to each other during the structural change.

Introduction

The majority of bacterial viruses, or bacteriophages, employ a special molecular device, called a tail, for recognition and attachment to a host cell, penetration of the cell envelope, and for DNA transfer from the virus capsid into the host cell cytoplasm (Ackermann, 2003). The infection process involving a tail is evolutionarily advantageous over other strategies of viral infection and may account for tailed phages being the most numerous biological entities on the planet (Hendrix, 2002). These phages serve as essential elements in horizontal gene transfer between different biomes (Desplats and Krisch, 2003; Wommack and Colwell, 2000).

Bacteriophage T4, which infects *Escherichia coli*, is a genetically and biochemically well-studied member of the *Myoviridae* family. Members of this family contract their tails during infection (Ackermann, 2003; Miller et al., 2003). More than 40 structural proteins, each with multiple copies, constitute the T4 virus particle, which consists of a 5-fold-symmetric head containing the 172 kbp genomic DNA, a 6-fold-symmetric bilayered tail, and six long fibers attached to the tail's baseplate (Eiserling and Black, 1994; Leiman et al., 2003a). In T4, as in most bacteriophages and in some eukaryotic viruses, there is a special pentagonal vertex that is utilized both for inserting the genome into the capsid and subsequently for ejecting the genome into a host (Gowen et al., 2003; Moore and Prevelige, 2002; Newcomb et al., 2001). The 1200 Å-long, 860 Å diameter head of T4 is a prolate icosahedron (Fokine et al., 2004) with the unique vertex occupied by the dodecameric portal connector, gene product (gp) 20 (Driedonks et al., 1981).

The 980 Å-long, 220 Å diameter central cylindrical segment of the tail consists of a rigid tube, composed of multiple copies of gp19, surrounded by the outer contractile sheath assembled from gp18 subunits (Amos and Klug, 1975; King, 1968; Moody and Makowski, 1981). At least 14 proteins, many of which are oligomeric, make up the 270 Å-long, 500 Å diameter baseplate that terminates the distal end of the tail (Kikuchi and King, 1975b, 1975c, 1975d; Kostyuchenko et al., 2003; Leiman et al., 2003a). The proximal end of the tail is terminated by gp3 and gp15, which form a lock that prevents depolymerization of the tube and sheath (King, 1968; Vianelli et al., 2000). The tail is connected to the gp20 dodecamer via oligomers of gp13 and gp14, which form the neck region (Coombs and Arisaka, 1994; Coombs and Eiserling, 1977). Six 350 Å-long "whiskers," composed of fibritin trimers (gp *wac*), are attached to the neck (Tao et al., 1997). A long tail fiber consists of a 700 Å-long proximal part formed by trimeric gp34, a flexible hinge formed by monomeric gp35, and a 700 Å-long distal part formed by trimeric gp36 and gp37 (Cerritelli et al., 1996). The complex contractile tail of bacteriophage T4, whose structure and function are probably similar to those of other *Myoviridae* phages (Monod et al., 1997), is responsible for an almost 100% success rate of T4 infection of gram-negative bacteria (Coombs and Arisaka, 1994; Goldberg et al., 1994).

The baseplate has a hexagonal shape in the mature virus but changes to a star shape after adsorption to a host cell (Crowther et al., 1977; Simon and Anderson, 1967a). The cryo-electron microscopy (cryoEM) map of the hexagonal baseplate (Kostyuchenko et al., 2003) was interpreted by using X-ray structures of some of the component proteins. The baseplate was found to have a dome-shaped structure with a central spike formed by the membrane-puncturing device composed of the gp5-gp27 complex. The short tail fibers (gp12) were found to interact with each other in a head-to-tail fashion around the periphery of the dome and were proposed to stabilize the baseplate in the hexagonal conformation. The short tail fibers bend by about 90° around gp11 trimers situated at the vertices of the hexagonal dome. The baseplate "pins," seen in earlier negative staining EM experiments (Watts et al., 1990), are formed by gp7, gp10, and gp11. The central portion of the baseplate's density was assigned to the proteins gp6, gp25, and gp53. In addition, the density for the tail

*Correspondence: mgr@indiana.bio.purdue.edu

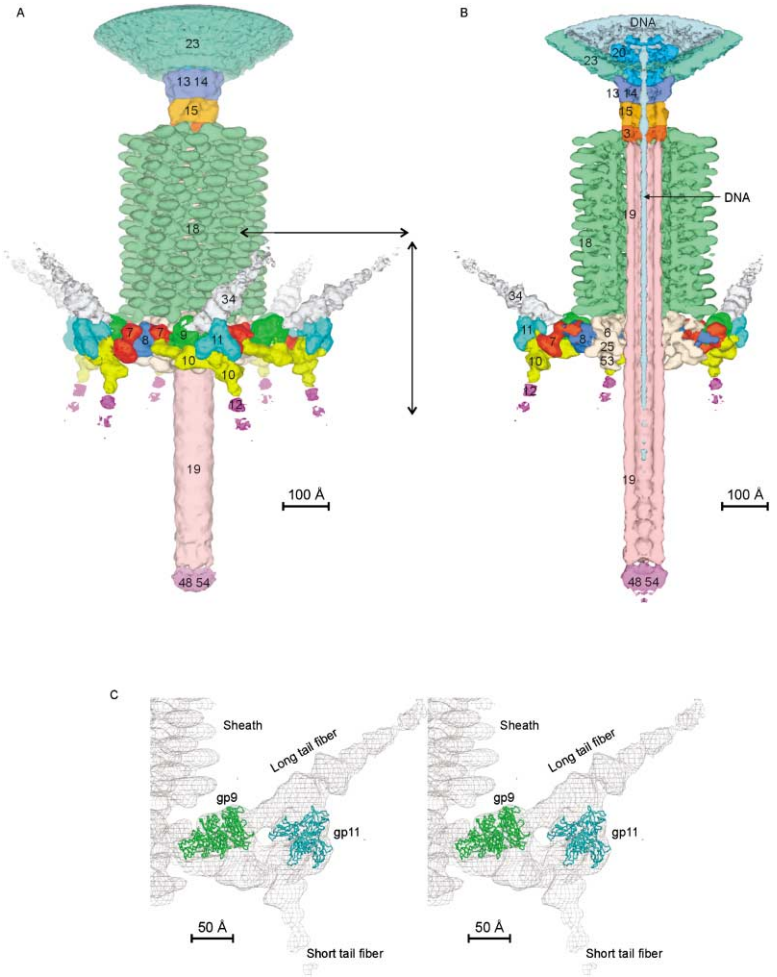


Figure 1. Structure of the Contracted T4 Tail

In (A) and (B), each identified protein, or protein assembly, is colored in a distinct color and identified with its respective gene number.

(A) Side view. The black arrows indicate the region shown in (C).

(B) Cut-away view. Note that the thickening of the tail tube wall toward its tip is probably an artifact of the image reconstruction caused by the lack of linearity of the tail tube in the different frozen phage particles.

(C) The crystal structures of gp9 (green) and gp11 (cyan), shown as C_α traces, are fitted computationally into the cryoEM density, highlighted in (A) with the black arrows. The long tail fiber binds collinearly with the trimer axis of gp9 and also interacts with the head and finger domains of gp11.

tube-associated proteins gp48 and gp54 was located at the top of the baseplate dome. The long tail fibers were proposed to attach collinearly to gp9, which was found to have a variable orientation in agreement with the swinging motion observed for the long tail fibers (Kellenberger et al., 1996).

