

Chapter 23

The dsDNA Packaging Motor in Bacteriophage ϕ 29

Marc C. Morais

Abstract The tailed dsDNA bacteriophage ϕ 29 packages its 19.3-kb genome into a preassembled prolate icosahedral procapsid structure using a phage-encoded macromolecular motor. This process is remarkable considering that compaction of DNA to near crystalline densities within the confined space of the capsid requires that the motor work against considerable entropic, enthalpic, and DNA bending energies. The heart of the bacteriophage ϕ 29 packaging motor consists of three macromolecular components: the connector protein, an RNA molecule known as the pRNA, and an ATPase. The pRNA is thus far unique to ϕ 29, but the connector and ATPase are homologous to portal and terminase proteins, respectively, in other tailed dsDNA bacteriophages. Despite decades of effort and a wealth of genetic, biochemical, biophysical, structural, and single particle data, the mechanism of DNA packaging in bacteriophage ϕ 29 remains elusive. In this chapter, we describe the development of a highly efficient in vitro DNA packaging system for ϕ 29, review the data available for each individual macromolecular component in the packaging motor, and present and evaluate various packaging mechanisms that have been proposed to explain the available data.

23.1 Introduction

The ability to interconvert various forms of energy is an essential feature of living systems. Biological molecular motors accomplish this task by coupling the making and breaking of high-energy covalent bonds to conformational changes in large macromolecules. These conformational changes in turn drive a diverse array of biological phenomena, including muscle contraction, protein synthesis, nucleic acid manipulation, and the generation and maintenance of concentration gradients and electrostatic potentials. Understanding how molecular motors interconvert energy at the atomic scale provides insight into the fundamental question of how energy is harnessed in biological systems. Furthermore, characterizing the assembly pathways of molecular motors illuminates the mechanisms by which energy transduction systems are constructed and maintained in nature. There are only a few well-characterized biological molecular motors. These generally fall into two classes (1) rotary motors, such as the F_1 ATPase and bacterial flagellar motor, which turn a spindle by a rotary

M.C. Morais (✉)

Department of Biochemistry and Molecular Biology, Sealy Center for Structural and Computational Biology, University of Texas Medical Branch at Galveston, Galveston, TX 77555, USA
e-mail: mcmorais@utmb.edu

conformational change mechanism and (2) linear motors, like those required for muscle contraction, where conformational fluctuations result in translation of the motor with each step. This chapter describes a third, hybrid class of molecular motors, viral DNA packaging motors, where rotary events are converted into translational movement. In this type of motor, several energy-generating ATPase subunits are arranged as a ring, and coordinated hydrolysis of ATP molecules around this ring induces a rotary pattern of conformational changes in the motor that are coupled to the translocation of double-stranded DNA (dsDNA) into preformed viral capsids.

The process of genome encapsidation is remarkable considering that the length of DNA to be packaged is approximately 150 times the width of the capsid itself (Earnshaw and Casjens 1980). Hence, considerable entropic, electrostatic, and bending energies must be overcome to package DNA to near-crystalline densities within the confines of the virus shell and against pressures estimated to exceed 6 Mpa (Smith et al. 2001). Indeed, viral DNA packaging motors are among the most powerful biological motors known, capable of generating forces up to nearly 100 pN (Smith et al. 2001; Fuller et al. 2007; Rickgauer et al. 2008). Furthermore, there is the nontrivial topological requirement that packaged DNA must also exit freely during genome ejection; knots and tangles cannot be tolerated since the portal for DNA entrance and exit is only large enough to accommodate passage of a single dsDNA helix. It is thus not surprising that early hypotheses regarding genome packaging proposed that the viral capsid must be assembled around condensed DNA (Kellenberger et al. 1959).

The virally encoded ATPases that power genome encapsidation are a subgroup of the large ASCE (Additional Strand Catalytic *E* (glutamate)) ATPase family whose members are involved in various macromolecular remodeling tasks including: cell division; chromosome segregation; DNA recombination, strand separation, and conjugation; and the generation and maintenance of concentration gradients and electrostatic potentials (Mitchell et al. 2002; Burroughs et al. 2007; Singleton et al. 2007; Thomsen and Berger 2008). Thus, understanding the mechanisms by which these motors operate will elucidate the general principles of molecular partitioning and energy transformation in biological systems. Furthermore, although the oligomeric rings of virally encoded ATPases are similar to the ubiquitous ASCE macromolecular translocases, they are capable of generating considerably greater forces than other ASCE ATPase-based motors. Hence, insights gained from the study of viral packaging motors will not only shed light on the basic mechanistic principles of a broad class of macromolecular motors but can also illuminate how these principles have been adapted by viruses to generate and control the large molecular forces necessary for genome encapsidation.

Research on virus DNA packaging motors have converged on a few probable motor mechanisms that differ primarily in (1) the identity of the force-generating component, (2) the mechanics of the force-generating step, (3) the mode of coordination between motor components, and (4) the nature of substeps in one complete mechanical cycle of the motor. This chapter will examine genome packaging in dsDNA viruses based on available structural, biophysical, computational, genetic, and biochemical data that describe the self-assembly and mechanism of viral DNA packaging motors. These aspects of motor function will be examined in the bacteriophage ϕ 29 DNA packaging motor, an outstanding model system for studying genome packaging due to its compositional simplicity, well-defined genetics, high in vitro packaging efficiency, and general experimental accessibility (Grimes et al. 2002).

A comprehensive understanding of the ϕ 29 dsDNA packaging motor requires examining its mechanism at three levels (1) the chemical level, i.e., the mechanism of catalysis employed by the ATPase; (2) the mechanical level, i.e., the conformational changes of motor components that drive DNA translocation; and (3) the mechanochemical level, i.e., how the catalytic cycle of the ATPase is linked to coordinated mechanical movement of the various motor components. In addition to inspiring rationally designed molecular motors and illuminating basic biological questions regarding (1) molecular recognition and self-assembly; (2) coordinated protein–protein and protein–nucleic acid interactions; (3) nucleic acid translocation and compaction; and (4) conversion of chemical energy into mechanical motion, research on the ϕ 29 packaging motor has the potential to inform rational design of therapeutics that inhibit genome packaging motors in medically important eukaryotic viruses such as herpesviruses, adenoviruses, and poxviruses.

23.2 Overview of Bacteriophage $\phi 29$

23.2.1 The $\phi 29$ Life Cycle and Assembly Pathways

Bacteriophage $\phi 29$ is a dsDNA bacteriophage belonging to the *Podoviridae* family. The mature phage consists of a prolate, $T=3$, $Q=5$ prolate icosahedral head elongated along a fivefold axis of symmetry and measuring 540 Å long by 450 Å wide with a shell thickness of approximately 16 Å. Attached to the head is a pseudo sixfold symmetric, 380-Å long tail (Tao et al. 1998; Morais et al. 2005; Xiang et al. 2006). The phage recognizes and binds to its *Bacillus subtilis* host by virtue of this short noncontractile tail. Hydrolytic enzymes incorporated into the tail structure help $\phi 29$ degrade and penetrate the approximately 250-Å thick external peptidoglycan cell wall and membrane of the Gram-positive *Bacillus* host (Xiang et al. 2008; Cohen et al. 2009). The phage then ejects its genome through the tail and into the bacterial cytoplasm where the genome is replicated, and the phage-encoded structural and nonstructural proteins are expressed, thus setting the stage for assembly of progeny $\phi 29$ virions (Fig. 23.1).

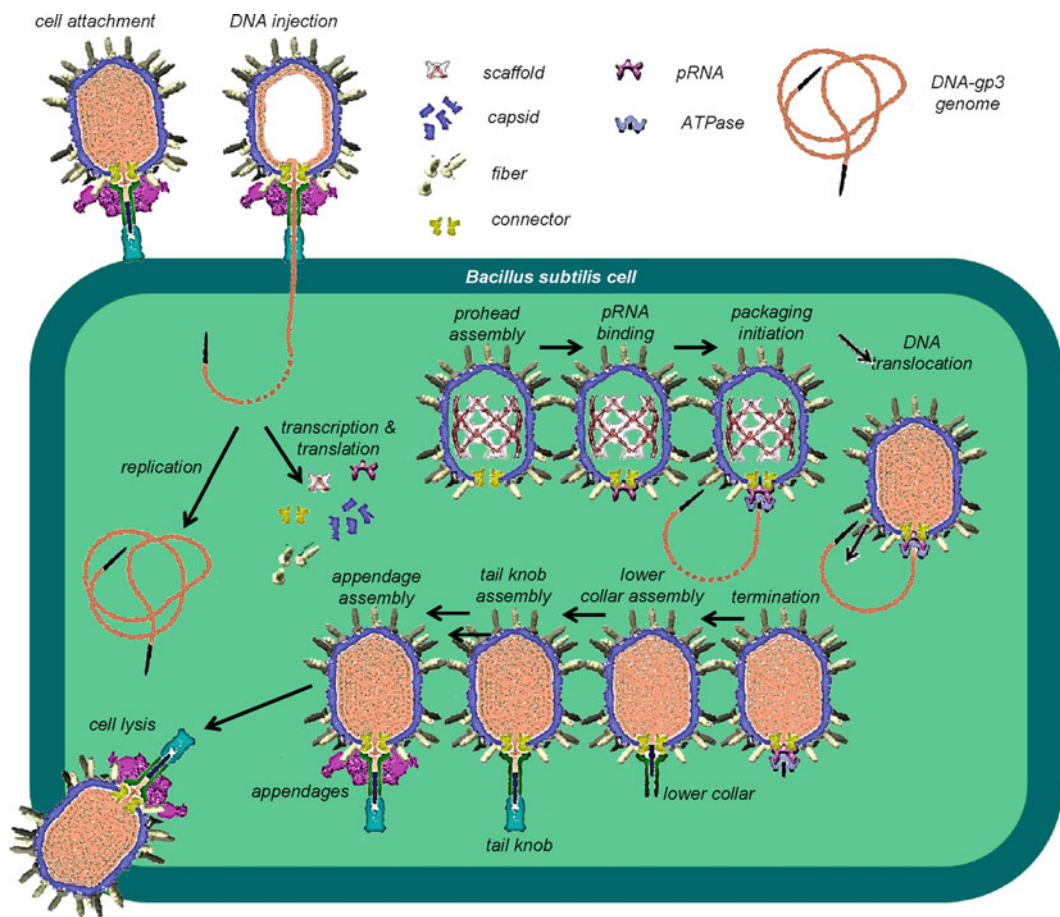


Fig. 23.1 Life cycle and assembly pathway of bacteriophage $\phi 29$. Sequential steps in the bacteriophage $\phi 29$ life cycle and assembly pathway are shown as a cartoon. A *Bacillus subtilis* host cell is represented as a green rectangle, various particles along the assembly pathway are shown as cross sections of cryoEM surface renderings, and individual phage components are either labeled directly or can be identified via the graphical key above the *B. subtilis* cell. Different steps in the assembly pathway are labeled accordingly, and the order of events indicated by black arrows

The assembly of $\phi 29$ particles occurs by a well-ordered pathway (Anderson and Reilly 1993; Tao et al. 1998; Grimes et al. 2002). Proheads, the first particles assembled during morphogenesis, consist of a dodecameric head–tail connector protein, or gene product 10 (gp10), approximately 150 copies of the scaffolding protein (gp7), 235 copies of the major capsid protein (gp8), 55 trimeric head fiber proteins (gp8.5), and a pentameric phage-encoded 174 base RNA (pRNA). The 19.3-kb phage genome, covalently attached at either end to gene product 3 (DNA-gp3), is then threaded through a central channel of the connector and encapsidated within the prohead. Genome packaging is powered by a phage-encoded ATPase (gp16), analogous to the large terminase subunit in other phage systems, which transiently assembles at the connector/pRNA vertex. The scaffolding protein exits as DNA enters the shell, although the route and mechanism of exit remain unclear (Morais et al. 2003). Unlike many other tailed dsDNA bacteriophages, $\phi 29$ packages a unit-length genome and does not undergo significant structural changes during genome packaging (Xiang et al. 2006; Tang et al. 2008). Upon completion of packaging, the ATPase and pRNA detach from the head, and the lower collar (gp11), appendages (gp12), and tail knob (gp9) subsequently attach to form the phage tail and thus complete virus maturation.

23.2.2 $\phi 29$ as a Model System for Genome Packaging

There are several advantages inherent to the $\phi 29$ system for studying genome packaging. First, it is relatively easy to produce large quantities of $\phi 29$ particles. The burst size of wild-type $\phi 29$ is greater than 1,000, and mutant phage infections can produce up to 5,000 proheads per cell (Grimes et al. 2002). In contrast, phage lambda produces only a few hundred proheads per cell. Additionally, $\phi 29$ is one of the smallest and simplest tailed dsDNA bacteriophages, and its 19.3-kb genome encodes only about 20 genes. Hence, the small size of the $\phi 29$ genome affords the opportunity to uncover the structure and function of every gene. For comparison, consider that phage T4's 170,000-kb genome codes for approximately 200 genes. Similarly, the availability of conditional lethal mutants in every gene makes $\phi 29$ easy to manipulate and adapt to emerging technologies. Furthermore, the $\phi 29$ connector and ATPase components are small (~60% the size of other phages), suggesting that they represent the essential minimum for motor operation. Perhaps most importantly, an extraordinarily efficient in vitro packaging system has been developed for $\phi 29$ (Guo et al. 1986) in which nearly every prohead is active and every DNA can be packaged. Thus, compositional simplicity, combined with a large burst size and the high efficiency of in vitro packaging make the $\phi 29$ system amenable to complete dissection of the motor mechanism.

DNA packaging, defined as production of DNA filled heads, can be initiated by mixing purified proheads, DNA-gp3, recombinantly expressed ATPase, and ATP. The extent of packaging can be determined by assaying for head filling via DNase protection; addition of DNase I will result in digestion of unpackaged DNA (Guo et al. 1986), whereas packaged DNA is protected in the viral capsid. Hence, the efficiency of packaging can be determined by extracting the packaged DNA and quantifying it on an agarose gel. It was shown that packaging takes about 4 min to complete, and that the motor translocates approximately two base pairs of dsDNA per ATP hydrolyzed. By analyzing packaging of restriction digests, it was demonstrated that $\phi 29$ preferentially packages the left end of DNA-gp3, suggesting that the left end is the first end to enter the capsid (Bjornsti et al. 1983, 1984; Grimes and Anderson 1989).

Using fiberless proheads that have ejected their scaffolding proteins, a minimal packaging system can be constructed that consists of only five macromolecular components: (1) the capsid protein, (2) the connector, (3) the pRNA, (4) the ATPase, and (5) the DNA-gp3. The relative arrangement of these components in $\phi 29$ particles (Fig. 23.2) was initially proposed from cryoEM reconstructions of wild-type $\phi 29$ proheads (Tao et al. 1998; Simpson et al. 2000; Morais et al. 2001, 2005, 2008).

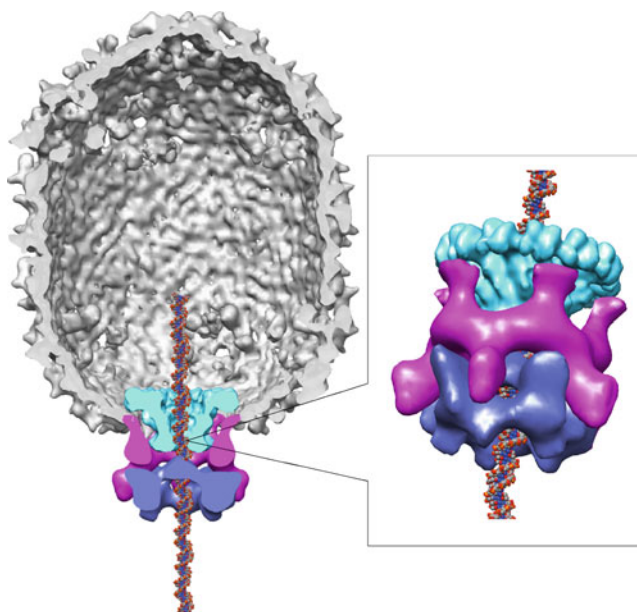


Fig. 23.2 Molecular envelopes of the bacteriophage ϕ 29 motor components. Three-dimensional reconstruction of a ϕ 29 prohead complexed with the packaging ATPase. The particle is shown in a cutaway view, where its front half has been removed to facilitate visualization of the dsDNA packaging motor. The capsid is shown in *gray*, the connector in *green*, the pRNA in *magenta*, and the ATPase in *blue*. The dsDNA genome, shown as *spheres* that have been colored by element, was modeled into the central channel of the connector. The inset shows a close-up of the packaging motor at a skewed angle. Molecular envelopes for various motor components were obtained via cryoEM difference maps (Morais et al. 2008)

The major capsid protein is arranged as a $T=3$, $Q=5$ prolate icosahedral lattice on the surface of the capsid. The dodecameric connector protein replaces a major capsid protein pentamer at a unique vertex situated in an icosahedral endcap along the long axis of the phage. The pRNA also binds this unique vertex, where it circumscribes the protruding narrow end of the connector as a five-membered ring. Five RNA spokes emanate from this central ring, extending away from the phage at a slight angle from the particle's central axis. The packaging ATPase binds the distal end of these pRNA spokes to form a presumably pentameric ring at the unique packaging vertex. These initial observations and predictions were later confirmed and refined in a series of cryoEM reconstructions that definitively located the different components of the packaging motor. Briefly, cryoEM reconstructions were determined for various ϕ 29 particles that were engineered to lack different components of the motor such as the connector, the ATPase, full-length 174-base pRNA, or to lack different domains within the pRNA. The resulting reconstructions were then subtracted from reconstructions of particles containing an intact motor such that the residual density in the difference maps provided molecular envelopes for individual motor components and domains within components (Fig. 23.2) (Morais et al. 2008).

Hence, the connector, pRNA and ATPase form three concentric rings, and the DNA-gp3 is believed to be threaded through a continuous channel along their shared central axis and into the phage capsid. This overall picture of the motor not only provides a general context for understanding the structure and function of individual macromolecular motor components but also provides a framework for understanding how different motor components interact with each other to coordinate the mechanochemical cycle of the motor. Below, the structures and functions of individual motor components are examined in greater detail to illuminate their potential roles in both motor assembly and DNA translocation.

