doi:10.1016/j.jmb.2003.[illegible] available online at http://www.sciencedirect.com on [illegible] J. Mol. Biol. (2003) 331, [illegible]–[illegible]

JMB

Detailed Architecture of a DNA Translocating Machine: The High-resolution Structure of the Bacteriophage ϕ29 Connector Particle

Alicia Guasch[1†], Joan Pous[1†], Borja Ibarra[2], F. Xavier Gomis-Rüth[1], José Maria Valpuesta[2], Natalia Sousa[3], José L. Carrascosa[3] and Miguel Coll[1*]

[1]Institut de Biologia Molecular de Barcelona, CSIC, Josh Clarens 18-26, E-08034 Barcelona, Spain

[2]Centro Nacional de Biotecnología, CSIC, Universidad Autónoma de Madrid, Cantoblanco E-28049, Madrid, Spain

[3]Centro Nacional de Biotecnología, CSIC, Universidad Autónoma de Madrid, Cantoblanco E-28049, Madrid, Spain

*Corresponding author

The three-dimensional crystal structure of the bacteriophage ϕ29 connector has been solved and refined to 2.1 Å resolution. This ~225 kDa oligomeric protein connects the head of the phage to its tail and translocates the DNA into the prehead during packaging. Each monomer has an elongated shape and is composed of a central, mainly α-helical domain that includes a three-stranded helix, a distal α/β domain and a proximal elongated β-barrel domain. The protomers assemble into a 12-mer propeller-like supramolecular with a 35 Å wide central channel. The surface of the channel is mainly electronegative, but it includes two lysine rings 20 Å apart. On the external surface of the particle a hydrophobic belt extends to the connector area below the β-barrel domain, which forms a crown that retains the particle in the head. The hydrophilic belt contacts the concentric symmetry vertex of the capsid and forms a bearing for the connector rotation. The structure suggests a translocation mechanism in which the longitudinal displacement of the DNA along its axis is coupled to connector spinning.

© 2003 Academic Press

Keywords: bacteriophage ϕ29; X-ray diffraction; connector particle; DNA packaging machinery

Introduction

DNA translocation is a central biological issue since it is essential to the processing and transfer of genetic information, including cell infection by viruses. It involves a number of basic mechanisms that are probably common to many systems (nucleic acid replication, transcription, translation, conjugation, etc). Many proteins involved in DNA translocation, or processive movements along the DNA, function as oligomers with a general toroidal morphology.[1] Among these toroidal proteins are the bacteriophage head-to-tail connectors, present in one particular vertex of viral heads. The head-to-tail connecting region of bacteriophage ϕ29 is one of the best known. ϕ29 is a small double-stranded DNA bacteriophage that infects Bacillus subtilis cells. The viral particles are formed by a prolate icosahedral capsid or head, which contains a 19 kb long DNA, and a tail of complex structure. Between these two structures there is a connecting region called portal vertex or connector. This head-to-tail connector plays an important role in the first steps of head assembly and is the central piece of the DNA packaging machinery. In addition, during viral infection, the DNA is ejected through the tail and hence traverses the connector in the opposite direction (for reviews see references[2,3]).

The packaging of DNA into the head involves, besides the connector, two other essential components that interact in its distal part: an RNA called pRNA,[4] and the ATPase gp16, required to provide energy to the translocation machinery. Both RNA and ATPase are of viral origin and, once the DNA has been packaged, they are released from the connector while the collar and tail components assemble to make distal parts.

Electron microscopy studies, based on two-dimensional projections and three-dimensional reconstructions, showed that the ϕ29 connector is an oligomeric protein built up from 12 copies of

†Both authors contributed equally to this study and share first authorship.

E-mail address of the corresponding author: mcoll@ibmb.csic.es

the 30 kDa gene 10 product.[5] This structure is shaped like a propeller, in two parts: a wide disk, with protruding loops that interacts with the capsid proteins, and a narrower cylinder that has a longitudinal channel along its axis and contacts the collar and tail proteins.[6]

In spite of the absence of sequence homology, viral connectors have similar structural features. The three-dimensional reconstructions of the connector from bacteriophages T3[7] and SPP1[8] have revealed striking similarities to the φ29 connector. This common architecture is also seen in other viruses: the connectors from bacteriophages λ,[9] p22[10] and T4[11] show projection images that are fully consistent with a ring of appendages surrounding a narrower channel with a channel, similar to those obtained from φ29,[1,12,13] T3,[7] or SPP1.[8] These connectors differ only in dimensions (which correlate to the molecular mass of the connector protein), and all of them present 12 outer appendages, except for SPP1, which has 13.

We have previously reported the crystallization of the φ29 connector particle in two crystal forms (C2 and $P4_2$).[14] The crystals diffracted to 3.6 and 3.0 Å, respectively, the second form showing merohedral twinning. These initial crystallographic studies, together with the cryo-electron microscopy data[12] and the atomic force microscopy images obtained from connector aggregates in solution,[15] offered a consensual view about the symmetry of the φ29 connector (reviewed in reference[1]). A cell-rotation function calculation demonstrated unambiguously the 12-fold local symmetry of the particle.[14] Recently, Simpson et al.[16] reported an unrefined low-resolution crystal structure of the connector, determined from C2 crystals. Although our present data fit the general outline of that structure, the previous model differs markedly in both the chain tracing and the side-chain positioning at the proximal part, where the β-strands in each of the 12 protomers are missing and the main chain is out of register from residue 271 onward. Therefore, the crucial SH3-like domain is not described in Simpson's model.

Here we report the high-resolution 2.1 Å structure of the bacteriophage φ29 connector particle, obtained from a new crystal form. The quality of these data allowed us to correctly trace the polypeptide chain and refine the structure. The resulting structural details reveal an amazing RNA-translocation device and suggest a simple rigid-body rotary mechanism of the connector coupled to the linear movement of the double helix through its central channel.

Results

Quality of the data

All previous attempts to obtain high-quality diffracting data on the connector particle were unsuccessful. The size of the particle and, most probably, the presence of flexible regions at the N and C termini of the protein have so far prevented the collection of high-resolution data. This problem was solved by the controlled proteolysis of p10 by the protease CfaC from *Staphylococcus aureus* V8, which removes the first 13 residues from the N terminus and the last 18 residues from the C terminus.[17] The C2 crystals obtained previously by us[14] and others[16] contain one or several particles in the asymmetric unit. This feature challenges the phase extension procedure, because 12-fold averaging can be applied, but the monoclinic crystals hardly diffract to 30 Å. On the other hand, the orthorhombic $P4_2$ crystals[14] diffracted to higher resolution (about 3 Å) but showed a merohedral twinning problem, where a 50% mixture of up and down connector orientations was present.

The crystals here reported have a new chiral space group $A22_1$. One connector particle occupies a special position with its 12-fold local axis coincident with the crystallographic 4-fold axis. Therefore, the asymmetric unit includes only three protomers. The particles are tightly packed with many intermolecular contacts and the crystals diffract to 2.1 Å, a remarkably high resolution for a particle of this size.

Structure determination

A 10 Å resolution electron cryomicrography model[12] was used for initial phasing in the averaging/phase extension procedure. The coincidence of the 12-fold axis of the connector with the crystallographic 4-fold axis and the fact that only 3-fold averaging could be applied resulted in an uninterpretable map. However, because high resolution was available, this map was used for an automatic structure determination with the program *Arp/Warp*. By alternating *Arp/Warp* cycles with 3-fold averaging cycles, the map improved substantially and some secondary structure elements showed up, in particular the long helices of the central channel. At this stage, however, the coordinates of the Simpson X-ray model became available (pdb code 1hp4) and they were also used for molecular replacement. The model was not good enough to provide a solution using a protein monomer, but a search with the whole connector (see Materials and Methods) gave an identical solution to that obtained with the electron cryomicrography model. From the resulting maps, given the high resolution of the data, the chain tracing could be corrected. Subsequently, the structure was fully refined, with the addition of solvent molecules (figure 1b).

