Maturation of phage T7 involves structural modification of both shell and inner core components

Xabier Agirrezabala, Jaime Martín-Benito, José R Castón, Roberto Miranda, José María Valpuesta and José L Carrascosa*

Centro Nacional de Biotecnología, Campus de la Universidad Autónoma de Madrid, Cantoblanco, Madrid, Spain

The double-stranded DNA bacteriophages are good model systems to understand basic biological processes such as the macromolecular interactions that take place during the virus assembly and maturation, or the behavior of molecular motors that function during the DNA packaging process. Using cryoelectron microscopy and single-particle methodology, we have determined the structures of two phage T7 assemblies produced during its morphogenetic process, the DNA-free prohead and the mature virion. The first structure reveals a complex assembly in the interior of the capsid, which involves the scaffolding, and the core complex, which plays an important role in DNA packaging and is located in one of the phage vertices. The reconstruction of the mature virion reveals important changes in the shell, now much larger and thinner, the disappearance of the scaffolding structure, and important rearrangements of the core complex, which now protrudes the shell and interacts with the tail. Some of these changes must originate by the pressure exerted by the DNA in the interior of the head.

The EMBO Journal (2005) **24**, 3820–3829. doi:10.1038/sj.emboj.7600840; Published online 6 October 2005

Subject Categories: microbiology & pathogens; structural biology

Keywords: bacteriophage; capsid maturation; cryoelectron microscopy; morphogenetic intermediate; single-particle reconstruction

Introduction

Assembly of simple viruses is based on the direct interaction between their genomes and multiple copies of a reduced number of structural proteins. In more complex viruses, the assembly pathway involves intermediate steps that have highlighted the basic processes of macromolecular interaction. In particular, the study of the morphogenesis of complex double-stranded (ds) DNA viruses, such as bacteriophages and herpesvirus, and also dsRNA retrovirus (Steven *et al*, 2005), has revealed a number of basic steps, including the

formation of a proteinaceous prohead, the packaging of the viral genome into the prohead, and the final maturation by interaction with other additional components (Dokland, 1999). Detailed studies on different model systems have revealed that these basic principles also incorporate differences in the way each virus solves its specific life cycle. However, despite the lack of any evident sequence homology, there is a strong structural convergence with regard to the basic aspects of the viral life cycle: the use of an internal scaffold to direct the shape and size of the virus shell, the common structure of the packaging machinery, and the use of common folds for the major shell protein are examples of this structural and functional similarity (Wikoff *et al*, 2000; Fokine *et al*, 2005).

The successful combination of cryoelectron microscopy and X ray diffraction has been instrumental for our present understanding of the viral structure. The use of icosahedral symmetry in the reconstruction methods based on cryoelectron microscopy has revealed dynamic structural transitions involved during the maturation in different viruses as HK97 (Conway *et al*, 2001), P22 (Jiang *et al*, 2003) and lambda (Dokland and Murialdo, 1993). Nevertheless, the use of icosahedral symmetry constraints has prevented the study of those components of the viral structure that do not follow this symmetry. A more recent approach, in which the viral particles were reconstructed without imposing icosahedral symmetry, has revealed the topology of the different components of structures, as well as fundamental features of viruses such as T4 (Fokine *et al*, 2004) and ϕ 29 (Tao *et al*, 1998; Ibarra *et al*, 2000).

The incorporation of the nucleic acid into the virus follows different strategies depending on the viral system, from the simple assembly of the viral shell around the nucleic acid up to its complex encapsidation into preformed proheads. One of the more interesting aspects of complex virus morphogenesis is the mechanism of DNA packaging inside the preformed viral prohead. The packaging machinery is located at a unique vertex of the prohead, and it comprises the connector (that builds a channel with precise characteristics to fit the DNA) and the terminases, that are involved not only in the selection and processing of the DNA to be packaged but also in the ATP-driven DNA translocation (reviewed in Valpuesta and Carrascosa, 1994). Different systems also incorporate other components that probably deal with peculiar aspects of their life cycles, such as the pRNA in the case of ϕ 29 (Guo *et al*, 1987), or the core protein in the case of T7 (Serwer, 1976; Cerritelli *et al*, 2003a). Recent results obtained using optical tweezers (Smith *et al*, 2001), together with the resolution of the structure of several components of the packaging machinery, have provided different models to explain the packaging mechanism (Simpson *et al*, 2000; Guasch *et al*, 2002).

The genus T7-like bacteriophages comprise the coliphage T7 and the closely related phage T3. These viruses are

*Corresponding author. Department of Structure of Macromolecules, Centro Nacional de Biotecnología (CSIC), Campus de la Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain.
Tel.: +34 91 5854509; Fax: +34 91 5854506;
E-mail: jlcarras@cnb.uam.es

Received: 8 June 2005; accepted: 16 September 2005; published online: 6 October 2005

isometric, with a noncontractile tail, and contain a dsDNA around 40 kb in size. The prohead is built from the structural proteins gp10A and gp10B (derived from a read-through), the scaffolding gp9, the connector gp8, and a conspicuous core (made by the proteins gp14, 15, and 16). This core is unique among the different viral systems described so far, and it is not required for the prohead morphogenesis, but it seems to be essential for infectivity (Garcia and Molineux, 1996). The initial prohead interacts with the major subunit of the terminase (gp19), while the DNA attaches the smaller terminase subunit (gp18). Prohead and DNA then interact and the packaging process starts until the unit length DNA is encapsidated. Concomitant with this process the prohead expands, the scaffolding is released from the capsid, and the encapsidated DNA is cut away from the concatemer formed during the replication (reviewed in Cerritelli *et al*, 2003a). The structures of the prohead and the final T7 head were solved at near 17 Å resolution using icosahedral symmetry (Cerritelli *et al*, 2003a). In an attempt to define the structure and relative topology of those T7 components that are not represented when applying symmetry constraints, we have obtained the three-dimensional (3D) reconstruction of the proheads and the final mature viral particles using cryoelectron microscopy and single-particle reconstruction methods without symmetry constraints. This approach has revealed the structure of the packaging portal complex, including the connector and the core, and their topological relations with the shell, the scaffolding, the DNA, and the tail.

