MOLECULAR BIOLOGY INTELLIGENCE UNIT

Carlos Enrique Catalano

Viral Genome Packaging Machines:
Genetics, Structure, and Mechanism

KLUWER ACADEMIC / PLENUM PUBLISHERS

LANDES BIOSCIENCE

MOLECULAR BIOLOGY
INTELLIGENCE
UNIT

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism

Carlos Enrique Catalano, Pharm.D., Ph.D.
Department of Pharmaceutical Chemistry
The University of Colorado School of Pharmacy
Denver, Colorado, U.S.A.

LANDES BIOSCIENCE / EUREKAH.COM
GEORGETOWN, TEXAS
U.S.A.

KLUWER ACADEMIC / PLENUM PUBLISHERS
NEW YORK, NEW YORK
U.S.A.

VIRAL GENOME PACKAGING MACHINES: GENETICS, STRUCTURE, AND MECHANISM

Molecular Biology Intelligence Unit

Landes Bioscience / Eurekah.com
Kluwer Academic / Plenum Publishers

Printed in the U.S.A.

Kluwer Academic / Plenum Publishers, 233 Spring Street, New York, New York, U.S.A. 10013
http://www.wkap.nl/

Please address all inquiries to the Publishers:
Landes Bioscience / Eurekah.com, 810 South Church Street
Georgetown, Texas, U.S.A. 78626
Phone: 512/ 863 7762; FAX: 512/ 863 0081
www.Eurekah.com
www.landesbioscience.com

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano, Landes / Kluwer dual imprint / Landes series: Molecular Biology Intelligence Unit

ISBN: 0-306-48227-4

Library of Congress Cataloging-in-Publication Data

Viral genome packaging machines : genetics, structure, and mechanism / [edited by] Carlos Enrique Catalano.
p. ; cm. -- (Molecular biology intelligence unit)
Includes bibliographical references and index.
ISBN 0-306-48227-4
1. Viruses--Reproduction. 2. Viral genetics. I. Catalano, Carlos Enrique. II. Title. III. Series: Molecular biology intelligence unit (Unnumbered) [DNLM: 1. Virus Assembly. 2. Bacteriophages. 3. DNA Packaging. 4. Genome, Viral. QW 160 V8124 2005]
QR470.V55 2005
571.8'292--dc22

2005005134

This book is dedicated to Maria Amparo,
who taught me that the difficult path is likely the most rewarding.

CONTENTS

EDITOR

Carlos Enrique Catalano
Department of Pharmaceutical Chemistry
The University of Colorado School of Pharmacy
Denver, Colorado, U.S.A.
Chapters 1, 2

CONTRIBUTORS

Dwight Anderson
Department of Oral Science
and Department of Microbiology
University of Minnesota
Minneapolis, Minnesota, U.S.A.
Chapter 7

Joel D. Baines
Department of Microbiology
and Immunology
Cornell University
Ithaca, New York, U.S.A.
Chapter 9

Dennis H. Bamford
Department of Biosciences and Institute
of Biotechnology
University of Helsinki
Helsinki, Finland
Chapter 8

Lindsay W. Black
Department of Biochemistry
and Molecular Biology
University of Maryland School
of Medicine
Baltimore, Maryland, U.S.A.
Chapter 3

Sherwood Casjens
Department of Pathology
University of Utah Medical School
Salt Lake City, Utah, U.S.A.
Chapter 5

Anja Dröge
Department of Neuroproteomics
Max-Delbrück-Centrum für Molekulare
Medizin (MDC)
Berlin-Buch, Germany
Chapter 6

Michael Feiss
Department of Microbiology
University of Iowa
Iowa City, Iowa, U.S.A.
Chapter 2

Shelley Grimes
Department of Oral Science
University of Minnesota
Minneapolis, Minnesota, U.S.A.
Chapter 7

Markus J. Pirttimaa
Department of Biosciences and Institute
of Biotechnology
University of Helsinki
Helsinki, Finland
Chapter 8

Minna M. Poranen
Department of Biosciences and Institute
of Biotechnology
University of Helsinki
Helsinki, Finland
Chapter 8

Venigalla B. Rao
Department of Biology
The Catholic University of America
Washington, District of Columbia,
U.S.A.
Chapter 3

Philip Serwer
Department of Biochemistry
The University of Texas Health Science
Center at San Antonio
San Antonio, Texas, U.S.A.
Chapter 4

Paulo Tavares
Unité de Virologie Moléculaire
et Structurale
Gif-sur-Yvette, France
Chapter 6

Peter Weigele
Biology Department
Massachusetts Institute of Technology
Cambridge, Massachusetts, U.S.A.
Chapter 5

Sandra K. Weller
Department of Microbiology
University of Connecticut Health Center
Farmington, Connecticut, U.S.A.
Chapter 9

CHAPTER 1

Viral Genome Packaging Machines:

An Overview

Carlos Enrique Catalano

A virus particle is a marvel of nature, designed to replicate with a minimal genetic repertoire. All viruses are, of course, obligate intracellular parasites and require an appropriate host in which to develop and multiply. They have evolved a variety of strategies to infect a host cell and to usurp the cellular machinery to manufacture the components required to construct a virion. These precursors are then assembled into an infectious virus particle within the cell.

Virus assembly is a complex process that requires the temporal and coordinated activities of numerous proteins of both viral and host origin. Assembly pathways vary among the virus types, but common features are observed within certain groups. For instance, double-stranded DNA (dsDNA) viruses include the poxviruses, adenovirus, the herpesvirus groups, and many of the bacteriophages. Despite their obvious differences, common development pathways exist among these viruses, as follows. Infection of the host cell ultimately leads to the synthesis of capsid proteins that are assembled into "procapsid" structures. Concurrently, viral DNA is replicated producing numerous copies of the viral genome. The assembly of an infectious virus requires that a single genome be "packaged" into the restricted confines of an empty procapsid. This extraordinary process represents the intersection of the capsid and DNA synthetic pathways, and is an essential step in virus assembly.

This book focuses on the process of viral genome packaging. Chapters 2 through 6 describe our current understanding of genome packaging in bacteriophages λ (Catalano and Feiss), T4 (Black and Rao), T7 (Serwer), P22 (Casjens and Weigele) and SPP1 (Dröge and Tavares). These chapters reveal common mechanisms for DNA packaging among the phage and establish the basic genetic and biochemical rules for the process. In these cases, viral DNA is replicated as linear concatemers of the viral genome. The assembly of an infectious virus requires that individual genomes be cut from the concatemer and concurrently packaged into an empty procapsid, much as one might cut an individual doll from a paper chain and package it into a box. Terminase enzymes are common to these viruses and play a direct role in genome packaging. All of the characterized terminase enzymes are composed of small (18-21 kDa) and large (49-72 kDa) subunits, and the functional holoenzyme is an oligomer of these subunits.

Packaging of viral DNA begins with specific binding of the terminase proteins to a packaging initiation site on the viral DNA concatemer (*pac*, Fig. 1). Specific recognition of viral DNA is mediated by the small terminase subunits. Once assembled, a nuclease activity centered in the large subunit cuts the duplex, thus forming a mature genome end in preparation for DNA packaging. This nucleoprotein complex then binds to a doughnut-shaped portal

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

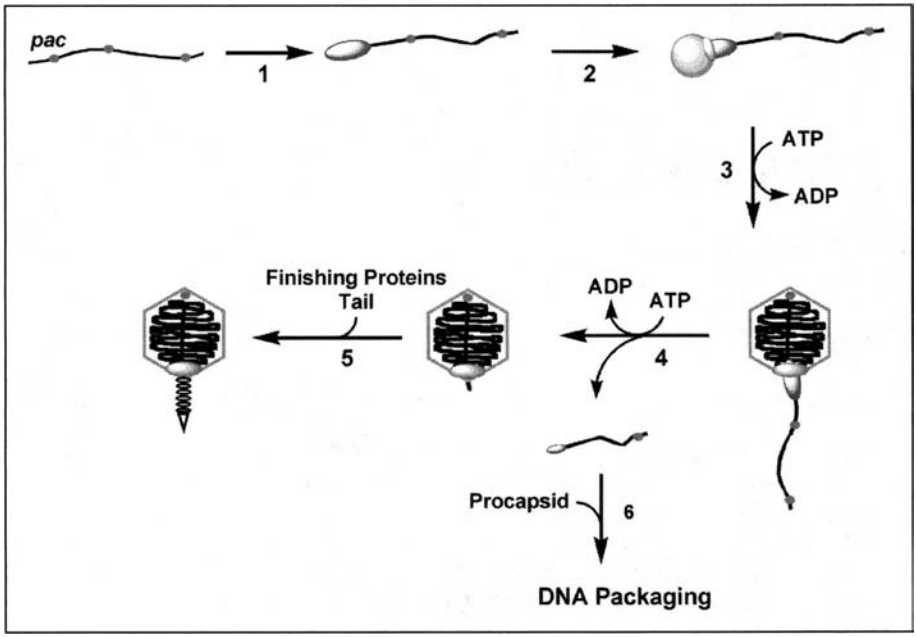


Figure 1. Generalized scheme for genome packaging in double-stranded DNA viruses. A viral genome concatemer is shown as a thick black line with repeated *pac* sites depicted as red dots. The terminase enzyme (blue oval) specifically binds to a *pac* site and cuts the duplex to generate a mature genome end in preparation for packaging (1). The terminase proteins bind to portal proteins located at a unique vertex in a procapsid (cyan sphere with portal shown in purple) (2). This interaction forms the packaging motor that translocates DNA into the capsid, fueled by ATP hydrolysis; DNA packaging triggers capsid expansion (3). Once the entire genome has been inserted into the capsid, terminase again cuts the duplex to complete the packaging process (4). Addition of "finishing" proteins, and a tail in the case of bacteriophages, complete the assembly of an infectious virus (5). The terminase•concatemer complex binds another procapsid to initiate a second round of packaging (6). A color version of this figure is available online at http://www.Eurekah.com.

complex that resides at a unique vertex in the procapsid shell. The portal forms a hole through which DNA enters the capsid during packaging; while details of the interaction remain obscure, it is likely that a combination of the terminase proteins and the portal proteins make up a DNA packaging motor that actively translocates viral DNA into the interior of the capsid. Packaging activity resides in the large terminase subunits and is fueled by ATP hydrolysis.

In many viruses, DNA packaging triggers procapsid expansion. This is a remarkable process where the roughly spherical procapsid undergoes an expansion step that increases the inner capsid volume and yields a more angularized capsid structure. Expansion requires significant reorganization of the capsid proteins and is typically followed by the addition of "stabilization" proteins to the capsid surface, or physical cross-linking of the capsid proteins to provide enhanced structural integrity.

The translocating motor ultimately fills the capsid with DNA, packaging a single viral genome condensed to near liquid crystalline density. This represents an energetically demanding process, and the DNA packaging motor is among the most powerful biological motors thus far characterized. Upon packaging a complete genome, terminase again cuts the duplex, which separates the DNA-filled capsid from the terminase•concatemer complex. The mechanism regulating this terminal cleavage event is unclear, but there is a universal "head-full"

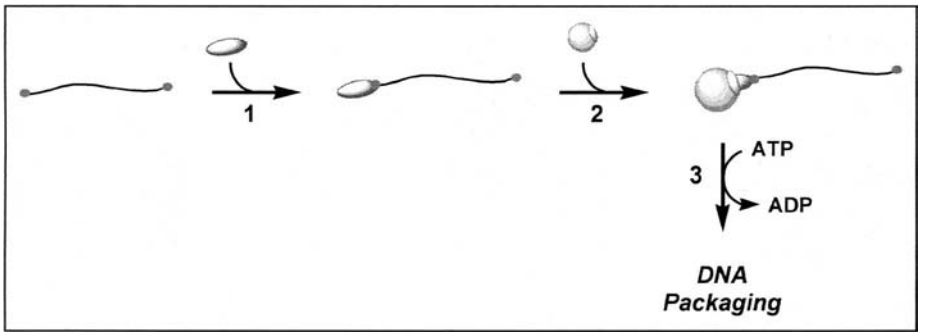


Figure 2. Genome packaging in bacteriophage ϕ29. Viral DNA is replicated as a monomer using a protein primed mechanism. The monomeric genome is shown as a thick black line, with terminal proteins covalently attached to each end (red dots). The terminal protein associates with the "packaging ATPase", shown as a blue oval (1). This complex binds to the portal proteins to complete the packaging motor (2). The motor translocates DNA into the capsid, powered by the hydrolysis of ATP (3). A unique feature of ϕ29 is the requirement for portal-associated RNA molecules in the packaging motor (not shown). A color version of this figure is available online at http://www.Eurekah.com.

component. That is, the nuclease activity of the terminase large subunit is in some way activated once the capsid is filled to capacity. Addition of "finishing" proteins and a viral tail completes the assembly of an infectious virus. The terminase•concatemer complex (described above) binds a second procapsid to initiate a second round of DNA packaging (Fig. 1). Thus, DNA packaging is a processive process, with multiple genomes in the concatemer packaged per DNA binding event.

Not unexpectedly, the details of DNA packaging vary with each virus; however, the similarities are striking and indicate a common strategy. Indeed, this strategy may be universal and traverse the prokaryotic-eukaryotic boundary. The herpesviruses are large eukaryotic dsDNA viruses that encompass many human and animal pathogens. Despite the obvious differences between herpesviruses and bacteriophages, common developmental pathways exist, especially as it relates to DNA packaging mechanisms. Bains and Weller describe our current understanding of herpesvirus genome packaging in chapter nine of this book.

An interesting variation on this general packaging strategy is found in bacteriophage ϕ29, which is described by Anderson and Grimes in Chapter 7 of this book. Unlike the phage described above, ϕ29 replicates its genome as a monomer. This is accomplished through a protein-primed DNA replication mechanism, which yields individual genomes with a terminal protein covalently attached to the 5' ends of the duplex (Fig. 2). Genome packaging requires a "packaging ATPase" protein that associates with the terminal protein. This enzyme also binds to the ϕ29 portal complex to complete the packaging motor. Despite the apparent difference, this mechanism is quite similar to the general packaging strategy, as follows. The small terminase subunits described above provide specific recognition of viral DNA, while the large subunits possess the ATP-powered packaging activity. In the case of ϕ29, the terminal protein (30 kDa) is strictly required for genome packaging and may be viewed as a small terminase subunit. Further, the packaging ATPase of ϕ29 is analogous to the large subunits found in the conventional terminase holoenzymes of λ, T4, etc. The general strategy for genome packaging is thus retained in ϕ29 despite the apparent divergence from the conventional model. Here too, packaging strategies may traverse the prokaryotic-eukaryotic boundary, as packaging in adenovirus may be analogous to that of ϕ29.

The majority of this book examines genome packaging in the dsDNA viruses. In reading these chapters, it becomes apparent that the basic mechanisms of energy transduction linked to DNA translocation are quite similar. This conceptual model is not limited to DNA packaging machines, however. The mechanism of genome packaging in ϕ6, a double-stranded RNA virus, is reviewed in Chapter 8 of this book (Poranen, Pirttimaa and Bamford). In this virus, a ring-shaped NTPase located at a procapsid vertex is responsible for packaging each of three dsRNA segments into the interior of a preformed ϕ6 procapsid. In a twist from the dsDNA viruses, this motor is also responsible for extrusion of newly synthesized message RNAs from the capsid upon the next round of infection. Importantly, the prokaryotic RNA packaging system shows functional similarity to the eukaryotic reoviruses, and again suggests that a general packaging mechanism traverses prokaryotic-eukaryotic boundaries. It is further clear that the ϕ6 packaging and replication machinery share many of the features common to the dsDNA packaging motors.

A coherent mechanistic model for any complex biological process requires (i) a description of the macromolecules involved, (ii) a detailed understanding of how these molecules interact in the formation of larger biological structures, (iii) a description of the catalytic activities associated with these complexes, and (iv) an accounting of the processes that link catalytic activity to structure and function. Genome packaging is a crucial step in virus assembly in a number of prokaryotic and eukaryotic viruses. The molecular motors responsible for this process show mechanistic similarity in viruses as distinct as bacteriophage λ, herpes virus and the dsRNA bacteriophage ϕ6. The chapters in this book provide a detailed summary of our current state of knowledge of the genetics, biochemistry and structure of these fascinating motors. The recent emergence of "new" viral scourges responsible for diseases such as SARS, West Nile fever, etc., and the increasing threat of biological weapons underscore the need to understand virus development at the most basic biological level. We hope that this book provides the experimental background and a philosophical roadmap towards this goal.

CHAPTER 2

Bacteriophage Lambda Terminase and the Mechanism of Viral DNA Packaging

Michael Feiss and Carlos Enrique Catalano

Abstract

The developmental pathways of many double-stranded DNA (dsDNA) viruses, both prokaryotic and eukaryotic, are remarkably similar. In viruses as diverse as bacteriophage λ and the herpesviruses, DNA replication proceeds through a rolling circle mechanism where the circular genome serves as a template for the synthesis of linear concatemers multiple genomes in length. Concurrently, viral gene expression produces structural proteins, which self-assemble into procapsids and, in the case of the bacteriophage, tails necessary to assemble an infectious virion. Virus assembly requires that monomeric virion DNA molecules be produced from concatemers during packaging of the DNA into a procapsid. Thus, packaging represents the convergence of the DNA replication and capsid shell assembly pathways. Genome packaging in bacteriophage λ has been extensively studied and this system has been used as a paradigm for virus assembly. Here we summarize current knowledge, present a working model, and indicate issues worthy of further investigation.

Bacteriophage Lambda Infection and DNA Replication

A λ virion consists of a 48.5 kb dsDNA genome tightly packaged within an icosahedral protein shell and a tail, which serves to deliver the linear genome through the cell envelope into the cytoplasm of an *Escherichia coli* cell. Virus infection initiates with adsorption of the virus particle to the surface of a cell; this interaction is mediated by the gpJ[a] protein of the viral tail and the LamB maltodextrin porin protein of the cell (Fig. 1A).[1] A partially understood series of events ultimately leads to "injection" of the genome into the cytoplasm. The linear genome immediately circularizes via 12-base, complementary "sticky" ends and the nicks are sealed by host ligase yielding a circular duplex. The annealed sticky ends form one subsite of *cos*, the cohesive end site of the λ genome.

[a]Proteins expressed from lambda genes are prefaced with a "gp", for gene product. For instance, the protein products of the *J*, *Nu1*, and *A* genes are gpJ, gpNu1, and gpA, respectively.

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

Table 1. Genes and proteins involved in lambda assembly

Gene/Sequence	Gene Product	Function
b	gpB	Portal protein
c	gpC	Protease
nu3	gpNu3	Capsid scaffold
e	gpE	Major capsid protein
FI	gpFI	Assembly catalyst
C, E	pX1	gpC-gpE fusion protein, portal
C, E	pX2	gpC-gpE fusion protein, portal
D_L	-	The left end of a mature λ genome
D_R	-	The right end of a mature λ genome
cos	-	DNA sequences required for λ DNA recognition, processing, and packaging into the capsid shell
cosN	-	The *cos* subsite where duplex nicking occurs
cosNL	-	The left *cosN* half-site
cosNR	-	The right *cosN* half-site
cosB	-	gpNu1 binding site for initiation of λ DNA packaging
cosQ	-	The *cos* subsite required for termination of λ DNA packaging
Nu1	gpNu1	The λ terminase small subunit
A	gpA	The λ terminase large subunit
himA, hip	IHF	*E. coli* site specific DNA bending protein
hupA,B	HU	*E. coli* histone-like protein

Bacteriophage λ is a temperate phage, which means that the virus may enter either of two infection pathways. The decision of which pathway to enter depends on the physiology of the host cell and the multiplicity of infection. In the lysogenic pathway, lytic genes are repressed and the viral chromosome integrates into the host chromosome by site-specific recombination, forming a repressed prophage. Lysogeny has been extensively described[2-4] and will not be considered here. The second fate is the lytic pathway.[5] In this case, the lambda *O* and *P* genes are expressed, yielding replication proteins that initiate viral DNA synthesis at *ori*. Initially, DNA synthesis by *E. coli* DNA polymerase III follows a classical Θ replication mechanism where bidirectional replication forks synthesize daughter circles (Fig. 1A). Later during infection, a rolling circle mechanism (σ replication) predominates, which produces linear end-to-end polymers of λ chromosomes, called concatemers. Circular concatemers are also produced by recombination between circular molecules, but linear concatemeric DNA is the major substrate for the assembly of infectious virions.

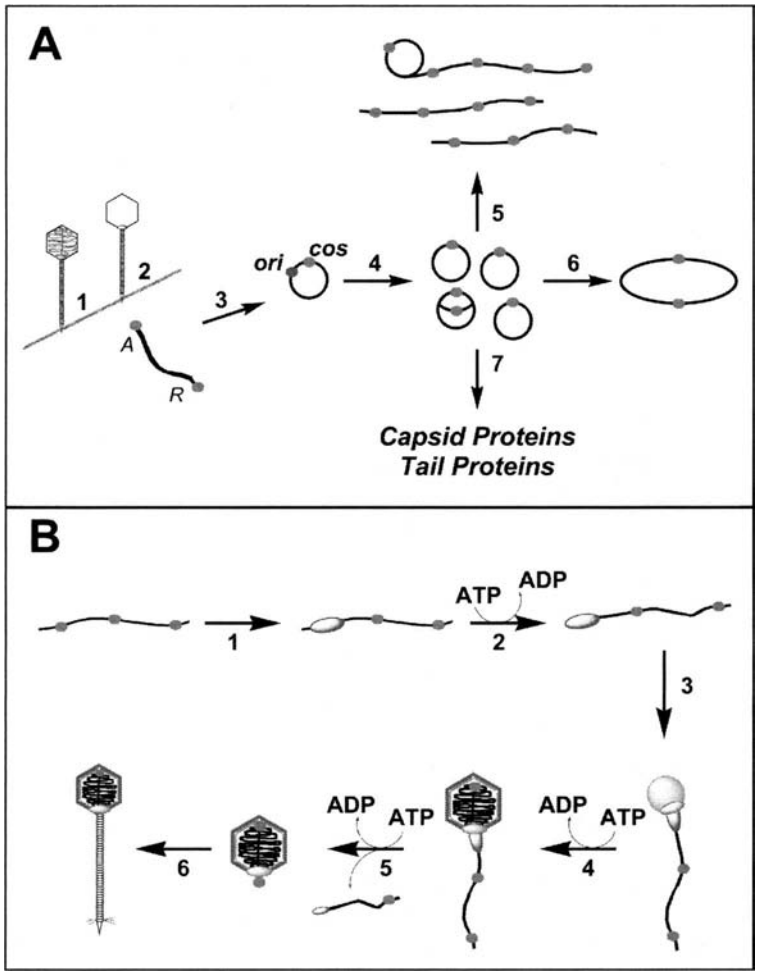


Figure 1. Developmental pathway for bacteriophage lambda. Panel A: Infection of an *E. coli* cell and replication of viral DNA. Infection initiates with adsorption of the virus to the cell surface (1) followed by "injection" of viral DNA into the cytosol of the host cell (2). By convention, the left and right ends of the genome are demarcated by the *A* and *R* genes, respectively. The linear genome circularizes via the 12-base single-stranded cohesive ends (red dots) and the nicks are sealed by host ligase (3). DNA replication initially proceeds via a bi-directional (Θ) replication mechanism yielding daughter circles (4). At later times, DNA is replicated via a rolling circle (σ) mechanism that produces end-to-end concatemers of the viral genome (5). The duplicated *cos* sites in concatemeric DNA are indicated with red dots. Circular concatemers are also produced via recombination (6). Transcription and translation of the viral genome (7) yields structural proteins required to assemble an infectious virus. Panel B: Processing and packaging of viral DNA into a procapsid, and tail attachment to yield an infectious virion. The terminase enzyme (blue oval) binds to a *cos* site in the concatemer (1). The holoenzyme nicks the duplex and separates the strands to "mature" the genome end for packaging (initial *cos*-cleavage) (2). Strand separation requires ATP hydrolysis. The binary terminase•DNA complex binds to a procapsid (cyan sphere) (3) which activates a packaging ATPase activity and translocation of DNA into the capsid (4). DNA packaging promotes capsid expansion, which increases the volume of the capsid and increase angularization of the capsid. The translocating terminase complex stops at the downstream *cos* site and again nicks the duplex (terminal *cos*-cleavage) to complete the packaging process. Strand separation releases the DNA-filled capsid (5) and tail attachment completes the virion (6). A color version of this figure is available online at http://www.Eurekah.com.

recruits a procapsid, sponsors insertion of the DNA into the procapsid, and finally cuts the end of the genome to complete the packaging process (Fig. 1B). As with other terminase enzymes, λ terminase is a heteroligomer composed of small (gpNu1) and large (gpA) subunits (see Fig. 3). GpA carries the DNA cutting activity required to initiate and terminate DNA packaging. This is accomplished through a site-specific endonuclease activity that introduces nicks into *cos* that are staggered by 12 bp. GpA also has a so-called "helicase" activity that separates the nicked strands thus generating the single-stranded "sticky" ends of the mature genome. The gpA subunit further contains a putative DNA translocase activity that is responsible for active DNA packaging, and an ATPase activity that powers translocation. While the large terminase subunit possesses all of the catalytic activities required to cut and package the viral genome, gpA alone exhibits low catalytic activity.[6-10] The gpNu1 subunit specifically recognizes *cos*, and is responsible for the assembly and stability of the packaging machinery. The biological activities of λ terminase are discussed in detail below.

In summary, the packaging pathway entails terminase assembly at a *cos* site in the concatemer and cutting of the duplex (the initial *cos* cleavage reaction), which yields the mature left end of the genome to be packaged. Upon binding a procapsid, the packaging machinery translocates DNA into the capsid through a capsid structure known as the portal vertex (active DNA packaging). Upon arrival at the next downstream *cos* site in the concatemer, the packaging machinery stops and terminase again cuts the duplex generating the mature right end of the genome (the terminal *cos* cleavage reaction); this process yields a single viral genome tightly packaged within the confines of the capsid as described in Figure 1B.

Bacteriophage λ *cos*: A Multipartite Assembly Site

cos

The *cos* site is a ≈200 bp long segment that is required to both initiate and terminate the packaging of a monomeric genome from concatemeric DNA. The site where terminase introduces staggered nicks to generate the cohesive ends is called *cosN* (Fig. 2). Early during the study of λ, it was thought that *cosN* was both necessary and sufficient for DNA packaging. Later studies showed, however, that *cos* is complex and consists of three and perhaps four distinct subsites.[11-15] Both initiation and termination of packaging require duplex nicking at the *cosN* site; additionally, efficient initiation requires the presence of the *cosB* subsite, which is directly downstream from *cosN*. Conversely, efficient termination requires the presence of *cosQ*, a subsite that is located upstream of *cosN* (Fig. 2).[16] The I2 sequence is located between *cosN* and *cosB*, and also plays a distinct role in efficient DNA packaging.[11] Thus, the complete *cos* sequence consists of several subsites, each of which plays a specific role in the recognition, processing, and packaging of viral DNA. Each of these subsites is discussed in detail below.

cosN

The terminase enzyme introduces nicks into the duplex at the *cosN* site to generate the cohesive ends of mature virion DNA. Many of the base pairs (bp) within *cosN* show two-fold rotational symmetry, which extends over 22 bp if one includes purine-purine and pyrimidine-pyrimidine symmetry (Fig. 2); this has been used as evidence that a symmetrically disposed enzyme complex (i.e., a gpA dimer) is responsible for duplex nicking. This argument is further supported by (i) analogies to the interactions of type II restriction endonucleases with their palindromic recognition sequences, and (ii) the presence of a leucine-zipper motif in the primary sequence of gpA.[17] We presume here that a gpA dimer is responsible for symmetric duplex nicking at *cosN*.

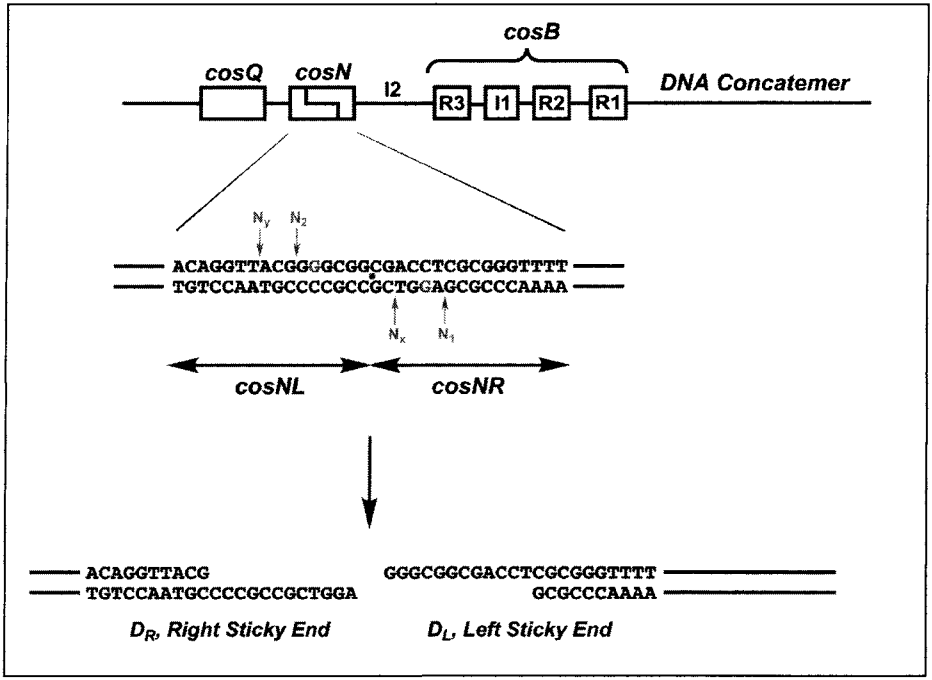


Figure 2. The *cos* region of a lambda concatemer. Upper panel: The *cosQ*, *cosN* and *cosB* subsites within a *cos* site in concatemeric DNA. The *cosB* subsite is composed of the I1 and R-elements, as indicated. The I2 region lies between *cosN* and the R3 element. Middle panel: The nucleotide sequence of *cosN*, with the *cosNL* and *cosNR* half-sites indicated. The center of symmetry of *cosN* is indicated with a dot. Terminase normally nicks the duplex at N_1 and N_2 sites indicated with arrows. In the absence of ATP, terminase incorrectly nicks the duplex at N_x and/or N_y sites. Lower panel: Strand separation by terminase yields the matured D_R and D_L ends of the lambda genome, as shown.

cosB

Terminase nicking of an isolated *cosN* site is error-prone, and the overall rate of nicking is slow.[18,19] The presence of *cosB* and the I2 element (discussed below) is required for efficient and accurate duplex nicking. The *cosB* subsite contains three 16 bp R elements that are specifically recognized by gpNu1 (Fig. 2);[12,33,43,55] specific *cosB*-gpNu1 interactions are likely required to properly position a gpA dimer at *cosN* for the initial *cos* cleavage event.[18,19] The *cosB* subsite also contains a consensus sequence for *Escherichia coli* integration host factor (IHF). This I1 site introduces an intrinsic bend into the duplex, and is also specifically recognized by IHF.[20,21] The role of IHF in DNA packaging and virus assembly is discussed more fully below.

cosQ

This seven base pair subsite is located 17 bp upstream of *cosN* (Fig. 2), and is essential for proper termination of the packaging reaction.[16,22-25] Severe *cosQ* mutations do not significantly affect the initial *cos* cleavage reaction, but abolish nicking of the bottom strand at the terminal *cos* site. Moreover, packaging is not arrested in the absence of *cosQ*, and additional DNA, including the downstream *cos*, is packaged until the capsid shell is filled to capacity.[26] This suggests that *cosQ* is required to stop the packaging machinery for appropriate cleavage at

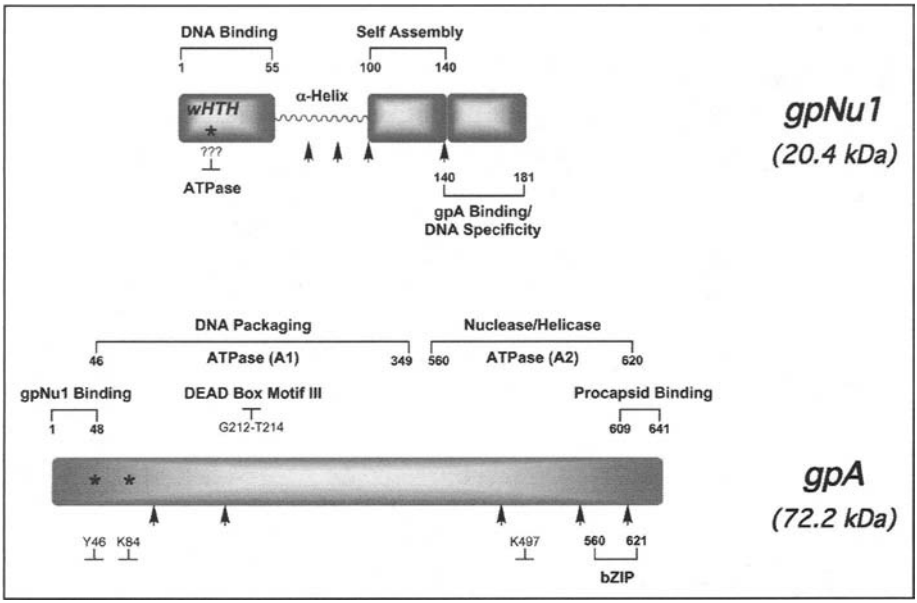


Figure 3. Domain organization of the terminase gpNu1 and gpA subunits. Upper panel shows gpNu1. wHTH indicates the winged helix-turn-helix motif. Lower panel shows gpA. In both panels, sites covalently modified with 8-azido ATP are indicated with asterisks, and specific proteolysis sites are indicated with arrows, respectively. Details are presented in the text.

the terminal *cosN* site; however, *cosQ* alone is incapable of arresting DNA packaging. Rather, *cosQ* acts in concert with *cosN* and I2 to promote efficient termination.[26]

The *I2* subsite, originally identified by sequence homology to the IHF binding site consensus sequence, is located between *cosN* and *cosB* (Fig. 2). I2 is not a functional IHF binding site, however, as evidenced from mutagenesis studies,[27] sequence information content analysis[28] and direct IHF-DNA binding studies.[29] Nevertheless, the I2 region appears to play a role in both initial and terminal *cos*-cleavage events as follows.[b] (i) Mutation of individual bases within the I2 site does not affect the initial *cos*-cleavage reaction, while small deletions reduce nicking accuracy in vitro and packaging initiation in vivo.[27,30] Based on these data, it was proposed that the role of I2 in the initial *cos* cleavage reaction may be to simply provide the proper spacing between the *cosN* and *cosB* subsites.[27,30] (ii) The role of I2 in the late steps of DNA packaging is more interesting. While the *cosB* sequence is not required for the terminal *cos*-cleavage reaction, efficient nicking of a downstream *cosN* site requires the I2 region.[11] Moreover, it has recently been found that replacing the I2 sequence with a randomized sequence that preserves the *cosN-cosB* spacing is profoundly lethal (B. Charbonneau and M. Feiss, unpublished). Thus, the I2 sequence serves a specific role in terminating the packaging reaction. In other words, *cosQ* and *cosN* alone are insufficient for terminal *cos* cleavage, and efficient termination requires that the entire *cosQ-cosN*-I2 segment be present.