Structure of the monomer

Each monomer consists of four domains (figure 2a): a central domain that extends to the internal face of the wide proximal end, an SH3-like domain at the exterior of the wide proximal end, and a distal domain at the tip of the narrow end. The central domain contains five α-helices (α1, α2, α3, α9 and α10) and two β-strands (β1 and β10).

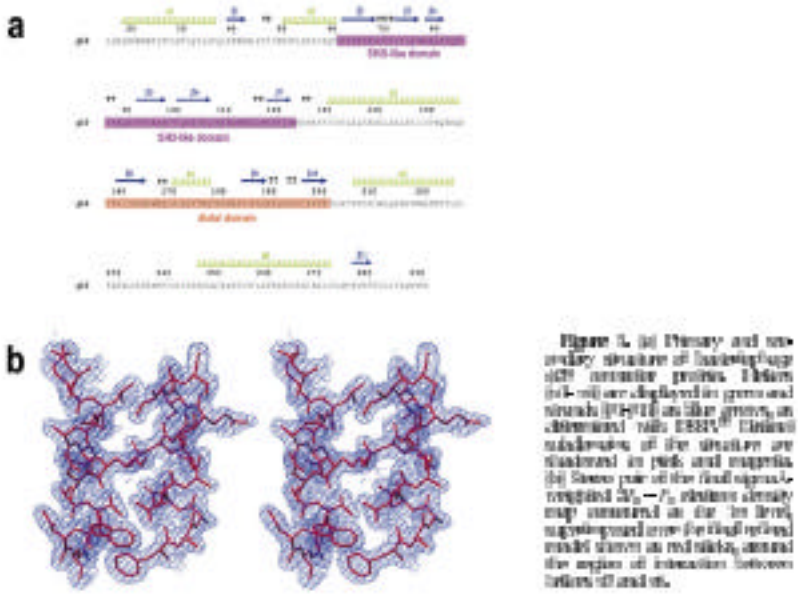

Figure 5. (a) Primary and secondary structure of bacteriophage σ32 connector protein. Helices (α1–α6) are displayed in green and strands (β1–β11) in blue, green, as determined with DSSP.[17] Distinct subdomains of the structure are shaded in pink and magenta. (b) Stereo pair of the final sigma-weighted $2F_o - F_c$ electron density map contoured at the 1σ level, superimposed over the final refined model shown as red sticks, around the region of interaction between helices α5 and α6.

Figure 5(a) shows the residues included in each secondary structure element. Helices α4, α5 and α6 form a bundle in which α5 and α6 are antiparallel and α4 forms an angle of 30° with the other two, running antiparallel to α5. Helix α2 is located at the C terminus of α1 and the N terminus of α3, almost perpendicular to both, while helix α6 is placed at the C terminus of α5. Helix α6 has a kink between residue Val256 and Ile257. This causes a deviation of the helical axis of about 30°. Interestingly, the axis of helices α5 and the second part of α6 (after the kink) are almost coincident and thus, this part of α6 appears as a continuation of α5. Finally, strands β1 and β11 form an antiparallel β-ribbon, at the proximal part of the central domain.

A gap of 14 residues, from 231 to 244, in the loop between α5 and α6, are not seen in the electron density maps. We checked the molecular weight of the crystallized protein by mass spectrometry and ruled out proteolytic cleavage at this site. Therefore, this loop is a flexible region, facing the interior of the channel, and it might play a role in closing the gate after DNA packaging. The rest of the loops of this central domain are well defined and include two β-turns, between β1 and α2 and between β7 (at the proximal SH3-like domain) and α3.

The SH3-like domain is inserted between α2 and α3 of the central domain, and forms the protruding lobe(s) of the proximal wide part of the particle. It consists of six β-strands (β2 to β7) with an SH3-like β-barrel topological arrangement, where β2, β3 and β7 form an antiparallel three-stranded β-sheet which lies perpendicular to the second antiparallel three-stranded β-sheet formed by β4, β5 and β6 (Figure 2(b)). The two sheets enclose a compact hydrophobic core that includes six aromatic residues, three of them (Tyr63, Phe104 and Phe108) perfectly stacked. The connections of β2 with β3, β3 with β4 and β5 with β6 are rather short, with a β-turn between β2 and β3. On the other hand, the connections from β4 to β5 and β6 to β7 are terminal long loops or hairpins, both including β-turns. A structural homology search with the program DALI[18] did not reveal any similarity with other known structures. Nevertheless, it is clear from Figure 2(b) and superposition studies that the topology is similar to the SH3 proteinogenic recognition domains. However, it does not form a cleft for polypeptide binding, because loop β6–β7 would occupy part of the cleft space and loop β3–β4 is too short to form one of the walls of the cleft, in contrast to the long equivalent RT-loop of typical SH3 domains.

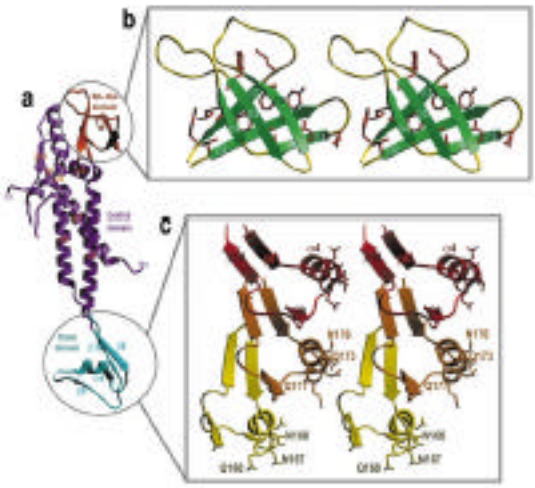

Figure 3. (a) Ribbon cartoon of a bacteriophage φ29 connector protein monomer displaying the labelled secondary structure elements. The three domains of the connector monomer are shown in different colours. (b) and (c) Close up views of the core domain and distal domain of the bacteriophage φ29 connector protein monomer, the SH3-like domain (b) and the distal domain (c). The regular secondary structure elements are labelled. In (b) the hydrophobic residue substitutions making up the filling of the β-sandwich are displayed as orange sticks. In (c) three visual distal domain are displayed in red, orange and yellow, respectively. In the latter two, the side-chains of the residues making up the Glu/Asn knob are displayed as sticks. Figure created with SETORS[39].

The distal domain, at the narrow tip of the connector, consists of three β-strands (β5, β9 and β10), one α-helix (α4) and a β-turn located between the last two β-strands (Figure 2(c)). Two strands (β5 and β10) of one monomer and one strand (β9) of an adjacent monomer form a mixed β-sheet where β5 and β9 are parallel and β10 antiparallel. Helix α4 runs antiparallel to β5 and also belongs to the adjacent monomer (Figure 2(c)). In this way the monomers are strongly intermonomer.

The architecture of the φ29 protein monomer does not show any similarity with other known protein structures. Only the alignment of the three-helix bundle in the central domain resembles of other three-helix bundles, like that found in the N-terminal domain of the low-density lipoprotein receptor adrenal protein (LDP (pdb code 1lrb).

Structure of the dodecamer

Twelve protomers assemble in a propeller-like supramolecular structure that has an external diameter of 146 Å at its proximal part and 77 Å at its distal end (Figure 3(a) and (b)). The height of the connector is 79 Å. The three hollow stacking the bundle of the central domain are not parallel with the 12-fold axis of the particle, but slightly inclined, about 45°. This implies that the proximal part of monomer n, including the SH3-like prolamerase, projects over the distal domain of monomer n + 2, as seen when looking down the 12-fold axis of the particle.

