Structural Characterization of the Bacteriophage T7 Tail Machinery**

Ana Cuervo†, Mar Pulido-Cid†, Mónica Chagoyen†, Rocio Arranz‡, Verónica A. González-García†, Carmela García-Doval§, José R. Castón†, José M. Valpuesta†, Mark J. van Raaij‡, Jaime Martín-Benito‡, and José L. Carrasco*†‡

From the †Structure of Macromolecules and §Systems Biology Departments, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Darwin 3 and the ‡Instituto Madrileño de Estudios Avanzados en Nanociencia, Cantoblanco, 28049 Madrid, Spain

Most bacterial viruses need a specialized machinery, called “tail,” to inject their genomes inside the bacterial cytoplasm without disrupting the cellular integrity. Bacteriophage T7 is a well characterized member of the Podoviridae family infecting Escherichia coli, and it has a short noncontractile tail that assembles sequentially on the viral head after DNA packaging. The T7 tail is a complex of around 2.7 MDa composed of at least four proteins as follows: the connector (gene product 8, gp8), the tail tubular proteins gp11 and gp12, and the fibers (gp17). Using cryo-electron microscopy and single particle image reconstruction techniques, we have determined the precise topology of the tail proteins by comparing the structure of the T7 tail extracted from viruses and a complex formed by recombinant gp8, gp11, and gp12 proteins. Furthermore, the order of assembly of the structural components within the complex was deduced from interaction assays with cloned and purified tail proteins. The existence of common folds among similar tail proteins allowed us to obtain pseudo-atomic threaded models of gp8 (connector) and gp11 (gatekeeper) proteins, which were docked into the corresponding cryo-EM volumes of the tail complex. This pseudo-atomic model of the connector-gatekeeper interaction revealed the existence of a common molecular architecture among viruses belonging to the three tailed bacteriophage families, strongly suggesting that a common molecular mechanism has been favored during evolution to coordinate the transition between DNA packaging and tail assembly.

Double-stranded DNA (dsDNA) bacteriophages actively package their genetic material inside the capsid using a protein motor that hydrolyzes ATP as an energy source (1). DNA enters the viral head through a channel formed by a protein named connector that sits at a unique 5-fold vertex of the icosahedral capsid (2), a vertex that is also involved in the delivery of the genome during DNA ejection. The ordered packaged DNA generates a high pressure inside the viral head requiring the presence of protein complexes helping to retain the nucleic acid inside the capsid, probably by closing and/or securing the connector channel. In most of the dsDNA viruses, this task is accomplished by a protein or proteins that form a complex that builds a plug at the end of the portal channel, the so-called gatekeeper (3). These proteins also play an important role during the transition between final DNA packaging steps and the subsequent assembly of the tail proteins (3). The tail is a protein complex present in the majority of bacteriophages, which is involved in host recognition and genome delivery. Three main morphologies have been described in tailed phages as follows: short and long noncontractile or long contractile corresponding to Podoviridae, Siphoviridae, and Myoviridae families, respectively (4). Although there is a necessary adaptation of the tail structure of the virus to each viral host family (5), two main features are shared by tail structures; they have a central tubular structure that forms the channel for DNA ejection, which is surrounded by fibers or spikes that are essential in the initial steps of host recognition (5).

The Podoviridae family constitutes the group with the simplest tail morphology, composed of a small number of proteins (6). Structural analysis of different viruses belonging to this group (φ29 (7), P22 (8), P-SPP7 (9), N4 (10), e15 (11), and K1E and K1–5 (12)) has shown that these structures are formed by a central knob or nozzle contoured by 6–12 trimeric appendages named fibers or spikes. The tail is attached to the head through the tail adaptor protein, a dodecameric ring with similarities to the gatekeeper proteins described for long-tailed phages (13). The short length of the Podoviridae requires additional components to drill through the whole bacterial envelope, and it has...
been proposed to be enlarged by internal capsid proteins, which
would be also ejected during infection (6, 14). This hypothesis
has been recently supported by a cryo-electron tomography
study, which has defined for the first time a detailed description
of the sequential steps involved in the interaction of the bacte-
riophage T7 with *Escherichia coli* (15). The bacteriophage T7
infects bacteria by interaction of the tail with one or more bac-
terial receptors, followed by DNA ejection, and accompanied
by extensive structural remodeling of the tail structure (15).
Nevertheless, the limited resolution of previous studies from
complete viral particles (16, 17) has not yet identified the loca-
tion and precise topology of the different proteins that constitu-
t the tail structure. Biochemical studies have proposed that
the T7 tail machine is formed by proteins gp7.3 (10 kDa), gp8
(59 kDa), gp11 (22 kDa), gp12 (89 kDa), and gp17 (61 kDa) (14).
The best structurally characterized proteins are the connector
(gp8) and the fibers (built by protein gp17). Purified gp8 protein
has been shown to assemble as a dodecamer (18), and gp17 has
been shown to assemble into trimers to form the T7 fibers (19–
21). Proteins gp11 and gp12 have been proposed to form the
tubular structure of the tail, but their precise location
remains unclear (14). Finally, 32 subunits of the protein gp7.3
have been predicted as a component of the tail, but its location
and function are controversial. This protein might be involved
either in helping the tail protein assembly process or as a part of
the tail channel or tip, thus contributing to the host cell inter-
action (6, 14), and it has also been proposed to be injected inside
the bacteria during infection (14).

The biochemical and structural studies carried out in this
work have permitted us to define the location and oligomeric
state of each of the proteins that form the tail of bacteriophage
T7 and to understand their sequential assembly order into the
mature DNA-filled viral heads.