We note that terminase remains bound to the concatemer after the DNA-filled capsid has been released (see Fig. 1B). This binary enzyme-DNA complex binds an empty procapsid to

[b]The I2 element is not a functional IHF binding site but this segment nevertheless plays a role in genome packaging. Here we refer to the region between *cosN* and *cosB*, namely bp 18 to 49, as the I2 subsite for historical reasons.

initiate a second round of packaging, and terminase thus packages successive genomes in the concatemer in a processive manner. While the terminal *cos*-cleavage reaction requires only the I2 element, processive packaging requires that the *cosB* subsite also be present.[11]

Components of the Packaging Machinery

As discussed above, terminase enzymes form an integral part of the DNA packaging motor in a variety of dsDNA viruses. λ terminase is a heteromultimer of gpNu1 and gpA subunits. Terminase holoenzyme possesses a gpA_1•$gpNu1_2$ subunit stoichiometry;[31] however, the terminase subunit composition may vary along the packaging pathway (*vide infra*). Here we discuss the structure, catalytic activities, and function of the individual subunits, and the holoenzyme complex.

gpNu1

The phage λ small terminase subunit (181 amino acids) is responsible for the assembly of the packaging machinery, and for the stability of the packaging initiation complexes.[32] gpNu1 assembles at *cosB*, and direct interactions with the three R-elements have been demonstrated by DNase protection assays.[33] ATP increases gpNu1 binding to *cos* containing DNA substrates without affecting nonspecific DNA binding (M. Ortega and C.E. Catalano, unpublished). Nucleotides do not affect the DNase protection pattern, however.[33] While the isolated subunit can bind ATP, significant hydrolysis by gpNu1 is observed only in the context of the holoenzyme (H. Gaussier and C.E. Catalano, unpublished).[9,34]

Structural and Functional Domains

A domain organization of gpNu1 was demonstrated using chimeric constructs of phage λ and the closely related phage 21. The capsid genes of phages λ and 21 descend from a common ancestor and are of similar size, function and genetic structure.[35,36] Furthermore, the λ and 21 terminase genes share about 60% sequence identity. Despite sequence homology, the λ and 21 gene products are not interchangeable due to divergent interaction specificities.[37] In other words, λ terminase packages λ DNA specifically into λ procapsids, while phage 21 terminase specifically utilizes phage 21 procapsids. Viable λ-21 constructs contained chimeric terminase genes and were used to show the locations of specificity domains.[38-40] These studies show that the N-terminal half of gpNu1 contains the *cosB* binding determinant, while the C-terminal half interacts with gpA (see Fig. 3).[39]

More recent biochemical and genetic studies have defined three structural and functional domains in gpNu1, as described in Figure 3. The C-terminal 40 amino acids of the protein are required for efficient gpA-binding interactions and holoenzyme formation.[41,42] This region of the protein may further play a role in discrimination between *cos* containing and nonspecific DNA substrates.[42] Residues Lys100 – Pro140 define a hydrophobic domain that is required for high-affinity DNA binding interactions; deletion of this self-association domain decreases DNA binding interactions by three orders of magnitude.[41,42] This region of the protein is also responsible for the observed aggregation of the protein in solution. The N-terminal 55 residues of gpNu1 define the minimal DNA binding domain of the protein.[41-43] Biochemical and biophysical studies suggest that residues Ala55 – Lys100 form an extended helix connecting the DNA binding domain and the self-association domain of the protein. It has been proposed that this helix forms a flexible linker between the two domains that alternately plays a role in (i) cooperative gpNu1 binding interactions at *cosB* and (ii) cooperative assembly of gpA at the *cosN* subsite.[44]

The Three Dimensional Structure of gpNu1

A high-resolution NMR solution structure for the gpNu1 DNA binding domain (gpNu1-DBD) has recently been solved (see Fig. 4A).[43] The structure, combined with NMR-monitored titration studies and genetic experiments, confirm sequence analysis predictions

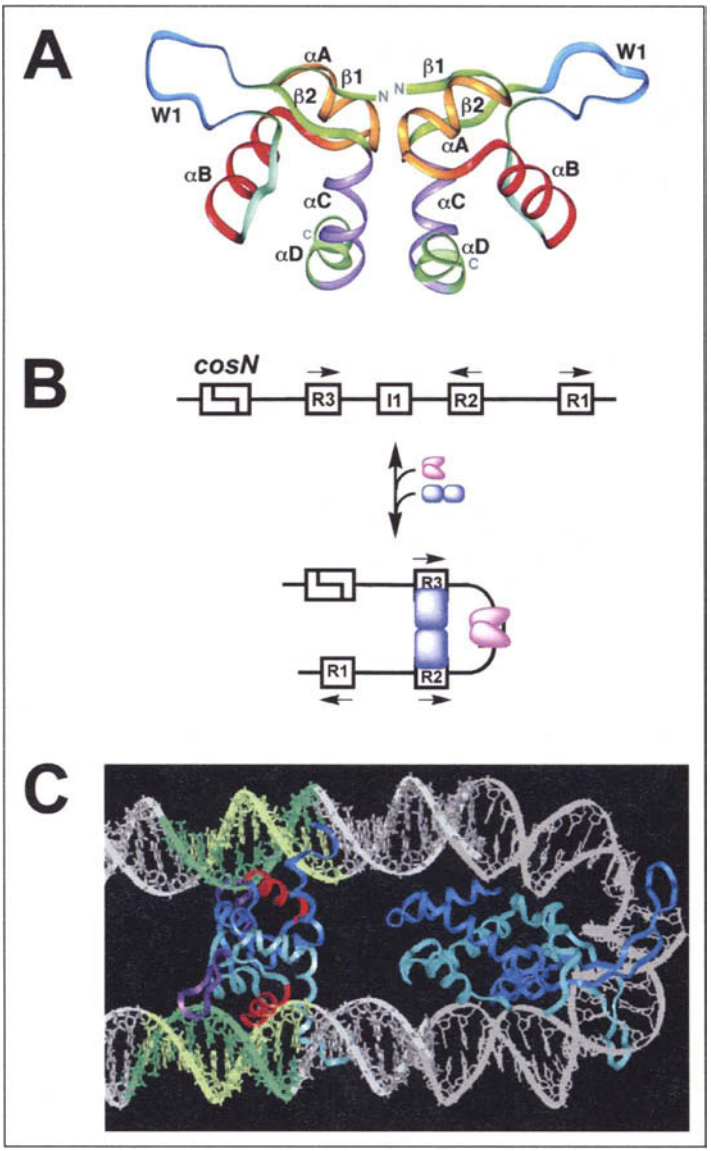


Figure 4. Structure of gpNu1-DBD and model for cooperative gpNu1-IHF DNA binding. Panel A: The High Resolution Structure of the DNA Binding Domain of gpNu1 (gpNu1-DBD). The average of the 20 lowest energy structures obtained from NMR is shown in ribbon representation. The winged HTH motif is made up of wing residues (W1, blue), helix αA (yellow) and the recognition helix αB (red). Panel B: Cartoon describing cooperative assembly of gpNu1 and IHF at *cosB*. The gpNu1 dimer (blue) binds to two R-elements (R3 and R2) spatially separated in the duplex. Duplex bending is supported by IHF (purple) binding to the I1 element. Panel C: Molecular model showing cooperative binding of gpNu1 and IHF. The R-elements in the DNA duplex are colored green/yellow. The α and β subunits of IHF, bound at I1, are colored blue/cyan. The gpNu1 dimer subunits are colored blue/cyan, with wing and recognition helices colored purple and red, respectively. Note that wing residues are positioned to directly interact with bases in the minor groove of the duplex. Reproduced from reference 43, with permission.

that DNA binding is mediated by a helix-turn-helix (HTH) DNA binding motif in this region of the protein (residues Lys5 – Glu24).[45-47] Consistently, substituting the recognition helix from the small subunit of phage 21 terminase into gpNu1 created a chimeric terminase specific for packaging of 21 DNA. Structural analysis of gpNu1-DBD further classified the motif as a winged helix-turn-helix (wHTH) DNA binding motif. The wHTH is present in a diverse family of DNA binding proteins, including rat hepatocyte nuclear factor-3,[48] *Drosophila* homeotic fork-head protein,[49] the bacterial OmpR response regulator protein,[50] and the phage Mu C repressor and transposase proteins.[51,52] These proteins bind DNA utilizing classical HTH-major groove binding interactions, with additional affinity and/or specificity modulated by interactions between wing residues and the minor groove of DNA.[53]

Cooperative gpNu1 DNA Binding to *cosB*

Based on similarities to the phage λ repressor-operator system,[54] it was initially proposed that gpNu1 dimers cooperatively assemble at the three R-elements of *cosB*.[55] Biophysical and structural analysis of gpNu1 confirm that the DBD is indeed a dimer;[43,44] however, the NMR structure of gpNu1-DBD reveals that the dimer possesses C2 symmetry in which the wHTH motifs face away from each other, frustrating a simple model for DNA binding. Based on the solution structure for gpNu1-DBD and the organization of *cosB*, a second model for gpNu1 assembly is proposed in Figure 4B. Importantly, this model accommodates the observed effects of IHF on lambda development, which are detailed below.

Escherichia coli Integration Host Factor (IHF)

IHF is a site-specific DNA binding protein that introduces a ≈180° bend into duplex DNA.[56,57] IHF often exerts its biological effect by forming nucleoprotein complexes that are conducive to the assembly of other DNA binding proteins at that site;[58] direct IHF-protein interactions in these higher-order complexes have not been demonstrated. IHF modulates λ development in vivo, with virus yields lowered 3 to 4-fold in its absence[c].[56,59] Importantly, in vitro DNA packaging is also lowered ≈4-fold in the absence of IHF,[60,61] consistent with a direct role for the host factor in virus assembly. An IHF consensus sequence, I1, has been identified between R3 and R2 in *cos* and direct IHF-I1 binding interactions have been demonstrated (Fig. 4B).[20,56,59]

Genetic and mutagenesis studies show that any of a number of mutations affecting the R-elements of *cosB* renders the virus IHF-*dependent* for growth,[62] and a direct role for IHF in gpNu1 assembly at *cosB* has been implicated.[63,64] It is presumed that cooperative assembly of gpNu1 at *cosB* is efficient under normal conditions, and IHF plays a supportive role only; however, when gpNu1-DNA interactions are attenuated (i.e., by mutation), gpNu1 assembly requires additional interactions, presumably provided by structural alterations in the DNA induced by IHF.

Model for an IHF-gpNu1 Nucleoprotein Complex

The high-resolution structure of gpNu1-DBD, while initially perplexing, led to an attractive model that nicely accommodates both gpNu1 assembly at *cosB*, and the supportive role of IHF in virus assembly (Fig. 4B,C). In this model, IHF binding to I1 introduces a strong bend in the duplex, which juxtaposes the R3 and R2 gpNu1 binding sites in the appropriate orientation. A single gpNu1 dimer spans these two "half-sites" that are separated by 44 bp in the duplex. This model predicts that both the DNA sequence (R-elements) and duplex structure

[c]IHF is not crucial for plaque formation in the laboratory. We recognize, however, that a λ mutant unable to utilize IHF would be at a disadvantage in the environment.

(bend) are critical determinants to gpNu1 assembly. Importantly, an intrinsic bend is observed at the I1 site,[20,65] which would promote gpNu1 binding in the absence of IHF. Mutation of an R-element decreases intrinsic gpNu1 binding affinity, which places increased emphasis on the bend; IHF binding provides the prerequisite DNA structure. While this model is fully consistent with all of the genetic, biochemical, and structural data available, it does not directly address the role of the R1-element in gpNu1 assembly. Of note, however, is that this element is dispensable in the presence of IHF in culture[d].[55,62,64]

gpA

The large terminase subunit is composed of 641 amino acids. Genetic and biochemical studies have defined several functional domains of the protein, including an N-terminal domain responsible for gpNu1 interactions in the holoenzyme complex,[39,40] and a C-terminal domain responsible for procapsid binding during initiation of DNA packaging[21,38,66] (Fig. 3). Other domains are described in detail below.

The Nuclease/Helicase Domain

The endonuclease activity of terminase is contained within the C-terminal half of gpA, based on the results of random mutagenesis studies.[17,67] Unexpectedly, further evidence for a C-terminal nuclease domain in gpA comes from mutational analysis of the ATPase activity of the enzyme, as follows. Walker and others have defined a common motif in a variety of ATP- and GTP-binding proteins;[68,69] central to this motif is a phosphate-binding loop, or P-loop, comprised of a number of glycine residues and a conserved lysine that is critical to NTP binding and/or hydrolysis (consensus, GXXGXGK[S/T]). Kinetic studies demonstrate that gpA possesses ATPase activity,[70,72] and primary sequence analysis indeed identified a P-loop sequence located between Gly491 - Lys497 of the protein.[71] Mutagenesis studies sought to provide a link between this putative P-loop, the ATPase activity of gpA, and the packaging activity of the holoenzyme; however, mutation of the "critical" lysine in the putative P-loop (Lys497) did not significantly affect the ATPase or the packaging activities of terminase.[72] Rather, these mutations simultaneously abolished the endonuclease and helicase activities of gpA.[73] The data are consistent with genetic experiments demonstrating that both nuclease and helicase activities localize to the C-terminal half of gpA, and further suggest that there is a structural and functional overlap between the two catalytic activities. We note that the endonuclease and helicase activities are stimulated and fueled, respectively, by ATP. We presume that a distinct ATPase site is associated with these catalytic activities; however, direct evidence for an independent nuclease/helicase ATPase site remains elusive.

The Translocase Domain

The DNA packaging activity that is required to translocate DNA into the procapsid is located between residues Tyr46 to Asp349 of gpA (Fig. 3).[74,75] This functional domain was identified through the analysis of mutants that had defects in virus development, but that retained normal endonuclease and helicase activities. All of these mutations are located in the amino terminal 60% of gpA, and all possess post-*cos* cleavage defects[e] which may be grouped as follows: (i) DNA packaging is completely deficient,[74] (ii) DNA is slowly and/or only partially

[d]We recognize that the R1 element likely provides a competitive advantage to virus in the environment. Indeed, studies suggest that the presence of R1 contributes about 5% of the yield of wild type lambda growing in IHF$^+$ *E. coli* (Cue and Feiss. J Mol Bio 1992a; 228:58-71).

[e]Post-*cos* cleavage defects are defined as those that do not affect duplex nicking or strand separation steps (i.e., *cos* cleavage), but that are defective in one or more aspects of the subsequent packaging of DNA into the capsid.

packaged, or (iii) DNA is packaged but infections virion assembly does not occur. Importantly, mutations that result in these post-*cos* cleavage defects are distinct from and do not overlap with those that result in *cos* cleavage and helicase defects described above. The genetic results thus indicate that two independent functional domains are responsible for (i) preparation of the genome for packaging (nuclease/helicase domain) and (ii) active DNA packaging (translocase domain) as illustrated in Figure 3.

The Self-Association Domain(s)

It is presumed that rotationally symmetric gpA subunits bound to *cosN* are responsible for duplex nicking based, in part, on the two-fold rotational symmetry of the *cosN* sequence (see Fig. 2). This model requires that gpA self-associates, and the gpA sequence Leu560 - Asp620 bears strong sequence homology to the consensus basic leucine zipper (bZIP) DNA binding motif[17] which mediates protein dimerization linked to DNA binding.[76] Consistently, mutation of Glu586 to lysine (gpA-E586K) specifically inactivates the endonuclease activity of terminase holoenzyme, presumably due to a defect in the assembly of a gpA dimer at *cosN*.

The Packaging ATPase

The failure of gpA "P-loop" mutations (i.e., Lys497 discussed above) to affect the ATPase and packaging activities of the holoenzyme led to the search for the "packaging ATPase" site. Photoaffinity labeling of terminase with 8-azidoATP identified gpA residues Tyr46 and Lys84 as being proximate to an ATP binding site,[75] a location quite distant from the putative P-loop centered at Lys497. Mutation of either Tyr46 or Lys84 strongly affects both ATPase and DNA packaging activities of terminase holoenzyme, but does not significantly affect *cos* cleavage or helicase activities.[75,77] These data provide additional support for the domain organization of gpA outlined in Figure 3, and further implicate Tyr46 and Lys84 as important residues associated with the packaging ATPase catalytic site.

The Packaging ATPase Is Conserved among Terminase Enzymes

The presence of an N-terminal packaging ATPase is consistent with comparative sequence analysis of terminase enzymes from a variety of dsDNA viruses.[78] Conserved sequences are observed in these viruses, and include the Walker A (P-loop) and B motifs, an adenine binding motif, and a conserved glutamate presumed to provide the catalytic carboxylate for ATP hydrolysis. All of these motifs are found in the primary sequence of the λ terminase gpA subunit.[f] Interestingly, the two mutations that abrogate the ATPase and DNA packaging activities of λ terminase discussed above correspond to conserved motifs found in all of the terminase large subunits: (i) Tyr46 is a strictly conserved residue found in the adenine binding motif and (ii) Lys84 immediately follows a conserved Walker A motif found in gpA. It has been proposed that these motifs represent a conserved ATPase catalytic site that is directly associated with the packaging activity of terminase enzymes (*ibid*). Rao and coworkers have demonstrated that mutation of an N-terminal Walker A sequence of the large subunit of bacteriophage T4 terminase abrogates the ATPase and DNA packaging activities of enzyme, without loss of DNA cutting functions.[79] This is identical to the observed effects of Tyr46 and/or Lys84 mutations in gpA, and the data are consistent with the notion of a conserved DNA packaging domain in the N-terminus of all terminase large subunits.

Comparative sequence analysis of terminase enzymes has also revealed the presence of "motif III" in the N-terminus of the large terminase subunits;[78] residues Gly212-Thr214 rep-

[f]Initial sequence analysis of gpA revealed a putative P-loop sequence in the C-terminus of the protein, centered about Lys497 as described above. Closer inspection reveals that gpA possesses a second, albeit weaker, match to the P-loop consensus sequence in the N-terminus of the protein, positioned at Lys84.

resent motif III in the λ gpA protein (Fig. 3). This motif is a common feature found in the ATPase domains of DEAD box helicases.[80,81] It has been postulated that motif III serves as part of an ATP transducing switch that couples ATP hydrolysis to translocation in the DEAD box helicases.[81] The presence of a conserved motif III in the putative packaging domain of terminase large subunits suggest that these residues may play a direct role in coupling ATP hydrolysis to DNA translocation and packaging.

Terminase Holoenzyme

Lambda terminase is isolated as a heteromultimer with a subunit ratio of gpA_1•$gpNu1_2$.[31] As outlined above, the holoenzyme possesses a site-specific endonuclease activity, a so-called helicase activity, and a DNA packaging activity that act in concert to package viral DNA. The holoenzyme also possesses multiple ATPase catalytic sites that play unique roles along the packaging pathway. Each of these catalytic activities is discussed in turn below.

Nuclease Activity

While the isolated gpA subunit will cut a *cos* containing DNA substrate, this nuclease activity is strongly stimulated by interactions with gpNu1.[10] Kinetic data suggest that terminase assembly at *cos* is the rate-limiting step in the *cos* cleavage reaction.[10,31] Once assembled, gpA subunits introduce symmetric nicks at the *cosN* subsite, twelve base-pairs separated in the duplex (see Fig. 2). Assembly of the holoenzyme at *cos* occurs efficiently in the absence of metals, but duplex nicking is strictly dependent on Mg^{2+} (or Mn^{2+}).[31] Kinetic data further suggest that two gpA molecules are required to nick the duplex, consistent with the assembly of a symmetric gpA dimer at *cosN*.[31]

ATP modulates the nuclease activity of the holoenzyme in two important ways. First, ATP strongly stimulates the rate of duplex nicking. It is noteworthy that 50 μM ATP is sufficient to fully stimulate nuclease activity[82] which is comparable to the K_m for ATP hydrolysis by the gpA subunit (K_m = 5 μM).[70,72] In contrast, millimolar concentrations of ATP are required stimulate DNA binding by the holoenzyme (M. Ortega and C.E. Catalano, unpublished), a value that is similar to the K_m for ATP hydrolysis by the gpNu1 subunit (K_m = 0.5 mM in the presence of DNA).[60,67,70] Thus, the data suggest that ATP plays distinct roles in both terminase assembly at *cos* (mediated through the gpNu1 subunit) and stimulation of nuclease activity (mediated through the gpA subunit). A second effect of ATP is the modulation of nuclease fidelity; in the absence of ATP terminase incorrectly nicks the duplex, predominantly generating a four base pair nick which is lethal to virus development (Fig. 2).[6] The effects of ATP on gpA nuclease activity presumably arise from interactions with the C-terminal ATPase catalytic site described above and illustrated in Figure 3.

Cooperative Assembly of Terminase at *cos*

A subtle interaction between gpNu1 assembly at *cosB* and gpA assembly at *cosN* has been exposed through mutagenesis studies, as follows. Introduction of mutations into the right *cosN* half site (cosNR, Fig. 2) has little effect on the *cos* cleavage reaction in an otherwise wild-type background.[83] Conversely, introduction of a symmetric mutation into the left *cosN* half site (cosNL) has dramatic effects, significantly decreasing the rate of the *cos* cleavage reaction in vitro[83] and virus yield in vivo.[84] Introduction of secondary mutations in *cosB*, or complete deletion of *cosB* removes this asymmetry, and cosNR and cosNL mutations have identical effects. These data suggest that intrinsic DNA binding interactions between gpA and cosNL are critical to efficient duplex nicking, and that introduction of mutations into this DNA site strongly affect the reaction. Conversely, gpA binding to the cosNR half site is supported by cooperative interactions with gpNu1 assembled at *cosB* (Fig. 5). Thus, introduction of

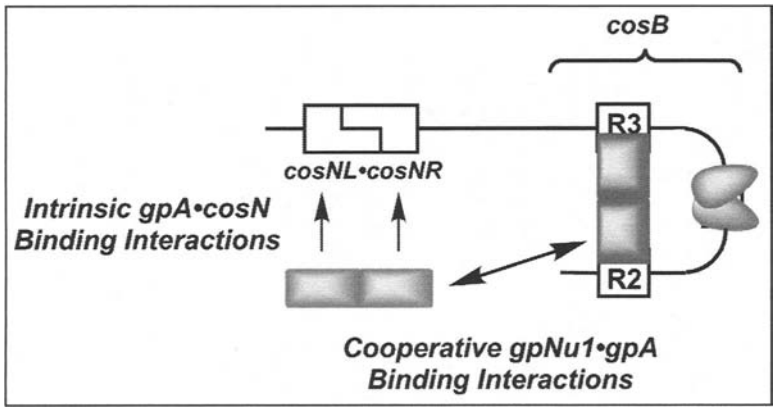


Figure 5. Model for cooperative assembly of gpA (red rectangle) and gpNu1 (blue rectangle) at *cos*. A symmetric gpA dimer binds to *cosN*. Binding of gpA to the *cosNL* half-site relies strictly on intrinsic binding interactions. Binding of gpA to the *cosNR* half-site is supported by cooperative interactions with gpNu1 and IHF (purple lobes) assembled at *cosB*. A color version of this figure is available online at http://www.Eurekah.com.

mutations into the cosNR half-site are "masked" by supportive interactions in the holoenzyme-DNA complex. Deletion (or mutation) of *cosB* DNA abrogates gpNu1 assembly, which in turn abolishes cooperative gpNu1-gpA binding interactions. The outcome is that gpA binding at both half-sites relies exclusively on intrinsic binding interactions, and symmetric *cosN* mutations have identical results.

Strand-Separation (Helicase) Activity

Subsequent to duplex nicking, the annealed strands are actively separated by the holoenzyme (Fig. 2).[7,73,85] The energy required to separate the annealed twelve bases of DNA is provided by ATP hydrolysis, presumably at the C-terminal ATPase catalytic site in gpA (A2, Fig. 3). While this reaction has been described as a helicase activity, processive duplex separation has not been demonstrated. Formally, this reaction constitutes a single catalytic turnover of a typical helicase enzyme,[86,87,88] and we therefore refer to the reaction as a strand-separation activity. The reaction is strictly dependent on divalent metal, with Mg^{2+} being most efficient, and ATP (or dATP) hydrolysis.[85]

Strand-separation releases the right chromosome end (D_R) from the nucleoprotein complex and yields a packaging intermediate composed of the terminase subunits tightly bound to the mature left end of the lambda genome (D_L, Figs. 1B and 2).[85] This intermediate is remarkably stable in vitro ($T_{1/2}$ ≈8 hours),[85] and because of this stability catalytic turnover is not observed in the absence of procapsids. Early studies isolated a stable intermediate in the packaging reaction in vivo. This species, known as complex I, could be isolated on sucrose gradients and chased into infectious virus with the addition of a cell extract containing λ tails.[89] We presume that the stable complex characterized in vitro is in fact complex I characterized in vivo, though this has not been rigorously demonstrated. The stability of complex I, presumably mediated by gpNu1 interactions with *cosB*, is likely critical for protection of the matured genome end prior to DNA packaging in vivo; should the complex prematurely dissociate, the 12 base single-stranded end is expected to be rapidly degraded by cellular nucleases, a lethal event. The nature of complex I is more fully described below.

ATPase Activity

It has long been appreciated that terminase possesses a DNA-stimulated ATPase activity, and early models proposed that ATP hydrolysis provided the energy to power translocation of the packaging machinery. Kinetic analysis of ATP hydrolysis by the holoenzyme identified two distinct catalytic sites, with K_m's of 5 μM and 1 mM.[70] Primary sequence analysis identified a putative P-loop motif in each subunit of the enzyme,[71,90] supporting the contention that both gpA and gpNu1 could bind and hydrolyze ATP. Subsequent mutational analysis indeed confirmed that the high- and low-affinity ATPase sites were located in the gpA and gpNu1 subunits, respectively.[67]

The gpNu1 ATPase Catalytic Site

Support for the hypothesis that gpNu1 possessed a functional P-loop was derived from affinity studies, which demonstrated that photoreactive ATP analogs covalently modified peptides in this region of the protein (Thr18-Lys35 and Met1 - Gln20.[91] Mutational analysis yielded conflicting data, however. Mutation of the "critical" P-loop lysine in gpNu1 (Lys35) indeed abrogated ATP hydrolysis by this subunit, but only when limiting concentrations of DNA were used; increasing the concentration of DNA in the reaction mixture restored the ATPase activity to wild type levels.[67] This suggested that Lys35 plays a role in DNA binding rather than ATP hydrolysis, a contention that was verified by structural analysis (*vide supra*). The data nevertheless established that the low affinity ATPase site ($K_m \approx 1$ mM) resides in the gpNu1 subunit, and that DNA stimulates the reaction both by increasing k_{cat} and decreasing K_m.[70] The gpNu1 subunit also hydrolyzes GTP with a similar kinetic profile,[9] but the biological relevance of this reaction remains uncertain.

The gpA ATPase Catalytic Sites

As discussed above, a P-loop motif was also identified in the gpA subunit; however, mutation of the "critical" lysine in the putative P-loop (Lys497) had only minor effects on the observed rate of ATP hydrolysis by this subunit.[72] Rather, this residue appears to be intimately involved in both nuclease and strand-separation activities. Clues to the location of the high-affinity ATPase catalytic site were provided by affinity labeling studies using photoreactive ATP analogs. These experiments demonstrated that peptides in the N-terminal region of gpA were modified (Ala59 - Lys84),[91] and that Tyr46 and Lys84 were specifically modified with 8-azidoATP.[75] Mutational analysis of these residues confirmed that both are required for high-affinity ATPase activity. Moreover, these residues are intimately involved in DNA packaging activity.[75,77]

Based on the aggregate data, we propose that two discrete ATPase catalytic sites reside in the gpA subunit. (i) A C-terminal ATP binding site regulates nuclease activity and hydrolyzes ATP to provide the energy required for strand-separation by the enzyme. We refer to this site as the nuclease/helicase ATPase site (A2, Fig. 3). (ii) A structurally and functionally distinct N-terminal ATPase site is responsible for the observed high-affinity ATP hydrolysis activity, and is directly linked to DNA packaging activity. We refer to this site as the packaging ATPase site (A1, Fig. 3). We note that the observed rate of ATP hydrolysis at this site ($k_{cat} \approx 40$ min^{-1}) is insufficient to package a full-length genome.[70,72] Recent kinetic studies have demonstrated an increase in the observed rate of ATP hydrolysis commensurate with that required to power DNA packaging.[61] We presume that the N-terminal ATPase is activated to sponsor translocation.

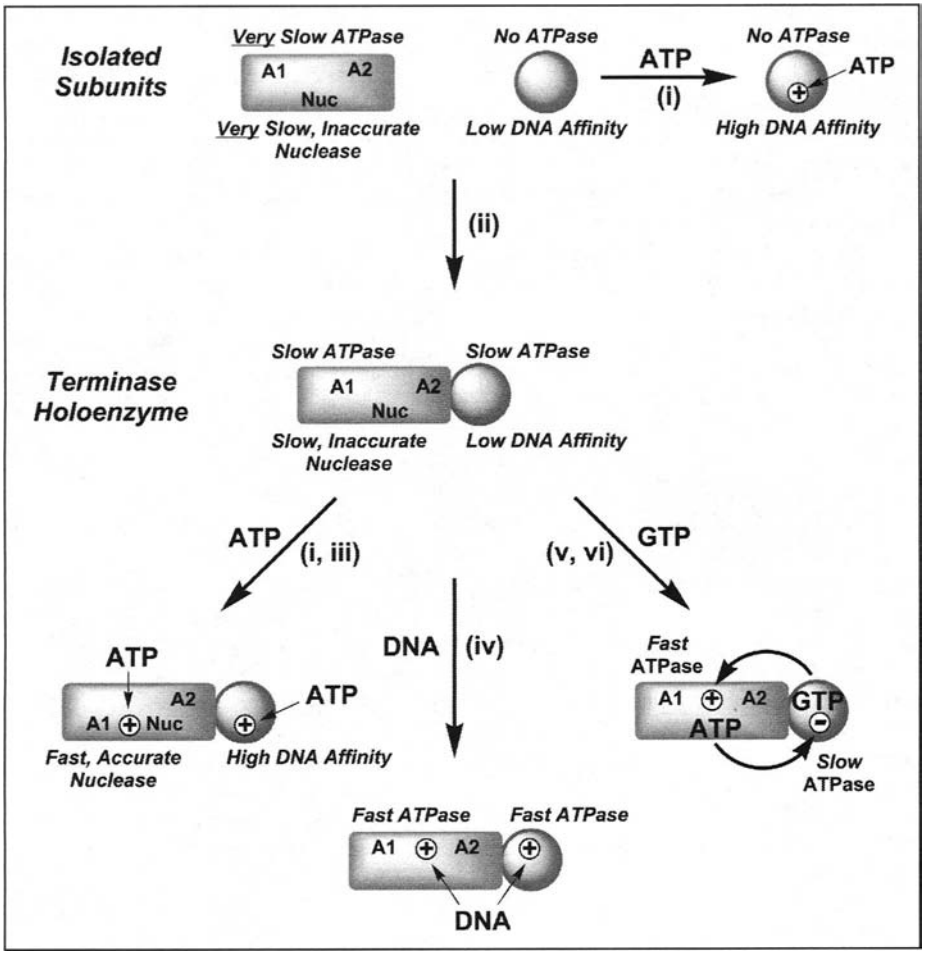


Figure 6. Summary of allosteric interaction between the catalytic sites of terminase holoenzyme. The catalytic activity of the isolated gpA (red rectangle) and gpNu1 (blue sphere) subunits is shown at top. Assembly into the holoenzyme activates catalytic activity as indicated in step (ii). Interaction between the multiple catalytic sites of the holoenzyme is shown in the lower half of the figure. Details are described in the text. A color version of this figure is available online at http://www.Eurekah.com.

Summary of the Allosteric Interactions between Terminase Catalytic Sites

Allosteric interactions between the multiple catalytic sites of terminase holoenzyme have been observed, and are summarized in Figure 6.[g] (i) ATP binding (but not hydrolysis) increases the affinity of gpNu1 for *cos* containing DNA.[32] This interaction is likely responsible for ATP-stimulated DNA binding by the holoenzyme.[82] (ii) The ATPase activity of the individual subunits is mutually stimulated in the holoenzyme complex; ATP hydrolysis by the gpNu1 subunit is dependent upon gpA, and the ATPase activity of gpA is stimulated by gpNu1.[10] Furthermore, the endonuclease activity of gpA is stimulated by interactions with the gpNu1

[g]The numbers presented here in the text refer to those displayed in Figure 6.

subunit.[10] (iii) ATP binding to the gpA subunit stimulates the rate of the nuclease reaction in the holoenzyme, but not by the isolated gpA subunit.[82] Similarly, (iv) DNA stimulates the ATPase activity of both terminase subunits in the holoenzyme complex, but not by the isolated subunits.[10,67,70] The above data suggest that quaternary interactions in the holoenzyme complex are required for full expression of catalytic activity, and provide a critical link for communication between the catalytic sites. (v) GTP binding to gpNu1 stimulates ATP hydrolysis at the gpA subunit of the holoenzyme.[92] Conversely, (vi) ATP binding to the gpA subunit of the holoenzyme *inhibits* ATP hydrolysis at gpNu1.[92]

It is thus clear that there are significant interactions between the multiple catalytic sites of terminase holoenzyme. It is likely that these interactions are central to the assembly of the packaging machinery onto viral DNA, to promoting stability of intermediate packaging complexes, and to modulating the procapsid-dependent transition to a mobile packaging machine.[93,94] It is further likely that ATP binding and hydrolysis play central roles in these processes. These concepts are incorporated into a model for terminase assembly into a DNA packaging machine, described below.

