The Bacteriophage DNA Packaging Motor

Venigalla B. Rao[1] and Michael Feiss[2]

[1]Department of Biology, The Catholic University of America, Washington, D.C. 20064; email: rao@cua.edu

[2]Department of Microbiology, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa 52242; email: michael-feiss@uiowa.edu

Annu. Rev. Genet. 2008. 42:647–81

First published online as a Review in Advance on August 7, 2008

The *Annual Review of Genetics* is online at genet.annualreviews.org

This article's doi: 10.1146/annurev.genet.42.110807.091545

0066-4197/08/1201-0647$20.00

Key Words

bacteriophage, virus assembly, DNA packaging, terminase, ATPase, molecular motor

Abstract

An ATP-powered DNA translocation machine encapsidates the viral genome in the large dsDNA bacteriophages. The essential components include the empty shell, prohead, and the packaging enzyme, terminase. During translocation, terminase is docked on the prohead's portal protein. The translocation ATPase and the concatemer-cutting endonuclease reside in terminase. Remarkably, terminases, portal proteins, and shells of tailed bacteriophages and herpes viruses show conserved features. These DNA viruses may have descended from a common ancestor. Terminase's ATPase consists of a classic nucleotide binding fold, most closely resembling that of monomeric helicases. Intriguing models have been proposed for the mechanism of dsDNA translocation, invoking ATP hydrolysis-driven conformational changes of portal or terminase powering DNA motion. Single-molecule studies show that the packaging motor is fast and powerful. Recent advances permit experiments that can critically test the packaging models. The viral genome translocation mechanism is of general interest, given the parallels between terminases, helicases, and other motor proteins.

INTRODUCTION

During large dsDNA virus assembly, viral DNA is translocated into preformed protein shells. The DNA packaging process compacts the highly negatively charged DNA to a density similar to that of crystalline DNA (45). DNA packaging is not spontaneous; rather, the DNA is driven into the shell by a translocating motor powered by ATP hydrolysis (**Figure 1**). The preformed empty shell is an icosahedron formed by many copies of the major capsid protein. One of the shell's 12 fivefold vertices is a special portal vertex formed by the dodecameric portal protein (42). During translocation, a viral enzyme, called terminase, is docked on the portal and the DNA is translocated through the portal channel. Terminase contains the ATPase that powers the translocation machine (**Figure 1**) (See sidebar: Terminase). Following DNA packaging, terminase undocks from the portal. The portal then serves as the site for tail attachment that completes virion assembly. At the start of an infection, the DNA exits the shell through the coaxial portal and tail channels into the host cell.

Many plausible models have been offered for how the DNA translocation motor might work. Here we discuss these ideas as well as recent results on the structures and dynamics of phage DNA packaging motors. Prior to DNA translocation, immature virus DNA is recognized and processed; aspects of virion assembly that have been discussed elsewhere (14, 21) are not emphasized here.

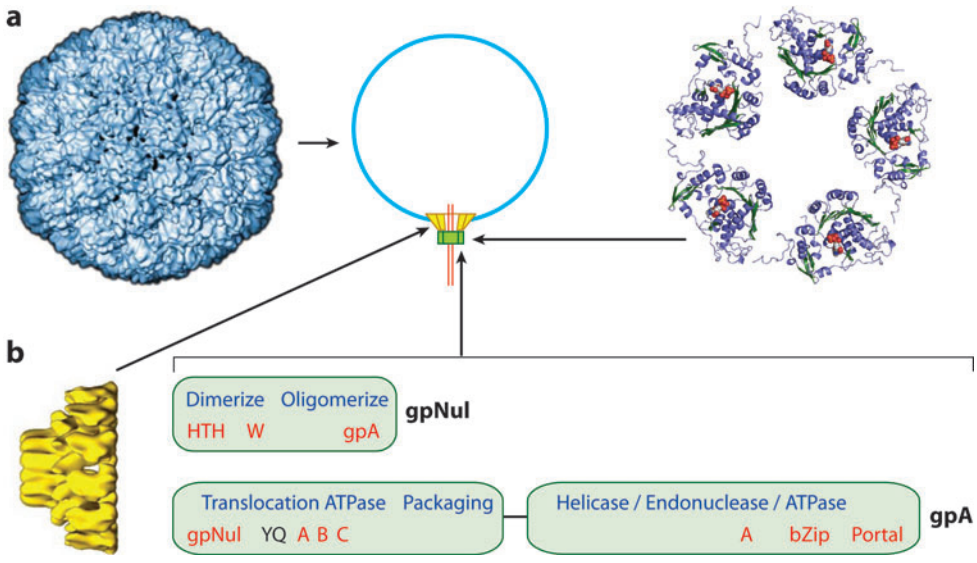


Figure 1

Components of the Phage DNA packaging machine. (*a*) *Left*: Prohead I of HK97. *Right*: A pentameric model of the T4 gp17 translocation ATPase domain. ATP is shown in red. *Center*: Schematic showing the packaging complex of a prohead (*blue*) with terminase (*green*) docked on the portal (*yellow*). DNA: red lines. (*b*) *Left*: CryoEM image of the φ29 portal, side view. *Right*: Functional map of the small (gpNu1) and large (gpA) subunits of phage λ terminase. gpNu1: HTH and W indicate the winged helix-turn-helix motif, and the segment marked gpA is the functional domain for interaction with the N terminus of gpA. gpNu1 segments involved in dimerization and oligomerization are indicated. Not shown is a low-affinity ATPase center in gpNu1 near the wing motif. gpA: The N-terminal 60% of the protein contains the translocation ATPase. ATPase motifs: YQ, adenine binding motif; A, B, C, the Walker A and B sequences and the coupling motif, respectively. At the N and C termini are functional domains for interacting with gpNu1 and the portal protein, respectively. The C-terminal domain contains the cohesive end-generating endonuclease, which includes a putative Walker A segment (A) and a basic leucine zipper (bZip).

DNA PACKAGING AND VIRUS ASSEMBLY: AN OVERVIEW

Chromosome

The virion DNAs of the tailed bacteriophages, adeno-, herpes-, and pox-viruses are linear ds-DNA molecules. The chromosomes of many tailed phages, e.g., the λ-like and P2/P4-like phages, have complementary cohesive ends, which anneal to cyclize the DNA upon injection into a host cell. The DNA packaging recognition site, *cos*, includes the cohesive end sequence. Virion genomes are generated from concatemers during packaging, when terminase introduces staggered nicks at *cos* sites to regenerate the cohesive ends (**Figure 2*a***).

The chromosomes of other tailed phages are terminally redundant, permitting cyclization by homologous recombination (e.g., phages P22 and SPP1), or end-to-end recombination to form concatemers (e.g., phages T4 and T7) (See sidebar, Concatemer). Processing is initiated when terminase binds to the *pac* recognition site on the concatemer and makes the initiating cut at a nearby sequence. Terminase remains bound to the newly created chromosome end, captures a prohead, and translocation of DNA into the prohead ensues. Head filling triggers terminase to make a second, nonspecific cut, which produces a terminally redundant virion DNA. The latter is due to a strict (evolutionary) linkage between capsid size and genome length, the capsid volume accommodating little over a unit length genome (102%–110%). Packaging is processive, so that the next chromosomes cut from the concatemer are likewise terminally redundant and circularly permuted (**Figure 2*a***). Phage T4 DNA processing is similar to that of P22 and SPP1, except that the initial cuts are not necessarily made near to or at a unique *pac* site (121).

Virus chromosomes with unique, terminally redundant sequences, e.g., T3 and T7, lose the terminal redundancy when newly replicated chromosomes undergo end-to-end recombination. Accordingly, generating virion DNA requires specific cutting and local replication to regenerate the terminal redundancy (52).

TERMINASE

Before it was discovered that DNA translocation involved a packaging enzyme, there was evidence that a viral enzyme was required to generate the cohesive (*cos*) ends in phage λ. This endonuclease function was ascribed to an enzyme called terminase, since the activity generated the termini (cohesive ends) of virion DNA (143). Subsequently, terminases have been found to consist of a large protein containing the ATPase activity that powers translocation along with the endonuclease activity, and a small protein for recognition of viral DNA.

CONCATEMER

Multimeric head-to-tail polymer of viral DNA produced by rolling circle replication and/or recombination. Concatemers are the substrate for the production of virion genomes by terminase during DNA packaging.

Herpes virus DNA molecules are unique and terminally redundant, though inversion of the two large unique segments creates four chromosomal isomers. Viruses with protein-primed DNA replication, including adenoviruses and φ29-like phages do not produce concatemers. Rather, these viruses replicate as monomers, with replication being initiated from the terminal proteins (127).

Prohead and Portal

The key component of the prohead that participates in the DNA translocation process is the portal protein. The portal protein of the tailed bacteriophages (147) and herpes simplex virus (146) is a cone-shaped dodecamer of subunits arranged radially, with a central channel for DNA passage (**Figures 1**, **2**). A narrow "stalk" domain protrudes outside the prohead, which provides the surface for assembly of the translocation motor. The central "stem" domain contains 2 long α-helices from each subunit, forming a central cylinder walled by 24 α-helices. Above the stem domain are massive "wing" and

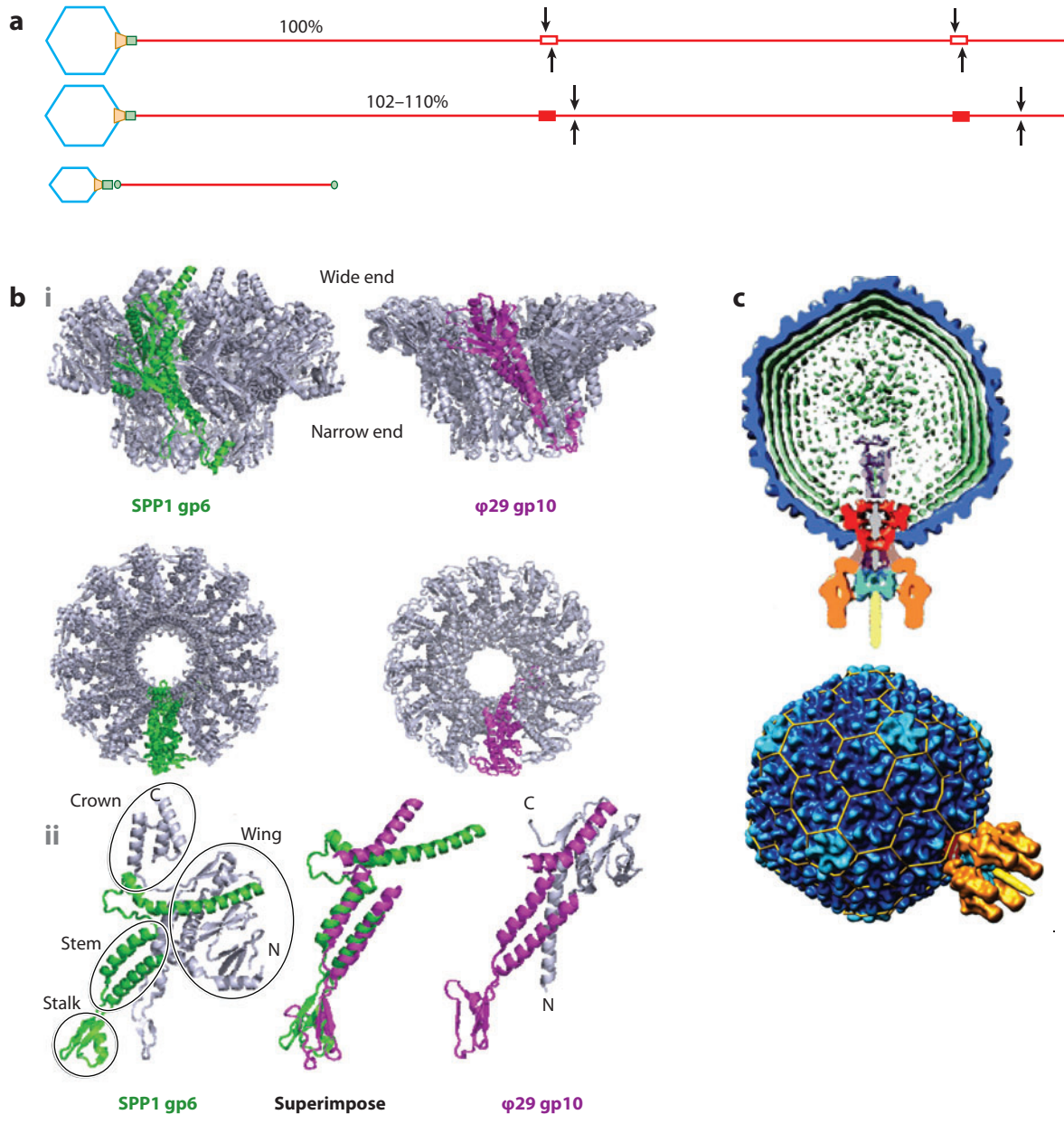


Figure 2

DNA processing by tailed dsDNA phages. (*a*) *Top*: Packaging-dependent concatemer processing by *cos*-containing phages (e.g., λ-like phages). *Middle*: Concatemer processing by *pac*-containing phages (e.g., P22, SPP1). *Bottom*: Protein-primed monomeric DNA replication and packaging by a strand displacement mechanism (φ29-like phages, adenoviruses) (*b*) *i*: Front and top views of SPP1 13-fold (PDB ID: 2jes) and φ29 12-fold (PDB ID: 1h5w) portal structures. *ii*: A single subunit of the SPP1 and φ29 portal rings is colored green and magenta, respectively. *ii*: *Left*: A single SPP1 portal subunit showing subdomains and the conserved core structure (*green*). *Right*: A single φ29 portal subunit showing the conserved structure (*magenta*). *Center*: Superimposition of the core structures shows striking structural conservation despite negligible sequence identity. All images were created with PyMol. (*c*) *Top*: CryoEM image of phage P22–cross section. Shell, blue; the portal protein, gp1, red; gp4, mauve; and DNA, green. The remaining structures are internal proteins and tail components. *Bottom*: P22 virion cryoEM image with head shell in blue and tail components in yellow.

"crown" domains that are internal to the prohead (**Figure 2*b***) (90).

Shell assembly is initiated on the portal protein. For some viruses, e.g., P22 and HK97, nonfunctional prohead shells assemble in the absence of the portal protein; these shells have normal lattice structure but are unable to package DNA. In functional proheads, there is a symmetry mismatch between the 12-fold symmetry of the portal protein, and the fivefold symmetry of the shell vertex occupied by the portal protein. Assembly of functional proheads requires the portal protein, the major capsid protein, and in many cases, a scaffold protein that assists prohead assembly but is not retained in the mature virion.

Terminase

Phage terminases are DNA packaging enzymes that contain the ATPase activity that powers DNA translocation. Most terminases also contain the endonuclease that, during DNA packaging, cuts concatemeric DNA into genome lengths. Terminases must also recognize viral DNA in a pool that may also include host DNA. Terminases generally are hetero-oligomers of a small protein involved in DNA recognition, and a large protein containing the translocation ATPase, the endonuclease, and a motif for docking at the portal vertex (14) (**Table 1**; **Figure 1**). Alignments show that terminase large subunits are descended from a common ancestor (19, 41, 105). Remarkably, a herpes virus DNA packaging ATPase falls into this family (120).

Phage φ29 is unusual in several respects. First, DNA replication is protein-primed by gp3, which is covalently joined to the viral DNA. Second, gp3 is a necessary component of the DNA packaging machinery and is analogous to the small terminase protein of other phages (62, 127). The large terminase protein equivalent is gp16, which possesses the packaging ATPase activity. Third, φ29's prohead contains a small 174-nt packaging RNA (pRNA). The pRNA has short, complementary segments that enable the formation of an oligomeric ring on the portal. The pRNA is required to dock gp16 on the portal protein, and hence may be considered as part of the large terminase protein of other phages.

Table 1 Terminases for packaging the dsDNA of tailed bacteriophages

	Small terminase component		Large terminase component	
Phage/virus	*Gene product*	*Mass*	*Gene product*	*Mass*
λ	gpNu1	20 kDa	gpA	73 kDa
T4	gp16	18 kDa	gp17	70 kDa
T3/T7	gp18	20 kDa	gp19	67 kDa
SPP1	gp1	21 kDa	gp2	49 kDa
P22	gp3	19 kDa	gp2	58 kDa
φ29	gp3	31 kDa	gp16 pRNA	39 kDa 58 kDa
T5	?	?	gp155	49 kDa
HSV1	UL28	85 kDa	UL15	81 kDa

DNA Packaging

To initiate DNA packaging, the small terminase subunit binds specifically to concatemeric DNA. The binding site is near the site of DNA cutting; for example, the P22 small subunit, gp3, binds to a 17-bp *pac* site located within the *3* gene, and the gp2 large subunit cuts the DNA at various sites in a 120-bp segment near *pac* (17). Following the initial DNA cut, terminase remains bound to the DNA end in a gp2-gp3-DNA complex that then docks on the prohead's portal vertex (112).