The phage initially recognizes the host via reversible interaction of the tips of its long tail fibers with lipopolysaccharide cell surface receptors (Crawford and Goldberg, 1980; Goldberg et al., 1994). In this configuration, the baseplate is about 1000 Å away from the cell surface (Simon and Anderson, 1967a, 1967b). Subsequently, the baseplate is brought closer to the cell surface, allowing the short tail fibers to bind irreversibly to the lipopolysaccharides and causing the baseplate to switch into the star conformation. This initiates sheath contraction, which propagates in a manner of falling dominos (Moody, 1973). The sheath then contracts to about 37% of its original length and drags the baseplate along the tail tube, making the tube protrude from beneath the baseplate by about half of the tube's length (Moody, 1967; Simon and Anderson, 1967a). The tube penetrates through the cell envelope with the help of the tail lysozyme gp5 and reaches to the cytoplasmic membrane (Kanamaru et al., 2002). Subsequently, the phage DNA is released into the cell through the tube (Simon and Anderson, 1967a).

Contraction of the T4 tail can be induced by treatment with a concentrated lipopolysaccharide solution or 3 M urea (Coombs and Arisaka, 1994). Such treatment does not cause the release of the phage DNA, indicating that the contraction and the DNA release are not linked (Goldberg et al., 1994). Phages with contracted tails can infect spheroplasts (cells with exposed cytoplasmic membrane) of the *E. coli*-like bacteria, suggesting that the release of the phage DNA can be triggered only by interaction of the tail tube with the *E. coli* cytoplasmic membrane (Goldberg et al., 1994).

We report here a 17 Å resolution, three-dimensional, cryoEM reconstruction of the contracted T4 tail and identify the locations and shapes of its component proteins. Using the 12 Å resolution structure of the hexagonal baseplate (Kostyuchenko et al., 2003), we identified the structural rearrangements, which occur in the baseplate during the hexagonal-to-star conformational switch leading to sheath contraction. We propose that this conformational transition is responsible for repositioning the baseplate from being 1000 Å away from the

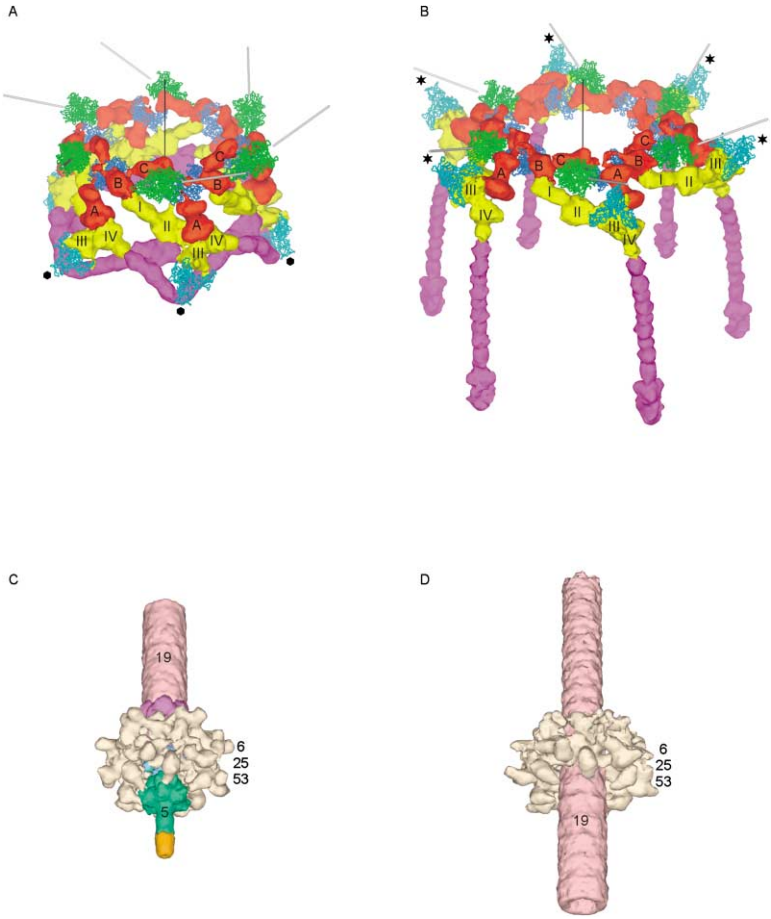


Figure 2. Comparison of the Baseplate in the Two Conformations

(A and B) Structure of the periphery of the baseplate in the hexagonal and star conformations, respectively. Colors identify different proteins as in Figure 1A: gp7 (red), gp8 (blue), gp9 (green), gp10 (yellow), gp11 (cyan), and gp12 (magenta). Three baseplate proteins (gp8, gp9, and gp11), with the available complete crystal structures, are shown as C_α traces. The density of the short tail fibers in the star conformation is based on the crystal structure of the receptor binding, C-terminal fragment of gp12 (Thomassen et al., 2003) and on the corresponding density from the hexagonal conformation of the baseplate. Directions of the long tail fibers are indicated with gray rods. The three domains of gp7 are labeled with letters A, B, and C. The four domains of gp10 are labeled with Roman numbers I through IV. The C-terminal domain of gp11 is labeled with a black hexagon or black star in the hexagonal or star conformations, respectively. The baseplate 6-fold axis is indicated by a black line.

(C and D) Structure of the proteins surrounding the hub in the hexagonal and star conformations. The proteins are colored as follows: spring green, gp5; pink, gp19; sky blue, gp27; violet, putative gp48 or gp54; beige, gp6-gp25-gp53; orange, unidentified protein at the tip of gp5. A part of the tail tube is shown in both conformations for clarity.

cell surface during the initial binding of the long tail fibers to being in close proximity with the cell surface. Furthermore, this conformational switch allows the short tail fibers to extend from the baseplate, thus preparing them for binding to the cellular receptors. These results visualize much of the infection process used by a complex macromolecular machine at quasi-atomic resolution. This information could form the basis for adapting this machine to become a highly specific, targeted gene delivery device.

Results

The Contracted Tail Structure

The contracted T4 tail consists of a 960 Å-long, 90 Å diameter tube, surrounded by a 420 Å-long, 330 Å diameter sheath that is terminated by a 120 Å thick, 610 Å diameter baseplate decorated with six long and six short tail fibers (Figure 1). The tail tube passes through the 95 Å wide hole in the center of the baseplate and protrudes by about 470 Å. The tail fibers, attached to the baseplate, are mostly disordered, and only 230 Å-long and 80 Å-long stretches of the proximal parts of the long and short tail fibers, respectively, can be seen in the cryoEM map. The sheath is connected to the head by a 110 Å-long, 110 Å diameter neck, which interacts with the dodecameric gp20 connector. Because the head has 5-fold symmetry but the reconstruction was calculated assuming 6-fold symmetry of the tail, the head has no discernable structural features.

The hand of the reconstruction was established by visual comparison of the arrangements of proteins in the hexagonal and star conformations of the baseplate. This was confirmed by fitting the crystal structures of gp8, gp9, and gp11 into the noninverted and inverted maps. Knowledge of the hand of the star-shaped baseplate thus established the absolute hand of the tail tube and sheath, which was found to be consistent with an earlier determination of the hand of the sheath (Amos and Klug, 1975).