Results

Structure of the prohead

Type I proheads were purified from cells infected with a mutant defective in gene 5 (the viral DNA polymerase) and observed by cryoelectron microscopy. The projection images were combined using single-particle reconstruction protocols without imposing any symmetry. The density map at 24 Å resolution reveals a $T=7$ shell (Figure 1). The size (510 Å in diameter) and the surface morphology of the shell, including the capsomeric corrugations and the morphology of the skewed hexamers, as well as the intercapsomeric distance

(110 Å), matched the features described previously for icosahedrally reconstructed T7 empty procapsids (Cerritelli *et al*, 2003a), as well as the icosahedral reconstruction of the prohead used as internal control (Figure 1a of Supplementary data). Application of five-fold symmetry along the longitudinal axis of the particle improved the quality of the reconstruction (up to 18 Å resolution) without a significant modification of the structure (Figure 1A and B). In fact, as an internal control we compared individual pentamers and hexamers from icosahedrally reconstructed proheads with their counterparts from five-fold symmetrized and nonsymmetrized reconstructions. The correlation coefficients obtained were always greater than 0.95, supporting the accuracy of the nonicosahedrally reconstructed maps. Central sections of the reconstructed density maps (Figure 1C) reveal that the capsid has an average thickness of 40–50 Å, from the inner protuberances to the highly corrugated outer side. The inner side of the shell shows a number of fine extensions projecting towards the center of the particle (see arrows in Figure 1C) and probably contacting the putative scaffolding structure, a discrete layer of protein (marked by a bracket in Figure 1C) that is placed between the shell and a conspicuous structure at the center of the particle. Projection images of T7 procapsids have previously shown the existence of such assembly, termed the core, which appears at one of the vertices of the icosahedral shell (Serwer, 1976). The core reconstructed without imposing any symmetry is 265 Å long, and is organized into three main domains enclosing a continuous channel that connects the outer side of the shell with the inner region of the prohead (Figure 2A): The region in contact with the shell vertex (encompassed by the green box in Figure 2A) has a conical profile (outer diameter around 215 Å) and extends up to 115 Å, with a narrow domain that interacts with the shell. The channel has a regular profile with an average diameter around 35 Å. The rotational analysis of the planes perpendicular to the longitudinal axis corresponding to this area shows two well-defined areas: The outer one, corresponding to the shell, shows a clear five-fold symmetry. The inner region shows a consistent 12-fold symmetry in the area corresponding to the wider side of the core (outlined in green in Figure 2A and B). The

Figure 1 3D reconstruction of the prohead: (A, B) Volume representation of the 3D reconstructed prohead viewed along the longitudinal and orthogonal axes, respectively. The density is displayed to emphasize the characteristic surface corrugation of the capsomers. The left half in each image corresponds to the nonsymmetrized reconstruction (p1), while the right half represents the reconstruction reinforced by five-fold symmetrization along the longitudinal axis of the virus. (C) Projection density maps of central sections of the corresponding reconstructions (p1 and p5), showing the internal core structure extending from the singular vertex of the prohead. Arrows point to presumptive connections between the shell and the scaffold. The bracket outlines the region corresponding to the scaffold layer.

Figure 2 3D reconstruction of the core in proheads. **(A)** Cross section of the nonsymmetrized prohead viewed along the two-fold axis of symmetry (sectioned surface is shown in gray). The density is displayed at 1σ above the mean to show the scaffolding lattice arrangement. For clarity, the residual pentameric density has been computationally removed. **(B)** Rotational analysis of the harmonic components of the plane groups shown by color-coded rectangles in (A). A representative plane of each group is shown at the right. The rotational analysis shown in the graphics corresponds to the inner region encircled by the corresponding color codes. The outer radii of the regions selected for the calculation of the rotational harmonics are 100 Å (green), 70 Å (blue), and 65 Å (red), respectively. The vertical axis in the graphics represents the percentage of rotational energy in each harmonic, while the horizontal axis represents the range of analyzed harmonics. **(C)** Surface representation of the four-fold symmetrized core complex reconstructed after penton-less shell subtraction, and final removal of the scaffolding moiety (see Materials and methods). A side view is shown on the left, and a view down the five-fold axis toward the particle base on the right. **(D)** Docking of the previously reported 8 Å resolution connector structure (Agirrezabala *et al*, 2005) into the reconstructed core complex. A cross section is shown on the left, and the bottom view on the right.

second core region (encompassed by the blue lines in Figure 2A and B) extends 70 Å towards the interior of the prohead, forming an inverted cone (external diameter from 175 to 140 Å) that embraces a wide cavity of nearly 110 Å inner diameter. This area shows a prominent eight-fold symmetry (Figure 2B). The more distal core region (encompassed by the red lines in Figure 2A and B) is more cylindrical with an average external diameter of 130 Å and a height around 80 Å, and it shows the presence of four- and eight-fold symmetries. The channel that runs along the longitudinal axis of this region has a variable diameter, from 35 Å in the narrowest region to 45 Å in the widest one.