The interactions between monomers are different in the three parts of the structure. For example (as indicated above), in the distal domain one of the β-strands of the three-stranded β-sheet belongs to the adjacent monomer and, therefore, multiple hydrogen bonds exist between monomers (Figure 2(c) and 4(c)). In addition, there are several hydrophobic contacts involving eight monomer residues: Tyr[illegible]–Ile[illegible], Leu[illegible]–Ile[illegible], Asp[illegible]–Ala[illegible], Thr[illegible]–Ala[illegible] and Leu[illegible]–Ile[illegible] (A, B and C denote adjacent monomers). Furthermore, there is an H-bond between Arg[illegible]A and Ser[illegible]B. Interestingly, residues of three consecutive monomers (Thr[illegible]A, Tyr[illegible] and Ile[illegible]C) converge at one point and form hydrophobic contacts between each other.

In the central domain most of the contacts are formed by residues belonging to the helices (Figure 4(b)). Helices α2 of two consecutive monomers run parallel along almost their entire length, establishing multiple contacts. Helix α2 contacts an adjacent α2 with its C-terminal half, while its N-terminal half contacts α1. The N-terminal part of the kinked α1 helix interacts with the last turns of α1 and α2 of a neighbouring protomer.

In the proximal part, the C-terminal half of helix α2 contacts the C-terminus of helix α2 and residues of strand β1 (Figure 4(a)). More important are the interactions of the SH3-like domain, loop β3–β4, together with the loop β1–α2 (which does not belong to the SH3-like domain), but which is also in the proximal part of the connector) define a hydrophobic cleft with strand β1 at the bottom. Loop β3–β4 of a neighbouring protomer interacts with this cleft, establishing hydrophobic interactions. Tyr[illegible], is at the tip of the loop and forms several hydro-

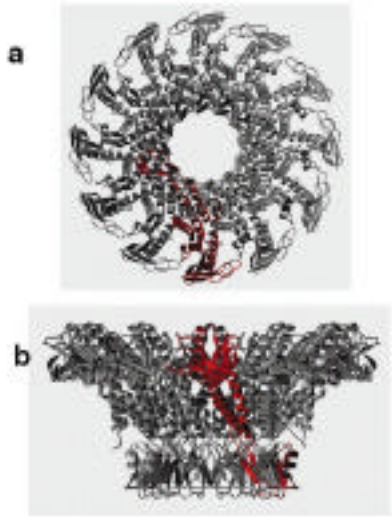

Figure 3. (a) and (b) Ribbon representations of a whole dodecamer of the bacteriophage φ29 connector protein. One of the 12 monomers is displayed in different red. The structure is shown in the usual (a) and turned by views. Figure made with VOLUMES/RIBBONS[19] and rendered with Ray-Shade[20] (c) Location of the connector structure in the viral prohead. The connector 3D-cryo microscopy was fitted into the density corresponding to the φ29 prohead obtained by cryo-electron microscopy at 23 Å resolution[4]. The fit of both volumes, and their graphic representation were generated using the VMD program[21]

---

polar contacts with Ala[238] and Ile[240], both from β2. It is adjacent to stacked on Phe[238]. Furthermore, the CH2 group forms H-bonds with the carbonyl oxygen of Thr[238] and the amide nitrogen of Ala[239]. Values of the same β-β[2] loop contacts Ile[240] and Val[241], while Gln[242] forms an H-bond with Asn[243]. Further contacts are observed between loop β6-β7 and loop β1-β2 of a neighbouring protomer, including Arg[104]-Glu[240] and Tyr[103]-Thr[238].

The buried surface of a monomer-monomer interface is 2139 Å^2, which means that, in the dodecamer, only one third of the small area of each monomer is accessible. Thus, a large proportion of the surface is involved in the intersubunit contacts, resulting in a strongly bound particle. This is in accordance with previous biochemical data indicating that single monomers are never observed, even at high salt concentrations. The connector structure is disassembled[17] only when incubated at high chaotropic agent concentrations (2.5 M guanidinium hydrochloride).

Internal channel and surfaces

A channel traverses the particle longitudinally. Its width is about 23 Å at the distal end and widens up at the proximal end. The width at the proximal end is difficult to estimate accurately, however, since the cleavage and hence not much dial residues 270–294 and the extend C terminus are present in that area. The proximal end of the channel may thus be much narrower than is shown in this model.

In Figure 5(a), a space filling representation highlights different features of the interior surface of the particle. Going from the proximal wide end to the distal narrow end, we first find the positive area that corresponds to the S2-DNA-binding domain with positively and negatively charged patch residues. At the subsequent narrow portion, there is a belt of hydrophobic residues, including Phe[58], Ile[23], Leu[28], Phe[61], Phe[238], Ile[239], and Ile[112]. Then, another area of mixed positively and negatively charged residues follows. This part corresponds to the N terminus of the protein, where 31 residues were ordered. Further down, a ring with exposed Asn and Gln residues forms the internal surface of the distal domain. It encompasses a large concentration of residues with side chains having amide functions, including Asn[186], Asn[187], Gln[188], Gln[122], Asn[176] and Gln[177]. All these residues are found at the exposed face of helix α4 (Figure 2(b)). Finally, at the apical end of the connector, surrounding the entrance to the channel, there is an acidic ring formed by the side chains of Glu[104], Asp[102] and Asp[106].

Figure 5(b) shows the electrostatic potential interior surface of the connector, which appears to be highly electronegative. However, closer inspection reveals two electropositive rings, due to the presence of residues Lys[188] and Lys[200], the ring in the middle of the narrow channel and the second at the beginning of the wide channel. The distance between the two rings (i.e. from Lys[188] Cζ to Lys[200] Cζ) is 20 Å, while the distance between adjacent lysine residues of the same ring is 9.0 Å.

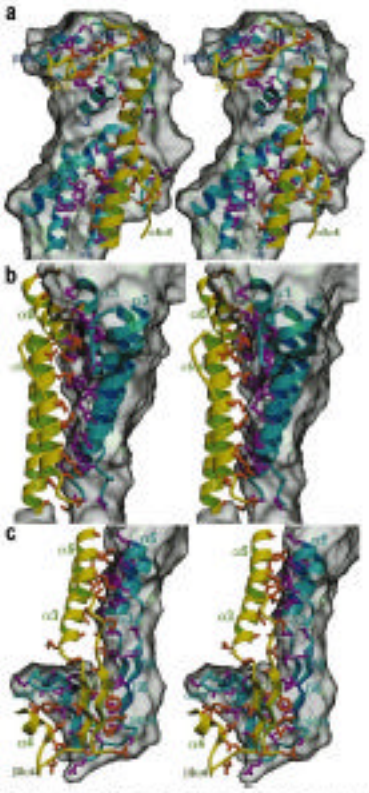

Figure 3. Close up stereo view of a protein monomer in light blue (ribbon) and magenta (helical side-chains), superimposed over its neighbouring. Centrally motion, and the appendix of a vinnil monomer (yellow ribbon, orange side-chains) taking part in inter-monomer contacts. The three subfigures display the β-like domain and the upper part of the central domain (a), the central helical domain (b) and the distal domain with the lower part of the central domain (c). figure prepared with VOLUMES/HOMEDOMT[illegible] and rendered with Raster3D.[illegible]

The electron density is well defined for most of these lysine side-chains. The distances of the rings are 36 and 41 Å, respectively. In addition, the clockwise 236-288 area contains two lysine residues, Lys234 and Lys238. Both residues 231 and 240 are at 30 Å from the nearest lysine ring towards the proximal end. Therefore, there are two local positively charged rings along the channel, at 30 Å intervals. Both Lys230 and Lys239 have their electronegative counterparts in close proximity, Asp232 and Asp240. It is interesting that the position of the negative side-chains is quite similar in both rings, with an alternation of Asp and Lys residues around a channel section. In the Lys230 ring, however, we observe a glue salt bridge, as long as one side of the sheet monomers of the crystal asymmetric unit, between the lysine and one of the flanking negative residues, while the other is at a distance of 6.8 Å. This is because the lysine side-chains are not oriented radially towards the centre of the channel. In contrast, in the Lys239-ring they point to the centre of the channel and are equidistant, 5 to 6 Å away, from the two flanking negative residues, without establishing any interaction with them.