For all the phages that process concatemeric DNA, there is little information about the stoichiometry and structure of (*i*) the initial precleavage complex, (*ii*) the postcleavage complex, and (*iii*) the ternary DNA-terminase-prohead complex. Major structural changes are expected to occur during these transitions. For example, cohesive ends like those of λ must be separated after being created through the introduction of staggered nicks by terminase. Separating the cohesive ends may also involve the driving apart of terminase protomers. Similarly, docking of the terminase-DNA complex onto the portal protein requires structural changes to activate the translocation ATPase. Additional terminase protomers are likely recruited

during assembly of the translocation complex at the portal. A major aspect of the transition to translocation is the movement of terminase away from the DNA recognition site—the translocating complex must then use nonspecific DNA-protein contacts. Little is known about these transitions. Whether the small terminase subunit is present during translocation is unclear—the findings that the T4 and λ large subunits are sufficient for translocation and high-force generation indicate that any role of the small subunit is not critical for translocation per se. On the other hand, packaging along a concatemer is processive, and the small subunit likely plays a role in initiating packaging of the next chromosome along the concatemer (35).

The second terminase cut is triggered by head filling, both for phages using the headful packaging scheme and for those with cohesive ends (See sidebar, Headful Packaging). The nature of the triggering event is not clear. For λ, the *cos* subsite *cosQ* is required for downstream cutting, and evidence indicates that *cosQ* recognition requires that the translocation rate be slow (32–35, 154, 155). Following translocation, downstream *cos* cleavage, and undocking of terminase from the portal, the terminase remains bound to the concatemer and ready to capture a prohead for the next round of packaging (47).

The Filled Head

As the prohead is being filled with DNA, each major capsid protein subunit forming the icosahedral lattice undergoes an in situ rearrangement to mature the shell. In most cases, maturation expands the shell, increasing the shell's DNA capacity by 50%–100%. In phages λ and T3, prohead expansion occurs when about 30% of the chromosome has been packaged (51, 54, 72). In T4, expansion can be uncoupled from packaging (123).

DNA packaged into isometric icosahedral shells is proposed to be cylindrically spooled about the axis of the portal channel (22, 82). DNA spooling could occur if the first packaged DNA forms an initial layer on the inside surface of the shell, followed by the formation of additional layers. The DNA layers account for much of the packaged DNA, but some randomly packed DNA, presumably the last packaged, is in the center of the DNA condensate (**Figure 2*c***). Recent φ29 studies show that the forces resisting DNA translocation rise to higher levels than predicted by spooling models, particularly in the early stages of packaging. CryoEM examination of packaging complexes at various stages of filling indicates that the DNA is not in an ordered state until the shell is about 70% filled. Rather, the DNA appears to be packaged without long-range order, and the layers form as a consequence of local hexagonal packing of DNA segments next to the inner shell surface as DNA packaging approaches the capacity of the shell (26, 126).

Single-molecule studies confirm that packaged DNA is under pressure. The estimated internal forces for φ29 are ~50 pN (136). Modeling indicates that these forces can be accounted for by extensive DNA bending and charge repulsion in packaged DNA. In addition to requiring that the translocation motor be powerful, the internal pressure has implications for shell strength and DNA ejection. In single-molecule experiments done with λ in the absence of the head stabilization protein, gpD, the internal pressure at late stages of packaging was sufficient to cause shell rupture (54). Head stabilization proteins, which stabilize the icosahedral shell lattice, are found for many phages, including the T4's Soc protein (49). In vitro studies indicate that the internal force is

HEADFUL PACKAGING

A packaging mode in which the prohead is filled to capacity with DNA, followed by a DNA-cut to free the virion genome from the concatemer. In phages that make sequence specific cuts (λ), a headful precisely equals one genome length DNA. In phages that make nonspecific headful cuts, the virion chromosomes contain a terminal redundancy between 1.02–1.1 times the unit length genome, and the end sequences in individual virions are not identical.

sufficient to drive ejection of part, but not all, of the λ and SPP1 chromosomes during infection.

Head Completion

There is little evidence about how terminase dismounts from the portal and cuts the concatemer following completion of translocation, or the order of these events. At present, there is no evidence that additional phage proteins are required to undock terminase. Following terminase's departure, the "neck" proteins assemble at the portal vertex to hold the DNA in, and to provide a binding site for the tail. In λ, gpW binding to the portal assists the filled head to retain the viral DNA and enables the next protein, gpFII, to bind. About six copies of gpFII are found in virions, so perhaps hexameric rings of gpW and gpFII assemble on the portal. In T4, the analogous proteins are gp13, gp14, and gp15. In P22, the tail addition protein, gp4, forms a 12-mer on the stalk domain of the dodecameric gp1 portal protein. The gp4 interaction changes the portal's conformation, which may signal a transition from translocation to virion completion (166).

PROPERTIES OF THE DNA PACKAGING MOTOR

Defined In Vitro Packaging

DNA translocation is measured by defined in vitro DNA packaging systems consisting of highly purified components (15, 62, 64, 76, 86, 113, 160). The packaged DNA is quantified by determining the amount of encapsidated DNA that is protected from DNAse digestion. Only two basic components are required: the prohead and the large terminase protein; the small terminase protein is not essential in most in vitro systems (124), though it is essential in vivo for recognition of viral DNA. Linear dsDNA molecules can be translocated, apparently from the end, and as many as 50%–100% of the proheads can be filled. If short, less than headful-length DNA molecules are added as the DNA substrate, the system fills the prohead with multiple molecules until the head is full (28, 93, 130a).

Packaging of individual DNA molecules can be studied in single-molecule experiments, in which the prohead-terminase complex is tethered to a microsphere coated with capsid protein antibody, and the biotinylated DNA is tethered to another microsphere coated with streptavidine. The microspheres containing either a single packaging motor or a single DNA molecule are held by optical tweezers and packaging is initiated by moving the packaging motor bead into near contact with the fixed DNA bead (136). The motor captures the DNA and as it translocates, the dynamics of the packaging motor such as force and speed can be quantified.

Step Size

Step size is defined as a single step taken by the packaging motor, which is equivalent to the number of base pairs of DNA translocated per ATP hydrolyzed (See sidebar, Step Size). The average step size is about 2 bp/ATP in phages φ29 and T3 (24, 62, 109), a value obtained by dividing the bp of DNA packaged by the number of P_i molecules produced in a defined packaging reaction. These bulk assay measurements by their very nature cannot be precisely applied to single molecules, and it is unknown whether the step size is

STEP SIZE

The prevailing hypothesis is that firing of each ATPase subunit of the packaging motor gives a "push" to translocate DNA; a pause would follow during which the DNA is handed over to the next subunit in sequence, and then the firing of the second ATPase. It is, however, possible that ATPase firings may be continuous and pause occurs not between each firing but after all (or a set of) ATPase subunits are fired. Also, contrary to the thinking that step size is a whole and fixed number of base pairs, the motor subunits may have to adjust to the imprecise DNA symmetry as well as the internal pressure built during translocation. Thus the step size may not be a whole number and may vary from step to step and depend on the extent of head filling.

universal for all packaging motors, or it varies at different stages of translocation. Nevertheless, the 2 bp/ATP step size fits with the 10_1-fold screw symmetry of double helical DNA.

Force

A surprising finding from single-molecule studies is that the phage packaging motor generates enormous force in order to package DNA. Forces as high as ~60 pN were measured in phages φ29, λ, and T4, thus making the packaging motor one of the strongest force generating biological motors reported to date (53, 54, 136). The force is 20–25 times that of myosin, 10 times that of kinesin, or >2 times that of RNA polymerase. Such high forces seem to be essential to pack the viral DNA against the enormous electrostatic repulsive forces (and bending and entropic energies) to confine a highly negatively charged DNA polymer within a limited volume of the capsid (55).

Velocity

The phage packaging motors show high rates of packaging as well as high processivity. The T4 motor can achieve rates as high as ~2000 bp/sec, the highest recorded to date (53). Slips and pauses do occur but these are relatively short and rare, and the motor recovers and recaptures DNA, continuing translocation. The translocation rate decreases as packaging proceeds and internal pressure builds. In λ, the motor velocity slows down by about three fold when ~90% of the genome is packaged (54). In φ29, the packaging rate nearly falls to zero when 100% of the genome is packaged (136).

The rate of translocation is in keeping with the genome size of the respective phage. Whereas the φ29 motor packages its 20-kb size genome at an initial rate of 100–150 bp/sec, the phage T4 motor packages a 170-kb size genome at an average rate of ~700 bp/sec. The phage λ motor packages the 48.5-kb chromosome at a rate of ~600 bp/sec. Assuming the same step size for T4 and φ29, the T4 packaging motor "burns" ATP fuel ~7 times as fast as the φ29 motor. Thus, although the phages package different-sized genomes, the motors may have evolved to adjust the rate and complete packaging in 2–5 min, as would be necessary to complete the infection cycle within 20–30 min (53).

Power

Phage packaging motors generate enormous power, with the T4 motor being the fastest and the most powerful. Even with a high external load force of 40 pN, the T4 motor can translocate DNA at a remarkable speed of ~380 bp/sec. This is equivalent to a power of 15,200 pN/bp/s, or 5.2×10^{-18} W. Scaling up the nanoscale T4 packaging motor to a macromotor, the motor power density is approximately twice that of a typical automobile engine (53).

Efficiency

The free energy of ATP hydrolysis, when expressed in units of force and displacement, is equivalent to 120 pN per nm. Since all three phage packaging motors studied generate ~60 pN force, the motor should have translocated 2 nm, or approximately 6 bp, of DNA if 100% of the free energy is converted into DNA motion. But with a 2-bp step size, assuming that it is the same in all the motors, the phage packaging motor converts about 30% of the ATP free energy into mechanical motion (24, 136). Of the remaining 70%, a fraction likely supports conformational transitions, whereas the rest is probably released as heat, thus making an enormous entropic contribution to offset the ordered packing of DNA.

PARTS AND FUNCTIONS OF THE DNA PACKAGING MOTOR

Molecular genetic analyses, biochemical studies, and sequence alignments led to the prediction of key functional motifs involved in the DNA packaging process. Although there is no significant overall sequence similarity, the terminase proteins from numerous phages contain well-conserved patches of amino acid sequences

or structural motifs that are required for packaging (see **Figure 1**).

The Small Terminase Protein

Helix-Turn-Helix (HTH) motif. The phage λ small terminase protein, gpNu1, consists of a HTH motif at the N terminus, which recognizes the R box sequences in the *cosB* region (10, 38, 88). In the NMR structure, the HTH motif is followed by a loop, which forms a "wing-like" structure (38). This type of architecture is similar to the atypical winged HTH motifs found in a diverse array of DNA binding proteins. The two HTH motifs of the gpNu1 dimer contact specific nucleotides in the R2 and R3 boxes in *cosB* that are about 45 nucleotides apart, sharply bending the DNA. The bend is stabilized by the binding of the protein IHF of *Escherichia coli* (104, 157, 158). Assembly of this site-specific nucleoprotein complex sets things in motion for assembly of the large terminase protein, gpA, and cleavage at the adjacent *cosN* sequence (29–31, 68–70, 130, 161). Although the recognition sequences (e.g., *pac* sites in phages P22 and SPP1) and details of the pathway vary, similar basic themes of DNA recognition and packaging initiation likely operate in other phage systems. Indeed, many small terminase proteins including T4's gp16 do contain a predicted HTH motif in the N terminus that is presumably involved in the recognition of the viral DNA (38).

Oligomerization domain. Small terminase proteins exist as stable oligomers in solution, forming rings of 8–10 subunits (23, 95, 112, 152). Sequence analyses identified the presence of one to three coiled-coil motifs in the central region of the small terminase protein (85). The T4 gp16 consists of two long central helices with—two to four heptad repeats, with each heptad having hydrophobic residues at positions *a* and *d*, and charged residues at positions *e* and *g*. The helices of neighboring subunits are predicted to form parallel coiled-coil interactions producing oligomers. Oligomerization is critical for function. The monomeric gp16, unlike the oligomer, cannot stimulate the gp17-ATPase activity, and the truncated λ gpNu1 dimer shows >100-fold reduction in *cosB* binding when compared to the wild-type protein (85, 159).

Large terminase protein binding. The small terminase protein interacts with the large terminase protein, modulating its functions. For instance, although the small terminase protein recognizes the viral DNA, the active center for DNA cleavage is located in the large terminase protein. Thus the small terminase protein binding to DNA must be accompanied by its interactions with the large terminase protein (holo-terminase assembly) in order to make a packaging initiation cut in a controlled fashion. Second, the T4 small terminase protein stimulates the ATPase and in vitro DNA packaging activities associated with the large terminase protein (92). These have also been well documented in phages λ, SPP1, T3/T7, and P22 [see (20)]. Genetic studies show that the C terminus of λ gpNu1 interacts with the N terminus of gpA, although the boundaries and precise interacting amino acids are undefined (50). Unlike in phage λ where a stable holo-terminase forms readily between gpNu1 and gpA, the T4 gp16-gp17 interaction is transient and dynamic. Stabilization of the complex may require interactions with the components of DNA transcription, replication, and repair, since packaging in T4 is more intimately linked to these processes (102).

ATP binding site. At least three small terminase proteins, gpNu1 (λ), gp16 (T4) and gp1 (SPP1), are reported to bind ATP (23, 57, 75, 95). The gpNu1 ATP binding site is referred to as the low-affinity site (K_d = 1 mM) with weak ATPase activity. Cross-linking studies with azido-ATP show that this segment (amino acids 18 to 35) includes part of the DNA binding helix and the accompanying wing (6). Mutational studies suggest that this site contributes to specific binding of gpNu1 to the *cos* recognition sequences (78). No ATPase activity is, however, detectable with the gp16 and gp1. Sequence alignments do not predict a classic Walker A- or Walker B-type ATP binding

NBD (NUCLEOTIDE BINDING DOMAIN)

First identified in dinucleotide binding proteins (128), numerous nucleotide binding proteins consist of the Rossmann fold, a β-sheet with three or more parallel β-strands interspersed by α-helices. The first solved RecA ATPase structure (140) has a modified version of this fold, a β-sheet with five parallel β-strands containing the highly conserved sequence signatures such as Walker A and Walker B. The "RecA-type ATPase fold" was subsequently discovered in F_1F_0-ATPases, helicases, chaperones, and a host of other ATPases with diverse functions.

motif in any of the small terminase proteins (105). Furthermore, the mostly α-helical secondary structure prediction suggests that the protein does not possess the classic nucleotide binding domain (NBD) (See sidebar, NBD), a β-sheet scaffold (128). Thus there appears to be a novel and a rather "simple" nucleotide binding motif in the small terminase protein. The nature of this motif and its role in DNA packaging remain to be investigated.

CM (COMBINATORIAL MUTAGENESIS)

In CM, the target amino acid is changed to all twenty amino acids using overlap extension PCR (73). A library of mutant clones is generated by insertion of the mutant DNA into the pET expression vector. Random screening of the library by recombinational marker rescue, using phage containing an amber mutation very close to the target amino acid, identifies a collection of phenotypes (null, functional, *ts*, *cs*) (122). Following DNA sequencing, selected mutant proteins are overexpressed, and the functional defect is determined by biochemical assays.

Large Terminase Protein

The large terminase protein consists of two domains, an N-terminal ATPase domain that powers DNA translocation and a C-terminal nuclease domain that generates the termini of the viral genome (44, 83). Though the individual domains exhibit various activities, a physical connection between the domains is critical for DNA translocation (83).