23.3 Components of the ϕ 29 DNA Packaging Motor

23.3.1 *The DNA-gp3*

As mentioned above, the ϕ 29 genome is one of the smallest and simplest dsDNA bacteriophage genomes, consisting of approximately 19.3 kb that code for about 20 genes. The early genes required for initial DNA replication are located at the ends of the genome, whereas the late genes coding for structural and morphogenetic proteins are expressed from the opposite strand in the middle of the genome (Mosharrafa et al. 1970; Schachtele et al. 1973). A distinguishing feature of ϕ 29 is the requirement for a protein primer to initiate DNA replication. As a result, ϕ 29 gene product 3, gp3 is covalently attached to the first adenine of each 5' end of the ϕ 29 genome via gp3 serine (Harding et al. 1978; Salas et al. 1978; Yehle 1978). Limited proteolysis of gp3 with trypsin or protease K reduces packaging two- and threefold, respectively, suggesting that gp3 is also necessary for efficient DNA packaging (Bjornsti et al. 1983; Grimes and Anderson 1989). Similarly, restriction fragments containing gp3 are packaged at least ten times more efficiently than fragments lacking gp3 (Grimes and Anderson 1989).

Available structural and biochemical data regarding genome organization within capsids (Xiang et al. 2006; Tang et al. 2008) indicate that the gp-3 attached to the left end of the genome is likely first to enter the capsid during packaging and last to leave upon ejection. Similarly, gp-3 bound to the right end should be last to enter the shell during encapsidation and first to leave upon genome ejection. The passage of gp3 through the central channel of the connector is remarkable in that the smallest dimension of gp3 is greater than the diameter of the connector channel (Kamtekar et al. 2006; Xiang et al. 2006). Presumably, gp3 must be at least partially unfolded to pass through the narrow channel of the connector. There is some controversy regarding the location of the right-end gp3 in mature phage; one group reports that gp3 is lodged in the central channel of the connector (Xiang et al. 2006), while another group reports that it is located in the central channel of the tail just below the connector (Tang et al. 2008). Regardless of the exact position of the right-end gp3 in mature phage, the ϕ 29 packaging motor must at least recognize and package the left-end gp3 along with the phage DNA, and subsequently inject it into the host cell cytoplasm upon genome ejection to initiate a new replication cycle.

23.3.2 *The Connector*

The ϕ 29 connector protein, analogous to the “portal” proteins in other phage systems, is remarkable in that it plays an essential role in nearly every step of the phage life cycle. In addition to providing a conduit for DNA entrance and egress, the connector nucleates capsid assembly, directs formation of an elongated prolate icosahedral shell (Hagen et al. 1976; Camacho et al. 1977; Guo et al. 1991a), and provides the site of attachment for the virus tail during phage maturation. As a consequence, the connector must selectively interact with the scaffolding protein, the capsid protein, the pRNA, the DNA, the lower collar, and the appendages at different stages of the assembly pathway (Guo et al. 1991a; Simpson et al. 2000; Morais et al. 2001, 2005, 2008; Xiao et al. 2005; Xiang et al. 2006; Atz et al. 2007; Tang et al. 2008; Fu and Prevelige 2009). Given the central importance of connector proteins in every aspect of the phage life cycle, it is not surprising that the structures of phage connector proteins are highly conserved in spite of often undetectable sequence homology (Valpuesta and Carrascosa 1994; Simpson et al. 2000; Lebedev et al. 2007; Zheng et al. 2008; Lander et al. 2009) (Fig. 23.3f). Indeed, cone-shaped connectors were some of the earliest defining morphological features observed in negative stained electron micrographs of tailed dsDNA bacteriophages.

The structure of isolated recombinantly expressed ϕ 29 connectors has been studied by a variety of biophysical methods, including atomic force microscopy (Muller et al. 1997), cryoEM of two-dimensional arrays (Valpuesta et al. 1999), immune-electron microscopy (Valle et al. 1999), and

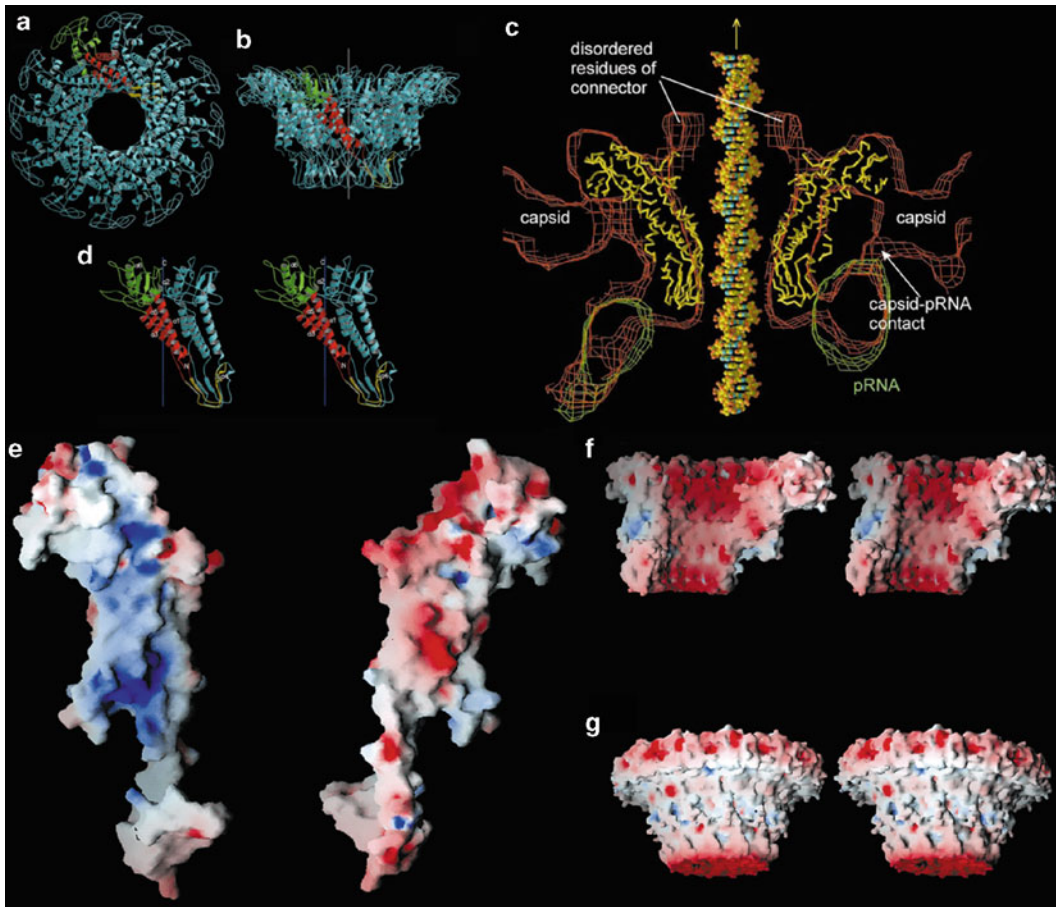


Fig. 23.3 Crystal structure of the bacteriophage ϕ 29 connector protein. (a) Three-dimensional structure of the ϕ 29 shown as a *ribbon diagram*. The direction of view is looking down the 12-fold symmetry axis from the narrow end of the connector. A single subunit is shown with its wide, central, and narrow domains in *green*, *red*, and *yellow*, respectively. (b) Side view of the connector with the central 12-fold symmetry axis shown as a white line. (c) Fit of the X-ray structure of the connector to cryoEM density of the prohead. The backbone of the connector is shown as *yellow wire*; a pRNA density is outlined with *green mesh*. A DNA molecule was modeled in the central channel of the connector. The capsid, one of the five contacts between the pRNA with the capsid, and the partially disordered residues 229–246 and 287–309 in the connector are indicated. (d) Stereodialog of a pair of monomers. One monomer is in *blue* and the other as described in (a). Six α -helices are labeled, along with the N- and C-termini. (e) Electrostatic surface potential of opposite faces of a monomer showing the complementarily charged surfaces that interact between the subunits. (f) Internal electrostatic surface of the connector showing its primarily negative charge. (g) External electrostatic surface showing negatively charged regions at opposite ends of the connector and the relatively hydrophobic region in the center that interacts with the capsid. Parts (a–c) were modified from Simpson et al. (2000); parts (d–f) were modified from Simpson et al. (2001)

X-ray crystallography (Simpson et al. 2000, 2001; Guasch et al. 2002). Although 13-fold symmetric connectors have been reported (Tsuprun et al. 1994), both X-ray crystallography and cryoEM indicate that connectors form 12-fold symmetric dodecamers (Simpson et al. 2000; Morais et al. 2001; Xiang et al. 2006; Tang et al. 2008). The 12 monomers assemble a hollow cone-like structure where the 12-fold symmetry axis is coincident with the conical axis (Fig. 23.3a, b). The connector structure can be further divided into three approximately cylindrical regions that differ in height and radius: the narrow end, central part, and wide end have external radii of 33, 47, and 69 Å and heights of 25, 28, and 23 Å, respectively. Hence, the connector measures ~75 Å along its conical axis, and its width varies from ~140 Å at its wide end to ~65 Å at the narrow end. The diameter of the connector's

central channel increases from approximately 36 Å at the narrow end to about 60 Å at the wide end, large enough to accommodate smooth passage of an extended dsDNA helix. Fitting the X-ray structure of the connector into cryoEM reconstructions of ϕ 29 proheads shows that the wide end of the connector resides within the viral head, the central region is adjacent to the virus shell and the pRNA, and the narrow end protrudes from the head and interacts with the central ring of the pRNA (Simpson et al. 2000; Morais et al. 2005, 2008; Ding et al. 2011) (Fig. 23.3c).

The 309 residue connector monomer resembles a seahorse, where the wide domain corresponds to the head; the central domain, the body; and the narrow domain, the tail (Fig. 23.3d). Residues 158–202 comprise the narrow-end domain, forming mostly an antiparallel β -sheet that interacts with the central ring of the pRNA. The central domain (residues 11–14, 130–157, and 203–226) consists of an antiparallel three-helix bundle twisted in a left-handed superhelix around the central 12-fold axis at an angle of approximately 40°. The wide-end domain (residues 1–10, 229–246, and 286–309) adopts a more globular mixed α/β -structure that is flexibly attached to the central domain, and contacts the inner surface of the head and the pRNA (Simpson et al. 2000; Ding et al. 2011) (Fig. 23.3c). Residues 1–10, 229–246, and 286–309 are not visible in the crystal structure, suggesting that these regions are disordered in recombinantly expressed connectors. However, the second and third disordered regions are visible in cryoEM maps where they are located on the inside of the channel at, and just above, the junction of the wide and narrow domains, respectively, consistent with the locations of analogous elements in the connector proteins from bacteriophages SPP1 and P22 (Orlova et al. 2003; Lebedev et al. 2007; Lander et al. 2009) (Fig. 23.3c).

Calculated electrostatic potentials show that the connector monomer has one predominantly negatively charged surface and a complementary net positively charged surface, promoting the circular interactions necessary for formation of a closed ring-like structure (Fig. 23.3e). The exterior surface of the central region is predominantly hydrophobic, whereas the exterior surfaces of the wide- and narrow ends have negatively charged patches (Fig. 23.3f, g). With the exception of two lysine rings, the surface corresponding to the central channel is negatively charged. It has been proposed that the net negative charge of this surface acts as a focus to center the opposing negatively charged DNA as it passes through the connector channel (Simpson et al. 2000), and that the positively charged lysine rings play a direct role in packaging by interacting with the negatively charged phosphate backbone of the translocating DNA (Guasch et al. 2002) (see also Sect. 23.5).

Comparison of connector structures fitted into cryoEM reconstructions of mature ϕ 29, wild-type proheads, and particles lacking pRNA indicates that the conformation of the connector is strongly influenced by the attachment/detachment of the pRNA and the tail structure (Fig. 23.4a–c). For example, in the absence of pRNA, the outer and inner diameter of the central region and narrow end increases compared to connectors in particles bound to pRNA (Morais et al. 2008). The narrow domain of the connector also flares out in mature phage. It has been suggested that connector expansion occurs in response to high pressure in fully packaged proheads, and that this expansion forces the pRNA off the capsid and results in an optimal connector conformation for tail attachment (Xiang et al. 2006). However, it is not clear if this conformational change occurs prior to tail assembly, or if tail binding induces conformational change in the connector.

Additional information regarding conformational flexibility of the connector can be obtained by comparing connector structures from different crystal forms, and by analyzing structural differences among monomers within the asymmetric unit of a given crystal form (Simpson et al. 2001) (Fig. 23.4d, e). Differences between different crystal forms are localized in the wide-end domain. Similarly, superposition of different monomers within a crystallographic asymmetric unit shows that while the central and narrow domains strictly obey 12-fold symmetry, the wide end can undergo rigid body rotations of up to 5.4° relative to the central and narrow domains. These conformational fluctuations correspond to a “rolling” motion of the wide-end domain and were proposed as an essential component of a connector-based mechanism for DNA packaging (Simpson et al. 2000) (see also Sect. 23.3.2).

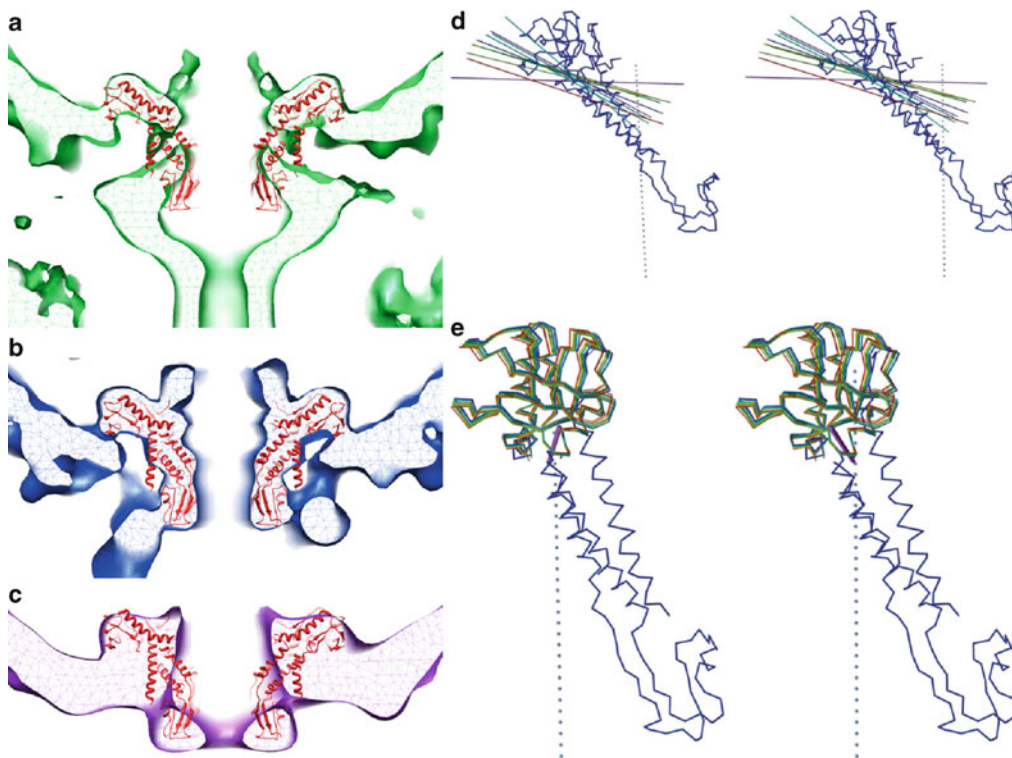


Fig. 23.4 Conformational flexibility of the connector. (a) Fit of the atomic coordinates of the connector (red ribbon, Protein Data Bank accession code: 1IJG) to cryoEM density (green density) corresponding to the connector in mature phage, (b) to prohead-bound phage (blue density), and (c) to pRNA-free proheads (purple density). Comparison of cryoEM densities indicates conformational changes involving the central and narrow domains during phage maturation. (d) Stereodigram based on the superposition of the helical domains within the crystallographic asymmetric unit of the connector. Different colored lines represent the different axes of rotation relating any one wide-end domain to a reference monomer. (e) The variation in position of the wide-end domain in the crystallographic unit. Parts (a–c) were inspired by Xiang et al. (2006); parts (d and e) were adapted from Simpson et al. (2001)

23.3.3 The pRNA

One of the more surprising observations to emerge from early experiments on the ϕ 29 packaging motor was the sensitivity of packaging activity to ribonuclease, as comparable ribonuclease sensitivity has not been observed in other phage systems. Perhaps equally surprising was the explanation for this sensitivity, i.e., the discovery of a 174-base phage-encoded RNA molecule that is absolutely required for packaging (Guo et al. 1987a). This RNA was shown to be a transcript from a promoter at the extreme left end of the genome (bases 320–147) (Guo et al. 1987a; Wichitwechkarn et al. 1989). Since the RNA was found to bind to proheads (Guo et al. 1987b; Garver and Guo 1997; Simpson et al. 2000), it was named “pRNA.” The pRNA is not needed to assemble proheads (Guo et al. 1991b), suggesting that it attaches to proheads after capsid assembly. It was shown that pRNA binds to proheads in an Mg^{2+} -dependent reaction (Wichitwechkarn et al. 1989; Reid et al. 1994a; Trottier and Guo 1997). The role of pRNA is transitory and likely limited to DNA packaging, as pRNA is not present in mature ϕ 29 (Guo et al. 1987a).