**EXPERIMENTAL PROCEDURES**

**Protein Cloning**—The genes 11 and 12 from T7 bacterio-
phage were amplified together or individually by polymerase
chain reaction (PCR) using forward and reverse primers con-
taining the BglIII-Ncol restriction sites, respectively, with the
Expand Long Template PCR System (Roche Applied Science).
The fragments were then cloned into the pRSET-B plasmid
(Invitrogen) using T4 ligase (Roche Applied Science) overnight
at 16 °C. The T7 gene 8 was amplified using the same system
with the primers containing the EcoRI-HindIII restriction sites,
and it was then cloned into the pRSET-B plasmid containing
the T7 11 and 12 genes. The gene 17, coding for the fiber pro-
tein, was amplified by PCR using forward and reverse primers
containing Ncol and HindIII restriction sites, respectively, and
cloned into the plasmid pCR2.1-TOPO using topoisomerase I
from vaccinia virus (TOPO TA cloning kit, Invitrogen). The
cloned into the plasmid pCR2.1-TOPO using topoisomerase I
containing NcoI and HindIII restriction sites, respectively, and
cloned into an ion exchange chromatography on a ResourceQ column
(Invitrogen) using T4 ligase (Roche Applied Science) overnight
for 15 h. The culture was harvested, and the 8-11-12 complexes were
removed by centrifugation at 25,000 × g for 15 min. gp12 was
precipitated overnight at 4 °C by addition of ammonium sulfate
to a final concentration of 40% and pelleted at 6500 × g for 1 h.
The pellet was resuspended in TMS buffer or 50 mM HEPES, pH 7.5, and 150 mM NaCl containing 10 mM imidazole
(no differences were observed between the two buffers). The
8-11-12 complexes, His-tagged at the N-terminal side of gp11
protein, were incubated overnight at 4 °C with Cobalt TALON
metal affinity resin (Clontech) and eluted at 1 M imidazole. The
enriched eluted fractions were then concentrated on Amicon
ultra centrifugal filter units (Millipore) and loaded onto a Super-
ose 6 10/300 GL column as described above. The purity of the
samples was checked by SDS-PAGE as described previously
(14).

**Protein Purification**—Bacteriophage T7 was produced as
described previously (22). Briefly *E. coli* BL21 strain was
infected with wild type phage, and the lysate was purified in a
cesium chloride step gradient (23) and dialyzed using VISKING
dialysis tubing (SERVA) on 50 mM Tris, pH 7.8, 10 mM MgCl₂,
and 100 mM NaCl (TMS buffer). The tail complexes were
obtained by incubating the T7 viruses in the presence of 100
mM EDTA and complete antiproteases mixture (Roche Applied
Science) for 45 min at 65 °C. The complexes were comple-
mented with 20 mM MgCl₂, incubated with 10 μg/ml DNase
(Sigma) for 30 min at 37 °C, and pelleted by ultracentrifugation
at 213,000 × g for 3 h. The complexes were resuspended in
TMS buffer, loaded onto a 10–40% sucrose gradient, and cen-
trifuged at 148,000 × g for 105 min. The enriched fractions
were concentrated using Amicon ultracentrifugal filter units
(Millipore) and loaded onto a Superose 6 10/300 GL column for
size exclusion chromatography (GE Healthcare). The sample
purity was checked by SDS-PAGE as described previously
(14).

The plasmid coding for the gp8, gp11, and gp12 proteins
(8-11-12 complex) was transformed into the *E. coli* C41 strain.
The culture was grown at 37 °C to an absorbance of 0.6 mea-
sured at 600 nm (A₆₀₀), and then protein expression was
induced with 1 mM isopropyl 1-thio-β-d-galactopyranoside for
3 h. The culture was harvested, and the 8-11-12 complexes were
found in the supernatant, which was supplemented with com-
plete EDTA-free protease inhibitors (Roche Applied Science)
and 1 mM phenylmethylsulfonyl fluoride (PMSF). The complex
was precipitated overnight at 4 °C by addition of ammonium
sulfate to a final saturation of 40% and pelleted at 6500 × g for 1 h.
The pellet was resuspended in TMS buffer or 50 mM
HEPES, pH 7.5, and 150 mM NaCl containing 10 mM imidazole
(no differences were observed between the two buffers). The
8-11-12 complexes, His-tagged at the N-terminal side of gp11
protein, were incubated overnight at 4 °C with Cobalt TALON
metal affinity resin (Clontech) and eluted at 1 M imidazole. The
enriched eluted fractions were then concentrated on Amicon
ultra centrifugal filter units (Millipore) and loaded onto a Super-
ose 6 10/300 GL column as described above. The purity of the
samples was checked by SDS-PAGE as described previously
(14).