Discussion

Requirements for DNA packaging

The packaging of DNA into preformed prolands is a mechanism exhibited by complex double-stranded DNA bacteriophages, as well as other animal viruses (adenovirus, herpesvirus).[illegible] The packaging process[illegible] has two important aspects. First, there is a progressive movement of the DNA through the packaging machinery.[illegible] The other characteristic feature of the viral DNA packaging process is the condensation of the DNA inside the viral head to crystalline concentrations. The translocation of a long, rather stiff DNA polymer into a proximal container imposes severe limitations on the function of the packaging machinery and the movement of the DNA, as the acquisition of the solenoid-like ordered parallel arrangement of the DNA inside the viral prohead generates pressure and a high concentration of cross and charges. Thus, the viral DNA packaging machinery has to deal with the general problems of translocation of nucleic acids by protein pumps, and at the same time has to find suitable solutions to release the tensions accumulated by the forced concentration of a highly charged polymer in a limited container.

The connector shows a unique architecture

It was reasonable to believe that the movement of the DNA would share some common mechanisms with other nucleic acid processive systems,[illegible] such as the sliding clamps,[illegible] motor helicases[illegible] or conjugation systems.[illegible] However, the φ29 connector analysed here has a completely novel architecture, which suggests differences in the way the motors translocate. For example, ring helicases and the conjugation protein TraB have a basic RecA/F1-ATPase architecture, including a core formed by a parallel doubly wound β-sheet with helices on both sides that encompass a nucleotide binding domain. These proteins may contain additional domains, but neither these nor

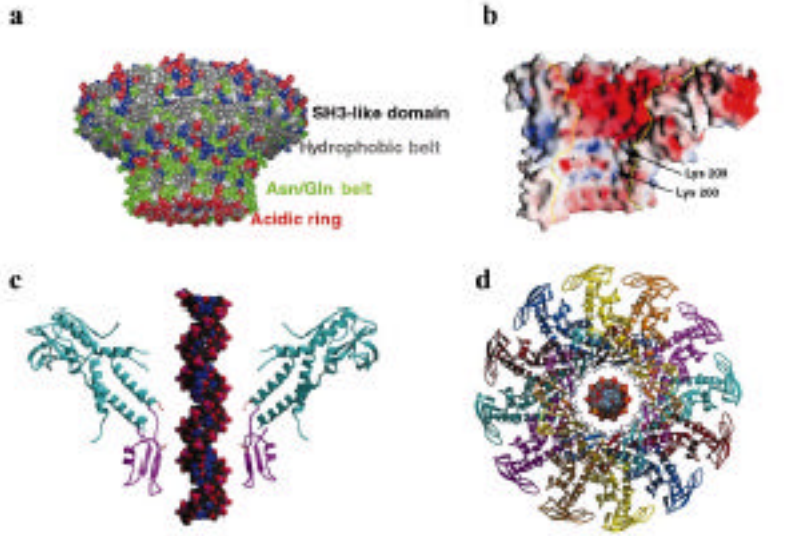

Figure 5. (a) Atomic-resolution model of the φ29 connector particle. Colour coding of the subdomains goes for Asn/Gln, red for Asp/Glu, grey for hydrophobic residues, and dark blue for Lys/Arg. The model was generated with CRASP.[30] (b) Molecular surface coloured according to electrostatic potential and the residues, blue for positive belonging to six viral protein monomers generated with CRASP.[30] The positions of the two terminal lysine rings (marked as Lys309 and Lys300 of each monomer, respectively) is indicated. (c) Ribbon plot showing a longitudinal cut of the connector particle. Note that the arrows of the distal domains are different from the medial and N-ribbon domains since they belong to different monomers. A dsDNA molecule is shown traversing the central channel. The subdomains of Lys309 and Lys300 are indicated. (d) View along the connector axis with the modelled DNA. The side chains of Lys309 and Lys300 are displayed as balls and sticks.

the central nucleotide binding domain share any resemblance with the connector p10 protein monomer. Regarding the size of the whole particle, the connector is also unique, larger than all the others. Ring helicases are normally hexamers, while all the functional connectors are formed by 12 subunits.[7] The shape of the connector is unique too, with its distal narrow cylinder and a wide proximal disc of a diameter twice as large as the DNA, quite different from the more oligomerical or cylindrical shape of helicases and TrwB. A notable functional difference with ring helicases is that the connectors do not contain ATPase activity, and thus need ancillary ATP hydrolases to function, which is similar to what happens in sliding clamps. The protomer of a sliding clamp has a completely different fold from a connector protomer, but their oligomeric structure shows striking similarities with the φ29 connector particle: the central channel is similarly defined by 12 subunits that are instead with respect to the particle axis, giving a propeller-like appearance.[8] The distribution is such that the helices would lie perpendicular to the supposed composite backbone of a putative DNA traversing the channel. Kong et al.[8] suggest that this feature prevents the entry of the protein into the DNA groove, facilitating its rapid motion. A further similarity between sliding clamps and the connector is the diameter of the channel, which is structurally determined by the diameter of the DNA double helix.

B-DNA fits smoothly into the channel

The connector defines a central channel of appropriate size to accommodate a double-stranded DNA molecule. The diameter of the channel is 35 Å at its narrowest part, while the DNA B-form has an average diameter of 23 Å. This leaves about 6 Å between the DNA and the channel wall. This wall is mostly electronegative (Figure 5(b)), as is the tip of the connector where the DNA enters the channel. The electronegative nature of the channel surface is not incompatible

with its direction since it does not have to bend tightly in DNA, but rather turns it to pass smoothly through at 36 Å pitch.[10,11] The DNA would traverse the channel leaving away from the repulsive walls. There is room for a layer of water molecules between the DNA and the protein. It is notable that the majority of gp10 proteins involved in single-stranded DNA translocation also displays a rather electronegative channel.[12]

The exception to the extended electronegative nature of the channel surface are the two lysine rings (310 and 209) and positively two more at Lys219 and Lys220. When fitting the DNA into the channel, if a particular backbone phosphate is aligned with one of the Lys310, another lysine of the same ring would be aligned with another phosphate from the other DNA strand (Figure 3b) and 3d). This second lysine is not opposite the first one, across the channel, but that belonging to the fifth monomer. Keeping these two residues of ring 310 aligned with two phosphates, we again observe that at the second lysine ring (209) two lysine side-chains, five monomers apart, lie in front of two phosphate groups. This is because the 36 Å that separates the two lysine rings is about the same distance that separates two phosphates of the DNA across the major groove, in the helix axis direction. In consequence, at any particular moment during the translocation, the DNA molecule seems to be able to remain electrostatically at four points of the channel and presumably two more at the most distal lysine rings.

Sliding groove and SH3-like ridge

The Navy structure of the φ29 connector fits very well in the three-dimensional reconstructions from cryo-electron microscopy of both prolateral[13] (Figure 3b) and prefilled connectors[14] (not shown). Although the docking does not reveal details due to the limited resolution and the averaging imposed in the microscopy reconstruction, it is clear that the wider part of the connector, comprising some elements of the central domain and the whole SH3-like domain, is buried inside the cavity of the head, while the helical bundle of the central domain is inserted into the dodecahedral portal vertex. In particular, the hydrophobic belt at the external connector area or groove of the connector is facing the corresponding area of the head protein building the vertex.