Comparison of the Star-Shaped Baseplate with Its Structure in the Hexagonal Conformation

The star-shaped baseplate (Figure 2) is flatter and wider than the 500 Å diameter, 270 Å high baseplate in the hexagonal dome-like conformation (Kostyuchenko et al., 2003). Although the overall conformational change of the baseplate is large with, for example, gp11 moving by as much as 190 Å, the component proteins or their domains move as rigid bodies. In both the hexagonal and the star-shaped reconstructions of the baseplate, gp10 and gp7 could be interpreted to consist of four (I, II, III, and IV) and three (A, B, and C) domains, respectively (Figures 2A and 2B). Their equivalence in the two cryoEM

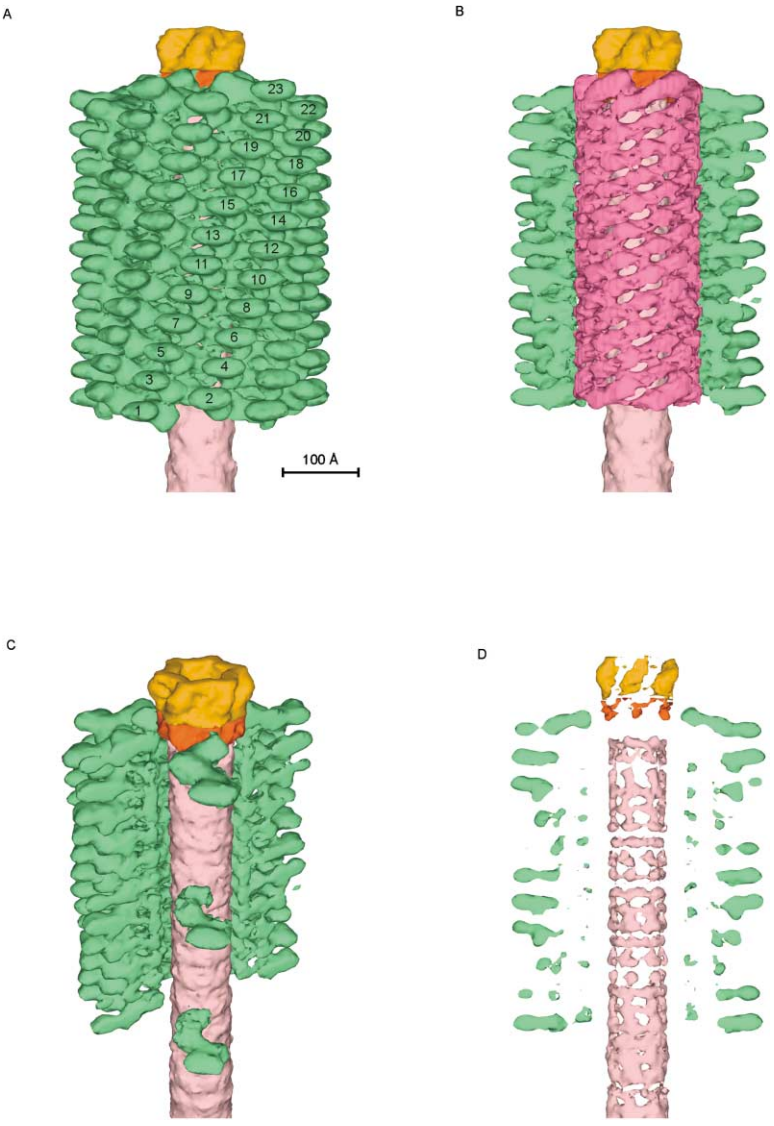


Figure 3. Structure of the Contracted Sheath

(A) The sheath is composed of 138 subunits of gp18 arranged into 23 hexameric rings. Other proteins are colored as in Figure 1A.

(B) The six start helix formed by the gp18 inner domains, shown in pink. The outer domains are colored in light green.

(C) The gp18 monomer is an S-shaped molecule. The structure is tilted toward the viewer by 30°.

(D) Fine structure of the tail tube seen at a high contour level. An about 50 Å-thick slice of the cryoEM map is shown.

maps was readily recognizable. Their positions and orientations within each protein are different, albeit domains I, II, and III of gp10 move as one unit.

In the hexagonal conformation, gp11 is located at the vertices of the baseplate's dome, helping to accommodate the short tail fibers along the rim of the dome (Figures 2A and 2B). The C-terminal domain of gp11 points away from the phage head, and its trimer axis makes a 144° angle with respect to the 6-fold axis of the baseplate. In the star conformation, however, the gp11 C-terminal domain points toward the phage head, and the trimer axis makes a 48° angle with respect to the baseplate 6-fold axis. Thus, upon completion of the baseplate's conformational change, each gp11 molecule rotates by almost 100° and associates with a long instead of a short tail fiber. The interaction between gp10 and gp11 is unchanged in the two conformations (Figures 2A and 2B). As a result, domain IV of gp10 changes its orientation, causing the associated short tail fiber to unfold from under the baseplate and point toward the potential host cell surface. During the hexagonal-to-star conformational transition, domain A of gp7 swivels outwards by about 45° and alters its association with gp10, making the baseplate structure flat (Figures 2A and 2B). This rearrangement brings domains I and II of gp10 into the proximity of gp9 and probably allows the latter to interact with gp8. The current reconstruction supports the hypothesis that the hexagonal-to-star conformational change of the baseplate is the result of a reorientation of the pins (gp7, gp10, gp11) (Watts et al., 1990) and additionally shows that the transformation also involves rearrangements of gp8, gp9, and gp12 situated around the periphery of the baseplate.

After gp7, gp8, gp9, gp10, gp11, and gp12 had been assigned to their specific positions, their densities were set to zero in the baseplate map. The remaining portion of the map was ascribed to the unassigned proteins gp6, gp25, and gp53. In the hexagonal baseplate, the gp6-gp25-gp53 density surrounds the hub (the gp5-gp27 complex) and the end of the tail tube (gp48 and/or