ATPase Catalytic Center

Walker A motif. A consensus Walker A sequence, [G/A-XXXXGK(T/S)] is present in virtually all ATPase structures (150). Considered to be an ancient motif, the Walker A residues form a fairly rigid phosphate-binding loop (P-loop), which captures ATP through interaction of the ε-amino group of Lys with the β- and γ-phosphates of ATP and the hydroxyl group of Thr (or Ser) with the Mg^{2+} of the Mg-ATP complex (**Figure 3**). Although initial analyses identified two or more putative Walker A motifs (62, 111), extensive sequence alignments show only one strictly conserved N-terminal Walker A motif in all large terminase proteins (105) (**Figure 3*a***). Combinatorial mutagenesis (See sidebar, CM) of the phage T4 Walker A ($SRQLGKT_{161-167}$) showed that no substitutions are tolerated at the highly conserved GKT signature. Conservative substitutions such as G165A, K166R, and T167A resulted in a null phenotype. Biochemical analyses showed that

Figure 3

The N-terminal gp17-ATPase center. (*a*) Sequence alignments showing that the functional signatures of the ATPase domain of phage T4 and other viral terminases are conserved and are similar to that of the translocating monomeric SF2 helicases, types I and III restriction endonucleases, and protein translocases. The alignments were generated by CLUSTALW and the consensus secondary structure predictions by Jpred. Numbers in parentheses represent the number of amino acids. (*b*) The T4 gp17-ATPase domain (Glu255Asp256 mutant) X-ray structure showing the classic NBD and RecA-type ATPase fold consisting of six parallel β-strands having 3,2,1,4,5,6 topology (shown in different colors). The bound ATP molecule is shown in black. The adenine binding loop (*red*) adopts a different conformation in the apo form (*cyan*). (*c*) The ATPase active center showing the catalytic residues mapped by mutagenesis studies; additional adenine binding residues that are not yet functionally mapped are: Ile127 (*green*), Gln138 (*magenta*), Arg140 (*red*), Tyr142 (*cyan*), and Gln143 (*orange*).

a

T4	NH_2	(137)	QLRD	**YQ**	RDML	(9)	VCN	**LSRQLGKT**	TVVA	(75)	NSFA	**MIYIDE**	CAFI	(20)	KIII	**TTT**	PNGL	(319)	$COOH_2$
RB49	NH_2	(135)	QLRD	**YQ**	KDML	(9)	AHK	**LSRQLGKT**	TAVA	(75)	NSFS	**FIYIDE**	CAFI	(20)	KMIM	**TTT**	PNGL	(318)	$COOH_2$
T7	NH_2	(33)	VPTK	**CQ**	IDMA	(11)	ILQ	**AFRGIGKS**	FITC	(82)	SRAD	**IIIADD**	VEIP	(29)	RVIY	**LGT**	PQTE	(381)	$COOH_2$
T5	NH_2	(37)	TPNG	**PQ**	IAII	(11)	TAC	**VSRRVGKS**	FIAY	(73)	RSYD	**FIIFDE**	AAIS	(20)	KALF	**IST**	PRGG	(248)	$COOH_2$
λ	NH_2	(41)	KESA	**YQ**	EGRW	(21)	NVV	**KSARVGYS**	KMLL	(82)	KSVD	**VAGYDE**	LAAF	(24)	KSIR	**GST**	PKVR	(423)	$COOH_2$
SPP1	NH_2	(7)	EKFT	**PH**	FLEV	(12)	VLK	**GGRGSAKS**	THIA	(81)	FPVA	**GMWIEE**	LAEF	(24)	IFFY	**SYN**	PPKR	(248)	$COOH_2$
P22	NH_2	(37)	APYS	**KQ**	REFI	(9)	CFM	**AGNQLGKS**	FTGA	(123)	DTIH	**GVWFDE**	EPPY	(15)	FSIL	**TFT**	PLMG	(265)	$COOH_2$
φ29	NH_2	(4)	LFYN	**PQ**	KMLS	(6)	FVI	**GARGIGKS**	YAMK	(72)	PNVS	**TIVFDE**	FIRE	(28)	RCIC	**LSN**	AVSV	(170)	$COOH_2$
HK97	NH_2	(27)	RLDP	**FQ**	KDFI	(13)	ILS	**IARKNGKT**	GLIA	(79)	LSPI	**LAILDE**	TGQV	(21)	LLIV	**IST**	QAAN	(284)	$COOH_2$
HSV-1	NH_2	(221)	GDHA	**EQ**	VNTF	(23)	VFL	**VPRRHGKT**	WFLV	(78)	QDFN	**LLFVDE**	ANFI	(17)	KIIF	**VSS**	TNTG	(350)	$COOH_2$

Secondary structure

SF2 Helicases

RAD3	NH_2	(15)	KIYP	**EQ**	YNYM	(13)	ILE	**MPSGTGKT**	VSLL	(173)	SKDS	**IVIFDE**	AHNI	(215)	VIIT	**SGT**	ISPL	(312)	$COOH_2$
NS3	NH_2	(181)	PVFT	**DN**	SSPP	(9)	HLH	**APTGSGKS**	TKVP	(66)	GAYD	**IIICDE**	CHST	(22)	VVLA	**TAT**	PPGS	(303)	$COOH_2$
EIF4A	NH_2	(43)	EPSA	**IQ**	QRAI	(9)	LAQ	**AQSGTGKT**	GTFS	(84)	DKIK	**MFILDE**	ADEM	(21)	VVLL	**SAT**	MPND	(188)	$COOH_2$

Type I restriction endonucleases

EcoKI	NH_2	(461)	GLRY	**YQ**	EDAV	(11)	LLA	**MATGTGKT**	RTAI	(83)	ARYD	**CIVVDE**	AHRG	(34)	KIAL	**TAT**	PALH	(546)	$COOH_2$
EcoR124I	NH_2	(270)	VMPR	**YQ**	IAAT	(23)	YIW	**HTTGSGKT**	LTSF	(80)	NQQV	**VFIFDE**	CHRS	(17)	QFGF	**TGT**	PIFP	(597)	$COOH_2$
KpnAI	NH_2	(293)	CCRY	**PQ**	YYAG	(20)	TYF	**GATGCGKS**	YTMQ	(92)	RSNI	**ICISDE**	AHRS	(35)	YVGF	**TGT**	PIDA	(621)	$COOH_2$

Type III restriction endonucleases

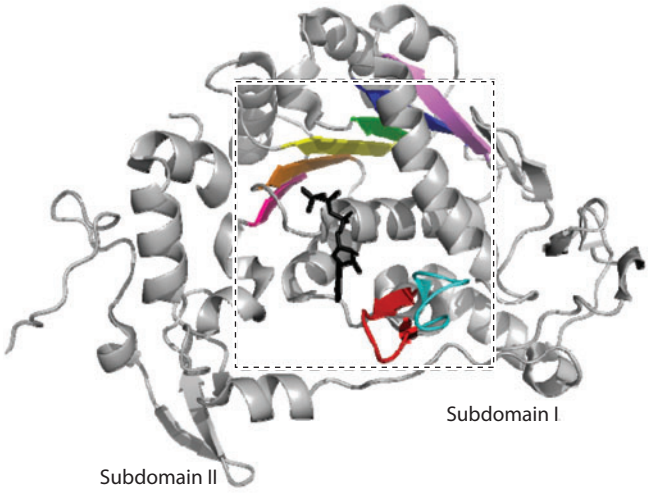
LlaFI	NH_2	(43)	AFRR	**FQ**	MQDN	(16)	LFN	**MATGSGKT**	MVMA	(99)	DEDI	**VILGDE**	AHHF	(29)	LLEF	**SAT**	INMD	(634)	$COOH_2$
EcoPI	NH_2	(55)	NIKK	**VQ**	ELNG	(15)	DVS	**METGTGKT**	YTYT	(122)	AVKP	**FIIIDE**	PHKF	(17)	IIRY	**GAT**	FSEG	(701)	$COOH_2$
StyLT1	NH_2	(24)	GIDH	**AQ**	ADHN	(17)	DVK	**METGTGKT**	YVYT	(123)	MTRP	**VVIIDE**	PHRF	(17)	IVRF	**GAT**	FPDI	(759)	$COOH_2$

Protein translocases

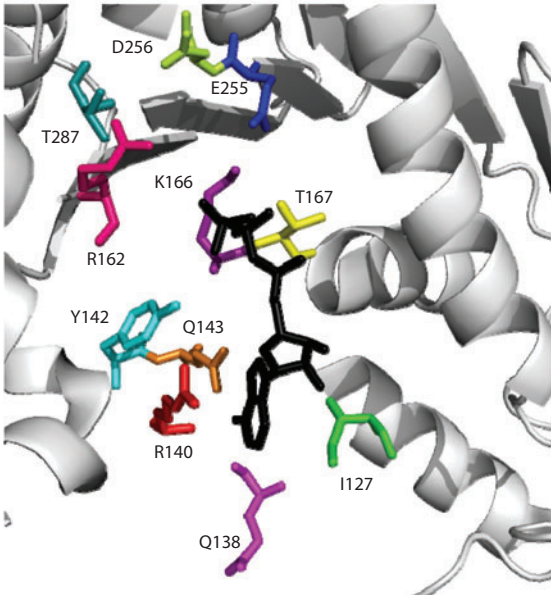
FtsH	NH_2	(176)	NPSR	**FH**	EMGA	(6)	LLV	**GPPGVGKT**	HLAR	(39)	HAPC	**IVFIDE**	IDAV	(35)	VVMA	**ATN**	RPDI	(319)	$COOH_2$
CLPA	NH_2	(210)	TTNL	**NQ**	LARV	(28)	LLV	**GESGVGKT**	AIAE	(51)	DTNS	**ILFIDE**	IHTI	(32)	TYQE	**FSN**	IFEK	(435)	$COOH_2$
SecA	NH_2	(79)	FPFK	**VQ**	LMGV	(7)	IAE	**MKTGEGKT**	LTST	(95)	DEVD	**SILIDE**	ARTP	(146)	LAGM	**TGT**	AKTE	(464)	$COOH_2$

Adenine binding **Walker A** **Walker B** **Catalytic carboxylate** **C-motif**

b



c



the Walker A mutants lost the ATPase and DNA packaging activities but retained the nuclease activity (122).

Phage λ gpA has two Walker A motifs, a high-affinity N-terminal Walker A, $KSARVGYS_{76–83}$, that is analogous to gp17's, and a putative C-terminal Walker A, $GASVYGKP_{491–498}$ (62, 75, 77, 79, 116, 129). The gpA and a few other large terminases show rare Walker A variants in which the Lys position is switched to the beginning of P-loop (106). Structural modeling shows that this Lys is similarly oriented into the ATP binding pocket like the classic Walker A (41). The importance of the N-terminal Walker A for DNA translocation was established by isolating an azido-ATP cross-linked adduct (39, 66), and a conservative null mutant, Lys76Arg, which lost DNA packaging activity but not the *cos* cleavage activity (44). On the other hand, the C-terminal Walker A is required for *cos* cleavage and melting of the 12-bp GC-rich cohesive ends (77, 79, 116) (see below).

Walker B motif. Four hydrophobic amino acids forming a β-strand with a strictly conserved Asp at the tip of the strand (ZZZZD; Z represents hydrophobic amino acid), at about 50–130 amino acid distance from the Walker A Lys, constitute the Walker B motif. The Asp coordinates with the Mg of the Mg-ATP complex, precisely orienting the β,γ-phosphates for nucleophilic attack. CM of the T4 motif ($MIYID_{251–255}$) showed that an array of substitutions are tolerated at the β-strand "Z" residues as long as the hydrophobicity and β-strand character are preserved (107). On the other hand, any substitution at the conserved Asp255, including Asn or Glu, resulted in a null phenotype. The mutants bind ATP as well as the wild-type, but apparently not in a functional orientation, because they show complete loss of ATPase and DNA packaging activities.

Catalytic carboxylate. A carboxylate group, usually Glu, is oriented into the ATPase pocket and a water molecule links it to the β,γ-phosphates of ATP. Acting as a general base, the carboxylate activates the water molecule for a nucleophilic attack on the γ-phosphorus, resulting in ATP hydrolysis (140). In numerous large terminase proteins, a conserved Glu is present immediately adjacent to the Walker B Asp residue, a pattern similar to that in helicase ATPase centers (in RecA and other ATPases such as F_1F_0 ATPase, the Glu is located between the Walker A and Walker B motifs) (105). CM showed that no substitutions are tolerated at the conserved Glu256 of phage T4 gp17 (56).

ATPase coupling motif. An ATPase coupling motif, or C-motif, is found in the large terminase proteins. First identified in helicases as motif III (59), the C-motif consists of a tripeptide sequence T/S-G/A-T/S, 20–30 amino acids downstream of the Walker B motif (105). The third amino acid is critical and almost always a Thr or Ser but Asn is seen in some ATPases. The helicase C-motif forms a network of hydrogen bonds connecting the γ-phosphate of ATP and DNA. The helicase C-motif mutants generally retain ATPase and DNA binding functions but lose helicase activity, i.e., the coupling between ATPase and movement on DNA is defective (137). The phage T4 C-motif mutants (Thr287Ala/Asp) exhibit a novel phenotype; they can bind and hydrolyze ATP at least once but are deficient in turnover, thus losing steady-state ATPase and DNA translocation activities (41). In λ, the gpA mutation Gly212Ser alters the first of the three C-motif residues. This mutant terminase can sponsor the packaging of only part of the chromosome, underscoring the importance of the C-motif (44).

Adenine binding motif. A conserved two amino acid sequence, YQ, is present in large terminase proteins ~15 amino acids upstream of the Walker A lysine (105). The variable Tyr apparently forms hydrophobic stacking interactions with the adenine base, whereas the conserved Gln forms hydrogen bonds with the N6 and N7 of the heterocyclic ring. The Tyr46 from phage λ gpA ($YQ_{46–47}$) cross-links with

azido-ATP demonstrating its close proximity to the ATPase pocket, and Tyr46Phe/Ala mutations reduce ATP binding, hydrolysis, and in vitro DNA packaging, whereas *cos* cleavage is unaffected (66). CM of T4 gp17 indicates that the Gln residue is more critical than the Tyr for DNA packaging (K.R. Kondabagil & V.B. Rao, unpublished data).

Switch residue. The above functional signatures are essential to capture ATP and precisely orient it in the catalytic pocket for nucleophilic attack. A switch is, however, necessary to trigger phosphoanhydride cleavage and product release. The ATPases encode a switch residue(s), Arg or Gln (rarely, His) that stabilizes the transition state as well as aids the separation of the leaving group, Pi (3). Sequence and structural analyses suggest that Arg162 in the gp17 Walker A and analogous Arg in most large terminase proteins are candidates for the switch residue (122, 142). CM data show that no substitutions are tolerated at Arg162 (122).

Atomic structure. Numerous attempts to crystallize phage terminases have failed presumably because the protein, which plays dynamic roles in DNA processing and translocation, is a flexible, conformationally heterogeneous molecule. However, the N-terminal ATPase domain of T4 gp17 Walker B double "flip" mutant, ED (Asp255Glu-Glu256Asp), crystallized in apo-, ATP-, and ADP-bound forms and the structure was determined to 1.8 Å resolution (142). The mutant protein binds ATP tightly (107) forming a novel salt bridge between the Walker A Lys and the mutant Walker B Glu, thus locking the ATPase in a compact conformation.

The ATPase is a rather flat structure, consisting of a core β-sheet with six parallel β-strands, a classic Rossmann NBD fold with a topology similar to that found in other ATPases such as RecA (**Figure 3*b***). All the functional signatures of the ATPase center are positioned into the catalytic pocket as predicted by CM. Structural modeling generated very similar ATPase β-sheet cores for λ, T3/T7, SPP1, P22, φ29, and HSV large terminase proteins (41). The apo- and ADP-bound structures are virtually identical to the ATP-bound structure, with two significant differences. The adenine binding YQ loop appears to move in response to ATP hydrolysis (**Figure 3*b***), which may reflect the opening and closing of the ATP binding pocket for product release and substrate capture.

Despite lacking significant overall sequence similarity, the terminase ATPase domain of T4 gp17 has the closest structural similarity to the monomeric helicases such as the PcrA helicase from *Bacillus stearothermophilus* (148). The order and positioning of the β-strands in the six-stranded β-sheet as well as the interactions between the catalytic residues and bound ATP are similar. Phylogenetic analyses further suggest that the terminase ATPase has the closest evolutionary relationship to the monomeric helicases belonging to "superfamily 2" (SF2) helicases, and to a lesser extent, type III and type I translocating restriction enzymes (41).