To obtain a 3D reconstruction of the core and the scaffolding that is independent of the shell, the projection components resulting from the prohead shell were subtracted from the original images, and the resulting difference projections of the core-scaffolding assembly were 3D reconstructed without imposing any symmetry. The existence of four-, eight-, and 12-fold symmetries observed in the reconstruction

of the core-containing prohead also allowed imposition of conservative four-fold symmetry. The reconstructed volumes were basically identical (the symmetrized one is shown in Figure 2C), and also revealed the three main domains observed in the 3D reconstruction of the prohead. Consequently, the rotational analysis of planes along the nonsymmetrized core-scaffolding complex yields results similar to those obtained in the full prohead reconstruction. As the core is connected with the vertex of the capsid where the packaging portal is presumably located, we performed a docking of the recently described structure of the T7 connector (Agirrezabala *et al*, 2005) into this volume. The connector fitted quite well into the lower domain of the core (the one exhibiting 12-fold symmetry), extending into the inner region well inside the second, central domain (Figures 2D and 3A). The shape of the core channel maps almost perfectly with the one described for the connector channel, as well as the stalk and lower wings of the connector fits with the outer profile of the core lower side.

One important aspect of the proposed core assembly is that the density of the pentamer below the connector is much lower than that of the rest of the other equivalent icosahedral pentamers (Figures 2A and 3A). Furthermore, there are no hints of density in that position in core–scaffold reconstructions obtained after penton-less shell subtraction, either in the nonsymmetrized or in the four-fold symmetrized reconstructions. Also, the subtraction of a fully icosahedral closed shell renders a core reconstruction lacking any density at the level of the stalk end. While the comparison of individual pentamers from icosahedrally reconstructed proheads with those derived from control icosahedral shell reconstructions yields correlation coefficients above 0.95, the comparison of the density at this singular vertex with any other prohead pentamer yields values lower than 0.56, supporting the intrinsic difference of this unique vertex. Thus, the residual density in this vertex of the nonsymmetrized prohead might be due to the reconstruction procedure, which is mainly directed by the matching of the shell projections, resulting in an artifactual inclusion of density in the hollow vertex.

The protein layer located between the core and the inner side of the shell can be assigned to the scaffolding protein (gp9). There are clear connections between the scaffold and the core, at the level of the wings of the connector (see arrows in Figure 3B). Also, there are connections between the scaffold and the inner side of the shell (Figure 1C, arrows). The reconstruction without imposing any symmetry of the core–scaffolding complex from images where the shell components were subtracted rendered a network pattern (Figure 3C) that resembles closely that described for the internal scaffold of phage $\phi 29$ (Morais *et al*, 2003). Furthermore, the existence of nubbins of density around 15 Å in size in the inner face of the shell (inset in Figure 3B) has been previously described as remnants of the interaction between the shell and the scaffolding (Cerritelli *et al*, 2003a), but we have not found any systematic interaction among these nubbins and the scaffold protein layer.

Structure of the mature virus

Mature viruses were obtained by purification from *Escherichia coli* infected with T7 wild-type strain. Projection images from cryoelectron micrographs were combined using single-particle reconstruction methods without any symmetry constraints to render a 3D density map at 24 Å resolution (see Materials and methods) (Figure 4). The virus showed a capsid with a $T = 7$ structure and a diameter of 600 Å, which explains the increase of the head volume by around 50% with respect to proheads (510 Å in diameter). The shell of the capsid is much thinner than that of the prohead (20–25 Å respect to 40–50 Å). The intercapsomeric distance is 140 Å, and the hexamers of the capsid exhibit a nondistorted six-fold symmetry, while those from the prohead are skewed (Figure 1), a trend that has been also previously reported for icosahedrally reconstructed T7 (Cerritelli *et al*, 2003a), P22 (Jiang *et al*, 2003), and lambda phages (Dokland and Murialdo, 1993). The use of five-fold symmetry along the tail–capsid axis produced a slight increment in the resolution of the reconstruction (up to 19 Å, see Materials and methods) and a better definition of the capsomers, but otherwise no major changes were observed. The comparison between any two hexamers or pentamers of the control icosahedral reconstruction (Figure 1b of Supplementary data) and nonicosahedral reconstructions always yields correlation coefficients greater than 0.89 and 0.88, respectively, indicative of the map's quality.

The virus reconstruction shows a tail, located at the same vertex where the connector core is assembled (Figure 4A and B). The tail, built by proteins gp11 and gp12 (Matsuo-Kato *et al*, 1981), extends 185 Å, and presents a cylindrical region (around 80 Å in length) that ends in an almost spherical tip (diameter around 75 Å). The reconstructed density corresponding to the tail was further averaged using the already described presence of six-fold symmetry in the tail region (Matsuo-Kato *et al*, 1981). Most of the tail shows an inner channel that is closed at least at two positions (Figure 4E, arrows), leaving the final third part of the channel devoid of any material. Around the middle of the outer face of the

Figure 3 Organization of the prohead. (A) Sectioned half volume of fitted 3D reconstructions of the five-fold symmetrized shell (displayed at $\sigma = 3.2$ to enhance the corrugations on the inner face of the shell, yellow), the four-fold symmetrized core (magenta), and the previously described connector (Agirrezabala *et al*, 2005) (green). The nonsymmetrized core–scaffolding complex (transparent gray representation) is displayed at high contour level (1σ) to facilitate visualization of molecular boundaries. (B) Close-up perspective view of the nonsectioned superimposed reconstructions shown in (A). This view is slightly tilted from the view in (A) to highlight the connections between the scaffold lattice and the core, at the level of the wings of the connector. Arrows point to connection points. Inset: view of the nubbins of density in the inner face of the shell. (C) Surface representation of the core–scaffold reconstruction after shell subtraction. No symmetry was imposed. The colored stripes and the long arrows highlight the suggested network arrangement of the scaffolding subunits.