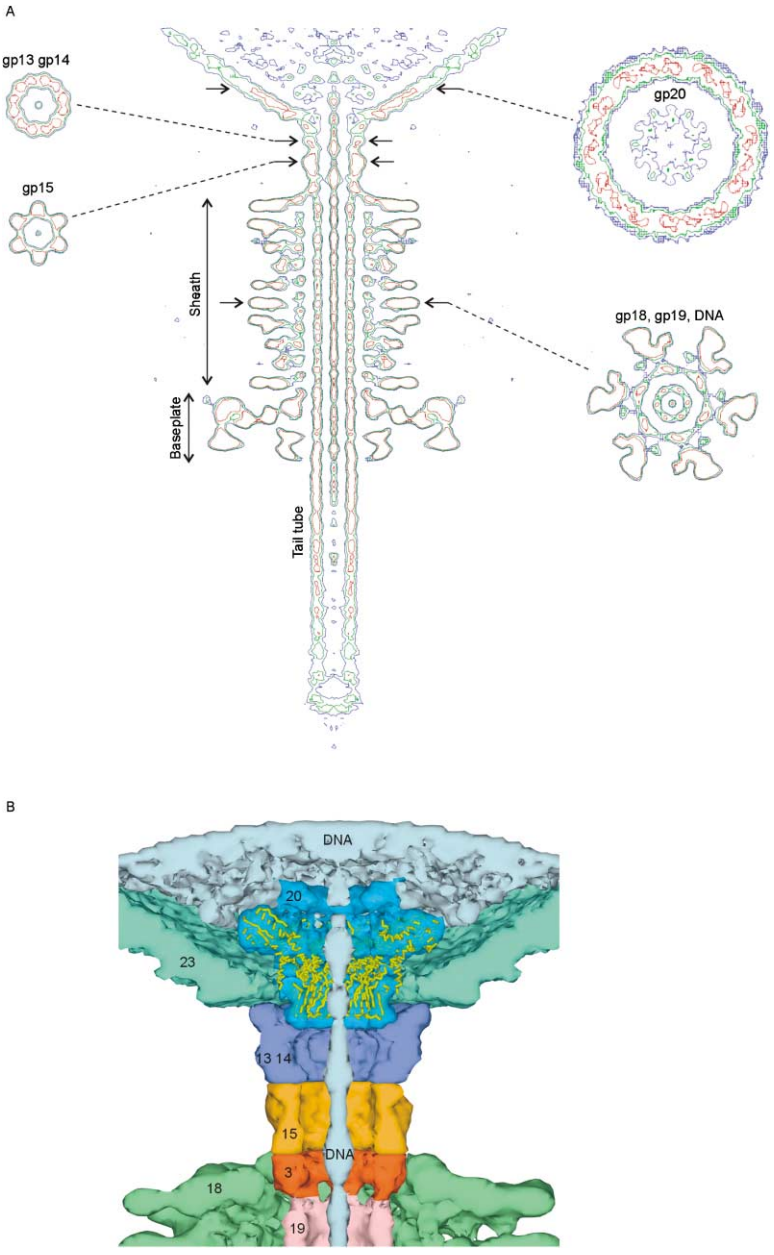


Figure 4. Structure of the Neck Region

(A) A cross-section along the 6-fold axis of the tail and four cross-sections orthogonal to it are shown. The long tail fiber density is beyond the borders of the chosen side-view slice. Three contour levels are given: 1.5 σ in blue, 3.0 σ in green, and 4 σ in red.

(B) A cut-away view of the neck region. The proteins are colored and labeled as in Figure 1A. For clarity, the density corresponding to the gp20 dodecamer is shown at a lower contour level than the rest of the map. The portal protein from phage ϕ29, shown as a yellow C_α trace, is fitted into the gp20 density. The assignment of densities to gp3, gp15, gp13, and gp14 is tentative due to a lack of additional structural and biochemical data.

gp54), whereas in the star conformation, this gp6-gp25-gp53 structure surrounds the tail tube (Figures 2C and 2D). The correlation coefficient found between the gp6-gp25-gp53 protein densities in the two conformations of the baseplate is only 0.45. This can be compared with a 0.92 correlation coefficient between the gp11 densities in both baseplate conformations, suggesting that the interactions among gp6, gp25, and gp53 are somewhat different in the two conformations of the baseplate. The distal part of the tail sheath connects to the gp6-gp25-gp53 complex, in agreement with earlier predictions (Kostyuchenko et al., 2003). Hence, the conformational change of this complex might initiate sheath contraction.

Six 230 Å-long stretches of the proximal part of the long tail fiber, which is composed of the gp34 trimer, emanate from the star-shaped baseplate (Figure 1). The fiber forms an angle of about 57° with respect to the 6-fold axis of the baseplate, consistent with earlier in vivo studies (Simon and Anderson, 1967a). The beaded structure of the visible part of the long tail fibers resembles that observed by negative staining EM of individual fibers (Cerritelli et al., 1996). The long tail fiber is attached to the baseplate via a collinear interaction with gp9, in agreement with earlier predictions (Kostyuchenko et al., 2003).

The Sheath

The cryoEM reconstruction of the sheath shows that it is composed of 138 subunits of gp18 arranged into 23 hexameric rings (Figure 3A). This number of gp18 subunits is different from that reported earlier (Abuladze et

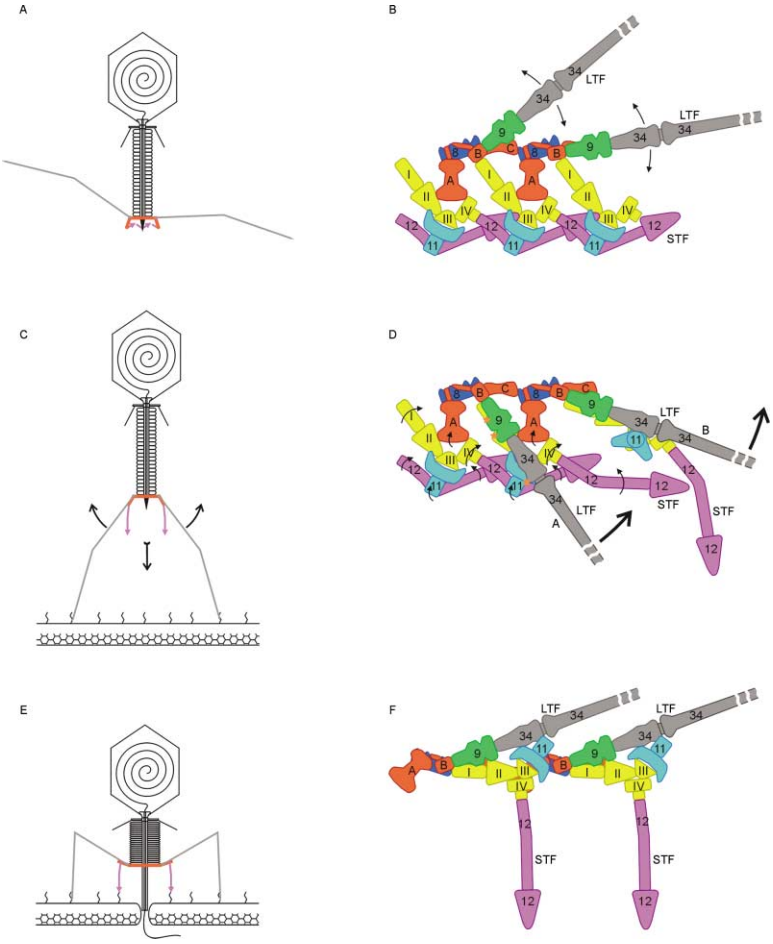


Figure 5. Baseplate Conformational Switch Schematics

(A and B) The phage is free in solution. The long tail fibers are extended and oscillate around their midpoint position. The movements of the fibers are indicated with black arrows. The proteins are labeled with their corresponding gene numbers and colored as in Figure 1A. Domains of gp7 and gp10 are labeled as in Figure 2A.

(C and D) The long tail fibers attach to their surface receptors and adapt the "down" conformation. The fiber labeled "A" and its corresponding attachment protein gp9 interact with gp11 and with gp10, respectively. These interactions, labeled with orange stars, probably initiate the conformational switch of the baseplate. The black arrows indicate tentative domain movements and rotations, which have been derived from the comparison of the two terminal conformations. The fiber labeled "B" has advanced along the conformational switch pathway so that gp11 is now seen along its 3-fold axis and the short tail fiber is partially extended in preparation for binding to its receptor. The thick red arrows indicate the projected movements of the fibers and the baseplate.

(E and F) The conformational switch is complete; the short tail fibers have bound their receptors and the sheath has contracted. The phage has initiated DNA transfer into the cell.

al., 1994; Amos and Klug, 1975), where the tail sheath was thought to be composed of 24 hexameric rings of gp18, resulting in 144 subunits. The extra ring is a 240 Å diameter platform on top of the baseplate's dome in the hexagonal conformation, formed by gp6, gp25, and gp53. As first noted by King (King, 1971), this platform resembles a ring of gp18 subunits of the extended sheath at low resolution. Apparently, this platform was included as a part of the sheath in subsequent studies (Abuladze et al., 1994). In the current cryoEM reconstruction, the diameter and structure of the contracted sheath are different from those of the platform, allowing for assignment of the sheath's terminal ring.