As indicated in Results, the SH3-like domain does not have the canonical SH3-like loops for polyproline recognition. Its topology, however, is similar to SH3 domains, with two sets of three-stranded β-sheets arranged at 90 degrees to each other, forming a barrel, which produces a tight hydrophobic core. A comparison with other structures with the same topology, but not strictly SH3, reveals that it closely resembles the DNA-binding domain of the HIV1 integrase.[15] This protein, also a viral protein like φ29 gp10, dimerizes using one of the β-sheets as a dimerization surface. In the φ29 connector the SH3-like domains from five various monomers also contact each other but in a different way. They assemble in a sliding handle-like fashion, using the protruding loops β5-β6 and β6-β7, which fit into the groove defined by β5-β6 and β1-α2 of the next monomer (Figures 4b) and 5b). Thus, the distinct arrangement of the loops in the SH3-like domain facilitates protein-protein contacts, which finally hold together the protruding lobules of the wide proximal end. The SH3-like domain is therefore a sort of protein recognition motif, since it serves to assemble the compact protuberances that form the outer crown of the connector.

Spinning mechanism

Hendrix[16,17] suggested a way to correlate translocation of DNA and its packaging into the viral prolateral head on the symmetry mismatch between the 5-fold vertex of the prolateral and the dodecahedral symmetry of the tail components. The hypothesis that the rotation of the connector is actively involved in the actual DNA translocation was strongly supported by the disruption of 12-fold symmetry in mutant head-tail connectors when extracted from viral particles or assembled from overexpressed proteins.[18,19,20] A longitudinal movement of the DNA, coupled to the connector rotation, has been proposed for both SPP1[21] and φ29[22] phage connectors.

The surface features as derived now in the high-resolution connector structure are consistent with the spinning hypothesis. As previously discussed[23] the rotation of the connector with respect to the non-rotating 5-fold symmetry head vertex could take advantage of the hydrophobic groove used as a sliding narrow. The adjacent crown of SH3-like domains would retain the connector inside the capsid while spinning.

Is the rotation of the connector enough to drive the DNA into the head? The finding of the two (or three, if we also assume the alleged loop to interact lysine rings at regular 36 Å spacings in an otherwise electronegatively coated channel, together with the helical arrangement of the negatively charged DNA phosphate groups, strongly suggests that both movements are coupled. A one nanometer longitudinal displacement of the DNA would result in the wandering of four (or six) hydrogen-phosphate interactions. The rotating of the four interactions, but this time involving four new lysine residues of adjacent monomers, would occur only if the connector rotated. In other words, the energy barrier for jumping from a particular set of hydrogen-phosphate interactions to the next set would be overcome by the rotary drive of the connector. For this model to work, the rotation of the connector has to be clockwise when looking from the wide domain, i.e. in the opposite direction of the DNA displacement. This is determined by the right-handed nature of the double helix, as the right-handed groove of a helix determines the clock-

value rotation of the rod in order to drive the bulk in the direction of the phage head.

A characteristic feature of this model is that the rotation of the connector has to match the symmetry of the DNA, rather than the symmetry of the connector itself. In that sense, a rotation of 6° would be appropriate. Such a rotation would align the position of the helix residues with respect to their phosphate counterparts in a way that the most conserved helix residues are exactly the right place for interacting with the next set of phosphates, providing a DNA translocation of one base-pair in the proposed (figure 6).

A spinning model has the beauty of simplicity. It does not need molecular rearrangements within the connector and thus the spool does not depend on transitional intermolecular movements. Other more complicated models, like the one put forward by Simpson et al.[8] assume, besides the rotation, a domain displacement within the gp10 protomers and a springlike elongation and contraction of the connector length, which would require the relative slipping of the central solution between monomers. The φ29 connector quaternary structure analyzed here appears to be very compact, with a high percentage of protomer surface involved in intersubunit contacts. This arrangement gives little ground for extensive structural rearrangements, limiting only as a possible hinge candidate the region that connects the central helix bundle domain with the narrow β-structured RNA-binding tip (figure 3b). However, even here, we do not observe a flexible connecting region, for there are no Gly residues. A superposition of the three monomers of the asymmetric unit does not reveal any substantial differences. Although we cannot rule out a movement between domains or between subunits, we do not find any indication for such a movement in the solved high-resolution structure.

Localization and possible role of the pRNA

The packaging motors of the φ29 system require a component protein, which contains, beside the capsid, the pRNA and the ATPase gp16.[1] These are indispensable components of the packaging machinery since they are, in our model, as a stator and ancillary motor, respectively. The pRNA required for the packaging of φ29 DNA has been located in the narrow tip of the connector.[6,8] The docking of the connector structure in the portal region reconstructed from protein containing pRNA clearly shows the pRNA moiety forming a protruding ring that enhances the distal domain of (right-handed) β-strands at the apical region of the gp10 dodecamer (figure not shown). A peculiar aspect of this region is the presence of a ring built by helices rich that presents a high content of exposed glutamine and asparagine residues. This unusual arrangement offers an unusual polar surface rich in hydrogen bond donors and acceptors that could play a role in the interaction with the pRNA.

The pRNA is a unique component of the φ29 packaging machinery that interacts with the outer end of the connector in a multimeric form.[8,10] Extensive analysis of this pRNA, based on transversal assays of directed mutant pRNA molecules indicated the importance of the interaction between complementary loops in adjacent pRNA molecules (inter-subunit complementarity).[8] The data that two or three different mutually complementary mutant pRNAs are required for restoring packaging activity in DNA packaging assays in vitro led to the conclusion that the pRNA should be arranged in a multimer containing 2n and 3n subunits. This conclusion is supported by the finding that a set of six mutually complementary mutants are fully active in packaging, strongly suggesting the hexameric arrangement of the pRNA, as the functional oligomer that is assembled in portals competent for DNA packaging.[8,10] A different pRNA arrangement has recently been proposed[11] where a pentamer has been found associated with the presence of pRNA in portals reconstructed from electron microscopy.

Assuming the rotation of the connector, there are two possible scenarios for the pRNA. In one of them, it rotates with the connector. In the other, it is rigid in the head and does not follow the connector movement. The hinge is an emerging hypothesis since the pRNA could act as a stator ring in the head vertex. A 5-fold symmetry would suit this case better. The alternative scenario is also acceptable: the pRNA would be fixed to the connector, constituting a washer to hold it in place, and would be the equivalent piece of the 12-His crown, but on the outside of the capsid wall. A 6-fold symmetry would suit this role better, to be compatible with the 12-fold symmetry of the connector. In this scenario, the head vertex would be the rotor stator.

Localization and role of ATPase gp16

The other basic component of the packaging machinery is protein gp16.[1] In our model, the energy necessary for the connector spinning is provided by this ancillary ATPase ATP hydrolase. This protein interacts transiently with the portal during packaging, but it is not present in the final viral particle.[12] The pRNA is required to couple the ATPase activity to packaging.