Figure 4 3D reconstruction of the mature virus. **(A)** Surface-shaded representation of the five-fold symmetrized complete virion viewed along a two-fold axis of symmetry. **(B)** Central section of the corresponding reconstruction. In this perpendicular view to the connector–core axis, the DNA projects punctuate patterns spaced 2–2.5 nm. The central, less ordered density probably corresponds to the last packaged segment of DNA, suggesting that it is collapsed on itself. **(C)** Central section viewed along the five-fold axis of symmetry to show the concentric pattern of the packaged DNA. The right-hand half shows the rotationally averaged section. Inset: The radial density plot of this section exhibits an outer dense peak corresponding to the viral capsid, and then at least six equally spaced rings. **(D)** Enlarged perspective view of the proximal part of the tail shown in **(A)**, and **(E)** central section after the local six-fold symmetrization of this domain. This allows to highlight the density corresponding to the partially reconstructed fiber attachment proteins (the densities protruding outside from the equatorial region of the tail, see the stars). The DNA is placed along the core–tail axis. Arrows mark the position where the channel is fully closed.

tail, a density was observed that can be assigned to the region of the tail fibers (made by protein gp17 (Steven *et al*, 1988)). The proximal side of the fibers is seen in the sections of the locally six-fold reconstructed tails (Figure 4E, stars), but probably these fibers are only partially reconstructed due to their flexibility. Near the top of the tail, a second prominent ring is located just below the attachment site to the shell (at 25 Å). This ring (160 Å diameter), together with the connector wings, seems to hold the connector–tail attachment in the pentameric shell vertex (Figure 4D and E).

The interior of the capsid shows a complex morphology due to the existence of the core complex and the density derived from the presence of the viral DNA. As previously described (Cerritelli *et al*, 1997), the DNA is organized in six concentric layers and a less structured central region where three additional layers could be hinted (Figure 4C, inset), consistent with a spool around the axis of the connector–core assembly (Figure 4B and C). Sections of the layers show a particulate structure with 20 Å diameter consistent with sectioned DNA molecules (Figure 4B), with a center-to-center distance of 25 Å. Inside the DNA spool, the connector–core complex is recognizable (Figure 4B). The complex has an overall shape similar to the one found in proheads, but it shows significant differences: The length of the complex is slightly shorter than that found in the prohead (245 versus

265 Å). The connections between the three main domains of the core complex are altered, and the internal cavity in the center of the structure is slightly compressed (from 50 to 30 Å in height) and partially occupied by material, indicating a rearrangement of the assembly after packaging of the DNA (compare Figure 4B with Figure 1C). The position of the core complex with respect to the shell layer is also different from that found in the prohead, which in the mature virion traverses the shell. One important feature of the core–connector complex in the mature virion is that it is directly attached to the tail assembly at a level where the pentameric shell empty vertex embraces it (Figure 4D and E). The fact that most of the channel that runs along the tail–connector–core assembly is occupied in the reconstructed virus by a well-defined mass (20 Å in diameter) suggests that the last segment of the DNA molecule is located along the translocating pathway, and ready to be ejected in the infection process. The presence of DNA near the tail assembly has been documented for other viral systems (see, for example, lambda (Padmanabhan *et al*, 1972) and T4 (Leiman *et al*, 2004)).

Discussion

The use of a reconstruction strategy which does not rely on the use of icosahedral symmetry has allowed us to generate

3D reconstructions for the proheads and mature phage T7 particles that, besides the icosahedral shell, reveals the structure of the tail and other important internal components. The prohead shows a thick shell organized following a $T=7$ lattice, with skewed hexameric capsomers (Figure 1). The presence of common folds in the major capsid proteins of different unrelated bacteriophages has been suggested based on the evidences obtained from different systems: HK97 (Wikoff *et al*, 2000; Conway *et al*, 2001), P22 (Jiang *et al*, 2003), T4 (Fokine *et al*, 2005), and $\phi 29$ (Morais *et al*, 2005). These viruses utilize similar morphogenetic pathways as T7, and they also exhibit an intercapsomeric distance of 140 Å, a feature that has been suggested as characteristic for the HK97-like motif of the major structural head proteins (Fokine *et al*, 2005).

A major characteristic of the prohead is the presence of an eccentric structure, called the core, connected to an unique vertex of the shell (Figures 1 and 2). Between this internal core and the shell, there is a layer of protein (around 65 Å in the thicker regions) that shows extensions towards the shell and the core, and represents the internal scaffold required for a functional prohead assembly (Cerritelli and Studier, 1996; Greene and King, 1996; Thuman-Commike *et al*, 1998). It has been demonstrated that internal scaffolding proteins participate in the recruitment of the connector for its correct assembly into the prohead (van Driel and Couture, 1978; Guo *et al*, 1991; Greene and King, 1996; Droge *et al*, 2000). Furthermore, scaffolding proteins share a number of common characteristics, as in the case of $\phi 29$ (Tao *et al*, 1998), P22 (Thuman-Commike *et al*, 1998), and T7 (Cerritelli *et al*, 2003a). These features include their assembly in nonicosahedral lattices, and radial profiles comprising regions with different densities. In most cases, interactions between the scaffold and the hexameric capsomers have been described, involving small, low-density connecting regions, following more or less ordered arrangements (reviewed by Dokland, 1999). In the case of $\phi 29$, a structure comprising several concentric layers has been described (Morais *et al*, 2003). Although the resolution of our reconstruction does not allow to unambiguously define a precise protein arrangement in this scaffolding protein layer, it resembles the network appearance described for $\phi 29$ (Morais *et al*, 2003), where dimers support the interior of the shell, with their axis parallel to the surface. In the T7 scaffold it is possible to find contacts with the core complex (at the level of the connector wings), as well as with the shell (Figures 1 and 3), supporting the idea that the scaffolding protein could act as a link between the capsid and the connector, as part of the control of the size and shape of the growing shell.