gp8 protein purification was described previously (24). The
plasmids coding for gp11 or gp12 were transformed into *E. coli*
C41 strain and grown as described above. The cultures were
harvested and resuspended into buffer 50 mM Tris, pH 7.8, 500
mM NaCl, 20 mM MgCl₂, 2.5 mM β-mercaptoethanol, 10 mM
imidazole, with complete EDTA-free protease inhibitor mix-
ture. Cells were lysed by sonication, and the cellular debris was
removed by centrifugation at 25,000 × g for 15 min. gp12 was
precipitated overnight at 4 °C by addition of ammonium sulfate
to a final concentration of 40%, pelleted at 25,000 × g for 15
min, and resuspended in the same buffer. The N-terminal His-
tagged proteins were purified in a TALON affinity resin as
described above and eluted at 1 M imidazole. gp11 protein was
pure enough after this step; gp12 protein was dialyzed in buffer
50 mM Tris, pH 7.8, 50 mM NaCl, 20 mM MgCl₂, and purified
into an ion exchange chromatography on a ResourceQ column
(GE Healthcare), where the protein was eluted at 250 mM NaCl.
The purity of the proteins was checked by SDS-PAGE. To study
its oligomerization state, gp11 was loaded into a Superdex 200
10/300 GL (GE Healthcare) column for size exclusion chroma-
tography. Bovine albumin, ovalbumin, anhydric carbomace, and
cytochrome c were used as protein markers.
In Vitro Fiber Assembly—gp17 fiber protein was produced in E. coli strain BL21(DE3). The culture was grown at 37 °C up to an \( A_{600} \) between 0.6 and 0.8, induced with 1 mM isopropyl 1-thio-\( \beta \)-D-galactopyranoside, and grown for 16 h at 16 °C. Harvested cells were resuspended in 50 mM Tris-HCl, pH 8.9, 4% glycerol, 50 mM ammonium chloride, 2 mM EDTA, 150 mM sodium chloride and then frozen at −20 °C. The cells were lysed by several freezing and thawing cycles and centrifuged to eliminate cell debris. A one-to-one culture volume ratio of the gp17 copy, the samples were applied to Quantifoil 2-m holey carbon film (Gatan Inc.) and frozen in liquid ethane and transferred to a FEI Tecnai FEG200 operated at 200 kV, using low dose protocols with a 4K \( \times \) 4K Eagle CCD camera (Gatan Inc.). A nominal magnification of \( \times 108,696 \) and a defocus range of 1.5–3.5 Å. The contrast transfer function was corrected in the micrographs using standard XMIPP software (25); the particles were manually selected with XMIPP (26) and down-sampled to a factor of 3 to a final pixel size ratio of 4.14 Å/pixel. Images were classified using XMIPP reference-free clustering approach classification methods (CL2D) (27), and 640 particles corresponding to side views were selected. A Gaussian blob was used as a starting model to perform a three-dimensional reconstruction using a restricted projection angle protocol with EMAN (28). Next the model was refined using XMIPP Projection Matching package (25). For cryo-electron microscopy, the samples were applied to Quantifoil 2-\( \mu \)m holey carbon grids coated with a thin carbon layer, frozen in liquid ethane, and transferred to a FEI Tecnai FEG200 operated as described above. Tail complex images were acquired at a nominal magnification of \( \times 50,000 \), and the selected micrographs were scanned on a Zeiss scanner (Zeiss SCAI scanner) with a step size of 7 \( \mu \)m. The images of the 8-11-12 complexes were acquired under the same conditions but using a 4K \( \times \) 4K Eagle CCD camera at a nominal magnification of \( \times 108,696 \). Next, images were contrast transfer function-corrected (25) and down-sampled by a factor of 2 to a final pixel size ratio of 2.8 and 2.75 Å/pixel for the tail and the 8-11-12 complex, respectively. To perform the three-dimensional reconstructions, 3056 and 1820 particles of the tail and the 8-11-12 complex, respectively, were extracted, normalized, and two-dimensionally aligned and classified using reference-free XMIPP maximum likelihood (ML2D) (29) and CL2D (27) procedures. Averages resulting from these classifications were used to generate an initial three-dimensional model based on the common line methods implemented in the EMAN software (28) applying 6-fold symmetry. These initial models were refined using SPIDER (30) or XMIPP Projection Matching package (25). The final resolutions using the 0.3 Fourier shell correlation criterion were ~16 Å for the tail complex and ~12 Å for the 8-11-12 complex. Three-dimensional volumes and images were generated with the University of California at San Francisco Chimera software (31).

Structural Modeling (Threading), Segmentation, and Fitting—Appropriate templates for model building of portal protein gp8 were found with HHpred (32). These corresponded to the enterobacteriophage P22 connector protein (E-value 5.5E-05, p value 2E-09, 10% identity) and the bacteriophage SPP1 portal protein (E-value 0.18, p value 5.4E-10, 12% identity). The final structural model of gp8 was built with I-TASSER (33), using the fully automatic procedure. Alignment of gp8 with P22 connector protein was produced by I-TASSER. Structural templates for gp11 were assessed with several sequence-based and threading algorithms providing no significant output. In the absence of a clear template for gp11, a sequence profile was created and further used to search against the InterPro database (34) using HHpred (32). The first two hits corresponded to Pfam domains of unknown function, and the third to Pfam P22 tail accessory factor (PF11650). None of these hits were statistically significant. Nevertheless, functional and structural data indicated that T7 gp11 and P22 gp4 have similar roles. Therefore, a structural model for gp11 was built using the atomic structure of gp4 from enterobacteria phage P22 (PDB code 1vt0 chain k). The sequence alignment provided by HHpred was used to guide the I-TASSER modeling (33). Segmentation of gp8 and gp11 oligomers was carried out manually using the University of California at San Francisco Chimera software (31). A 12-fold symmetry was imposed to the maps using proc3d EMAN package (28). The monomer models were manually fitted as rigid bodies into the EM electron density using Chimera, and the oligomers were built using pdb2sym program from the SITUS package (35). Atomic structures were presented using PyMOL (Delano Scientific LLC, San Francisco) and EM maps using Chimera (31).