The gp18 subunits form a six start right-handed helix with a pitch of 16.4 Å and a twist angle of 32.9°, situated between radii of 60 Å and 165 Å (Figures 3A and 4A). These parameters resemble those of the polysheath helix (pitch 15 Å, twist angle 32°) (Amos and Klug, 1975), which was used as the starting model for the present reconstruction. The sheath has an inner diameter of 120 Å and does not interact with the 90 Å diameter tail tube, in agreement with previous observations (Amos and Klug, 1975). Upon superimposing the midsection of the sheath onto itself using the helical transformation, the correlation coefficient was found to be 0.98, showing that there is little variation in the structure of the gp18 subunits and that the sheath contracts uniformly.

The molecular envelope of a gp18 subunit was derived from analysis of the first and last terminal rings of the sheath (Figure 3C). These rings were superimposed onto each other, and 1/6th of the common volume was assigned to a gp18 monomer based on the connectivity seen in the cryoEM density. Each gp18 monomer was found to contain three lobes, with the outer two connected to the inner lobe by a weak density located at about 100 Å radius from the 6-fold axis (Figures 3C and 4A). Proteolysis studies suggest that the larger and more exposed lobe of gp18 is in the N-terminal domain (residues 82 through 320), whereas the other two lobes correspond to the C-terminal domain (residues 361 through 659) (Arisaka et al., 1990; Takeda et al., 1990).

The two external lobes of each gp18 subunit (except for the terminal gp18 rings) interact with their four nearest neighboring subunits, whereas the internal lobes form a six start helix with few interactions between neighboring strands (Figure 3B). The contacts between the inner lobes within each strand appear to be more extensive than those between the outer lobes, which is consistent with the observation that the sheath can be depolymerized by osmotic shock into individual strands of gp18 subunits in which the interactions between the inner lobes would be preserved (Coombs and Arisaka, 1994). It is possible, therefore, that the interactions between the inner lobes of gp18 persist in both conforma-

Table 1. Uniqueness of the Location of the Fitted Proteins in the Star-Shaped Baseplate

	Best Fit			Second Best Fit		
	COLORES	EMfit		COLORES	EMfit	
Protein	Best Peak (σ)	Sumf (%)	Negative Density	Highest Noise Peak (σ)	Sumf	Negative Density
gp8	10.5	29.7	1.1	9.8	12.5	35.7
gp9	18.7	27.4	0.5	10.8	16.1	16.3
gp11	10.0	26.1	0.0	9.6	25.7	11.1
gp7	9.3	20.4	13.2	7.1	15.3	26.5
gp10	8.5	18.0	13.4	7.7	12.5	31.2

A comparison is given between the best and second best fits. The program COLORES (Chacon and Wriggers, 2002) represents the fit in terms of the number of standard deviations above background. The program EMfit (Rossmann et al., 2001) measures the fit (Sumf) in terms of the average density at all fitted Cα atoms (where the structure is known) or pseudo atoms distributed over the designated volume (where the structure is unknown) normalized by the maximum density fixed to 100. The EMfit program also measures the percentage of atoms in negative density.

tions of the sheath. Thus, in the noncontracted sheath, this helix should have a smaller diameter, causing it to interact with the tail tube (Amos and Klug, 1975).

The Tail Tube

The tail tube is a 920 Å-long cylinder with external and internal diameters of about 40 Å and 90 Å, respectively (Figures 1 and 4). The distal end of the tail tube is terminated by a 120 Å diameter, 80 Å-long cap (Figures 1B and 4A). The cell-puncturing device formed by the gp5-gp27 complex (Figure 2C) is not present in the cryoEM reconstruction, which is consistent with earlier results (Duda et al., 1985). These EM and immunological studies showed that this complex dissociates from the tail tube in the presence of 3 M urea and that gp48 and, possibly, gp54 are situated at the tube's tip. Therefore, the cap terminating the tube is probably composed of gp48 and, possibly, gp54 and may act as a plug to stop DNA exit in the mature phage.

There is a continuous density within the tube channel with a diameter of about 15 Å (Figures 1B and 4A). This density can be attributed to either the tape measure protein gp29 or the phage DNA. Contrary to earlier observations (Abuladze et al., 1994; Duda et al., 1985), there was no density in the tail tube channel of the hexagonal baseplate-tail tube cryoEM reconstruction (Kostyuchenko et al., 2003). Thus, the density observed in the tube of the urea-treated particles may be phage DNA, one end of which extends from the capsid in preparation for injection into the host cell.

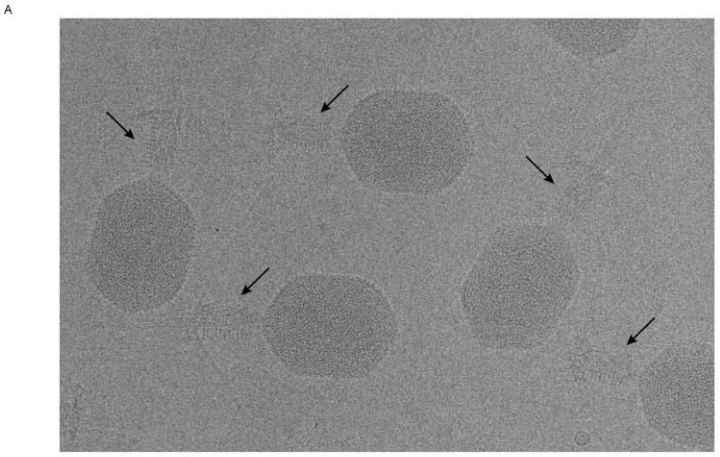
Using 80 Å resolution X-ray diffraction of tubes oriented in gels, Moody and Makowski (1981) had shown that the tail tube is a six start helix with a 41 Å-long repeating unit and a twist angle of about 18° between the successive repeats. Because these helical parameters coincide with those of the noncontracted sheath, it had been suggested that the sheath is assembled onto the tail tube with each gp18 subunit interacting with a corresponding gp19 subunit (Coombs and Arisaka, 1994). However, the reconstruction presented here shows that the tube is constructed of 61.8 Å-long repeating units arranged with a twist angle of 2.5° (Figure 3D). These helical parameters bear no relationship to those of Moody and Makowski (1981), but visual inspection shows that a translation of 41 Å and a rotation of 18° are inconsistent with the cryoEM results. The absence of the symmetry match between the tail tube and the sheath allows the sheath to move along the tube without a potential energy barrier. Such symmetry mismatches are frequently encountered in biological macromolecular assemblies when one part of the assembly moves relative to another (Simpson et al., 2000).

The Neck

The neck region, connecting the tail to the dodecameric portal protein gp20 situated at the unique head vertex, is composed of gp3, gp15, gp13, gp14, and gp *wac* (fibritin). Earlier results had suggested that gp3 terminates the tail tube and interacts with gp15 and gp18 (Kikuchi and King, 1975a; Vianelli et al., 2000; Zhao et al., 2003). Hence, the hexameric ring that terminates the proximal end of the tail tube, whose structure is different than the rings of gp18, is probably gp3 (Figures 3 and 4). The dimensions of this ring are consistent with those of recombinant gp3 hexamers (Zhao et al., 2003).

In the absence of gp15, the tail and head do not join and the sheath depolymerizes from the tube (King, 1968). Gp15 extends from the top of the sheath by about 60–80 Å (Coombs and Eiserling, 1977; King, 1968) and forms a 130 Å diameter hexamer (Zhao et al., 2003). At about this distance from the top of the sheath, the symmetry of the density in the neck region switches from being 6-fold to 12-fold, although only 6-fold symmetry was applied during the reconstruction. Thus, the 6-fold-symmetric density extending from the gp3 hexamer was assigned to gp15 (Figures 4A and 4B).