The internal core shows a complex morphology that is 265 Å high and an average of 175 Å wide. There is a conspicuous tunnel that runs open along the whole assembly, with 35 Å diameter in the narrower area and 110 Å in the central wider cavity. The structure comprises three domains, each one exhibiting a different symmetry: 12-, eight-, and four-fold (Figure 2). The presence of eight-fold symmetry was reported previously in projection core images (Cerritelli *et al*, 2003b). The docking of the structure of the T7 connector into the core supports the interpretation that the connector builds the lower domain that presents a consistent 12-fold symmetry, which is in contact with the inner side of the shell (Figures 2A and 3A). On top of the connector and embracing it at the level

of the crown domain, there is an intermediate domain with two subareas: the lower one in contact with the connector, which has 12-fold symmetry, and the upper one with a lobulated aspect in top views (Figure 2C and D). Finally, there is an upper cylinder with a major four-fold symmetry. The eight- and four-fold domains would then be built by proteins gp14, gp15, and gp16 (reviewed by Cerritelli *et al*, 2003a). The estimated number of copies for each of these core proteins in the prohead (12, 8, and 3–5, respectively) can be accommodated in our model of the core: 12 copies of gp14 (20.8 kDa) in the core–connector interface, eight copies of gp15 (84.2 kDa) in the intermediate domain, and four copies of gp16 (143.8 kDa) in the cylinder.

The existence of an open channel along the connector–core assembly is consistent with the general model proposed for the translocation of DNA inside the prohead during the packaging process. This channel would drive the DNA into the capsid, allowing the formation of an ordered spool around the core complex (Cerritelli *et al*, 1997). It is noteworthy that the connector is not embedded in the prohead shell as in $\phi 29$ (Tao *et al*, 1998; Ibarra *et al*, 2000), but it is rather contacting its inner side (Figure 1C). This interaction is held in place by the inner extensions of the shell subunits around the pentameric vertex, which forms a ring embracing the stalk of the connector (Figure 5A). This would be the actual surface where the symmetry mismatch between the five-folded vertex of the prohead and the 12-folded stalk of the connector takes place, and it would be functionally equivalent to the interaction between the sliding groove of the $\phi 29$ connector and the shell of the prohead (Guasch *et al*, 2002). The connector–shell contact area is leaving an open cavity of 95 Å diameter and 40 Å height in the pentameric vertex below the stalk of the connector (Figure 5B). As this is the location for the interaction of the terminases with the connector, it is tempting to suggest that the major terminase protein (gp19), probably assembled as an hexameric ring (Morita *et al*, 1995), could fit well into this cavity. This type of arrangement could be similar to that found for the hexameric ATPase P4 from the dsRNA phage $\phi 12$ (Mancini *et al*, 2004). This packaging motor is also located at the vertices of the icosahedral capsid, building part of the shell (de Haas *et al*, 1999).

The maturation of the prohead involves a number of dramatic changes, including the packaging of the DNA, the release of the scaffolding, drastic changes in the structure of the shell and the core complex, and the attachment of the tail. While in T4 maturation takes place upon the independent formation of both the prohead and the tail, in the case of T7 it occurs upon assembly of tail proteins gp11 and gp12 into the preformed head (Matsuo-Kato *et al*, 1981). The structural changes at the shell level have been well defined: The shell increases in diameter by 15% (around a 50% increment in volume; Figure 5B) and becomes much thinner (from 40–50 to 20–25 Å). The hexamers of the capsid become fully six-folded in a similar way as happens in other viruses, suggesting a common mechanism for maturation and stabilization that probably involves rotation of domains of the major capsid protein, as well as some degree of local refolding (Conway *et al*, 2001).

The mature virus particle shows a tight compaction of the DNA in the form of a series of concentric layers (Figure 4B and C). The packaging model of a DNA molecule spooled

Figure 5 Domain model for the core rearrangement during viral maturation. (A) Gray level central sections of the core region reconstructed from proheads (left) and final virus (right). (B) Schematic model of the core structural transition showing the central sections of the reconstructed prohead (left) and the mature virus (right). The different domains are coded using different colors: connector, green; wider toroid, blue; eight-folded toroid, yellow; and upper cylinder, red. The arrows represent the possible direction of the movements of the complex. The assignment of the structural proteins to each of the domains is based in the predominant rotational symmetry of the domains determined in this work and the previously proposed stoichiometry (Cerritelli *et al*, 2003b). (C) 3D structure of the prohead (left) and the mature virus (right) shown along the longitudinal core-vertex five-fold symmetry axis. The connector structure (highlighted in green) is visible through the open pentameric vertex of the prohead and mature virus capsid reconstructions.

around the five-fold axis of the core, proposed using cryo-electron microscopy images and computer modelling (Cerritelli *et al*, 1997), is fully consistent with our 3D reconstruction, comprising six clearly visible concentric layers (and three more additional less well-defined layers) with center-to-center spacings of 25 Å. Towards the core region, the DNA becomes less ordered, although the whole available space is filled with DNA. It has been proposed that the curvature stress is related to the strong bending forces that induce the DNA molecule to crowd into itself, resulting in a smaller DNA-DNA spacing. Therefore, the observed structur-

al pattern of the packaged DNA could be imposed by the equilibrium between the bending forces and the pressure exerted by the DNA at the shell walls (Tzllil *et al*, 2003). Also, the empty channel present in the prohead core-connector complex is filled in the mature virion with density that could be compatible with the last segment of the packaged DNA. This density is filling the channel up to the tail attachment site, suggesting that the end of the DNA to be ejected from the capsid is actually in contact with the injection apparatus (Padmanabhan *et al*, 1972; Leiman *et al*, 2004). The initial injection is a passive process driven by the energy stored upon DNA packaging (Garcia and Molineux, 1996; Tzllil *et al*, 2003).