Bioinformatic analysis and secondary structure prediction indicated that the pRNA consists of two domains; domain I consists of the first 117 bases and is separated from domain II (bases 131–174) by

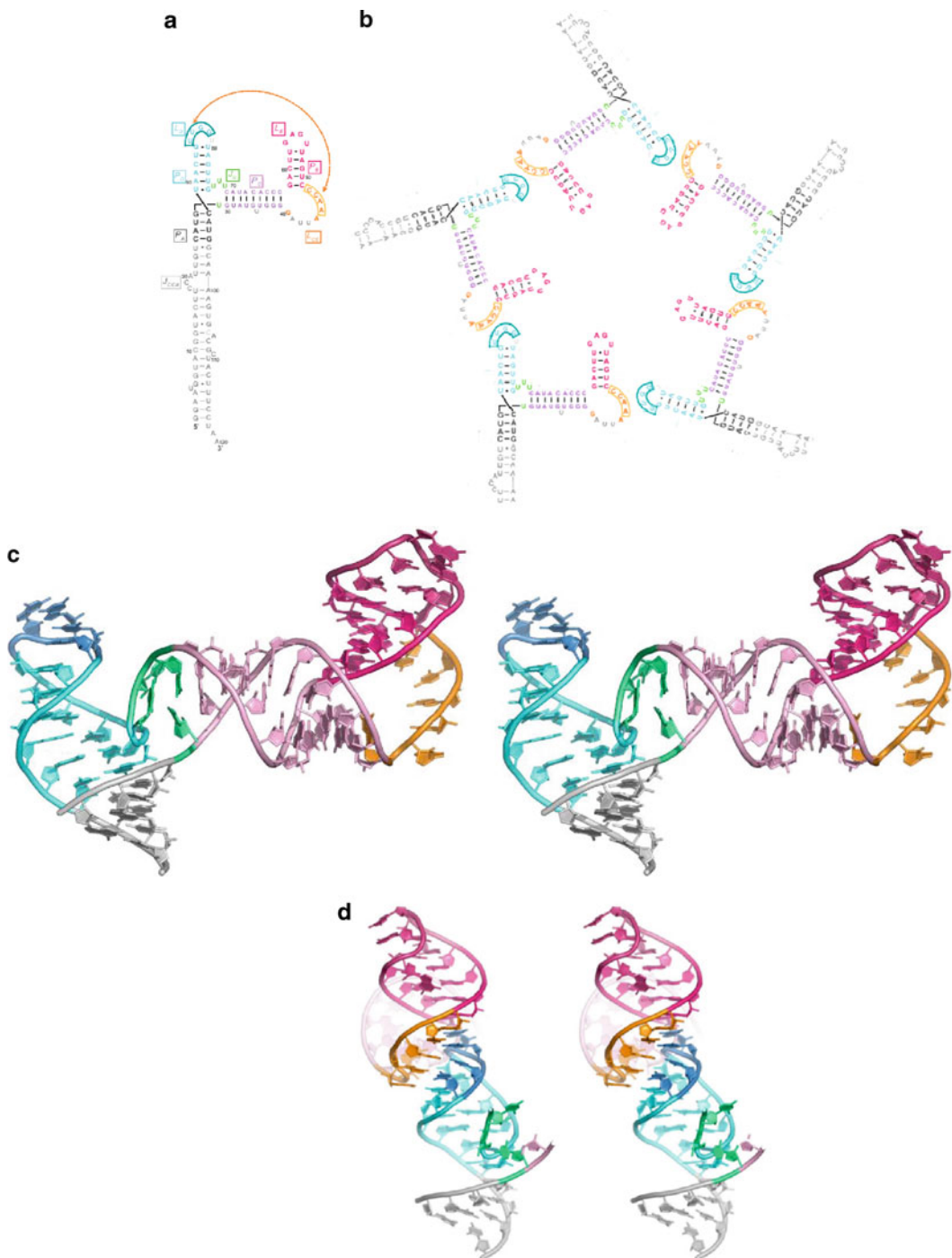


Fig. 23.5 Crystal structure of the pRNA prohead-binding domain. **(a)** Predicted secondary structure of pRNA domain I. Nucleotides are colored: *gray*, A-helix; *green*, three-way junction; *violet*, C-helix; *orange*, CE-loop; *magenta*, E-helix and E-loop; and *cyan*, D-helix and D-loop. In *light gray* are nucleotides deleted in the final crystallization construct. **(b)** Proposed cyclic oligomerization of the pRNA based on intermolecular base pairing at the pseudoknot. Although hexameric oligomerization was initially predicted, a pentameric structure is shown to reflect the likely oligomerization state on the prohead. **(c)** Stereodigram of the crystal structure of the pRNA-binding domain. The phosphate backbone is shown as a coil, and the bases as filled slabs colored as above. **(d)** Stereoview of the intermolecular pseudoknot. The CE–D loop interaction mediates the continuous base stacking from the E-helix of one pRNA monomer to the A-helix of its neighbor, creating an RNA superhelix. This figure was adapted from Ding et al. (2011)

a 13 base single-stranded region (Bailey et al. 1990) (Fig. 23.5a). Although the full-length wild-type pRNA is a two-domain 174-base transcript, a 120-base construct encompassing domain I and lacking the 54 nucleotide 3'-domain II is fully competent for packaging and assembly *in vitro*. Although the function of domain II is not currently understood, its conservation in all ϕ 29 relatives suggests that it has an essential function *in vivo*. Seven RNA helices, designated A–G, were also predicted, with helices B and F residing in domain II, and helices A, C, D, and E in domain I. Three conserved loops, CE, D, and E, named for their position relative to the similarly named helices, were also identified.

Packaging activity is eliminated upon removal of pRNA, but can be restored by reconstituting RNA-free particles with recombinantly expressed pRNA (Guo et al. 1987b). Reconstitution of RNA-free proheads with mutant pRNAs is a remarkably powerful technique, and in combination with bioinformatics and phylogenetic analysis, it has been used to exhaustively dissect structural and functional elements of the pRNA (for a comprehensive review, see Grimes et al. 2002). Early experiments distinguished two functional modules within pRNA domain I, one necessary for prohead binding (bases 22–84), the other required for DNA packaging activity (bases 1–20 and 92–117) (Reid et al. 1994a, b, c; Zhang et al. 1994, 1995). The prohead-binding domain contains the C, E, and D helices and the CE, E, and D loops; all of these structural elements were shown to be important for prohead binding, as were the base sequences within each element (Reid et al. 1994b). Phylogenetic analysis, combined with mutagenesis and biochemistry identified an essential U-rich three-way junction that was predicted to join helices A, D, and E in domain I (Bailey et al. 1990). The DNA packaging domain corresponds to the majority of the pRNA A-helix; mutants in this region bind proheads as tightly as wild type, but are defective in DNA packaging (Reid et al. 1994b, c; Zhang et al. 1994, 1995). Since many of these mutants were also defective in binding the ATPase, it was proposed that the pRNA A-helix provides the attachment site for the packaging ATPase.

Considering that quaternary structures larger than homodimers had never been observed for RNA (Ding et al. 2011), it was surprising that recombinantly expressed pRNA formed higher-order oligomers in solution (Guo et al. 1998; Zhang et al. 1998; Chen et al. 1999, 2000). This oligomerization was demonstrated to arise from a novel intermolecular RNA pseudoknot. From examination of the predicted secondary structure of the pRNA, it was clear that bases 45–AACC–48 from the C–E loop could form Watson–Crick base pairs with bases 85–UUGG–82 from the D loop (Reid et al. 1994c) (Fig. 23.5a, b). Furthermore, phylogenetic analysis indicated that comparable base-pairing potential is conserved in pRNAs from ϕ 29 relatives. To test the biological significance of this potential pseudoknot, the CE-loop was mutated from AACC to GCGA in one mutant, and the D-loop mutated from UUGG to CGCU in another mutant. As expected, neither mutant alone could package DNA, yet a double mutant, which restores the base-pairing potential, regained full biological activity. Interestingly, a mixture of the two single mutants also restored packaging (Zhang et al. 1998). Since formation of intramolecular pseudoknots would not be possible in this experiment, it was concluded that the pseudoknot must be an intermolecular one between neighboring pRNAs. It was further proposed that formation of intermolecular pseudoknots between adjacent pRNAs would promote circular interactions resulting in an oligomeric ring-like structure (Fig. 23.5b).

Structural analysis has now confirmed many of the early predictions regarding pRNA structure and function. CryoEM analysis has definitively shown that (1) bases 22–84 form a pentameric ring-like structure that binds to the prohead; (2) five pRNA A-helices extend as spokes from this central ring; and (3) the packaging ATPase attaches to the distal end of the A-helix spokes (Morais et al. 2008) (Fig. 23.2). The recently determined X-ray structure of the prohead-binding domain of the pRNA confirms the presence of the U-rich junction and, via crystal contacts between pRNA monomers, the intermolecular pseudoknot (Ding et al. 2011) (Fig. 23.5c, d). That such sophisticated and prescient structure/function predictions could be made in the absence of any direct structural information is a testament to the power of carefully conceived genetic and biochemical experiments when applied to a well-characterized system.

Of course, structural data on the pRNA has also provided additional information that would not be available using other experimental approaches. The crystal structure of the prohead-binding

domain shows that pRNA monomers adopt an extended conformation, and that the three major helical elements are organized around the U-rich three-way junction (J_u); helices A and D stack coaxially, whereas the C-helix extends outward at an approximately 70° angle (Ding et al. 2011) (Fig. 23.5c). At the distal end of the C-helix, the CE-loop causes the direction of the E-helix to change by about 110° . The conformation of the E-loop is similar to an NMR-derived structure of an E-loop hairpin (Harris and Schroeder 2010). Analysis of electron density and temperature factors in the crystal structure suggests the presence of two flexible joints at the J_u bulge and the CE-loop. In addition to confirming the presence of all four potential Watson–Crick base pairs in the intermolecular pseudoknot, the structure showed that there is an additional fifth interaction, a base triple between A_{44} in the CE-loop and G_{86} and U_{80} in the D-loop (Fig. 23.5d).

The crystal structure of the prohead-binding domain of the pRNA fits remarkably well into its corresponding density in a cryoEM reconstruction of the $\phi 29$ prohead (Fig. 23.6a). The excellent agreement between the two structures not only confirms that each is essentially correct but also provides insight as to how the pRNA interacts with the connector and the prohead. The structure fits into the density region corresponding to the central ring, where it makes contact with protruding narrow end of the connector (Fig. 23.6b). The pRNA bases adjacent to the connector are consistent with hydroxyl radical probing studies that showed increased cleavage sensitivity in the ring region of the pRNA when an Fe-BABE probe was tethered to the narrow end of the connector (Atz et al. 2007). Furthermore, the fitted pRNA is close to the expected position of the positively charged RKR residues at the N-terminus of the connector that were shown to be important for pRNA binding (Xiao et al. 2005; Atz et al. 2007). Additionally, this fitting placed the E-loop in density that connects the pRNA central ring to the prohead (Fig. 23.6a). This density had been identified as pRNA in previous cryoEM reconstructions (Morais et al. 2005, 2008), but existing pRNA models were unable to explain which bases of the pRNA accounted for this density. The current fitting clearly shows that E-loop bases G_{55} and A_{56} make contact with a unique-binding site on the capsid protein that is exposed as a result of rotation of the Ig-like C-terminal domains in capsid proteins circumscribing the unique motor vertex (Morais et al. 2005, 2008; Ding et al. 2011) (Fig. 23.6c). The importance of these bases in prohead binding was also confirmed via mutagenesis (Ding et al. 2011).

Although pRNA monomers were independently fitted into the cryoEM density without any prior assumptions, the intermolecular base-pairs were preserved in the resulting pRNA pentamer. This fitting also positions the small portion of the A-helix present in the truncated construct in cryoEM density identified as the A-helix by cryoEM (Morais et al. 2008; Ding et al. 2011), thus allowing modeling of the remainder of the A-helix and construction of pseudoatomic model of entire pRNA domain I. This model revealed a remarkable RNA superhelix that dominates the pRNA structure and arises from continuous stacking of multiple helical elements from adjacent pRNA monomers; an E-helix from one monomer shares a common superhelical axis with a D-helix (via the intermolecular pseudoknot) and an A-helix from a neighboring pRNA monomer (Fig. 23.6d). Hence, the basic structural unit of the pRNA spans two topological units, an arrangement that might mediate coordination between adjacent motor subunits.

The oligomerization state of the pRNA on the prohead has been a source of considerable controversy, and thus merits discussion here. Although the pRNA forms tetramers in the crystallographic unit cell (Ding et al. 2011), this arrangement is not believed to reflect the oligomerization state on the prohead. A truncated construct was crystallized, with additional mutations introduced to reduce molecular flexibility. Furthermore, the fourfold molecular symmetry axis is coincident with the crystallographic fourfold axis. Hence, since there is no evidence for wild-type pRNA tetramers, it seems likely that this packing arrangement might be an artifact of either/both crystallization and/or genetic manipulation of the construct.

Wild-type pRNA oligomerizes in solution, and both hexamers and dimers have been detected by analytical ultracentrifugation (Guo et al. 1998; Zhang et al. 1998; Chen et al. 1999, 2000). Initial genetic investigations into the nature of the intermolecular pseudoknots showed that two-, and three-way

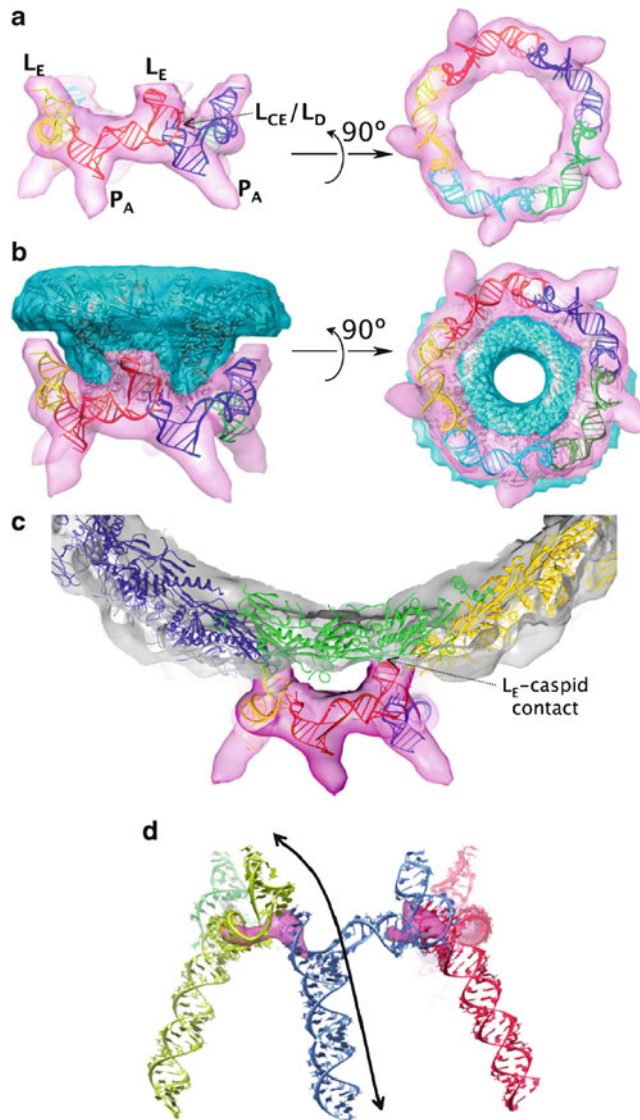


Fig. 23.6 Pseudoatomic structure of the pRNA on proheads. **(a)** Side and bottom views of the pentameric pRNA ring structure docked into its cryoEM envelope. **(b)** Side and bottom views of the pentameric pRNA ring and the dodecameric connector structures fit into their corresponding cryoEM envelopes. **(c)** Side view of the pRNA ring and a pseudoatomic model of the capsid protein fit into the cryoEM envelope showing that the pRNA contacts the capsid via the E-loop. The connector structure was omitted for clarity. **(d)** Pseudoatomic structure of full-length pRNA. Three of the five pRNA monomers are colored *blue*, *red*, and *yellow*, respectively, and part of a fourth monomer is shown in green. The A-helix was modeled as an ideal RNA helix and positioned via fitting to cryoEM density. The E-helix from one monomer shares a common super helical axis with a D-helix (via the intermolecular pseudoknot) and an A-helix from a neighboring pRNA monomer. The approximate direction of the superhelical axis is shown as a *double-headed arrow*. The strongest density in cryoEM maps is shown as a *magenta* isosurface; note that this density is near the intermolecular pseudoknot, demonstrating that this region is highly ordered in proheads. Parts **(a–c)** of this figure were modified from Ding et al. (2011)

complementation of CE/D-loop mutants could rescue packaging, suggesting that both multiples of two and three are compatible with an active pRNA oligomer. Hence, the pRNA was believed to be hexameric because (1) six is the smallest integer that is a multiple of both two and three, (2) hexamers were observed in analytical ultracentrifugation, and (3) the pRNA was initially believed to bind to

the connector, an oligomer with 12-, and thus also sixfold symmetry. This view was recently reinforced by single-particle fluorescence quenching experiments where statistical analysis of photobleaching events was used to conclude that there is an average of six pRNAs per prohead (Shu et al. 2007). However, recent analytical ultracentrifugation experiments show that head-to-tail dimers dominate in solution, not hexamers (Chen et al. 2000), and it is likely these dimers are artifacts of recombinant overexpression since comparable dimers are not observed in X-ray or cryoEM structures of the pRNA (Ding et al. 2011).

A hexameric pRNA also does not agree with structural results obtained from either cryoEM or X-ray crystallography, which clearly show pentameric and tetrameric pRNAs, respectively. Fitting of the X-ray structure of a pRNA into cryoEM density results in a pentameric pRNA, even for carefully calculated asymmetric reconstructions where no symmetry is imposed (Simpson et al. 2000; Morais et al. 2001; Ding et al. 2011). This fitting also indicates that although the pRNA is close to the connector, it makes much more extensive contacts with the fivefold symmetric head than the 12-fold symmetric connector. Similarly, it has been shown that pRNA binds to connectorless proheads with an occupancy comparable to that seen for wild-type prohead binding, indicating that the pRNA's primary-binding site is on the fivefold symmetric capsid rather than the 12/6-fold symmetric connector (Morais et al. 2008). Furthermore, the excellent fit of the X-ray structure of the pRNA to the cryoEM structure demonstrates that two independent methods give rise to essentially identical structures; it is difficult to account for such a result if the structures are not truly similar. Thus, it seems that the pentameric pRNA observed in cryoEM reconstructions reflects its true oligomeric state.

The argument has been made that cryoEM reconstructions of $\phi 29$ will always result in unintentional imposition of fivefold symmetry, and thus that the observed pentameric pRNA is an artifact of unintentional symmetry averaging (Shu et al. 2007). The underlying logic is that individual particles are aligned for averaging during reconstruction by virtue of their relatively massive fivefold symmetric capsids. Hence, small components of the motor that do not have the same symmetry as the prohead might not be aligned for meaningful averaging during the reconstruction procedure, see (Morais et al. 2001) for a more extensive discussion. However, it is important to point out that improper symmetry averaging, whether intentional or unintentional, results in smeared out uniform density without any distinguishing features. In contrast, the pRNA structure determined by cryoEM has distinct features that agree with the crystallographically determined structure of the pRNA (Ding et al. 2011). Furthermore, it was shown that subtle yet predictable structural changes can be introduced into the pRNA, and that these changes can then be readily observed in cryoEM reconstructions of proheads (Morais et al. 2008; Zhao et al. 2008), strongly indicating that the reconstruction procedure faithfully reproduces the molecular structure of the pRNA. Finally, the dodecameric connector provides an internal control to verify that symmetry mismatches have been resolved in cryoEM reconstructions of $\phi 29$ proheads. In particular, the ability to resolve 12 monomers, as well as secondary structural elements within monomers, in a recent ~ 8 Å asymmetric reconstruction of the packaging motor demonstrates that even relatively small nonpentameric components are properly aligned and averaged in asymmetric prohead reconstructions (Morais, unpublished).

The question of how pRNA assembles on the head also remains controversial. Based on the early analytical centrifugation results and pseudoknot complementarity experiments, it was assumed that the pRNA forms a hexameric ring which then docks onto the connector (Guo et al. 1998; Zhang et al. 1998). Later results suggested that pRNA forms head-to-tail dimers in solution, but that this dimer is dissembled as the pRNA binds to the connector to form a hand-to-hand circular ring. Additional evidence for connector-based docking came from experiments showing that the first 14 amino acids of the connector are important for pRNA binding, especially the short RKR sequence (Xiao et al. 2005; Atz et al. 2007). However, it was recently shown that pRNA can bind particles that lack connectors as well as bind wild-type particles (Morais et al. 2008).