Procapsid Assembly

The λ procapsid, also known as the prohead, is an icosahedron composed primarily of gpE, the major capsid protein. One vertex contains the doughnut-shaped portal complex, which forms a hole (the portal vertex) through which DNA enters the capsid during packaging, and exits during infection. The portal also serves to nucleate capsid assembly, and it is likely that portal protein(s) are an active part of the DNA packaging motor. Procapsid assembly is a step-wise process that, in addition to gpE, requires the phage proteins gpB, gpC and gpNu3, and host groELS chaperonins.[95] The process begins with the oligomerization of gpB monomers into a preconnector, a 25 S dodecameric ring with a ~25 Å hole at its center (Fig. 7).[96] Proper assembly in vivo requires host groELS chaperonins, and perhaps gpNu3.[95] The groELS

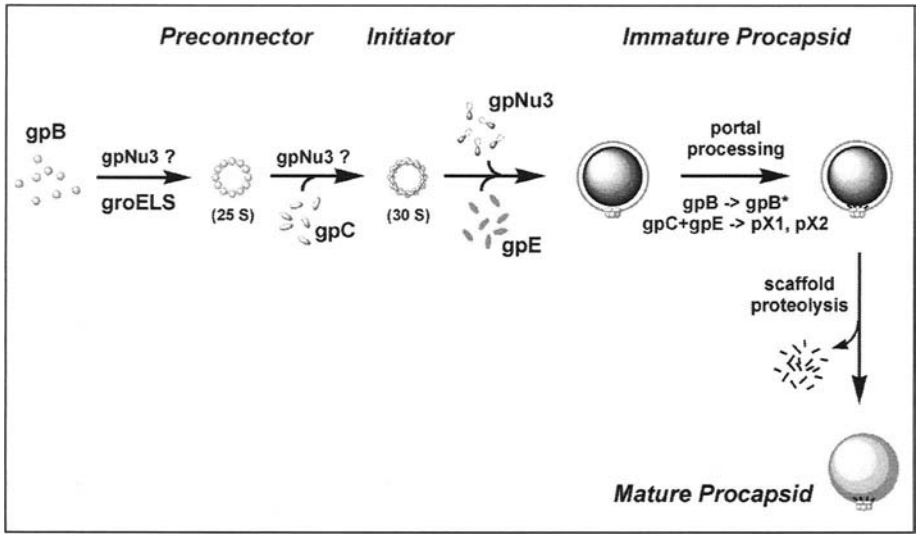


Figure 7. Procapsid assembly and processing. Details are described in the text.

proteins are presumably required for proper folding of gpB, but the role of gpNu3 in assembly of the ring is not clear. The phage-encoded gpC protein next adds to the preconnector to yield a 30S initiator structure; this reaction may also involve gpNu3.[95] Neither the stoichiometry nor the structure of gpC in the initiator complex is known.

The initiator serves to nucleate copolymerization of gpNu3 and gpE, which yields an immature procapsid (Fig. 7). GpNu3 serves as a typical scaffolding protein, directing the polymerization of gpE into an icosahedral shell structure.[95] Thus, the gpB/gpC initiator forms a hole, or portal to the capsid interior, which resides at a unique vertex of the procapsid (the portal vertex); the remaining 11 vertexes of the procapsid are constructed of gpE pentamers.

Processing of the λ Portal

The λ portal, unlike the simple portals found in other viruses, is quite complex both in terms of protein composition and protein modification. First, "processing" of the initiator involves proteolysis of gpB to yield gpB*, a protein in which the N-terminal 20 residues of gpB have been deleted.[97] The molecular weight of gpB is 59.4 kDa (based on sequence) and proteolytic digestion of the N-terminal 20 residues would yield a protein with mass 57.2 kDa. The mass of gpB* is 53-56 kDa (based on SDS-PAGE),[97,98,99] suggesting that additional residues are removed from the C-terminus of gpB and/or that gpB* migrates anomalously on SDS-PAGE.

A second step in processing of the λ portal is formation of the pX1 and pX2 proteins (referred to collectively as pX proteins). Remarkably, these proteins appear to derive from an uncharacterized covalent fusion product of gpC and gpE proteins.[100,101] A transient intermediate in portal processing (the pY protein, ≈87 kDa) has also been described. This intermediate has a molecular weight consistent with a direct fusion of gpC and gpE;[102] presumably, proteolysis of pY yields both pX1 (31 kDa) and pX2 (29 kDa). The chemical nature of these proteins and the mechanism of their formation remain completely unknown. It is noteworthy that the gpC protein shares sequence identity with the S49 family of serine proteases,[103] and we postulate that gpC is the viral protease responsible for the formation of both gpB* and the pX proteins. The temporal relationship between gpB proteolysis, cross-linking of gpC and gpE, proteolysis of pY to yield pX1 and pX2, and gpE assembly into a procapsid is unclear. Moreover, the roles of gpB*, pX1, and pX2 in the structure and function of the portal complex remains unknown.

The final step in procapsid maturation is proteolysis of the gpNu3 scaffolding protein and exit of the products from the procapsid (Fig. 7). The mature procapsid contains 12 copies of gpB*, ≈6 copies each of pX1 and pX2, and ≈420 copies of gpE.[104]

A Working Model for Lambda DNA Packaging

As with most complex biological processes, viral genome packaging can be separated into initiation, propagation, and termination events. The initiation of DNA packaging in λ is considered to include (i) assembly of the packaging machinery at a *cos* site in the concatemer, (ii) duplex nicking at *cosN*, and (iii) strand separation to yield complex I (Fig. 8). DNA translocation into the capsid defines propagation, which includes those steps that (i) promote the transition from the exceptionally stable complex I to a mobile packaging motor, and (ii) active translocation of DNA into the capsid (Fig. 9). Finally, termination includes those events responsible for (i) recognition of the terminal *cos* sequence by the translocating complex, (ii) duplex nicking and strand separation to release the DNA-filled capsid from the terminase-concatemer complex, and (iii) addition of accessory proteins and a tail to yield an infectious virus (Fig. 11). Each of these processes is discussed in turn below.

Initiation of Packaging

Assembly of the Packaging Machinery

gpA and gpNu1 cooperatively assemble at *cosN* and *cosB*, respectively. This involves binding of a gpNu1 dimer to the R3 and R2 elements of *cosB*, supported by bending of the duplex via IHF interactions with I1 (shown in Fig. 4). As discussed above, the role of R1 in this initial assembly is unclear. Interestingly, however, model-building studies show that the R1 element of *cosB* is juxtaposed to the *cosNR* half site in the bent duplex (C.E. Catalano, unpublished). Whatever the case, gpNu1 bound at *cosB* serves to anchor a symmetric gpA assembly at the *cosN* half-sites yielding a prenicking complex (Fig. 8). We presume that a gpA dimer assembles at *cosN*, but the structural details and stoichiometry of the subunits assembled at *cos* remain speculative. The presence of a bZIP protein dimerization motif in gpA suggests that an even number of subunits is involved, however.

ATP modulates the assembly and stability of the terminase subunits bound at *cos* as follows. First, ATP increases the affinity of gpNu1 and the holoenzyme for *cos* DNA (but does not affect gpA binding).[32,82] Second, ATP affects the DNase protection pattern of terminase assembled at *cos*, yielding an "extensive footprint" that extends from *cosQ* through the R-elements of *cosB*.[105] Once assembled, the complex is extraordinarily stable, with a half-life of greater

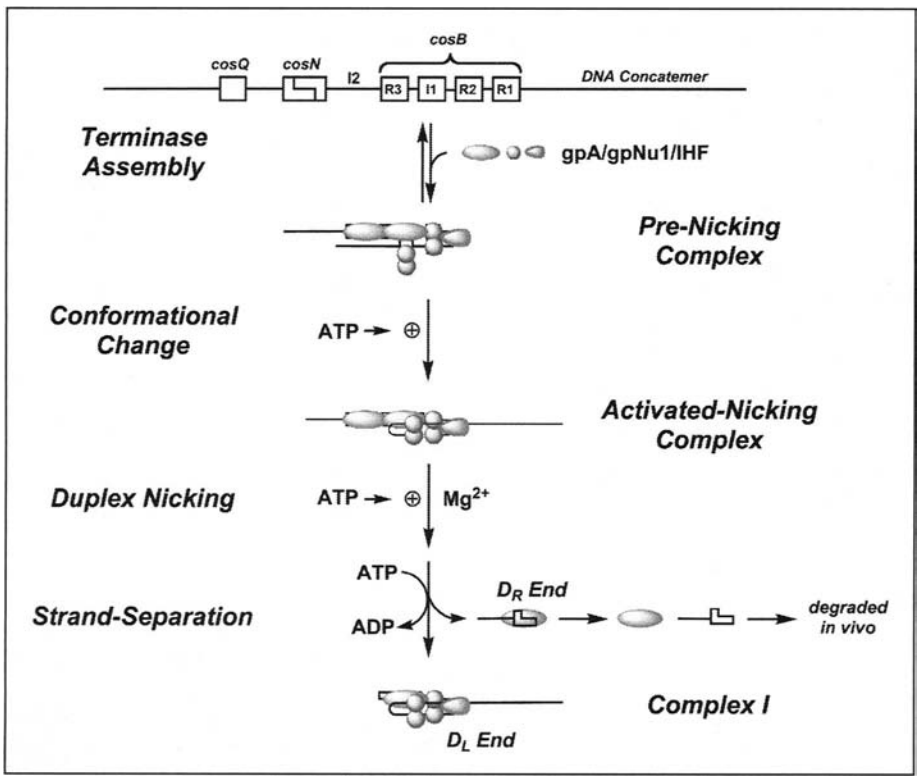


Figure 8. Initiation of DNA Packaging. Initiation steps include terminase assembly at a *cos* site in the concatemer and maturation of the D_L end. The terminase gpA and gpNu1 subunits are depicted as red ovals and blue spheres, respectively. IHF is shown as a purple lobe. A color version of this figure is available online at http://www.Eurekah.com.

than 12 hours.[32] Finally, ATP stimulates the rate and fidelity of the nuclease complex.[6,82] In aggregate, the data are consistent with an ATP-induced reorganization of the proteins assembled in the prenicking complex to yield an activated nicking complex, perhaps inducing wrapping of the DNA by terminase (Fig. 8).

Duplex Nicking and Strand Separation

In the presence of Mg^{2+}, the gpA subunit(s) introduce symmetric nicks into the duplex at *cosN*. Separation of strands by the so-called "helicase" activity of gpA leads to release of the right cohesive end (D_R) from the complex (Fig. 8).[32] It is clear that this reaction requires ATP hydrolysis, but an increase in the rate of ATP hydrolysis during strand separation has not been detected. The products of the helicase reaction have not been characterized in detail and it is interesting to consider the fate of the terminase proteins bound at *cos* upon strand separation. *A priori*, there is no reason to expect that the structure or stoichiometry of IHF and gpNu1 bound to *cosB* would change upon strand separation. Conversely, it is feasible that separation of the cohesive ends requires remodeling of gpA assembled at the two *cosN* half-sites. The two obvious outcomes are that (i) gpA bound to the cosNL half-site separates from the complex along with the D_R fragment (as shown in Fig. 8), or that (ii) gpA initially bound to the cosNL half-site is retained in the complex after strand separation. While the structural details of this transition must await further biochemical analysis, we propose that strand separation ultimately yields two products in which terminase proteins remain bound at each chromosomal end. This is discussed further below.

What Is Complex I?

This stable intermediate was originally described in vivo, and was defined as terminase bound to concatemeric DNA at uncut *cos* sites.[89,106] This definition was based on the properties of the complex that was formed by incubation of terminase with concatemeric DNA and analyzed using sucrose gradients. That is, the terminase•DNA complex migrated at a rate similar to that of the input concatemeric DNA, and at a position distinct from slower sedimenting monomeric virion DNA. Based on these data, it was concluded that the so-called "complex I" was composed of terminase bound to concatemeric DNA, and it was presumed that the *cos*-sites were not yet cut; these authors proposed that *cos* cleavage occurred subsequent to procapsid binding.[89,106] We suggest that these early sedimentation experiments could not discriminate between cut and uncut DNA. We further propose that "complex I" formed in vivo is in fact identical to the stable complex that we have characterized in vitro; specifically, terminase tightly bound to the D_L fragment formed by duplex nicking and strand separation, as shown in Figure 8.[32] This posit is supported by the following lines of evidence. (i) The *cos*-cleavage reaction is stoichiometric with respect to terminase, and the reaction stalls after a single round of duplex nicking and strand separation; catalytic turnover by the enzyme is not observed.[32] In the absence of packaging, the *cos*-cleavage reaction is thus limited by the concentration of terminase present, both in vitro and in vivo. (ii) Terminase is poorly expressed during lytic infection,[107] and it is likely that the in vivo concentration of terminase is low relative to the number of *cos* sites in concatemeric DNA. Stoichiometric *cos*-cleavage would thus result in only a limited number of *cos*-sites cleaved by the enzyme, assuming that packaging were interrupted in vivo. (iii) Terminase enzymes with mutations in the procapsid-binding domain of gpA have been constructed. The mutant enzymes fail to package DNA, presumably a result of attenuated procapsid binding interactions.[21,65,66,108] The mutations are specific for packaging and do not affect the *cos*-cleavage activity of the enzyme in vitro. Nevertheless, the extent of *cos*-cleavage by these packaging-defective terminases in vivo is reduced to roughly one third the wild type level. This discrepancy may be explained as follows. Processive packaging by terminase results in the sequential packaging of ≈3 genomes per DNA binding event.[109] Thus, three *cos* sites, on aver-

age, are cut in a concatemer once the enzyme has initiated packaging in a wild-type infection. The 30% decrease in *cos*-cleavage activity observed with the packaging-defective terminases in vivo closely matches the value expected from an enzyme that can efficiently cut the duplex, but that cannot not processively package concatemeric DNA.

The ensemble of data fully support a model where stoichiometric duplex nicking and strand separation by terminase in vivo yields a stable complex in a manner identical to what is observed in vitro; we refer to this intermediate as complex I. It is likely that the stability of this complex is required to protect the single-stranded left cohesive end (D_L) from nuclease damage in vivo, though this has not been directly demonstrated. Conversely, the D_R complex is relatively unstable, and dissociation of the terminase protein(s) leads to degradation of the right DNA end by host nucleases (Fig. 8), including the RecBCD nuclease.[110]

Transition to a Packaging Machine

cos Clearance

One of the most fascinating and yet ill-characterized aspects of the λ packaging pathway is the transition from the stable complex I to a highly mobile packaging machine that translocates DNA into the capsid. This transition is analogous to promoter clearance by RNA polymerase enzymes, a prelude to active RNA synthesis.[111] In the λ packaging pathway, "*cos* clearance" requires activation of the packaging ATPase in gpA, and a switch from a terminase complex specifically-bound to *cos* to a motor complex that binds tightly, but nonspecifically to DNA. This series of events is likely initiated by interactions of complex I with the portal proteins in the procapsid, and is also modulated by the phage gpFI protein, as discussed in detail below.

Interaction between the C-terminal 32 residues of gpA and the portal gpB* protein has been demonstrated, and it is likely that these interactions play a direct role in "docking" terminase to the procapsid.[21,38,112] Additional interactions between gpA and gpE on the procapsid surface may also be involved prior to the docking of complex I at the portal vertex.[113] Whatever the case, these initial interactions lead to a pretranslocation complex, which we propose is related to complex II previously described in vivo (Fig. 9).[89] The structural details of the complex remain unknown; however, it is likely that this interaction is mediated by gpA in complex I and gpB* in the portal complex (*vide infra*). The next step in the packaging pathway is the transition to a translocation complex, a reaction that is also mediated in an ill-characterized manner by the phage gpFI protein.

The Roles of gpFI and Procapsids in cos Clearance in Vivo

The gpFI protein is abundantly produced during the λ lytic cycle, but it is not a component of the virion. Mutations in the *FI* gene are lethal to the virus. *FI* mutant phage produce normal amounts of viral DNA, but the concatemers are not processed to mature length (i.e., *cos* cleavage is not observed in vivo).[112] Interestingly, this phenotype is also observed in phages carrying mutations in any of the procapsid assembly genes, including the host genes for the GroESL chaperonin (see Fig. 7).[112,114,115] These data have been interpreted to indicate that gpFI and mature procapsids are required for efficient *cos* cleavage in vivo, and several models have been proposed in which gpFI and/or procapsid proteins directly control the *cos* cleavage reaction. One model invokes a negative regulation mechanism where unassembled capsid proteins directly inhibit the endonuclease activity of terminase; assembly of these proteins into a procapsid effectively removes the inhibitor(s) from solution, relieving the repression of *cos* cleavage.[116] The negative regulation model has been eliminated by the following experiments. Phage with amber mutations in *FI* and/or any of the procapsid genes (i.e., mutations that result in elimination of the putative regulatory proteins) are still defective in nuclease activity.[112,114,115] Indeed, deletion of the entire block of genes including *FI* and all of the procapsid genes yields

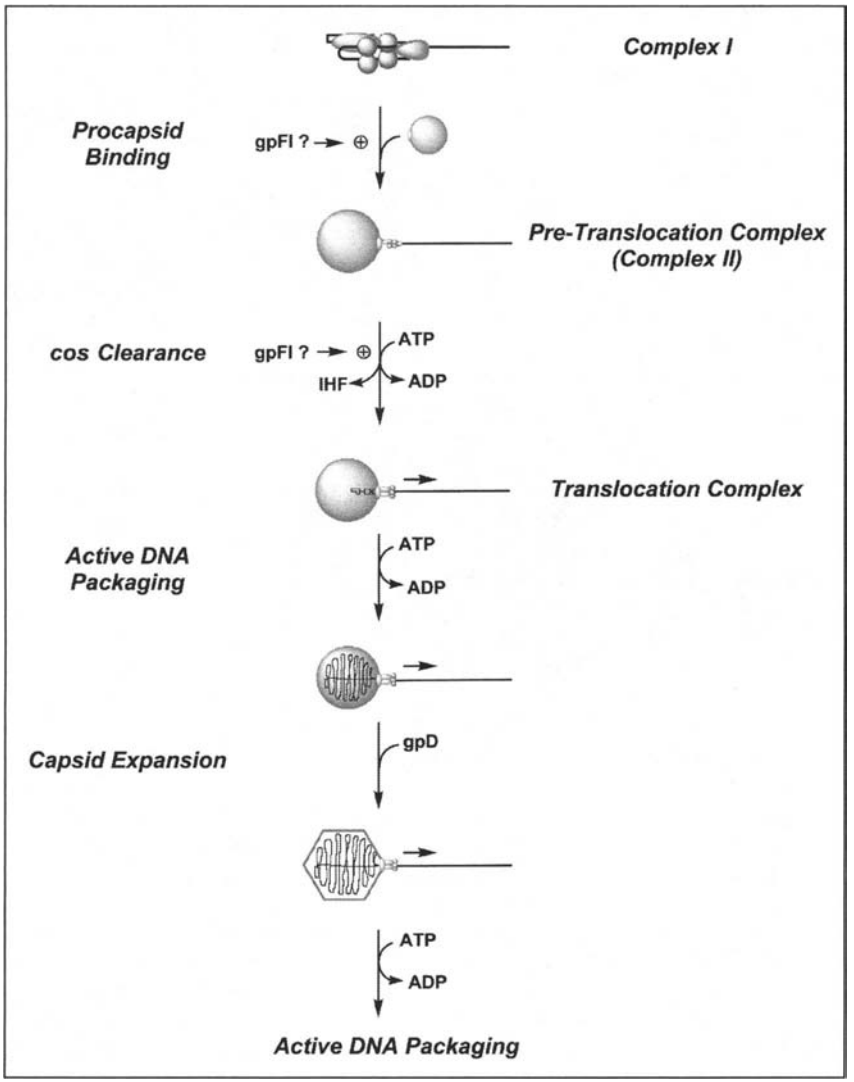


Figure 9. Translocation of DNA into the Procapsid. "Propagation" steps require *cos*-clearance, followed by active DNA packaging. The procapsid is shown as a large cyan sphere containing a portal complex (purple oval). A color version of this figure is available online at http://www.Eurekah.com.

phage that similarly possess an in vivo *cos* cleavage defect.[108] Thus, negative regulation of terminase activity by gpFI and/or procapsid precursor proteins in vivo is unlikely. A second model invokes positive regulation and proposes that gpFI and/or mature procapsids directly stimulate the endonuclease activity of terminase. This model suggests that terminase binds *cos* and is poised to nick the duplex, but requires gpFI and procapsids to stimulate the reaction. Central to this model is the presumption that complex I formed in vivo is composed of terminase bound to an *uncut cos* site in the concatemer,[89,106,117] a model which we disfavor (*vide supra*).

The in vivo *cos* cleavage results described above are in direct contrast to in vitro studies showing that terminase can efficiently cut *cos*-containing DNA in the absence of gpFI and

procapsids.[6,18,31,118,119] In fact, gpFI only modestly stimulates the *cos* cleavage reaction in vitro, presumably by promoting enzymatic turnover rather than by direct stimulation of the nicking reaction; procapsids alone or in combination with gpFI have little, if any effect.[120] Thus, neither procapsids nor gpFI affect duplex nicking by terminase in vitro, and we propose that terminase may similarly cut *cos* in the absence of these factors in vivo (see the discussion of complex I above).

A Reversibility Model for the Procapsid Requirement for cos*-Cleavage in Vivo*

If terminase does not need to be activated to cut *cos*, how might the in vivo requirement for procapsids be explained? We suggest that the in vitro results accurately reflect the activity of terminase in vivo; that is, the enzyme can efficiently assemble and nick the duplex at *cos* in the absence of procapsids. This predicts that during an infection by a mutant phage unable to produce procapsids, duplex nicking and strand separation should proceed normally. A recently proposed reversibility model accounts for the apparent procapsid dependence of terminase in vivo.[108] Here we present a variation of that model in which the *cos*-cleavage reaction yields an intermediate where the nicked-annealed duplex is bound by terminase in a strand separation complex ($T•D_R•D_L$, Fig. 10). Physical separation of the strands requires binding of the 12 base single-stranded D_L and D_R ends by individual gpA subunits in the complex, a process that is driven by an ATP-dependent conformational change. We propose that the strand separation step is reversible within this nucleoprotein complex, and that the nicked, annealed duplex may dissociate from the enzyme if packaging does not proceed in timely fashion. In the absence of procapsids in vivo, the packaging pathway stalls at the strand-separation complex and the nicked duplex is slowly released from the enzyme (Fig. 10). Repair of the nicks by host ligase results in an apparent lack of *cos*-cleavage in vivo. We suggest that procapsid binding to $T•D_L$ in the strand separation complex draws the intermediates towards packaging by mass action, thus avoiding significant dissociation of nicked DNA from the complex. This model may also explain why complex I ($T•D_L$) is efficiently formed in vitro in the absence of procapsids. A key point is that elevated concentrations of terminase are utilized in the in vitro assays. This would have the effect of driving the reaction forward by mass action, yielding complex I.

Reversibility of the *cos* cleavage reaction has been previously invoked to explain related phenomena. For instance, certain lethal mutations in the R sequences of *cosB* were found to completely block *cos* cleavage in vivo, and yet only mildly affect *cos* cleavage in vitro.[62] To resolve the discrepancy, it was proposed that the *cosB* mutations primarily affect the formation and/or stability of a post-cleavage intermediate; failure to form the intermediate leads to DNA dissociation from terminase, reannealing of the cohesive ends, and religation of the nicked duplex in vivo.[62] The post-cleavage intermediate is presumably the strand-separation complex shown in Figure 10. To further emphasize this point, a series of suppressor mutations to the *cosB* mutant phage were isolated and characterized; all of the pseudorevertants contained mutations in gpNu1.[63] Though it was anticipated that the mutant gpNu1 terminase enzymes would be more efficient in utilizing the mutant *cosB* DNA as a substrate in the *cos*-cleavage reaction, this was not the case. The mutant enzymes were no more efficient than wild type terminase at cutting *cosB* mutant DNA; rather, the mutant enzymes were more efficient than the wild type enzyme in packaging of the *cosB* mutant DNA. Here again it was argued that the *Nu1* mutations increased the efficiency of a post-cleavage transition that was made inefficient by the *cosB* mutations.[121]

The proposed reversibility of *cos* cleavage in vivo depends on ligation to reseal the nicks of annealed cohesive ends. It is interesting that in phage P2, which also has cohesive ends, efficient production of cohesive ends is similarly procapsid-dependent.[122] In contrast are phages P22 and SPP1 which use a head-full packaging mechanism and do not produce cohesive ends.

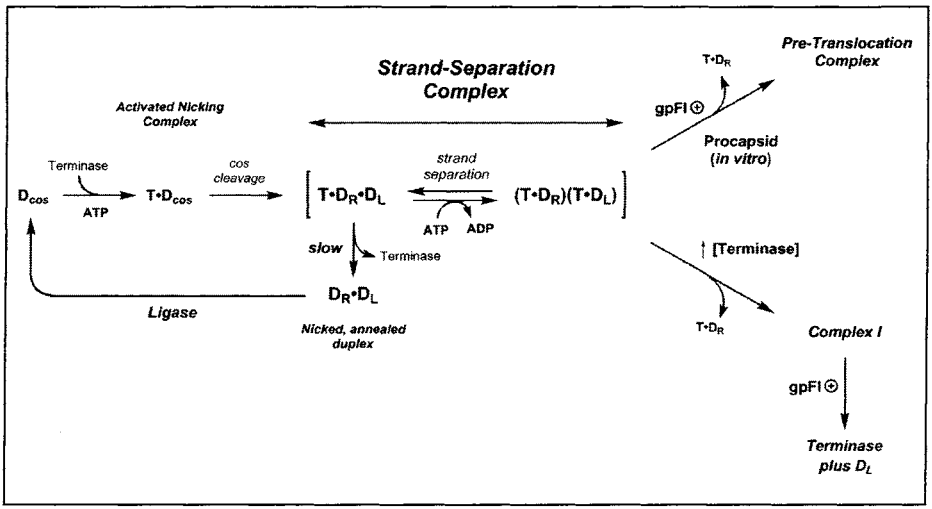


Figure 10. Model for gpFI and Procapsids in the *cos*-Cleavage Reaction. Details are described in the text.

For concatemer processing by these phages, an initial cut is made at a defined site called *pac*. Following *pac* cleavage to initiate packaging, downstream cutting occurs nonspecifically upon filling the procapsid with DNA; the resulting virion chromosomes lack cohesive ends. For both P22 and SPP1, *pac* site cutting is procapsid-*independent*.[123,124] Thus, based on this limited sample, there is a correlation that only in phages with cohesive ends is concatemer cutting procapsid dependent. DNA cutting may be irreversible when the concatemer cleavage products lack cohesive ends.

Model for gpFI Requirement in Vivo

As described above, phage with mutations in the *FI* gene also exhibit an apparent lack of *cos*-cleavage activity in vivo. The gpFI protein stimulates virion assembly in vitro, specifically when terminase and procapsids are at dilute concentrations.[117,125,126] The authors suggested that gpFI increases the efficiency of procapsid binding to the terminase proteins assembled at *cos*.[90,117] We have come to a related conclusion by examining the effect of gpFI on the *cos*-cleavage reaction in vitro, and have suggested that gpFI promotes *cos*-clearance.[120] In this model, procapsids are necessary, but not sufficient to promote *cos*-clearance and gpFI stimulates the process, as shown in Figure 10.

While the mechanism for gpFI action remains obscure, genetic studies provide insight into the nature of the protein-protein interactions between terminase, gpFI and procapsids. λFI_{amber} mutants are "leaky", which means that the virus yield is 10^{-3} compared to wild type virus. This is in contrast to the $\approx 10^{-5}$ reduction caused by nonsense mutations that inactivate a gene encoding a virion structural component. Bypass mutations, called fin mutations, allow λFI^- to form plaques in the absence of gpFI. Two types of fin mutations have been described. The first are finA mutants that have changes affecting terminase, and generally cause the production of increased levels of the gpA protein.[112,113] An increase in the concentration of gpA is expected to increase the in vivo concentration of terminase, which may increase the flux towards packaging, independent of gpFI (see Fig. 10). This is similar to the mechanism proposed above for the formation of complex I in vitro. The second are finB mutations that alter a 26-residue-long segment of gpE, which has been termed the Efi domain.[112] It has been

proposed that the Efi domain is located on the surface of the procapsid and that complex I directly interacts with the Efi domain *en route* to docking at the portal. In this model, gpFI directly modulates this initial interaction and guides the procapsid-terminase docking step.[113] An alternative model suggests that the Efi domain is, in fact, part of the gpE protein retained in the portal pX proteins (see Fig. 7). In this model, the Efi domain represents a portal docking site that makes direct contacts with terminase in complex I. The finB mutations presumably allow this interaction in the absence of gpFI.

The Fate of IHF and gpFI during cos-Clearance

One final point to consider is the fate of IHF and gpNu1 assembled at *cos* during *cos*-clearance. Both of these proteins bind to specific sites in *cosB*, and are likely central to the stability of complex I. It thus clear that a major alteration in protein-DNA binding interactions must take place prior to translocation. IHF most likely dissociates from complex I to allow passage of gpA and the procapsid, as shown in Figure 9. Indeed, based on the effect of gpFI on the *cos*-cleavage reaction in vitro, we have suggested that gpFI may act antagonistically to IHF in modulating the stability of complex I.[120] The fate of gpNu1 is less clear. One possibility is that gpNu1 switches from a site-specific DNA binding protein to a nonspecific DNA binding protein that is an active part of the packaging motor (shown in Fig. 9). This model is consistent with the observation that small terminase subunits from bacteriophage T4 and SPP1 form oligomeric rings in solution, which may indicate their role as a "sliding clamp" during translocation.[127,128] It is also feasible that gpNu1 releases DNA, but remains bound to gpA and simply "goes along for the ride". A final possibility is that gpNu1 is ejected from the translocating complex. This model suggests that the role of gpNu1 is to site-specifically assemble gpA at *cosN*, and maintain the integrity of complex I until the procapsid arrives. Thus, gpNu1 plays a role analogous to that of sigma factors and transcription factors in the assembly of the transcription complex. Once the appropriate nucleoprotein complex has been assembled, movement of the protein machinery results in ejection of the assembly protein(s) from the DNA. Again, the answers to these questions must await further experimentation.

DNA Translocation: Active DNA Packaging

The Translocation Complex

Subsequent to *cos*-clearance, the packaging motor translocates DNA into the confines of the procapsid interior, a reaction fueled by the hydrolysis of ATP. The components and stoichiometry of the λ translocation machine remain unknown, but we speculate that a hexameric gpA ring is arrayed in a polar fashion in direct contact with the gpB* dodecamer of the portal complex. A hexameric gpA structure is based on a variety of data; (i) studies in phages T3 and ϕ29 have demonstrated the presence of 6 copies of the respective large terminase subunit in the packaging complex,[71,129] (ii) electron microscopy has shown that the small terminase subunits of T4 and SPP1 terminase enzymes form oligomeric ring structures,[127,128] and (iii) recent electron micrographs of λ holoenzyme clearly demonstrate that the purified enzyme forms rings in solution (E. Bogner, unpublished). A hexameric gpA ring is pleasing in that it accommodates symmetric interactions with a dodecameric portal complex, and shows mechanistic similarity to the translocating hexameric helicase enzymes.[86,88] It is noteworthy that a gpA hexamer in the packaging complex has important implications for the nature of the activated nuclease complex described above. It has been presumed that a symmetrically disposed gpA dimer is responsible for duplex nicking. This would require recruitment of additional gpA subunits to complex I prior to translocation in order to complete the gpA hexamer. Alternatively, a symmetrically disposed gpA *hexamer* may be the relevant nuclease complex. The answers to this dilemma must await further experimentation.

The Packaging ATPase

It is presumed that ATP hydrolysis provides the energy required for inserting DNA into the confines of the capsid, and purified terminases from a number of viruses possess ATPase activity.[130] There is general agreement that, for the tailed dsDNA phages, the ATPase center resides in their respective large terminase subunits. Consistently, sequence analysis of terminase enzymes from phage to the herpesviruses indicates the presence of conserved ATPase motifs in these proteins.[69,78] In λ, a number of studies link the ATPase site located in the N-terminus of gpA to DNA translocation (summarized above).[74,75,77]

Active DNA packaging in the phage T3 and ϕ29 systems consumes about one ATP per two base pairs packaged. The observed stoichiometry in the λ system is two ATP's consumed per base pair packaged, though this may represent a significant overestimation.[61] If we take the conservative estimate, packaging of a λ genome (48,502 bp) would require the hydrolysis of ≈24,250 moles of ATP. Packaging requires 2-3 minutes in vivo,[14,60] so the estimated ATPase rate would be ≈10,000 min^{-1} per packaging motor. This is significantly greater than the basal rate of 50 min^{-1} for gpA in the holoenzyme.[60,70] Recent studies have demonstrated a packaging-specific ATPase activity that hydrolyzes ATP at a rate of 600 min^{-1}.[61] While this value is still lower than expected, it is clear that assembly of the translocation complex must activate a translocation ATPase catalytic site, presumably the N-terminal ATPase site in gpA; however, present data do not rigorously exclude the presence of a cryptic ATPase site that is located elsewhere in terminase holoenzyme, in the portal complex, or that is formed by interaction of the two structures.

Models for DNA Translocation

Once released from the *cos* site, the packaging machinery translocates along the duplex, actively inserting DNA into the capsid. How might such a translocation machine work? The goal of understanding a biological motor at the molecular level challenges us to mechanistically link the energy of ATP hydrolysis to physical changes in protein structure that lead to translocation. Significant progress has been made towards our understanding of a number of motors, including myosins, kinesins, and the rotary F_1F_0 and flagellar motors.[131] Viral DNA packaging motors have been particularly difficult to dissect at the molecular level, however, for the following reasons. (i) With very few exceptions, a complete list of the components of the motors remains speculative. It is likely that both terminase proteins and portal proteins assemble to complete a translocating motor, but the exact nature of the complex is unknown in all cases. (ii) Once assembled, the motors act quickly and transiently. Dissociation of the components occurs upon completion of the packaging process. (iii) There is little structural information available for any of the components of these machines, and none on the actively packaging motors.