RESULTS

Structure and Protein Composition of the T7 Tail Complex—T7 virions incubated with 100 mM EDTA results in the disruption of the viral particles into separate heads and tails (data not shown). The tail complexes, purified by sucrose gradient centrifugation and size exclusion chromatography, consist of the connector protein (gp8) and the tail proteins gp11, gp12, and
Biochemical and Structural Characterization of T7 Tail

To obtain a direct mapping of the different structural proteins in the tail, we cloned and expressed each protein alone or in combination with the others, as block-cloning strategies have been successfully used to obtain other tail components (such as the base plates) in lactococcal phages (36). The joint expression of the genes 8, 11, and 12 resulted in the assembly of fiber-less tail complexes, referred to in this study as the 8-11-12 complex. The absence of the flexible fibers facilitated the cryo-electron microscopy data acquisition with respect to the complete tail complex, and it was possible to reconstruct the volume of this fiber-less complex at a resolution of ~12 Å (Fig. 1D).

The averaged projections of the side and end-on views of the complex revealed a very similar structure with respect to the tail complex (Fig. 1B, lower row), with identical dimensions and overall shape, and with the channel also closed at the nozzle end. The main difference is located at the level of the fibers protruding from the sides of the main body of the tail. In this case, the absence of the fibers allows us to distinguish a new toroidal region at the intermediate position between the connector, in the upper part, and the conical region, in the lower part. This intermediate structure also shows a clear 12-fold symmetry. Additionally, the higher resolution of this reconstruction permits us to discriminate two regions into the conical domain, one with six protrusions extending outward from the complex axis at an angle similar to the tail fibers and another, at the vertex, defining a nozzle, which serves as the channel closure (Fig. 1D).

The difference map between the 3DR of the complete tail and the 8-11-12 complex (Fig. 1E) revealed that both structures were very similar with the exception of the density that corresponds to the fiber proteins (Fig. 1E). As the difference between both complexes resides in the presence of the protein gp17, this protein should build the six well defined fibers. A study of isolated fibers by Steven et al. (20) demonstrated that the protein gp7.3 was not detected in any of our tail complex purifications. Cryo-EM of the vitrified tail complexes (Fig. 1B) showed images that, after two-dimensional averaging, revealed projections characteristic of side views of the tails as follows: a 29-nm-long tubular structure with a central channel that is closed at the narrower end (Fig. 1B, upper row, left panel). The averaged projections corresponding to the end-on view of the tail complex clearly shows the existence of six fibers, radially extending, which are better defined in the domain proximal to the tail axis (Fig. 1B, upper row, right panel). Three-dimensional reconstruction (3DR) from the vitrified tail complexes using single particle image reconstruction methods produced a volume showing a roughly conical structure with two well defined morphological domains along the longitudinal axis (Fig. 1C).

The upper domain is a toroidal region with a clear 12-fold symmetry, which showed an overall similarity to other viral connector structures (2, 18). The lower domain of the complex has a conical shape, with six conspicuous upward pointing fibers, and ends in a closed nozzle structure (Fig. 1, B and C). This overall structure is very similar to the one described for the tail in the complete phage (15, 16), strongly suggesting that the complex has not suffered detectable structural changes during detachment from the phage. The ~15-nm-long thin fibers are tilted by around 42°, and this topology was similar to the so-called upward conformation described in intact virions from tomographic reconstructions (15). In addition, the overall structure of the T7 tail described here notably resembles the cyanophage P-SPP7 tail (9), also a member of the Podoviridae family. This structural similarity suggests that both viruses may have a common infection mechanism.

FIGURE 1. Structural and biochemical characterization of T7 tail complexes. A, Coomassie-stained SDS-PAGE showing the proteins present in the phage-extracted tail complex, the recombinant 8-11-12 complex, and the complete T7 phage particle. The position of the T7 proteins is marked with an arrow. B, two-dimensional averages of electron microscopy images showing side (left panel) and end-on views (right panel) of the extracted tail complexes (top panel) and recombinant 8-11-12 complexes (bottom panel). C, three-dimensional volume of the phage-extracted tail complex in side (left panel) and end-on (right panel) views. D, three-dimensional volume of the 8-11-12 complexes showing side and end-on views, (left and right panels, respectively). E, side view of the difference map of the three-dimensional volumes described in C and D. The reference volume is shown in contour lines. The scale bar is 20 nm.

The abbreviations used are: 3DR, three-dimensional reconstruction; PDB, Protein Data Bank.

SEPTEMBER 6, 2013 • VOLUME 288 • NUMBER 36

JOURNAL OF BIOLOGICAL CHEMISTRY 26293
strated that the gp17 trimer forms a kinked thin fiber 2 nm in diameter and a length of 32 nm. As the length of the fibers we have reconstructed in the tail is around 15 nm, it is clear that this density corresponds to the proximal N-terminal half-fiber (16.4 nm long) described in the isolated gp17 trimer (20). The lack of the distal carboxyl part of gp17 in our assembled fibers is probably due to the flexibility of the fibers around its hinge, which prevents their averaging. This flexibility has been proposed as a critical factor in the structural transitions involved in the phage-bacteria interaction, and their conformation might be stabilized by its interaction with the capsid (15).

Comparison between the tail and the fiber-less models (Fig. 1, C--E) allowed us to clearly define the docking point of the fibers in specific positions of the middle and the lower tail domains and to unambiguously identify the composition of these two domains of the tail complexes formed by proteins gp11 and gp12. It has been previously discussed that the protein gp7.3 could be present in the central tail channel or in the tail tip (14), but we have not detected it in any of the complexes described here. Although we cannot exclude that the protein has been lost from the tail during the chemical treatment needed to extract the tail complex from the virion, the structural similarity between the structures characterized here and the one present in the entire phage (15, 16) suggests that gp7.3 protein is not a structural protein of the mature T7 tail.