An initial clue regarding pRNA assembly came from analysis of cryoEM reconstructions of the prohead, which show that the E-loop of the pRNA makes extensive contacts with the capsid that would be sterically impossible if the Ig-like domain of the capsid protein had not rotated $\sim 180^\circ$ from its expected position (Morais et al. 2005, 2008; Ding et al. 2011). Hence, an assembly scenario compatible with all the conflicting data is one in which the pRNA initially recognizes and binds to the RKR residues at the N-terminus of the connector, followed by a conformational change in the capsid proteins surrounding the motor vertex that creates a new, tight-binding site on the capsid for the pRNA. This proposed conformational change likely involves the previously described rotation of the Ig-like C-terminal domains in capsid proteins surrounding the pRNA (Morais et al. 2005), which would uncover the pRNA-binding site on the capsid. It was demonstrated that this conformational change is present in prohead particles whose pRNA has been removed by RNase digestion (Morais et al. 2008). Further support for this assembly scenario comes from cryoEM reconstructions of proheads that have never seen pRNA. It would be expected that the C-terminal Ig-like domain would not be rotated from its typical position since these particles are pRNA naive and thus would not have undergone the conformational change necessary for binding pRNA; this is indeed the case. Furthermore, it has been shown that when pRNA binds these particles, the Ig-like domain rotates from its expected position as in wild-type proheads (Morais, unpublished data). Thus, it may be that pRNA assembly proceeds via attachment to the procapsid first, followed by formation of intermolecular pseudoknots. However, since pseudoknots are absolutely required for packaging, they may play a role in the mechanism of translocation. Their location at the junction of helices contributed from neighboring monomers in the pRNA superhelix may provide a means of communication around the motor ring (Fig. 23.6d).

23.3.4 The Packaging ATPase

Packaging dsDNA into a preformed viral capsid is an inherently unfavorable energetic process. The dsDNA helix is a relatively rigid molecule, with a persistence length of approximately 50 nm. It is also highly negatively charged due to its phosphate backbone. Hence, there are considerable enthalpic and entropic costs associated with confining DNA within the limited space of the viral capsid. Coarse-grained simulations of DNA packaging and subsequent decomposition of packaging energy into its principle components suggests that electrostatic repulsions contribute most of the enthalpic penalty associated with packaging the ϕ 29 DNA-gp3, whereas the elastic cost of bending the DNA helix is modest (Petrov and Harvey 2008). These analyses also indicate that the entropic cost of packaging accounts for approximately half of the total free energy cost of genome encapsidation. This is not surprising since ϕ 29 DNA-gp3 has a radius of gyration of ~ 340 nm in solution compared to only ~ 15 nm inside proheads, corresponding to an $\sim 10,000$ -fold volume compression; clearly, there must be a considerable entropic price associated with such a dramatic reduction of conformational space. Regardless of relative percentages, the enthalpic and entropic costs of genome encapsidation must be paid for via the input of energy.

ATP is the universal energy currency of the cell, and several protein families have evolved that couple ATP hydrolysis to otherwise thermodynamically unfavorable reactions. The net effect of this coupling is to drive the linked reactions in a direction necessary to do chemical or mechanical work. The dsDNA viruses, including tailed bacteriophages like ϕ 29, have co-opted polymer-dependent ATPases from the ASCE P-loop NTPase superfamily to couple hydrolysis of host-supplied ATP to the translocation of the dsDNA viral genome into the preassembled procapsid (Mitchell et al. 2002; Burroughs et al. 2007; Singleton et al. 2007; Thomsen and Berger 2008) (Fig. 23.7). The energy released by hydrolysis of the high-energy phosphoanhydridic bond between the β - and γ -phosphates induces a conformational change in the ATPase that exerts a mechanical force that results in translocation of the DNA into the head.

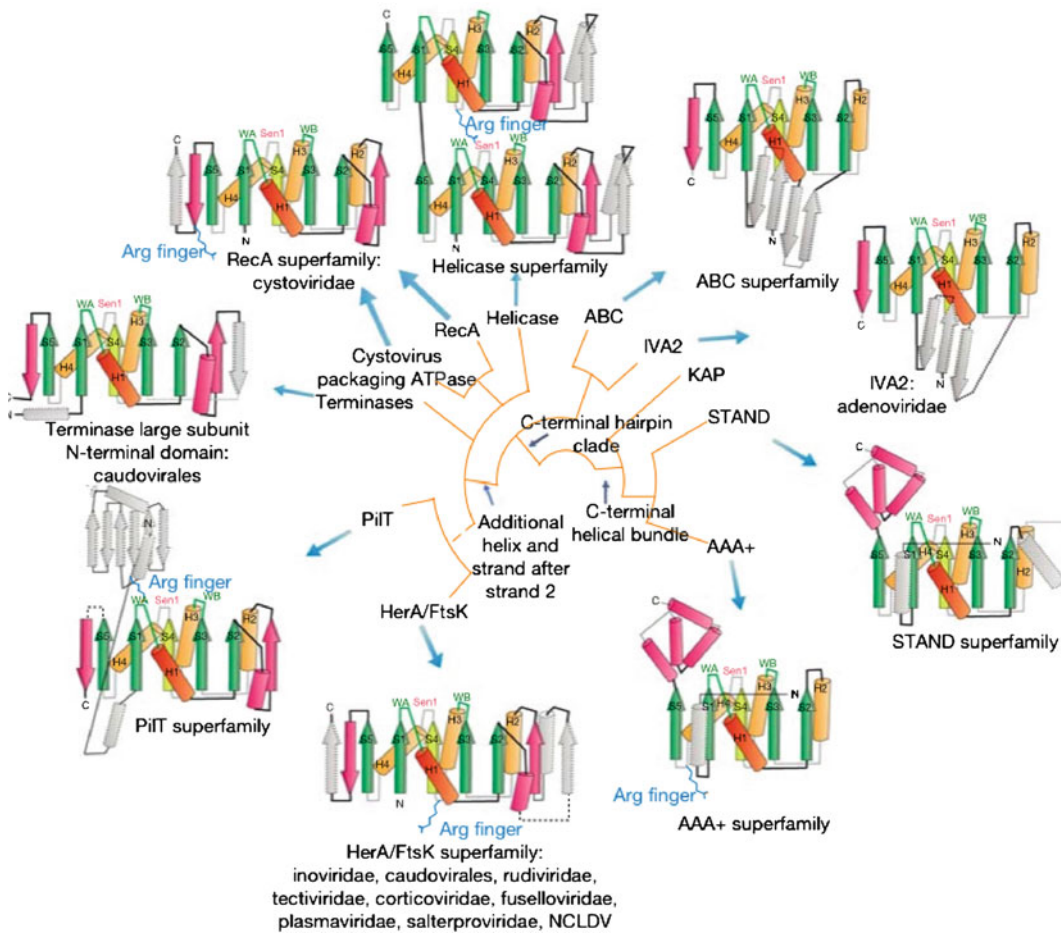


Fig. 23.7 ACSE superfamily. Proposed evolutionary relationships between ASCE P-loop NTPases are shown. Topology diagrams of different representative family members are superimposed on a cladogram depicting higher-order relationships. The highly conserved central strands and helices forming the core of the ASCE domain are numbered and colored. Strands are shown in green with the central strand S4 colored yellow and helices colored orange. Synapomorphies shared across different lineages are shown in pink, whereas nonconserved features are colored gray and outlined in broken lines. Lines connecting different lineages represent higher-order relationships constructed by comparison of shared structural and/or sequence similarities. Broken lines represent more tenuous relationships. Viral lineages with packaging ATPases from a specific superfamily are listed following a colon below the superfamily name. Abbreviations: WA Walker A, WB Walker B, and Sen1 sensor-1. This figure is reproduced from Burroughs et al. (2007)

The 39-kDa 332-residue $\phi 29$ packaging ATPase consists of two domains (Koti et al. 2008). The first ~200 N-terminal amino acids correspond to the ASCE P-loop ATPase domain, and the last ~130 residues comprise the C-terminal domain. Sequence analysis suggests that the C-terminal domain adopts a five-stranded oligonucleotide/oligosaccharide-binding fold, possibly indicating that the C-terminal domain could either play a role in pRNA binding (Koti et al. 2008) or DNA translocation (Sun et al. 2007, 2008). Both dimers and trimers of recombinantly expressed $\phi 29$ ATPase have been reported (Lee and Guo 2006; Koti et al. 2008), although larger oligomeric rings are likely assembled on proheads (Simpson et al. 2000; Koti et al. 2008; Morais et al. 2008). Since the ATPase has been shown to bind to the A-helix of the pRNA, it seems likely that it would also form pentamers on the prohead. This prediction is supported by cryoEM reconstructions that show a pentameric ring of ATPases bound to distal ends of the pRNA A-helices (Koti et al. 2008;

Morais et al. 2008) (Fig. 23.2). It has been shown that the ATPase can either assemble on the pRNA and then bind the DNA-gp3 substrate, or first interact with DNA-gp3 and then attach to the pRNA (Koti et al. 2008). It is not clear which assembly pathway is utilized in infected *Bacillus* cells.

Similar to other phage packaging ATPases, the ϕ 29 ATPase has only weak ATPase activity in the absence of other motor components (Grimes and Anderson 1990). This activity is modestly stimulated by the addition of pRNA and increases even further in the presence of pRNA–prohead complexes. There is no increase in activity when the ATPase is mixed with pRNA-free proheads, consistent with structural and biochemical results indicating that the ATPase binds to the A-helix of the pRNA (Reid et al. 1994b, 1994c; Zhang et al. 1994, 1995; Simpson et al. 2000; Morais et al. 2008). The increased activity observed for prohead–pRNA–ATPase mixtures compared to pRNA–ATPase mixtures indicates that the ring-like oligomerization of the ATPase upon prohead binding is important for catalysis. However, maximum activity is only observed in the presence of both pRNA–prohead complexes and the DNA-gp3 substrate. The mechanism of DNA-enhanced ATPase activity is not well understood, although it has been proposed that the translocation of the helically symmetric dsDNA substrate through the ring channel coordinates ATPase activity around the ring (Morais et al. 2008).

Although no high resolution structural information is available for the ϕ 29 ATPase, considerable insight can be obtained from analysis of structurally homologous ATPases. The N-terminal domain of the ϕ 29 ATPase contains several short sequence motifs characteristic of the ASCE P-loop NTPases, including an adenine-binding motif, Walker A and B motifs, a catalytic glutamate, a “phosphate sensor” motif, and an arginine finger (Fig. 23.7). The ASCE P-loop NTPases are built upon a Rossmann fold (Rossmann et al. 1974). First identified in dinucleotide-binding proteins, the Rossmann fold consists of three or more parallel β -strands interspersed by α -helices. This arrangement results in the C-terminal ends of each strand forming one edge of a β -sheet that is sandwiched between α -helices. The conserved nucleotide-binding site is located at this edge and consists of loops connecting β -strands to downstream α -helices. The ASCE P-loop NTPases adopt a modified version of this fold, with five conserved parallel β -strands forming a central β -sheet that contains all of the sequence signatures listed above (Story and Steitz 1992; Story et al. 1992). Consistent with the position of the nucleotide-binding site in the Rossmann fold, residues responsible for ATPase function are located almost exclusively on loops connecting the C-termini of the β -strands to α -helices. The position, sequence, and function of each of these motifs during catalysis are summarized below.

The Walker A and B motifs, named for their identification in the sequence of the bovine mitochondrial F_1 -ATPase (Walker et al. 1982), are the most highly conserved sequences in the ASCE P-loop NTPase superfamily. Both are involved in nucleotide-binding and activation of the nucleophilic water molecule. The Walker A motif, also referred to as the P-loop, is located between strand β -1 and helix α -1 and is defined by the consensus sequence (G/A)XXXXGK(T/S), where X can be any amino acid. The backbone amides of the glycine residues coordinate the phosphate groups of the ATP, and the ϵ -amino group of the conserved lysine forms ion pairs with the β - and γ -phosphates. The hydroxyl group of the conserved threonine/serine coordinates a critical catalytic Mg^{2+} ion that coordinates the β - and γ -phosphates of ATP and activates the catalytic water nucleophile. The Walker B motif, located between strand β -3 and helix α -3, contains the consensus sequence ZZZZD/E, where Z is a hydrophobic residue. The conserved acidic group provides another coordinating ligand for the catalytic Mg^{2+} , thus also playing an essential role in positioning both the ATP substrate and the water nucleophile for in-line attack of the activated water molecule at the γ -phosphate. The adenine-binding motif, ZQ, is located ~15 amino acids upstream of the Walker lysine (Mitchell et al. 2002), where the conserved glutamine forms hydrogen bonds with the N6 and N7 nitrogens of the heterocyclic adenine ring (Sun et al. 2007).

Although the spatial position of the catalytic glutamate is highly conserved, its sequence position varies among different branches of the ASCE superfamily. For example, in the F_1 /RecA, HerA/FtsK, and PilT ATPases, the catalytic glutamate is located on a loop between strand β -2 and helix α -2 (Story et al. 1992). In contrast, the catalytic glutamate in the helicase and ABC ATPases is positioned

immediately after the aspartate of the Walker B motif (Subramanya et al. 1996; Hung et al. 1998; Geourjon et al. 2001). Sequence analysis of the ϕ 29 ATPase and the bacteriophage terminase large subunits (TLS) ATPases indicates that their catalytic glutamates are also positioned immediately following the aspartate in the Walker B motif (Mitchell et al. 2002; Burroughs et al. 2007; Koti et al. 2008), a prediction confirmed by the X-ray crystal structure of the ATPase domain of the T4 large terminase domain (Sun et al. 2007). Regardless of the location of the catalytic glutamate in the primary amino acid sequence, its position in three-dimensional space is highly conserved among all known ASCE ATPase structures. In each case, the catalytic glutamate is positioned to act as a general base and activate a water molecule for nucleophilic attack on the γ -phosphate of ATP (Geourjon et al. 2001; Sun et al. 2007).

Another essential motif conserved in the active site of many ASCE ATPases is a polar residue at the C-terminal end of strand β -4. This residue, asparagine 159 in ϕ 29, is proposed to function as a “ γ -phosphate sensor” that plays a role in transmitting conformational changes to other parts of the protein upon ATP binding (Story and Steitz 1992; Yoshida and Amano 1995; Iyer et al. 2004a, b). A similar function was assigned to the switch II region of distantly related G proteins, members of the broader P-loop ATPase superfamily (Leipe et al. 2002). In ring-forming ATPases like the ϕ 29 ATPase pentamer, this sensor might function to communicate the catalytic state of a particular monomer to its neighbors, and hence play an essential role in coordinating the activity of the ring.

Similarly, most ASCE P-loop ATPases also utilize an “arginine finger” to couple hydrolysis of ATP to large scale conformational changes between adjacent monomers and/or domains. The arginine finger from one monomer/domain is inserted into the active site of its neighbor, where the positively charged guanidinium group of the arginine stabilizes the pentavalent phosphoanhydride transition state that forms during ATP hydrolysis and promotes separation of the phosphate leaving group (Nadanaciva et al. 1999; Braig et al. 2000; Ogura et al. 2004). This *in trans* activation of substrate in a neighboring subunit/domain is reminiscent of the arginine (or lysine) fingers observed in GTPase-activating proteins (Wittinghofer et al. 1997).

As with the catalytic glutamate, the primary sequence position of the arginine finger varies in different branches of the ASCE superfamily, but the spatial position is highly conserved. In the HerA/FtsK superfamily, the arginine finger is at the N-terminal end of strand β -4 and is inserted into the active site of the adjacent monomer in an oligomeric ring. Similarly, the arginine finger in the Helicase superfamily is also located at the base of strand β -4 in one of the two nucleotide-binding domains that comprise the overall fold of this superfamily. In this case, the arginine finger participates in an interdomain rather than intersubunit engagement of a neighboring active site. In the TLS superfamily branch of the ASCE ATPases, the arginine finger is in the third position of the Walker B motif following strand β -3 (Burroughs et al. 2007; Sun et al. 2007). Interestingly, the crystal structure of the T4 large terminase domain indicates that this arginine finger acts *in cis*, via an intrasubunit interaction rather than *in trans* (Sun et al. 2007).

Bacteriophage ϕ 29 and its relatives have conserved arginines at both positions, arginine 26 in the Walker A motif and arginine 146 at the base of strand β -4, an observation that has led to differing classifications of the ϕ 29 ATPase within the ASCE superfamily (Mitchell et al. 2002; Burroughs et al. 2007). This distinction is significant, as it has important implications regarding the mechanism of the ϕ 29 ATPase, i.e., whether it functions more like a monomeric helicase, or like an oligomeric ring helicase/DNA translocase. Based on secondary structure prediction and identification of the characteristic sequence motifs, it was initially assumed that the ϕ 29 packaging ATPase belonged to the same superfamily as the large terminase proteins from other tailed dsDNA viruses (Mitchell et al. 2002). More recently, it was proposed that the ϕ 29 ATPase is more closely related to the HerA/FtsK ATPases within the broader ASCE superfamily (Burroughs et al. 2007). This assignment is based on not only the position of the arginine finger but also the absence of a nuclease domain and the topology of the accessory domain inserted between strands β -2 and helix α -2. However, this

classification is problematic in that all other viral ATPases in the HerA/FtsK superfamily come from viruses that have internal lipid membranes, lack portal/connector proteins, and are constructed from capsid proteins utilizing the jelly roll fold. Like all tailed dsDNA bacteriophages, ϕ 29 lacks internal membranes, depends on portal/connector proteins for packaging, and utilizes the HK97-fold for their major capsid proteins (Wikoff et al. 2000; Fokine et al. 2005; Morais et al. 2005). Hence, the logical conclusion of grouping ϕ 29 with the HerA/FtsK superfamily is that dsDNA packaging motors evolved from ancestral ASCE ATPases at two separate occasions during the evolution of tailed bacteriophages.

Regardless of the exact location of the arginine finger in the primary amino acid sequence, in three-dimensional space, it is typically positioned on the edge of the β -sheet opposite the edge where the ATP-binding loops are situated. This arrangement suggests an intriguing mechanism for coupling ATP hydrolysis to force generation in ring ATPases. In some members of the ASCE ATPase family, movement of the substrate-binding loops is coupled to twisting of the central β -sheet. This β -sheet deformation has been proposed to store elastic energy that can be used to power large motions in accessory domains. (Abrahams et al. 1994; Kabaleeswaran et al. 2006) Hence, controlling sheet deformation might provide a means for coordinating ATP hydrolysis and conformational change. Since the arginine finger and substrate-binding loops within a subunit are on opposite sides of the central β -sheet, the arginine finger might also help transmit information regarding the catalytic state of the neighboring active site across this sheet to the substrate-binding loops on the opposite edge, thus providing a route for communication between adjacent subunits.