There are no direct data on the function of gp16 in the packaging machinery, although Simpson et al.[8] provide evidence for its presence in the pRNA domain of portals competent for packaging. We have found that gp16 complexes with portals in gp10-gp16 (BL, DBV, and BL2C, unpublished results). Our results also indicate that gp16 interacts with the connector. The interaction with the isolated connector is dependent on the presence of the carboxy-terminal region, as connectors lacking the 120-terminal amino acid residues cannot interact with gp16. Taken together, these results support the hypothesis that gp16 contacts the narrow part of the

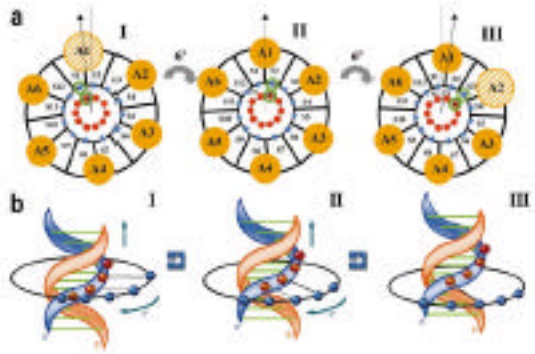

Figure 6. The spinning connector model. (a) Schematic arrangement of a portal vertex in projection view from the viral portal. A1 to A6 represent the ATPase (gp16). S1 to S12 represent the connector subunits, and P1, 2, 3, etc. represent the DNA phosphates of a DNA chain. For clarity, the model is restricted to the description of one single DNA chain and one lysine ring. The starting position (I) aligns a Lys (blue sphere) of subunit S1 with the phosphate P1 lining of the pilot ATPase (A1) rotates the connector 30° clockwise while the DNA moves linearly along the connector axis by 3 bps. A transient interaction among the corresponding Lys from S1 and the phosphate P3 (II) occurs at 0° rotation. The final stage (III) is characterized by the interaction of Lys S1 and the phosphate P5, which is accomplished after a further clockwise rotation of the connector by 30°. This movement positions the next pilot ATPase (A2) in an equivalent position with respect to the DNA, in that of A1 in step I. (b) Schematic representation of a side view of the model, corresponding to the same steps as in (a). The lysine residues are depicted as blue spheres (labelled K1, K2, ...). In step I there is a strong interaction between lysine K1 and phosphate P1 (continuous line), while the interaction of lysine K2 with phosphate P2 is still weak (because P2 is not a lower plane than the ring of lysine residues (broken line). In step II the 30° rotation of the connector places lysine K2 in an optimal position to interact with phosphate P3, provoked there is a longitudinal displacement of the DNA by 3 bps. In step III there has been a further 30° rotation of the connector and another 3 bps displacement of the DNA, implying that lysine K3 is now aligned with phosphate P5. The movement of two base-pairs of DNA, from I to III in (b), is correlated with a rotation of the connector by 120°.

connector, which remains accessible from outside the prohead. As the pRNA is arranged in a ring around the outer end of the dodecameric connector, this system comprises a complex set of concentric structures with different orientations the DNA inside the channel (Ile-tRNA), the connector walls (Ile-tRNA), and the outer ring of pRNA coupled to gp16 (most probably, 6-fold). As the gamma ring, the other side of the connector (Ile-tRNA) is embedded in the 5-fold vertex of the prohead.

At present there are insufficient data to confirm a direct interaction of gp16 with the DNA. Therefore, we cannot discern whether the ATPase motor pushes the DNA directly while the connector rotates passively, or whether the motor only drives the connector rotation. In any case, as explained above, the connector would have to rotate at least 60° until the next base-pair establishes an electrostatic interaction with the following connector subunit. If the pRNA-gp16 complex were a homomer, the connector would have to rotate twice 60° (Ile. 12) to place the gp16/pRNA into an equivalent position with respect to the DNA. This would combine the three main factors in the model: the linear translocation of two DNA base-pairs, a clockwise rotation of 120° of the connector, and the activation of one gp16 ATPase. Each packaging step would then involve two base-pairs of DNA, two gp16 subunits and one ATPase/pRNA subunit. Every ATPase would then be sequentially used, followed by the clockwise rotation of the connector. According to our model, the activation of one ATPase would induce the translocation of two base-pairs into the prohead, the rotation of the connector by 120°, and the pinning of the next ATPase in an equivalent position with respect to the DNA. Even if the connector rotates passively to the DNA translocation, its rotation would control the sequential firing of the ATPase.

The hydrolysis of one ATP molecule to package two base-pairs fits with the experimental data obtained for packaging λ29 DNA and other viruses in vitro.[5,6,8,9] In fact, this ATP consumption is slightly lower than that associated with the walking of helicases along the DNA, just to two ATP molecules per base-pair, but in the latter case these proteins separate DNA strands besides carrying it. However, despite the structural and mechanical differences mentioned above, the translocation of DNA by motors and helicases or sliding clamps could have its energy the hydrolysis of ATP as a power stroke to progressively move the DNA in a confined channel contained by the overall topology of these protein assemblies.

Other alternatives for the packaging of DNA into the viral heads could involve the rotation of the DNA and the connector (reviewed in reference[3]). Yet, any model involving rotation of the DNA would be precluded by the additional torsion generated during its packaging into the prohead container. The sum of the strain imposed

by the coiling of the DNA within the head, following the inner wall of the protein, together with the twisting forces induced by the insertion of DNA during packaging would lead to irreversible strain in the DNA, frustrating both the viral head packaging and partial strand separation, and, eventually, in the arrest of the packaging process.

Further studies on the structure of the components of the various components of the packaging machinery and the connector will be required to provide deeper insight into the molecular events involved in the processive translocation of DNA_{ss} but the present high-resolution structure already suggests a simple and elegant mechanism that can now be validated with other complementary techniques.

Materials and Methods

A detailed description of the protein expression, production, connector particle preparation, crystallization and preliminary crystallographic analysis has been published elsewhere.[20] A new crystal form was obtained using the vapour diffusion method, at 20 °C, from sitting drops containing 20% MPD (Sigma, Spectrum Kit), 0.1 M Tris–HCl (pH 8.0). Typically, crystals appear after five days and reach a maximum size of approx. 0.9 mm × 0.1 mm × 0.1 mm. These crystals were picked up with a cryoloop and directly flash-cryocooled by immersion in liquid propane. Native diffraction data were collected using synchrotron radiation at the ESRF ID14-B BI beamline in Grenoble, France. The crystals diffract to 3.9 Å resolution, although the very high-resolution shell always equally dilute reflecting one frame. Complete data were collected up to 3.1 Å resolution (Table 1). Crystals are tetragonal, with space group $I422$, and large unit cell dimensions of $a = b = 199.7$ Å and $c = 188.4$ Å. The calculated V_m value[21] is 2.8 Å³/Da, assuming one quarter of a connector particle per asymmetric unit. Data were evaluated with MOSFLM[22] and SCALA, of the CCP4 suite.[23] All attempts to obtain suitable heavy-atom derivatives were unsuccessful. Crystals obtained from SeMet preparations did not diffract appreciably either.

A reflection function, calculated with GRIP[24] showed a 3-fold axis parallel to the crystal c-axis. This and packing considerations suggested that the particle was located so that its centre of mass was approximately at $z = 0$, $y = 0$, $x = 0.25$. A starting model for phasing was obtained from the phases corresponding to a cylindrical shape, positioned as mentioned above and adjusted by two-dimensional correlation fitting, using the 6Å reflections between 40 and 15 Å from programs. A phase extension protocol was carried out using DM[25] as follows. The phase polynomial started from a ring calculated at 8 Å and the successive steps were of 0.1 Å. The DM-generated values for the figure-of-merit (FOM) were not used. Instead, FOM-values were set, at each extension cycle, to a low value in the high-resolution range, in order to avoid an overestimation of the likelihood. The typical values for the manually set FOM ranged between 0.6 and 0.9 for unbiased resolution shells and 0.1 to 0.2 for the new shell extensions. The mask used for the phase extension was made initially from a 20° cylindrical sector of the DM model. This mask was manually edited four programs after three or four cycles of phase extension according to the maps. The extension was continued in this way until 3.1 Å. Although the global shape of the molecular connector was identifiable in the resulting maps, no clear secondary structure elements could be located. This map, however, was used to build a pseudo-atomic model with MODELLA[26] filling the connector core. This model was then docked in three 20° cylindrical sectors, which were then separately used, resulting in an averaged pseudo-atomic model (four programs) that was used both to generate a new mask and to further improve the phases with ARP/WARP[27] using the whole resolution of the data to 3.1 Å. Each cycle consisted of an unrestrained water mask and refinement using REFMAC. After each ARP/WARP cycle the model was reconstructed from an averaged pseudo-atomic model as before.