**gp17 Fiber Assembles into the Tubular Tail Complex in Vitro**—To study the assembly of the different structural components of the tail complex, we tested whether the 8-11-12 complex was able to incorporate gp17 protein in vitro. The extracts containing the 8-11-12 complexes and the extract derived from the separately cloned gp17 were incubated together at 30 °C, purified, and then analyzed for the presence of complexes. Although purified free gp17 trimers and control complexes (not incubated in the presence of gp17) eluted in the size exclusion chromatography column at a volume of 14 and 12 ml, respectively, the complexes assembled in the presence of gp17 eluted at a volume of 8 ml (Fig. 2A). This difference in elution volume suggested an increase in the size of the gp17-incubated complexes. SDS-PAGE of the peak fractions showed the presence of gp17 in the larger complexes (Fig. 2A). Analysis by electron microscopy of these samples showed that although no fibers were observed in the control complexes (Fig. 2B, middle panel), the complexes incubated in the presence of gp17 contained the typical fibers, and they were identical to the ones extracted from the viral particles (Fig. 2C, right and left panels, respectively). These experiments strongly suggested that the His-tagged 8-11-12 complexes present a native structure, as they are able to assemble the fibers. Furthermore, the binding of the gp17 fibers to the pre-assembled tubular 8-11-12 complex suggests that this protein is the last protein to be incorporated into the virus during the morphogenetic pathway. This result matches previous experiments in which fiber-less viruses incorporated fibers using gp17 extracts (14).

**Localization of gp11 and gp12 Tail Proteins**—Cloning and purification of the isolated gp8, gp11, and gp12 proteins allowed us to carry out biochemical assays to identify the interactions that take place inside the tail complex. Pulldown assays using Dynabeads coated with anti-gp8 antibody allowed us to assign the interaction between gp8 and gp11 proteins (Fig. 3A). Although purified gp11 alone was not able to bind to the anti-gp8 antibody-coated beads, a significant amount of this protein interacted with gp8 as seen in the precipitation with the gp8
Biochemical and Structural Characterization of T7 Tail

FIGURE 3. Localization of the tail proteins inside the complex. A, Coomassie-stained SDS-PAGE (left panel) showing the purification of connector (gp8) and minor tail (gp11) proteins. Western blot was developed with an anti-His tag antibody showing the pulldown assays used to characterize the gp8-gp11 interaction (right panel). B, chromatogram showing the elution volume of the gp11-purified protein. The arrows show the elution volume of the void volume (V₀) and the protein markers as follows: 1, bovine albumin (67 kDa); 2, ovalbumin (43 kDa); 3, carbonic anhydrase (29 kDa), and 4, cytochrome c (14 kDa). AU, absorbance units. C, side section of the tail complex structure showing the localization of gp8 (red), gp11 (dark blue), gp12 (green), and gp17 (orange) proteins. D, Coomassie-stained SDS-PAGE (left panel) showing the purification of the major tail protein (gp12). Middle panel, micrograph of a negatively stained sample showing a gp12 representative field. Side views and end-on views are marked with a red square and a circle, respectively. Right panel, two different side views of the reconstructed gp12 monomer structure (green) docked inside of the gp12 hexamer volume segmented from the 8-11-12 complex structure (gray contour lines). The scale bar, 4 nm.

antibody (Fig. 3A). This interaction clearly identifies the location of gp11 inside the reconstructed tail complex, as forming the 12-fold symmetric toroidal domain immediately below the connector (Fig. 3C), an oligomeric state previously proposed for gp11 (14). To study the oligomeric state of gp11 in solution, we performed size exclusion chromatography of the purified protein (Fig. 3B). The protein eluted between 43- and 29-kDa protein markers, which clearly suggests that the protein is present as a monomer in solution when it is not interacting with gp8 connector protein. The location of gp11 in the tail, together with its oligomerization change upon assembly (between monomer and dodecamer), strongly suggests that gp11 could be a gatekeeper protein that acts as an adaptor protein. In other phages these proteins also show a monomeric form that acquires a dodecameric architecture after interacting with the connector (3, 13, 37, 38).

The third structural component of the tail tube, gp12, would then form the end of the tail, including the conical tube, the nozzle, and the small extensions below the fibers (Fig. 3C). The existence of 6-fold symmetry in this domain (previously suggested in Ref. 14) supports a fixed fitting of this domain with the upper gatekeeper. The size of the gp12 protein (89 kDa) allowed its direct observation in the electron microscope (Fig. 3D, middle panel). Image processing of negatively stained samples led us to obtain the 3DR of the gp12 protein. The volume presented an elongated shape (around 14 nm long and 4 nm wide), whose structural features are in agreement with the volume of each of the six monomers observed in the conical domain of the tail (Fig. 3D, right panel). The fitting of the gp12 negatively stained volume in one of the six lobes of the gp12 hexamer segmented from the 8-11-12 cryo-EM map (Fig. 3D, right panel) strongly suggests that gp12 is also present as a monomer in solution and that it is its interaction with the gp11 gatekeeper that drives its assembly as a hexameric nozzle containing the tail end.

Protein gp11 Shares Structural Similarities with Adaptor-Gatekeeper Proteins from Other Phages—gp4 from phage P22 and gp11 from phage T7 have similar locations, and therefore both proteins might share structural features. Even though no sequence homology was found between these proteins, protein structure prediction methods suggested that they indeed might have a common fold. The model of the gp11 T7 protein was thus obtained by threading methods using as a template the gp4 protein of P22 (13). The overlay of the gp11 model with the atomic structures of proteins gp4 from P22 and also gp6 from Siphoviridae phage HK97 pointed to the presence of four α-helices preserved in all three structures (Fig. 4A) (13, 37). The N-terminal end of the gp11 structure is nevertheless unique, and its atomic structure could not be predicted by the threading. Our results indicated that gp11 would be formed by two different modules as follows: one, which is more conserved with respect to other tailed phages (13, 37, 39, 40), builds the central channel of the gatekeeper structure, and second, another new module that might be involved in the interaction with the fibers for their subsequent binding. The differences between fiber and spike structures presented in T7 and P22 (41), respectively, may explain the modifications in the corresponding folds of the gatekeeper proteins.