Structural information regarding arginine finger-dependent coordination in ASCE ATPases comes from the ssRNA bacteriophage ϕ 12 packaging motor P4 (related to SF4 helicases) (Lisal and Tuma 2005; Burroughs et al. 2007) and from PcrA and UvrD, members of the SF1 helicase branch of the ASCE superfamily (Velankar et al. 1999; Lee and Yang 2006). For these motors, structures of motor nucleic acid complexes in multiple ATP-bound states have been determined. These structures provide examples of how the relatively small intradomain conformational changes that occur upon ATP binding can be amplified in an oligomeric motor. In the empty state, the two ASCE nucleotide-binding domains are separated so that the arginine finger cannot reach its neighbor's active site. Upon binding ATP, one of these domains rotates toward the other, facilitating insertion of the arginine finger into the neighboring domain's active site and subsequent hydrolysis of the ATP. In P4, this movement is coupled to the movement of a luminal lever that interacts with the ssRNA to propel it into the capsid. In the SF1 helicases, this movement is coupled to unwinding the dsDNA duplex and translocation along the resulting ssDNA.

Several mechanistic generalizations can be drawn from the above analyses of the characteristic functional motifs in ASCE ATPases. When applied to ϕ 29 packaging ATPase, a picture of its role in genome encapsidation begins to emerge. By analogy to other ASCE ring ATPases, ATP and Mg^{2+} are bound at an intersubunit interface such that adjacent monomers both contribute residues to the active site. Hence, ATP binding and hydrolysis directly affect adjacent subunits, and the motor can thus respond to allosteric interactions between multiple ATP and DNA-binding sites. In a ring-like structure, this arrangement can facilitate coordination of conformational changes required for DNA translocation. The glycine-rich Walker A P-loop and the PQ adenine-binding motif within the active site cleft recognize and bind the adenine base in ATP and thus help correctly orient the nucleotide substrate for hydrolysis. Walker A and Walker B loops both contribute to binding the catalytic Mg^{2+} ion. The catalytic glutamate immediately following the Walker B aspartate helps to activate the Mg^{2+} -bound water molecule for nucleophilic attack on the γ -phosphate of ATP. In order to couple coordinated conformational changes to particular nucleotide states, the ϕ 29 ATPase likely uses the phosphate sensor at the tip of strand β -4 to distinguish ATP from its hydrolysis products. The arginine finger at the base of strand β -4 stabilizes the pentavalent transition state and phosphate leaving group during hydrolysis and likely plays a role in coupling hydrolysis of ATP to the conformational rearrangements that drive DNA translocation. Because the arginine finger belongs to a neighboring

subunit, this in trans activation might also be important in initiating the next ATP hydrolysis event in a neighboring subunit in the oligomeric ring. Although clearly important for directional transport of DNA, the underlying mechanism for coordinating the sequential firing of ATP subunits around the ring is not well understood. It has been suggested that the inherent 10_1 helical symmetry present in B-form DNA might also play a role in determining the order of subunit firing, i.e., whichever subunit is aligned with a particular structural feature of DNA will be triggered to initiate ATP hydrolysis (Simpson et al. 2000; Morais et al. 2008; Sun et al. 2008).

23.4 Single Particle Analysis of the ϕ 29 DNA Packaging Motor

A complete description of genome encapsidation by the ϕ 29 dsDNA packaging motor remains elusive. In addition to capturing high-resolution structures of each intermediate step during DNA translocation, it is also necessary to correlate each catalytic step of ATP hydrolysis with the sequence of conformational changes that move the dsDNA substrate into the capsid. Although structural, genetic, and bulk biochemical assays have been instrumental in providing a framework for understanding the overall mechanochemistry of genome encapsidation, none of these techniques is particularly well suited for capturing the underlying dynamics of DNA packaging. Fortunately single-particle experiments have recently allowed researchers to directly examine the dynamic properties of individual ϕ 29 motor complexes at extraordinarily high spatial and temporal resolutions (Smith et al. 2001; Chemla et al. 2005; Fuller et al. 2007; Hugel et al. 2007; Rickgauer et al. 2008; Aathavan et al. 2009; Moffitt et al. 2009).

In particular, sophisticated optical tweezer measurements have provided essential information regarding the mechanochemical cycle of the motor (Smith et al. 2001). In these experiments, force-measuring laser tweezers are used to follow packaging activity of a single complex in real time. Briefly, a stalled, partially packaged complex is attached to one polystyrene microsphere via the unpackaged end of the DNA, and this microsphere is captured in an optical trap. The phage capsid is then attached to a second bead, held by a pipette, which is coated with antibodies that selectively bind to the phage capsid. Hence, a stable tether is formed between the two beads; the virus head is bound to one bead, and the unpackaged end of the genome is bound to the other. Because the positions and forces applied to each bead can be precisely controlled, it is possible to monitor the changes in DNA length as a function of time, as well as quantitate the forces necessary to stall the motor. From this force–velocity data, it has been possible to extract detailed information regarding the mechanochemistry of the ϕ 29 packaging motor.

Remarkably, it was found that the ϕ 29 can generate approximately 100 pN of force, making it the most powerful biological motor reported to date (Rickgauer et al. 2008). Initial optical tweezer experiments showed that DNA shortened in bursts exceeding 5 μ m, indicating that the motor is highly processive (Smith et al. 2001). Pauses and slips were also observed, occurring more frequently as the motor works against higher forces. From the force–velocity data, it was shown that the rate-limiting step of the motor's cycle is force dependent even at low loads; hence, the packaging rate decreases as the prohead is filled (Smith et al. 2001). Later experiments using higher resolution optical tweezers indicated that the motor moves in steps of ten base pairs with intervening dwells of varying intervals (Moffitt et al. 2009). Closer analysis indicates that each ten-base pair step consists of four 2.5 base-pair substeps. Single particle packaging experiments using chemically modified DNA substrates show that the motor is somewhat “promiscuous” and can tolerate short inserts of nonnucleic acid polymers in one of the two DNA strands in the double helix (Aathavan et al. 2009), suggesting that the motor tracks along the phosphate backbone of a single DNA strand. Optical tweezers experiments also suggest that the DNA is not simply translated into the capsid, but is also negatively rotated (in the underwinding direction) during packaging (Yu et al. 2010).

Force–velocity data was also obtained in the presence of varying concentrations of ATP, ADP, inorganic phosphate, and various ATPase inhibitors, allowing for the determination of kinetic parameters for the motor as well as their dependence on external load (Chemla et al. 2005). The motor was shown to have a V_{\max} of 103 bp/s, a K_m of $\sim 30 \mu\text{M}$, and a Hill coefficient of 1, showing that ATP binding is not cooperative. ADP was found to act as a competitive inhibitor for ATP, indicating that they share the same binding site. However, while affinity for DNA is high in the ATP-bound state, it is low in the ADP-bound state. In contrast to ADP, there is almost no rate dependence on the concentration of inorganic phosphate, reflecting the large equilibrium constant for product release and suggesting that this step is essentially irreversible. Based on these observations, the authors conclude that DNA translocation does not occur during ATP binding but rather is triggered by phosphate release (Chemla et al. 2005). Furthermore, although the motor has a Hill coefficient of 1, analysis of ATPase inhibitor data indicates that the motor subunits act in a coordinated, successive fashion with high processivity.

23.5 Proposed Packaging Mechanisms

Despite the wealth of genetic, biochemical, biophysical, structural, and single particle data accumulated regarding genome encapsidation by ϕ 29, the fundamental mechanism of DNA translocation remains unknown. Proposed models typically fall into one of two basic categories (1) connector-driven models, where rotation of the connector is coupled to DNA translocation and (2) ATPase-driven models where different domains within the ATPase alternately grab and release DNA. It's important to note that in all proposed models, genome translocation depends on conformational changes in the ATPase induced by hydrolysis of ATP. In ATPase-driven models, these conformational changes directly translocate DNA, whereas in connector-driven models, conformational changes in the ATP module the connector structure to varying degrees depending on the proposed role of the connector in DNA translocation.

23.5.1 Connector Driven Packaging Models

Connector-driven models were initially based largely on the ideas of Hendrix (Hendrix 1978), who proposed that the symmetry mismatch between the fivefold symmetric prohead shell and 12-fold symmetric connector would facilitate rotation of the connector. Hendrix astutely realized that given a pair of concentric rings with differing symmetries, the energy minima for pairwise interactions between subunits in adjacent rings cannot add in phase, and hence only a small energy barrier prevents rotation of one ring with respect to the other. In contrast, if the number of subunits in the two rings is the same, then the energy minima for pairwise interactions between subunits in adjacent rings can add in phase, resulting in a deep energy minimum that essentially locks the two rings in place. Secondly, Hendrix recognized that as a helically symmetric molecule, dsDNA is a threaded object. Combining these observations led to the idea that the 12-fold symmetric connector would be relatively free to rotate in the fivefold symmetric prohead vertex where it sits, and that rotation of the connector around the helically threaded DNA would thus “screw” the DNA into the capsid much as a nut moves on a bolt. Of course, this type of model requires that the lumen of the connector is also threaded in a complementary way to the helical DNA. Since there is no apparent luminal threading observed in any connector structures determined so far, alternative connector-based packaging mechanisms have been developed.

One variation of the Hendrix model is based on observed flexibility of the connector protein and combined results from X-ray crystallography and cryoEM demonstrating that the DNA, connector,

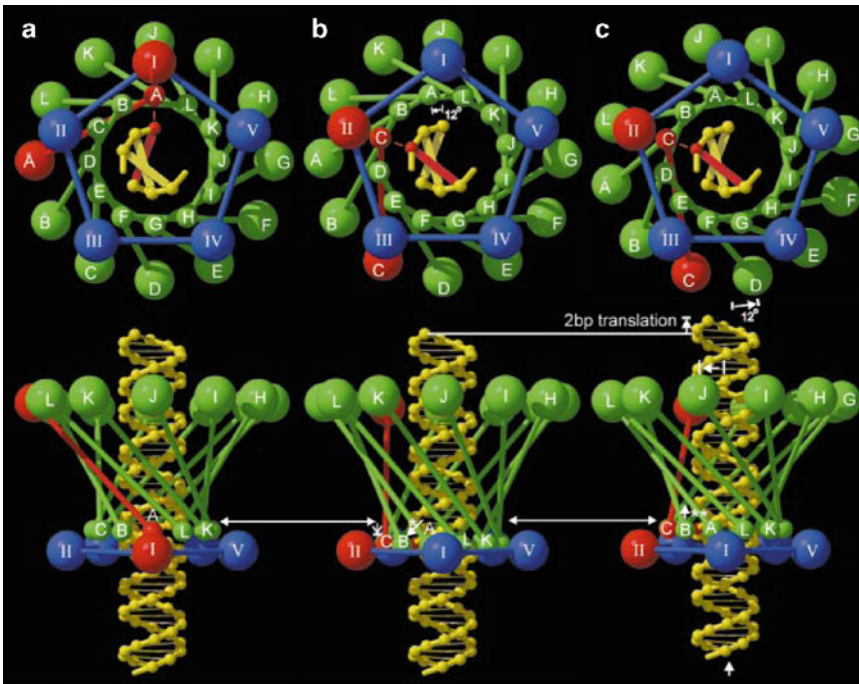


Fig. 23.8 Packaging mechanism based on connector expansion and contraction. A single cycle in the proposed packaging mechanism is depicted. The connector is shown as a set of *small spheres* at the narrow end and a set of *larger spheres* at the wide end joined by a thin tube corresponding to the central helical region. One “active” connector monomer is shown in *red*, the rest in *green*. The pRNA–ATPase complex is shown by a set of four *blue spheres* and one active *red sphere*. A DNA base that is aligned with the active connector monomer is also shown in *red*. The images in the *top* panel are end-on views looking down the wide end of the connector; the bottom images are side views corresponding to each top view. **(a)** The active pRNA–ATPase I interacts with the adjacent connector monomer that is properly aligned with the DNA. **(b)** The narrow end of the connector rotates anticlockwise by 12° , such that the narrow end of monomer C is aligned with ATPase II, the next ATPase to fire; this rotation causes the connector to expand lengthwise by slightly changing the angle of the helices in the central domain (*white arrow with asterisk*). **(c)** In the final step, to return to a low energy state, the wide end of the connector also rotates 12° , causing the connector to contract back to its original conformation, albeit rotated 12° (*white arrow with two asterisks*). Two base pairs of DNA are translated into the phage head during the contraction event. ATPase II and connector (C) are now position for the next cycle. This figure is reproduced from Simpson et al. (2000)

and prohead–pRNA–ATPase complex form a series of concentric structures with 10_1 -, 12_1 -, and 5-fold symmetries, respectively (Simpson et al. 2000, 2001; Morais et al. 2001). In this model, ATP hydrolysis promotes lengthwise expansion and contraction of the connector that is coupled to DNA translocation; the connector essentially acts as an oscillating helical spring that pumps DNA into the prohead (Fig. 23.8). To achieve this ATPase induced expansion, the antiparallel three-helix bundle that is twisted around the central 12_1 -fold axis in the crystal structure is proposed to “straighten” out, thus extending its length and becoming more parallel with the 12_1 -fold axis. The net result is that the central and narrow domains of the connector extend approximately two base pairs and rotate 12° relative to wide end. Upon extension and rotation, one connector monomer becomes aligned to bind DNA. During relaxation, the wide end of the connector follows the narrow end, rotating 12° and causing a contraction of the connector and subsequent translocation of two base pairs of DNA into the prohead. Because of the 10_1 helical symmetry of B-form DNA, translocation of two base pairs is equivalent to a 72° rotation; thus, the DNA would be aligned with the next ATPase in the pentameric ring. Furthermore, because the 12_1 -fold symmetric connector has rotated 12° , it is also aligned with

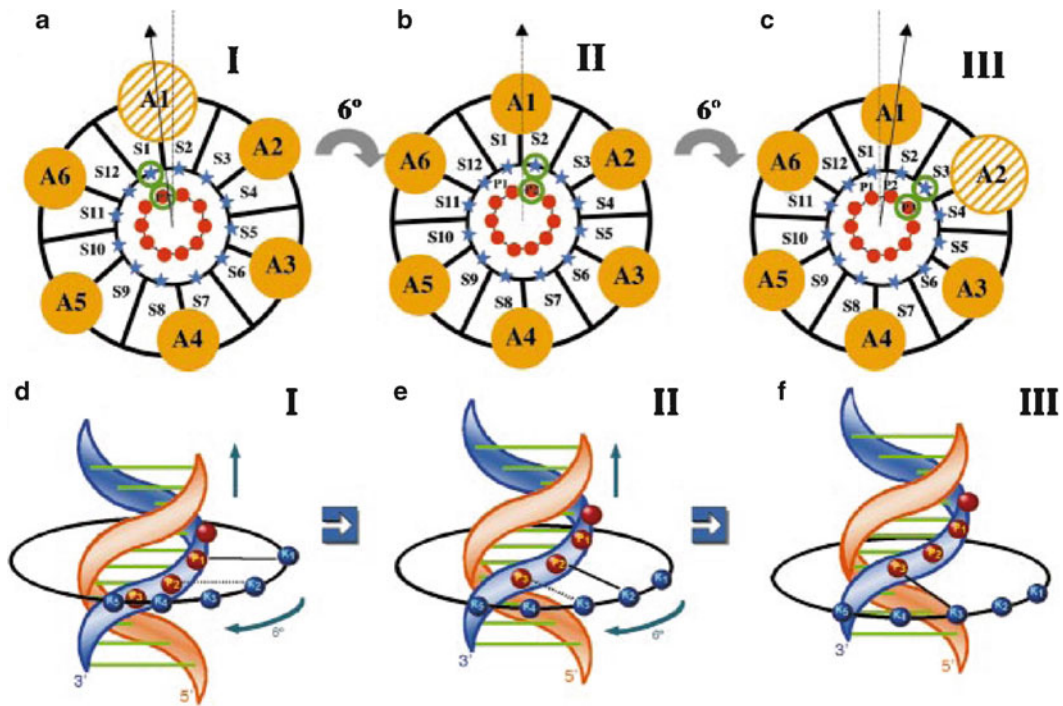


Fig. 23.9 Packaging mechanism based on luminal lysine rings. A schematic of the packaging motor is shown looking down the packaging axis from inside the procapsid. A hexameric ATPase is depicted as six yellow circles numbered A1–A6; the connector is depicted as a disk with individual subunits numbered S1–S12 and corresponding lysines numbered K1–K12 (blue star, top panel; blue ball, bottom panel), and the DNA is represented as red balls (top panel) or a cartoon ribbon (bottom panel). (a) The cycle begins with a lysine aligned with connector subunit S1 and phosphate P1. Upon hydrolysis, ATPase A1 rotates the connector 12° clockwise while the DNA translocates along the connector axis by two base pairs. (b) At 6° rotation, there is a transient interaction between the lysine from S2 and the DNA P2. (c) An additional 6° rotation allows the lysine from S2 to interact with DNA P3. This movement aligns ATPase A2 for the next cycle. In (a) there is a strong interaction between lysine K1 and phosphate P1 (continuous line) and a weak interaction between K2 and P2 because P2 is still below the ring of lysine residues (broken line). In (b), the 6° rotation of the connector allows K2 to tightly bind P2 if DNA is translocated by one base pair. In (c), an additional 6° rotation of the connector and another one base pair movement of DNA aligns K3 with phosphate P3. Hence, the two-base pair step size is accompanied by 12° rotation of the connector. Panels (d–f) correspond to side views of (a–c). For simplicity, only a single DNA chain and one lysine ring are considered. This figure was modified from Guasch et al. (2002)

the DNA and the ATPase for the next mechanochemical cycle. Hence, the connector functions much like a ratchet in this two-state model, and overall rotation of the connector is not necessarily a driving force for translocation, but rather needs to occur to maintain alignment between the 12-fold symmetric connector, the fivefold symmetric ATPase, and the 10_1 -helically symmetric B-form DNA. After all five ATPases have fired, the connector will have rotated 60° , and the DNA will have been translated the length of one pitch of its helix, thus packaging ten base pairs.

Another variation of connector-driven packaging mechanisms envisions that the two lysine rings in the connector lumen alternately grip and release DNA to ensure unidirectional transport into the phage capsid. These two positively charged luminal lysine rings are formed by lysine 200 and lysine 209 (Fig. 23.9). The two rings are separated by approximately 20 \AA , and neighboring lysines within the ring are about 10 \AA apart (Guasch et al. 2002). Considering the symmetry mismatch between the connector and the DNA, it can be shown that the environment of any particular connector monomer is necessarily similar to that of another monomer five subunits around the ring. Hence, if lysine 200

It was suggested that these luminal loops interact directly with the DNA and undergo coordinated sequential conformational changes that are coupled to connector rotation and DNA translocation. Orchestration of these conformational changes gives rise to an “undulating belt” around the DNA with 12 connector–DNA contacts engaged in tight grip. Here again, ATP hydrolysis causes connector rotation, and sequential conformational rearrangements of the luminal loops that translocate DNA and help align the motor for the next cycle. Symmetry mismatches between the connector, and the DNA are also accommodated by loop flexibility. Although there is considerable structural homology between phage connector/portal proteins, it is unlikely that this mechanism could apply to ϕ 29 since the homologous loops in ϕ 29 can be deleted with a minimal effect on packaging activity (S. Grimes, personal communication).