As in this protocol the FOM-values and the crystallographic R-factors are unreliable, checks were made between cycles by varying different combinations of FOM-values, looking for features in maps, and mainly checking the correlation between the cylindrical sectors of the maps. This protocol was successful, and after a few more cycles of well-connected electron density appeared, with several alpha helices visible.

At this stage, the coordinates of the Simpson et al. model (PDB code 1fmi) were the best binary available in the PDB. They were used as an initial model for molecular replacement with AMoRe.[28] No solution, however, could be obtained using a monomer as a starting model. A

Table 1. Data collection and refinement statistics

Data collection	
Wavelength (Å)	[illegible]
Temperature (K)	[illegible]
Resolution range (Å)	[illegible]
No. of measured reflections	[illegible]
No. of unique reflections	[illegible]
Completeness (%)	[illegible]
R_{sym}[a]	[illegible]
R_{cryst}[b]	[illegible]
Multiplicity	[illegible]
Refinement	
Resolution range used (Å)	[illegible]
No. of reflections in working set	[illegible]
No. of reflections in test set	[illegible]
No. of protein atoms in the asymmetric unit (kDa)	[illegible]
No. of solvent molecules in the asu.	[illegible]
No. of MPD molecules in the asu.	[illegible]
ΔB ($Å^2$)[c]	[illegible]
Root deviations from ideal values	
Bond lengths	[illegible] Å
Bond angles	[illegible] degrees
Average B factors ($Å^2$)	
Protein atoms	[illegible]
Water molecules	[illegible]
All atoms	[illegible]
Wilson plot	[illegible]

The maximum resolution shell (3.2–3.1 Å) values are shown in parentheses.

[a] $R_{sym} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where I_i is the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is its mean value.

[b] $R = \sum_{hkl}||F_{obs}| - |F_{calc}|| / \sum_{hkl}|F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively.

[c] ΔB_{rms} = B-value calculated for 95% of reflections not used for refinement.

different strategy was then adopted: a whole asymmetric model, with 12 monomers, was fitted into the density of the phase extension/WARP map. A cross-correlation translation search was then carried out using the model, followed by rigid-body refinement. This gave a solution with an R-factor of 41% and a correlation of 31% (next peak: 27% and 20%, respectively). Sigma-A-weighted $2F_o - F_c$ and $F_o - F_c$ difference density maps calculated with this solution showed that it was correct, although in many areas, and in particular in the peripheral α-helix part, the density did not coincide with the model. These areas were manually rebuilt with TURBO[19] including the fitting of the open β-strands and the correct placing of the amino and carboxyl ends, which in the initial model was one or more residues away. The structure was refined with X-PLOR[20] and further manual interventions. Finally, 240 NCS restraints were applied. These were relaxed and finally superimposed, over the 20Å to 3.8Å range, at full resolution. All this steps, water molecules were added to the model. Atoms outside the $2F_o - F_c$ and $F_o - F_c$ sigma-A-weighted maps were left in one conformation. This mainly affected the N- and C-terminal residues and some residues before β10. The sequence of 10 amino acid residues between residues 208 and 218 of each monomer was not seen on the maps and was not included, as explained in Results. The final model is of excellent geometry as measured by PROCHECK.[21] The final refinement statistics are summarized in Table 1.

Atomic coordinates

The atomic coordinates have been deposited with the PDB under accession number 1AHV.

Acknowledgements

This study was supported by grants from the Ministerio de Educación y Cultura of Spain (PB95-1031 and 2FD97-0918 to M.G., PB95-0229-C02 to EM.V., and PB96-0918 to B.F.) and the Generalitat de Catalunya (1999SGR-00 and Centre de Bioinformàtica en Biotecnologia to M.G. Synchrotron data collection was supported by the EU grants BIO4-CT98-0311 and BIO4-CT-1999-0017 to the EMBL-EMBY and by the DESY. B.V. acknowledges a fellowship from the Ministerio de Educación y Cultura. We are indebted to Rosa Pérez-Luque for assistance in crystallization experiments and Ignacio Ellis and Jaime Martín-Benito for useful suggestions.

References

1. Bhuyan, M. M. & O'Donnell, M. (1998). Topoisomerase nicking: steps around DNA. Curr. Biol. 8, R814-816.
2. Gaspar, S. de Haenen, R. W. (1998). Central conformation in dsDNA bacteriophage assembly. In The Bacteriophages (Calendar, R., ed), vol. 1, pp. 19-40, Plenum Press, New York.
3. Valpuesta, J. M. & Carrascosa, J. L. (1994). Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. Quart. Rev. Biophys. 27, 107-123.
4. Guo, P. X., Erickson, S. & Anderson, D. (1987). A small viral RNA is required for in vitro packaging of bacteriophage phi 29 DNA. Science, 236, 690-694.
5. Valpuesta, J. M., Fernandez, J. J., Carazo, J. M. & Carrascosa, J. L. (1998). The three-dimensional structure of a DNA translocating motor at 30 Å resolution. Struct. Fold. Des. 7, 539-546.
6. Ian, Y., Olfatz, N. H., Nix, W., Anderson, D. L., Baumann, M. G. & Baker, T. S. (1998). Assembly of a tailed bacterial virus and its genome release studied in three dimensions. Cell, 95, 171-182.
7. Carrascosa, J. L. & Valpuesta, J. M. (1999). Bacteriophage connectors: structural features of a DNA translocating motor. Trends Biochem. Sci. 24, 483-488.
8. Valpuesta, J. M., Sousa, N., Bartolomé, B., Fernandez, J., Fujisawa, H., Carazo, J. & Carrascosa, J. L. (1998). Structural analysis of the bacteriophage T3 head-to-tail connector. J. Struct. Biol. 123, 116-125.
9. Olveira, S. V., Olveira, S., Bartolomé, B., Alonso, J., Rodríguez, J., Tenallony, T. A. et al. (1999). Structure of the T3-tail symmetric portal protein of bacteriophage SPP1. Virology Struct. Biol. 6, 1230-1236.
10. Carrascosa, J. L., Carrascosa, J. L. & Mendiola, M. (1994). Bacteriophage baseplate parameters. Purification and properties. J. Mol. Biol. 241, 123-137.
11. Bazinet, C., Bertrand, J., Bing, J., Carazo, J. M. & Carrascosa, J. L. (1988). Purification and organization of the gene 12 portal protein required for phage T3 DNA delivery. EMBO J. 7, 1849-1856.
12. Dobrodumov, K. A., Rippel, A., ten Heggeler, B. & van Etten, J. L. (1998). Gene 20 product of bacteriophage T4: its purification and structure. J. Mol. Biol. 280, 519-533.
13. Carrascosa, J. L., Carazo, J. M., Rocha, E. de Santisteban, A. (1983). Structure of phage phi 29 connector protein assembled in vitro. Virology, 131, 196-203.
14. Carrascosa, J. L., Mirabella, G., Garcia, N. de Santisteban, A. (1982). Structure of the bacteriophage connector of bacteriophage phi 29. J. Mol. Biol. 154, 519-534.
15. Carazo, J. M., Donate, L. G., Herranz, L., Stella, J. L. & Carrascosa, J. L. (1986). Three-dimensional reconstruction of the connector of bacteriophage phi 29 at 35 nm resolution. J. Mol. Biol. 192, 933-947.
16. Dube, D., Tavares, P., Lurz, R. & van Heel, M. (1993). The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. EMBO J. 12, 1303-1309.
17. Grimes, S., Jardine, P., Anderson, D., Valpuesta, J. M., Carrascosa, J. L. & Cerritelli, M. (1998). Purification, crystallization and preliminary X-ray diffraction studies of the bacteriophage phi29 connector particle. Acta Cryst. D54, 1315-1317.
18. Valpuesta, J. M., Carrascosa, J. L. & Henderson, R. (1994). Analysis of electron microscope images and electron diffraction patterns of thin crystals of phi29 connectors in ice. J. Mol. Biol. 240, 281-287.
19. Mallet, F., Rippel, A., Carrascosa, J. L. & Vellieux, M. (1997). The bacteriophage phi29 head-tail connector imaged at high resolution with the cryo-electron microscope in Saclay laboratory. Acta Cryst. D 53, 224-233.
20. Grimes, S., Jardine, P., Anderson, D., Valpuesta, J. M., Carrascosa, J. L. & Cerritelli, M. (1998). Crystallographic analysis reveals the 12-fold symmetry of the bacteriophage phi29 connector particle. J. Mol. Biol. 284, 239-251.
21. Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J. et al. (2000). Structure of