The presence of two cysteine residues in the gp11 sequence (C121 and C159, Fig. 4B, left panel) allowed testing the validity of the threaded gp11 model. According to our proposal, these residues are placed close enough in the folded structure to form a disulfide bond under nonreducing conditions. Gel electrophoresis analysis of purified gp11 in the absence of reducing agent showed the presence of a band migrating faster (gp11-SS) than the reduced gp11 (Fig. 4B, right panel). This could indicate that an intra-molecular disulfide bond is indeed formed. A similar result was obtained when the protein was assembled inside the 8-11-12 complex, indicating that this disulfide bond was
formed when the protein is in its native structure (Fig. 4B, right panel). In the latter case, a high molecular mass band (gp12-SS, with a relative mass higher than 150 kDa) was also observed when the reducing agent was absent (Fig. 4B, right panel). The partial disappearance of the gp12 monomer band and the presence of five cysteine residues inside its sequence suggested that this band could correspond to inter-molecular gp12 disulfide cross-linking.

To generate a pseudo-atomic model of the gp11 toroidal complex, the volume corresponding to this protein was segmented out of the 8-11-12 complex structure. Then 12-fold symmetry was imposed to the map, and the threaded gp11 model was docked inside this volume (Fig. 4C). The precise docking of the model confirmed that gp11 was present in the phage as a dodecamer, as is the case for gp4 from P22 (13) and gp6 from HK97 (37). The docking shows that the oligomer is assembled in the same way as in P22, with the α4 helix (the most C-terminal of the helices) building the internal surface of the central channel (Fig. 4C, right panel, contoured in red). The C-terminal domain ends in a long loop that might be implicated in the interaction with the connector (Fig. 4B, left panel, and C, arrows). The nonmodeled N-terminal side of the protein is located at the outside part of the ring, and it might form an interaction surface with the fibers.

**Protein gp8 Presents an Internal Channel Formed by a 24x-Helix Stretch Characteristic of Connector Proteins**—The 3DR of the 8-11-12 complex also allowed us to obtain a detailed model of the structure of the connector (gp8) and its interaction with the gatekeeper (gp11). The model of gp8, containing residues 1–497, was obtained by threading using the P22 connector (gp1) as template (Fig. 5A) (13). The threaded model presented the overall structure previously described for other connector proteins (2), composed by the stem, crown, and wing domains (Fig. 5A, right panel). The superposition of the gp8 model with the atomic structures of p10 from φ29 (42, 43), gp1 from P22 (13), and gp6 from SPP1 (44) showed that the topology of the helices that form the internal surface of the channel is conserved among the four structures (Fig. 5A, left panel). The lower end of these structures, which can be either formed by a helix-β-sheet motif or an unstructured loop, is highly variable. This structural variability is probably related to the multitask activity of this region, which has to interact with the viral ATPase (terminase) in the early steps of the assembly pathway (the terminase is later detached from the complex and replaced by the gatekeeper protein). Another region that showed variability was the C-terminal region, where a group of three helices adopt a more open or close conformation depending on the connector protein. Segmentation of the gp8 volume from the 8-11-12 complex map allowed docking the model inside (Fig. 5B). A good fit is observed between the model and the cryo-EM map, confirming the quality of the structural model obtained. The only discrepancy is located at the bottom part (Fig. 5B, middle and right panels), suggesting that this region may have a different conformation in the assembled tail complex for the reasons discussed above. The segmented volume presents some differences with the structure previously described for the isolated connector, notably in the C-terminal domain (crown) (18). This region was found more prominent in the isolated connector dodecamer, suggesting that it might present some structural flexibility when gp8 is assembled in the tail structure.

The fitting of the pseudo-atomic models of gp8 and gp11 inside the 8-11-12 complex 3DR showed that both proteins

---

**FIGURE 4. Bacteriophage T7 gp11 protein.** A, ribbon representation of the gp11 T7 model overlapped with the atomic structures of the gp4 P22 (PDB code 3LJ4) and gp6 HK97 (PDB code 3JVO). N- and C-terminal ends and α4 helix are marked with arrows. B, ribbon representation of gp11 showing the position of the cysteine residues Cys-121 and Cys-159. The square represents the zoomed-in area (bottom panel). Coomassie-stained SDS-PAGE showing the behavior of the proteins of the 8-11-12 complex and gp11 in the presence (+) or absence (−) of the β-mercaptoethanol-reducing agent (right panel). The bands corresponding to the gp8, gp11, and gp12 monomers and intra-molecularly disulfide-bonded gp11 (gp11-SS) and inter-molecularly disulfide bonded gp12 (gp12-SS) are marked with arrowheads. C, docking of the gp11 pseudo-atomic structure inside of the volume segmented from the 8-11-12 complex structure showing an end-on (left panel), side (middle panel), and side section views (right panel). The location of the internal α4 helix is contoured with a red dashed line and an arrow. The location of the N-terminal and C-terminal ends are marked with arrows.
present an overall assembly and interaction very similar to the one previously described in P22 (13), with the characteristic helices forming the central channel of the tail assembly (Fig. 6A). The electrostatic isosurface of the complex showed that the internal channel presents an overall negative charge (Fig. 6B). This highly electronegative environment has also been described for other portal channels as an essential characteristic to avoid DNA interaction with the complex during the packaging and ejection processes (2).