23.5.2 *ATPase-Driven Packaging Models*

In all the connector-based mechanisms described above, the inherent symmetry mismatch between the connector and the DNA requires that motor rotates in order maintain alignment of the motor and the DNA. However, recent experimental evidence from single particle and biochemical experiments indicates that the connector does not rotate during packaging (Baumann et al. 2006; Hugel et al. 2007). In light of this evidence, a connector/ATPase hybrid model was proposed that does not require connector rotation (Morais et al. 2008). In this model, conformational changes in the ATPase generate the power stroke that translocates DNA into the procapsid, and the connector provides a valve-like function holding the DNA in place while the ATPase reloads. The primary role of the connector in this third type of model is to prevent slippage of the DNA via nonspecific connector–DNA interactions and thus maintain unidirectional transport. It is not clear how this valve-like function of the connector would be regulated, but since the pRNA bridges the connector and ATPase, it also might also provide a means of communication between the two; one could imagine, for example, that the ATPase uses the pRNA superhelix as a lever to exert force on the connector.

In further contrast to rotary connector-based mechanisms, the ϕ 29 motor was suggested to function more like a linear motor driven directly by conformational changes in the ATPase. Genetic analysis, sequence alignment, and structural homology between phage ATPases and monomeric helicases suggest that phage DNA packaging motors could be better explained by inchworm-type mechanisms, in which opening and closing of the ATP-binding cleft cause movement of a DNA-binding domain in the ATPase that is coupled to DNA translocation (Draper and Rao 2007; Sun et al. 2007, 2008). In these types of models, the connector functions as little more than a passive portal to provide a conduit for DNA to enter the phage shell.

A detailed mechanism for dsDNA packaging by bacteriophage T4 was recently proposed based on ATPase domain motions inferred from structural analysis of the T4 ATPase (Sun et al. 2008). The T4 ATPase consists of two functionally distinct domains. In addition to an N-terminal ASCE P-loop ATPase fold, the T4 ATPase also possesses a C-terminal nuclease domain necessary for cleavage of the concatenated T4 genome at the termination of packaging. The X-ray crystal structure of the N-terminal domain indicates that it is more closely related to the monomeric helicases than to hexameric ATPases, whereas the C-terminal nuclease domain is a member of the Rnase H/resolvase/integrase superfamily (Sun et al. 2007). The crystal structure of the full-length protein shows that the two domains pack tightly together, with extensive interactions between them including six charge pairs and a hydrophobic core. The structure of the ATPase bound to proheads was analyzed via cryoEM, which, like ϕ 29, showed that five copies of the ATPase bind at the unique connector vertex of the phage. However, unlike ϕ 29, there is very little contact between adjacent subunits, and thus the ATPase forms a pentameric ring only by virtue of binding to a pentameric vertex of the procapsid. In contrast to the crystal structure, where the two domains interact extensively to form a

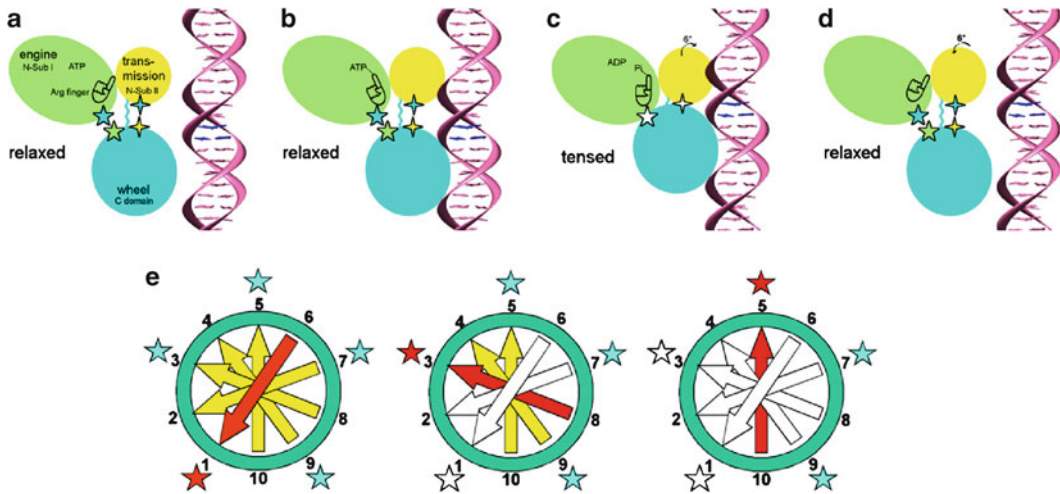


Fig. 23.11 T4 packaging mechanism based on electrostatic interactions. The T4 ATPase N-terminal subdomain I, subdomain II, and C-terminal domain are shown as *green*, *yellow*, and *cyan* ovals, respectively. Electrostatic interactions between the C-terminal domain and the N-terminal subdomains I and II are shown as five- and four-pointed stars, respectively. The linker between the N- and C-terminal domain is shown as a *cyan line*. (a) The C-terminal domain waits to bind DNA. (b) DNA binding causes the C-terminal domain to approach the N-terminal domain. A conformational change in the N-terminal domain inserts the arginine finger into the ATPase active site to trigger hydrolysis. (c) Hydrolysis of ATP causes the N-terminal subdomain II to rotate by about 6° , resulting in alignment of charge pairs and induction of an electrostatic attraction that moves the C-terminal domain and the bound DNA two base pairs. (d) Upon release of ADP and P_i , the C-terminal domain returns to its original position and DNA is released. The next subunit in the pentameric ring is now aligned for DNA translocation. (e) DNA base pairs outside, entering, and inside the procapsid are shown as *yellow*, *red*, and *white* arrows, respectively. Subunits in the pentameric ATPase are shown as *stars*, colored according to whether they are hydrolyzing ATP (*red*), ready to hydrolyze ATP (*blue*), or have already hydrolyzed ATP (*white*). Hydrolysis of ATP at position 1 translocates two base pairs of DNA (*left*), hydrolysis at position 3 translocates two more base pairs (*middle*), and the ATPase at position 5 is now ready to fire. This figure is reproduced from Sun et al. (2008)

relatively closed structure, the cryoEM structure suggested that the two domains are well separated, and thus each ATPase adopts a more open conformation.

Based on these observations, it was proposed that the two different conformations of the ATPase correspond to two distinct states in the DNA packaging pathway (Sun et al. 2008) (Fig. 23.11). The closed conformation present in the crystal structure is believed to correspond to the posttranslocation state, whereas the relaxed open state, obtained via independent fitting of the two ATPase domains into cryoEM density, represents the pretranslocation state. Comparison of the two structures shows that the C-terminal domain in the open state is translated approximately 7 \AA along the packaging axis, similar to the two-base pair step size of the motor. Hence, if the C-terminal domain movements were coupled to binding and release of DNA, then this two state conformational switch might provide a means of translocating the dsDNA genome into the procapsid. It was proposed that the C-terminal domain has a high affinity for DNA in the open state. Upon binding DNA to the C-terminal nuclease domain, a *cis*-arginine finger located at the base of strand β -3 is inserted into the active site to trigger ATP hydrolysis. Hydrolysis is believed to be coupled to conformational change that aligns complementary charges in the N- and C-terminal domains; electrostatic forces then cause the subunit to close, and the DNA-bound C-terminal domain moves 7 \AA along the packaging axis to package two base pairs of dsDNA. Since translocation of two base pairs is equivalent to a 72° rotation, the next subunit in the pentameric ATPase ring will be aligned with DNA for the next translocation cycle (Fig. 23.11b).

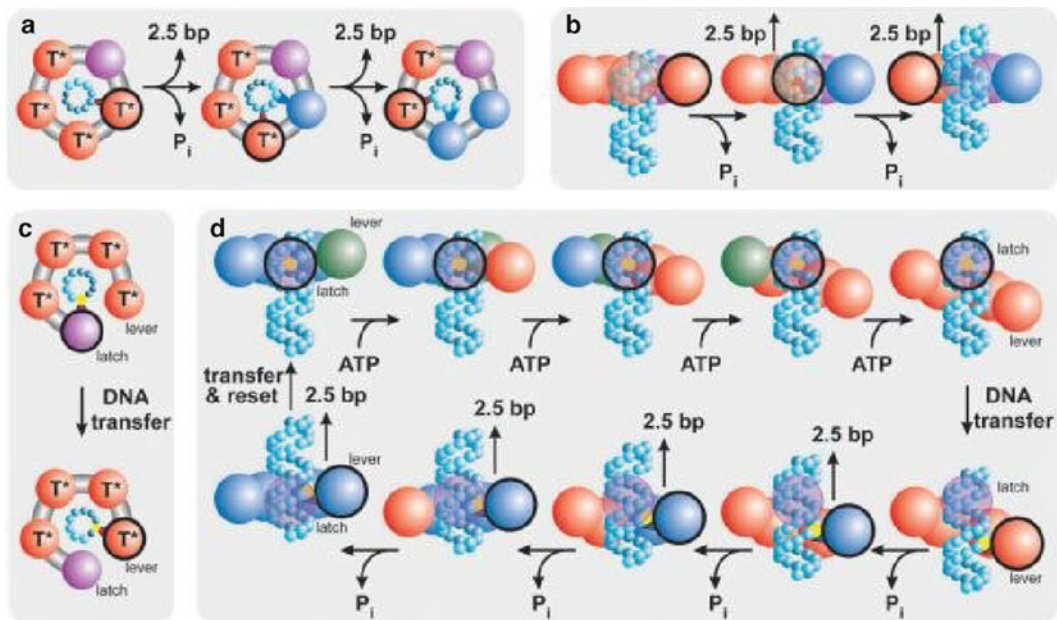


Fig. 23.12 Packaging mechanisms that accommodate a noninteger step size. **(a)** Schematic top view of a translocation model wherein every subunit interacts with the DNA (cyan spheres). The subunit presently interacting with DNA is outlined in black (top view). **(b)** Side view of **(a)** showing how the size of internal conformational changes determines the step size. **(c)** Top view of a schematic of a translocation model in which only two subunits contact the DNA (black outline). **(d)** Side view of the model shown in **(c)**; one subunit maintains contact with the DNA (the latch) while binding of ATP induces relative subunit-subunit rotations in the other subunits that distort the ring. This distortion causes one subunit (the lever) to move along the DNA by ten base pairs. The DNA is then handed from the latch to the lever, and the release of hydrolysis products causes the ring to relax, retracting the lever and the DNA. The DNA contact is then handed back to the latch, the ring resets, and the cycle begins again. Because there are four pairs of subunits, the ring is retracted in four steps, dividing a ten-base pair step into four 2.5-base pair substeps. This figure is reproduced from Moffitt et al. (2009)

23.5.3 Models with Noninteger Step Sizes

Applying the T4 packaging model described above to $\phi 29$ is problematic for a number of reasons. Firstly, the $\phi 29$ ATPase does not have a C-terminal nuclease domain since it packages a unit-length genome. Given the central importance of this domain to the proposed packaging mechanism, it is hard to imagine how a homologous mechanism might work in $\phi 29$. It is possible that the small C-terminal oligonucleotide-binding domain of $\phi 29$ fulfills a function similar to that of the nuclease domain in the T4 ATPase, but additional research is required to refute or confirm this possibility. Secondly, single particle experiments have demonstrated that the $\phi 29$ packaging motor moves in ten-base pair bursts, where each burst consists of four noninteger 2.5-base pair steps (Moffitt et al. 2009). A step size that is a noninteger number of base pairs is incompatible with any mechanism wherein each subunit in the ATPase ring makes specific and interactions with the DNA. Thus, resolving the mechanistic details of $\phi 29$ DNA translocation may be more complex as it requires integration of a noninteger step size and the reconciliation of a four-step cycle within the structural framework of a pentameric motor.

Two alternative models incorporating four 2.5-base pair steps within a pentameric motor accompanied the initial report of a noninteger step size (Moffitt et al. 2009). The first model (Fig. 23.12a, b) proposes that the motor does not make any specific contacts with the DNA, but instead generates

force via nonspecific steric interactions with the major groove of the DNA. Hence, the step size of the motor need not coincide with the helical symmetry of the DNA and is thus solely determined by the extent of the conformational changes that generate the power stroke. To accommodate four translocation steps, this model assumes asymmetry within the pentameric ring in that one of the five subunits is different from the other four. Because the nucleotide-free state has a low affinity for DNA (Chemla et al. 2005), the unique subunit may be required to hold the DNA at the end of each cycle, preventing slippage ATP during the subsequent dwell phase.

In the second model (Fig. 23.12c, d), specific interactions between the motor and the DNA occur but only for a select subset of the subunits. In the proposed model, two subunits make specific contact with the DNA, but a different number of DNA-binding subunits can also be accommodated in the same general framework. The DNA is pumped into the procapsid via an “inchworm-like” movement of these two subunits that is driven by sequential distortions around the ring. In short, the arrangement of the five ATPase subunits in the motor oscillates between a closed planar ring and open single turn helical structures. In this way, the ten-base pair burst can be divided into noninteger steps. This mechanism also suggests how a four-step cycle arises from a pentameric motor; one subunit interface sees a unique environment due to the accumulated distortion of the other four subunits, perhaps resulting in its inactivation. The proposed motor structures are similar to the quaternary structures of other ring ATPases, but were not implicated as part of any translocation mechanism (Singleton et al. 2000; Skordalakes and Berger 2003).

A comprehensive “push-and-roll” model was recently proposed that explicitly couples ATP hydrolysis with the coordinated conformational changes that drive packaging (Fig. 23.13). Although this model is primarily based on results from single particle experiments, it is also compatible with available structural, genetic, and biochemical data (Yu et al. 2010). In this model, genome encapsidation is not simply the result of translating DNA along the central axis of the phage, but rather requires that the DNA also roll around the inner lumen of the ATPase ring; hence, packaging combines a “push” and a “roll.” Based on geometric arguments and packaging simulations, it was demonstrated that this push-and-roll mechanism can explain how a planar pentameric packaging motor translocates the DNA in bursts of four 2.5-base pair power strokes.

Lacking a high-resolution structure of the ATPase, the authors constructed a homology model based on the P4 ssRNA packaging motor in bacteriophage ϕ 12 (Mancini et al. 2004; Lisal and Tuma 2005). Predicated on this model, it was suggested that during each power stroke, a positively charged luminal lever loop from one subunit is electrostatically steered to the DNA backbone as in P4 (Fig. 23.13a). Upon engagement with the DNA, the lever then pushes sterically, orthogonal to the backbone axis; assuming B-form DNA geometry, this push will cause the right-handed DNA helix not only to be translated in the packaging direction but also to be rotated in a left-handed direction (Fig. 23.13b, c). Thus, the DNA movement consists of two steps, a vertical step along the packaging direction and a horizontal step in the tangential direction of the DNA cross section (Fig. 23.13b). It is not entirely clear how ATP hydrolysis is coupled to force generation by the luminal levers, but it was proposed that, similar to the F_1 -ATPase, movement of loops in the active site is coupled to twisting of the central β -sheet, and that this deformation is stored as elastic energy that can be used by the luminal levers to apply a force on the DNA (Abrahams et al. 1994; Kabaleeswaran et al. 2006; Yu et al. 2010).

The proposed reaction scheme begins with one subunit bound to DNA (Fig. 23.14). Assuming that the subunit initially interacting with DNA is in an ATP-bound state, then the electrostatic coupling between the lever and the DNA will be tight (Chemla et al. 2005; Moffitt et al. 2009), thus holding the DNA in place and preventing slippage. It has been shown that affinity for DNA decreases in the ADP-bound state (Chemla et al. 2005). Hence the lever affinity for DNA would decrease upon ATP hydrolysis at end of the power stroke that accompanies P_i release (Chemla et al. 2005) causing the DNA to roll to the lever on the adjacent subunit. If the next subunit in the ring is in the ATP-bound

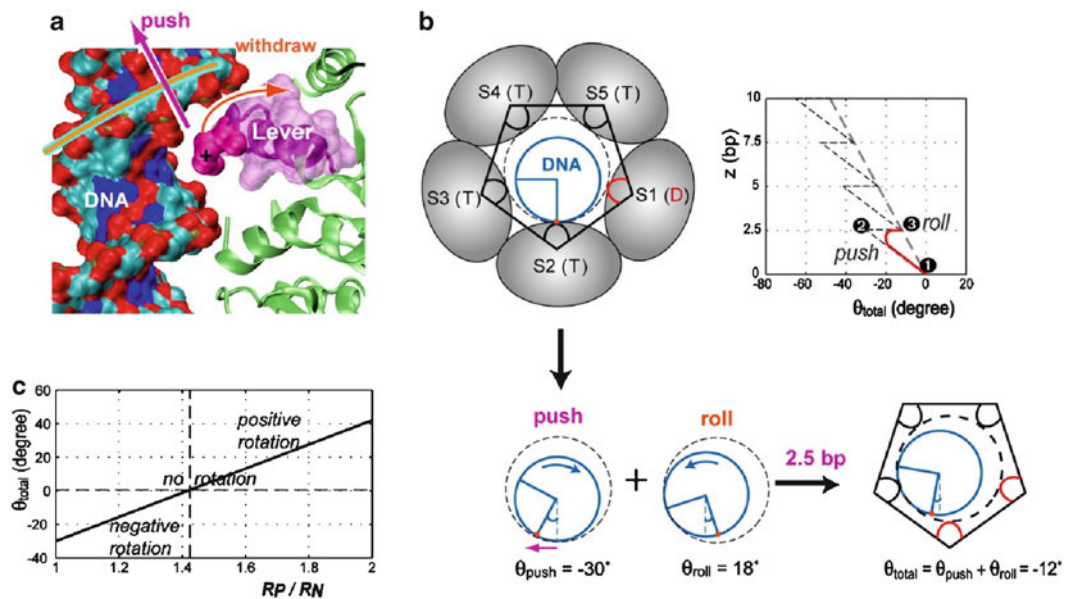


Fig. 23.13 Push-and-roll packaging mechanism. (a) Model of the proposed ATPase (green ribbon) luminal lever (magenta isosurface) interacting with DNA; this model is based on the structure of the ssRNA packaging motor P4. (b) DNA movement generated by the power stroke. A schematic top view of the DNA and the motor ring is shown (left) with ATPase subunits labeled S1–5, and the hydrolysis state of each is in parentheses: T, ATP-bound; D, ADP-bound. During the power stroke (T \rightarrow D), the motor lever pushes the DNA up by 2.5 base pairs (out of the paper) and rotates it by $\theta_{\text{push}} \approx -30^\circ$. DNA rolls to the next subunit upon disengaging the lever resulting in an additional DNA rotation of $\theta_{\text{roll}} = +18^\circ$ (when $R_p/R_n = 5/4$, see below); hence, each power stroke results in a total DNA rotation of -12° . A coordinate trace of a point on the DNA during packaging is shown on the right for the special case when the ratio of the radial cross section of the DNA (R_p) to the radius of the lumen of the ATPase ring (R_n) is $5/4$: the DNA is rotated by θ_{push} (1 \rightarrow 2) + θ_{roll} (2 \rightarrow 3) = $\theta_{\text{total}} \approx -12^\circ$ for every 2.5 base pairs translocated. The actual trajectory is likely to be more like the red line where the rolling commences before the push has completed. (c) The net DNA rotation predicted for each power stroke depends on the relative size of the motor ring cross section to that of the DNA (R_p/R_n): a negative (underwinding) rotation of the DNA occurs if $R_p/R_n < 1.4$, there is no rotation when $R_p/R_n \approx 1.4$, and a positive (overwinding) DNA rotation when $R_p/R_n > 1.4$. This figure was adapted from Yu et al. (2010)

state, it has a high affinity for DNA and can thus accept the incoming DNA. The proposed rolling motion thus carries the DNA from one subunit to the next. Similar to P4, each power stroke also causes the arginine finger in the active subunit to be inserted into the active site of its neighbor, thus promoting hydrolysis in the next subunit around the ring (Lisal and Tuma 2005; Kainov et al. 2008). However, because ADP release is slow, when it comes time for the fifth subunit to fire, subunit 1, the next subunit in a pentameric ring is still bound to ADP, thus having a low affinity for DNA and rendering it incapable of accepting the rolling DNA substrate. As a result, at the end of every four power strokes, the cycle pauses, constituting a dwell phase before the fifth subunit can fire, during which time four ATPs are loaded into the catalytic sites. The next burst phase of four power strokes starts once spontaneous ATP hydrolysis takes place in the fifth site without insertion of an arginine finger. A necessary consequence of this mechanism is that the order of subunit firing changes each cycle. For example, if the subunit firing order of a particular cycle is 1, 2, 3, 4 then the firing order in the subsequent cycle will be 5, 1, 2, 3, and so on. In summary, the push-and-roll model naturally accounts for the observed frequency and duration of the dwell phase, explains how a four 2.5-base pair step cycle arises from a planar pentameric motor, and provides a new perspective on how a multimeric ATPase might transport DNA.