Translocation ATPase. More than a dozen highly purified conservative mutants in the N-terminal ATPase center of T4 gp17 and λ gpA, for example Lys166Arg, Asp255Glu, Glu256Asp, and T287A in gp17, show a strikingly consistent biochemical phenotype: loss of ATPase, loss of DNA translocation, and retention of DNA cleavage. Combined with insights gained from the ATP-liganded T4 gp17-ATPase X-ray structure, the evidence overwhelmingly identifies this as the ATPase that powers viral DNA translocation. The phenotype of a putative second ATP binding site in the λ gpA is distinct as it is required for DNA cleavage, not translocation. An alternative possibility, however remote, is the appearance of a novel translocating ATPase upon assembly of the packaging motor, but this is unlikely since the T4 gp17 N-terminal ATPase mutants exhibit no prohead/DNA-stimulated ATPase activity or DNA translocation in a defined system. However, the basal ATPase activity associated with the large terminase proteins is weak, on the order of 1–2 ATP hydrolyzed/molecule/min

in T4 gp17 (8, 92). This must be stimulated by ~2000-fold to sustain the measured rates of translocation in phage T4. How this occurs is a fundamental question that remains unanswered.

Nuclease Center

Catalytic metal center. All well-characterized large terminase proteins (λ gpA, T4 gp17, SPP1 gp2, and T3 gp19) except φ29's gp16 have an endonuclease activity, which generates the termini of packaged genomes (11, 12, 37, 79, 151). Concatemeric DNA of *cos* and *pac* phages is first cut at or near a specific sequence and the end is inserted into the prohead to initiate packaging. Specific cutting requires the presence of the small terminase subunit (31, 65, 71). Mutations that inactivated the endonuclease activity in λ and T4 terminases affected residues in the C-terminal domains, showing that the endonuclease resides in the C-terminal domain (29, 37, 87).

CM of T4's gp17 identified a 10 amino acid segment containing acidic residues that are critical for nuclease activity (125). Secondary structure predictions suggest that the Asp and Glu residues are clustered in a hairpin loop created by two antiparallel β-strands; these are predicted to coordinate a metal ion (Mg^{2+}) to form a catalytic nuclease center. The transition intermediate is predicted to be a pentacovalent phosphorane that is common to diverse nucleases such as DNase I, restriction enzymes, integrases, RNAseH, and Holliday junction resolvases (87, 125) (**Figure 4**). The terminase endonucleases align most closely with the resolvase-type nucleases, and structural modeling of the phage T5 endonuclease showed an α,β-domain similar to that of RuvC resolvase with three Asp residues, a catalytic triad, coordinating with Mg (119). More extensive alignments and modeling show the presence of a RNAseH/resolvase-type β-sheet core in several phage terminase C-domains including phages T4, T7, and λ (**Figure 4**) (B. Draper & V.B. Rao, unpublished data).

The λ-like phages require an additional ATP-powered helicase-like activity to separate the annealed 12-base cohesive ends following cutting. There is evidence for an ATPase center in λ's gpA C-terminal domain, which is required for cutting and strand separation (77, 116).

Headful nuclease. A second, downstream cut is made following encapsidation of a unit-length viral chromosome. Whether the endonuclease cuts specifically, as in *cos* phages, or nonspecifically, as in *pac* phages, the downstream cut is tightly linked to capsid filling, which is equivalent to one (for *cos* phages) or 102–110% (for *pac* phages) genome (97, 98, 141, 149). Evidence suggests that the same C-terminal nuclease center carries out both initiation and termination cutting (48). Explanations for headful nuclease control include the idea that the nuclease is not able to cut DNA that is being translocated at a high rate, and that resistance to further packaging eventually slows or stalls translocation, permitting cutting to occur (32–34, 153, 155). Another proposal is that head filling causes a conformational change in the portal, causing terminase dissociation and DNA cutting (4).

Several observations implicate portal involvement in headful cutting. Certain *siz* mutations in phages P22 and SPP1 either overfill or underfill the head, consistent with the portal acting as a headful packaging sensor (18, 144). The structure of the virion portal differs from that of the purified, ectopically expressed portal, suggesting that a DNA packaging–induced conformational switch could trigger the nuclease to cut, or terminase to release (156, 166). Also, in the virion, a loop of DNA is tightly wrapped around the wing domain of the portal. This DNA, presumably the last bit packaged, might be involved in flipping the switch (89).

An obvious problem with implicating a causative relationship between a portal structural change and headful terminase cutting is that the change is observed in the final mature virion whereas all the DNA packaging steps and tail attachment happen between the points

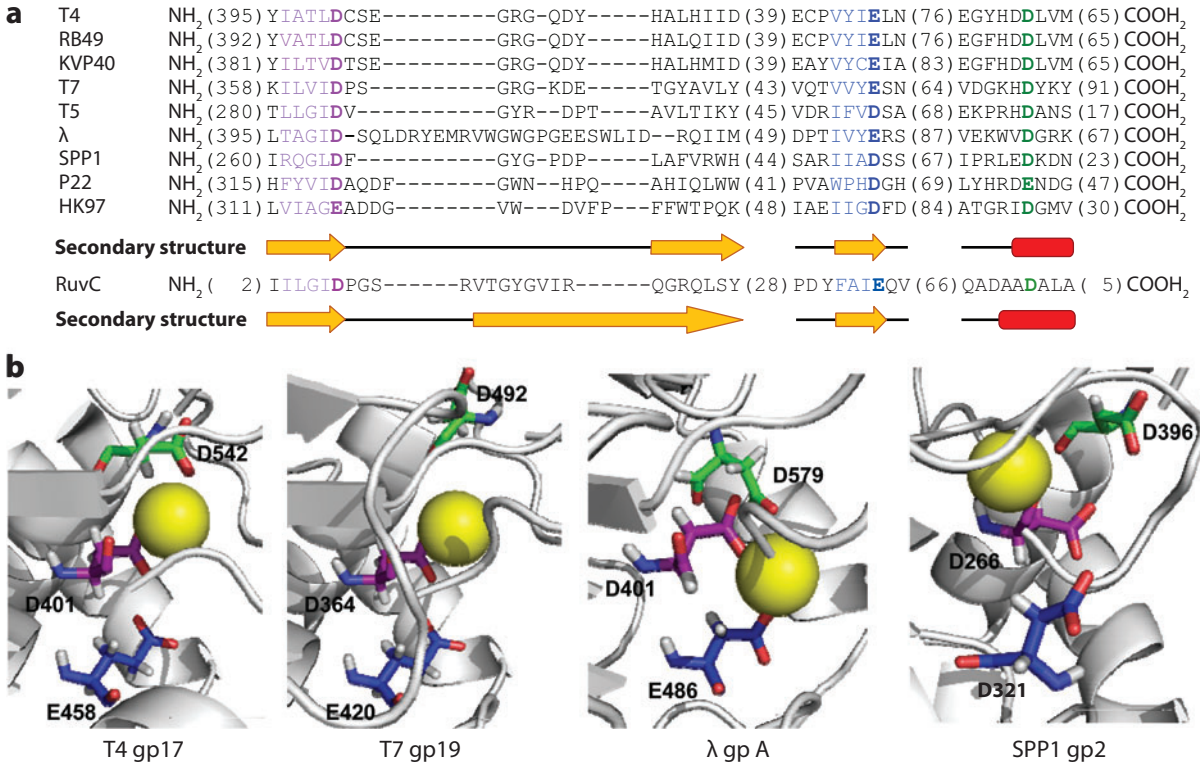


Figure 4

The C-terminal nuclease center. (*a*) The C-terminal nuclease center of phage terminases contains a catalytic triad of Asp/Glu residues similar to that of resolvase RuvC. The C-terminal nuclease domain of phage terminases were aligned with CLUSTALW. The default alignment was then optimized based on the secondary structure prediction obtained from Jpred. Numbers in parentheses represent the number of amino acids. The first conserved Asp (*dark purple*) is located at the tip of a conserved β-strand. The second Asp (*dark blue*) is located approximately 60 residues downstream, at the tip of a short conserved β-strand. The third Asp (*green*) is within a structurally adjacent α-helix. Structural models of phage terminase nuclease domains were obtained by homology modeling using SWISS-MODEL. The target sequence was restricted to approximately 150 amino acids encompassing the three conserved residues, in order to optimize the sequence alignment to the template. The Mg ion was modeled by superimposition with the RuvC resolvase structure (PDB ID: 1HJR). All images were created with PyMol.

when the portal is part of a naïve prohead and a mature virion. Furthermore, late DNA packaging events appear to be rather complex and not well characterized for any virus. In phage λ where there is some detailed information, cutting the downstream *cosN* requires recognition of an upstream *cosQ* sequence (32). *cosQ* mutations lead to failure to arrest translocation, and there is a failure to nick the bottom strand in the rotationally symmetric *cosN* (34). *cosQ* is recognized in context with *cosN* (34). Leaky *cosQ* mutations are suppressed by increasing chromosome length or portal mutations, suggesting that the translocation rate affects the efficiency of *cosQ* recognition. Fast translocation may account for the absence of *cos* cleavage for short, *cosQ*$^+$ chromosome lengths (33, 84, 155). The bottom strand nicking failure suggests that *cosQ* may sponsor remodeling of terminase to present a properly oriented gpA (+gpNu1?) for nicking the bottom strand. The gpA presentation might occur by a reconfiguring of a gpA of the translocation complex or by recruiting a gpA from solution. Whether a similar remodeling occurs in *pac* phages with headful cutting mechanisms is unclear. The idea that

the endonuclease is in an inactive configuration during translocation is attractive.

DNA Binding Motifs

At least three key steps of the DNA packaging process require dynamic interactions between the terminase and DNA. First, a DNA binding motif in terminase recognizes the viral DNA. As described above, the HTH motif in the small terminase protein is required for DNA recognition. Second, a precisely aligned nuclease-DNA complex forms in order to cut concatemers to initiate and terminate DNA packaging. Third, the packaging motor must interact with the DNA during translocation, trigger directional movement, and release the DNA. Numerous such bind-release cycles processively translocate the viral genome. Different functional requirements of these steps would dictate that these are distinct DNA binding sites. The ATPase domain of T4 gp17 binds DNA (4); so do the large terminase proteins gpA of λ (116) and gp16 of φ29 (91). Whether these sites are involved in DNA cutting or DNA translocation remains unknown.—3

STOICHIOMETRY AND SYMMETRY

Portal Vertex

Despite having no significant sequence similarity, all portals from mature phage virions as well as the baculovirus-expressed herpes virus portal show a conserved dodecameric ring structure with 12-fold rotational symmetry (**Figure 1**) (40, 90, 147). Variants such as 11-, 13- and 14-fold symmetries have been documented, but these are seen only in the heterologously overexpressed proteins (**Figure 2*b***) (43, 115). Comparison of 13-fold and 12-fold SPP1 portal structures shows flexible intersubunit interactions (90), but the fact that only the 12-fold symmetry is incorporated into mature phage suggests that the dodecamer is a functional necessity and that portal assembly is tightly regulated in vivo.

Small Terminase Protein

All the *E. coli*-overexpressed small terminase proteins show high propensity to self-associate, forming oligomers of various stoichiometries. The T4 gp16 forms a donut-shaped oligomer with a central open cavity and no discernible surface features. STEM measurements estimate 8 ± 3 subunits per ring (95) but the subunits could not be counted by rotational symmetry analysis (A. Fokine & M.G. Rossmann, personal communication). Double rings, which are likely figure 8-shaped helical structures, and strings of single and double rings are also seen (85, 95). The phage λ gpNu1 forms large amorphous aggregates of no particular stoichiometry. However, gpNu1 lacking the C-terminal half, or in the presence of gpA at low protein concentrations, forms dimers that retain *cos* binding activity (7, 159). T7 gp18 and SPP1 gp1 oligomerize into octamers and decamers, respectively, whereas P22 gp3 oligomerizes into nanomers (23, 112, 152). But a gp3 mutant, A112T, oligomerizes into a decamer. The gp16 and gp3 oligomers have a central channel of approximately 2 nm diameter (95, 112).

Although oligomerization is important (85), stoichiometry or a defined central cavity does not appear to be strictly essential for function. However, interactions with other components of the packaging machine may select for a defined stoichiometry in vivo.

Large Terminase Protein

The *E. coli*-overexpressed and purified large terminase proteins from many phages exhibit different oligomeric states. T4 gp17, SPP1 gp2, T3 gp19, and P22 gp3 exist essentially as monomers in solution (60, 92, 111, 112). φ29 gp16 and λ gpA oligomerize into heterodisperse complexes (62, 132). Diluted gpA-gpNu1 complexes dissociate into a stable heterotrimer of one gpA and two gpNu1 subunits (99). Upon concentration, the heterotrimer oligomerizes into a tetramer of heterotrimers (101). The latter can catalyze *cos* cleavage and in vitro DNA packaging, whereas the simple

heterotrimer requires the assistance of the *E. coli* host factor, IHF. However, during assembly of the packaging machine the stoichiometry of the large terminase protein may be remodeled.

Assembly of motor complex. The stoichiometry of the large terminase protein is thought to vary depending on the functional state of the packaging complexes, whether it is the packaging initiation complex, the translocation motor complex, or the packaging termination complex. Perhaps the most critical of these is the motor complex formed in association with the prohead portal and DNA, which powers translocation. Genetic and biochemical evidence in λ, T3, SPP1, and T4 suggest that the large terminase protein interacts with the stalk of the portal, forming a ring of subunits (58, 94, 110, 162, 163). The loop residues of the outer αβ-domain of the portal, e.g., Asn290 of SPP1 gp6 (36), Pro331 of λ gpB, and amino acids 300–310 of T4 gp20 (the gpB and gp20 residues are presumed to be in a similar position as that of the gp6 X-ray structure), appear to provide docking sites for the large terminase protein.

The portal binding site in the large terminase protein is localized to the last 15 amino acids of λ gpA and T3 gp19. The putative binding motif corresponds to $LYWEDD_{571-576}$ and $LSGEDE_{636-641}$ in gpA and gp19, respectively, implicating hydrophobic and charge-charge interactions between terminase and portal (110, 162, 163). In phage T4, however, the extreme C-terminal sequence of gp17 is not required for portal interaction. Deletion of the last 33 amino acids results in no loss of ATPase and DNA packaging functions (56). Second site suppressors of gp20 *cs* mutants map to the central (Ser336Asn) and C-terminal (Ser583Asn) regions of gp17 (95). Both the N-terminal ATPase domain and the C-terminal nuclease domain inhibit in vitro DNA packaging, presumably by competing with the full-length gp17 for portal binding (83). Therefore, unlike in λ and T3, the portal binding determinants in the phage T4 large terminase subunit are not restricted to a narrowly defined region at the C terminus, although it is unclear exactly where the sites are located.

Stoichiometry of motor complex. Various stoichiometries of the large terminase protein assembled on the dodecameric portal have been suggested. In phage T3, direct binding experiments estimated that 6 molecules of gp19 bind to a single prohead particle (111). In phage λ, a tetrameric holo-terminase composed of four heterotrimers of 2 gpNu1:1gpA has been proposed to be the active oligomer (101). In phage SPP1, a complex containing a monomer of gp2 and a dimer of gp1 decamer rings was suggested to be the portal interacting unit (16).

In phage φ29, a unique 174-nt pRNA molecule is the primary portal-interacting component (63). The packaging ATPase, the large terminase protein equivalent, is attached to the pRNA but not to the portal (135; M.C. Morais & M.G. Rossmann, personal communication). The pRNA is a Y-shaped, mostly double-stranded, molecule with two upper arms consisting of complementary single-stranded loops at the ends. Base pairing between the loops interlocks the monomers into a ring structure. Biochemical experiments originally suggested that the pRNA ring formed in the absence of proheads is a hexamer (145, 165). The prohead-bound pRNA is also reported to be a hexamer; the stoichiometry was inferred by imaging the total number of quantized steps required to photo-bleach the prohead-bound fluorophore-coupled pRNA (134). However, cryoEM reconstruction of the prohead-bound packaging motor showed a pentameric pRNA ring with the base-paired upper arms in contact with the portal and five spokes of the lower arm projecting out below the portal vertex (135). Recent more refined and higher-resolution reconstructions of prohead-pRNA-ATPase complexes confirmed the pentamer stoichiometry of both pRNA and gp16 ATPase (M.C. Morais & M.G. Rossmann, personal communication).

Symmetry Matches and Mis-Matches

There are basically four key essential parts in the actively packaging motor: prohead, portal vertex, large terminase protein, and DNA. In φ29, the pRNA and the gp16 ATPase together constitute the structural equivalent of the large terminase protein. The small terminase protein, although essential in vivo, is not essential in many in vitro DNA packaging systems and thus is not an integral part of the translocation mechanism per se. The symmetries of three of four essential parts of the motor are well established; the prohead's portal vertex—5-fold, the portal—12-fold, and the DNA—10_1-fold. However, the stoichiometry of the large terminase protein remains uncertain and controversial. The two basic possibilities are: (*i*) terminase stochiometry matches with the portal symmetry, e.g., hexameric pRNA/ATPase (φ29) or tetrameric gpA (λ); in this scenario, the portal and terminase symmetries do not match with the 10_1-symmetry of DNA; and (*ii*) terminase stoichiometry is mismatched to the portal symmetry, e.g., pentameric pRNA/ATPase, but it matches with the DNA symmetry.