The dodecameric portal protein gp20 occupies the unique vertex of the capsid (Black et al., 1994). The portal connector assembly was found to have a similar core structure for two unrelated tailed phages ϕ29 and SPP1 (Orlova et al., 2003). Hence, the assumption that the T4 connector has also a similar core structure would seem reasonable. Thus, the density, corresponding to the gp20 connector (MW = 61.0 kDa), was identified by fitting the crystal structure of the ϕ29 connector (MW = 35.9 kDa) (Simpson et al., 2000) into the cryoEM map (Figure 4B). Gp20, however, is a larger protein, and some



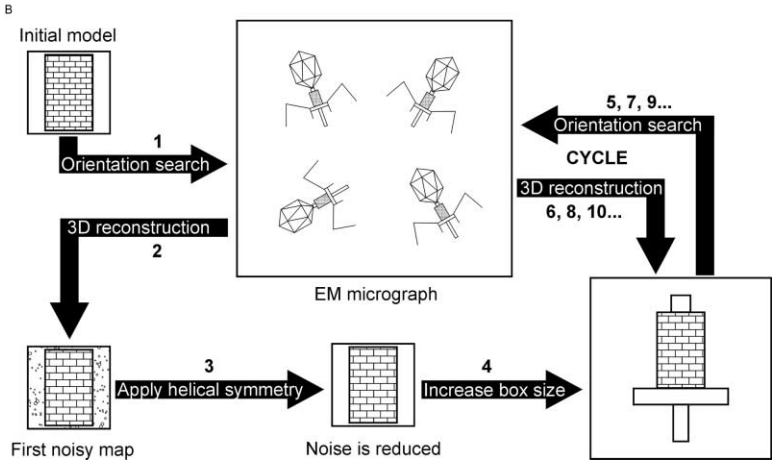


Figure 6. Cryo-Electron Microscopy Image Reconstruction

(A) Electron micrograph of urea-treated T4 particles, out of focus 2.6 μm. Black arrows point to the contracted tails.

(B) Flowchart of the contracted tail reconstruction procedure.

of the gp20 density is not occupied by the ϕ29 portal protein structure.

The region of the neck situated between the putative gp15 and gp20 densities has pseudo 12-fold symmetry and, based on its volume, would have a mass of about 230 kDa, which corresponds to 19 kDa per monomer assuming 12-fold symmetry or 36 kDa per monomer assuming 6-fold symmetry. This density could represent either gp13 (MW = 35 kDa) or gp14 (MW = 30 kDa) proteins, which participate in attachment of the assembled tail to the head (Figures 4A and 4B). The reconstruction shows very little density that can be attributed to the whiskers (Coombs and Arisaka, 1994).

Discussion

It has been shown that purified baseplates switch spontaneously into the star conformation (Watts et al., 1990), and in the absence of either the baseplate or the tail tube, the sheath assembles into a long tubular structure similar to that of the contracted sheath (Amos and Klug, 1975). Furthermore, the tail sheath contraction is irreversible, and the contracted tail structure is resistant to 8 M urea (Arisaka et al., 1981). These observations suggest that the baseplate in the hexagonal conformation with its extended sheath can be compared to an extended spring ready to be triggered (Caspar, 1980) and thus represent a high energy, metastable state of the structure.

The wealth of experimental information on T4 infection, combined with the structural results presented here, permits the analysis of the mechanism that controls the concerted macromolecular interactions, which occur in the baseplate when the phage attaches to the host cell (Figure 5). The attachment of at least three long tail fibers to their host cell receptors results in these fibers pointing toward the host cell surface (Figure 5C) (Simon and Anderson, 1967a). This initiates the infection process (Goldberg et al., 1994) and the associated conformational changes by the concerted effect of these fibers altering their position relative to the proteins in the baseplate (Supplemental Movie S1 at http://www.cell.com/cgi/content/full/118/4/419/DC1). The axes of the gp9 trimers are then directed along the axes of the proximal halves of the long tail fibers and bring gp9 into contact with domains I and/or II of gp10 (Figure 5D). The long tail fibers are then able to interact with the gp11 trimers, allowing the long tail fibers and gp9 to become attached to the baseplate pins (Figure 5D). The initial binding of gp9 to gp10 and the long tail fibers to gp11 may cause a rotation of gp10, which, in turn, triggers the receptor-attachment domains of the short tail

fibers to dissociate from under the baseplate, thus preparing them for binding to the host cell receptors (Figure 5D). The rotation of gp10 results in a simultaneous reassociation of gp10 with the A domain of gp7, causing the latter to twist and the pins, with their attached long tail fibers, to swivel outwards. Thus, the long tail fibers, bound to the host cell receptors, are being used as levers to move the baseplate toward the cell surface (Figures 5E and 5F). The overall change in the angle of the proximal part of the long tail fibers is about 100°, which moves the baseplate into close proximity with the host cell surface.

The conformational changes, which have been initiated at the periphery of the baseplate, probably spread toward the gp6-gp25-gp53 assembly. The latter, in turn, alters its conformation, thus initiating sheath contraction. The contracting sheath pushes the tail tube, carrying the baseplate hub containing the gp5 membrane-puncturing needle and its three lysozyme domains, through the baseplate into the periplasm (Kanamaru et al., 2002). The lysozyme domains digest the peptidoglycan, thus allowing the tube to reach the cytoplasmic membrane and initiate DNA transfer (Figure 5F and Supplemental Movie S2 on *Cell* website).

The execution of the events described above has been perfected to serve the complex processes required by the highly efficient T4 machine. The building blocks used in T4 tail are the individual protein domains that move relative to each other during the infection process. Similar rigid body rearrangements of proteins employed in the performance of other essential biological functions are used in the maturation of viruses during genome packaging (Conway et al., 2001), in some complex chaperones for protein folding (Xu and Sigler, 1998), and in ATPases for pumping protons across cellular membranes (Abrahams et al., 1994). The evolution of such complex biological machines results in a gradual increase in complexity as their function improves.

Experimental Procedures

Preparation of Phages with Contracted Tails for cryoEM

A T4 phage sample of 1 ml with a titer of about 10^{11} plaque forming units per ml was diluted 10-fold by 3 M urea buffered with 50 mM Tris-HCl at pH 8.0 complemented with 1 mM $MgCl_2$ and incubated for 2 hr at +4°C.

DNase I was added to the sample to obtain a final concentration of 30 μg/ml in order to remove DNA from broken phage particles. The phage was precipitated by centrifugation for 1 hr at 75,000 × g and 4°C. The pellet was subsequently dissolved in 100 μl of water.

Electron Microscopy

Low-dose cryoEM was performed as described by Baker et al. (1999). Images of frozen-hydrated samples were recorded on Kodak film using a Philips CM300 FEG microscope at a magnification of 47,000 and a radiation dose of 21–31 $e^-/Å^2$. The images (Figure 6A) were digitized with a ZI scanner.

CryoEM and Image Reconstruction

The three-dimensional reconstruction of the bacteriophage T4 contracted tail has been obtained using the program SPIDER (Frank et al., 1996). The initial model contained only the sheath portion of the tail, which was modeled as a six start helix with the helical pitch and twist parameters taken from the structure of the polysheath obtained by negative staining EM (Amos and Klug, 1975). The external features were derived from a series of sections of the polysheath drawn by Amos and Klug (Amos and Klug, 1975). The resulting model, consisting of 144 subunits of gp18 represented with uniform density, was used to find the orientations of the phage particles for an initial image reconstruction (Figure 6B). This noisy map was improved by averaging using the apparent helical parameters and then refined with the help of the programs IMP and AVE (Kleywegt and Jones, 1999).