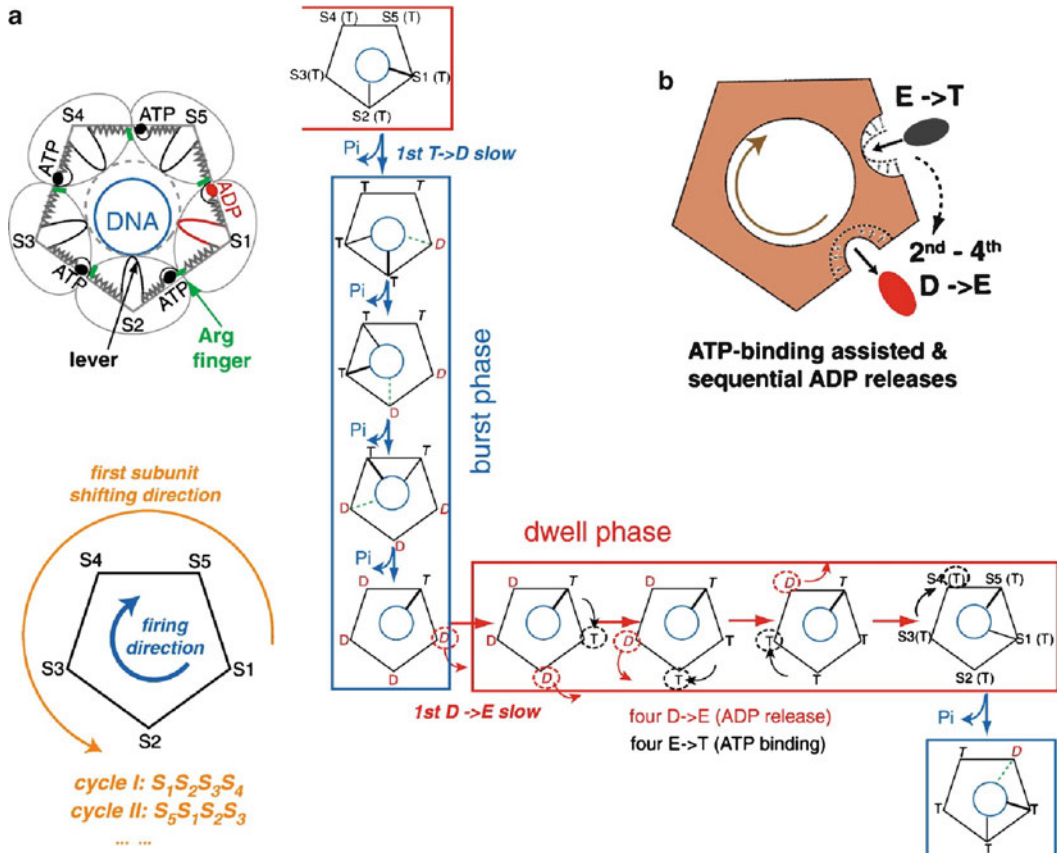


Fig. 23.14 Dominant reaction scheme of the push-and-roll model. **(a)** A schematic top view of the ATPase pentamer (*upper left*) and the direction of power stroke firing within a cycle versus the direction of the first subunit shifts in subsequent cycles (*lower left*) proposed in the model. **(b)** A cooperative ADP release mechanism would link four ADP release and ATP-loading events sequentially around the motor ring. The reaction cycle consists of a burst phase and a dwell phase. The burst phase includes four sequential 2.5-base pair power strokes at four consecutive catalytic sites. During the dwell phase, four consecutive ATP loadings and several non-ATP-binding events occur. A power stroke accompanies the T → D transition (Pi release) when the subunit is bound to DNA (*thick continuous line*). Each power stroke requires the next subunit (*thin continuous line*) is in the ATP-bound state (T) to receive the "rolling" DNA when it arrives at the end of the previous power stroke. After four rapid sequential power strokes in the burst phase, the motor must pause in the dwell phase because the next subunit is still in the low DNA affinity ADP-bound state (D). The first ADP release in the dwell phase is slow, but the following ADP releases are rapid if ATP binds quickly and accelerates ADP release at the next site (see **b**). Although the first power stroke after the four ATPs are loaded is also slow, the ensuing power strokes (second to fourth T → D) happen very quickly in the next burst phase since insertion of an arginine finger from the preceding subunit promotes catalysis. This figure was adapted from Yu et al. (2010)

23.6 Summary and Prospectus

The $\phi 29$ motor is arguably the best characterized DNA packaging motor, and new results regarding the mechanism of the motor can be interpreted in light of a rich background of sophisticated genetic, biochemical, structural, and single particle data. Research on the $\phi 29$ packaging motor is thus poised to move the DNA packaging field from speculation regarding the nature of the force-generating step to a detailed description of the coordinated conformational changes that drive DNA packaging. As our description of the mechanochemical cycle of the $\phi 29$ packaging motor becomes more and more

complete, it is worth asking which aspects of the motor mechanism and assembly pathway are common to all dsDNA viruses and which are uniquely adapted to challenges faced by ϕ 29. Ignoring close relatives, ϕ 29 shares very little sequence similarity with other viruses or bacteriophages. Furthermore, interpreting bacteriophage relationships in the context of Darwinian evolution is complicated by the considerable influence of horizontal gene transfer on phage evolution (Hendrix 2002). Nonetheless, recent structural data shows that the tailed dsDNA bacteriophages and herpesviruses utilize the same basic protein folds for their capsid (Zhou et al. 2000; Jiang et al. 2003, 2008; Fokine et al. 2005; Morais et al. 2005; Agirrezabala et al. 2007; Dai et al. 2010) and portal proteins (Simpson et al. 2000; Orlova et al. 2003; Lebedev et al. 2007; Lander et al. 2009; Chen et al. 2011), whereas bioinformatics analysis indicates that the virus/phage-encoded ATPases that power dsDNA packaging all adopt the characteristic the ASCE superfamily fold (Mitchell et al. 2002; Burroughs et al. 2007). Hence, regardless of whether vertical or horizontal gene transfer has dominated phage evolution, it seems clear that all tailed phages utilize similar macromolecular components to package their DNA, pointing to a common mechanism of genome encapsidation.

There are, however, features of the ϕ 29 packaging motor that are unique to ϕ 29. These include the requirement for a phage-encoded pRNA, packaging of a unit-length genome, and the covalent attachment of gp3 to the ends of the ϕ 29 genome. Clearly, other phages have evolved such that the role of the pRNA is carried out by proteins, or domains within proteins. Although it has been speculated that the pRNA is a vestige from an RNA-dominated world, it is not clear whether the pRNA evolved to replace a protein component, or whether the analogous protein components in other phages evolved to replace a pRNA. Similarly, regarding the requirement for gp3, it is not clear if the ϕ 29 motor evolved to accommodate a polymerase that requires a protein primer (Blanco and Salas 1996), or if the genome replication machinery evolved to accommodate the specificity requirements of the packaging motor. Furthermore, because ϕ 29 replicates nonconcatenated unit-length DNA, the ϕ 29 ATPase does not require or possess any nuclease activity, and thus the ϕ 29 motor must also utilize a different mechanism to terminate packaging than phages that encapsidate a concatenated multimeric genome. Indeed, this relative simplicity is a feature that makes ϕ 29 an excellent model system for studying packaging; in addition to utilizing motor components that represent the bare minimum for a functional packaging motor, ϕ 29 also carries out a simplified, and thus more easily dissected, packaging reaction.

As described above, genetic studies have identified all the components of the ϕ 29 packaging motors (Grimes et al. 2002); structural, biochemical, and bioinformatics analyses have identified and characterized specific residues within these components that are critical for motor function, and biochemical, structural, and single particle analyses suggest ways that these components operate in a highly coordinated manner (Chemla et al. 2005; Aathavan et al. 2009; Moffitt et al. 2009; Yu et al. 2010). Yet, despite the wealth of accumulated data on genome packaging in ϕ 29 and other dsDNA viruses, several aspects of DNA packaging remain obscure. Most importantly, it is still not known exactly how these molecular components interact with each other in a coordinated fashion to drive genome translocation.

As a first step toward understanding the conformational changes that drive packaging, it is necessary to determine high-resolution structures of the entire motor complex in different steps of the mechanochemical cycle. These structures would help identify the interactions between specific amino/nucleic acids within and between motor components. Since these contacts likely guide both motor assembly and coordination, this structure would provide valuable information regarding these two critical aspects of motor function. Due to the size and complexity of a multisubunit packaging motor, the most tractable approach to obtaining its atomic structure is to determine high-resolution structures of isolated motor components by X-ray crystallography and then fit these structures into subnanometer resolution cryoEM reconstructions of the intact motor complexes trapped in different stages of the packaging pathway.

As discussed above, atomic resolution structures of the $\phi 29$ connector and the prohead-binding domain of the $\phi 29$ pRNA have been determined (Simpson et al. 2000; Ding et al. 2011), and thus only the atomic structure of the ATPase remains unknown. Using cryoEM, the structure of a $\phi 29$ connector–pRNA complex has been determined to sub-nanometer resolution (Morais et al., in preparation), and thus it is likely that a similar complex that includes the ATPase can be reconstructed to a comparable resolution. At this resolution, secondary structural elements are easily recognized, and atomic resolution structures can be fit with great accuracy and reliability (Rossmann et al. 2005). Hence, should an atomic resolution structure of the ATPase become available, it will likely be possible to quantitatively fit the structures of the ATPase, connector, and pRNA into their corresponding densities in a cryoEM reconstruction of the entire motor complex for at least the ground state of the motor. The resulting pseudoatomic structure would map amino acid residues at component interfaces, identify interactions that guide motor assembly and coordinate its activity, suggest biochemical experiments to test specific packaging hypotheses, and serve as a reference for determining the conformational changes that drive packaging.

As with any mechanical device, the $\phi 29$ DNA packaging motor cannot be fully appreciated without some sense of how the different components and domains within components move relative to one another. The ability to visualize the packaging motor during DNA translocation would provide invaluable insight into the mechanism of packaging by showing how the different motor components move, interact with each other, and interact with the DNA. Direct visualization of conformational change has been extremely challenging. Although NMR provides information regarding molecular dynamics, it is limited to molecules with molecular weights less than ~ 30 kDa. Conventional X-ray crystallography and cryoEM are not restricted by size limitations, but produce only static images representing the average of a large ensemble of individual molecules. Thus, unless dynamic behavior can be synchronized in all molecules, it is not possible to simultaneously align all the moving parts in different individual molecules for averaging. Hence, macromolecules and their complexes are typically stalled and then imaged at different points in their enzymatic or mechanical cycle, and conformational changes are then inferred by interpolating between distinct static structures. For example, addition of γ -S ATP or AMP-PNP to $\phi 29$ particles would trap the motor in the ATP-bound, prehydrolysis state, whereas stalling with ADP-vanadate mimics the pentavalent transition state of ATP hydrolysis; comparing these structures would reveal any significant conformational differences between substrate and transition state-bound intermediates. An advantage of this approach is that by stopping packaging at defined points, it is possible to correlate specific structures with discrete steps in the mechanochemical pathway.

An alternative and particularly enticing approach for visualizing the conformational changes that drive packaging is to flash-freeze actively packaging particles, image them using cryoEM, and then computationally sort and reconstruct different motor conformations present in the sample. Upon freezing, motors in different particles would be trapped in different stages of the mechanochemical cycle, and thus the relative positions of the motor's moving parts will vary in different particles imaged. Until recently, this conformational heterogeneity would have been an obstacle in cryoEM analysis, i.e., unless dynamic behavior could be synchronized in all molecules (as described above), it would be impossible to simultaneously align all the moving parts in different individual molecules for image reconstruction. However, recent advances in image processing have demonstrated that it is possible to computationally sort different conformational species present in cryoEM micrographs and extract meaningful biological information from their differences (Leschziner and Nogales 2007). In these approaches, particles are initially classified according to direction of view by model-based projection matching and/or other well-established methods (Baker et al. 1999; Ludtke et al. 1999; van Heel et al. 2000; Frank et al. 2002). Each class of particles would then be analyzed for the presence of subclasses corresponding to different conformations of the various motor components.

Of course, structural data alone is not sufficient for a complete description of the motor, and a comprehensive understanding of the coordinated mechanochemical cycle of the bacteriophage

ϕ 29 DNA packaging motor will require a highly integrated genetic, biochemical, biophysical, computational, as well as structural approach. In addition to identifying mutations that completely abolish packaging, it would be informative to have a library of mutants that modulate, rather than eliminate, DNA packaging. Phage particles with altered activities could then be subjected to additional analysis via structural or single particle techniques. In this way, it may be possible to unravel more subtle aspects of the motor function and begin the task of determining the complete sequence of conformational changes that constitute the mechanochemical cycle of the packaging motor.

Insight provided by current and future results on ϕ 29 and other model systems for genome encapsidation will inform us as to the fundamental mechanism of packaging in dsDNA viruses, as well as the operation of the broader class of ATP-driven, force-generating ring motors from which they evolved. Hence, future research on the ϕ 29 DNA packaging motor promises to describe principles of the mechanochemistry of a broad range of molecular motors, provide insight into how these principles are optimized in viral packaging motors to generate large forces, inspire novel therapeutics that selectively target genome packaging to inhibit medically relevant dsDNA viruses, and stimulate work in the developing field of rational design of molecular machines.

Acknowledgments I am indebted to Dr. Michael Rossmann, who first introduced me to ϕ 29, mentored much of my own research on the DNA packaging motor, and always graciously shared his considerable wisdom. I would also like to thank my collaborators: Drs. Dwight Anderson, Paul Jardine, Dr. Jaya Koti and Shelley Grimes for many thoughtful discussions regarding the mechanism of genome packaging in ϕ 29, as well as for their invaluable contributions to the field of DNA packaging by dsDNA bacteriophages.