Structural changes are particularly relevant in the connector-core complex. While the different domains comprising the connector-core complex can be observed in both proheads and mature heads, their relative topology is different (Figure 5). The connector (colored green), that is held above the pentameric vertex hole in the prohead by the extensions of the shell protein, moves down and fully traverses the thin layer of the shell in the mature capsid, which allows its interaction with the tail (Figure 4D). The intermediate region just above the connector can be divided into two domains: one, probably in direct contact with the connector, exhibits 12-fold symmetry (colored blue in Figure 5). The other (colored yellow), is the eight-folded domain, and comprises a wide cavity around 110 Å in diameter. These two domains, arranged as two consecutive toroids of different diameters, undergo an important reorganization during head maturation: the subunits of the upper one tilt outwards, while those of the lower and wider domain collapse towards the channel axis of the complex. The result of such movements is a partial closure of the middle cavity (final height around 30 Å), which brings the 12-folded domain of the core in contact with the crown of the connector. This large conformational transition might be triggered by the internal pressure of the DNA once it reaches the full occupancy of the capsid. The new interaction between the core intermediate domain and the connector (that brings into contact a completely new set of interactions between the connector crown and the intermediate domain protein) could play some role in the finalization of the packaging reaction by inactivating the connector rotation coupled to packaging and/or the interaction of the connector with the terminase. It has been previously proposed for phage lambda that the increase of the packaged DNA length may alter the connector structure (Cue and Feiss, 1997), reducing the rate of DNA translocation and thereby increasing the efficiency by the terminase of the recognition of the packaging termination sequence (pac or cos sites). This would be a more sophisticated variant of the headful mechanism proposed for other systems such as SPP1 (Tavares *et al*, 1992). At this resolution level, we cannot discard that these rigid body transitions might be coupled to minor changes at the level of the individual structure of the connector and the other components of the core complex.

The movement of the core complex towards the outside of the capsid caused after DNA packaging, together with the widening of the pentameric vertex where the connector attaches due to the transformation of the shell (Figure 5C), has important consequences: The narrow end of the connector (the stalk domain) probably pushes the terminase complex out of the pentameric cavity, thus becoming accessible to

the tail components, to form the connector–tail structure of the mature virion (Figure 4D and E).

The existence of two structural conformations of the connector–core complex can be correlated with the packaging of the DNA, as the prohead packaging-competent machinery shows the central channel fully open for the translocation of the DNA (Figures 1 and 5A). The complex from the mature virus has the channel occupied by a density, probably corresponding to DNA. Additionally, the channel of the mature virus complex has several places where it is further constricted, probably due to the rearrangement of the domains caused by the DNA pressure (Figures 4B and 5A and B). These constrictions could be sufficient so as to initially maintain the DNA inside the capsid and therefore prevent its rapid ejection, as tail-deletion mutants of phage T7 produce DNA-filled heads (Cerritelli *et al*, 1997). Nevertheless, an additional stopper must be added, as these tail-less particles lose their DNA under a variety of mild conditions. The fact that the tail shows two regions where the channel is completely closed (Figure 4E, arrows) provides such a permanent closure of the exit way of the DNA until the moment it will be properly injected during the infection of the host bacteria. Similar closing strategies are exhibited by other phages such as T4 (Leiman *et al*, 2004), SPP1 (where the gp16 protein closes the connector channel (Orlova *et al*, 2003)), and $\phi 29$, in which the lower collar protein p11 induces a conformational change of the connector generating the closure of the channel (Carazo *et al*, 1985; Tao *et al*, 1998). Although noncontractile, the T7 tail structure is very similar to the tip of the SPP1 and T4 tails (Orlova *et al*, 2003; Leiman *et al*, 2004), and it is tempting to propose that this tail region must be responsible in T7 for the penetration of the bacterial outer membrane, followed by a structural transition in the channel that releases the DNA inside the infected cell. Higher-resolution studies of these viral structures, as well as the production of the atomic structures of the different components of the DNA translocating machinery, will be required to get more detailed information on how these fascinating processes take place at the molecular level.

Materials and methods

Sample preparation

Mature viruses were purified from *E. coli* lysates of BL21 cells infected with a wild-type strain of T7 using two consecutive centrifugations on sucrose gradients. Type I proheads were purified from *E. coli* lysates of BL21 cells infected with a mutant in gene 5 (the viral DNA polymerase) as described previously (Nakasu *et al*, 1985). The composition of the purified proheads and mature viruses was assessed by SDS–PAGE analysis, their homogeneity checked by nondenaturing agarose gel migration (Cerritelli and Studier, 1996) and their morphology by negative staining electron microscopy. Prior to examination in the microscope, the samples were dialyzed overnight against 50 mM Tris–HCl, pH 7.7, 10 mM MgCl₂, and 0.1 M NaCl solution (TMS).

Electron microscopy and image processing

Low-dose images ($<10\text{e}^-/\text{Å}^2$) of purified proheads and viruses were taken on a FEI TecnaiG² FEG200 electron cryomicroscope at 200 kV using a Gatan side-entry cryo-holder and a nominal magnification of $\times 50\,000$ and $\times 29\,000$ for proheads and mature viruses, respectively. Quantifoil grids with 2 μm holes were used. The selected micrographs were scanned on a Zeiss scanner (Photoscan TD, Z/I Imaging Corporation) with step sizes of 21 and 7 μm for the proheads and mature viruses, respectively. The final pixel size corresponds to 4.06 Å in the case of the proheads,

and 4.66 Å in the mature viruses after a downsampling process using the XMIPP software package (Sorzano *et al*, 2004). The number of particles incorporated into each reconstruction was 4460 for the proheads and 4785 for the mature virus. For all the reconstructions, images were preprocessed to normalize mean intensities and variances and to remove linear background gradients. The image processing included a 3D projection alignment procedure (real space projection matching) with correction of the contrast transfer function (CTF) in defocus groups, varying the defocus of the selected images from 1 to 3 μm in both cases. The CTF correction was performed using the Wiener filter approach. Preliminary icosahedral reconstructions of prohead and mature viruses obtained by common lines procedures at low resolution (34 Å), and modified by the incorporation of the T7 connector (Agirrezabala *et al*, 2005) in one pentameric vertex using XMIPP tools, were used as initial 3D references.