The best available evidence points to the pentamer stoichiometry of terminase, as observed in the cryoEM reconstruction of the φ29 packaging motor. The pentamer stoichiometry fits well with the 10_1-fold DNA symmetry and 2-bp step size, but it is in conflict with the 12-fold portal symmetry. The latter can be largely reconciled if there were interactions between the terminase and the capsid. Indeed, this is evident from the overlapping capsid and pRNA densities in φ29 reconstructions (135). Therefore, it is probable that, following initial interaction with the portal, the pRNA/large terminase protein establishes contact with the fivefold capsid vertex, thus maintaining overall symmetry between the capsid, the packaging motor, and the DNA. But how the pentameric large terminase protein dynamically interacts with the dodecameric portal during translocation, as evident from the portal's ATPase stimulation and headful sensing roles, remains an interesting and puzzling question.

PACKAGING MODELS

Several packaging models have been proposed to define the basic mechanism of DNA translocation. Most models assume that a prohead-bound packaging motor, using ATP as fuel, generates motion of certain part(s) of the motor that is coupled to DNA movement. All models agree that the DNA motion must be translational (linear), involving no significant rotation of DNA, which would otherwise introduce supertwists and eventually stall the motor, as well as present serious problems for delivery of DNA into the host cell. Slight DNA rotation, likely of the DNA outside the capsid, would, however, be necessary to compensate for the torsion introduced as the DNA is wound around the capsid interior. The main differences are with respect to which part of the motor moves DNA and the precise mechanism that causes motions. The models fall into two basic types: (*i*) Rotary motors—the motor consists of a stator and a rotor, the portal, and rotary motion of the rotor is coupled to linear motion of DNA; although rotary motors have been well documented in biological systems (e.g., F_1F_0 ATP synthase; flagellum), coupling between rotary and linear motions is unique and has not yet been experimentally demonstrated; (*ii*) linear motors—the motor causes linear motion of a part of the motor, the terminase, which is coupled to linear movement of its partner, the DNA; many such motors have been described in biological systems (e.g., myosin, helicase).

Rotary Motors

Nut and bolt model. In 1978 it was proposed that the portal is not merely a passive conduit for DNA entry and exit but an active packaging machine, a rotary motor that transports DNA into the capsid (67) (**Figure 5*a***). The basic features of the model are: (*i*) The symmetry mismatch between fivefold icosahedral capsid and sixfold portal vertex allows portal rotation imposing minimal energy barriers; (*ii*) the portal channel must be threaded to match the DNA structure, analogous to a nut that rotates on a

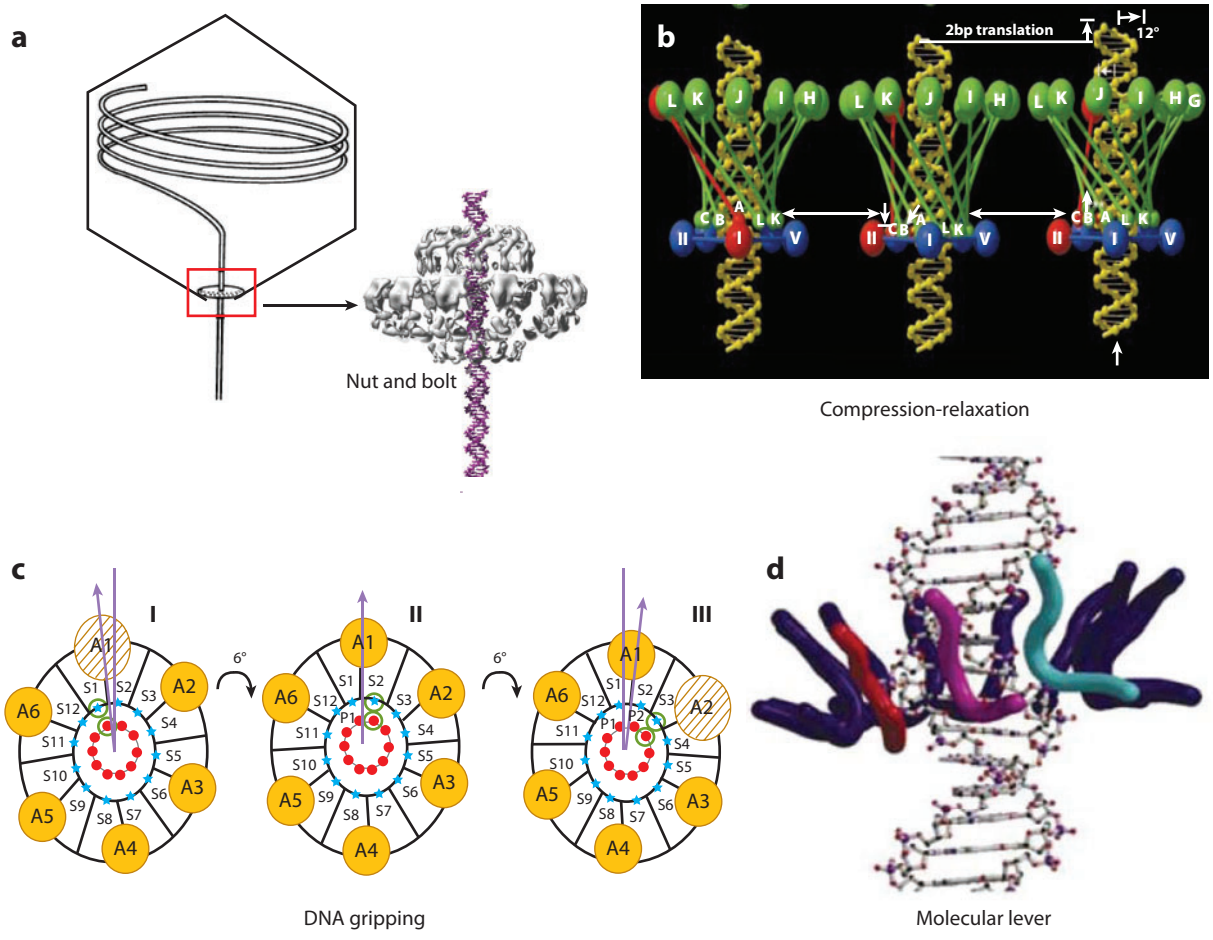


Figure 5

Portal rotation models.

bolt; (*iii*) directional rotary motion of portal (nut) causes linear motion of DNA (bolt) into the capsid; and (*iv*) ATP hydrolysis powers directional rotation; 30 equivalent positions are expected where a rotating portal subunit comes into registry with a capsid subunit, triggering ATP hydrolysis. The free energy of hydrolysis drives portal rotation through one to several of these equivalent positions, coupling it to translocation of a proportional number of base pairs. This could be tailored, or regulated, to generate different gearing ratios.

Insights from portal structure. The portal rotation model attracted wide attention, has been refined as details emerged (43), and subjected to extensive genetic, structural, and biochemical analyses (147), but testing of rotation has not been possible until recently because of large gaps in the available structural information and technical barriers. The structural data, however, provided an initial test of the model.

As described above, cryoEM reconstructions and/or X-ray structures of phage portals and herpes virus portals verified the symmetry mismatch and showed remarkable conservation of the overall structure (2, 40, 42, 43, 146) (**Figure 2**). The portal vertex turned out to be a dodecamer, not hexamer as was originally predicted. A dodecamer fits the rotation model just

as well as a hexamer except that there will be 60 equivalent positions, instead of 30, where the portal and capsid subunits come into registry.

The X-ray structures of both φ29 and SPP1 portals exhibit conserved features (61, 90, 135). The φ29 portal is a 75 Å long cone-shaped structure with three basic parts (**Figure 2*b***); a 70 Å wide end that is inserted into the icosahedral vertex, a 28 Å long central stem, and 25 Å narrow stalk that protrudes out of the capsid. A central channel, also cone-shaped, with 35 Å inner radius traverses the portal at the narrowest outer end, gradually widening to ∼60 Å at the inner end. The narrow end, consisting of a α/β domain, is connected to the central stem of two long α-helices, α3 and α5, which in turn is connected to the wide end formed by additional α-helices and a β-sheet. In the SPP1 portal, the wide end has a more massive ∼150 Å wing and crown (**Figure 2*b***).

The channel, the passage-way for translocating DNA, appears to be the most conserved feature of the portal. Its wall is lined by 24 α-helices radiating from the center at a ∼45° angle relative to the central axis. Acidic residues coat the channel making it highly electronegative, a feature consistent with repulsion and smooth passage of negatively charged DNA. There are two constrictions, one at the wide end and another at the narrow end, areas where the channel surface can come into close proximity to translocating DNA. Disappointingly, however, and inconsistent with the nut and bolt model, the portal channel features no obvious "threads" that are complementary to the double helical grooves. Nevertheless, the structural details are consistent with portal rotation and provided insights for developing newer rotation models.

Compression-relaxation model. In this model, ATP hydrolysis drives lengthwise expansion and contraction of the portal, which causes portal rotation and coupling to DNA motion (**Figure 5*b***) (135). The cryoEM reconstruction of φ29's packaging motor shows a stoichiometry of 12 subunits of portal, 5 subunits of pRNA, and 5 subunits of gp16 ATPase. The pRNA ring, present in the middle, interacts with the capsid and portal on one side, and the ATPase on the other, "gluing" the parts together. The portal channel thus is extended by the pRNA and ATPase rings, with the DNA placed in the center. The portal is the rotor, and the capsid, pRNA, and ATPase together constitute the stator. At any time, only one portal subunit (subunit 1, for example) can be in alignment with one pRNA/ATPase complex due to the symmetry mismatch. The aligned portal-pRNA/ATPase interactions trigger "firing" of that ATPase molecule causing lengthwise expansion of the portal through slight rearrangement of the channel helices, allowing the portal to grab the DNA by making nonspecific contacts with it. Relaxation to normal length causes upward movement of DNA by 2 bp and passive counterclockwise rotation of the portal by 12°. The rotation brings another portal subunit (subunit 3) into alignment with the adjacent ATPase molecule, causing its firing and DNA translocation. Thus portal rotation is necessary primarily to engage all the ATPase molecules of the motor.

A central feature of this model is that the symmetry of pentameric pRNA and ATPase is matched with the capsid but mismatched with the portal, as evident from the cryoEM reconstructions. While this needs to be reconciled with the fluorescence imaging results that suggest a hexameric pRNA, the model offers a plausible role for the conserved angular disposition of channel helices. There is evidence for inter-portal subunit flexibility and motion (90); variant portal oligomers of 11, 13, or 14 subunits, which must involve slightly different intersubunit interactions, readily form in the ectopically expressed protein (43). Disulfide cross-linking of adjacent portal subunits that curtail motion abolish DNA packaging in SPP1 (36). However, there is as yet no evidence for lengthwise expansion of the portal structure, a substantial >6 Å movement of each channel helix, and it is conceptually unclear how the free energy from a single ATPase firing can be transmitted to all the portal subunits. If 12° portal

rotation is strictly followed, half of the portal subunits would never engage in translocation, a feature at odds with the highly conserved 12-fold portal symmetry.

DNA gripping model. The φ29 portal consists of 12 alternating Lys and Asp residues lining the narrow end constriction of the portal. In this model, positively charged lysines are proposed to form an electrostatic grip on the negatively charged DNA phosphate backbone, which is coupled to portal rotation and DNA movement (**Figure 5*c***) (61). In this model the stoichiometry of the φ29 packaging motor is assumed to be 12 subunits of portal, 6 subunits of pRNA, and 6 subunits of ATPase. Thus all parts of the motor are symmetrically matched and the whole packaging motor is a rotor and the capsid becomes the stator.

The two electropositive rings formed by Lys200 and Lys209 are at a distance of 20 Å with a 10 Å spacing between adjacent lysines. When a Lys200 residue is aligned with a backbone phosphate, another Lys200 belonging to the fifth portal subunit is also aligned with another phosphate of the opposite strand. Similarly, 20 Å (6 bp) down the channel, two Lys209 residues align with two additional phosphates. This configuration brings into alignment a portal subunit, four DNA phosphates, a pRNA subunit, and an ATPase subunit, causing ATPase firing, 12°-degree clockwise rotation, and 2-bp DNA translocation. Rotation disrupts the lysine-phosphate interactions allowing DNA movement, and brings four neighboring lysines in alignment with another set of backbone phosphates to reform the electrostatic grip. Either the ATPase firing causes the portal to rotate, which pushes the DNA, or the ATPase itself pushes the DNA and the portal rotates passively.

In this model, rotation is necessary to match the motor symmetry to 10_1-DNA screw symmetry, but the hexamer stoichiometry of pRNA and ATPase (80) is in conflict with the cryoEM evidence (135). A single ATPase subunit must support rotation of the whole motor - not just the portal, pushing of DNA, and disruption of lysine-phosphate bonds, a rather tall order. The 35 Å diameter of the portal channel at the narrow end is quite large for maintaining a firm grip on DNA that has a diameter of 20–23 Å, leaving a ~6 Å gap between the lysine and the phosphate. Nonradial orientation of Lys200 and the observed salt bridge between Lys200 and Asp202 (135) are also inconsistent with the model.

Molecular lever model. The pseudo atomic structure of dodecameric SPP1 portal shows 12 loops projecting into the wide end constriction (**Figure 2**), compressing the open radius of the portal channel to just about 18 Å, 5 Å less than the diameter of the DNA (**Figure 5*d***) (90). These ~15 amino acid-long "tunnel" loops, also predicted in other phage portals, are disordered and not seen in the φ29 X-ray structure. Each loop connects two helices, the α5 channel helix and the α6 helix, that is perpendicular to α5 with an unusual 135° kink in the middle. This substructure is held together by hydrophobic interactions and is hypothesized to function as a movable molecular lever to translocate DNA (90). There is structural evidence that the tunnel loops exist in at least two conformations, up and down.

The molecular lever model proposes that the tunnel loops tightly grip DNA and undergo sequential conformational changes that are coupled to portal rotation and DNA translocation (90). The axis of the DNA would be slightly tilted to accommodate the 23 Å wide DNA in the 18 Å portal constriction. The tunnel loops must be pushed in and asymmetrically arranged around the DNA structure. The relatively hydrophobic tunnel loops are not expected to interact with the backbone phosphates, instead form nonspecific shape-matching van der Waals interactions with the major groove. The loops adjust to the 10_1-fold symmetry of the DNA by assuming different conformations at different contact points. An "undulating belt" is formed around DNA with 12 contact surfaces forming a tight embrace, and even squeezing the DNA. ATP hydrolysis causes a 12° counterclockwise portal rotation

and loop sliding and repositioning in a new conformation, which again matches the DNA structure. This process, sliding and repositioning of one loop causing the same in the adjacent loop and so on, spreads like a Mexican wave, resulting in directional translocation of DNA.

Several SPP1 mutations that interfere with packaging were mapped to loop residues (amino acids 343–360) (81). One of these mutants, Glu352Gly, showed ~50% reduction in packaging efficiency, whereas other mutants such as Val347Met/Ala showed no packaged DNA, and the Ser343Phe mutant packaged a smaller viral chromosome producing a *siz* phenotype (114). However, in phage φ29, loop mutations did not significantly interfere with DNA packaging. Substitution of 1, 2, or 3 loop residues with Ala, or even deletion of the entire 18 amino acid loop, showed no major reduction in packaging efficiency. However, unlike the wild-type motor, the loop mutants released the packaged DNA during gradient centrifugation, suggesting that the loops may be involved in stabilizing the packaged DNA but not translocation per se (R. Atz, S. Grimes & D. Anderson, personal communication).

Test of portal rotation. The portal structural features unfortunately do not compellingly support the specific packaging models invoking portal rotation. Therefore, genetic and biophysical approaches were used for testing portal rotation.

Using a tethering approach, the phage T4 portal was attached to the capsid through Hoc interactions (9). Hoc is a nonessential T4 outer capsid protein that binds as a monomer at the center of the major capsid protein hexon (49). Hoc binding sites are not present in the unexpanded proheads but are exposed following capsid expansion. With 155 hexons forming the shell, the T4 virus is decorated with 155 copies of Hoc. To tether the portal, unexpanded proheads were first prepared with 1 to 6 of the 12 portal subunits replaced by Hoc-portal fusion proteins (the C terminus of Hoc is fused to the N terminus of portal) (at least 6 of the 12 portal subunits must be wild-type for phage viability). The proheads were then expanded to expose Hoc binding sites. The Hoc portion of the portal fusion is expected to bind to the center of the nearest hexon, tethering 1–6 portal subunits to the capsid. It was argued that the Hoc-capsid bonding is equivalent to a covalent cross-link and thus must curtail portal rotation. If portal rotation were to be central to DNA packaging, the tethered expanded proheads should show very little or no packaging activity. However, the efficiency and rate of packaging of tethered proheads were comparable to those of wild-type proheads, suggesting that portal rotation is not an obligatory requirement for packaging.