DISCUSSION

Our results reveal the detailed location of gp8, gp11, gp12, and gp17 proteins in the tail of bacteriophage T7 (Fig. 3C). The overall architecture of the T7 tail is very similar to the one of...
Biochemical and Structural Characterization of T7 Tail

The T7 tail complex consists of a connector and gatekeeper proteins, which are essential for bacterial DNA packaging. The connector, gp11, and gatekeeper, gp12, are key components in the viral assembly process. In this study, we investigated the structural similarities and interactions between the connector and gatekeeper proteins, as well as the formation of the T7 tail complex.

Key findings include:

1. The connector and gatekeeper proteins are conserved among virus families, indicating their importance in viral assembly.
2. Structural changes in the gp12 protein upon DNA packaging suggest a role in the transition of the packaging motor.
3. The gp17 tail fiber binds to the gp11 connector, facilitating the alignment of the viral genome.

The results contribute to our understanding of the molecular mechanisms involved in bacterial DNA packaging and viral assembly.

Acknowledgments—We thank the Protein Tools Unit Facility at the CNB for gp8 antibody production. We are indebted to Alina Ionel for previous work on T7 tail complex preparation.

REFERENCES

1. Casjens, S. R. (2011) The DNA-packaging nanomotor of tailed bacteriophages. Nat. Rev. Microbiol. 9, 647–657
2. Cuervo, A., and Carrascosa, J. L. (2012) Viral connectors for DNA encapsulation. Curr. Opin. Biotechnol. 23, 529–536
3. Tavares, P., Zinn-Justin, S., and Orlova, E. V. (2012) Genome gating in tailed bacteriophage capsids. Adv. Exp. Med. Biol. 726, 585–600
4. Ackermann, H. W. (2007) 5500 phages examined in the electron microscope. Arch. Virol. 152, 227–243
5. Cuervo, A., and Carrascosa, J. L. (2012) Bacteriophages: Structure, pp. 1–7, John Wiley & Sons Ltd., Chichester, UK
6. Casjens, S. R., and Molinex, I. J. (2012) Short noncontractile tail machinery: adsorption and DNA delivery by podoviruses. Adv. Exp. Med. Biol. 726, 143–179
7. Xiang, Y., Morais, M. C., Battisti, A., I. Grimes, S., Jardine, P. J., Anderson, D. L., and Rossmann, M. G. (2006) Structural changes of bacteriophage 429 upon DNA packaging and release. EMBO J. 25, 5229–5239
8. Lander, G. C., Tang, L., Casjens, S. R., Gilcrease, E. B., Prevelige, P., Poliakov, A., Potter, C. S., Carragher, B., and Johnson, J. E. (2006) The structure of an infectious P22 virion shows the signal for headful DNA packaging. Nature 312, 1791–1795
9. Liu, X., Zhang, Q., Murata, K., Baker, M. L., Sullivan, M. B., Fu, C., Dougherty, M. T., Schmid, M. F., Osborne, M. S., Chisholm, S. W., and Chiu, W. (2010) Structural changes in a marine podovirus associated with release of its genome into Prochlorococcus. Nat. Struct. Mol. Biol. 17, 830–836
10. Choi, K. H., McPartland, J., Kaganam, L., Bowman, V. D., Rothman-Denes, L. B., and Rossmann, M. G. (2008) Insight into DNA and protein transport in double-stranded DNA viruses: the structure of bacteriophage N4. J. Mol. Biol. 378, 726–736
11. Jiang, W., Chang, J., Jakana, J., Weigels, P., King, J., and Chiu, W. (2006) Structure of e15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. Nature 439, 612–616
12. Leiman, P. G., Battisti, A. J., Bowman, V. D., Stummeyer, K., Mühlenhoff, M., Gerardy-Schahn, R., Scholl, D., and Molinex, J. I. (2007) The structures of bacteriophages K1E and K1–5 explain progressive degradation of polysaccharide capsules and evolution of new host specificities. J. Mol. Biol. 371, 836–849
13. Olia, A. S., Prevelige, P. E., Jr., Johnson, J. E., and Cingolani, G. (2011) Three-dimensional structure of a viral genome-delivery portal vertex. Nat. Struct. Mol. Biol. 18, 597–603
14. Kemp, P., Garcia, L. R., and Molinex, I. J. (2005) Changes in bacteriophage T7 virion structure at the initiation of infection. Virology 340, 307–317
15. Hu, B., Margolin, W., Molinex, I. J., and Liu, J. (2013) The bacteriophage T7 virion undergoes extensive structural remodeling during infection. Science 339, 576–579
16. Agirrezabala, X., Martín-Benito, J., Castón, J. R., Miranda, R., Valpuesta, J. M., and Carrascosa, J. L. (2005) Maturation of phage T7 involves structural modification of both shell and inner core components. EMBO J. 24, 3820–3829
17. Serwer, P., and Jiang, W. (2012) Dualities in the analysis of phage DNA packaging motors. Bacteriophag. 2, 239–255
18. Agirrezabala, X., Martín-Benito, J., Valle, M., González, J. M., Valencia, A., Valpuesta, J. M., and Carrascosa, J. L. (2005) Structure of the connector of bacteriophage T7 at 8 Å resolution: structural homologies of a basic component of a DNA translocating machinery. J. Mol. Biol. 347, 895–902
19. García-Doval, C., and van Raaij, M. J. (2012) Crystallization of the C-terminal domain of the bacteriophage T7 fibre protein gp17. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 68, 166–171
20. Steven, A. C., Trus, B. L., Maizel, J. V., Unser, M., Parry, D. A., Wall, J. S., Hainfeld, J. F., and Studier, F. W. (1998) Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. J. Mol. Biol. 200, 351–365
21. García-Doval, C., and van Raaij, M. J. (2012) Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. Proc. Natl. Acad. Sci. U.S.A. 109, 9390–9395
22. Fumagalli, L., Esteban-Ferrer, D., Cuervo, A., Carrascosa, J. L., and Gomila, G. (2008) Label-free identification of single dielectric nanoparticles and viruses with ultrafast polarization forces. Nat. Mater. 11, 808–816
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, pp. 74–76, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. Daudén, M. I., Martín-Benito, J., Sánchez-Ferrero, J. C., Pulido-Cid, M., Valpuesta, J. M., and Carrascosa, J. L. (2013) Large terminase conformational change induced by connector binding in bacteriophage T7. J. Biol. Chem. 288, 16998–17007
25. Scheres, S. H., Núñez-Ramírez, R., Sorzano, C. O., Carazo, J. M., and Marabini, R. (2008) Image processing for electron microscopy single-particle analysis using XMIIP. Nat. Protoc. 3, 977–990
Biochemical and Structural Characterization of T7 Tail