For reconstruction of the core–scaffolding region, the density derived from the capsids from the original images of proheads was removed by projection of the corresponding density in the known orientations of the single particles and subtracting these projections computationally from the corresponding original particles following a procedure similar to that described by Morais *et al* (2003). The corresponding steps were as follows: A 3D alignment of images was performed using as initial models icosahedral volumes, where the structure of the connector was placed in a pentameric vertex. Then, the density corresponding to the capsid was projected into the same orientations of each experimental image. The density exclusively corresponding to the capsid was extracted using SITUS from the five-fold symmetrized reconstructed prohead. The projections of that capsid were subtracted from their corresponding images (preserving the same assigned orientations) without any further re-normalization. Finally, the resulting difference projections were back projected to reconstruct a new volume. After reaching the convergence of the refinement in this core–scaffold reconstruction, the core was extracted for clearer visualization.

The initial models of the mature virus contained only a partially reconstructed tail due to the smaller 3D reconstruction radius used, thus avoiding the noise amplification and interference due to the real space projection matching procedure used. In the initial refinement cycles, the individual images were also band-pass filtered, subtracting the frequencies corresponding to a range of 23–27 Å, and high inner radii were also used in alignments, attenuating in this way the contribution of the highly compacted DNA to the complete signal. After reaching the convergence of these refinements, these reconstructions yield resolutions of 24 and 19 Å for nonsymmetrized and five-fold symmetrized reconstruction, respectively. Once the features of the shell and the partially reconstructed tail became clear, the size of the reconstructed volumes was increased to recover the whole particle. The final resolutions of the reconstructed viruses were 25 and 21 Å, corresponding to nonsymmetrized and five-fold symmetrized reconstructions, respectively. The resolution of the nonsymmetrized prohead yields resolutions of 24 and 18 Å for the five-fold symmetrized prohead, 26 Å for the shell-subtracted core–scaffolding reconstruction, and 23 Å when four-fold symmetry was applied during the refinement. The resolution of these reconstructions was determined by the FSC 0.5 criterion. The hand of the volumes was assumed to be the same as that defined for previously presented icosahedral volumes (Cerritelli *et al*, 2003a). Data processing was performed using the SPIDER image-processing system (Frank *et al*, 1996). General visualization and the manual fitting of the connector volume to the core complex were performed using Amira (<http://amira.zib.de>).

The icosahedral maps used as internal control were built following the same general procedure as described above for nonicosahedral reconstructions, except that icosahedral symmetry was imposed throughout the procedure. For the quantitative comparison of the reconstructions obtained using different symmetries, we determined the correlation coefficients between the reconstructed capsomers. To this end, the envelopes of individual capsomers from different reconstruction procedures were determined by visual inspection of maps contoured at different sigma levels, and were extracted and 3D aligned using XMIPP tools.

Rotational power spectra of the sections from the nonsymmetrized reconstructed prohead volume, the shell-subtracted scaffold–core reconstruction, and the mature virus were obtained in different

planes of the reconstructions using different angular sectors to include the main morphological features in each domain. XMIPP tools were used to center and align the areas selected in each plane to calculate their rotational symmetry.

The maps have been deposited with the European Bioinformatics Institute under accession numbers EMD-1161, EMD-1162, EMD-1163 and EMD-1164.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

References

- Agirrezabala X, Martin-Benito J, Valle M, Gonzalez JM, Valencia A, Valpuesta JM, Carrascosa JL (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
- Carazo JM, Santisteban A, Carrascosa JL (1985) Three-dimensional reconstruction of bacteriophage phi 29 neck particles at 2 nm resolution. *J Mol Biol* **183**: 79–88
- Cerritelli ME, Cheng N, Rosenberg AH, McPherson CE, Booy FP, Steven AC (1997) Encapsidated conformation of bacteriophage T7 DNA. *Cell* **91**: 271–280
- Cerritelli ME, Conway JF, Cheng N, Trus BL, Steven AC (2003a) Molecular mechanisms in bacteriophage T7 procapsid assembly, maturation and DNA containment. In *Advances in Protein Chemistry*, Chiu W, Johnson JE (eds), pp 301–323. New York: Academic Press
- Cerritelli ME, Studier FW (1996) Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. *J Mol Biol* **258**: 286–298
- Cerritelli ME, Trus BL, Smith CS, Cheng N, Conway JF, Steven AC (2003b) A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. *J Mol Biol* **327**: 1–6
- Conway JF, Wikoff WR, Cheng N, Duda RL, Hendrix RW, Johnson JE, Steven AC (2001) Virus maturation involving large subunit rotations and local refolding. *Science* **292**: 744–748
- Cue D, Feiss M (1997) Genetic evidence that recognition of cosQ, the signal for termination of phage lambda DNA packaging, depends on the extent of head filling. *Genetics* **147**: 7–17
- de Haas F, Paatero AO, Mindich L, Bamford DH, Fuller SD (1999) A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage phi6. *J Mol Biol* **294**: 357–372
- Dokland T (1999) Scaffolding proteins and their role in viral assembly. *Cell Mol Life Sci* **56**: 580–603
- Dokland T, Murialdo H (1993) Structural transitions during maturation of bacteriophage lambda capsids. *J Mol Biol* **233**: 682–694
- Droge A, Santos MA, Stiege AC, Alonso JC, Lurz R, Trautner TA, Tavares P (2000) Shape and DNA packaging activity of bacteriophage SPP1 procapsid: protein components and interactions during assembly. *J Mol Biol* **296**: 117–132
- Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG (2004) Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci USA* **101**: 6003–6008
- Fokine A, Leiman PG, Shneider MM, Ahvazi B, Boeshans KM, Steven AC, Black LW, Mesyanzhinov VV, Rossmann MG (2005) Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc Natl Acad Sci USA* **102**: 7163–7168
- Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**: 190–199
- Garcia LR, Molineux IJ (1996) Transcription-independent DNA translocation of bacteriophage T7 DNA into *Escherichia coli*. *J Bacteriol* **178**: 6921–6929
- Greene B, King J (1996) Scaffolding mutants identifying domains required for P22 procapsid assembly and maturation. *Virology* **225**: 82–96
- Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, Sousa N, Carrascosa JL, Coll M (2002) Detailed architecture of a DNA