The caveats of this test are that (*i*) the number of tethers may be lower than expected due to the symmetry mismatch at the portal vertex, and (*ii*) the packaging force, supposed to be very powerful, may disrupt the tethers. Recent evidence suggests that Hoc's capsid binding site is located in the C terminus, and fusions to the C terminus show greatly reduced binding affinity when compared to the N-terminal fusions (131, 133). Thus, while the data argue against portal rotation, rotation cannot be ruled out (100).

Single molecule fluorescence spectroscopy was used to test portal rotation (74). Recombinant empty proheads were prepared with a single cysteine per portal subunit and no other cysteines in any of the other prohead proteins. Six such cysteines were independently inserted into the exposed regions of the portal, in both narrow and wide ends. Fluorescent Cy3 dye was coupled so that only one of the 12 portal cysteines was labeled and the mutant proheads are fully functional for DNA packaging. ATP-γ-S stalled single packaging complexes were prepared on a mounted slide and packaging was resumed by exchanging with excess ATP. If the portal were to rotate, the fluorescence intensity must change at a frequency of 1-2 Hz due to polarization, at the determined packaging rate of ~60 bp/sec. Attempts to identify a signal indicative of portal rotation during active filling of proheads were negative. Fifty independent measurements of actively packaging motors using Cy3-cysteine reporters at six different locations of the portal structure

did not show any incidence of portal rotation relative to the capsid.

In conclusion, in addition to providing a docking site for the large terminase protein, the highly conserved structural features imply that the portal is an important part of the packaging motor. Complete loss of DNA packaging by disulfide cross-linking of adjacent portal channel helices and resumption to packaging following reversal of cross-links signify an active role (36). Portal alone, or portal-containing proheads, stimulate the ATPase activity of the large terminase protein, and a strict correlation between ATPase stimulation and DNA packaging was observed (8, 114). Thus, dynamic movement of channel helices and communication between the portal and the large terminase protein seem to be central to the DNA packaging mechanism. But, the evidence argues against directional portal rotation being the driving force for DNA movement.

Linear Motors

Beginning in 1978 (13), models implicating terminase directly in the translocation mechanism have been introduced (45, 46, 52, 122). The underlying themes in these models are that (*i*) there is no obligatory requirement for, or coupling to, portal rotation; (*ii*) terminase is both the provider of ATP energy and the DNA pump; and (*iii*) there is a parsimony in proposing that both the ATPase center and the mechanical components of the motor are located in the same molecule. Models include supercoiling-relaxation, DNA tracking, conformational switching, or inchworming by terminase linked to translocation. Basically, in these models linear DNA translocation is achieved through linear motion, not rotation, of terminase domains.

Supercoiling model. This model proposes that the first DNA end entering the capsid would be fixed to the portal or the prohead, allowing the terminase to introduce supercoils much like an ATP-dependent DNA gyrase, resulting in transient storage of part of the ATP hydrolysis free energy as DNA torsion (13). Relaxation of the supercoils unleashes this energy, propelling DNA into the capsid. Supporting this model is the observation that phage T4 DNA ligase mutants are blocked in DNA packaging, apparently because of failure to repair nicks (164). Since nicks dissipate torsional energy, DNA translocation would be blocked; in principle, a single nick can arrest DNA packaging. However, in several defined in vitro phage DNA packaging systems including T4, T3, and φ29, nicked DNA can be packaged (51, 108; L. Black, personal communication). Packaging seems unaffected when the motor has to pass several nicks in a single packaging event. Cross-linked DNA can also be packaged (51). Thus there is as yet no evidence that the ATP free energy is stored in the form of torsion or an altered DNA structure. The data, however, suggest that the packaging motor does require a double helical DNA structure, although it can tolerate minor perturbations such as nicks. The in vivo packaging block in ligase mutants may be due to larger problems in DNA repair, not necessarily simple nicks. For instance, gaps and large heteroduplex loops block DNA packaging (51, 118).

Inchworm model. Although the packaging ATPase features a prototypical ATPase structure, it exhibits closest similarity to the ATPase domain of monomeric SF1 and SF2 helicases (41, 142). Disposition of certain ATPase signatures is also very similar; the catalytic glutamate of the terminases and helicases is immediately adjacent to the Walker B aspartate, whereas in RecA, F_1-ATPase and many other ATPases it resides between the Walker A and Walker B motifs (105, 140). A *trans*-peptide is present next to the Walker B Asp in the gp17-ATPase and monomeric helicases, whereas a *cis*-peptide is present in other ATPases, orienting an upstream glutamic acid into the active center (142). An ATPase coupling C-motif is present at the tip of the β_5 strand in both gp17 and monomeric helicases, orienting a threonine or serine into the catalytic center for interaction with the γ-phosphate of ATP (41, 105).

These structural and sequence similarities imply that the terminase is an active translocase achieving DNA translocation by an inchworm-type mechanism similar to that observed in helicase translocation on DNA (or RNA) (139, 148).

In helicases (and other ATPases), an ATP binding cleft is formed by two NBDs, each featuring a 5- or 6-stranded parallel β-sheet (128) and connected by a hinge (27). One of the NBDs (NBD-1) possesses the canonical ATPase signatures such as Walker A, Walker B, and catalytic carboxylate, whereas the second NBD (NBD-2) provides the arginine switch residue(s) and additional ATP interactions. The cleft exists in at least two conformations: a "closed" conformation as a result of ATP binding and closure of domains, and "open" conformation following ATP hydrolysis and product release. The T4 gp17-ATPase structure consists of helicase NBD-1 domain (subdomain I) and a smaller subdomain consisting of residues from both the N-terminal (residues 1–58) and C-terminal (residues 314–360) domains (subdomain II; **Figure 3*b***) (142). The subdomain II is linked to the nuclease domain (residues 361-610) that has a similar structural fold as the RuvC resolvase (**Figure 4**).

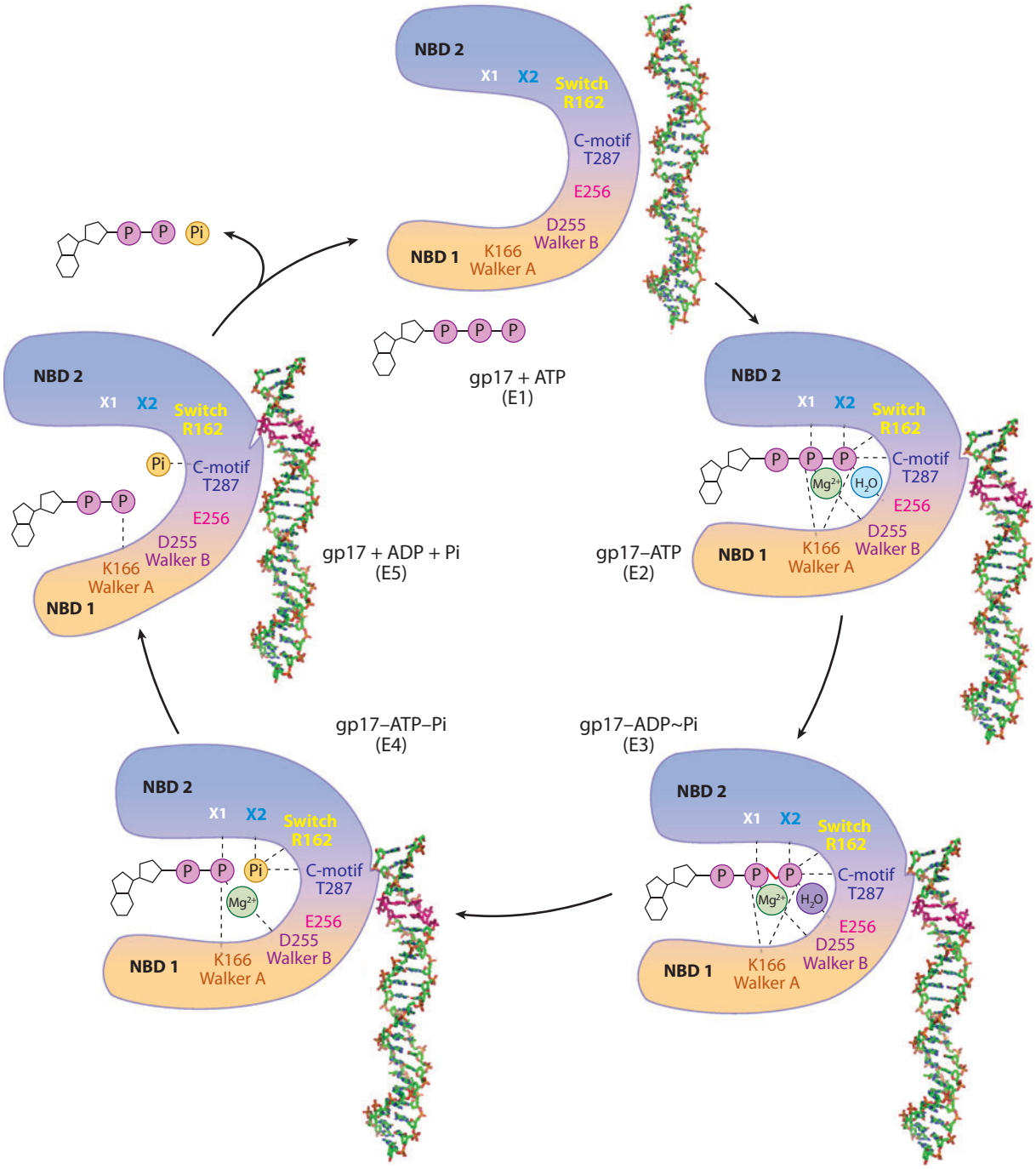
These features and biochemical evidence led to the proposal of an inchworm-type mechanism for viral DNA translocation (41). In the model shown (**Figure 6**), opening and closure of the ATP binding cleft cause movement of terminase domains, which is coupled to DNA motion. In the apo state, the large terminase protein exists in an open conformation (E1). ATP capture facilitates large-scale interdomain interactions, closing the cleft and probably introducing a degree of strain into the structure (E2) (117). In this conformation, the terminase will have the highest affinity for DNA. With the ATP and DNA bound and precisely oriented, the arginine switch is engaged and the transition state (E3) is stabilized. Nucleophilic attack by an activated water molecule hydrolyzes ATP (E4), pulling ADP and Pi apart and untethering the interdomain interactions. Opening of the cleft (E5), a force generating conformational change, causes DNA translocation and product release. The protein resets to apo state to reload ATP for the next catalytic cycle.

The DNA binding site involved in the translocation mechanism must display unique features. Unlike the *cos/pac* DNA recognition site, the translocation site should form transient, yet tight, interaction with certain groups of DNA, for example the backbone phosphates, and release the DNA after each translocation step. Thus this site functions as a "hook," a mechanical part that grabs and moves DNA, but its binding affinity must dynamically oscillate between high- and low-affinity states in coordination with ATP binding and hydrolysis, respectively (4).

The large terminase protein likely assembles as a ring of five subunits around the portal vertex. The pentamer stoichiometry fits with the 10_1-fold symmetry of DNA and 2-bp step size (135, 142). Pushing DNA by 2 bp by one subunit brings another terminase subunit into registry with the DNA, triggering the next cycle of DNA binding, ATP hydrolysis, and conformational transitions causing DNA movement and product release. Sequential firing of the

Figure 6

An inchworm-type viral DNA translocation model. An ATP binding cleft is formed between two NBDs of the T4 large terminase protein, gp17. NBD 1, containing the well-conserved core with classic ATPase functional signatures, corresponds to subdomain I of the ATPase structure (see **Figure 3*b***). NBD 2, which is not yet known (subdomain II, or C-domain?) is presumed to contain the Arg162 switch residue and additional interacting residues yet to be identified (X1 and X2). In the apo state, the terminase has low affinity for DNA (E1). ATP binding (E2) facilitates interaction between the two NBDs; closing the cleft and opening a high-affinity dsDNA binding site in the gp17 molecule. At the transition state (E3), the catalytic carboxylate activates a water molecule (*purple*) and cleaves the β,γ-phosphoanhydride bond of ATP (E4). The C-motif disengages the cleaved phosphate, untethering the interdomain interactions and opening the ATP binding cleft (E5). This conformational change causes product release and the associated DNA is translocated by 2-bp into the prohead. ATP binding resets the motor for the next translocation cycle.



DNA SYMMETRY

The packaging models assume that the B-form double helical DNA has perfect tenfold helical (screw) symmetry. In the literature, both 10- and 10.5-fold symmetries are used. However, the DNA does not have a fixed fold symmetry. The angle of rotation to observe symmetry does not repeat exactly in each rotation. Thus the double helical B-DNA has "nonrepeating" helical symmetry. Therefore the motor subunits must be flexible to adjust to the imprecise symmetry as they grab the DNA and translocate it. Strand switching at certain points during translocation may also allow the motor to compensate for the symmetry imprecision.

subunits hands over DNA from one subunit to another, translocating 10 bp, or one helical turn per round. The subunits must be flexible enough to adjust to the imprecise nonrepeating helical symmetry of DNA (See sidebar, DNA Symmetry). In addition, concerted movements of portal helices, or slight expansion and relaxation of the portal channel, might be necessary to correctly position DNA in the channel, as well as prevent its exit during the handovers. The latter becomes particularly important in the late stages of genome packaging when the internal force works against the packaging force and slows down the motor.

Although largely speculative, the inchworm model is consistent with the available body of information on DNA packaging motors, and certain aspects of the model are similar to the recently described φ12 ssRNA packaging motor (103), and other motor proteins such as helicases (138) and F_1F_0 ATPases (1). There is evidence linking phosphate release to force generation and DNA motion in the φ29 packaging motor (24). The T4 gp17 mutants in the ATPase coupling motif can hydrolyze ATP at least once but are unable to catalyze DNA translocation or catalytic turnover (41). The mutant protein appears to be "frozen" in a proteolysis-resistant product conformation (E4) (B. Draper & V.B. Rao, unpublished data), unable to sense the ATP hydrolysis state and sustain ATP turnover and DNA translocation. In helicases, this coupling motif is part of a network of interactions that allow communication between the ATPase and DNA binding sites, achieving translocation in response to ATP hydrolysis.

On the other hand, the presence of a DNA translocation groove with unique binding properties, a center piece of the inchworm mechanism, has not yet been identified. A DNA binding site was found in the ATPase domain of T4 gp17 (4), and another DNA binding site must be present in the C-terminal domain in order to cut the DNA. It is unclear if either of these sites are directly involved in the translocation mechanism. The enormous force generated by the packaging motors cannot be readily explained by simple cleft closure or terminase subunit-DNA interactions. These aspects of the model might be illuminated when the structure of the full-length terminase protein is solved. Nevertheless, the inchworm model provides an experimentally testable framework to elucidate the detailed mechanism of the fascinating viral genome packaging motors.

CONCLUDING REMARKS

From the stride of an elephant to the extension of an amoeba's pseudopod, movements of cellular organisms are powered by nanoscale biological motors. Within cells, motors transport organelles and macromolecules, remodel cytoskeleton, and power progression through the cell cycle. As emphasized years ago, although DNA is the cell's store of genetic information, DNA is metabolically dynamic, as the center of transcription, recombination, replication, repair, partition, translocation, and so on (25). These processes require motor proteins, including polymerases, nucleases, helicases, and translocases. Translocases that transport DNA include ATPases that move DNA in the cells, like FtsK (5) and SpoIIIE (96). Viral DNA translocases move viral genomes into shells during virus assembly. Viral dsDNA translocation is coordinated with the processing of concatemeric DNA to produce unit-length virion chromosomes. Efficient viral assembly requires that the DNA packaging

motor be very fast and very powerful. The viral DNA packaging machine is an ancient invention that is found in all kingdoms, and consists of the terminase with a translocation ATPase and an endonuclease, and the icosahedral shell with its portal protein. How the components of the motor assemble and change configurations during DNA processing and translocation are under intense study. Although the location and nature of the translocation ATPase are largely understood, the mechanics of translocation are yet to be defined in detail. Furthermore, how the translocation complex is assembled/disassembled, how the packaging ATPase is stimulated, and how the endonuclease and translocase activities are coordinated to orchestrate DNA processing and packaging are issues about which we know little. Understanding the biochemical and structural basis for force generation, and the dynamics of DNA compaction as well as the precise measurement of viral genome are sure to elicit some surprises. Happily, remarkable progress in understanding the structure and biophysics of the packaging motor, founded on strong genetics and biochemistry, is leading to a clearer picture of the motor and its dynamics. Sufficient knowledge is available now for researchers to design experiments that can critically address the mechanism of viral DNA translocation.