Nevertheless, it is clear that there are three parts to the motor: terminase, the portal vertex, and the procapsid shell. There are a number of creative models describing how the components of the packaging complex sponsor DNA movement. The first proposes that the terminase enzyme is directly responsible for translocation of DNA into the capsid. These models propose that the terminase subunits physically translocate via flexible, DNA-contacting domains that cyclically contract, and then undock from the DNA.[94,130,132] The conformational changes required in this model are presumably driven by the hydrolysis of ATP. These models find mechanistic similarity to the "inch worm" mechanism proposed for translocation of hexameric helicases.[86,87] A second class of models is portal-centric, which rely on the concept of a symmetry mismatch between the twelve-fold symmetry of the portal and the five-fold symmetry of the capsid shell vertex to which it is attached. In one version, which is actually the earliest model, ATP hydrolysis drives rotation of the portal protein with respect to the capsid shell, which "screws" the DNA into the capsid.[133] A more recent portal model is inspired by the

structure of the ϕ29 portal (Anderson and Grimes, Chapter 7 of this work). This model proposes that the procapsid-terminase complex (which possesses ATPase activity) acts as a stator, the DNA as a spindle, and the portal complex as a ball-race. The DNA helix is proposed to convert the rotary action of the portal into translocation of the DNA.[134] In the third class of models, the capsid shell acts as a reciprocating gated pump that pulls the DNA into the procapsid interior[135] (Serwer, Chapter 4 of this work).

For the λ system, available evidence is most consistent with the first class of models in which terminase subunits constitute the DNA translocating motor, though other models are not rigorously excluded. The strongest evidence comes from genetic studies in which lethal mutations in the virus were screened to identify those that affected DNA packaging.[74] Ten of these mutants were characterized, and all had mutations that mapped to a gene *A* segment extending from codons 18 to 349 (summarized above). These mutants provide evidence for the involvement of the terminase gpA subunit in translocation and active DNA packaging. A detailed biochemical analysis of these mutants in vitro will undoubtedly provide further insight into the role of gpA in the packaging motor.

Capsid Expansion and the Phage gpD Protein

Translocation of the packaging machinery along the duplex results in packaging DNA into the procapsid. Upon packaging of 10-50% of the viral genome, the procapsid undergoes an expansion process whereby the 25 nm spherical capsid expands to a radius of 32 nm, roughly doubling its volume and acquiring its mature icosahedral shape.[136,137] The molecular basis for expansion is not known. The viral gpD protein, which is a monomer in solution, adds to the surface of the expanded icosahedral capsid as a trimer, and in numbers similar to the major capsid protein (Fig. 9).[138,139] The presumed role of gpD is to stabilize the partially filled capsid, and prevent DNA release.[136,138-140] Indeed, D^- phages are viable if the viral chromosome is less than 80% of wild type chromosome length, and if the virions are kept in medium supplemented with 10^{-2} M Mg^{++}.[140] Within this context, an interesting chimeric phage is λ *shp*, a phage in which the *D* gene has been replaced by the analogous *shp* gene of phage 21. The Shp protein has a lower affinity than gpD for the λ gpE shell, so that 15-20% of the binding sites on λ *shp* capsid shells are unfilled. λ *shp* has a stability intermediate between λ wild type and λ D^-, such that λ *shp* virions with full-length chromosomes require supplemental Mg^{++} ions (10^{-2} M) for viability. The intermediate stability of λ *shp* virions indicates that the *shp* protein imparts considerable stability to the capsid shell, even though the shell is not completely populated by Shp.[141]

Termination of Packaging

Genome packaging is terminated when the packaging motor encounters the next downstream *cos* in the concatemer. When terminase encounters this site, translocation stops and symmetric nicks are introduced into the *cosN* subsite (Fig. 11); this terminal *cos* cleavage reaction is less dependent on *cosB*, but requires that the *cosQ* and I2 subsites be present for efficient termination (described in detail below). The strand separation activity of terminase separates the cohesive ends, releasing the DNA-filled capsid from the concatemer. Terminase remains bound to the matured (D_L) end of the concatemer, which represents the next λ chromosome to be packaged. The complex, which is functionally-related to complex I, can bind a new procapsid and initiate packaging of the second chromosome in the concatemer. Terminase-mediated genome packaging is thus processive, and it has been estimated that 2-3 genomes are packaged per DNA binding event.[109]

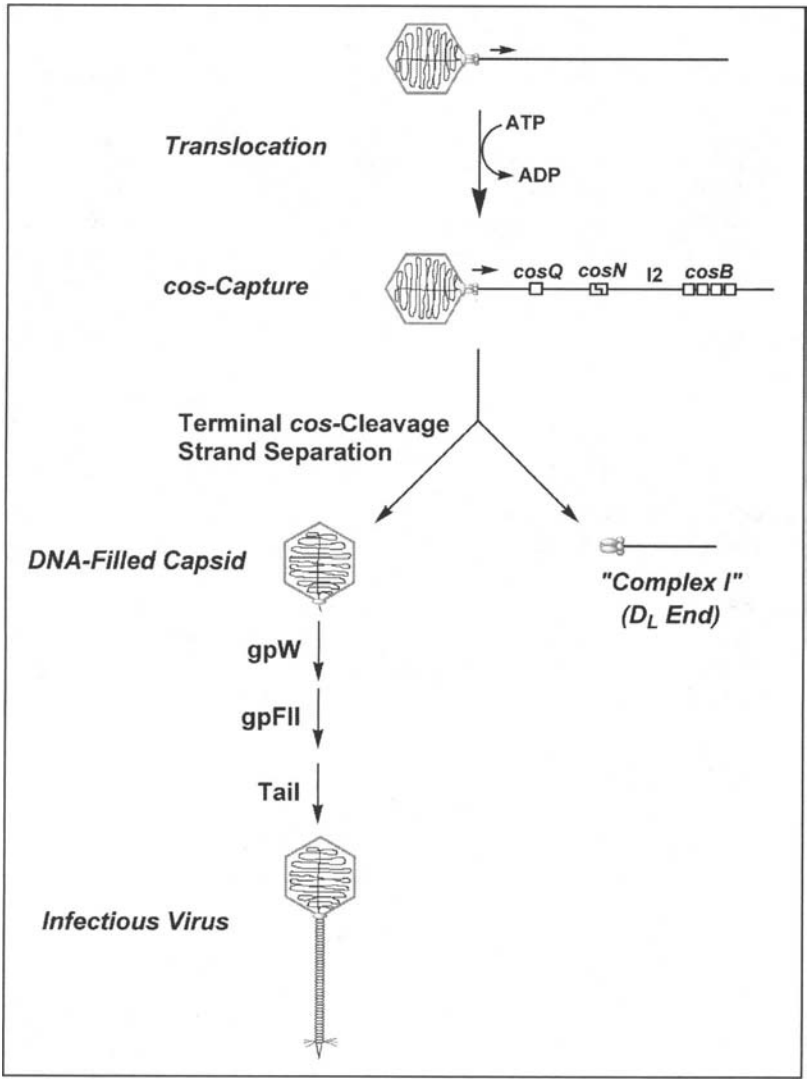


Figure 11. Termination of DNA Packaging. Termination requires recognition of *cosN* by the translocating complex. The *cosQ-cosN*-I2 segment is central to capture of the packaging machinery. Duplex nicking and strand separation complete DNA packaging. Addition of gpW, gpFII and a tail complete the assembly of an infectious virus. Strand separation also releases terminase bound to the D_L end of the concatemer, which is utilized for a second round of packaging.

cosQ and Recognition of the Terminal cos Sequence

Termination is a complex process that likely involves significant modification of the translocation complex. Proper termination requires that the downstream *cos* contain *cosQ*, *cosN* and the I2 segment.[11] When *cosQ* is mutant, there is a failure to nick the bottom strand of *cosN*, and translocation does not arrest. Rather, packaging continues so that the singly-nicked duplex is

further inserted until the capsid is filled to capacity; this is a lethal event because the protruding, uncut DNA prevents virus maturation (*vide infra*).[26] Our interpretation of the bottom strand nicking failure of λ *cosQ* is based on symmetry issues, as follows. We presume that the gpA oligomer (hexamer?) responsible for translocation in the packaging motor is arranged in a unidirectional manner through interactions between the C-termini of gpA subunits and the gpB* portal protein.[24,26] Conversely, the symmetry of *cosN* suggests that symmetrically arranged (head-to-head) gpA subunits are required to nick each *cosN* half-site.[24,26] Thus, reorganization of the unidirectional gpA complex in the packaging motor is required to assemble the symmetric nicking complex required to cut the terminal *cos* site. We propose that *cosQ* plays a direct role in this reorganization, and is required to sponsor the presentation of a properly oriented terminase complex to the *cosN* half-sites. The mechanism for this is obscure, but could involve a ***bone fide*** reorganization of the proteins in the translocation complex, and/or recruitment of additional proteins from solution.[24,26]

The packaging protein that binds *cosQ* has not been identified, despite extensive studies of pseudorevertants of *cosQ* mutants. While not revealing the *cosQ* binding protein, studies of nonallele specific *cosQ* suppressors indicate a relationship between the extent of DNA packaged and the efficiency of recognition of mutant *cosQ* sites, as follows. Mild *cosQ* mutations can be suppressed by (i) lengthening the chromosome, and/or (ii) missense mutations altering the portal protein, gpB.[22,23] It is argued that chromosome lengthening or altering gpB slows the translocation rate at the time the downstream *cos* is encountered by the packaging motor; this may increase the efficiency of recognition and cleavage of a mutant downstream *cos* site. While there is no direct information about the relationship between chromosome length and the rate of λ DNA translocation, data obtained in the phage ϕ29 system indicates that translocation rate declines dramatically as the end of the chromosome is approached.[142]

The Extent-of-Packaging Sensor

The efficiency with which a wild type downstream *cos* is cut to terminate packaging also depends on the length of DNA to be packaged. When the genome is shortened to lengths <80% of wild type, *cos* cleavage efficiency in vivo declines sharply.[22] The *cos*-cleavage reaction is insensitive to duplex length in vitro,[10] suggesting that the terminal *cos* cleavage reaction in λ is affected by the amount of DNA packaged into the capsid. This is similar to phages such as T4, P22 and SPP1 that use a head-full activated DNA cleavage mechanism for chromosome processing (Chapters 3, 5 and 6, respectively, of this work). Genetic observations in the phage P22 and SPP1 systems have lead these researchers to propose that the packaging motor has a "sensor" that detects when the capsid is full of DNA, which then activates the endonuclease catalytic site.[143,144] This putative sensor might detect (i) the extent of DNA packaged, (ii) the rate of DNA translocation, and/or (iii) the energy required for translocation. For λ terminal *cos* cleavage, we propose that the efficiency of *cosQ* recognition, and hence cleavage, is inversely dependent on translocation rate. Thus, *cosQ* recognition is proposed to represent the sensor.

Processive Genome Packaging

Following the downstream *cos* cleavage event, the terminase proteins dissociate from the portal complex of the DNA-filled capsid, but remain bound to the matured (D_L) end of the next genome in the concatemer (Fig. 11). This terminase-DNA complex, which is functionally related to complex I, captures a procapsid to sponsor the packaging of the next chromosome in the concatemer. While the *cosB* subsite is not essential for the terminal *cos* cleavage reaction, it is required for processive genome packaging. Surprisingly, however, some mutations of *cosB* that are defective in the initial *cos* cleavage reaction retain a processive packaging phenotype; these include a triple point mutant *cosB* $R3^-$ $R2^-$ $R1^-$ allele, and *cosB*ϕ21, a λ genome with the *cosB* of phage 21.[16] These data indicate that the terminase-DNA interactions required for

processive packaging are "relaxed" in comparison with the requirements for the initial assembly of terminase at *cos* (see packaging initiation, above). Structural studies on these complexes are needed to understand the DNA site requirements for each of the packaging complexes.

Virion Completion

Subsequent to separation of the DNA-filled capsid from the concatemer, the minor capsid proteins gpW and gpFII are sequentially added to the portal (Fig. 11). Presumably, gpW and gpFII are required to prevent DNA loss from the filled capsid[139] and to provide an attachment site for tail addition, respectively.[145] A high-resolution 3D structure of gpW (68 residues) has recently been solved by NMR.[146] The solution structure represents a novel fold, consisting of two α-helixes and a two-stranded β-sheet, arranged around a hydrophobic core. The 14 C-terminal residues, which are essential to virus development, are disordered in the structure. These residues presumably contact gpB* in the portal complex, and in addition may provide contacts for subsequent gpFII addition to the capsid.

A high-resolution 3D structure of the gpFII protein (117 residues) has also been solved by NMR.[147] This structure is composed of seven β-strands and a short α-helix, with two unstructured regions extending between residues 1-24 and 46-62. GpFII has homologues in other lambda-like phages, such as 21 and φ80, and it is possible to compare sequence alignments with binding specificity. Of particular interest is the gpFII analogue from φ80, because this protein is tail specific. That is, gpFII φ80 forms infectious virions with φ80 tails, but not with λ tails.

Concluding Remarks

Assembly of an infectious λ virus starts with the products of the DNA replication and procapsid assembly pathways. DNA processing and packaging constitute major morphogenetic pathways where λ DNA is "matured" and translocated into the procapsid by terminase, assisted by gpFI and the host factors IHF and HU. Addition of the "finishing proteins" gpW and gpFII, followed by tail attachment complete the infectious virion. An ordered and essentially irreversible series of macromolecular assembly steps are required to carry out the interdependent processes of (i) cutting concatemeric DNA into unit-length virion chromosomes, (ii) packaging the chromosomes into procapsid shells and (iii) stabilization of the DNA-filled capsid and tail attachment. Since the last review in 2000, the enzymology of DNA packaging by λ terminase has been extensively studied which has led to significant insights into DNA-protein and protein-protein interactions involved in this complex process. Further, high-resolution structures of several assembly proteins, including gpD, gpW, gpFII and the DNA binding domain of gpNu1 have been determined, providing a glimpse into the structure-function relationships of these critical proteins. As an experimental system, λ is highly developed, with excellent genetics and strong biochemistry. Each of the proteins involved in λ assembly has been cloned and purified which sets the stage for a detailed characterization of virus development at the molecular level. We are thus challenged to characterize the biochemical, structural and functional aspects of each step along the developmental pathway leading to an infectious virus. These studies remain a challenging area of research, but will undoubtedly lead to significant insight into the molecular mechanisms of virus assembly.

Acknowledgements

We are thankful for a number of talented and dedicated students, post-doctoral fellows and coworkers that have contributed to much of the experimental work and ideas from our laboratories that are described in this chapter; most recently, Tonny de Beer, Alok Dhar, Carol Duffy, Helene Gaussier, Karl Maluf, Marcos Ortega, Kristen Potratz, Jean Sippy, Jennifer Wendt,

Douglas Wieczorek and Qin Yang. We further wish to acknowledge the seminal work in the Toronto labs of Andy Becker, Marvin Gold and Helios Murialdo, which forms the basis for many of the models presented in this review. Special thanks go to Alan Davidson and Karen Maxwell for ensuring that λ research continues in Toronto.

References

1. Katsura I. Tail assembly and injection. In: Hendrix RW, Roberts JW, Stahl FW et al, eds. Lambda II. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1983:331-346.
2. Hendrix RW, Roberts JW, Stahl FW et al. Lambda II. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1983.
3. Herskowitz I, Hagen D. The lysis-lysogeny decision of phage lambda: Explicit programming and responsiveness. Ann Rev Genetics 1980; 14:399-445.
4. Ptashne M, Gann A. Genes & Signals. Cold Spring Harbor: Cold Spring Harbor Press, 2001.
5. Friedman D, Gottesman M. Lytic mode of lambda development. In: Hendrix RW, Roberts JW, Stahl FW et al, eds. "Lambda II". Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1983:21-51.
6. Higgins RR, Lucko HJ, Becker A. Mechanism of cos DNA cleavage by bacteriophage lambda terminase: Multiple roles of ATP. Cell 1988; 54(6):765-75.
7. Rubinchik S, Parris W, Gold M. 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 1994; 269(18):13586-93.
8. Rubinchik S, Parris W, Gold M. The in vitro translocase activity of lambda terminase and its subunits. Kinetic and biochemical analysis. J Biol Chem 1995; 270(34):20059-66.
9. Woods L, Catalano C. Kinetic characterization of the GTPase activity of phage lambda terminase: Evidence for communication between the two "NTPase" catalytic sites of the enzyme. Biochemistry 1999; 38:4624-4630.
10. Woods L, Terpening C, Catalano CE. Kinetic analysis of the endonuclease activity of phage lambda terminase: Assembly of a catalytically competent nicking complex is rate-limiting. Biochemistry 1997; 36(19):5777-85.
11. Cue D, Feiss M. Bacteriophage λ DNA packaging: DNA site requirements for termination and processivity. J Mol Biol 2001; 311:233-240.
12. Feiss M, Kobayashi I, Widner W. Separate sites for binding and nicking of bacteriophage lambda DNA by terminase. Proc Natl Acad Sci USA 1983; 80(4):955-9.
13. Feiss M, Widner W, Miller G et al. Structure of the bacteriophage lambda cohesive end site: Location of the sites of terminase binding (cosB) and nicking (cosN). Gene 1983; 24(2-3):207-18.
14. Hohn B. DNA sequences necessary for packaging of bacteriophage λ DNA. Proc Nat Acad Sci USA 1983; 80:7456-7460.
15. Miwa T, Matsubara K. Lambda phage DNA sequences affecting the packaging process. Gene 1983; 24:199-206.
16. Cue D, Feiss M. A site required for termination of packaging of the phage lambda chromosome. Proc Natl Acad Sci USA 1993b; 90(20):9290-4.
17. Davidson A, Gold M. 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 1992; 161:305-315.
18. Cue D, Feiss M. The role of cosB, the binding site for terminase, the DNA packaging enzyme of bacteriophage lambda, in the nicking reaction. J Mol Biol 1993a; 234(3):594-609.
19. Higgins RR, Becker A. 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 1994a; 13(24):6152-61.
20. Xin W, Feiss M. 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 1993; 230:492-504.
21. Yeo A, Feiss M. Specific interaction of terminase, the DNA packaging enzyme of bacteriophage lambda, with the portal protein of the prohead. J Mol Biol 1995; 245(2):141-50.

22. Cue D, Feiss M. Genetic evidence that recognition of cosQ, the signal for termination of phage λ DNA packaging, depends on the extent of head filling. Genetics 1997; 147:7-17.
23. Wieczorek D, Didion L, Feiss M. Alterations of the portal protein of bacteriophage λ suppress mutations in cosQ, the site required for termination of DNA packaging. Submitted Genetics 2002; 161:21-31.
24. Wieczorek D, Feiss M. Defining cosQ, the site required for termination of bacteriophage lambda DNA packaging. Genetics 2001; 158:495-506.
25. Wieczorek D, Feiss M. Genetics of cosQ, the DNA packaging termination site of phage λ: A study of local suppression and methylation effects. Genetics 2003; In press.
26. Cue D, Feiss M. Termination of packaging of the bacteriophage lambda chromosome: cosQ is required for nicking the bottom strand of cosN. J Mol Biol 1998; 280(1):11-29.
27. Miller G, Feiss M. The bacteriophage lambda cohesive end site: Isolation of spacing/substitution mutations that result in dependence on Escherichia coli integration host factor. Mol Gen Genet 1988; 212(1):157-65.
28. Goodrich JA, Schwartz ML, McClure WR. Searching for and predicting the activity of sites for DNA binding proteins: Compilation and analysis of the binding sites for Escherichia coli integration host factor (IHF). Nucleic Acids Res 1990; 18:4993-5000.
29. Mendelson I, Gottesman M, Oppenheim AB. HU and integration host factor function as auxiliary proteins in cleavage of phage lambda cohesive ends by terminase. J Bact 1991; 173:1670-1676.
30. Higgins RR, Becker A. 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 1994b; 13(24):6162-71.
31. Tomka MA, Catalano CE. Physical and kinetic characterization of the DNA packaging enzyme from bacteriophage lambda. J Biol Chem 1993b; 268(5):3056-65.
32. Yang Q, Hanagan A, Catalano CE. Assembly of a nucleoprotein complex required for DNA packaging by bacteriophage lambda. Biochemistry 1997; 36(10):2744-52.
33. Shinder G, Gold M. The Nu1 subunit of bacteriophage lambda terminase binds to specific sites in cos DNA. J Virology 1988; 62:387-392.
34. Parris W, Rubinchik S, Yang YC et al. A new procedure for the purification of the bacteriophage lambda terminase enzyme and its subunits. Properties of gene product A, the large subunit. J Biol Chem 1994; 269(18):13564-74.
35. Smith MP, Feiss M. Sequence analysis of the phage 21 genes for prohead assembly and head completion. Gene 1993a; 126(1):1-7.
36. Smith MP, Feiss M. Sites and gene products involved in lambdoid phage DNA packaging. J Bacteriol 1993b; 175(8):2393-9.
37. Siegele DA, Frackman S, Sippy J et al. The head genes of bacteriophage 21. Virology 1983; 129(2):484-9.
38. Frackman S, Siegele DA, Feiss M. A functional domain of bacteriophage lambda terminase for prohead binding. J Mol Biol 1984; 180(2):283-300.
39. Frackman S, Siegele DA, Feiss M. The terminase of bacteriophage lambda. Functional domains for cosB binding and multimer assembly. J Mol Biol 1985; 183(2):225-38.
40. Wu WF, Christiansen S, Feiss M. Domains for protein-protein interactions at the N and C termini of the large subunit of bacteriophage lambda terminase. Genetics 1988; 119(3):477-84.
41. Yang Q, Beer TD, Woods L et al. Cloning, expression, and characterization of a DNA binding domain of gpNu1, a phage lambda DNA packaging protein. Biocemistry 1999a; 38:465-477.
42. Yang Q, Berton N, Manning M et al. Domain structure of gpNu1, a phage lambda DNA packaging protein. Biochemistry 1999b; 38:14238-14447.
43. de Beer T, Meyer J, Ortega M et al. Insights into specific DNA recognition during assembly of a viral genome packaging machine; structure and genetics of the DNA binding domain of gpNu1. Molecular Cell 2002; 9:981-991.
44. Bain D, Berton N, Ortega M et al. Biophysical characterization of the DNA binding domain of gpNu1, a viral DNA packaging protein. J Biol Chem 2001; 276:20175-20181.
45. Becker A. (cited in Feiss, M). Terminase and the recognition, cutting and packaging of λ chromosomes. Trends Genet 1986; 2:100-104.

46. Feiss M. Terminase and the recognition, cutting and packaging of λ chromosomes. Trends Genet 1986; 2:100-104.
47. Kypr J, Mrazek J. Lambda phage protein Nu1 contains the conserved DNA binding fold of repressors. J Mol Biol 1986; 91:139-140.
48. Clark K, Halay E, Lai E et al. Cocrystal structure of HNF-3/forkhead DNA-recognition motif resembles histone H5. Nature 1993; 364:412-420.
49. Weigel D, Jackle H. The fork head domain: A novel DNA binding motif of eukaryotic transcription factors. Cell 1990; 63:455-456.
50. Martinez-Hackert E, Stock A. Structural relationshiip in the OmpR family of winged-helix transcription factors. J Mol Biol 1997; 269:301-312.
51. Clubb R, Omichinski J, Savilahti H et al. A novel class of winged helix-turn-helix protein: The DNA-binding domain of Mu transposase. Structure 1994; 2:1041-1048.
52. Ilangovan U, Wojciak J, Connolly K et al. NMR structure and functional studies of the Mu repressor DNA binding domain. Biochemistry 1999; 38:8367-8376.
53. Wintjens R, Rooman M. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. J Mol Biol 1996; 262:294-313.
54. Gussin G, Johnson A, Pabo C et al. Repressor and cro protein: Structure, function, and role in Lysogenization. In: Hendrix RW, Roberts JW, Stahl FW et al, eds. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1983:93-120.
55. Bear S, Court D, Friedman D. An accessory role for Escherichia coli integration host factor: Characterization of a lambda mutant dependent upon integration host factor for DNA packaging. J Virol 1984; 52:966-972.
56. Kosturko L, Daub E, Murialdo H. The interaction of E. coli intergration host factor and lambda cos DNA multicomplex formation and protein-induced bending. Nucleic Acids Res 1989; 17:329-334.
57. Rice PA, Yang S, Mizuuchi K et al. Crystal structure of an IHF-DNA complex: A protein-induced DNA U-turn. Cell 1996; 87(7):1295-306.
58. Friedman D. Integration host factor: A protein for all reasons. Cell 1988; 55:545-549.
59. Xin W, Cai Z-H, Feiss M. 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 1993; 230:505-515.
60. Hwang Y, Feiss M. A defined system for in vitro lambda DNA packaging. Virology 1995; 211(2):367-76.
61. Yang Q, Catalano C. Biochemical characterization of bacteriophage lambda genome packaging in vitro. Virology 2003; 305:276-287.
62. Cue D, Feiss M. Genetic analysis of cosB, the binding site for terminase, the DNA packaging enzyme of bacteriophage lambda. J Mol Biol 1992a; 228(1):58-71.
63. Cue D, Feiss M. Genetic analysis of mutations affecting terminase, the bacteriophage lambda DNA packaging enzyme, that suppress mutations in cosB, the terminase binding site. J Mol Biol 1992b; 228(1):72-87.
64. Granston AE, Alessi DM, Eades L et al. A point mutation in the Nu1 gene of bacteriophage λ facilitates phage growth in Escherichia coli with himA and gyrB mutations. Mol Gen Genet 1988; 212:149-156.
65. Yeo A, Kosturko LD, Feiss M. Structure of the bacteriophage lambda cohesive end site: Bent DNA on both sides of the site, cosN, at which terminase introduces nicks during chromosome maturation. Virology 1990; 174(1):329-34.
66. Sippy J, Feiss M. Analysis of a mutation affecting the specificity domain for prohead binding of the bacteriophage lambda terminase. J Bacteriol 1992; 174(3):850-6.
67. Hwang Y, Catalano CE, Feiss M. Kinetic and mutational dissection of the two ATPase activities of terminase, the DNA packaging enzyme of bacteriophage λ. Biochemistry 1996; 35(8):2796-803.
68. Saraste M, Sibbald P, Wittinghofer A. The P-loop - A common motif in ATP and GTP-binding proteins. Trends in Biochemical Sciences 1990; 15:430-434.
69. Walker JE, Saraste M, Runswick MJ et al. 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 1982b; 8:945-951.

70. Tomka MA, Catalano CE. Kinetic characterization of the ATPase activity of the DNA packaging enzyme from bacteriophage lambda. Biochemistry 1993a; 32(45):11992-7.
71. Guo P, Peterson C, Anderson D. Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage ϕ29. J Mol Biol 1987; 197:229-236.
72. Hwang Y, Feiss M. 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 1996; 261(4):524-35.
73. Hwang Y, Feiss M. The endonuclease and helicase activities of Bacteriophage λ? terminase: Changing nearby residue 515 restores activity to the gpA K497D mutant enzyme. Virology 2000; 277:204-214.
74. Duffy C, Feiss M. The large subunit of bacteriophage lambda's terminase plays a role in DNA translocation and packaging termination. J Mol Biol 2002; 316:547-561.
75. Hang Q, Tack B, Feiss M. An ATPase center of bacteriophage λ terminase involved in post-cleavage stages of DNA packaging: Identification of ATP-interactive amino acids. J Mol Biol 2000; 302:777-795.
76. Pu W, Struhl K. The leucine zipper symmetrically positions the adjacent basic regions for specific DNA binding. Proc Nat Acad Sci USA 1991; 88:6901-6905.
77. Dhar A, Feiss M. Mutations in the ATP reactive center of λ terminase and its effect on DNA packaging. Unpublished observations 2003.
78. Mitchell M, Matsuzaki S, Imai S et al. 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 2002; 30:4009-4021.
79. Rao V, Mitchell M. The N-terminal ATPase site in the large terminase protein Gp17 is critically required for DNA packaging in bacteriophage T4. J Mol Biol 2001; 314:411-421.
80. Gorbalenya A, Koonin E. Curr Opin Struct Biol 1993; 3:419-429.
81. Story R, Li H, Abelson J. Crystal structure of a DEAD box protein from the hyperthermophile Methanococcus jannaschii. Proc Natl Acad Sci USA 2001; 98:14650-1470.
82. Yang Q, Catalano C. A minimal kinetic model for a viral DNA packaging machine. Biochemistry 2004; 43:289-299.
83. Hang J, Catalano C, Feiss M. The functional asymmetry of cosN, the nicking site for bacteriophage λ DNA packaging, is dependent on the terminase binding site, cosB. Biochemistry 2001; 40:13370-13377.
84. Xu SY, Feiss M. Structure of the bacteriophage lambda cohesive end site. Genetic analysis of the site (cosN) at which nicks are introduced by terminase. J Mol Biol 1991; 220(2):281-92.
85. Yang Q, Catalano CE. Kinetic characterization of the strand separation ("helicase") activity of the DNA packaging enzyme from bacteriophage λ. Biochemistry 1997; 36:10638-10645.
86. Delagoutte E, Hippel PV. Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. Quart Rev Biophys 2002; 35:431-478.
87. Lohman T. Helicase-catalyzed DNA unwinding. J Biol Chem 1993; 268:2269-2272.
88. Patel S, Picha K. Structure and function of hexameric helicases. Ann Rev Biochem 2000; 69:651-697.
89. Becker A, Murialdo H, Gold M. Studies on an in vitro system for the packaging and maturation of phage λ DNA. Virology 1977; 78:277-290.
90. Becker A, Gold M. Prediction of an ATP reactive center in the small subunit, gpNu1, of the phage lambda terminase enzyme. J Mol Biol 1988; 199(1):219-22.
91. Babbar BK, Gold M. ATP-reactive sites in the bacteriophage λ packaging protein terminase lie in the N-termini of its subunits, gpA and gpNu1. Virology 1998; 247:251-264.
92. Catalano C, Woods L. Kinetic characterization of the GTPase activity of phage lambda terminase: Evidence for communication between the two "NTPase" catalytic sites of the enzyme. Biochemistry 1999; 38:4624-4630.
93. Catalano C. The terminase enzyme from bacteriophage lambda: A DNA-packaging machine. Cell Mol Life Sci 2000; 57:128-148.
94. Catalano CE, Cue D, Feiss M. Virus DNA packaging: The strategy used by phage lambda. Mol Microbiol 1995; 16(6):1075-86, [Review] [87 refs].

95. Georgopoulos C, Tilly K, Casjens S. Lambdoid phage head assembly. In: Hendrix RW, Roberts JW, Stahl FW et al, eds. Lambda II. Cold Spring Harbor: Cold Spring Harbor Press, 1983:279-304.
96. Kochan J, Carrascosa JL, Murialdo H. Bacteriophage lambda preconnectors: Purification and structure. J Mol Biol 1984; 174:433-447.
97. Walker JE, Aufferet AD, Carne A et al. Solid-phase sequence analysis of polypeptides eluted from polyacrylamide gels: An aid to interpretation of DNA sequences as exemplified by Escherichia coli unc operon and bacteriophage lambda. Eur J Biochem 1982a; 123:23-260.
98. Hendrix R, Casjens S. Locations and amounts of major structural proteins in bacteriophage lambda. J Mol Biol 1974a; 88:535-545.
99. Hohn B. DNA as substrate for packaging into phage lambda in vitro. J Mol Biol 1975; 98:93-106.
100. Hendrix RW, Casjens SR. Protein fusion during the assembly of phage lambda heads. Journal of Supramolecular Structure 1974b; 2(2-4):329-36.
101. Hendrix RW, Casjens SR. Protein fusion: A novel reaction in bacteriophage lambda head assembly. Proc Natl Acad Sci USA 1974c; 71(4):1451-5.
102. Hohn T, Hohnl F. Petit lambda, a family of particles from coliphage lambda-infected cells. J Mol Biol 1975; 98:107-120.
103. Baird L, Lipinska B, Raina S et al. Identification of the Escherichia coli sohB gene, a multicopy suppressor of the HtrA (DegP) null phenotype. J Bacteriol 1991; 173:5763-70.
104. Hendrix RW, Casjens SR. Assembly of bacteriophage lambda heads: Protein processing and its genetic control in petit lambda assembly. J Mol Biol 1975; 91(2):187-99.
105. Higgins RR, Becker A. Interaction of terminase, the DNA packaging enzyme of phage lambda, with its cos DNA substrate. J Mol Biol 1995; 252(1):31-46.
106. Becker A, Murialdo H, Gold M. Early events in the in vitro packaging of bacteriophage DNA. Virology 1977; 78:291-305.
107. Murialdo H, Fife W. The maturation of coliphage lambda DNA in the absence of its packaging. Gene 1984; 30:183-194.
108. Sippy J, Feiss M. Initial cos cleavage of bacteriophage λ concatemers requires proheads and gpFI in vivo. Mol Microbiol 2004; In press.
109. Emmons SW. Bacteriophage lambda derivatives carrying two copies of the cohesive end site. J Mol Biol 1974; 83(4):511-25.
110. Kuzminov A, Schabtach E, Stahl FW. Chi sites in combination with RecA protein increase the survival of linear DNA in Escherichia coli by inactivating the ExoV activity of RecBCD nuclease. EMBO J 1994; 13:2764-2776.
111. Borukhov S, Severinov K. Role of the RNA polymerase sigma subunit in transcription initiation. Res Microbiology 2002; 153:557-562.
112. Murialdo H, Tzamtzis D. Mutations of the coat protein gene of bacteriophage λ that overcome the necessity for the FI gene. The EFi domain. Mol Microbiol 1997; 24:341-53.
113. Murialdo H, Tzamtzis D, Berru M et al. Mutations in the terminase genes of bacteriophage λ that bypass the necessity for FI. Mol Microbiol 1997; 24:937-952.
114. MacKinlay AG, Kaiser AD. DNA replication in head mutants of bacteriophage?λ. J Mol Biol 1969; 39:679-683.
115. Wake R, Kaiser A, Inman R. Isolation and structure of phage lambda head-mutant DNA. J Mol Biol 1972; 64:519-540.
116. Murialdo H, Fife WL. Synthesis of a trans-acting inhibitor of DNA maturation by prohead mutants of phage λ. Genetics 1987; 115:3-10.
117. Becker A, Murialdo H, Lucko H et al. Bacteriophage lambda DNA packaging. The product of the FI gene promotes the incorporation of the prohead to the DNA-terminase complex. J Mol Biol 1988; 199(4):597-607.
118. Becker A, Gold A. Enzymatic breakage of the cohesive end site of phage lambda DNA: Terminase (ter) reaction. Proc Natl Acad Sci USA 1978; 4199-4203(75).
119. Chow S, Daub E, Murialdo H. The overproduction of DNA terminase of coliphage lambda. Gene 1987; 60:277-289.
120. Catalano CE, Tomka MA. Role of gpFI protein in DNA packaging by bacteriophage lambda. Biochemistry 1995; 34(31):10036-42.