Acknowledgements

We acknowledge the support of Alasdair C Steven, Benes Trus, Juan Fontana, and Mario Cerritelli during different steps of this work. We also express our recognition to Miyo T Morita for the generous gift of virus and bacterial strains. This work was supported by Grants BMC2002-00996 from the DGI from the Ministerio de Ciencia y Tecnología (MCyT), and EU contract LSHG-CT-2004-502828. JRC holds a Ramon y Cajal Contract from the MCyT. XA acknowledges a fellowship from the Gobierno Vasco/Eusko Jaurlaritza.

- translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle. *J Mol Biol* **315**: 663–676
- Guo PX, Erickson S, Anderson D (1987) A small viral RNA is required for *in vitro* packaging of bacteriophage phi 29 DNA. *Science* **236**: 690–694
- Guo PX, Erickson S, Xu W, Olson N, Baker TS, Anderson D (1991) Regulation of the phage phi 29 prohead shape and size by the portal vertex. *Virology* **183**: 366–373
- Ibarra B, Caston JR, Llorca O, Valle M, Valpuesta JM, Carrascosa JL (2000) Topology of the components of the DNA packaging machinery in the phage phi29 prohead. *J Mol Biol* **298**: 807–815
- Jiang W, Li Z, Zhang Z, Baker ML, Prevelige Jr PE, Chiu W (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**: 131–135
- Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118**: 419–429
- Mancini EJ, Kainov DE, Grimes JM, Tuma R, Bamford DH, Stuart DI (2004) Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **118**: 743–755
- Matsuo-Kato H, Fujisawa H, Minagawa T (1981) Structure and assembly of bacteriophage T3 tails. *Virology* **109**: 157–164
- Morais MC, Choi KH, Koti JS, Chipman PR, Anderson DL, Rossmann MG (2005) Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol Cell* **18**: 149–159
- Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BA, McMurray CT, Anderson DL, Rossmann MG (2003) Bacteriophage phi29 scaffolding protein gp7 before and after prohead assembly. *Nat Struct Biol* **10**: 572–576
- Morita M, Tasaka M, Fujisawa H (1995) Structural and functional domains of the large subunit of the bacteriophage T3 DNA packaging enzyme: importance of the C-terminal region in prohead binding. *J Mol Biol* **245**: 635–644
- Nakasu S, Fujisawa H, Minagawa T (1985) Purification of characterization of gene 8 product of bacteriophage T3. *Virology* **143**: 422–434
- Orlova EV, Gowen B, Droge A, Stiege A, Weise F, Lurz R, van Heel M, Tavares P (2003) Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. *EMBO J* **22**: 1255–1262
- Padmanabhan R, Wu R, Bode VC (1972) Arrangement of DNA in lambda bacteriophage heads. 3. Location and number of nucleotides cleaved from lambda-DNA by micrococcal nuclease attack on heads. *J Mol Biol* **69**: 201–207
- Serwer P (1976) Internal proteins of bacteriophage T7. *J Mol Biol* **107**: 271–291
- Simpson AA, Tao Y, Leiman PG, Badasso MO, He Y, Jardine PJ, Olson NH, Morais MC, Grimes S, Anderson DL, Baker TS, Rossmann MG (2000) Structure of the bacteriophage phi29 DNA packaging motor. *Nature* **408**: 745–750
- Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C (2001) The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **413**: 748–752
- Sorzano CO, Marabini R, Velazquez-Muriel J, Bilbao-Castro JR, Scheres SH, Carazo JM, Pascual-Montano A (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**: 194–204
- Steven AC, Heymann JB, Cheng N, Trus BL, Conway JF (2005) Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr Opin Struct Biol* **15**: 227–236

- Steven AC, Trus BL, Maizel JV, Unser M, Parry DA, Wall JS, Hainfeld JF, Studier FW (1988) Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J Mol Biol* **200**: 351–365
- Tao Y, Olson NH, Xu W, Anderson DL, Rossmann MG, Baker TS (1998) Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95**: 431–437
- Tavares P, Santos MA, Lurz R, Morelli G, de Lencastre H, Trautner TA (1992) Identification of a gene in *Bacillus subtilis* bacteriophage SPP1 determining the amount of packaged DNA. *J Mol Biol* **225**: 81–92
- Thuman-Commike PA, Greene B, Malinski JA, King J, Chiu W (1998) Role of the scaffolding protein in P22 procapsid size determination suggested by $T=4$ and $T=7$ procapsid structures. *Biophys J* **74**: 559–568
- Tzlil S, Kindt JT, Gelbart WM, Ben Shaul A (2003) Forces and pressures in DNA packaging and release from viral capsids. *Biophys J* **84**: 1616–1627
- Valpuesta JM, Carrascosa JL (1994) Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Quart Rev Biophys* **27**: 107–155
- van Driel R, Couture E (1978) Assembly of the scaffolding core of bacteriophage T4 preheads. *J Mol Biol* **123**: 713–719
- Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**: 2129–2133