the bacteriophage φ29 DNA packaging motor. *Nature*, **388**, 793–796.

21. Simon, L. D., Valpuesta, J. M., Muñoz, A., Mínguez, E., Bravo, R., Salas, M. & Carrascosa, J. L. (1998). Role of the amino-terminal domain of bacteriophage φ29 connector in DNA binding and packaging. *J. Mol. Chem.* **273**, 10918–10924.

22. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.

23. Valpuesta, M. A., Bravo, R., Carrascosa, J. L. & Valpuesta, J. M. (1994). An immunocytochemical study of phage φ29 connector/nucleic acid interactions. *Eur. J. Biochem.* **220**, 787–793.

24. Cuiyun, S., van Vliet, F., Billy, E., Bras, P. A. & Stevenson, R. (1997). Homology throughout the multigene family encoding capsid proteins of bacteriophage φ29. *J. Virol.* **71**, 7317–7327.

25. Wikoff, W. R. & Johnson, J. E. (1999). Virus assembly: imaging a molecular machine. *Curr. Biol.* **9**, 276–278.

27. Black, L. W. (1988). DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.* **43**, 267–292.

28. Kong, X. P., Cornu, D., Gottesman, M. & Hanson, J. (1992). Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell*, **69**, 425–437.

29. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C. & Ellenberger, T. (1999). Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell*, **99**, 167–177.

30. Ban, N., Begley, D. R. & Harpp, J. M. (1999). Crystal structure of the N-terminal domain of the DnaB hexameric helicase. *Struct. Fold. Des.* **7**, 891–899.

31. Gonzalez-Huici, B. X., Menendez, C., Hernandez-Lagos, R., Gonzalez, J., Chinea, G., de la Cruz, F. & Goff, M. (1999). The bacterial conjugation protein TrwB assembles into hexamers and DNA-ATPase. *Biochem.*, **38**, 8274–8281.

32. Sabatte, P., Rajagopal, H. & Mengaud, T. (1997). Characterization of the bacteriophage T3 DNA packaging machinery in vitro in a cell-free system. *J. Mol. Biol.* **266**, 815–831.

33. Calendar, G. R., Gay, D. & Poteu, M. (1999). Virus DNA packaging: the strategy used for phage lambda. *Mol. Microbiol.* **32**, 875–886.

34. Bravo, R., Carrascosa, J. L., Elorza, G., Valle, M., Valpuesta, J. & Carrascosa, J. L. (1998). Topology of the components of the DNA packaging machinery in the phage φ29 prohead. *J. Mol. Biol.* **280**, 837–851.

35. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase similarity to other polynucleotidyl transferases. *Science*, **266**, 1981–1986.

36. Hendrix, R. W. (1998). Bacteriophage DNA packaging: RNA gears in a DNA transport machine. *Cell*, **94**, 147–150.

37. Hendrix, R. W. (1978). Symmetry mismatch and DNA packaging in large DNA bacteriophages. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4779–4783.

38. Carazo, J. M., Santisteban, A. & Carrascosa, J. L. (1983). Three-dimensional reconstruction of bacteriophage φ29 neck particles at 2·5 nm resolution. *J. Mol. Biol.* **169**, 779–798.

39. Rivera, L. E., Fernandez, L., Sevilla, J. P., Carazo, J. M., Valpuesta, J. & Carrascosa, J. L. (1988). Bacteriophage T3 connector three-dimensional structure and comparison with other small dodecamer connecting regions. *J. Mol. Biol.* **203**, 99–108.

40. Valpuesta, J. M., Serrano, M., Chinea, G., Carrascosa, J. L. & Carrascosa, J. L. (1992). DNA-induced change induced by the bacteriophage φ29 connector. *Nucl. Acids Res.* **20**, 5983–5984.

41. Guo, P., Zhang, C., Chen, C., Garver, K. & Trottier, G. (1998). Inter-RNA interaction of phage φ29 pRNA to form a hexameric complex for viral RNA transportation. *Mol. Cell.* **2**, 149–155.

42. Zhang, R., Lewinter, S., Wu, X., Shu-Jennel, H., McMurray, C. T. & Anderson, D. (1998). Function of hexameric RNA in packaging of bacteriophage φ29 DNA in vitro. *Mol. Cell.* **2**, 149–155.

43. Hermoso, M. A., Philip, R. B. & Anderson, D. (1992). Morphogenesis of bacteriophage φ29 of *Bacillus subtilis* examined and quantified in vitro packaging of DNA gene gp8. *J. Virol.* **66**, 3024–3034.

44. Villafane, R., Carrascosa, A., Jimenez, E., Carrascosa, J. L., Martinez, C. & Salas, M. (1994). Structure and assembly of phage φ29. *Philo. Trans. Phys. Sci. Lond. A, Math. Sci.* **348**, 29–33.

45. Chen, E., Freimann, C. & Anderson, D. (1987). Structural and RNA-gp8-gp10-gp11 ATPase activity of the DNA packaging protein gp16 of bacteriophage φ29. *J. Mol. Biol.* **197**, 239–246.

46. Matsuo, M., Tanaka, M. & Fujisawa, H. (1998). DNA packaging ATPase of bacteriophage T3. *Virology*, **242**, 276–283.

47. Malkowski, R. W. (1998). Silicon control of protein crystals. *J. Mol. Biol.* **32**, 474–487.

48. Leslie, A. G. W. (1991). Macromolecular data processing. In *Crystallographic Computing* (Moras, D., Podjarny, A. D. & Thierry, J. C., eds), vol. 5, pp. 27–38, Oxford University Press, Oxford.

49. CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. sect. D*, **50**, 760–763.

50. Terwilliger, T. (1998). RESOLVE, a suite of computer programs for molecular-replacement calculations. *J. Appl. Crystallogr.* **31**, 744–747.

51. Klepeis, G. J. & Floudas, C. A. (1999). Software for handling macromolecular envelopes. *Acta Crystallogr. sect. D*, **55**, 1456–1464.

52. Brunger, A., Moore, R. & Lynn, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 786–791.

53. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. sect. A*, **50**, 157–163.

54. Brunell, A. & Gottschalk, G. (1998). Tutorial-Guide to Silicon Graphics Geometry Partners Director, pp. 27–29, Silicon Graphics, Mountain View, CA.

55. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. sect. D*, **54**, 905–921.

56. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.

57. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Edited by R. Huber

(Received 17 September 2001; revised in revised form 13 November 2001; accepted 16 November 2001)