121. Cai Z-H, Hwang Y, Cue D et al. Mutations in Nu1, the gene encoding the small subunit of bacteriophage? λ??terminase, suppress the postcleavage DNA packaging defect of cosB mutations. J Bacteriol 1997; 179:2479-2485.
122. Pruss G, Calendar R. Maturation of bacteriophage P2 DNA. Virology 1978; 86:454-467.
123. Chai S, Bravo A, Luder G et al. Molecular analysis of the Bacillus subtilis bacteriophage SPP1 region encompassing genes 1 to 6. The products of gene 1 and gene 2 are required for pac cleavage. J Mol Biol 1992; 224:87-102.
124. Laski F, Jackson E. Maturation cleavage of bacteriophage P22 DNA in the absence of DNA packaging. J Mol Biol 1982; 154:565-79.
125. Davidson A, Gold M. A novel in vitro DNA packaging system demonstrating a direct role for the bacteriophage λ FI gene product. Virology 1987; 161:305-315.
126. Murialdo H, Fife W, Becker A et al. Bacteriophage lambda DNA maturation. The functional relationships among the products of genes Nul, A and FI. J Mol Biol 1981; 145(2):375-404.
127. Lin H, Simon M, Black L. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. J Biol Chem 1997; 272:3495-3501.
128. Lurz R, Orlova E, Gunther D et al. Structural organisation of the head-to-tail interface of a bacterial virus. J Mol Biol 2001; 310:1027-1037.
129. Fujisawa H, Shibata H, Kato H. Analysis of interactions among factors involved in the bacteriophage T3 DNA packaging reaction in a defined in vitro system. Virology 1991; 185:788-794.
130. Fujisawa H, Morita M. Phage DNA packaging. Genes to Cells 1997; 2:537-545.
131. Schliwa M, Woelke G. Molecular motors. Nature 2003; 422:759-765.
132. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21(2):319-31.
133. Hendrix R. Symmetry mismatch and DNA packaging in large bacteriophages. Proc Natl Acad Sci USA 1978; 75:4779-4793.
134. Simpson A, Tao Y, Leiman P et al. Structure of the bacteriophage phi29 DNA packaging motor. Nature 2000; 408:745-750.
135. Serwer P. Models of bacteriophage DNA packaging motors. J Struct Biol 2003; 141:179-188.
136. Dokland T, Murialdo H. Structural transitions during maturation of bacteriophage lambda capsids. J Mol Biol 1993; 233(4):682-94.
137. Murialdo H. Bacteriophage lambda DNA maturation and packaging. Ann Rev Biochem 1991; 60:125-153.
138. Imber R, Tsugita A, Wurtz M et al. Outer surface protein of bacteriophage lambda. J Mol Biol 1980; 139(3):277-95.
139. Perucchetti R, Parris W, Becker A et al. Late stages in bacteriophage lambda head morphogenesis: In vitro studies on the action of the bacteriophage lambda D-gene and W-gene products. Virology 1988; 165(1):103-14.
140. Sternberg N, Weisberg R. Packaging of coliphage lambda DNA: II. the role of the gene D protein. J Mol Biol 1977; 117:733-759.
141. Wendt J, Feiss M. A fragile lattice: Replacing bacteriophage?λ's head stability gene D with the shp gene of phage 21 generates the Mg^{++}-dependent virus, λ shp. Virology 2004; 326:41-46.
142. Smith D, Tans S, Smith S et al. The bacteriophage phi29 portal motor can package DNA against a large internal force. Nature 2001; 413:748-52.
143. Casjens S, Wyckoff E, Hayden M et al. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. J Mol Biol 1992; 224(4):1055-74.
144. Tavares P, Santos MA, Lurz R et al. Identification of a gene in Bacillus subtilis bacteriophage SPP1 determining the amount of packaged DNA. J Mol Biol 1992; 225(1):81-92.
145. Casjens S. Bacteriophage lambda FII gene protein: Role in head assembly. J Mol Biol 1974; 90:1-20.
146. Maxwell K, Yee A, Booth V et al. The solution structure of bacteriophage lambda protein W, a small morphogenetic protein possessing a novel fold. J Mol Biol 2001; 308:9-14.
147. Maxwell K, Yee A, Arrowsmith C et al. The solution structure of the bacteriophage lambda head-tail joining protein, gpFII. J Mol Biol 2002; 318:1395-1404.

CHAPTER 3

DNA Packaging in Bacteriophage T4

Venigalla B. Rao and Lindsay W. Black

Introduction

Double-stranded (ds) DNA packaging in phage T4 and other icosahedral viruses is a fascinating biological problem. During packaging, a complex, metabolically active, concatemeric DNA is translocated into an empty prohead in an ATP-driven process and condensed as a highly ordered structure of near crystalline density.[1-3] dsDNA packaging serves as an excellent model system to understand fundamental biological mechanisms such as the reversible condensation and decondensation of DNA, DNA movement along protein complexes, and transduction of ATP hydrolysis energy into mechanical motion of DNA.

The phage T4 head assembly pathway produces a complex prohead consisting of six essential proteins and at least seven nonessential proteins.[4] The T4 prohead maturation protease degrades the scaffolding core into peptides and cleaves off the N-termini of the major capsid protein (gp23) and the vertex protein (gp24) to generate an empty mature prohead that is competent for DNA packaging. In parallel, the T4 DNA replication pathways generate a highly branched "endless" concatemeric DNA, which is associated with a myriad of protein complexes involved in replication, transcription, recombination and repair. A terminase complex of gp16 (18 kDa) and gp17 (70 kDa) links these two pathways by recognizing the viral concatemer, making an endonucleolytic cut, and joining it to the prohead through specific interactions with the dodecameric portal vertex constituted by gp20 (Table 1). Consequently, a DNA packaging machine is assembled, with the terminase as one of the key components (Fig. 1). This machine translocates an intact, unit-length, linear dsDNA genome into the capsid to form a highly ordered condensed structure. Terminase apparently also makes the second cut terminating DNA packaging, dissociates from the packaged structure, and repeats the DNA linkage to another prohead in a processive fashion.

The above represents an overall pathway for DNA packaging in T4, which, in many respects, is a common DNA packaging pathway among a number of well-characterized dsDNA phages. Common features among the dsDNA packaging phages include the dodecameric ring structure of the portal, a terminase complex formed by multiple copies of one small and one large subunit, the enzymatic functions associated with the terminase proteins, and the headful nature of the packaging process.[2] Recent evidence suggests that the ATPase motifs in terminase are particularly well conserved among the phage packaging proteins,[5] indeed, phage genomics reveals that these are among the signature proteins defining a phage "quasi-species".[6] The conservation appears to extend to the putative terminase from herpes viruses, which shows a particularly close resemblance to bacteriophage T4.[5,7]

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

Table 1. Major T4 packaging and prohead proteins

Gene Product	Mass	Function
gp23*	48 kDa	Major Capsid Protein
gp24*	46 kDa	Vertex Protein
gp20	61 kDa	Portal Protein
gp16	18 kDa	Small Terminase Protein
gp17	70 kDa	Large Terminase Protein

*represents the mature protein following N-terminal cleavage by T4 prohead protease

Thus, it is likely that the dsDNA bacteriophages and herpes viruses package DNA by a common mechanism. Phage T4, with its large genome (171 kb), complex genetic make-up and rich details of head morphogenesis, provides an excellent, though challenging, model system to elucidate the molecular details of the DNA packaging machine. In this review, we focus on the components of the T4 DNA packaging machine, in particular, the terminase proteins. At the end, we discuss the emerging details of the T4 packaging mechanism and provide evidence for a DNA translocating ATPase in phage T4.

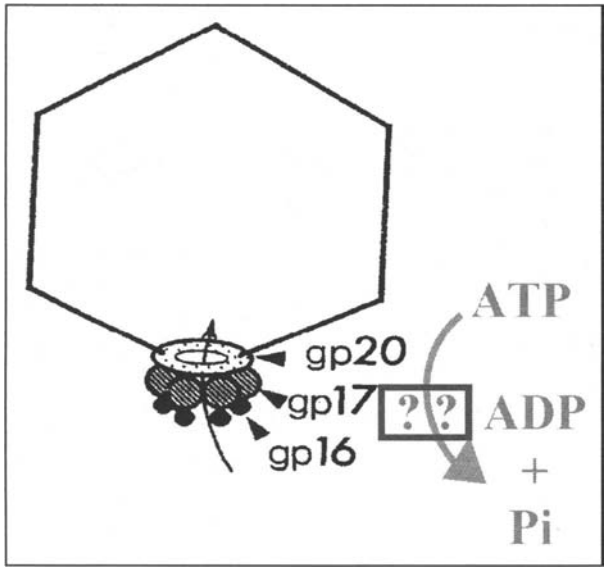


Figure 1. Schematic representation of the putative DNA packaging machine. The terminase proteins gp16 and gp17 cut the DNA concatemer and attach it to the empty prohead through interactions with the gp20 portal vertex. The packaging machine drives translocation of DNA utilizing the free energy of ATP hydrolysis. The question marks indicate that the ATPase motor for the machine is still unknown although evidence suggests that the large terminase protein gp17 is a strong candidate (see text). The arrow within the prohead represents the DNA end generated by the terminase.

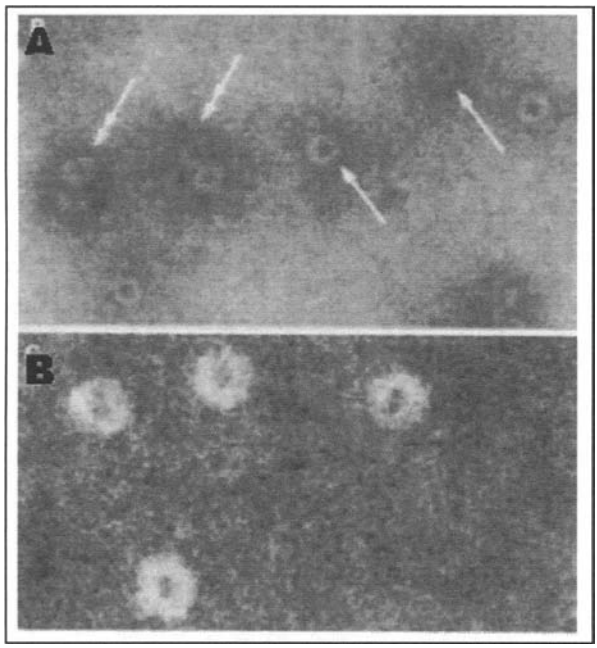


Figure 2. STEM micrographs of gp16 oligomers.[11] A) Bright field micrograph of gp16 stained with methylamine vanadate. Single- and double-headed arrows indicate the single and double rings, width of the ring monomer ~8 nm. B) Same as (A), except at higher magnification to show the central channel of the rings, ~2 nm diameter.

Components of the Phage T4 DNA Packaging Machine

Gp16: The Small Terminase Protein

Gp16 is an 18 kDa protein.[8,9] It is the small subunit of the terminase holoenzyme. Gp16 is dispensable for in vitro DNA packaging although packaging efficiencies are low in its absence. It is essential in vivo since *16am* mutations are lethal, although microscopy reveals that, in double *am* mutants of gene *16*, packaging begins late after infection and proceeds slowly and incompletely.[10] Gp16 appears to play an important role in DNA recognition and in modulating the functions of the terminase/packaging complex. It consists of the following binding sites and associated activities.

Oligomerization Site

Gp16 forms stable oligomeric rings and double rings with a diameter of ~8 nm and a central channel of ~2 nm[11] (Fig. 2). Each ring apparently consists of about 8 gp16 monomers. Since expression vector synthesis of gp16 with or without a his-tag, and refolding of gp16 from 6M urea lead to the ring and double-ring structures and little, if any, monomeric gp16, these structures are strongly and preferentially formed.[11,12] They are also observed to be active for in vitro DNA packaging. Sequence analysis identifies a conserved stretch of hydrophobic amino acid sequence in the center of the protein, which may be responsible for the protein-protein interactions that lead to oligomerization (Fig. 3).[5,11] Consistent with this hypothesis, truncation experiments show that the N-terminal 55 residues do not oligomerize whereas the COOH terminal 121 residues do oligomerize (N. Malys, unpublished). An analogous site was recently

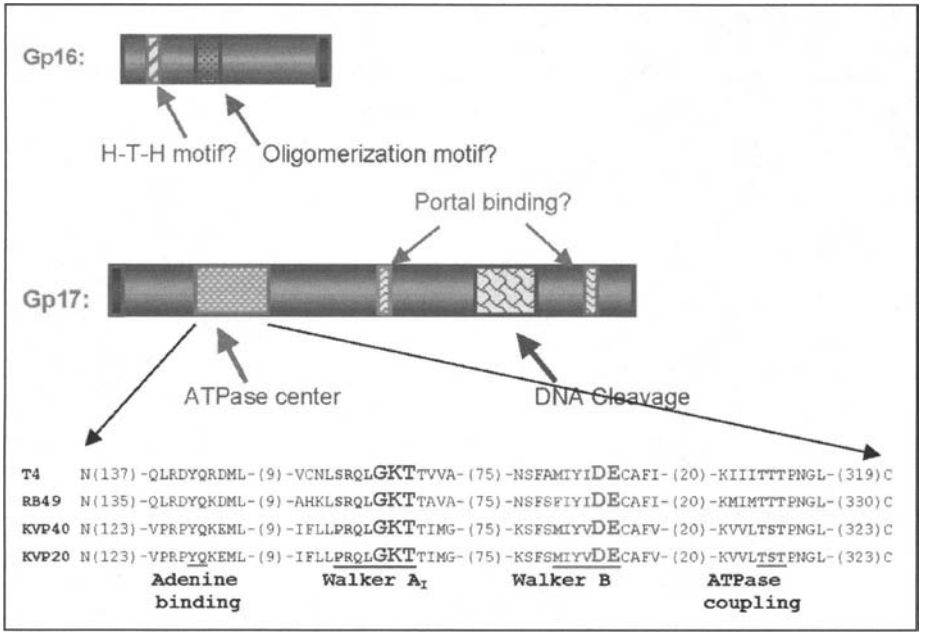


Figure 3. Schematic representation of putative functional sites in gp16 and gp17. Top) The functional motifs in gp16 and gp17 are labeled as shown. Note that an out-of-frame 15 bp sequence codes for 5 amino acids at the amino-terminal-end of gp17 as well as for 5 amino acids at the carboxy-terminal-end of gp16. The functional sites are represented by different shaded boxes. The insert shows alignment of the putative ATPase domain of T4 family terminases; Phages RB49, KVP40 and KVP20 are related to T4; all are large-tailed phages with an elongated prolate head and a contractile tail.[5] The sequence similarity among their genomes suggests that they most likely diverged from a common ancestor. The critical amino acids as discussed in the text are shown with a larger Font. Amino acids are shown using the single letter code. The numbers in parentheses represent number of amino acids. N and C correspond to the amino and carboxy termini respectively of gp17. See text and Mitchell et al[5] for details.

identified in the phage λ small terminase subunit gpNuI.[13] One model for the ring and double ring structures consistent with a single multimerization interface is that the structures are helical lock washers and double washers, where the double washers unstack. Gp16 rings may play a role in the formation of higher order packaging initiation complexes consisting of the holo-terminase and DNA bound to the portal ring (Fig. 1).

ATP Binding Site

There are two forms of gp16 synthesized in vitro and in vivo, most likely because the out-of-frame partial overlap region between genes *16* and *17* induces translational frameshifting at the ribosome binding site for gene *17*, which is located within the 3'-end of gene *16* (Fig. 3). The short form gp16 of 155 residues is truncated at a C-terminal R residue, 9 amino acids from the end of the full-length 164 residue protein.[11] The short form predominates in vivo. In fact, expression of the short form gp16 alone is sufficient to complement gene *16 am* mutants (H. Lin, unpublished observations). Gp16 is shown to have a weak ATP binding activity with the longer form gp16 showing greater ATP binding.[11] A Walker-B type ATP binding motif was predicted in the central region of gp16,[14] which is analogous to the predicted ATP reactive sites within the small terminase subunits of λ and SPP1. However, neither of the potential D

residues appear likely to be the metal-chelating D of a Walker-B motif. In addition, the mostly α-helical secondary structure prediction for gp16 suggests that the protein does not contain the classic nucleotide-binding fold, which consists of β-strands at its core. The gp16 protein does not display an ATPase activity,[11,12] among terminase small subunits, only the λ Nu1 protein has been reported to display an ATPase activity.[15] The observed weak ATP-binding activity of gp16 is likely due to an anomalous ATP-binding site whose biological function, if any, is uncertain.

DNA Binding Site

T4 gp16 oligomer displays little or no DNA binding, but when renatured from urea together with DNA, gp16 is observed to bind to dsDNA but not ssDNA.[11] There appears to be preferential binding to DNA containing gene *16*. Analogous to the small terminase subunits of phages λ and SPP1, a helix-turn-helix (H-T-H) DNA binding motif is predicted in the N-terminus of gp16. This motif appears to be a variant of the classical H-T-H signature structure determined for many DNA binding proteins that has been characterized structurally as a winged helix-turn-helix in the case of the λ Nu1 protein.[16]

Stimulation of Gp17 ATPase and Packaging Activities

Gp16 stimulates the gp17-associated ATPase and in vitro DNA packaging activities by at least 50-fold[12] (also, see below). Costimulation of ATPase and DNA packaging activities suggested a linkage between these two functions. Although the mechanism is unknown, it was proposed that gp16 interaction with gp17 induces a conformational change in the large subunit, transforming a basal ATPase into a stimulated ATPase with high catalytic capacity. This may be analogous to the stimulation of GTPases by interaction with the GAPs[17] and hence may represent a common molecular switch (see below). Experiments are currently underway to test this hypothesis.

Gp17: The Large Terminase Protein

Gp17 is the 70 kDa large subunit of the terminase holoenzyme.[8,9] Gp16 and gp17 together form the terminase holoenzyme complex although the stoichiometry of the subunits in the holoenzyme is unknown. Unlike in the case of the small subunit gp16, gp17 alone is sufficient for packaging DNA in vitro.[8,12] A number of functional motifs and associated activities have been recognized, which shed light on the roles of gp17 in the DNA packaging pathway.

ATP Binding Site I

The consensus sequence of the Walker-A nucleotide binding motif, (G/A)XXXXGK(T/S), is present in a large number of enzymes capable of nucleotide binding and/or hydrolysis. Two Walker-A motifs have been identified in gp17,[9,18] the N-terminus proximal $SRQLGKT_{161\text{-}167}$ (Walker-A_I) and a centrally located $TAAVEGKS_{299\text{-}306}$ (Walker-A_{II}). As shown recently, Walker-A_I is highly conserved among all the four T4-family (T4D, RB49, KVP40, KVP20) terminase sequences (Fig. 3) as well as in numerous other phage terminases and herpes virus terminases.[5,19] Extensive combinatorial mutagenesis analyses of this site revealed a striking conservation of its features. No substitutions were tolerated in the highly conserved GKT signature sequence of the T4 gp17 Walker-A, including the conservative substitutions G165A, K166R, and T167A[19] (Fig. 4, also, see below).

ATPase sites also contain a Walker-B motif, with the consensus sequence ZZZZD (Z represent a hydrophobic amino acid), which is generally located 50-130 residues downstream of the Walker-A_I lysine. The four hydrophobic amino acids of the Walker-B motif form a β-strand that ends with the highly conserved aspartate. This well-characterized aspartate is responsible for chelating Mg^{2+} of the bound Mg-ATP complex and for orienting the substrate

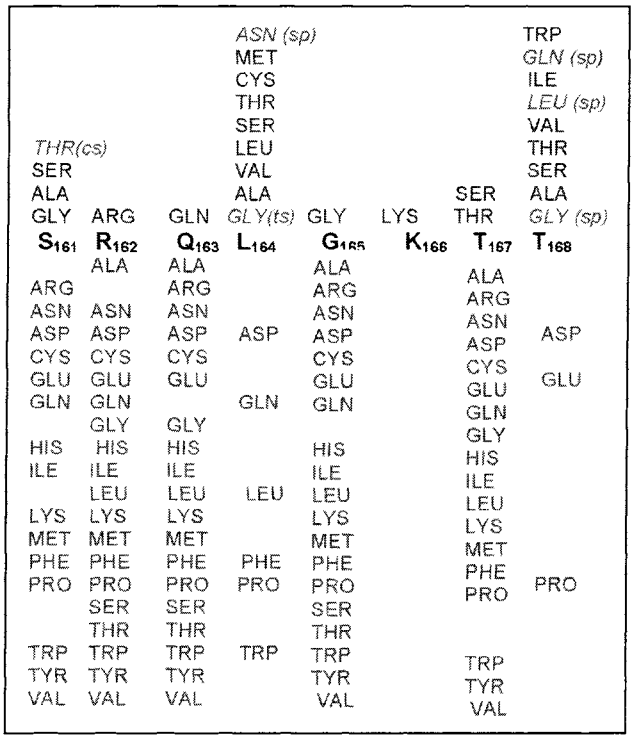


Figure 4. The critical Walker A motif in gp17.[19] Phenotypes of Walker-A mutants.[19] Amino acid substitutions below the native sequence resulted in a null phenotype and the ones above resulted in a functional phenotype. Some of the null phenotypes shown were inferred from the marker rescue data. The conditional phenotypes sp (small plaque, small to minute plaques at 37°C, ts (temperature sensitive, no plaques at 42°C) and cs (cold sensitive, no plaques at 20°C), are shown in italics (see ref. 19 for further details). Each substitution shown corresponds to a single amino acid substitution from independent single amino acid libraries constructed for each residue.

for nucleophilic attack by an activated water molecule. Sequence analysis of the T4 large terminase shows that a potential Walker-B motif, $MIYID_{251\text{-}255}$, is located 94 amino acid residues downstream of the Walker A lysine[5] (Fig. 3). The critical elements of this Walker-B motif not only appear to be strictly conserved in the four T4-family terminases, but also in HSV-1 U_L15 and other terminases.[5] Recent mutagenesis and biochemical studies of the gp17-D_{255} further provided evidence supporting the Walker-B assignment to this sequence (M. Mitchell and V. Rao, unpublished data). In addition, a catalytic carboxylate residue (E_{256}) and a novel helicase motif III that is presumably involved in ATPase coupling have been identified.[5] The catalytic carboxylate residue is required to activate a water molecule for an in-line attack on the γ-phosphate of ATP, whereas the motif III in helicases makes a direct contact with the nucleic acid backbone and triggers subsequent conformational changes resulting in ATP hydrolysis.[5]

ATP-Binding Site II

A second Walker-A P-loop has been proposed in a number of terminases.[18] In T4 gp17, the second ATP-binding site is represented by the sequence $TAAVEGKS_{299\text{-}306}$ with its apparent Walker-B motif close to the C-terminus $GVSVAKSLYMD_{468\text{-}478}$.[20] Unlike the striking conservation of Walker-A_I, the sequence conservation of Walker-A_{II} and its surrounding region

is quite poor. Moreover, the gp17s from phages KVP40 and KVP20 substitute isoleucine and valine respectively for the critical lysine of the canonical Walker-A motif (K_{305} in T4 gp17).[5] Such a substitution would severely compromise ATP binding and hydrolysis. Recent mutagenesis data further showed that nonconservative single and double substitutions are tolerated at the residues K_{305} and S_{306}; in fact, screening of a library consisting of all possible double substitutions showed that >30% of the substitutions, including PG, VG, RN and SL, were tolerated, suggesting that this site is not critical for gp17 function.[19]

Terminase Cutting Site

One or more H-X_2-H type metal-binding motifs with suggested roles in DNA-binding and endonucleolytic activity have been identified in terminases.[18] The histidine-rich motif in the C-terminal half of T4 gp17, H_{382}-X_2-H_{385}-X_{16}-C_{402}-X_8-H_{411}-X_2-H_{414}-X_{15}-H_{430}-X_5-H_{436}, appeared more complex consisting of three H-X_2-H type motifs.[19] Site-directed mutagenesis of His_{436} showed that none of the twelve substitutions tested were tolerated, which is a clear indication that this is a critical residue.[21] Biochemical analyses of H436R gp17 mutant further showed that this mutant lost both the DNA packaging and terminase cutting activities. Recent combinatorial mutagenesis experiments identified certain strictly conserved aspartate residues within the "histidine-rich" motif (D_{401} and D_{409}) that are critical for gp17 function and some of these mutants exhibited a loss of terminase cutting activity, but not the in vitro DNA packaging activity[22] (F. Rentas and V. Rao, in preparation). We suggest that the terminase DNA cutting site consists of a cluster of histidine and aspartic acid residues that are involved in metal-coordinated acid-base catalysis of DNA cleavage, as was reported in a number of nucleases.[23]

Prohead Binding Site

The portal-terminase interactions appear to be mediated by the negatively charged and hydrophobic residues at the carboxy-terminus of the large terminase protein. Analogous to the phages λ and T3, the negatively charged sequence $ELQDMSDDYAP_{578\text{-}588}$ at the carboxy-terminus of gp17 was considered a good candidate for the portal binding site. However, it turned out that this sequence is not critical for function since a recombinant gp17 that is truncated after K_{577} is fully functional for in vitro DNA packaging (V. Rao, unpublished data). Based on recent genetic data, other negatively charged clusters, $LYNDEDIFDD_{322\text{-}331}$, and $IDYADKDD_{560\text{-}567}$ are likely candidates for the portal binding site.[5,24] A peptide from phage T4 portal protein gp20 (residues 281-308), which encompasses the packaging-defective cold sensitive (*cs*) mutations in *g20*, was shown to interact with gp17 and block DNA packaging in vitro.[24] Second site mutations that suppress *cs20* mutations map in *g17*. These substitutions, I364F (*ts*R1) and S583N (*ts*L51), map close to the proposed sites, which are also well conserved among the T4 family terminases, suggesting that additional residues in the flanking sequence also provide specificity to the terminase-portal protein interactions.[24]

Activities Associated with gp17

DNA Packaging

Packaging efficiencies measured at 10% wild type (>10^8/ml infected bacteria) can be obtained using purified terminase proteins.[12,25] Gp16 enhances gp17-dependent packaging activity by about 100-fold at low concentrations of gp17.[8] DNA packaging in vitro normally requires both gp16 and gp17, but gp17 alone may suffice at high concentrations.[11,12] Endogenous phage T4 concatemeric DNA that accumulates in packaging defective phage infections is packaged at about 100-fold greater efficiency than externally added mature DNA.[12] Phenol extracted (protein-free) mature and concatemeric DNAs are packaged with comparable efficiency and predominantly by recombination into the active concatemeric DNA as judged by

formation of phage recombinants in vitro.[12,25] The very high competence of the endogenous DNA suggests that packaging occurs hand-in-hand with other DNA metabolic processes such as recombination and repair. Certain structural features of the metabolically active DNA and/or the associated protein complexes may provide sites for recognition and docking of the terminase onto the DNA substrate.[25] Thus concatemeric DNAs produced by genes *55* and *33* am defective infection are inefficiently packaged, apparently at least in part because of the absence of gp55 which interacts with the gp17 terminase subunit and may dock it to DNA.[25] The gp55 and gp33 are the T4 late transcription activator (σ-factor) and coactivator, respectively, and link replication and late transcription through attachment to the processivity sliding clamp, gp45, loaded onto the DNA.

ATPase

In concurrence with the DNA packaging results, gp17, but not gp16, exhibits an ATPase activity.[12] This activity hydrolyzes the β-γ phosphoanhydride bond of ATP generating ADP and P_i. The gp17 ATPase is highly specific to ATP; dATP is also cleaved, but at a reduced efficiency. However, none of the other NTPs/dNTPs are hydrolyzed to a significant extent. The k_m for ATP hydrolysis is about 110 μM and the k_{cat} is about 2 ATPs hydrolyzed/gp17 subunit/min. These data, in particular the low k_{cat}, are consistent with the notion that the basal ATPase activity of gp17 is maintained at a low level when it is not coupled to DNA translocation.[12]

Gp16-Stimulated ATPase

As mentioned above, the gp17-dependent in vitro DNA packaging activity is enhanced by gp16 by about 100-fold at limiting concentrations of gp17. Gp16 does not possess an ATPase activity, but stimulates the gp17-associated ATPase by about 50-fold.[12] The stimulation is specific and shows dependence on the gp16:gp17 ratio, implicating a specific gp16-gp17 interaction to form a holo-terminase complex. The gp16:gp17 ratio for maximal stimulation is estimated to be 6-8:1, although further biochemical studies are underway to determine the stiochiometry of an active holoterminase complex. Analysis of the catalytic parameters of the gp16-stimulated ATPase activity showed that the increase is due to an increase in the catalytic capacity of gp17, but not due to an increase in the affinity towards the ATP substrate;[12] the k_m and k_{cat} for ATP hydrolysis in the presence of gp16 are about 256 μM and 107 ATPs hydrolyzed/gp17 subunit/min, respectively (k_m: 110 μM and k_{cat}: 2 for gp17 alone).

Phosphorylation

Incubation of gp17 with ATP results in phosphorylation of gp17. About 10% of the total gp17 is estimated to be in the phosphorylated form, which seemed to be trapped only under conditions of low catalytic rates for ATP hydrolysis such as the absence of gp16, and incubation at low temperatures.[12] This and the fact that some of the bound P_i is released in the presence of gp16 suggests that the phosphorylated gp17 is an intermediate of ATP hydrolysis rather than gp17 having an independent protein kinase activity. It is important to determine the phosphorylation site in gp17 and further characterize its role in the catalytic process. A phosphorylated intermediate of gp17 shows mechanistic similarity to some motor proteins, such as the SR ATPase, but not others, such as myosin or F_1F_0ATPase, which do not display a covalently attached P_i intermediate.[26]

DNA Cleavage (Terminase)

In phages with specific ends (e.g., λ, T7, T3) and phages with nonsequence specific *pac* and headful cutting ends (e.g., P22, SPP1, P1), the nonstructural packaging proteins cleave the newly replicated viral concatemeric DNA and generate the termini of the packaged DNA molecule.[2,3] The small terminase protein recognizes the viral DNA substrate and directs the

large terminase protein to the cleavage site, which will then be cleaved by an endonucleolytic activity associated with the large protein. Phage T4 appears to have a similar functional division, since it packages DNA by a headful mode yielding circularly permuted and terminally redundant ends, i.e., the ends of the packaged genome are widely dispersed over the genome with an estimated 3.3 kb (~2% of the genome) repetition at the ends. As in the *pac* site phages, T4 packaging is terminated at a random sequence following packaging of ~102% (one headful) of the viral genome, to yield the terminal redundancy of 3.3 kb. Because phenol extracted mature T4 DNA can be efficiently ligated by T4 ligase in vitro,[27] either the mature DNA is blunt ended, or the same local nucleotide sequence is cleaved by terminase in packaging, the latter possibility appearing less likely. Overall, the T4 terminase must be a nonspecific but strictly regulated endonuclease, which allows only limited headful length cuts coupled to DNA packaging. The terminal headful cutting may be regulated by portal protein interaction, as is thought to be the case for phages P22 and SPP1.

Evidence thus far suggests that the large terminase protein indeed exhibits a nonspecific endonuclease activity.[28,29] Expression of gp17 in *E. coli* from strong promoters such as phage T7 and λ pL promoters results in extensive cleavage of both the resident plasmid DNA and the *E. coli* genomic DNA. That this activity is associated with the large terminase protein gp17 was rigorously characterized using a number of control recombinant constructs. The cleaved DNA exhibits a characteristic smear that extends throughout the lane upon agarose gel electrophoresis. Mapping the ends of isolated cleaved DNA showed no sequence specificity. Further characterization of this activity revealed that the T4 terminase exhibits a preference towards the cleavage of transcriptionally active DNA.[28,29]

Since gp17 nonspecifically cleaves the *E. coli* genomic DNA, significant basal level expression of gp17 in leaky expression strains such as *E. coli* BL21 (DE3) is lethal. This characteristic was exploited to select point mutations that render this activity defective and hence would allow *E. coli* BL21 (DE3) to survive and form colonies.[21] Such mutants, which are no longer lethal to *E. coli*, lack the characteristic cleavage of DNA upon gp17 expression. This strategy, as described above, allowed mapping of a terminase cutting site in the C-terminus of gp17.[21,22]

Franklin et al reported an in vitro terminase activity associated with the full-length gp17, which degrades single stranded DNA.[20] The gp17 preferentially binds to the single stranded DNA at the junctions of single and double stranded DNA (putative recombinational or replicative intermediates) and degrades the single stranded portion of the junction in a nonspecific manner. While these data in some ways mimic the in vivo terminase activity discussed above, a significant reservation should be noted. The gp17 purified by Franklin et al binds to ssDNA cellulose columns tightly, an activity not reported with the other gp17 preparations.[8,12,30] Also, their purified gp17 was not characterized by the in vitro DNA packaging assay in order to compare it with the other independently purified preparations, which were characterized by the bench-mark in vitro DNA packaging assay. The in vitro DNA packaging assay remains the only specific biological assay available to assess the functionality and identity of the purified protein.

Gp20: The Portal Protein

The T4 portal protein gp20 (61 kDa) was first determined to be directly connected to packaging by gene *20 cs* mutations that blocked packaging initiation at low temperature. The prohead defect was reversed upon temperature shift or could be suppressed by specific gene *17* terminase mutations, thereby showing intimate association between terminase and portal proteins in packaging.[31,32] As discussed above, a number of clustered *cs* mutations in the portal gene and their terminase suppressors have been sequenced.[24] Early electron microscopic structure determination showed that the T4 portal is a dodecamer.[33] In fact, it is likely that its 3D structure is similar to the ϕ29 and SPP1 portal proteins.[34] Assembly of the T4 portal is unusual

in requiring a specific assembly factor or chaperone which leads to assembly on the cytoplasmic membrane.[35-37] Assembly of the T4 prohead is in turn strongly dependent upon assembly of the membrane–linked portal, since in the absence of the portal protein, tubes rather than proheads form with a delay.[35] Proteolytic processing of the prohead releases the prohead from the membrane—among the prohead structural proteins only the portal protein is not processed—and presumably this maturation frees the portal for interaction with the terminase in the cytoplasm, where electron microscopy reveals packaging occurs.[4] In relation to the mechanism of DNA packaging and portal packaging function, modifications of the T4 portal protein by means of gene fusions reveal that bulky portal fusion proteins extending inside (gp20-gfp) or outside (HOC-gp20) the prohead can be assembled together with truncated gp20.[30] Formation of active phage heads containing these altered portals poses a problem for DNA translocation models that invoke the necessity of portal rotation.[38]

Major Events during Phage T4 DNA Packaging

Packaging Initiation

In a common pathway for packaging initiation in dsDNA phages, the small terminase subunit recognizes the viral genome, facilitates the assembly of a holoterminase complex on the recognition site, the large terminase subunit then cuts the DNA, and the DNA end is linked to the empty prohead via interactions with the portal protein.[2] Phage T4 follows this common pathway. Like the *pac* site phages, which recognize a unique sequence in the phage genome and cut near it to initiate a series of processive, random sequence ended phages, phage T4 generates circularly permuted ends. The evidence suggests that the phage T4 small terminase subunit gp16 also recognizes a *pac* site near the 3' end of its structural gene to create packaging initiation sites.