References

- Aathavan K, Politzer AT, Kaplan A, Moffitt JR, Chemla YR, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2009) Substrate interactions and promiscuity in a viral DNA packaging motor. *Nature* 461(7264):669–673
- Abrahams JP, Leslie AG, Lutter R, Walker JE (1994) Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* 370(6491):621–628
- Agirrezabala X, Velazquez-Muriel JA, Gomez-Puertas P, Scheres SH, Carazo JM, Carrascosa JL (2007) Quasi-atomic model of bacteriophage ϕ 7 procapsid shell: insights into the structure and evolution of a basic fold. *Structure* 15(4):461–472
- Anderson DL, Reilly BE (1993) Morphogenesis of bacteriophage ϕ 29. In: Hoch JA, Losick R, Sonenshein AL (eds) *Bacillus subtilis* and other gram positive bacteria: physiology, biochemistry and molecular genetics. ASM Press, Washington, DC, pp 859–867
- Atz R, Ma S, Gao J, Anderson DL, Grimes S (2007) Alanine scanning and Fe-BABE probing of the bacteriophage ϕ 29 prohead RNA-connector interaction. *J Mol Biol* 369(1):239–248
- Bailey S, Wichitwechkarn J, Johnson D, Reilly BE, Anderson DL, Bodley JW (1990) Phylogenetic analysis and secondary structure of the *Bacillus subtilis* bacteriophage RNA required for DNA packaging. *J Biol Chem* 265(36):22365–22370
- Baker TS, Olson NH, Fuller SD (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol Mol Biol Rev* 63(4):862–922, table of contents
- Baumann RG, Mullaney J, Black LW (2006) Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. *Mol Microbiol* 61(1):16–32
- Bjornsti MA, Reilly BE, Anderson DL (1983) Morphogenesis of bacteriophage ϕ 29 of *Bacillus subtilis*: oriented and quantized in vitro packaging of DNA protein gp3. *J Virol* 45(1):383–396
- Bjornsti MA, Reilly BE, Anderson DL (1984) Bacteriophage ϕ 29 proteins required for in vitro DNA-gp3 packaging. *J Virol* 50(3):766–772
- Blanco L, Salas M (1996) Relating structure to function in ϕ 29 DNA polymerase. *J Biol Chem* 271(15):8509–8512
- Braig K, Menz RI, Montgomery MG, Leslie AG, Walker JE (2000) Structure of bovine mitochondrial F1-ATPase inhibited by Mg(2+) ADP and aluminium fluoride. *Structure* 8(6):567–573
- Burroughs AM, Iyer LM, Aravind L (2007) Comparative genomics and evolutionary trajectories of viral ATP dependent DNA-packaging systems. *Genome Dyn* 3:48–65

- Camacho A, Jimenez F, De La Torre J, Carrascosa JL, Mellado RP, Vasquez C, Vinuela E, Salas M (1977) Assembly of *Bacillus subtilis* phage phi29. 1. Mutants in the cistrons coding for the structural proteins. *Eur J Biochem* 73(1):39–55
- Chemla YR, Aathavan K, Michaelis J, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2005) Mechanism of force generation of a viral DNA packaging motor. *Cell* 122(5):683–692
- Chen C, Zhang C, Guo P (1999) Sequence requirement for hand-in-hand interaction in formation of RNA dimers and hexamers to gear phi29 DNA translocation motor. *RNA* 5(6):805–818
- Chen C, Sheng S, Shao Z, Guo P (2000) A dimer as a building block in assembling RNA. A hexamer that gears bacterial virus phi29 DNA-translocating machinery. *J Biol Chem* 275(23):17510–17516
- Chen DH, Baker ML, Hryc CF, DiMaio F, Jakana J, Wu W, Dougherty M, Haase-Pettingell C, Schmid MF, Jiang W, Baker D, King JA, Chiu W (2011) Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci USA* 108(4):1355–1360
- Cohen DN, Sham YY, Haugstad GD, Xiang Y, Rossmann MG, Anderson DL, Popham DL (2009) Shared catalysis in virus entry and bacterial cell wall depolymerization. *J Mol Biol* 387(3):607–618
- Dai W, Hodes A, Hui WH, Gingery M, Miller JF, Zhou ZH (2010) Three-dimensional structure of tropism-switching Bordetella bacteriophage. *Proc Natl Acad Sci USA* 107(9):4347–4352
- Ding F, Lu C, Zhao W, Rajashankar KR, Anderson DL, Jardine PJ, Grimes S, Ke A (2011) Structure and assembly of the essential RNA ring component of a viral DNA packaging motor. *Proc Natl Acad Sci USA* 108:7357–7362
- Draper B, Rao VB (2007) An ATP hydrolysis sensor in the DNA packaging motor from bacteriophage T4 suggests an inchworm-type translocation mechanism. *J Mol Biol* 369(1):79–94
- Earnshaw WC, Casjens SR (1980) DNA packaging by the double-stranded DNA bacteriophages. *Cell* 21(2):319–331
- 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(20):7163–7168
- Frank J, Wagenknecht T, McEwen BF, Marko M, Hsieh CE, Mannella CA (2002) Three-dimensional imaging of biological complexity. *J Struct Biol* 138(1–2):85–91
- Fu CY, Prevelige PE Jr (2009) In vitro incorporation of the phage Phi29 connector complex. *Virology* 394(1):149–153
- Fuller DN, Rickgauer JP, Jardine PJ, Grimes S, Anderson DL, Smith DE (2007) Ionic effects on viral DNA packaging and portal motor function in bacteriophage phi 29. *Proc Natl Acad Sci USA* 104(27):11245–11250
- Garver K, Guo P (1997) Boundary of pRNA functional domains and minimum pRNA sequence requirement for specific connector binding and DNA packaging of phage phi29. *RNA* 3(9):1068–1079
- Geourjon C, Orelle C, Steinfeld E, Blanchet C, Deleage G, Di Pietro A, Jault JM (2001) A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends Biochem Sci* 26(9):539–544
- Grimes S, Anderson D (1989) In vitro packaging of bacteriophage phi 29 DNA restriction fragments and the role of the terminal protein gp3. *J Mol Biol* 209(1):91–100
- Grimes S, Anderson D (1990) RNA dependence of the bacteriophage phi 29 DNA packaging ATPase. *J Mol Biol* 215(4):559–566
- Grimes S, Jardine PJ, Anderson D (2002) Bacteriophage phi 29 DNA packaging. *Adv Virus Res* 58:255–294
- Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, Sousa N, Carrascosa JL, Coll M (2002) Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle. *J Mol Biol* 315(4):663–676
- Guo P, Grimes S, Anderson D (1986) A defined system for in vitro packaging of DNA-gp3 of the *Bacillus subtilis* bacteriophage phi 29. *Proc Natl Acad Sci USA* 83(10):3505–3509
- Guo P, Peterson C, Anderson D (1987a) Initiation events in in-vitro packaging of bacteriophage phi 29 DNA-gp3. *J Mol Biol* 197(2):219–228
- Guo P, Peterson C, Anderson D (1987b) Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage phi 29. *J Mol Biol* 197(2):229–236
- Guo PX, Erickson S, Xu W, Olson N, Baker TS, Anderson D (1991a) Regulation of the phage phi 29 prohead shape and size by the portal vertex. *Virology* 183(1):366–373
- Guo PX, Rajagopal BS, Anderson D, Erickson S, Lee CS (1991b) sRNA of phage phi 29 of *Bacillus subtilis* mediates DNA packaging of phi 29 proheads assembled in *Escherichia coli*. *Virology* 185(1):395–400
- Guo P, Zhang C, Chen C, Garver K, Trottier M (1998) Inter-RNA interaction of phage phi29 pRNA to form a hexameric complex for viral DNA transportation. *Mol Cell* 2(1):149–155
- Hagen EW, Reilly BE, Tosi ME, Anderson DL (1976) Analysis of gene function of bacteriophage phi 29 of *Bacillus subtilis*: identification of cistrons essential for viral assembly. *J Virol* 19(2):501–517
- Harding NE, Ito J, David GS (1978) Identification of the protein firmly bound to the ends of bacteriophage phi 29 DNA. *Virology* 84(2):279–292
- Harris S, Schroeder SJ (2010) Nuclear magnetic resonance structure of the prohead RNA E-loop hairpin. *Biochemistry* 49(29):5989–5997

- Hendrix RW (1978) Symmetry mismatch and DNA packaging in large bacteriophages. *Proc Natl Acad Sci USA* 75(10):4779–4783
- Hendrix RW (2002) Bacteriophages: evolution of the majority. *Theor Popul Biol* 61(4):471–480
- Hugel T, Michaelis J, Hetherington CL, Jardine PJ, Grimes S, Walter JM, Falk W, Anderson DL, Bustamante C (2007) Experimental test of connector rotation during DNA packaging into bacteriophage ϕ 29 capsids. *PLoS Biol* 5(3):e59
- Hung LW, Wang IX, Nikaido K, Liu PQ, Ames GF, Kim SH (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* 396(6712):703–707
- Iyer LM, Leipe DD, Koonin EV, Aravind L (2004a) Evolutionary history and higher order classification of AAA+ATPases. *J Struct Biol* 146(1–2):11–31
- Iyer LM, Makarova KS, Koonin EV, Aravind L (2004b) Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res* 32(17):5260–5279
- Jiang W, Li Z, Zhang Z, Baker ML, Prevelige PE Jr, Chiu W (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* 10(2):131–135
- Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W (2008) Backbone structure of the infectious epsilon 15 virus capsid revealed by electron cryomicroscopy. *Nature* 451(7182):1130–1134
- Kabaleeswaran V, Puri N, Walker JE, Leslie AG, Mueller DM (2006) Novel features of the rotary catalytic mechanism revealed in the structure of yeast F1 ATPase. *EMBO J* 25(22):5433–5442
- Kainov DE, Mancini EJ, Telenius J, Lisal J, Grimes JM, Bamford DH, Stuart DI, Tuma R (2008) Structural basis of mechanochemical coupling in a hexameric molecular motor. *J Biol Chem* 283(6):3607–3617
- Kamtekar S, Berman AJ, Wang J, Lazaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2006) The ϕ 29 DNA polymerase: protein-primer structure suggests a model for the initiation to elongation transition. *EMBO J* 25(6):1335–1343
- Kellenberger E, Sechaud J, Ryter A (1959) Electron microscopical studies of phage multiplication. IV. The establishment of the DNA pool of vegetative phage and the maturation of phage particles. *Virology* 8:478–498
- Koti JS, Morais MC, Rajagopal R, Owen BA, McMurray CT, Anderson DL (2008) DNA packaging motor assembly intermediate of bacteriophage ϕ 29. *J Mol Biol* 381(5):1114–1132
- Lander GC, Khayat R, Li R, Prevelige PE, Potter CS, Carragher B, Johnson JE (2009) The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. *Structure* 17(6):789–799
- Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova EV, Turner J, Dodson EJ, Tavares P, Antson AA (2007) Structural framework for DNA translocation via the viral portal protein. *EMBO J* 26(7):1984–1994
- Lee TJ, Guo P (2006) Interaction of gp16 with pRNA and DNA for genome packaging by the motor of bacterial virus ϕ 29. *J Mol Biol* 356(3):589–599
- Lee JY, Yang W (2006) UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* 127(7):1349–1360
- Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* 317(1):41–72
- Leschziner AE, Nogales E (2007) Visualizing flexibility at molecular resolution: analysis of heterogeneity in single-particle electron microscopy reconstructions. *Annu Rev Biophys Biomol Struct* 36:43–62
- Lisal J, Tuma R (2005) Cooperative mechanism of RNA packaging motor. *J Biol Chem* 280(24):23157–23164
- Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97
- 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(6):743–755
- Mitchell MS, Matsuzaki S, Imai S, Rao VB (2002) Sequence analysis of bacteriophage T4 DNA packaging/terminase genes 16 and 17 reveals a common ATPase center in the large subunit of viral terminases. *Nucleic Acids Res* 30(18):4009–4021
- Moffitt JR, Chemla YR, Aathavan K, Grimes S, Jardine PJ, Anderson DL, Bustamante C (2009) Intersubunit coordination in a homomeric ring ATPase. *Nature* 457(7228):446–450
- Morais MC, Baker AS, Dunaway-Mariano D, Allen KN (2000) Crystallization and preliminary crystallographic analysis of phosphonoacetaldehyde hydrolase. *Acta Crystallogr D Biol Crystallogr* 56(Pt 2):206–209
- Morais MC, Tao Y, Olson NH, Grimes S, Jardine PJ, Anderson DL, Baker TS, Rossmann MG (2001) Cryoelectron-microscopy image reconstruction of symmetry mismatches in bacteriophage ϕ 29. *J Struct Biol* 135(1):38–46
- Morais MC, Kanamaru S, Badasso MO, Koti JS, Owen BA, McMurray CT, Anderson DL, Rossmann MG (2003) Bacteriophage ϕ 29 scaffolding protein gp7 before and after prohead assembly. *Nat Struct Biol* 10(7):572–576
- 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 ϕ 29. *Mol Cell* 18(2):149–159
- Morais MC, Koti JS, Bowman VD, Reyes-Aldrete E, Anderson DL, Rossmann MG (2008) Defining molecular and domain boundaries in the bacteriophage ϕ 29 DNA packaging motor. *Structure* 16(8):1267–1274

- Mosharrafa ET, Schachtele CF, Reilly BE, Anderson DL (1970) Complementary Strands of Bacteriophage phi29 Deoxyribonucleic Acid: Preparative Separation and Transcription Studies. *J Virol* 6(6):855–864
- Muller DJ, Engel A, Carrascosa JL, Velez M (1997) The bacteriophage phi29 head-tail connector imaged at high resolution with the atomic force microscope in buffer solution. *EMBO J* 16(10):2547–2553
- Nadanaciva S, Weber J, Wilke-Mounts S, Senior AE (1999) Importance of F1-ATPase residue alpha-Arg-376 for catalytic transition state stabilization. *Biochemistry* 38(47):15493–15499
- Ogura T, Whiteheart SW, Wilkinson AJ (2004) Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. *J Struct Biol* 146(1–2):106–112
- 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(6):1255–1262
- Petrov AS, Harvey SC (2008) Packaging double-helical DNA into viral capsids: structures, forces, and energetics. *Biophys J* 95(2):497–502
- Reid RJ, Bodley JW, Anderson D (1994a) Characterization of the prohead-pRNA interaction of bacteriophage phi 29. *J Biol Chem* 269(7):5157–5162
- Reid RJ, Bodley JW, Anderson D (1994b) Identification of bacteriophage phi 29 prohead RNA domains necessary for in vitro DNA-gp3 packaging. *J Biol Chem* 269(12):9084–9089
- Reid RJ, Zhang F, Benson S, Anderson D (1994c) Probing the structure of bacteriophage phi 29 prohead RNA with specific mutations. *J Biol Chem* 269(28):18656–18661
- Rickgauer JP, Fuller DN, Grimes S, Jardine PJ, Anderson DL, Smith DE (2008) Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage phi29. *Biophys J* 94(1):159–167
- Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* 250(463):194–199
- Rossmann MG, Morais MC, Leiman PG, Zhang W (2005) Combining X-ray crystallography and electron microscopy. *Structure* 13(3):355–362
- Salas M, Mellado RP, Vinuela E (1978) Characterization of a protein covalently linked to the 5' termini of the DNA of *Bacillus subtilis* phage phi29. *J Mol Biol* 119(2):269–291
- Schachtele CF, De Sain CV, Anderson DL (1973) Transcription during the development of bacteriophage phi29: definition of “early” and “late” phi29 ribonucleic acid. *J Virol* 11(1):9–16
- Shu D, Zhang H, Jin J, Guo P (2007) Counting of six pRNAs of phi29 DNA-packaging motor with customized single-molecule dual-view system. *EMBO J* 26(2):527–537
- 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(6813):745–750
- Simpson AA, Leiman PG, Tao Y, He Y, Badasso MO, Jardine PJ, Anderson DL, Rossmann MG (2001) Structure determination of the head-tail connector of bacteriophage phi29. *Acta Crystallogr D Biol Crystallogr* 57(Pt 9):1260–1269
- Singleton MR, Sawaya MR, Ellenberger T, Wigley DB (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* 101(6):589–600
- Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 76:23–50
- Skordalakes E, Berger JM (2003) Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* 114(1):135–146
- 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(6857):748–752
- Story RM, Steitz TA (1992) Structure of the recA protein-ADP complex. *Nature* 355(6358):374–376
- Story RM, Weber IT, Steitz TA (1992) The structure of the *E. coli* recA protein monomer and polymer. *Nature* 355(6358):318–325
- Subramanya HS, Bird LE, Brannigan JA, Wigley DB (1996) Crystal structure of a DExx box DNA helicase. *Nature* 384(6607):379–383
- Sun S, Kondabagil K, Gentz PM, Rossmann MG, Rao VB (2007) The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. *Mol Cell* 25(6):943–949
- Sun S, Kondabagil K, Draper B, Alam TI, Bowman VD, Zhang Z, Hegde S, Fokine A, Rossmann MG, Rao VB (2008) The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* 135(7):1251–1262
- Tang J, Olson N, Jardine PJ, Grimes S, Anderson DL, Baker TS (2008) DNA poised for release in bacteriophage phi29. *Structure* 16(6):935–943
- 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(3):431–437
- Thomsen ND, Berger JM (2008) Structural frameworks for considering microbial protein- and nucleic acid-dependent motor ATPases. *Mol Microbiol* 69(5):1071–1090

- Trottier M, Guo P (1997) Approaches to determine stoichiometry of viral assembly components. *J Virol* 71(1): 487–494
- Tsuprun V, Anderson D, Egelman EH (1994) The bacteriophage phi 29 head-tail connector shows 13-fold symmetry in both hexagonally packed arrays and as single particles. *Biophys J* 66(6):2139–2150
- Valle M, Kremer L, Martinez AC, Roncal F, Valpuesta JM, Albar JP, Carrascosa JL (1999) Domain architecture of the bacteriophage phi29 connector protein. *J Mol Biol* 288(5):899–909
- Valpuesta JM, Carrascosa JL (1994) Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Q Rev Biophys* 27(2):107–155
- Valpuesta JM, Fernandez JJ, Carazo JM, Carrascosa JL (1999) The three-dimensional structure of a DNA translocating machine at 10 Å resolution. *Structure* 7(3):289–296
- van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, Patwardhan A (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* 33(4):307–369
- Velankar SS, Soutanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97(1):75–84
- Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1(8):945–951
- Wichitwechkarn J, Bailey S, Bodley JW, Anderson D (1989) Prohead RNA of bacteriophage phi 29: size, stoichiometry and biological activity. *Nucleic Acids Res* 17(9):3459–3468
- Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* 289(5487):2129–2133
- Wittinghofer A, Scheffzek K, Ahmadian MR (1997) The interaction of Ras with GTPase-activating proteins. *FEBS Lett* 410(1):63–67
- Xiang Y, Morais MC, Battisti AJ, Grimes S, Jardine PJ, Anderson DL, Rossmann MG (2006) Structural changes of bacteriophage phi29 upon DNA packaging and release. *EMBO J* 25(21):5229–5239
- Xiang Y, Morais MC, Cohen DN, Bowman VD, Anderson DL, Rossmann MG (2008) Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage phi29 tail. *Proc Natl Acad Sci USA* 105(28): 9552–9557
- Xiao F, Moll WD, Guo S, Guo P (2005) Binding of pRNA to the N-terminal 14 amino acids of connector protein of bacteriophage phi29. *Nucleic Acids Res* 33(8):2640–2649
- Yehle CO (1978) Genome-linked protein associated with the 5' termini of bacteriophage phi29 DNA. *J Virol* 27(3):776–783
- Yoshida M, Amano T (1995) A common topology of proteins catalyzing ATP-triggered reactions. *FEBS Lett* 359(1):1–5
- Yu J, Moffitt J, Hetherington CL, Bustamante C, Oster G (2010) Mechanochemistry of a viral DNA packaging motor. *J Mol Biol* 400(2):186–203
- Zhang C, Lee CS, Guo P (1994) The proximate 5' and 3' ends of the 120-base viral RNA (pRNA) are crucial for the packaging of bacteriophage phi 29 DNA. *Virology* 201(1):77–85
- Zhang C, Tellinghuisen T, Guo P (1995) Confirmation of the helical structure of the 5'/3' termini of the essential DNA packaging pRNA of phage phi 29. *RNA* 1(10):1041–1050
- Zhang F, Lemieux S, Wu X, St-Arnaud D, McMurray CT, Major F, Anderson D (1998) Function of hexameric RNA in packaging of bacteriophage phi 29 DNA in vitro. *Mol Cell* 2(1):141–147
- Zhao W, Morais MC, Anderson DL, Jardine PJ, Grimes S (2008) Role of the CCA bulge of prohead RNA of bacteriophage ϕ 29 in DNA packaging. *J Mol Biol* 383(3):520–528
- Zheng H, Olia AS, Gonen M, Andrews S, Cingolani G, Gonen T (2008) A conformational switch in bacteriophage p22 portal protein primes genome injection. *Mol Cell* 29(3):376–383
- Zhou ZH, Dougherty M, Jakana J, He J, Rixon FJ, Chiu W (2000) Seeing the herpesvirus capsid at 8.5 Å. *Science* 288(5467):877–880