Subsequently, the model was modified to include the baseplate, a part of the tail tube, and the neck (Figure 6B). The baseplate was modeled as a 600 Å diameter, 250 Å-long disk of density, attached to one of the sheath's ends. The tail tube and the neck were also modeled as disks with diameters of 90 Å and 140 Å, respectively. The size of the reconstructed volume was also increased to 169 × 169 × 169 pixels at 5.96 Å/pixel. The sheath structure was then used to find the particle orientations. Once the features of the baseplate and the neck became apparent, they were also included in the orientation searches. Based on the crosscorrelation coefficient between the observed projections, the data set was eventually reduced from 2500 to 1965 particles. The reconstruction was refined using a pixel size of 2.98 Å. A reconstruction containing the entire tail tube, which required a box size of 1877 Å, was also calculated using a 4.51 Å pixel size. The final resolution of the reconstruction was found to be 17 Å for the baseplate and 16 Å for the tail sheath using the Fourier shell correlation coefficient method, assuming the limit of resolution to be where the correlation coefficient drops below 0.5.

Fitting the Crystal Structures into the CryoEM Map

The crystal structures of gp8 (Leiman et al., 2003b), gp9 (Kostyuchenko et al., 1999), and gp11 (Leiman et al., 2000) and the cryoEM densities of gp7 and gp10 derived from the hexagonal baseplate (Kostyuchenko et al., 2003) were computationally fitted into the star-shaped baseplate map (Table 1). Due to the 6-fold symmetry of the baseplate, the searches were limited to about 1/4th of the baseplate volume.

The analysis of the contracted tail was initiated from interpretation of the baseplate density. Three (gp8 [Leiman et al., 2003b], gp9 [Kostyuchenko et al., 1999], and gp11 [Leiman et al., 2000]) out of six available crystal structures of the baseplate proteins (gp5 [Kanamaru et al., 2002], gp8 [Leiman et al., 2003b], gp9 [Kostyuchenko et al., 1999], gp11 [Leiman et al., 2000], gp12 [Thomassen et al., 2003], and gp27 [Kanamaru et al., 2002]) and the densities derived for gp7 and gp10 from the cryoEM map of the hexagonal baseplate (Kostyuchenko et al., 2003) were fitted into the star-shaped baseplate (Figure 1C). The part of the short tail fiber that would correspond to the available crystal structure of its C-terminal domain (Thomassen et al., 2003) is disordered in the reconstruction of the star-shaped baseplate (Figures 1 and 2). The density comprising the central portion of the baseplate, assigned to gp6, gp25, and gp53 in the hexagonal conformation of the baseplate, was then compared to that in the star-shaped structure. Once the map corresponding to the baseplate had been interpreted, the limits of the repetitive oligomeric sheath were readily established.

Acknowledgments

We thank Sharon Wilder and Cheryl Towell for help in the preparation of the manuscript. The work was supported by grants from the Keck foundation (to M.G.R. and Tim Baker), the National Science Foundation (to M.G.R.), the Human Frontier Science Program (to M.G.R., V.V.M., and Fumio Arisaka), and from the Howard Hughes Medical Institutes (to V.V.M.). We thank W. Scott Meador, Lee Gooding, and James A. Bartek of Purdue University for the creation of the two movies provided in Supplemental Data online.

Received: April 18, 2004
Revised: June 23, 2004
Accepted: June 28, 2004
Published: August 19, 2004

References

Abrahams, J.P., Leslie, A.G., Lutter, R., and Walker, J.E. (1994). Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature *370*, 621–628.

Abuladze, N.K., Gingery, M., Tsai, J., and Eiserling, F.A. (1994). Tail length determination in bacteriophage T4. Virology *199*, 301–310.

Ackermann, H.W. (2003). Bacteriophage observations and evolution. Res. Microbiol. *154*, 245–251.

Amos, L.A., and Klug, A. (1975). Three-dimensional image reconstructions of the contractile tail of T4 bacteriophage. J. Mol. Biol. *99*, 51–64.

Arisaka, F., Engel, J., and Klump, H. (1981). Contraction and dissociation of the bacteriophage T4 tail sheath induced by heat and urea. Prog. Clin. Biol. Res. *64*, 365–379.

Arisaka, F., Takeda, S., Funane, K., Nishijima, N., and Ishii, S. (1990). Structural studies of the contractile tail sheath protein of bacteriophage T4. 2. Structural analyses of the tail sheath protein, Gp18, by limited proteolysis, immunoblotting, and immunoelectron microscopy. Biochemistry *29*, 5057–5062.

Baker, T.S., Olson, N.H., and Fuller, S.D. (1999). Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol. Mol. Biol. Rev. *63*, 862–922.

Black, L.W., Showe, M.K., and Steven, A.C. (1994). Morphogenesis of the T4 head. In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (Washington, D.C.: American Society for Microbiology), pp. 218–258.

Caspar, D.L. (1980). Movement and self-control in protein assemblies. Quasi-equivalence revisited. Biophys. J. *32*, 103–138.

Cerritelli, M.E., Wall, J.S., Simon, M.N., Conway, J.F., and Steven, A.C. (1996). Stoichiometry and domainal organization of the long tail-fiber of bacteriophage T4: a hinged viral adhesin. J. Mol. Biol. *260*, 767–780.

Chacon, P., and Wriggers, W. (2002). Multi-resolution contour-based fitting of macromolecular structures. J. Mol. Biol. *317*, 375–384.

Conway, J.F., Wikoff, W.R., Cheng, N., Duda, R.L., Hendrix, R.W., Johnson, J.E., and Steven, A.C. (2001). Virus maturation involving large subunit rotations and local refolding. Science *292*, 744–748.

Coombs, D.H., and Eiserling, F.A. (1977). Studies on the structure, protein composition and assembly of the neck of bacteriophage T4. J. Mol. Biol. *116*, 375–405.

Coombs, D.H., and Arisaka, F. (1994). T4 tail structure and function. In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (Washington, D.C.: American Society for Microbiology), pp. 259–281.

Crawford, J.T., and Goldberg, E.B. (1980). The function of tail fibers in triggering baseplate expansion of bacteriophage T4. J. Mol. Biol. *139*, 679–690.

Crowther, R.A., Lenk, E.V., Kikuchi, Y., and King, J. (1977). Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4. J. Mol. Biol. *116*, 489–523.

Desplats, C., and Krisch, H.M. (2003). The diversity and evolution of the T4-type bacteriophages. Res. Microbiol. *154*, 259–267.

Driedonks, R., Engel, A., tenHeggeler, A., and van Driel, R. (1981). Gene 20 product of bacteriophage T4 its purification and structure. J. Mol. Biol. *152*, 641–662.

Duda, R.L., Wall, J.S., Hainfeld, J.F., Sweet, R.M., and Eiserling, F.A. (1985). Mass distribution of a probable tail-length-determining protein in bacteriophage T4. Proc. Natl. Acad. Sci. USA *82*, 5550–5554.

Eiserling, F.A., and Black, L.W. (1994). Pathways in T4 morphogenesis. In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (Washington, D.C.: American Society for Microbiology), pp. 209–212.

Fokine, A., Chipman, P.R., Leiman, P.G., Mesyanzhinov, V.V., Rao, V.B., and Rossmann, M.G. (2004). Molecular architecture of the prolate head of bacteriophage T4. Proc. Natl. Acad. Sci. USA, in press.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J. Struct. Biol. *116*, 190–199.

Goldberg, E., Grinius, L., and Letellier, L. (1994). Recognition, attachment, and injection. In Molecular Biology of Bacteriophage T4, J.D. Karam, ed. (Washington, D.C.: American Society for Microbiology), pp. 347–356.