SUMMARY POINTS

1. The DNA packaging machinery of the tailed dsDNA bacteriophages and the herpes viruses may have been derived from that of a common ancestor virus.
2. The DNA packaging machine utilizes energy from ATP hydrolysis to translocate DNA into a preformed empty shell. In the packaging machine, the packaging enzyme terminase docks on the special portal vertex of the icosahedral shell. The portal vertex is occupied by the dodecameric portal protein. The translocation ATPase and the concatemer-processing endonuclease reside in the large terminase subunit.
3. The packaging motor is an extraordinarily powerful biological motor, generating forces of about 60 pN. The bulk energy budget is about 2 bp translocated/ATP hydrolyzed. Translocation proceeds against a force that rises sharply as the shell is filled, resulting from extensive DNA bending and charge repulsion. The internal pressure is sufficient to power injection of much of the viral DNA during an infection.
4. Packaging models focus on terminase and/or portal protein as the mechanical center of the motor. Models suggest that conformational changes brought about by the ATP hydrolysis cycle cause domains of the terminase and/or the portal protein to translocate DNA into the shell.
5. The translocation ATPase center has a classic nucleotide binding fold and an ATPase catalytic pocket similar to that found in RecA and other ATPases. Structure and sequence alignments show the closest similarity to the ATPase domain of monomeric helicases.
6. For the first time, in recent years, researchers have been able to design and execute hypothesis-driven experiments testing the predictions of translocation models. For example, genetic and biophysical experiments indicate that the portal protein does not rotate relative to the capsid shell during translocation, challenging models that invoke portal rotation during translocation.
7. An inchworm translocation model is proposed: Series of ATP hydrolysis cycles by a pentamer of the large terminase protein are accompanied by domain movements that translocate DNA into the shell.

DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors thank Dr. Bonnie Draper for assistance in preparing **Figures 2–6**; Dr. Siyang Sun for the ATPase model in **Figure 1*a***; Drs. Fred Antson (University of York), Lindsay Black (University of Maryland), Jose Carrascosa (Centro Nacional de Biotecnologia, CSIC), Sherwood Casjens (University of Utah), Carlos Catalano (University of Washington), Bonnie Draper and Kiran Kondabagil (The Catholic University of America), Shelly Grimes and Paul Jardine (University of Minnesota), Michael Rossmann (Purdue University), Doug Smith (University of California, San Diego), and Paulo Tavares (Unité de Virologie Moléculaire et Structurale, CNRS), for critically reading the manuscript and providing thoughtful suggestions; and Drs. Jose Carrascosa, Elena Orlova (University of London), and Benes Trus (National Institutes of Health) for the portal images. M.F. thanks Andy Becker, Hisao Fujisawa, Marvin Gold, Helios Murialdo, Choon-Seok Oh, and Jean Sippy for lively discussions along the way. The research in the authors' laboratories has been funded by National Science Foundation (VBR MCB-0110574, 423528 and M.F. MCB-0717620) and National Institutes of Health (MF GM-51611). Special thanks for our present and former lab members for their contributions, and our apologies to colleagues whose work was not cited due to space limitations.

LITERATURE CITED

1. Abrahams JP, Leslie AG, Lutter R, Walker JE. 1994. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* 370:621–28
2. Agirrezabala X, Martin-Benito J, Caston JR, Miranda R, Valpuesta JM, Carrascosa JL. 2005. Maturation of phage T7 involves structural modification of both shell and inner core components. *EMBO J.* 24:3820–29
3. Ahmadian MR, Stege P, Scheffzek K, Wittinghofer A. 1997. Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nat. Struct. Biol.* 4:686–89
4. Alam TI, Rao VB. 2008. The ATPase domain of the large terminase protein, gp17, from bacteriophage T4 binds DNA: implications to the DNA packaging mechanism. *J. Mol. Biol.* 376:1272–81
5. Aussel L, Barre FX, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D. 2002. FtsK Is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108:195–205
6. Babbar BK, Gold M. 1998. ATP-reactive sites in the bacteriophage λ packaging protein terminase lie in the N-termini of its subunits, gpA and gpNu1. *Virology* 247:251–64
7. Bain D, Berton N, Ortega M, Baran J, Yang Q, Catalano C. 2001. Biophysical characterization of the DNA binding domain of gpNu1, a viral DNA packaging protein. *J. Biol. Chem.* 276:20175–81
8. Baumann RG, Black LW. 2003. Isolation and characterization of T4 bacteriophage gp17 terminase, a large subunit multimer with enhanced ATPase activity. *J. Biol. Chem.* 278:4618–27
9. Baumann RG, Mullaney J, Black LW. 2006. Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. *Mol. Microbiol.* 61:16–32
10. Becker A, Murialdo H. 1990. Bacteriophage lambda DNA: the beginning of the end. *J. Bacteriol.* 172:2819–24
11. Bhattacharyya SP, Rao VB. 1993. A novel terminase activity associated with the DNA packaging protein gp17 of bacteriophage T4. *Virology* 196:34–44
12. Bhattacharyya SP, Rao VB. 1994. Structural analysis of DNA cleaved in vivo by bacteriophage T4 terminase. *Gene* 146:67–72

13. Black L, Silverman D. 1978. Model for DNA packaging into bacteriophage T4 heads. *J. Virol.* 28:643–55
14. Black LW. 1989. DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.* 43:267–92
15. Black LW, Peng G. 2006. Mechanistic coupling of bacteriophage T4 DNA packaging to components of the replication-dependent late transcription machinery. *J. Biol. Chem.* 281:25635–43
16. Camacho AG, Gual A, Lurz R, Tavares P, Alonso JC. 2003. *Bacillus subtilis* bacteriophage SPP1 DNA packaging motor requires terminase and portal proteins. *J. Biol. Chem.* 278:23251–59
17. Casjens S, Weigele P. 2005. DNA packaging by bacteriophage P22. See Ref. 20, pp. 80–88
18. Casjens S, Wyckoff E, Hayden M, Sampson L, Eppler K, et al. 1992. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. *J. Mol. Biol.* 224:1055–74
19. Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, et al. 2005. The generalized transducing Salmonella bacteriophage ES18: complete genome sequence and DNA packaging strategy. *J. Bacteriol.* 187:1091–104
20. Catalano C, ed. 2005. *Viral Genome Packaging Machines: Genetics, Structure and Mechanism*. Georgetown, TX: Landes Biosci.
21. Catalano CE, Cue D, Feiss M. 1995. Virus DNA packaging: the strategy used by phage lambda. *Mol. Microbiol.* 16:1075–86
22. Cerritelli ME, Cheng N, Rosenberg AH, McPherson CE, Booy FP, Steven AC. 1997. Encapsidated conformation of bacteriophage T7 DNA. *Cell* 91:271–80
23. Chai S, Lurz R, Alonso JC. 1995. The small subunit of the terminase enzyme of *Bacillus subtilis* bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J. Mol. Biol.* 252:386–98
24. Chemla YR, Aathavan K, Michaelis J, Grimes S, Jardine PJ, et al. 2005. Mechanism of force generation of a viral DNA packaging motor. *Cell* 122:683–92
25. Clark AJ. 1974. Progress toward a metabolic interpretation of genetic recombination of *Escherichia coli* and bacteriophage lambda. *Genetics* 78:259–71
26. Comolli LR, Spakowitz AJ, Siegerist CE, Jardine PJ, Grimes S, et al. 2008. Three-dimensional architecture of the bacteriophage phi29 packaged genome and elucidation of its packaging process. *Virology* 371:267–77
27. Cordin O, Banroques J, Tanner NK, Linder P. 2006. The DEAD-box protein family of RNA helicases. *Gene* 367:17–37
28. Coren JS, Pierce JC, Sternberg N. 1995. Headful packaging revisited: the packaging of more than one DNA molecule into a bacteriophage P1 head. *J. Mol. Biol.* 249:176–84
29. Cue D, Feiss M. 1992. Genetic analysis of *cosB*, the binding site for terminase, the DNA packaging enzyme of bacteriophage lambda. *J. Mol. Biol.* 228:58–71
30. Cue D, Feiss M. 1992. Genetic analysis of mutations affecting terminase, the bacteriophage lambda DNA packaging enzyme, that suppress mutations in *cosB*, the terminase binding site. *J. Mol. Biol.* 228:72–87
31. Cue D, Feiss M. 1993. The role of *cosB*, the binding site for terminase, the DNA packaging enzyme of bacteriophage lambda, in the nicking reaction. *J. Mol. Biol.* 234:594–609
32. Cue D, Feiss M. 1993. A site required for termination of packaging of the phage lambda chromosome. *Proc. Natl. Acad. Sci. USA* 90:9290–94
33. Cue D, Feiss M. 1997. Genetic evidence that recognition of *cosQ*, the signal for termination of phage DNA packaging, depends on the extent of head filling. *Genetics* 147:7–17
34. Cue D, Feiss M. 1998. Termination of packaging of the bacteriophage lambda chromosome: *cosQ* is required for nicking the bottom strand of *cosN*. *J. Mol. Biol.* 280:11–29
35. Cue D, Feiss M. 2001. Bacteriophage λ DNA packaging: DNA site requirements for termination and processivity. *J. Mol. Biol.* 311:233–40
36. Cuervo A, Vaney MC, Antson AA, Tavares P, Oliveira L. 2007. Structural rearrangements between portal protein subunits are essential for viral DNA translocation. *J. Biol. Chem.* 282:18907–13
37. Davidson A, Gold M. 1992. Mutations abolishing the endonuclease activity of bacteriophage λ terminase lie in two distinct regions of the *A* gene, one of which may encode a leucine zipper DNA binding domain. *Virology* 161:305–15

38. de Beer T, Meyer J, Ortega M, Yang Q, Maes L, et al. 2002. Insights into specific DNA recognition during assembly of a viral genome packaging machine; structure and genetics of the DNA binding domain of gpNu1. *Mol. Cell* 9:981–91
39. Dhar A, Feiss M. 2005. Bacteriophage lambda terminase: alterations of the high-affinity ATPase affect viral DNA packaging. *J. Mol. Biol.* 347:71–80
40. Donate LE, Herranz L, Secilla JP, Carazo JM, Fujisawa H, Carrascosa JL. 1988. Bacteriophage T3 connector: three-dimensional structure and comparison with other viral head-tail connecting regions. *J. Mol. Biol.* 201:91–100
41. 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.* 369:79–94
42. Driedonks RA, Caldentey J. 1983. Gene 20 product of bacteriophage T4. II. Its structural organization in prehead and bacteriophage. *J. Mol. Biol.* 166:341–60
43. Dube P, Tavares P, Lurz R, van Heel M. 1993. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. *EMBO J.* 12:1303–10
44. Duffy C, Feiss M. 2002. The large subunit of bacteriophage lambda's terminase plays a role in DNA translocation and packaging termination. *J. Mol. Biol.* 316:547–61
45. Earnshaw WC, Casjens SR. 1980. DNA packaging by the double-stranded DNA bacteriophages. *Cell* 21:319–31
46. Feiss M. 1986. Terminase and the recognition, cutting and packaging of λ chromosomes. *Trends Genet.* 2:100–4
47. Feiss M, Sippy J, Miller G. 1985. Processive action of terminase during sequential packaging of bacteriophage lambda chromosomes. *J. Mol. Biol.* 186:759–71
48. Feiss M, Widner W. 1982. Bacteriophage lambda DNA packaging: scanning for the terminal cohesive end site during packaging. *Proc. Nat. Acad. Sci. USA* 79:3498–502
49. Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG. 2004. Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* 101:6003–8
50. Frackman S, Siegele DA, Feiss M. 1985. The terminase of bacteriophage lambda. Functional domains for *cosB* binding and multimer assembly. *J. Mol. Biol.* 183:225–38
51. Fujisawa H, Hamada K, Shibata H, Minagawa T. 1987. On the molecular mechanism of DNA translocation during in vitro packaging of bacteriophage T3 DNA. *Virology* 161:228–33
52. Fujisawa H, Morita M. 1997. Phage DNA packaging. *Genes Cells* 2:537–45
53. Fuller DN, Raymer DM, Kottadiel VI, Rao VB, Smith DE. 2007. Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proc. Natl. Acad. Sci. USA* 104:16868–73
54. Fuller DN, Raymer DM, Rickgauer JP, Robertson RM, Catalano CE, et al. 2007. Measurements of single DNA molecule packaging dynamics in bacteriophage lambda reveal high forces, high motor processivity, and capsid transformations. *J. Mol. Biol.* 373:1113–22
55. 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* 104:11245–50
56. Goetzinger K, Rao V. 2003. Defining the ATPase center of bacteriophage T4 DNA packaging machine: requirement for a catalytic glutamate residue in the large terminase protein gp17. *J. Mol. Biol.* 331:139–54
57. Gold M, Becker A. 1983. The bacteriophage terminase: partial purification and preliminary characterization of properties. *J. Biol. Chem.* 258:14619–25
58. Golz S, Kemper B. 1999. Association of holliday-structure resolving endonuclease VII with gp20 from the packaging machine of phage T4. *J. Mol. Biol.* 285:1131–44
59. Gorbalenya A, Koonin E. 1993. Helicases: amino acid sequence comparison and structure-function relationships. *Curr. Opin. Struct. Biol.* 3:419–29
60. Gual A, Camacho AG, Alonso JC. 2000. Functional analysis of the terminase large subunit, G2P, of *Bacillus subtilis* bacteriophage SPP1. *J. Biol. Chem.* 275:35311–19

61. Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, et al. 2002. Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle. *J. Mol. Biol.* 315:663–76
62. Guo P, Peterson C, Anderson D. 1987. Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage ϕ29. *J. Mol. Biol.* 197:229–36
63. Guo PX, Erickson S, Anderson D. 1987. A small viral RNA is required for in vitro packaging of bacteriophage phi 29 DNA. *Science* 236:690–94
64. Hamada K, Fujisawa H, Minagawa T. 1986. A defined in vitro system for packaging of bacteriophage T3 DNA. *Virology* 151:119–23
65. Hang J, Catalano C, Feiss M. 2001. The functional asymmetry of *cosN*, the nicking site for bacteriophage λ DNA packaging, is dependent on the terminase binding site, *cosB*. *Biochemistry* 40:13370–77
66. Hang Q, Tack B, Feiss M. 2000. An ATPase center of bacteriophage λ terminase involved in postcleavage stages of DNA packaging: identification of ATP-interactive amino acids. *J. Mol. Biol.* 302:777–95
67. Hendrix RW. 1978. Symmetry mismatch and DNA packaging in large bacteriophages. *Proc. Natl Acad. Sci. USA* 75:4779–83
68. Higgins RR, Becker A. 1994. Chromosome end formation in phage lambda, catalyzed by terminase, is controlled by two DNA elements of *cos*, *cosN* and R3, and by ATP. *EMBO J.* 13:6152–61
69. Higgins RR, Becker A. 1994. The lambda terminase enzyme measures the point of its endonucleolytic attack 47 +/– 2 bp away from its site of specific DNA binding, the R site. *EMBO J.* 13:6162–71
70. Higgins RR, Becker A. 1995. Interaction of terminase, the DNA packaging enzyme of phage lambda, with its *cos* DNA substrate. *J. Mol. Biol.* 252:31–46
71. Higgins RR, Lucko HJ, Becker A. 1988. Mechanism of *cos* DNA cleavage by bacteriophage lambda terminase: multiple roles of ATP. *Cell* 54:765–75
72. Hohn B. 1983. DNA sequences necessary for packaging of bacteriophage λ DNA. *Proc. Natl Acad. Sci. USA* 80:7456–60
73. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77:61–68
74. Hugel T, Michaelis J, Hetherington CL, Jardine PJ, Grimes S, et al. 2007. Experimental test of connector rotation during DNA packaging into bacteriophage phi29 capsids. *PLoS Biol.* 5:e59
75. Hwang Y, Catalano CE, Feiss M. 1996. Kinetic and mutational dissection of the two ATPase activities of terminase, the DNA packaging enzyme of bacteriophage lambda. *Biochemistry* 35:2796–803
76. Hwang Y, Feiss M. 1995. A defined system for in vitro lambda DNA packaging. *Virology* 211:367–76
77. Hwang Y, Feiss M. 1996. Mutations affecting the high affinity ATPase center of gpA, the large subunit of bacteriophage lambda terminase, inactivate the endonuclease activity of terminase. *J. Mol. Biol.* 261:524–35
78. Hwang Y, Feiss M. 1997. Mutations affecting lysine-35 of gpNu1, the small subunit of bacteriophage λ terminase, alter the strength and specificity of holoterminase interactions with DNA. *Virology* 231:218–30
79. Hwang Y, Feiss M. 2000. The endonuclease and helicase activities of bacteriophage λ terminase: changing nearby residue 515 restores activity to the gpA K497D mutant enzyme. *Virology* 277:204–14
80. Ibarra B, Caston JR, Llorca O, Valle M, Valpuesta JM, Carrascosa JL. 2000. Topology of the components of the DNA packaging machinery in the phage phi29 prohead. *J. Mol. Biol.* 298:807–15
81. Isidro A, Henriques AO, Tavares P. 2004. The portal protein plays essential roles at different steps of the SPP1 DNA packaging process. *Virology* 322:253–63
82. Johnson JE, Chiu W. 2007. DNA packaging and delivery machines in tailed bacteriophages. *Curr. Opin. Struct. Biol.* 17:237–43
83. Kanamaru S, Kondabagil K, Rossmann MG, Rao VB. 2004. The functional domains of bacteriophage t4 terminase. *J. Biol. Chem.* 279:40795–801
84. Kobayashi I, Murialdo H, Crasemann JM, Stahl MM, Stahl FW. 1982. Orientation of cohesive end site *cos* determines the active orientation of chi sequence in stimulating recA. recBC-mediated recombination in phage lambda lytic infections. *Proc. Natl. Acad. Sci. USA* 79:5981–85