26. Sorzano, C. O., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H., Carazo, J. M., and Pascual-Montano, A. (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* 148, 194–204

27. Sorzano, C. O., Bilbao-Castro, J. R., Shkolnisky, Y., Alcorlo, M., Meleo, R., Caffarena-Fernández, G., Li, M., Xu, G., Marabini, R., and Carazo, J. M. (2010) A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Struct. Biol.* 171, 197–206

28. Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* 128, 82–97

29. Scheres, S. H., Valle, M., and Carazo, J. M. (2005) Fast maximum-likelihood refinement of electron microscopy images. *Bioinformatics* 21, Suppl. 2, 243–244

30. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadi, 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

31. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612

32. Söding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248

33. Zhang, Y. (2008) 1-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40

34. Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A., Bernard, T., Binns, D., Bork, P., Burge, S., de Castro, E., Coggill, P., Corbett, M., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Fraser, M., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., McMahon, C., Milos, M., Mulder, N., Natale, D., Orengo, C., Pesseat, S., Punta, M., Quelle, A. F., Rivoire, C., Sangrador-Vegas, A., Selengut, J. D., Sigrist, C. J., Scheremetjew, M., Tate, J., Thimmajanarthan, M., Thomas, P. D., Wu, C. H., Yeats, C., and Yong, S. Y. (2012) InterPro in 2012: new developments in the family and domain prediction database. *Nucleic Acids Res.* 40, D306–D312

35. Wriggers, W., Milligan, R. A., and McCammon, J. A. (1991) Siteus: A package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.* 125, 185–195

36. Campanacci, V., Veessler, D., Lichière, J., Blangy, S., Sciara, G., Moineau, S., van Sinderen, D., Bron, P., and Cambillau, C. (2010) Solution and electron microscopy characterization of lactococcal phage base plates expressed in *Escherichia coli*. *J. Struct. Biol.* 172, 75–84

37. Cardarelli, L., Lam, R., Tuite, A., Baker, L. A., Sadowski, P. D., Radford, D. R., Rubinstein, J. L., Battaile, K. P., Chigradze, N., Maxwell, K. L., and Davidson, A. R. (2010) The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins. *J. Mol. Biol.* 395, 754–768

38. Olia, A. S., Al-Bassam, J., Winn-Stapley, D. A., Joss, L., Casjens, S. R., and Cingolani, G. (2006) Binding-induced stabilization and assembly of the phage P22 tail accessory factor gp4. *J. Mol. Biol.* 363, 558–576

39. Lhuillier, S., Gallopin, M., Gilquin, B., Brasiles, S., Lancelot, N., Letellier, G., Gilles, M., Dethan, G., Orlova, E. V., Couprie, J., Tavares, P., and Zinn-Justin, S. (2009) Structure of bacteriophage SP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8507–8512

40. Veessler, D., and Cambillau, C. (2011) A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* 75, 423–433

41. Tang, J., Lander, G. C., Olia, A. S., Olia, A., Li, R., Casjens, S., Proveilige, P. J., Cingolani, G., Baker, T. S., and Johnson, J. E. (2011) Peering down the barrel of a bacteriophage portal: the genome packaging and release valve in P22. *Structure* 19, 496–502

42. Guasch, A., Pous, J., Ibarra, B., Gomis-Ruth, F. X., Valpuesta, J. M., Sousa, N., Carrascosa, J. L., and Coll, M. (2002) Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage φ29 connector particle. *J. Mol. Biol.* 315, 663–676

43. 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., Baker, T. S., and Rossmann, M. G. (2000) Structure of the bacteriophage φ29 DNA packaging motor. *Nature* 408, 745–750

44. Lebedev, A. A., Krause, M. H., Isidro, A. L., Vagin, A. A., Orlova, E. V., Turner, J., Dodson, E. J., Tavares, P., and Antson, A. A. (2007) Structural framework for DNA translocation via the viral portal protein. *EMBO J.* 26, 1984–1994

45. Lorenzen, K., Olia, A. S., Uetrecht, C., Cingolani, G., and Heck, A. J. (2008) Determination of stoichiometry and conformational changes in the first step of the P22 tail assembly. *J. Mol. Biol.* 379, 385–396

46. Olia, A. S., Bhardwaj, A., Joss, L., Casjens, S., and Cingolani, G. (2007) Role of gene 10 protein in the hierarchical assembly of the bacteriophage P22 portal vertex structure. *Biochemistry* 46, 8776–8784