Gowen, B., Bamford, J.K., Bamford, D.H., and Fuller, S.D. (2003). The tailless icosahedral membrane virus PRD1 localizes the proteins involved in genome packaging and injection at a unique vertex. J. Virol. *77*, 7863–7871.

Hendrix, R.W. (2002). Bacteriophages: evolution of the majority. Theor. Popul. Biol. *61*, 471–480.

Kanamaru, S., Leiman, P.G., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2002). Structure of the cell-puncturing device of bacteriophage T4. Nature *415*, 553–557.

Kellenberger, E., Stauffer, E., Haner, M., Lustig, A., and Karamata, D. (1996). Mechanism of the long tail-fiber deployment of bacteriophages T-even and its role in adsorption, infection and sedimentation. Biophys. Chem. *59*, 41–59.

Kikuchi, Y., and King, J. (1975a). Assembly of the tail of bacteriophage T4. J. Supramol. Struct. *3*, 24–38.

Kikuchi, Y., and King, J. (1975b). Genetic control of bacteriophage T4 baseplate morphogenesis. I. Sequential assembly of the major precursor, in vivo and in vitro. J. Mol. Biol. *99*, 645–672.

Kikuchi, Y., and King, J. (1975c). Genetic control of bacteriophage T4 baseplate morphogenesis. II. Mutants unable to form the central part of the baseplate. J. Mol. Biol. *99*, 673–694.

Kikuchi, Y., and King, J. (1975d). Genetic control of bacteriophage T4 baseplate morphogenesis. III. Formation of the central plug and overall assembly pathway. J. Mol. Biol. *99*, 695–716.

King, J. (1968). Assembly of the tail of bacteriophage T4. J. Mol. Biol. *32*, 231–262.

King, J. (1971). Bacteriophage T4 tail assembly: four steps in core formation. J. Mol. Biol. *58*, 693–709.

Kleywegt, G.J., and Jones, T.A. (1999). Software for handling macromolecular envelopes. Acta Crystallogr. D *55*, 941–944.

Kostyuchenko, V.A., Navruzbekov, G.A., Kurochkina, L.P., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. (1999). The structure of bacteriophage T4 gene product 9: the trigger for tail contraction. Struct. Fold. Des. *7*, 1213–1222.

Kostyuchenko, V.A., Leiman, P.G., Chipman, P.R., Kanamaru, S., van Raaij, M.J., Arisaka, F., Mesyanzhinov, V.V., and Rossmann, M.G. (2003). Three-dimensional structure of bacteriophage T4 baseplate. Nat. Struct. Biol. *10*, 688–693.

Leiman, P.G., Kostyuchenko, V.A., Shneider, M.M., Kurochkina, L.P., Mesyanzhinov, V.V., and Rossmann, M.G. (2000). Structure of bacteriophage T4 gene product *11*, the interface between the baseplate and short tail fibers. J. Mol. Biol. *301*, 975–985.

Leiman, P.G., Kanamaru, S., Mesyanzhinov, V.V., Arisaka, F., and Rossmann, M.G. (2003a). Structure and morphogenesis of bacteriophage T4. Cell. Mol. Life Sci. *60*, 2356–2370.

Leiman, P.G., Shneider, M.M., Kostyuchenko, V.A., Chipman, P.R., Mesyanzhinov, V.V., and Rossmann, M.G. (2003b). Structure and location of gene product 8 in the bacteriophage T4 baseplate. J. Mol. Biol. *328*, 821–833.

Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Ruger, W. (2003). Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. *67*, 86–156.

Monod, C., Repoila, F., Kutateladze, M., Tetart, F., and Krisch, H.M. (1997). The genome of the pseudo T-even bacteriophages, a diverse group that resembles T4. J. Mol. Biol. *267*, 237–249.

Moody, M.F. (1967). Structure of the sheath of bacteriophage T4. I. Structure of the contracted sheath and polysheath. J. Mol. Biol. *25*, 167–200.

Moody, M.F. (1973). Sheath of bacteriophage T4. 3. Contraction mechanism deduced from partially contracted sheaths. J. Mol. Biol. *80*, 613–635.

Moody, M.F., and Makowski, L. (1981). X-ray diffraction study of tail-tubes from bacteriophage T2L. J. Mol. Biol. *150*, 217–244.

Moore, S.D., and Prevelige, P.E., Jr. (2002). DNA packaging: a new class of molecular motors. Curr. Biol. *12*, R96–R98.

Newcomb, W.W., Juhas, R.M., Thomsen, D.R., Homa, F.L., Burch, A.D., Weller, S.K., and Brown, J.C. (2001). The UL6 gene product

forms the portal for entry of DNA into the herpes simplex virus capsid. J. Virol. *75*, 10923–10932.

Orlova, E.V., Gowen, B., Droge, A., Stiege, A., Weise, F., Lurz, R., van Heel, M., and Tavares, P. (2003). Structure of a viral DNA gatekeeper at 10 A resolution by cryo-electron microscopy. EMBO J. *22*, 1255–1262.

Rossmann, M.G., Bernal, R., and Pletnev, S.V. (2001). Combining electron microscopic with x-ray crystallographic structures. J. Struct. Biol. *136*, 190–200.

Simon, L.D., and Anderson, T.F. (1967a). The infection of Escherichia coli by T2 and T4 bacteriophages as seen in the electron microscope. I. Attachment and penetration. Virology *32*, 279–297.

Simon, L.D., and Anderson, T.F. (1967b). The infection of Escherichia coli by T2 and T4 bacteriophages as seen in the electron microscope. II. Structure and function of the baseplate. Virology *32*, 298–305.

Simpson, A.A., Tao, Y., Leiman, P.G., Badasso, M.O., He, Y., Jardine, P.J., Olson, N.H., Morais, M.C., Grimes, S., Anderson, D.L., et al. (2000). Structure of the bacteriophage phi29 DNA packaging motor. Nature *408*, 745–750.

Takeda, S., Arisaka, F., Ishii, S., and Kyogoku, Y. (1990). Structural studies of the contractile tail sheath protein of bacteriophage T4. 1. Conformational change of the tail sheath upon contraction as probed by differential chemical modification. Biochemistry *29*, 5050–5056.

Tao, Y., Strelkov, S.V., Mesyanzhinov, V.V., and Rossmann, M.G. (1997). Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure *5*, 789–798.

Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J.P., Miller, S., and van Raaij, M.J. (2003). The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. J. Mol. Biol. *331*, 361–373.

Vianelli, A., Wang, G.R., Gingery, M., Duda, R.L., Eiserling, F.A., and Goldberg, E.B. (2000). Bacteriophage T4 self-assembly: localization of gp3 and its role in determining tail length. J. Bacteriol. *182*, 680–688.

Watts, N.R., Hainfeld, J., and Coombs, D.H. (1990). Localization of the proteins gp7, gp8 and gp10 in the bacteriophage T4 baseplate with colloidal gold:F(ab)2 and undecagold:Fab' conjugates. J. Mol. Biol. *216*, 315–325.

Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. *64*, 69–114.

Xu, Z., and Sigler, P.B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. J. Struct. Biol. *124*, 129–141.

Zhao, L., Kanamaru, S., Chaidirek, C., and Arisaka, F. (2003). P15 and P3, the tail completion proteins of bacteriophage T4, both form hexameric rings. J. Bacteriol. *185*, 1693–1700.

Accession Numbers

The coordinates of the fitted proteins gp8, gp9, and gp11 have been deposited with the Protein Data Bank, accession number 1TJA. The cryoEM density has been submitted to the Macromolecular Structure Database at the European Bioinformatics Institute, accession number EMD-1086.