Evidence for T4 *pac* site recognition includes small terminase subunit gp16 binding to dsDNA containing the gene *16 pac* site, analogous to the other phage terminase small subunits.[11] A variant H-T-H DNA binding motif is proposed in the gp16 sequence.[5,11] In vivo genetic evidence for gp16 binding to the *pac* site includes the observations that: (i) under selection for higher copy numbers of gene *17* to increase its expression, gp16 is required for amplification of a gene *16* to gene *19* region of the phage T4 chromosome bounded by the *pac* site in gene *16* and a homologous region in gene *19;* site directed mutagenesis of a 24 base pair homology region in the gene *16 pac* site can eliminate the amplification,[39,40,43] (ii) recombination between genes *16* and *19* in plasmids at the preferred amplification junction likewise requires synthesis of gp16;[41] western blotting confirms that gp16 synthesis occurs from these gene *16* containing plasmids (H. Lin, unpublished data); (iii) plasmids and prophages containing the gene *16 pac* site show enhanced transduction by the T4 transducing phage variant T4GT7, whereas site directed mutagenesis of the *pac* site lowers the transduction;[42] and (iv) a *pac* site DNA fragment of ~10 kb extending from the unique *Bam*HI site of phage T4 to gene *16* is found in mature phage T4 DNA.[42] Thus, overall, a gene *16 pac* sequence promotes transduction of *pac*-containing DNA, is found frequently to terminate mature T4 DNA, and gp16 is required to recognize the gene *16 pac* sequence to promote genetic recombination in vivo between it and a homologous sequence in gene *19* to produce terminase amplification mutants.

The above observations strongly suggest that phage T4 gp16 recognizes and binds a *pac* sequence near the 3' end of gene *16*. A model for gp16 DNA recognition and binding is that, as synthesized, the gp16 monomer binds to a preferred *pac* region toward the end of gene *16*. Oligomerization releases gp16 from DNA binding and allows the gp16 ring to slide along the DNA and to participate in mature DNA end formation and packaging. Formation of the double ring could lead to the observed in vivo recombination-driven amplification of gp16 binding sequences due to synapsis of DNA segments each binding a single gp16 ring.[42-44] This model proposes that the gp16 synapsis-related packaging function is to regulate DNA end formation by

making it dependent upon DNA synthesis and accumulation, i.e., upon concatemerization of identical sequences.[44] The model is consistent with the observation that the *pac* fragment in mature T4 DNA is more abundant when multiple plasmid gene *16* sequences are also present in the infected host,[42] and with the absence of DNA binding by the gp16 multimer.[11] Overall, this model for how gp16 recognizes the viral DNA substrate for packaging conforms to functions conserved among small terminase subunit proteins, including binding to a *pac* or *cos* region located within or near the small subunit gene. However, the mode of DNA binding by gp16 and the existence of a DNA bound gp16 ring form require experimental demonstration.

The old phage T4 literature is probably inadequate to distinguish between *pac* site and "random headful packaging" modes, which are mechanistically similar, especially if head filling is highly processive; i.e., an infrequent cut near to a *pac* site followed by multiple processive packaging rounds yields headfuls of apparently "randomized" DNA-containing phages. But, in fact, in the old phage T4 literature the most complete electron microscopic survey of heteroduplex DNAs formed from mature phage T4 particle DNAs concluded that although phage T4 mature DNA end sequences are dispersed over the chromosome, they are not "randomized" but distributed in a manner suggesting one or more *pac* sites in the genome.[45] Given a gene *16* *pac* site, it is not excluded that other modes of end formation and packaging are also utilized; indeed, the observation that site directed mutagenesis of the gene *16 pac* sequence which eliminates gene *17* amplification is not lethal,[43] suggests that other packaging modes can intervene, as does the slow accumulation of filled heads in a gene *16* defective infection.[10]

Prohead Expansion

In T4 and other dsDNA bacteriophages, procapsid expansion accompanying DNA packaging is an important event, resulting in an increase in the inner capsid volume by about 50% in T4. In phage T4 head assembly, expansion is a striking transformation resulting in the stabilization of capsid subunit interactions, reorganization of capsid protein epitopes between the outer and inner surfaces of the capsid shell, and exposure of binding sites for the outer capsid accessory proteins Soc and Hoc.[46,47] Changes in ATP- and DNA-binding properties of the assembled capsid protein have also been reported.[48,49] This remarkable transformation of the capsid structure during expansion, and the timing of this transformation which coincides with DNA packaging, have raised interesting biological questions.

What is the trigger for prohead expansion? A number of studies indicate that the expansion transformation is an inherent property of the capsid protein, and can occur both in vitro and in vivo in isolation from DNA packaging.[46,50,51] It is also clear that expansion is not energetically coupled to DNA packaging since much of the DNA enters an already expanded prohead and that DNA can be packaged in vitro into an expanded prohead.[52] However, in vivo, it is likely that interaction of the terminase-DNA complex with the portal vertex and possibly the entrance of some DNA into the prohead, is the natural trigger for expansion. Jardine et al have identified an unexpanded, presumably cleaved, T4 prohead containing some DNA, which apparently can be chased into an expanded prohead and eventually phage.[53,54] The amount of DNA associated with expansion in T4 is not characterized. There is better evidence for the above hypothesis in phages λ and T3 than in T4. It is well documented in the λ and T3 defined DNA packaging systems that expansion occurs after packaging of about 11% and 25% of the genomes, respectively.[55,56] These data suggest that the terminase-prohead interactions, packaging of a small fraction of the DNA, prohead expansion, and continued packaging into an expanded head, occur in that order. Nevertheless, in both T4 and T3/T7, efficient initiation and completion of packaging into expanded proheads can be demonstrated, showing that there is no necessary mechanistic coupling between expansion and packaging.[52,56,57]

What is the biological significance of prohead expansion? The overall expansion transformation appears to be an energetically favorable process although it must be triggered.[51] The initial

thermodynamic barrier is probably overcome by specific terminase-DNA-portal interactions. It has been demonstrated that expanded T4 proheads are stabilized relative to the procapsid precursors.[4] In addition to this function, expansion could play a role in DNA packaging or ejection. According to one hypothesis, the DNA entering the prohead interacts with the inner surface of the unexpanded capsid, serving as a "nucleator" for organization of the incoming DNA.[1] The following expansion would weaken these interactions, allowing organization of the packaged DNA and its delivery during the infection process.[1] Alternatively, expansion changes the inner capsid surface from a core-interacting surface to a DNA-interacting surface.[52] The inner surface of the capsid, which is built with a core-interacting capability, should be transformed at the time of DNA packaging since a core-interacting surface would be irrelevant, if not detrimental, for the accompanying DNA packaging event. Consistent with these hypotheses, there is evidence in phage T4 for a shift in capsid protein epitopes from outside to inside of the capsid surface during expansion.[58] How this bears on the structure of the packaged DNA is unknown. Phage T4 DNA is packed to ~500 mg/ml, the same density as other dsDNA phages. The condensed T4 DNA is oriented parallel to the long axis of the prolate head. Proteins encapsidated with the DNA display mobility within the condensate, since e.g., Staphlococcal nuclease can hydrolyze the packaged DNA.[59]

DNA Translocation

Although a number of interesting models have been proposed, viz., portal protein as a translocating rotor, topisomerase nicking and resealing, osmotic pump, tracking along the DNA, and conformational switching, the basic mechanism of DNA translocation is still very much a mystery. Although recent cryo-EM and X-ray structural analyses of the ϕ29 portal ring allowed postulation of explicit details which link portal compression and rotation to translational movement of DNA,[38] there is no evidence yet for a rotating portal. It is however clear that, regardless of the DNA translocation mechanism and which component does it, it is powered by an ATPase "motor". The defined in vitro DNA packaging systems and structural data estimate that hydrolysis of one ATP is coupled to translocation of two base pairs of dsDNA.[60,61] This means that roughly 10^5 molecules of ATP are expended to package one molecule of T4 DNA at an estimated catalytic rate of 10^4 ATPs per min per packaging unit. It is important to bear in mind however that these are merely estimates based on evidence from defined in vitro DNA packaging systems, which are inherently subject to a number of uncertainties.

A key question that needs resolution to further analyze the mechanistics of DNA packaging is: which component of the packaging machine (Fig. 1) is the ATPase motor that powers DNA translocation? A strong case can be made for the large terminase subunit gp17. Recent studies with T4 have taken the lead to address this question rigorously using a combination of molecular genetic and biochemical approaches. A combinatorial paradigm was developed to analyze the ATPase site I (Fig. 3), which, as described above, appeared to be a critical motif in preliminary studies.[19] Every possible amino acid substitution was introduced at every residue of the putative Walker-A motif $SRQLGKT_{161\text{-}167}$ (Fig. 4). Substitutions were either not tolerated or highly restricted at any of the invariant residues of the canonical G(A)XXXXGKT(S) Walker-A P-loop. Most strikingly, no substitutions, other than a conservative T167S, were tolerated at the GKT stretch, which is known to be responsible for binding ATP substrate (Fig. 4). Atomic structures of a number of ATPases show that the ε-amino group of lysine interacts with the β-, γ-, phosphates of ATP, whereas the hydroxyl group of T coordinates with the Mg of the ATP-Mg complex. In the context of these well-established roles, the lethality of "conservative" substitutions such as G165A, K166R and T167A is likely due to a loss of specific and essential interactions between the ATP-Mg substrate and the catalytic center residues rather than to a major perturbation of gp17 structure or folding.[19]

This conclusion was further supported by the consistent biochemical phenotype exhibited by the ATP binding site I mutants. The purified K166G gp17 showed a complete loss of

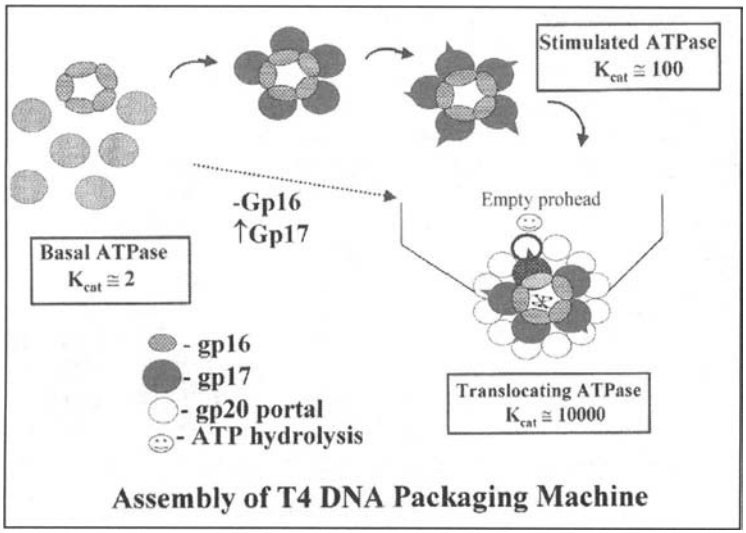


Figure 5. Model for the assembly of a functional packaging machine. The small T4 terminase protein gp16 exists as eight-subunit single or sixteen subunit double ring multimers in solution (see Fig. 2), whereas the large terminase protein gp17 exists largely as a monomer. The stoichiometry of the terminase subunits in the holoenzyme is unknown. In the speculative model depicted above, gp16 and gp17 are shown as pentamers. The main point made is that gp16 multimer interacts with gp17 subunits facilitating the formation of a holoenzyme complex. This results in a conformational change in gp17 converting a weak gp17-ATPase into a stimulated ATPase. Further stimulation to a translocating ATPase is proposed upon docking of the holoenzyme complex to the portal ring thus assembling a functional packaging machine.

DNA packaging activity and the gp16-stimulated ATPase activity. The conservative mutants G165A, K166R, and T167A, showed a loss of in vitro DNA packaging activity but not the terminase (DNA cutting) activity. Thus, the two major functions of gp17 can be separated leading to the inference that the ATPase I site is required for the DNA packaging function. Recent data further show that the mutants in the proposed downstream Walker-B ($MIYID_{251-255}$) residues, which are also predicted to be part of the same ATPase center, exhibited the same phenotypic pattern (M. Mitchell and V. Rao, unpublished data).

Thus, the evidence strongly supports the hypothesis that the T4 large terminase protein gp17 provides the ATPase motor function for the DNA packaging machine. In fact, a common ATPase center with conserved functional signatures has recently been discovered in the large terminase subunit of numerous dsDNA viruses, suggesting this site as the elusive translocating ATPase.[5] Recent mutational and biochemical evidence on the analogous ATPase site in the large terminase subunit gpA of phage λ also corroborates this conclusion.[62] However, more direct evidence is necessary in order to assign a causal role. It can be argued that DNA translocation could be catalyzed by a second "cryptic" ATPase that is activated upon the assembly of the complete packaging machine. Although unlikely, this possibility cannot be ruled out by the available evidence. Biochemical analyses of the conditionally lethal mutants are currently underway to rigorously assess the proposed linkage.

Based on the evidence presented above, a simple model for the assembly of the DNA translocating ATPase is proposed in Figure 5. This model invokes a trigger (stimulation) of the ATPase motor when it is connected to DNA translocation.[12,19] Analysis of a number of packaging systems suggests that the rate of ATP consumption during DNA packaging is substantially

greater than would be predicted from the k_{cat} of the ATPase associated with the large terminase subunit. Assuming that the large terminase subunit alone is the translocating ATPase, this suggests a coupling of ATPase catalysis to DNA packaging, resulting in an increase in the catalytic capacity of the ATPase upon the assembly of a functional packaging machine and ensuing of DNA translocation. Perhaps the best evidence for the stimulated ATPase state is obtained in the T4 terminase system. As discussed above, the gp17-associated ATPase and DNA packaging activities are stimulated by about 50-fold in the presence of the small terminase subunit gp16.[12,19] Considering the evidence that gp17 is hetero-disperse in solution and gp16 forms multimeric rings (or lock washers), it appears likely that the stably assembled gp16 facilitates multimerization of gp17 to form a gp16-gp17 holo-terminase complex. These interactions followed by conformational changes in gp17 result in the stimulation of the ATPase activity. In fact, gp16 dependent gp17 ATPase can be stimulated by an antiserum prepared against denatured gp17 to an activated enzymatic form hydrolyzing ~400 ATPs/gp17/min, an activity consistent with estimates of ATP hydrolysis requirements for DNA translocation (5,000 ATP/min), if gp17 forms an active ~12mer (~4800 ATP/min). And, in fact, it appears that this antibody-activated gp17 is a high molecular weight complex of roughly this dimension that does not require continued gp16 interaction for high turnover ATPase.[74] Stimulation must occur also upon docking of the terminase-DNA complex with the gp20 portal since packaging in vitro (and in vivo) can occur in the absence of gp16. The proposed stimulations may represent common regulatory switches that couple ATP hydrolysis to a biological function in many systems.[12,19]

Interactions of DNA Packaging with Other DNA Processes

Packaging in vivo must interact with numerous other DNA processes, suggesting that regulatory mechanisms governing these interactions are necessary. In fact, in several phages initiation of DNA packaging requires terminase interactions with other proteins, which help to control the process or facilitate it. For example, in phage λ, the host coded proteins IHF and HU apparently bend DNA to promote terminase binding.[3] Phage T4 generally operates relatively independently of host components thus suggesting terminase phage protein interactions may predominate. In fact, phage T4 terminase interacts with numerous phage proteins. In addition to interacting with itself, the small terminase subunit (gp16), and the portal (gp20), the phage T4 terminase gp17 subunit is known to interact with a number of phage T4 DNA binding proteins, namely gp32 (single strand DNA binding protein), gp55 (late T4 σ factor), as well as possibly gp45 (DNA sliding clamp[25]). The terminase interaction with the late T4 σ-factor is intriguing in view of the participation of the T7/T3 RNA polymerase in DNA packaging and DNA repair synthesis. Since gp17 shows little or no affinity for DNA, these interactions with DNA directed proteins may serve to promote terminase loading onto DNA for packaging as well as to regulate packaging initiation. Additionally, it is probable that the DNA replication machinery should be mobilized to the site of packaging initiation in order to repair the concatemer following the double strand DNA break required for DNA packaging. It is tempting to speculate that the terminase is a criticial regulatory protein whose structure and activity depend upon these multiple protein interactions.

DNA Structural Requirements for Packaging

Probably the best-understood example of DNA structure representing a roadblock to packaging arises in the case of T4 gene *49* (endonuclease VII) mutations. T4 gene *49* mutations lead to the accumulation of partially filled heads, apparently because of the accumulation of branches (recombinational intermediates) which can be visualized outside the partially filled head.[63,64] The gene *49* endonuclease is able to resolve Holliday structures in DNA as well as display other activities consistent with a role in removing such branches in the concatemeric

DNA. The phenotype of temperature-sensitive gene *49* mutations therefore demonstrates that DNA damage can arrest DNA packaging, and DNA repair processes can rescue arrested intermediates. Interestingly, the gene *49* endonuclease binds to the gp20 portal protein, positioning it favorably for resolution of branched structures as these encounter the prohead.[65]

DNA ligase is similarly required to complete DNA filling of T4 proheads in vivo. When both T4 and host DNA ligase temperature-sensitive mutant enzymes are inactivated, following the accumulation of a large DNA concatemer pool in the presence of ligase, full heads accumulate to only about 3% of the total. Restoration of ligase activity by temperature shift allows the major product, partially filled heads, to be filled to active heads.[66] These experiments demonstrate that either the T4 or the *E. coli* ligase can satisfy the packaging requirement. Whether the ligase lesion is a nick or more extensive discontinuity in duplex DNA is unknown, although the duplex concatemer must be interrupted to allow late transcription.[66] In phage T4 a number of "early" DNA replication-recombination functions, when altered by mutation, lead to defects in packaging as revealed by the accumulation of partially filled heads as the predominant product, suggesting that defects in the concatemer block continuation of DNA translocation.[67] On the other hand, heteroduplex loops of 19 bases can be packaged into λ heads, a surprising degree of deviation from dsDNA for models that gear dsDNA translocation into the head.[68] Can a single strand extension of a translocating double strand also be translocated into the prohead? Greater knowledge of the DNA structural requirements for packaging would provide necessary checks on models for the translocation mechanism.

Discontinuous Headful Packaging

An interesting twist to the classical headful packaging model has been discovered during the large DNA cloning experiments in phage T4 and phage P1 in vitro packaging.[69,70] Both the cloning systems package linear headful size foreign DNA molecules cloned within a unique site of the vector followed by transduction of these molecules into *E. coli*. Surprisingly, a fraction of the clones showed very small, less than headful length, inserts. Subsequent experiments using only the 7-30 kb size vector DNAs as the packaging substrate revealed that these phages can indeed package and transduce DNA molecules that are much smaller than the headful length DNA. But, these particles have the same density as the wild type suggesting that multiple plasmid molecules are packaged within the same head. Alternate mechanisms such as recombination or rolling circle replication that might convert the less than headful length DNA to a headful size have been eliminated.[70] These results have some interesting implications: (i) the packaging machinery does not have a "ruler" capability to measure the length of the packaging substrate before initiation of packaging and discriminate against packaging of small molecules; (ii) the partially packaged, presumably expanded, head is capable of initiating and packaging a second DNA molecule; and (iii) ds DNA packaging is a remarkably flexible process, and can occur by multiple modes; normally, the unit lengths are cut from concatemers (in vivo), occasionally, unit length linear molecules are packaged from end-to-end (both in vitro and in vivo), and rarely, multiple molecules are packaged to fill the head volume.

Conclusions

Phage T4 fundamentally employs the same mechanisms as the other well characterized dsDNA bacteriophages for packaging DNA. However, it also presents some unique challenges. For instance, the mechanism of DNA substrate recognition and regulation of concatemer maturation and cutting represents one such problem. This is complicated by the fact that the T4 packaging substrate is an "endless" concatemer with multiple branches that are associated with protein complexes. T4 may exploit multiple modes of substrate recognition and cutting, but a close collaboration between packaging and DNA recombination and repair processes is likely

essential to repair the packaging concatemer and to resolve competing DNA packaging complexes. The fact that T4 terminase associates with components (gp32, gp55) of the late transcription and replication pathways is direct evidence for probable coupling of packaging and DNA repair processes. Understanding the dynamics of these interactions and elucidating an integrated mechanism may generate novel insights into the overall mechanisms of regulation of DNA metabolic pathways.

With its similarity to the putative herpes viral terminases and the presence of a single critical ATPase center (other terminases reportedly have two or more ATPase sites[18,71,72]), the proposed ATPase motor in the large terminase subunit gp17 offers an excellent model to elucidate the DNA packaging mechanism. The collection of unique mutants generated by the powerful molecular genetic and biochemical approaches, including the very rare conditionally lethal mutants, offer a unique resource to dissect the molecular details of ATP energy transduction into mechanical movement of DNA. The data may also have broad implications to the general understanding of energy and signal transduction mechanisms. The ATP (GTP) consensus motif is apparently one of the most common motifs found in genomes (up to 5-10% of all expressed proteins), and may represent one of the fundamental (and ancient) motifs in biological systems.[73] Numerous ATPase (GTPase) systems and molecular motors use this motif to trap the energy-rich ATP and regulate and couple its hydrolysis to power numerous biological processes. Biochemical analyses of the mutant collection of gp17 may provide a more precise understanding of the common molecular switches involved in energy/signal regulation and coupling.

Acknowledgements

The research in VBR's laboratory is supported by National Science Foundation (MCB-0110574) and that in LWB's laboratory is supported by National Institutes of Health (AI11676).

References

1. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21:319-331.
2. Black LW. DNA packaging in dsDNA bacteriophages. Annu Rev Microbiol 1989; 43:267-292.
3. Catalano CE, Cue D, Feiss M. Virus DNA packaging: The strategy used by phage lambda. Mol Microbiol 1995; 16:1075-1086.
4. Black LW, Showe MK, Steven AC. Morphogenesis of the T4 head. In: Karam JD, ed. Molecular biology of bacteriophage T4. Washington, DC: ASM Press, 1994:218-258.
5. Mitchell M, Matsuzaki S, Imai S et al. 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 2002; 30:4009-4021.
6. Smith MCM, Burns RN, Wilson SE et al. The complete genome sequence of the streptomyces temperate phage ϕC31: Evolutionary relationships to other viruses. Nucleic Acids Res 1999; 27(No. 10):2145-2155.
7. Davison AJ. Channel catfish virus: A new type of herpesvirus. Virology 1992; 186:9-14.
8. Rao VB, Black LW. Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4: Construction of a defined in vitro packaging system using purified terminase proteins. J Mol Biol 1988; 200:475-488.
9. Powell D, Franklin J, Arisaka F et al. Bacteriophage T4 DNA packaging genes 16 and 17. Nucl Acids Res 1990; 18:4005.
10. Wunderli H, van den Brock J, Kellenberger E. Studies related to the head-maturation pathway of bacteriophage T4 and T2. I. Morphology and kinetics of intracellular particles produced by mutants in the maturation genes. J Supramol Struct 1977; 7:135-161.

11. Lin H, Simon MN, Black LW. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. J Biol Chem 1997; 272:3495-3501.
12. Leffers G, Rao VB. Biochemical characterization of an ATPase activity associated with the large DNA packaging unit gp17 from bacteriophage T4. J Biol Chem 2000; 275:37127-37136.
13. Yang Q, Berton N, Manning MC et al. Domain structure of gpNuI, a phage lambda DNA packaging protein. Biochemistry 1999; 38:14238-14247.
14. Franklin J, Mosig G. Expression of the bacteriophage T4 DNA terminase genes *16* and *17* yields multiple proteins. Gene 1996; 177:179-189.
15. Becker A, Gold M. Prediction of an ATP reactive center in the small subunit, gpNuI, of the phage lambda terminase enzyme. J Mol Biol 1988; 199:219-222.
16. de Beer T, Fang J, Ortega M et al. Insights into specific DNA recognition during the assembly of a viral genome packaging machine. Molecular Cell 2002; 9:981-991.
17. Via A, Ferre F, Brannetti B et al. Three-dimensional view of the surface motif associated with the P-loop structure: Cis and trans cases of convergent evolution. J Mol Biol 2000; 303:455-465.
18. Guo P, Peterson C, Anderson D. The DNA packaging protein gp16 of bacteriophage ϕ29 is a prohead- and DNA-gp3-dependent ATPase. J Mol Biol 1987; 197:229-236.
19. Rao VB, Mitchell MS. The N-terminal ATPase site in the large terminase protein gp17 is critically required for DNA packaging in bacteriophage T4. J Mol Biol 2001; 314:401-411.
20. Franklin J, Haseltine D, Davenport L et al. The largest (70 kDa) product of the bacteriophage T4 DNA terminase gene 17 binds to single stranded DNA segments and digests them towards junctions with double-stranded DNA. J Mol Biol 1998; 277:541-557.
21. Kuebler D, Rao VB. Functional analysis of the DNA-packaging/terminase protein gp17 from bacteriophage T4. J Mol Biol 1998; 281:803-814.
22. Rentas F. Molecular genetic and biochemical analysis of a putative terminase cutting site in the large DNA packaging protein gp17 from bacteriophage T4. PhD thesis. Washington, DC: The Catholic University of America, 2001.
23. Gerit JA. Mechanistic principles of enzyme catalyzed cleavage of phosphodiester bonds in nucleases. 2nd ed. NY: Cold Spring Harbor Press, 1993:1-34.
24. Lin H, Rao VB, Black LW. Analysis of capsid portal protein and terminase functional domains: Interaction Sites Required for DNA Packaging in Bacteriophage T4. J Mol Biol 1999; 289:249-260.
25. Malys N, Chang D-Y, Baumann RG et al. A bipartite bacteriophage T4 SOC and HOC randomized peptide display library: Detection and analysis of phage T4 terminase (gp17) and late σ factor (gp55) interaction. J Mol Biol 2002; 319:289-304.
26. Zhang Z, Sumbilla C, Lewis D et al. Mutational analysis of the peptide segment linking phosphorylation and Ca(2+)-binding domains in the sarcoplasmic reticulum Ca(2+)-ATPase. J Biol Chem 1995; 270:16283-16290.
27. Louie D, Serwer P. Blunt-ended ligation can be used to produce DNA ladders with rung spacing as large as 0.17 Mb, Nucleic Acids Res 1990; 18:3090.
28. Bhattacharyya SP, Rao VB. A novel terminase activity associated with the DNA packaging protein gp17 of bacteriophage T4. Virology 1993; 196:34-44.
29. Bhattacharyya SP, Rao VB. Structural analysis of DNA cleaved in vivo by bacteriophage T4 terminase. Gene 1994; 146:67-72.
30. Baumann RG. The role of portal and terminase proteins in the mechanism of DNA packaging in the dsDNA bacteriophage T4. PhD thesis. Baltimore, Md: University of Maryland Med Sch, 2002.
31. Hsiao CL, Black LW. Head morphogenesis of bacteriophage T4. III. The role of gene 20 in DNA packaging. Virology 1978; 91:26-38.
32. Hsiao CL, Black LW. DNA packaging and the pathway of bacteriophage T4 head assembly. Proc Natl Acad Sci USA 1977; 74:3652-3656.
33. Driedonks RA, Engel A, ten Heggeler B et al. Gene 20 product of bacteriophage T4 its purification and structure. J Mol Biol 1981; 152:641-662.
34. Valpuesta JM, Carrascosa JL. Structure of viral connectors and their function in bacteriphage assembly and DNA packaging. Q Rev Biophys 1994; 27:107-155.
35. Hsiao CL, Black LW. Head morphogenesis of bacteriophage T4. II. The role of gene 40 in initiating prehead assembly. Virology 1978; 91:15-25.

36. Michaud G, Zachary A, Rao VB et al. Membrane associated assembly of a phage T4 DNA entrance vertex structure studied with expression vectors. J Mol Biol 1989; 209:667-681.
37. Yap N, Rao VB. Novel mutants in the 5'-upstream region of the portal protein gene 20 overcome a gp40-dependent prohead assembly block in bacteriophage T4. J Mol Biol 1996; 263:539-550.
38. Simpson AA, Tao Y, Leiman PG et al. Structure of the bacteriophage ϕ29 DNA packaging motor. Nature 2000; 408:745-750.
39. Wu DG, Black LW. Gene amplification mechanism for the hyperproduction of T4 bacteriophage 17 and 18 proteins. J Mol Biol 1987; 195:769-783.
40. Wu DG, Wu CH, Black LW. Reiterated gene amplifications at specific short homology sequences in phage T4 produce Hp17 mutants. J Mol Biol 1991; 218:705-721.
41. Wu CHH, Lin H, Black LW. Bacteriophage T4 gene 17 amplification mutants: Evidence for initiation by the T4 terminase subunit gp16. J Mol Biol 1995; 247:523-528.
42. Lin H, Black LW. DNA requirements in vivo for phage T4 packaging. Virology 1998; 242:118-127.
43. Wu CHH, Black LW. Mutational analysis of the sequence-specific recombination box for amplification of gene 17 of bacteriophage T4. J Mol Biol 1995; 247:604-617.
44. Black LW. DNA packaging and cutting by phage terminases: Control in phage T4 by a synaptic mechanism. Bioessays 1995; 17:1052-1060.
45. Grossi GF, Macchiato MF, Gialanella G. Circular permutation analysis of phage T4 DNA by electron microscopy. Z Naturforsch 1983; 38:294-296.
46. Carrascosa JL. Head maturation pathway of bacteriophage T4 and T2. IV. In vitro transformation of T4 head-related particles produced by mutants in gene *17* to capsid-like structures. J Virol 1978; 26:420-428.
47. Steven AC, Greenstone H, Bauer AC et al. The maturation-dependent conformational change of the major capsid protein of bacteriophage T4 involves a substantial change in secondary structure. Biochemistry 1990; 29:5556-5561.
48. Rao VB, Black LW. Evidence that a phage T4 DNA packaging enzyme is a processed form of the major capsid gene product. Cell 1985; 185:565-578.
49. Xue MQ, Black LW. Role of the major capsid protein of phage T4 in DNA packaging from structure-function and site-directed mutagenesis studies. J Struct Biol 1990; 104:75-83.
50. Galisteo ML, King J. Conformational transitions in the protein lattice of phage P22 capsids. Biophys J 1993; 65:227-235.
51. Steven AC. Conformational change–an alternative energy source? Exothermic phase transition in phage capsid maturation. Biophys J 1993; 65:5-6.
52. Rao VB, Black LW. DNA packaging of bacteriophage T4 proheads in vitro: Evidence that prohead expansion is not coupled to DNA packaging. J Mol Biol 1985; 185:565-578.
53. Jardine PJ, McCormick MC, Lutze-Wallace C et al. The bacteriophage T4 DNA packaging apparatus targets the unexpanded prohead. J Mol Biol 1998; 284:647-659.
54. Jardine PJ, Coombs DH. Capsid expansion follows the initiation of DNA packaging in bacteriophage T4. J Mol Biol 1998; 284:661-672.
55. Hohn B. DNA sequences necessary for packaging of bacteriophage lambda DNA. Proc Natl Acad Sci USA 1983; 80:7456-7460.
56. Shibata H, Fujisawa H, Minagawa T. Characterization of the bacteriophage T3 DNA packaging reaction in vitro in a defined system. J Mol Biol 1987; 196:845-51.
57. Son, Watson RH, Serwer P. The direction and rate of bacteriophage T7 DNA packaging in vitro. Virology 1993; 196:282-289.
58. Steven AC, Bauer AC, Bisher ME et al. The maturation-dependent conformational change of phage T4 capsid involves the translocation of specific epitopes between the inner and outer capsid surfaces. J Struct Biol 1991; 106:221-236.
59. Mullaney J, Black LW. Activity of foreign proteins targeted within the bacteriophage T4 head and prohead: Implications for packaged DNA structure. J Mol Biol 1998; 283:913-929.
60. Morita M, Tasaka M, Fujisawa H. DNA packaging ATPase of bacteriophage T3. Virology 1993; 193:748-752.
61. Guo P, Peterson C, Anderson D. Initiation events in in-vitro packaging of bacteriophage phi 29 DNA-gp3. J Mol Biol 1987; 197:219-228.

62. Duffy C, Feiss M. The large subunit of bacteriophage lambda's terminase plays a role in DNA translocation and packaging termination. J Mol Biol 2002; 316:547-561.
63. Luftig RB, Wood WB, Okinaka R. Bacteriophage T4 head morphogenesis. On the nature of gene 49 defective heads and their role as intermediates. J Mol Biol 1971; 57:555-573.
64. Kemper B, Brown DT. Function of gene 49 of bacteriophage T4. II. Analysis of intracellular development and the structure of very fast-sedimenting DNA. J Virol 1976; 18:1000-1015.
65. Golz S, Kemper B. Association of holliday-structure resolving endonuclease VII with gp20 from the packaging machine of phage T4. J Mol Biol 1999; 285:1131-1144.
66. Zachary A, Black LW. DNA ligase is required for encapsidation of bacteriophage T4 DNA. J Mol Biol 1981; 149:641-658.
67. Zachary A, Black LW. Topoisomerase II and other DNA-delay and DNA-arrest mutations impair bactreriophage T4 DNA packaging in vivo and in vitro. J Virol 1986; 60:97-104.
68. Pearson RK, Fox MS. Effects of DNA heterologies on bacteriophage λ packaging. Genetics 1988; 118:5-12.
69. Coren J, Pierce J, Sternberg N. Headful packaging revisited: The packaging of more than one DNA molecule into a bacteriophage P1 head. J Mol Biol 1995; 249:176-184.
70. Leffers G, Rao V. A discontinuous headful packaging model for packaging of less than headful length DNA molecules by bacteriophage T4. J Mol Biol 1996; 258:839-850.
71. Hang JQ, Tack BF, Feiss M. ATPase center of bacteriophage lambda terminase involved in post-cleavage stages of DNA packaging: Identification of ATP-interacting amino acids. J Mol Biol 2000; 302:777-795.
72. Morita M, Tasaka M, Fujisawa H. Analysis of functional domains of the packaging proteins of bacteriophage T3 by site-directed mutagenesis. J Mol Biol 1994; 235:248-259.
73. Koonin EV, Tatusov RL, Rudd KE. Sequence similarity analysis of Escherichia coli proteins: Functional and evolutionary implications. Proc Natl Acad Sci USA 1995; 92:11921-11925.
74. Baumann RG, Black LW. Isolation and characterization of T4 bacteriophage gp17 terminase: A large subunit multimer with enhanced ATPase activity. J Biol Chem 2003; 278:4618-4627.