85. Kondabagil KR, Rao VB. 2006. A critical coiled coil motif in the small terminase, gp16, from bacteriophage T4: insights into DNA packaging initiation and assembly of packaging motor. *J. Mol. Biol.* 358:67–82
86. Kondabagil KR, Zhang Z, Rao VB. 2006. The DNA translocating ATPase of bacteriophage T4 packaging motor. *J. Mol. Biol.* 363:786–99
87. Kuebler D, Rao VB. 1998. Functional analysis of the DNA-packaging/terminase protein gp17 from bacteriophage T4. *J. Mol. Biol.* 281:803–14
88. Kypr J, Mrazek J. 1986. Lambda phage protein Nu1 contains the conserved DNA binding fold of repressors. *J. Mol. Biol.* 91:139–40
89. Lander GC, Tang L, Casjens SR, Gilcrease EB, Prevelige P, et al. 2006. The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312:1791–95
90. Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova EV, et al. 2007. Structural framework for DNA translocation via the viral portal protein. *EMBO J.* 26:1984–94
91. Lee TJ, Guo P. 2006. Interaction of gp16 with pRNA and DNA for genome packaging by the motor of bacterial virus phi29. *J. Mol. Biol.* 356:589–99
92. Leffers G, Rao V. 2000. Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. *J. Biol. Chem.* 275:37127–36
93. Leffers G, Rao VB. 1996. A discontinuous headful packaging model for packaging less than headful length DNA molecules by bacteriophage T4. *J. Mol. Biol.* 258:839–50
94. Lin H, Rao VB, Black LW. 1999. Analysis of capsid portal protein and terminase functional domains: interaction sites required for DNA packaging in bacteriophage T4. *J. Mol. Biol.* 289:249–60
95. Lin H, Simon MN, Black LW. 1997. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. *J. Biol. Chem.* 272:3495–501
96. Liu NJ, Dutton RJ, Pogliano K. 2006. Evidence that the SpoIIIE DNA translocase participates in membrane fusion during cytokinesis and engulfment. *Mol. Microbiol.* 59:1097–113
97. Luftig RB, Ganz C. 1972. Bacteriophage T4 head morphogenesis. IV. Comparison of gene 16-, 17-, and 49-defective head structures. *J. Virol.* 10:545–54
98. MacKinlay AG, Kaiser AD. 1969. DNA replication in head mutants of bacteriophage lambda. *J. Mol. Biol.* 39:679–83
99. Maluf N, Yang Q, Catalano C. 2005. Self-association properties of the bacteriphage lambda terminase holoenzyme: implications for the DNA packaging motor. *J. Mol. Biol.* 347:523–42
100. Maluf NK, Feiss M. 2006. Virus DNA translocation: progress towards a first ascent of Mount Pretty Difficult. *Mol. Microbiol.* 61:1–4
101. Maluf NK, Gaussier H, Bogner E, Feiss M, Catalano CE. 2006. Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine. *Biochemistry* 45:15259–68
102. Malys N, Chang DY, Baumann RG, Xie D, Black LW. 2002. A bipartite bacteriophage T4 SOC and HOC randomized peptide display library: detection and analysis of phage T4 terminase (gp17) and late sigma factor (gp55) interaction. *J. Mol. Biol.* 319:289–304
103. Mancini EJ, Kainov DE, Grimes JM, Tuma R, Bamford DH, Stuart DI. 2004. Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* 118:743–55
104. Mendelson I, Gottesman M, Oppenheim AB. 1991. HU and integration host factor function as auxiliary proteins in cleavage of phage lambda cohesive ends by terminase. *J. Bacteriol.* 173:1670–76
105. Mitchell M, Matsuzaki S, Imai S, Rao V. 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.* 30:4009–21
106. Mitchell MS, Rao VB. 2004. Novel and deviant Walker A ATP-binding motifs in bacteriophage large terminase-DNA packaging proteins. *Virology* 321:217–21
107. Mitchell MS, Rao VB. 2006. Functional analysis of the bacteriophage T4 DNA-packaging ATPase motor. *J. Biol. Chem.* 281:518–27
108. Moll WD, Guo P. 2005. Translocation of nicked but not gapped DNA by the packaging motor of bacteriophage phi29. *J. Mol. Biol.* 351:100–7

109. Morita M, Tasaka M, Fujisawa H. 1993. DNA packaging ATPase of bacteriophage T3. *Virology* 193:748–52
110. Morita M, Tasaka M, Fujisawa H. 1995. Analysis of the fine structure of the prohead binding domain of the packaging protein of bacteriophage T3 using a hexapeptide, an analog of a prohead binding site. *Virology* 211:516–24
111. Morita M, Tasaka M, Fujisawa H. 1995. Structural and functional domains of the large subunit of the bacteriophage T3 DNA packaging enzyme: importance of the C-terminal region in prohead binding. *J. Mol. Biol.* 245:635–44
112. Nemecek D, Gilcrease EB, Kang S, Prevelige PE Jr, Casjens S, Thomas GJ Jr. 2007. Subunit conformations and assembly states of a DNA-translocating motor: the terminase of bacteriophage P22. *J. Mol. Biol.* 374:817–36
113. Oliveira L, Alonso JC, Tavares P. 2005. A defined in vitro system for DNA packaging by the bacteriophage SPP1: insights into the headful packaging mechanism. *J. Mol. Biol.* 353:529–39
114. Oliveira L, Henriques AO, Tavares P. 2006. Modulation of the viral ATPase activity by the portal protein correlates with DNA packaging efficiency. *J. Biol. Chem.* 281:21914–23
115. Orlova EV, Dube P, Beckmann E, Zemlin F, Lurz R, et al. 1999. Structure of the 13-fold symmetric portal protein of bacteriophage SPP1. *Nat. Struct. Biol.* 6:842–46
116. Ortega ME, Gaussier H, Catalano CE. 2007. The DNA maturation domain of gpA, the DNA packaging motor protein of bacteriophage lambda, contains an ATPase site associated with endonuclease activity. *J. Mol. Biol.* 373:851–65
117. Oster G, Wang H. 2003. Rotary protein motors. *Trends Cell Biol.* 13:114–21
118. Pearson RK, Fox MS. 1988. Effects of DNA heterologies on bacteriophage lambda packaging. *Genetics* 118:5–12
119. Ponchon L, Boulanger P, Labesse G, Letellier L. 2006. The endonuclease domain of bacteriophage terminases belongs to the resolvase/integrase/ribonuclease H superfamily: a bioinformatics analysis validated by a functional study on bacteriophage T5. *J. Biol. Chem.* 281:5829–36
120. Przech AJ, Yu D, Weller SK. 2003. Point mutations in exon I of the herpes simplex virus putative terminase subunit, UL15, indicate that the most conserved residues are essential for cleavage and packaging. *J. Virol.* 77:9613–21
121. Rao V, Black L. 2005. DNA packaging in bacteriophage T4. See Ref. 20, pp. 40–58
122. Rao V, Mitchell M. 2001. The N-terminal ATPase site in the large terminase proteingp17 is critically required for DNA packaging in bacteriophage T4. *J. Mol. Biol.* 314:411–21
123. Rao VB, Black LW. 1985. DNA packaging of bacteriophage T4 proheads in vitro. Evidence that prohead expansion is not coupled to DNA packaging. *J. Mol. Biol.* 185:565–78
124. Rao VB, Black LW. 1988. Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4. Construction of a defined in vitro DNA packaging system using purified terminase proteins. *J. Mol. Biol.* 200:475–88
125. Rentas FJ, Rao VB. 2003. Defining the bacteriophage T4 DNA packaging machine: evidence for a C-terminal DNA cleavage domain in the large terminase/packaging protein gp17. *J. Mol. Biol.* 334:37–52
126. 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.* 94:159–67
127. Rodriguez I, Lazaro JM, Salas M, De Vega M. 2004. Phi29 DNA polymerase-terminal protein interaction. Involvement of residues specifically conserved among protein-primed DNA polymerases. *J. Mol. Biol.* 337:829–41
128. Rossmann MG, Moras D, Olsen KW. 1974. Chemical and biological evolution of nucleotide-binding protein. *Nature* 250:194–99
129. Rubinchik S, Parris W, Gold M. 1994. The in vitro ATPases of bacteriophage lambda terminase and its large subunit, gene product A. The relationship with their DNA helicase and packaging activities. *J. Biol. Chem.* 269:13586–93
130. Rubinchik S, Parris W, Gold M. 1994. The in vitro endonuclease activity of gene product A, the large subunit of the bacteriophage lambda terminase, and its relationship to the endonuclease activity of the holoenzyme. *J. Biol. Chem.* 269:13575–85

130a. Sabanayagam CR, Oram M, Lakowicz JR, Black LW. 2007. Viral DNA packaging studied by fluorescence correlation spectroscopy. *Biophys. J.* 93:L17–19
131. Sathaliyawala T, Rao M, Maclean DM, Birx DL, Alving CR, Rao VB. 2006. Assembly of human immunodeficiency virus (HIV) antigens on bacteriophage T4: a novel in vitro approach to construct multicomponent HIV vaccines. *J. Virol.* 80:7688–98
132. Shinder G, Gold M. 1988. The Nu1 subunit of bacteriophage lambda terminase binds to specific sites in *cos* DNA. *J. Virology* 62:387–92
133. Shivachandra SB, Li Q, Peachman KK, Matyas GR, Leppla SH, et al. 2007. Multicomponent anthrax toxin display and delivery using bacteriophage T4. *Vaccine* 25:1225–35
134. 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.* 26:527–37
135. Simpson A, Tao Y, Leiman P, Badasso M, He Y, et al. 2000. Structure of the bacteriophage phi29 DNA packaging motor. *Nature* 408:745–50
136. Smith D, Tans S, Smith S, Grimes S, Anderson D, Bustamante C. 2001. The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature* 413:748–52
137. Soultanas P, Dillingham MS, Velankar SS, Wigley DB. 1999. DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J. Mol. Biol.* 290:137–48
138. Soultanas P, Wigley DB. 2000. DNA helicases: 'inching forward'. *Curr. Opin. Struct. Biol.* 10:124–28
139. Soultanas P, Wigley DB. 2002. Site-directed mutagenesis reveals roles for conserved amino acid residues in the hexameric DNA helicase DnaB from *Bacillus stearothermophilus*. *Nucleic Acids Res.* 30:4051–60
140. Story RM, Steitz TA. 1992. Structure of the recA protein-ADP complex. *Nature* 355:374–76
141. Streisinger G, Emrich J, Stahl MM. 1967. Chromosome structure in Phage T4, III. Terminal redundancy and length determination. *Proc. Natl. Acad. Sci. USA* 57:292–95
142. 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* 25:943–49
143. Szpirer J, Brachet P. 1970. Physiological relationship between the temperate phages lambda and phi80. *Mol. Gen. Genet.* 108:78–92
144. Tavares P, Santos MA, Lurz R, Morelli G, de Lencastre H, Trautner TA. 1992. Identification of a gene in *Bacillus subtilis* bacteriophage SPP1 determining the amount of packaged DNA. *J. Mol. Biol.* 225:81–92
145. Trottier M, Mat-Arip Y, Zhang C, Chen C, Sheng S, et al. 2000. Probing the structure of monomers and dimers of the bacterial virus phi29 hexamer RNA complex by chemical modification. *RNA* 6:1257–66
146. Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, Steven AC. 2004. Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J. Virol.* 78:12668–71
147. Valpuesta JM, Carrascosa JL. 1994. Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Q. Rev. Biophys.* 27:107–55
148. Velankar SS, Soultanas 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* 97:75–84
149. Wake R, Kaiser A, Inman R. 1972. Isolation and structure of phage lambda head-mutant DNA. *J. Mol. Biol.* 64:519–40
150. Walker JE, Saraste M, Runswick MJ, Gay NJ. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 8:945–51
151. Wang J, Kaiser A. 1973. Evidence that the cohesive ends of mature λ DNA are generated by the gene *A* product. *Nat. New Biol.* 241:16–17
152. White JH, Richardson CC. 1987. Gene 18 protein of bacteriophage T7. Overproduction, purification, and characterization. *J. Biol. Chem.* 262:8845–50
153. Wieczorek D, Didion L, Feiss M. 2002. Alterations of the portal protein of bacteriophage λ suppress mutations in *cosQ*, the site required for termination of DNA packaging. *Genetics* 161:21–31
154. Wieczorek D, Feiss M. 2001. Defining *cosQ*, the site required for termination of bacteriophage lambda DNA packaging. *Genetics* 158:495–506
155. Wieczorek DJ, Feiss M. 2003. Genetics of *cosQ*, the DNA-packaging termination site of phage lambda: local suppressors and methylation effects. *Genetics* 165:11–21

156. Xiang Y, Morais MC, Battisti AJ, Grimes S, Jardine PJ, et al. 2006. Structural changes of bacteriophage phi29 upon DNA packaging and release. *EMBO J.* 25:5229–39
157. Xin W, Cai Z-H, Feiss M. 1993. Function of IHF in λ DNA packaging. II. Effects of mutations altering the IHF binding site and the intrinsic bend in *cosB* on λ development. *J. Mol. Biol.* 230:505–15
158. Xin W, Feiss M. 1993. Function of IHF in λ DNA packaging. I. Identification of the strong binding site for integration host factor and the locus for intrinsic bending in *cosB*. *J. Mol. Biol.* 230:492–504
159. Yang Q, Berton N, Manning M, Catalano C. 1999. Domain structure of gpNu1, a phage lambda DNA packaging protein. *Biochemistry* 38:14238–447
160. Yang Q, Catalano C. 2003. Biochemical characterization of bacteriophage lambda genome packaging in vitro. *Virology* 305:276–87
161. Yang Q, Hanagan A, Catalano CE. 1997. Assembly of a nucleoprotein complex required for DNA packaging by bacteriophage lambda. *Biochemistry* 36:2744–52
162. Yeo A, Feiss M. 1995. Mutational analysis of the prohead binding domain of the large subunit of terminase, the bacteriophage lambda DNA packaging enzyme. *J. Mol. Biol.* 245:126–40
163. Yeo A, Feiss M. 1995. Specific interaction of terminase, the DNA packaging enzyme of bacteriophage lambda, with the portal protein of the prohead. *J. Mol. Biol.* 245:141–50
164. Zachary A, Black LW. 1981. DNA ligase is required for encapsidation of bacteriophage T4 DNA. *J. Mol. Biol.* 149:641–58
165. Zhang F, Lemieux S, Wu X, St-Arnaud D, McMurray CT, et al. 1998. Function of hexameric RNA in packaging of bacteriophage phi 29 DNA in vitro. *Mol. Cell* 2:141–47
166. 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* 29:376–83

Annual Review of Genetics

Volume 42, 2008

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at http://genet.annualreviews.org/errata.shtml