CHAPTER 4

T3/T7 DNA Packaging

Philip Serwer

Summary

During formation of a mature bacteriophage particle, a procapsid of protein packages the linear double-stranded DNA genome of the related bacteriophages, T3 and T7. Initiation of T3/T7 DNA packaging in vivo occurs near the genetic right end of a concatemer-associated genome. Initiation in vivo requires transcription near the initiation site and results in the formation of the genome's right end by cleavage of the concatemer. The initiation is co-operative among capsids that are packaging the genomes of a single concatemer, based on observations of single in vitro DNA packaging events. An ATP fueled DNA packaging motor then drives the mature DNA genome into a capsid. Genetic analysis of a T3 DNA packaging ATPase reveals a capsid-binding site and also ATP binding sites that appear to be part of the DNA packaging motor. The packaged T7 DNA molecule is wound in concentric coils around an internal protein cylinder. The cylinder has a strong component of 8-fold rotational symmetry and is mounted on a 12-fold symmetric connector that sits on a 5-fold vertex of the procapsid. The DNA molecule enters the capsid through an axial hole in the cylinder-connector. The procapsid has recently been isolated in an altered state. In vitro T3 DNA packaging has been achieved in a purified system that has been used to define the following interaction among components: connector-large accessory protein (ATPase)-small accessory protein-DNA. Procedures are being developed for analysis of the cycle of the T3/T7 DNA packaging motor at the level of a single motor.

Introduction

Historically, research on bacteriophages began with an agenda that had a focus on bacteriophage-based cures of bacterial diseases. Bacteriophage T7 was isolated during propagation of a mixed culture originally used for bacteriophage therapy. T7 outgrew the other bacteriophages present in the culture, although a mixed culture was intended.[1] Recently, the research agenda has begun to refocus on bacteriophage-based cures, because of the increasing resistance of bacteria to known chemical antibiotic-based cures.[1-3] After the initial focus on bacteriophages-as-antibacterials, the bacteriophage research agenda developed a focus on the chemistry of biological information transfer.[4-6] This latter research was assisted by the low time of the generation of bacteriophages. For example, bacteriophages T3 and T7 infect a host and form over 100 progeny bacteriophages in a total time of about 13 minutes at 37°C. This low generation time was ideal for genetic studies.

After the fundamental events of biological information transfer were discovered, the bacteriophage research agenda developed a focus on the biochemical/biophysical mechanisms

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

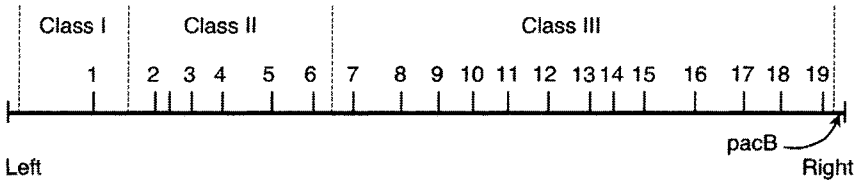
involved in these events. The primary events of information transfer were DNA replication, DNA recombination, RNA synthesis and protein synthesis. However, the following secondary event also attracted attention, in part because of its apparent susceptibility to empirical investigation: the packaging of a double-stranded DNA genome in the protein capsid of a bacteriophage particle. This "DNA packaging" occurs by assembly of a procapsid that subsequently binds a bacteriophage genome and, then, draws the bacteriophage genome into a cavity surrounded by the outer shell of the capsid. The energy for DNA packaging is derived from cleavage of ATP. The capsid's outer shell changes its structure during DNA packaging (reviewed in refs. 7-9).

For some bacteriophages, including T3/T7, the in vivo DNA substrate for packaging is an end-to-end polymer (concatemer) of the mature genome. The concatemer is cut to mature size during packaging. A pathway proposed for in vivo T7 DNA packaging is in Figure 1 (discussed further below). The length of the mature genome is 39.937 kilobase pairs (kb) for T7[10] and 38.208 kb for T3.[11] The maturation cleavage of a concatemer produces a genome that has the nucleotide sequence at one end repeated at the other end (terminally repetitious). The nucleotide sequence is unique (nonpermuted). The terminal repeat is 160 nucleotide pairs long for T7[10] and 231 nucleotide pairs long for T3.[11] Even without the complexity of producing a terminally repetitious, nonpermuted genome, DNA packaging appeared to have a level of biochemical complexity at least equal to that of other fundamental events. But, DNA packaging also appeared to have more physical simplicity because the DNA molecule did not undergo either strand separation or strand exchange during DNA packaging. The primary change in the DNA molecule during packaging is its dramatic collapse in a cavity that has a volume about twice the volume of the DNA double helix packaged.[12,13]

In the case of bacteriophages, studies of fundamental events were assisted by development of bacteriophage genetics. For T7, these studies initially produced (1) a map of genes, (2) conditional-lethal mutations in each of most essential genes, and (3) deletion mutations in each of most nonessential genes.[14-16] The map for T3 is close to the map for T7.[11] The T3/T7 nucleotide sequence homology is extensive; sequences as long as 177 are identical.[11] The system for naming genes is the same for T3 as it is for T7.[11,17] The genes are named by numbering from one end of the genome to the other. The left end is the first end injected into a host cell at the beginning of an infection.[18,19] Therefore, this end has genes with the lowest number. Newly discovered genes receive non-integral numbers, if they are between genes that already have neighboring integral numbers (for example, gene 2.5 in Table 1). Conditional-lethal mutants are used to help determine the function of a gene. For example, infection of a non-permissive host with some amber mutants results in no DNA packaging, but the production of all easily visible components needed for DNA packaging. This type of result was used to discover the accessory proteins for DNA packaging that are discussed below.

For analysis of DNA packaging, each of the various double-stranded DNA bacteriophages has advantages and disadvantages. An advantage of T3/T7 is the presence of the following two characteristics that assist in the isolation and characterization of particles in the DNA packaging pathway (DNA packaging intermediates): (1) Both monomeric and concatemeric DNA genomes are short enough so that hydrodynamic shear-induced breakage is not a major problem during pipeting. The longer genome of bacteriophage T4, for example, is more of a problem in this area. (2) During infection, the host genome is degraded, host DNA synthesis is stopped and host protein synthesis is stopped (reviewed in ref. 20). The second characteristic is shared only by bacteriophages in the T7-like and T4-like families, among the well-studied bacteriophages.

Table 1. The T3/T7 genome. An abridged list of the T3/T7 genes is presented. The pacB site is indicated.[11,14,20]



1 RNA Polymerase	8 Connector protein
2 Anti-host RNA Polymerase	9 Scaffolding protein
2.5 ssDNA-binding protein	10 Major outer shell protein
3 Endonuclease	11,12 Tail proteins
4 Primase-helicase	13-16 Proteins of the internal cylinder
5 DNA polymerase	17 Tail fiber protein
6 5' to 3' exonuclease	18 Small accessory protein for DNA packaging
7 ?	19 Large accessory protein for DNA Packaging

The usefulness of these two characteristics is illustrated by the early isolation from infected cells of hard to isolate T7 DNA packaging intermediates. These intermediates had a capsid bound near the right end of a T7 DNA molecule that was sometimes monomeric and sometimes concatemeric. These intermediates were explained by initiation of packaging near the right end of the T7 genome,[21] as illustrated in Figure 1. The first section below will review studies in which genetic/DNA sequence analysis was pursued beyond determination of the effects of mutations that completely inactivate the product of a gene. This genetic analysis has precisely defined the initiation site near the right end of both the T3 and the T7 genomes.

The isolation of DNA packaging intermediates raises the question of how the DNA packaging intermediates progress in the DNA packaging pathway. To answer this question, the structure of the intermediates must be determined at a level that provides constraints for models that describe the dynamics. Specifically, an ATP-dependent motor drives the entry of a bacteriophage DNA molecule into a procapsid, in the case of both T3/T7[22,23] and other bacteriophages (reviewed in refs 9, 24). Understanding the mechanisms of this motor will require knowledge of the structure of the parts of both the capsid and its accessory proteins. The second section below will review the current state of knowledge of the structure of T3/T7 procapsids and mature capsids.

Finally, for the following reasons, complete analysis of the DNA packaging pathway requires cell-free, in vitro systems that package T3/T7 DNA: (a) Some analysis requires control of the pathway beyond what can be achieved by genetic manipulation of in vivo DNA packaging. The components of in vitro systems can be independently varied at will. (b) The DNA packaging motor has multiple functional groups linked together in a single particle. The groups must become active in a defined order during each cycle of the motor. This order becomes obscured by asynchrony of motors during conventional biochemical analysis. The motor does not work if the functional groups are separated from each other. Therefore, the motor cannot be analyzed by the conventional reductionist techniques of solution biochemistry. Observation of DNA packaging at the level of single DNA molecules solves this problem. Single-particle (molecule) analysis of DNA packaging requires an in vitro system, at least at present. The third section below will review some aspects of the current state of knowledge obtained from in vitro T3/T7 DNA packaging.

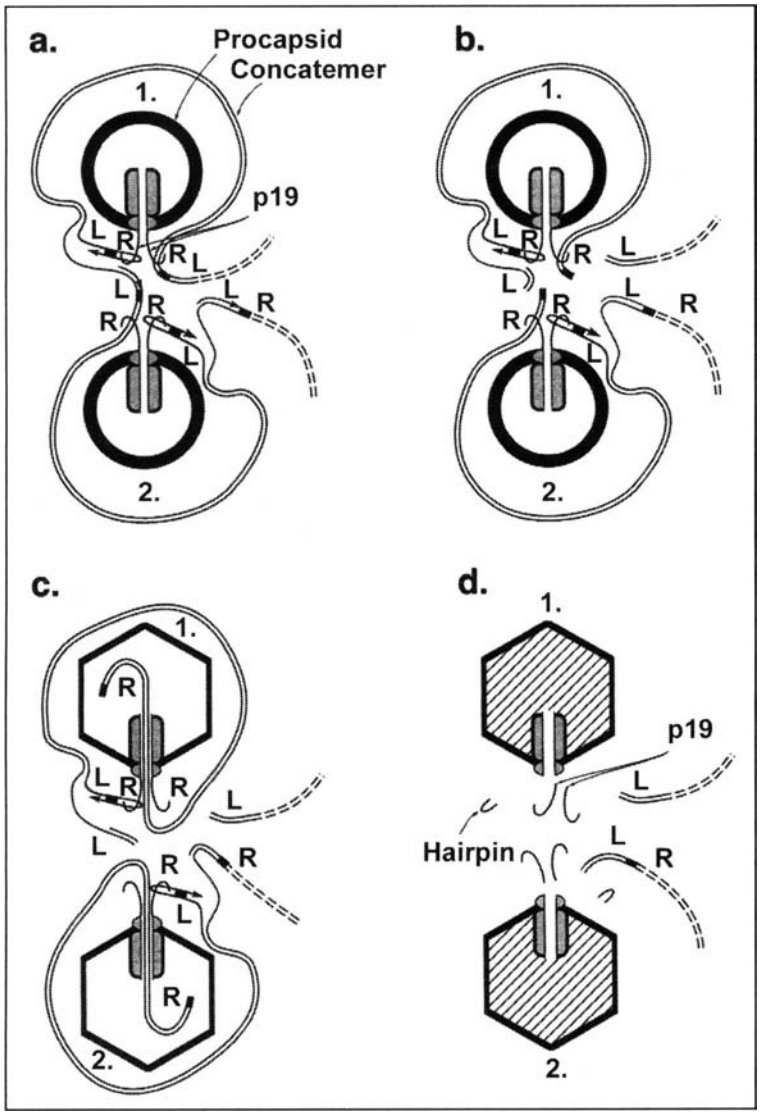


Figure 1. A simplified version of the T3/T7 DNA packaging pathway in vivo. a) Two procapsids (labeled 1 and 2) bind a concatemer at two identical sites. Each site is a pacB site near the right end of each of two successive genomes in a concatemer. A hairpin-primed replicative branch selectively duplicates the terminal repeat. An arrow within a DNA strand indicates replication of this strand. One copy of the terminal repeat is single-stranded until lagging strand synthesis occurs. Filling the space between the two strands of a DNA double helix indicates the terminal repeat in a double-stranded region. For visibility, the terminal repeat is drawn larger than it is in reality. b) The first of two cleavages occurs to initiate DNA packaging. c) A freshly cleaved right end enters the procapsid while the procapsid (capsid I) converts to a larger, more angular capsid (capsid II). d) A DNA molecule both finishes entry and undergoes a terminal cleavage near the hairpin formed by replication. L, left; R, right end of a T7 genome. Additional capsids are usually present beyond the two shown (adapted from ref. 100). In Figure 1, only one of the two accessory proteins (gp19) is shown for simplicity. Some details of the capsid structure have been simplified. A more detailed view of both the accessory proteins and the components of the capsid is in Figures 2 and 3b.

Genetics/Genome Sequence

Genome Sequence and Evolution

Complete nucleotide sequences have been determined for both the T7 genome[10] and the T3 genome.[11] These two bacteriophages appear to be representatives of a much larger family of bacteriophages. For example, both Vibrio parahaemolyticus bacteriophage VpV262[25] and *Roseophage* SIO1[26] have distant homology to T7, seen primarily in the nucleotide sequence and genetic organization of genes encoding proteins that participate in morphogenesis. These studies found no T3/T7-like RNA polymerase in either VpV262 or SIO1. In the study of VpV262, this absence is interpreted to mean that these two bacteriophages are descended from bacteriophage ancestors that existed before T3/T7-like RNA polymerases evolved. These ancestors would have existed before the emergence of the ancestors of *Cyanophage* P60, the most divergent of the T7-like bacteriophages that encode a T3/T7-like RNA polymerase.[27] Sequence divergence of the P60 RNA polymerase, as well as the time of the appearance of cyanobacteria, are explained by the hypothesis that ancestral T3/T7-like bacteriophages already existed at least 1.6 billion years ago (Hardies S., personal communication).

Homologues of at least one T7 gene are not confined to prokaryotes. Both linear mitochondrial plastids and eukaryotic chromosomes encode genes with homology to T3/T7 RNA polymerase. These genes are thought to have entered eukaryotes during mitochondrial endosymbiosis.[28] Speciation via "punctuated equilibrium"[29] is explained by the assumption that exchange of either genes or gene modules from prokaryotes to eukaryotes did, in fact, occur during the speciation of eukaryotes. The unique handedness of chromosomal DNA molecules is explained by the assumption that genetic exchange occurred among all organisms during at least early evolution.[30]

Assuming that some genes of the T3/T7 ancestral line are as old as they appear to be, the following question is asked: How much of current eukaryotic genomes began as a bacteriophage gene that was subsequently transferred? Does a precursor-product relationship exist between genes required for bacteriophage DNA packaging and genes required for other energy transducing events? In other words, was our convenient experimental subject, the T3/T7-like double-stranded DNA bacteriophage, also a convenient subject for rapidly selecting new genes that were precursors of the genes of higher organisms?

Whatever the history of T3/T7-like bacteriophages, distant relatives (like SIOI and VpV262) have been found to be the most abundant members of two independent ocean-borne bacteriophage communities. This conclusion is drawn from the results of DNA sequencing, without propagating members of the community.[31] Thus, the T3/T7-like bacteriophages appear to be a major factor in the microbial communities of the ocean.

Organization of the Genome

The genes of T7 (and presumably T3) are expressed in the order of injection into a host. An abridged summary of T7 genes (see ref. 14) is in Table 1. T3 genes are identified with the same nomenclature.[11] The leftmost (early or class I) genes are the first genes expressed. Some class I genes (with non-integral numbers, not indicated in Table 1) have protein products that inactivate host proteins, thereby producing an environment favorable for reproduction of bacteriophage particles. The class I genes also include the gene for T3/T7 RNA polymerase (gene 1). After the host RNA polymerase transcribes the gene for bacteriophage RNA polymerase, the T7 (T3) gene 1 RNA polymerase transcribes, for the most part, the remaining viral genes (reviewed in refs. 11, 14).

The next genes injected (class II) encode proteins that participate in the replication of the T3/T7 genome. These include gene 2.5 single-stranded DNA binding protein, gene 3 endonuclease, gene 4 primase-helicase, gene 5 DNA polymerase and gene 6 exonuclease (5' to 3')

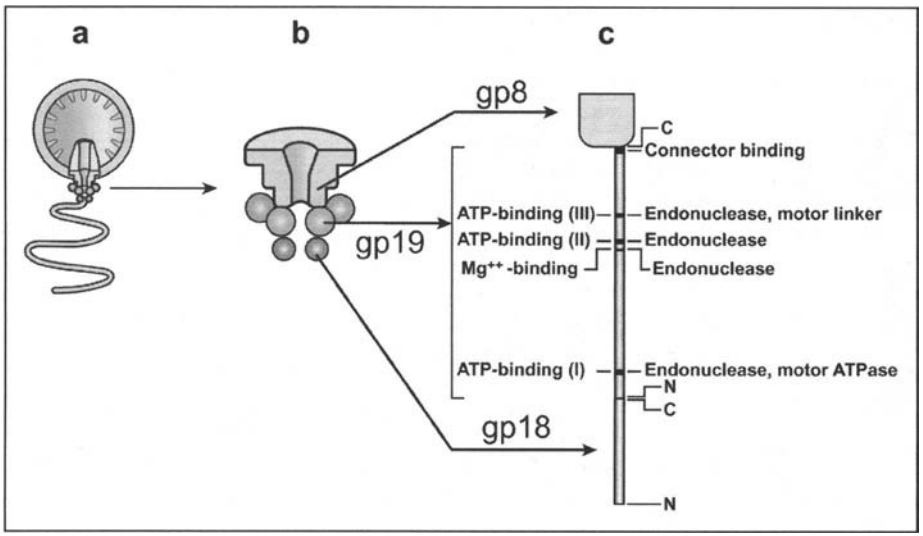


Figure 2. An illustration of the subdivision of an accessory protein. a) A T3/T7 procapsid is shown after binding accessory proteins and monomeric DNA molecule. This complex has been identified during T3 in vitro DNA packaging,[8] as discussed in the text. This in vitro complex is different from the in vivo complex illustrated in Figure 1. A more complete illustration of the in vitro DNA packaging pathway is in Figure 3b. b) The connector and its bound accessory proteins (gp18 and gp19) is removed from (a) and magnified. c) The complex in (b) is further magnified and the accessory proteins are unraveled to form a straight chain. The sites of significance discussed in the text are indicated.

(reviewed in refs. 11, 14). The predominant product of T7 (and presumably T3) DNA replication is a left end-to–right end joined concatemer[32] that has short, left end hairpin-containing branches that probably replicate the terminal repeat.[33] The concatemer will be cut to mature size during DNA packaging (illustrated in Fig. 1).

After the genes for DNA replication are expressed, then the genes (class III) for the proteins of both capsid assembly and DNA packaging are expressed. The name of a protein is the number of the protein's gene, preceded by gp. The proteins encoded by class III genes include (1) the major protein of the outer shell of the procapsid (gp10A and a longer frameshift/read-through product, gp10B; A. Rosenberg and F. W. Studier quoted for T7 in ref. 34), (2) a scaffolding protein (gp9) that assists assembly of gp10A and gp10B, (3) a connector protein (gp8) that forms a dodecameric DNA channel in the procapsid, (4) internal proteins that form a cylinder around which the packaged DNA molecule is wrapped (gp14-16), (5) a comparatively small accessory protein (gp18) that is not needed for procapsid assembly, but is needed for DNA packaging, and (6) a comparatively large accessory protein (gp19) that is not needed for procapsid assembly, but is also needed for DNA packaging. The DNA enters the capsid through a channel in the dodecameric connector during packaging. Details are in the section on structure and Figure 3, below. The identification of proteins is reviewed in references 11, 14, 20 and 35.

In the case of T7, details of both promoters and ribosome binding sites are known. These elements have been used to create both expression vectors[36] and a bacteriophage-display system.[37] Usually, T7 proteins are synthesized in rough proportion to their content in the bacteriophage capsid. For example, the promoter for gene 10 is more efficient than the promoter for other T7 genes. The former promoter has been used for high efficiency expression vectors.[11,36]

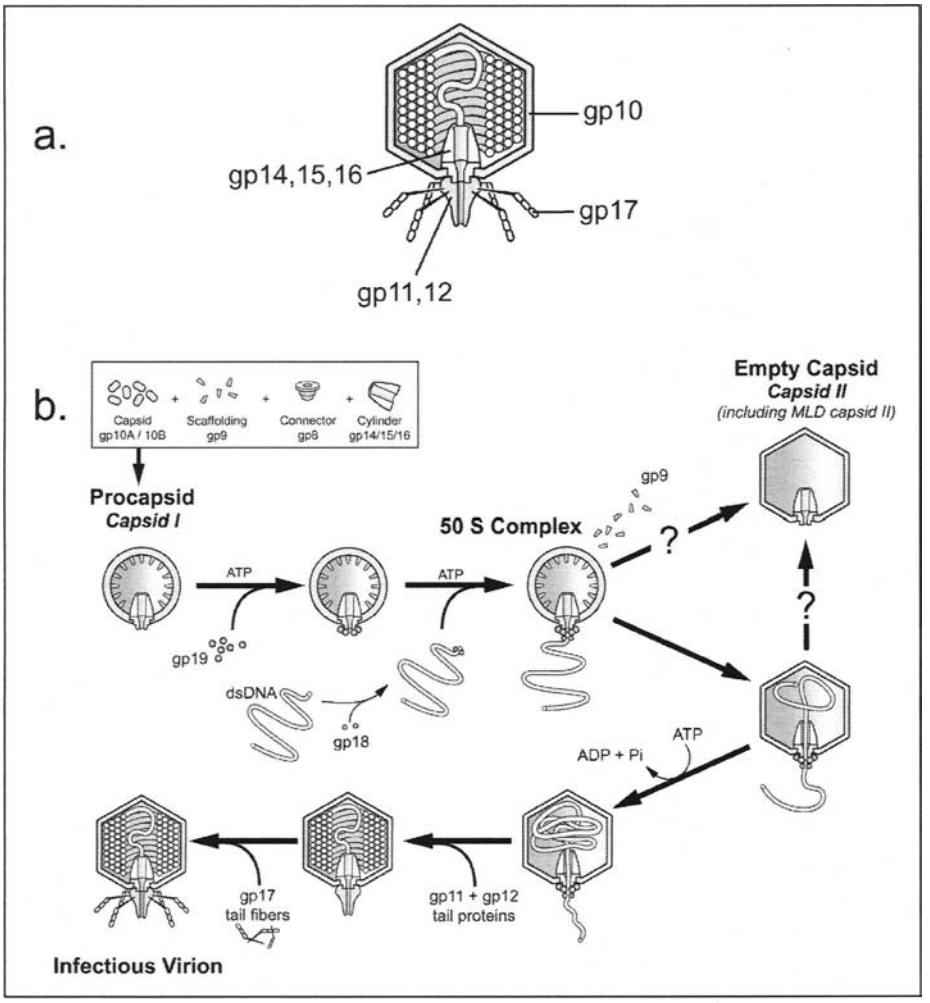


Figure 3. An illustration of T3/T7 capsids and accessory proteins during in vitro DNA packaging. a) The structure of bacteriophage T7 is illustrated. This drawing was adapted from reference 94 with changes to both the connector and the internal cylinder to increase the correspondance with some data; see references 72, 77 and 94. b) The in vitro pathway for packaging of monomeric DNA molecules is sketched. Most aspects of the pathway were obtained from a purified in vitro system. This pathway and the data used to derive it are reviewed in reference 8. The structure of the procapsid is deduced from a series of studies, the last of which are references 83 and 94. This drawing was adapted from reference 94. The connector and cylinder of the procapsid have been modified as described under (a).

More Detailed Genetic Analysis of DNA Packaging: A New Accessory Protein

Genetic analysis can be used to either (a) discover new proteins involved in DNA packaging, as done for gp18 and gp19, or (b) subdivide known DNA packaging genes/proteins by function. In the former category, in vivo analysis of mutant T7 gene 1 (RNA polymerase) produced mutant RNA polymerase that (a) is active in RNA synthesis, but (b) prevents initiation of DNA packaging.[38] This observation is explained by the hypothesis that limited transcription is required for DNA packaging in vivo.

The following observations support this hypothesis: (a) A bacteriophage promoter (also an origin of DNA replication in the case of T7) near the right end of either a T3[39] or a T7[40] genome is part of a site that is necessary for the packaging-dependent transduction of plasmid-associated DNA. (b) This site must be in its native orientation relative to the eventual cleavage sites for packaging the plasmid.[41] This site is called pac binding or pacB. The plasmid-associated presence of T3 pacB in a host cell inhibits growth of T3, but not T7, and vice versa.[42] Furthermore, transcription at the pacB-associated promoter is required for packaging a T3/T7 concatemer-associated genome in an in vitro system of unpurified components.[42] The T3 pacB does not work for T7 and vice versa.[43] Thus, in vivo packaging of an intact genome is assumed to require transcription at pacB. This transcription is omitted in Figure 1 for simplification of the figure. The pacB site identified in the study of T3 begins approximately 733 base pairs from the right genome end (Fig. 5 in ref. 39); the T7 pacB begins between 575 and 709 base pairs from the right genome end.[40]

The transcription phase of T3/T7 DNA packaging initiation can be bypassed both in vivo and in vitro. The basis for this conclusion is that DNA without T3/T7 base sequence homology can be packaged either (1) in vitro in a T7 system that, nonetheless, favors the packaging of concatemers,[44] (2) in vivo when either T7 gene 3 endonuclease (which degrades host DNA in vivo)[44] or T3 gene 3 endonuclease[45] is absent, and (3) in vitro in a purified T3 system that packages monomeric T3 DNA.[45] Transcription is not necessary for packaging monomeric T3 DNA in the latter in vitro system. These data raise the question of why transcription is necessary for in vivo wild-type T3/T7 DNA packaging if so easily bypassed.

In answer to this question, the author presents here the following (unproven) hypothesis: The transcription dependence of the initiation of T3/T7 DNA packaging is for the purpose of controlling the number of genomes packaged. The following are the details: (1) After initiation of DNA packaging, production of viable progeny bacteriophages requires that sufficient ATP be available to drive an entire genome into each of several capsids. One ATP molecule is cleaved for every 1.8 base pairs packaged on average, for T3 during in vitro packaging of monomeric DNA.[46] (2) Sufficient ATP for all possible progeny will not necessarily be present at the end of an infection when the bacterial cell has lost its ATP-generating capacity and has begun to leak.[47,48] (3) Thus, the following assumption seems reasonable: Evolutionary pressure exists to restrict DNA packaging if the supply of ATP is limiting for the number of genomes that can be packaged. For example, assume that the ATP concentration is high enough so that the host cell has enough ATP to package 10 genomes. If initiation is successful for 20 genomes, then 20 procapsids will cleave ATP molecules. In the absence of further control, packaging will continue until 20 DNA molecules are, on average, half-packaged. The ATP is now exhausted. In this case, the odds favor the outcome that no genomes complete packaging. (4) Packaging initiation via limited transcription introduces a threshold that must be crossed before the more ATP-hungry event of DNA packaging starts. Thus, the most packaging-favorable (ATP-rich) regions of the infected cell will initiate packaging most quickly and, by drawing ATP from the pool, further decrease the chance of initiating packaging in other regions. Possibly, transcription is needed to make the threshold sufficiently high and co-operative. This type of regulation is based on group, rather than individual, survival.

Control of which genomes complete packaging is also possible post-initiation of packaging. For example, ATP deficiency might induce a response that includes abortion of some DNA packaging events so that others can go to completion. This hypothesis explains the following observations that are difficult to explain by other means: (a) Premature termination of packaging becomes increasingly more frequent as a T7 infection progresses. The premature termination yields packaged DNA molecules that are shorter than a mature T7 DNA molecule. These shorter-than-mature DNA molecules form bands during gel electrophoresis, but the underlying

cleavage does not occur at a unique nucleotide sequence.[49] Even packaged monomeric T7 DNA molecules have low level single-stranded breaks that cause the formation of bands during gel electrophoresis of single DNA strands from mature DNA molcecules.[50] (3) The endonuclease/accessory protein that specifically cleaves T7 genomes from a concatemer (gp19; see Fig. 1; described in more detail below) cleaves non-specifically in vitro when ATP is removed.[51] Thus, the packaging aborting cleavages are explained by premature endonuclease activity of gp19. By this hypothesis, the premature gp19 cleavage becomes more frequent as the ATP concentration decreases. The low ATP, gp19-induced premature cleavage is a potential mechanism for post-initiation aborting of the packaging of the most vulnerable genomes when the supply of ATP is dwindling.

The initiation of T7 DNA packaging includes gp18-, gp19- and procapsid-dependent specific cleavage of a concatemer to form the genomic right end.[52,53] As previously reviewed and emphasized[8], the specificity of packaging is derived from the specificity of this cleavage. Any blunt ended DNA molecule can be packaged. Packaging of both T3[51] and T7[52] DNA terminates with the formation of the genomic left end. However, the left end is formed by a process that is apparently more complicated and includes the gp6 exonuclease-dependent selective duplication of the sequence terminally repeated in the mature genome.[33,52,54] The in vitro packaging of a T7 concatemer-derived genome occurs in a right-to-left direction, as judged by probing partially-packaged DNA with either right- or left-end specific oligonucleotides. For this analysis, a partially packaged DNA molecule had been either expelled from its capsid or left in the capsid before fractionation by use of gel electrophoresis, followed by probing.[55] The data of this section and other data are explained by the model of in vivo T3/T7 DNA packaging in Figure 1.

More Detailed Genetic Analysis of DNA Packaging: Subdivision of an Accessory Protein

The T3, and presumably T7, accessory proteins have several functions, including (1) binding a procapsid to a concatemer at the beginning of packaging, (2) conducting the initiating and terminating cleavage of concatemer-associated genomes, and (3) driving a DNA packaging motor. The pattern of binding is DNA-gp18-gp19-connector of the procapsid. This pattern is shown for a procapsid that has initiated packaging of a monomeric DNA molecule in Figure 2a,b.[56-58] Similar observations have made for bacteriophages λ[59] and T4.[60,61] Analysis of ATP-driven motors is a central problem in studies of eukaryotes, as well as prokaryotes (reviewed in refs. 62-64).

For developing a detailed understanding of the mechanisms of biological motors, bacteriophage systems have the advantage of a highly developed platform for genetics. Thus far, genetic analysis has been targeted primarily toward gene 19, because gp19 appears to be the ATPase that powers the DNA packaging motor. Both gp18 and gp19 have been purified for both T7[65,66] and T3.[67,68] Beginning at the C-terminus (amino acid 586) of T3 gp19, the deletion of the C-terminal 10-15 amino acids reduces the binding of gp19 to the procapsid-associated T3 connector (gp8), as does competition with a peptide from the C-terminus. These deletions do not alter either the ATPase activity or the responsiveness to ATP-induced conformational change of T3 gp19. Thus, the C-terminus of gp19 is considered to a procapsid-binding domain.[69] Accessory protein domains are illustrated in Figure 2.

However, the remaining functions of gp19 appear not to be so neatly differentiated by location in gp19. The endonuclease function of T3 gp19 can be selectively removed by mutations in (1) any of three domains (amino acids 54-66, 360-372 and 422-432) that have an ATP-binding motif, and also (2) a domain (amino acids 344-349) that has a magnesium-binding motif.[70,71] These mutations do not prevent the packaging of monomeric T3 DNA, but prevent

the terminal cleavage of packaging when the DNA substrate is a plasmid DNA with T3 pac sequences inserted.

In pursuit of the "spark plug" of the T3 DNA packaging motor, genetic alteration of the three ATP binding domains has been performed. Mutating gp19 in the ATP binding domain closest to the C-terminus (domain III) causes (1) an increase in packaging-specific ATPase activity, and (2) a decrease in DNA packaging efficiency. The mutations involved cause an apparent slowing of the conversion of initiation complexes to mature T3 bacteriophage particles. This slowing is best explained by a decrease in the rate of entry into the T3 capsid. This decrease is possibly caused by loss of linkage between the ATPase and the DNA-driving part of the motor.[70] If so, then the sites of endonuclease activity overlap an ATP-binding site that is part of the DNA packaging motor. This type of overlap for bacteriophage λ accessory proteins has been assumed to help construct a model of a bacteriophage DNA packaging motor,[72] based on more recent analysis of the phenotypes of bacteriophage λ DNA packaging mutants.[73,74]

The gp19 ATP-binding domain closest to the N-terminus (called I^{70}) appears to be part of the T3 DNA packaging motor (presumably the motor ATPase) because one mutation in this domain stops the motor. This mutation also causes a loss of ATP-dependent control of the assembly of p19 on the connector. About 20, rather than 6, gp19 molecules assemble in either the presence or absence of ATP (Assembly of gp19 on the connector is illustrated in Fig. 3b). Thus, binding of ATP by this site appears to be essential to even assemble the T3 DNA packaging motor.[70,71] ATP-binding domain I also appears to be present in the large accessory proteins of bacteriophages ϕ29, λ, T4 and T7.[70,75] In the case of λ, the DNA packaging motor can be made to either slow or prematurely stop by mutating the equivalent of T3 ATP-binding domain I.[76]

The above results of genetic analysis are suggestive, but require additional biochemical/biophysical analysis of phenotype in order to provide accurate models for the DNA packaging motor and its accompanying endonucleolytic activity. However, determination of phenotype cannot be performed by use of the classical procedures of biochemistry. The reason is that the motor's cycle occurs in stages that are not separable by use of classical biochemistry. This is the case for any sequence of events linked by the physical association of active components in a single particle. The asynchrony of the motor's cycle among several motor-complexes further limits analysis. These problems can, in theory, be solved if one adopts a single-particle analysis of DNA packaging, rather than the ensemble-averaged analysis used in the studies described above. Single-particle analysis is discussed in the last section.

Finally, the genetic analysis is an example of research whose direction cannot be easily either restricted or biased by the design of the experiment. Thus, the expectation is that the genetic analysis will (a) help to minimize the development of poorly targeted research strategies, and (b) eventually help fill the gaps left by biochemical/biophysical analysis.

Structure of Capsids

The Mature Bacteriophage Capsid

The capsid of the mature bacteriophage T3/T7 particle has the following components: (a) Most mass is associated with an icosahedral outer shell (T=7) that is made of gp10 (Fig. 3a).[77-79] (b) One 5-fold symmetric vertex (of twelve 5-fold symmetric vertices) of the outer shell has a 12-fold symmetrical ring attached.[77-80] This ring is made of gp8 and is called either the connector or the portal; the connector has been purified in the case of both T3[80] and T7.[81] All connectors have an axial hole through which a DNA molecule enters the capsid during DNA packaging (Fig. 3a,b). (c) Coaxially mounted on the interior side of the connector is a roughly cylindrical complex of three proteins, gp14, gp15 and gp16. (d) Coaxially mounted on the

external side of the connector is an external projection called the tail. The tail has six fibers, each a trimer of gp17.[82] The remainder of the tail is made of gp11 and gp12. The tail also has an axial hole (Fig. 3a).

The predominant symmetry of the T7 internal cylinder is 8-fold rotational symmetry, as determined by cryoelectron microscopy with averaging among 213 particles. This symmetry appears to be associated primarily with gp15.[83] Thus, two symmetry mismatches occur in at least the case of T7: The first mismatch is between the locally 5-fold symmetric outer shell and the 12-fold symmetric connector. The second mismatch is between the connector and the 8-fold symmetric internal cylinder.

A T7 infection is initiated when a packaged double-stranded DNA genome is injected into a host cell through the axial hole of the tail. The data are interpreted by the formation of a trans-host membrane bridge by gp16.[11,84] Images of the apparent bridge have been obtained.[85]

The Packaged Double-Stranded DNA Genome

The outer shell of the T3/T7 capsid encloses the T3/T7 double-stranded DNA genome. The packaged genome is wrapped around the internal cylinder,[86,87] as illustrated in Figure 3a. The volume of the DNA packaged is about 0.5x the volume of the cavity in which the DNA is packaged.[12,88] Similar packing density has been observed for other double-stranded DNA bacteriophages.[89] The wrapping of the DNA can, in theory, be either (a) unidirectional, like thread on a spool, or (b) bi-directional, like rope on the deck of a ship. Bi-directional wrapping implies kinks. Kinks, if they exist, would create non-B structure. The upper limit for non-B DNA is 2% of the total, as determined by Raman spectroscopy of T7.[90]

Packaged bacteriophage DNA is so tightly packed that the enthalpy of packaging is, in theory, orders of magnitude greater than the entropy of packaging.[91] In practice, the enthalpy per mole of base pair of packaged T7 DNA is 0.47 ± 0.05 kcal, as determined by the heat released when T7 DNA is expelled from its capsid.[92] The nanometry-derived work per nucleotide pair used to package bacteriophage ϕ29 DNA is 0.56 kcal per mole of base pair, not much different.[93] The conclusion drawn is that the heat dissipated during DNA packaging is less than the energy needed to package the DNA molecule.[72,93]

The Procapsid

The T3/T7 procapsid, as usually isolated, has a structure that differs from the mature capsid structure. As usually isolated, the T3/T7 procapsid is smaller in radius, less angular[77,78,94] and more electrically charged (negative) at its surface.[95] Similar differences exist in the case of most other bacteriophage procapsids (reviewed in ref. 96), possibly excluding the procapsid of bacteriophage ϕ29. The ϕ29 procapsid can appear larger than the mature bacteriophage capsid (see Fig. 3 in ref. 97). The T3/T7 procapsid has the internal cylinder. No known genetically unrelated bacteriophage has this cylinder.

Like all other studied double-stranded DNA bacteriophages,[96] the T3/T7 procapsid also has a scaffolding protein (gp9) that is required for assembly of the outer shell. In the case of T3/T7, the scaffolding protein leaves the procapsid after the initiation of DNA packaging (reviewed in ref. 94). The scaffolding protein is bound to the inner surface of the outer shell of the T7 procapsid as judged by both protease sensitivity[98] and image reconstruction of cryoelectron micrographs.[94] The cryoelectron microscopy also showed that (a) the gp9 molecules project like stalactites (called nubbins) into the interior of the procapsid, and (b) are not in contact with each other, assuming one gp9 molecule per nubbin.[94] The first particle in the pathway of Figure 3b is the T7 procapsid as it is typically isolated.

The number of gp9 molecules per procapsid is 360 based on the image reconstruction from cryoelectron micrographs.[94] As indicated in reference 94, this number is higher by a

factor of at least 2 than the number of gp9 molecules per procapsid that had previously been determined by direct gel electrophoretic assay for protein. This inconsistency might be explained by the formation of two projecting nubbins by a single molecule of gp9. A second unexplained aspect of the structure is the expulsion of gp9 during DNA packaging in the apparent absence of a hole large enough to release gp9. Denaturation, followed by end-first motion of an elongated protein molecule, is presented as one possible explanation.[94] However, reasons exist to believe that a procapsid can achieve states that are far removed from the state of the procapsid that has been investigated in previous studies (next section). Thus, the possibility remains that comparatively large holes arise in the lattice of the procapsid's outer shell during DNA packaging.

Procapids in Altered States

The procapsids used in the studies of structure described above are usually assumed to be the major, if not the only, T3/T7 procapsids. However, this assumption appears not always to be accurate for procapsids isolated from extracts of T7-infected *Escherichia coli* that had been developed for use in the high efficiency packaging of T7 DNA in vitro. Initially, the discovery was made that the DNA packaging activity of these procapsids migrated more rapidly than radiolabeled procapsids obtained from a different lysate, during electrophoresis in a density gradient.[99] This observation was mentioned without interpretation in reference 99 and then, forgotten.

Recently the following even more unusual results were obtained with extracts[100] similar, but not identical, to those used in reference 99 (P. Serwer, S. J. Hayes and E. T. Wright, unpublished data): Most procapsids did not penetrate beyond a density of 1.05 g/ml when layered on a Nycodenz density gradient and centrifuged (major peak in Fig. 4). The time of centrifugation was more than 5x what is needed to centrifuge the mature bacteriophage capsid to a position at 1.28 g/ml, isodense with capsid protein (indicated with an arrow in Fig. 4). Some procapsids in Figure 4 did penetrate further than the procapsids in the major peak. But, these denser procapsids did not reach 1.28 g/ml. Further purification of the procapsids of Figure 3 has yielded additional surprises that are under investigation.

The point here is that only one explanation exists for a density as low as 1.05 g/ml: The procapsid is impermeable to Nycodenz and, therefore, has an interior that contains Nycodenz-depleted, possibly Nycodenz-free, water. This internal water is the source of the low density. A T7 (and T3[101]) capsid with impermeability-based low density has previously been documented.[101,102] This latter capsid is called Metrizamide low density (or MLD) capsid II. MLD capsid II is in the in vivo T7 DNA packaging pathway, based on the kinetics of its appearance.[102] However, no form of T7 capsid II is active post-isolation during in vitro DNA packaging.[99]

The density of MLD capsid II is 1.09 g/ml and the radius of MLD capsid II is indistinguishable from the radius of the mature T7 capsid.[102] Thus, if the even less dense Nycodenz low density particles of Figure 4 are made only of bacteriophage capsid proteins, they have an outer shell with a radius larger than the radius of the mature bacteriophage capsid. Equation 2 of reference 102 yields a radius that is ~25% larger, assuming the composition of MLD capsid II (no gp9) and no permeability to Nycodenz. If gp9 were present, then the radius would be even larger. The denser procapsids of the minor peak in Figure 4 may be either smaller, more massive or more leaky than the procapsids that form the major peak in Figure 4. Capacity for change in both radius and permeability are aspects of a DNA packaging motor that has previously been proposed.[72]

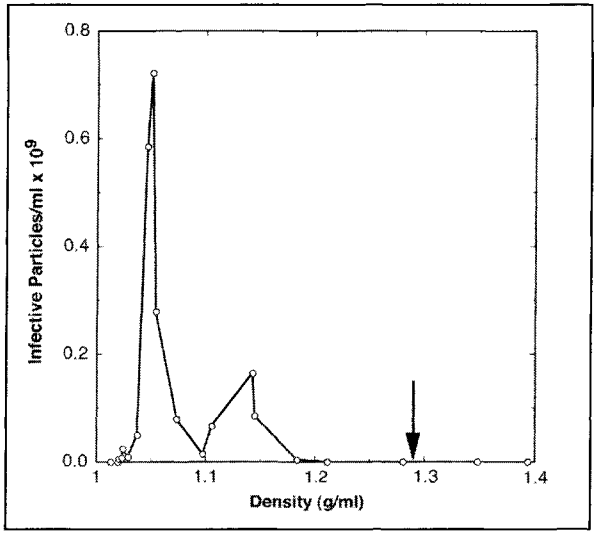


Figure 4. A new state of T7 procapsids isolated from an extract of concentrated T7-infected cells. Cells of *Escherichia coli* BL21 were infected with T7 amber mutant in genes 5 and 19. These cells are non-permissive for amber mutants. The amber mutation in gene 5 prevents T7 DNA synthesis; the amber mutation in gene 19 prevents DNA packaging, but does not prevent the formation of procapsids. These infected cells were both concentrated (100x) and lysed by use of lysozyme digestion/freeze-thawing, according to procedures that have been previously described.[100] The lysed infected cells were twice clarified by centrifugation in a Beckman TL-100 tabletop ultracentrifuge, in a TLA100.3 fixed angle rotor, at 60,00 rpm, at 2°C, for 5 min. Next, the clarified lysate was centrifuged in the same rotor at 90,000 rpm, at 2°C, for 60 min. This higher speed centrifugation is more than sufficient to pellet the T7 capsids that have already been investigated. However, the DNA packaging activity (>99%) did not pellet and was found in the bottom (yellow fraction) of two apparent phases that formed during centrifugation.
For fractionation of procapsids, the yellow fraction was layered (without dilution) on a Nycodenz gradient in a conical tube for the TLA100.3 rotor. The Nycodenz gradient was formed by layering the following Nycodenz solutions in the conical tube (volume of solution, followed by density of solution): 0.30 ml, 1.015 g/ml; 0.20 ml, 1.105 g/ml; 0.1 ml, 1.365 g/ml. The buffer for the Nycodenz solutions was the following: 0.02 M Tris-phosphate, pH 7.4, 0.1 M NaCl, 0.006 M $MgSO_4$, 400 mg/ml gelatin. The gradient was centrifuged at 60,000 rpm, at 2°C for 7 hours. Then, the gradient was fractionated by pipeting from the top. Each fraction was assayed for procapsid activity by dilution into a mixture of the following: (a) a DNA packaging extract made from cells infected by T7 amber mutant in genes 4 and 9 (no DNA synthesized and no capsids made), (b) DNA extracted from bacteriophage T7 particles. This mixture was incubated and, then, diluted and plated for infective particle count.[100] Infective particle titer is plotted as a function of the density of a fraction. Host proteins are present in all fractions.

Biochemistry

Some Past Studies of the Optimization of in vitro Systems

In vitro systems for bacteriophage DNA packaging provide the investigator with improved control of the conditions of DNA packaging. These systems also isolate the process of packaging from other cellular events. Finally, in vitro systems provide the investigator with the potential for single-particle observation of bacteriophage DNA packaging, i.e., observation of DNA packaging one packaging event at a time.

One of the conditions tested in the case of both T7[103] and P22[104] in vitro DNA packaging was the presence of electrically neutral compounds. The neutral compounds included polymers, sugars and polyols. This condition was tested for the following reasons: (a) The inside of a bacterial cell has both a raised osmotic pressure (lowered water activity)[105] and a raised excluded volume.[106] Both the lowered water activity and the raised excluded volume are caused primarily by the presence of intracellular protein and RNA.[105,106] (b) The neutral compounds lower the water activity during in vitro DNA packaging. Some of these compounds significantly raise the excluded volume. Empirically, the presence of neutral polymers, sugars and polyols caused a dramatic (several orders of magnitude) stimulation of in vitro DNA packaging, as measured by the formation of infective particles of both P22[104] and T7.[103]

The T7 in vitro DNA packaging efficiency is a very sensitive function of the concentration of polyethylene glycols,[107] frequently used polymers for generating either osmotic or excluded volume effects. This function has a peak so that a polyethylene glycol concentration (or quality) variation, either to the high side or to the low side, can cause an undesirable loss of DNA packaging efficiency. The polymer concentration-sensitivity of in vitro T7 DNA packaging is lower for dextran of molecular weight 10,000.[107] The dextran has, therefore, been favored for T7. Nonetheless, a purified T3 in vitro DNA packaging system uses polyethylene glycol, apparently without difficulty. The concentration used for T3 is consistent with the optimization performed for T7.[108,109]

The source of the neutral compound-induced stimulation of in vitro DNA packaging appears to be lowered water activity (rather than raised excluded volume), for the following reason: The neutral compounds used also inhibited both (a) the elevated temperature, chelating agent-induced expulsion of DNA from the capsids of bacteriophages P22 and T7[103] and (b) the receptor protein-induced expulsion of a packaged bacteriophage λ DNA molecule.[110] The cause of the expulsion was assumed to be the pressure exerted by the packaged DNA molecule against the outer shell of the capsid.[103,110] A polymer-induced excluded volume effect on a packaged DNA molecule seemed unlikely because an external polymer would not have access to the DNA molecule. More likely, an osmotic pressure difference across the capsid's outer shell was the cause of the stabilization.

Similarly, an osmotic pressure difference across the capsid's outer shell was the best explanation for the neutral compound-induced stimulation of in vitro DNA packaging. The reason is that quantification of excluded volume revealed that excluded volume effects were inadequate to explain the stimulatory effect of neutral compounds on in vitro T7 DNA packaging.[111] An ATP-maintained osmotic pressure gradient is potentially a source of bias for a ratchet-like DNA packaging motor. By this hypothesis, the osmotic pressure gradient is maintained, in part, by ATP-dependent pumping of non-DNA molecules out of the capsid.[72] A permeability coupled expansion/contraction cycle of the capsid's outer shell is a feature for one proposed pump.[72]

Furthermore, a T3 in vitro DNA packaging system with purified components (procapsids, DNA ATP, polyethylene glycol, buffer) has been developed.[108,109] The state of the T3 procapsid during these studies was the conventional smaller, rounded particle state (first particle in the pathway of Fig. 3b), as judged by both electron microscopy and nondenaturing gel electrophoresis.[109] This is the state originally observed, not the more recently observed apparently larger, impermeable particle state that is described in the previous section.

The T3 purified in vitro system has been used to isolate a monomeric DNA-procapsid complex whose formation requires the presence of either ATP or a non-cleavable ATP analogue.[108,109] Use of a non-cleavable analogue to form this complex is a procedure for synchronizing the start of the entry of a T3 monomeric DNA molecule into a capsid (illustrated in Fig.

3b). The entry occurs at an average rate of 21 Kb/min at 30 °C.[109] Within experimental error, this rate is equal to the rate of entry for the last 20-50% of T7 DNA packaged after cleavage from a concatemer in an unpurified in vitro system at 30°C (28 ± 6 Kb/min in the case of T7).[55]

The data for T7 do not indicate any significant decrease in the rate of entry as DNA packaging is completed. Only a slight decrease in the rate of entry occurs as more than 85% of T3 DNA is packaged (Fig. 2b in ref. 109). These observations contrast the following observation made by nanometry of a latex bead attached to a single bacteriophage ϕ29 DNA molecule that is being packaged in vitro: The rate of DNA entry decreases (to roughly 40% of its initial value when 85% of the genome is packaged) and the resistance to in vitro packaging increases as more DNA is packaged. A possible reason for the more constant T3/T7 rate of entry is that the T3 (but not the ϕ29) in vitro system retains feedback control of the speed of DNA packaging (see further discussion in ref. 72). This hypothesis predicts that ATP utilization increases as a function of the fraction of DNA packaged. Measurement has not yet been performed for ATP utilization as a function of the fraction of a DNA molecule that has been packaged.

Ensemble Averaging vs. Single-Molecule Analysis

Measurement should be possible for ATP utilization rate as a function of the fraction of a DNA molecule that has been packaged. In analogy with combustion rate in an internal combustion motor, ATP utilization rate is one of several motor state-dependent characteristics that, together, define the cycle of DNA packaging motors. The other characteristics also will eventually be measured as a function of the fraction of a DNA molecule that has been packaged. Based on past experimentation, two types of procedure can be used: (1) The start of DNA entry is synchronized. Then, ATP utilization is measured as a function of time by use of conventional, ensemble-averaging techniques. The ATP utilization by the T3 DNA packaging motor has been rigorously shown to be the difference between the total ATP utilization and the utilization in the presence of actinomycin. Actinomycin is an inhibitor of the motor-dependent, but not the motor-independent, ATP cleavage activity of T3 gp19.[46] (2) To bypass the need for synchronization, ATP utilization is measured at the level of a single DNA packaging event. In preparation for analysis by single-molecule techniques, both single capsids and single DNA molecules have been detected without difficulty by use of fluorescence microscopy (see Fig. 5). Single, fluorescently labeled ATP molecules have been detected by use of total internal reflection fluorescence microscopy.[112]

The advantage of synchronization-based ensemble averaging procedures is that they are already a part of conventional biochemistry. On the other hand, a major disadvantage of synchronization-based ensemble averaging procedures is that the synchronization will be progressively lost as more DNA is packaged. The data will average results at several stages of DNA packaging. Therefore, the results will not be accurate enough to answer some, if not most, questions. Loss of synchrony is one of the intrinsic difficulties of analyzing any biochemical pathway by use of conventional ensemble averaging techniques.

More specifically, loss of synchrony during synchronization-based ensemble averaging procedures results in the loss of ordering in time for ATP cleavage in relation to the structural changes of the motor's cycle. Loss of ordering in time also occurs for the various structural changes of the motor's cycle (see ref. 112). These losses cripple attempts to analyze the motor's cycle because ordering in time is critical for understanding cause and effect. On the other hand, a disadvantage of single-particle analysis is that single-particle analysis is not yet a mature discipline and is, currently, difficult to apply. That, of course, should not prevent the building of a platform for single-particle studies of DNA packaging.

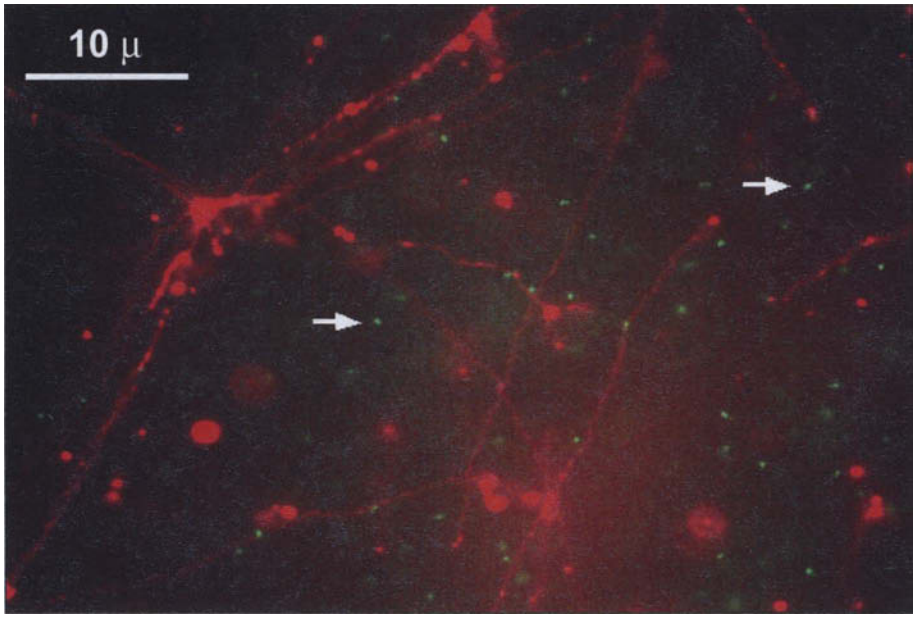


Figure 5. Fluorescence microscopy of a single T7 100S+ DNA molecule in the presence of T7 procapsids. Fluorescence microscopy was performed of the following mixture: (i) Alexa 488-stained (green fluorescence; see ref. 114) particles from a procapsid fraction of a Nycodenz gradient, (ii) T7 100S+ DNA that was stained with 1 mg/ml ethidium bromide (orange fluorescence), and (iii) 1.1% low-gelling agarose (SeaPrep, Biowhittaker, Rockland, ME). The buffer was 0.039 M NaCl, 0.77 mM Tris-Cl, pH 7.4, 0.077 mM $MgCl_2$, 0.077 μg/ml gelatin, 0.34 mg/ml ethidium bromide, 4.5% β-mercaptoethanol. This mixture was placed on a glass microscope slide as a 22 ml drop. A cover glass was placed on the drop. To gel the agarose, the slide was placed for 15 minutes on an aluminum block pre-chilled to 0°C. The gelled agarose-embedded specimen was observed by fluorescence microscopy. The green particles in the figure are undergoing thermal motion that is damped by the presence of the agarose gel. The orange strands are segments of a 100S+ T7 replicating/recombining DNA complex. Gelation of the agarose has stretched the DNA segments. Note that several DNA segments are concatemeric in length. This specimen was not active in DNA packaging.

Single-Particle Analysis in Vitro: Some Past Studies by Fluorescence Microscopy

The following procedure achieved single-particle analysis of in vitro bacteriophage T7 DNA packaging: (a) Monomeric T7 DNA molecules were added to an extract of T7-infected cells. The DNA was stained with 4',6-diamidino-2-phenylindole (DAPI). (b) The extract was observed by use of fluorescence microscopy. While in the extract, the monomeric DNA molecules formed concatemers via the 5' to 3' exonuclease activity of gp6 exonuclease. This activity made the terminal repeat single-stranded. The single-stranded terminal repeats joined to form concatemers.

The fluorescence microscopy yielded the following surprising observation: The concatemers partitioned to form a DNA network that was embedded in solution that was more DNA poor. Each strand of the network had several DNA molecules. The network was eventually cleaved by p19. Single events of p18-, p19-dependent cleavage were observed by fluorescence microscopy.[113] (see also www.biochem.uthscsa.edu/~serwer). The cleavage products were eventually packaged.

The second surprise was that, after release from the network by p18-, p19-dependent cleavage, concatemer-associated T7 genomes were not packaged at random. Instead, packaging events were clustered in both time and space on the concatemers. Thus, the conclusion was drawn that the packaging of T7 concatemer-associated genomes is co-operative among the different capsids.[100] Before the single-event studies by fluorescence microscopy, the possibility of co-operative DNA packaging had not previously been considered, to the author's knowledge. Nonetheless, co-operative packaging provides a potential answer to an old, troublesome question: How does T3/T7 avoid packaging a concatemer-associated genome (Fig. 1) that does not have a duplicated terminal repeat? For that matter, does T3/T7 (or do other viruses) have a mechanism for selecting the more promising DNA molecules for packaging? Binding more than one genome in a concatemer during co-operative packaging (see Fig. 1) is a possible component of this pre-packaging assessment.[100]

In the case of T7 DNA packaging in vivo, the concatemeric DNA substrate is usually attached to additional replicating-recombining, branched DNA. This multigenomic DNA complex is called 100S+ DNA because of its high sedimentation rate during rate zonal centrifugation.[32] A fluorescence micrograph of a stretched T7 100S+ DNA complex is in Figure 5 (orange, ethidium-stained). Embedding in an agarose gel was used to stretch the strands of this DNA complex during specimen preparation. This specimen also had particles from a procapsid fraction of a Nycodenz density gradient. These particles had previously been stained with Alexa 488 (see ref. 114). Some of these particles (green) are indicated with arrows. The thermal motion of T7 capsids in buffered aqueous solution is so high that the capsids can be tracked only for about one second on average.[114] However, the agarose gel used to stretch DNA also reduced the thermal motion of smaller particles, including capsids (see also ref. 114). Specimens like the specimen in Figure 5 are being prepared as part of an effort to achieve real time observation of both single capsids and single DNA molecules during in vitro T7 DNA packaging.

Acknowledgments

The author thanks Stephen C. Hardies for critical comments on a draft of this manuscript. Research in the author's laboratory was supported by The National Institutes of Health (GM24365), The Welch Foundation (AQ-764) and the Texas Coordinating Board (003659-0071-2001).

References

1. Summers WC. Bacteriophage therapy. Ann Rev Microbiol 2001; 55:437-451.
2. Sulakvelidze A, Alavidze Z, Morris Jr JG. Bacteriophage therapy. Antimicrobial Agents and Chemotherapy 2001; 45(3):649-659.
3. Biswas B, Adhya S, Washart P et al. Bacteriophage therapy rescues mice bacteremic from a clinical isolate of vancomycin-resistant Enterococcus faecium. Infect Immun 2002; 70 (1):204-210.
4. Gottesman M. Bacteriophage lambda: the untold story. J Mol Biol 1999; 293(2):177-180.
5. Stent, GS, Calendar, R. Molecular Genetics:An Introductory Narrative. WH Freeman and Company, San Francisco 1978.
6. Doermann, AH. Introduction to the early years of bacteriophage T4. In: Mathews CK, Kutter EM, Mosig G et al, eds. Bacteriophage T4. Washington, DC: American Society for Microbiology, 1983;1-7.
7. Kellenberger E, Wunderli-Allenspach H. Electron microscopic studies on intracellular phage development-history and perspectives. Micron 1995; 26(3):213-245.
8. Fujisawa H, Morita M. Phage DNA packaging. Genes to Cells 1997; 2(9):537-545.
9. Catalano CE. The terminase enzyme from bacteriophage lambda: a DNA-packaging machine. Cell Mol Life Sci 2000; 57(1):128-148.
10. Dunn JJ, Studier FW. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J Mol Biol 1983; 166:477-535.

11. Pajunen MI, Elizondo MR, Skurnik M et al. Complete nucleotide sequence and likely recombinatorial origin of bacteriophage T3. J Mol Biol 2002; 319(5):1115-1132.
12. Serwer, P. Buoyant density sedimentation of macromolecules in sodium iothalamate density gradients. J Mol Biol 1975; 92(3):433-448.
13. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21(2):319-331.
14. Studier FW, Dunn JJ. Organization and expression of bacteriophage T7 DNA. Cold Spring Harbor Symp Quant Biol 1983; 47(2):999-1007.
15. Hausmann R. Bacteriophage T7 genetics. Curr Topics Microbiol Immunol 1976; 75:77-110.
16. Studier FW. Genetic analysis of non-essential bacteriophage T7 genes. J Mol Biol 1973; 79(2):227-236.
17. Yamada H, Fujisawa H, Kato H et al. Cloning and sequencing the genetic right end of bacteriophage T3 DNA. Virology 1986; 151(2):350-361.
18. Molineux IJ. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. Mol Microbiol 2001; 40(1):1-8.
19. Czaika G, Mamet-Bratley. Defective DNA injection by alkylated and nonalkylated bacteriophage T7. Biochim Biophys Acta 1992; 1130(1):52-62.
20. Studier FW. Bacteriophage T7. Science 1972; 176(33):367-376.
21. Serwer P, Watson RH. Capsid-DNA complexes in the DNA packaging pathway of bacteriophage T7:characterization of the capsids bound to monomeric and concatemeric DNA. Virology 1981; 108(1):164-176.
22. Masker WE. In vitro packaging of bacteriophage T7 DNA requires ATP. J Virol 1982; 43(1):365-367.
23. Hamada K, Fujisawa H, Minagawa T. A defined in vitro system for packaging bacteriophage T3 DNA. Virology 1986; 151(1):119-123.
24. Black LW. DNA packaging in dsDNA bacteriophages. Ann Rev Microbiol 1989; 43:267-292.
25. Hardies SC, Comeau AM, Serwer P et al. The complete sequence of marine bacteriophage VpV262 infecting Vibrio parahaemolyticus indicates that an ancestral component of a T7 viral supergroup is widespread in the marine environment. Virology 2003; 310:359-371.
26. Rohwer F, Segall A, Steward G et al. The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with non-marine phages. Limnol. Oceanogr 2000; 42:408-418.
27. Chen F, Lu J. Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. Appl Environ Microbiol 2002; 68(5):2589-2594.
28. Cermakian N, Ikeda TM, Miramontes P et al. On the evolution of the single-subunit RNA polymerases. J Mol Evolution 1997; 45(6) 671-681.
29. Gould SJ. Gulliver's further travels:the necessity and difficulty of a hierarchical theory of selection. Philos Trans Royal Soc London-Series B: Biological Sciences 1998; 353(1366):307-314.
30. Gold T. The Deep Hot Biosphere. New York:Springer-Verlag, 1999.
31. Breitbart M, Salamon P, Andresen B et al. Genomic analysis of uncultured marine viral communities. Proc Natl Acad Sci USA 2002; 99:14250-14255.
32. Serwer P, Watson RH, Hayes SJ. Multidimensional analysis of intracellular bacteriophage T7 DNA:effects of amber mutations in genes 3 and 19. J Virol 1987; 61(11):3499-3509.
33. Chung YB, Nardone C, Hinkle DC. Bacteriophage T7 DNA packaging. III. A "hairpin" end formed on T7 concatemers may be an intermediate in the processing reaction. J Mol Biol 1990; 216(4):939-948.
34. Gabashvili IS, Khan SA, Hayes SJ et al. Polymorphism of bacteriophage T7. J Mol Biol 1997; 273(3):658-667.
35. Steven AC, Trus BL. The structure of bacteriophage T7. In: Harris JR, Horne RW, eds. Electron Microscopy of Proteins. Vol 5. London: Academic Press, 1986:1986:1-35.
36. Studier FW. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J Mol Biol 1991; 219(1):37-44.
37. Rosenberg A, Griffin K, Studier FW et al. T7Select phage display system: A powerful new protein display system based on bacteriophage T7. Novations, Newsletter of Novagen, Inc. 1996; 6:1-6.
38. Zhang X, Studier FW. Isolation of transcriptionally active mutants of T7 RNA polymerase that do not support phage growth. J Mol Biol 1995; 250(2):156-168.

39. Hashimoto C, Fujisawa H. DNA sequences necessary for packaging bacteriophage T3 DNA. Virology 1992; 187(2):788-795.
40. Chung YB, Hinkle DC. Bacteriophage T7 DNA packaging. I. Plasmids containing a T7 replication origin and the T7 concatemer junction are packaged into transducing particles during phage infection. J Mol Biol 1990; 216(4):911-926.
41. Chung YB, Hinkle DC. Bacteriophage T7 DNA packaging. II. Analysis of the DNA sequences required for packaging using a plasmid transduction assay. J Mol Biol 1990; 216(4):927-938.
42. Hashimoto C, Fujisawa H. Transcription dependence of DNA packaging of bacteriophages T3 and T7. Virology 1992; 191:246-250.
43. Tsuchida SI, Kokubo H, Tasaka M et al. DNA sequences responsible for specificity of DNA packaging and phage growth interference of bacteriophages T3 and T7. Virology 1996; 217(1):332-337.
44. Son M, Serwer P. Role of exonuclease in the specificity of bacteriophage T7 DNA packaging. Virology 1992; 190(2):824-833.
45. Hashimoto C, Fujisawa H. Packaging and transduction of non-T3 DNA by bacteriophage T3. Virology 1988; 166(2):432-439.
46. Morita M, Tasaka M, Fujisawa H. DNA packaging ATPase of bacteriophage T3. Virology 1993; 193(2):748-752.
47. Condit RC. F factor-mediated inhibition of bacteriophage T7 growth:increased membrane permeability and decreased ATP levels following T7 infection of male Escherichia coli. J Mol Biol 1975; 98(1):45-59.
48. Britton JR, Haselkorn R. Permeability lesions in male Escherichia coli infected with bacteriophage T7. Proc Natl Acad Sci USA 1975; 72(6):2222-2226.
49. Khan SA, Hayes SJ, Watson RH et al. Specific, nonproductive cleavage of packaged bacteriophage T7 DNA in vivo. Virology 1995; 210(2):409-420.
50. Khan SA, Hayes SJ, Wright ET et al. Specific single-stranded breaks in mature bacteriophage T7 DNA. Virology 1995; 211(1):329-331.
51. Fujisawa H, Kimura M, Hashimoto C. In vitro cleavage of the concatemer joint of bacteriophage T3 DNA. Virology 1990; 174, 26-34.
52. White JH, Richardson CC. Processing of concatemers of bacteriophage T7 DNA in vitro. J Biol Chem 1987; 262(18):8851-8860.
53. Serwer P, Watson RH, Hayes SJ. Formation of the right before the left mature DNA end during packaging-cleavage of bacteriophage T7 DNA concatemers. J Mol Biol 1992; 226(2):311-317.
54. Serwer P, Watson RH, Son M. Role of gene 6 exonuclease in the replication and packaging of bacteriophage T7 DNA. J Mol Biol 1990; 215(2):287-299.
55. Son M, Watson RH, Serwer P. The direction and rate of bacteriophage T7 DNA packaging in vitro. Virology 1993; 196(1):282-289.
56. Morita M, Tasaka M, Fujisawa H. 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 1995; 245(5):635-644.
57. Morita M, Tasaka M, Fujisawa H. Analysis of functional domains of the packaging proteins of bacteriophage T3 by site-directed mutagenesis. J Mol Biol 1994; 235(1):248-259.
58. Fujisawa H, Shibata H, Kato H. Analysis of interactions among factors involved in the bacteriophage T3 DNA packaging reaction in a defined in vitro system. Virology 1991; 185(2):788-794.
59. Yeo A, Feiss M. Specific interaction of terminase, the DNA packaging enzyme of bacteriophage lambda, with the portal protein of the prohead. J Mol Biol 1995; 245(2):141-150.
60. Leffers G, Rao VB. Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. J Biol Chem 2000; 275(47):37127-37136.
61. Lin H, Rao, VB, Black LW. Analysis of capsid portal protein and terminase functional domains:interaction sites required for DNA packaging in bacteriophage T4. J of Mol Biol 1999; 289(2):249-260.
62. Howard J. Mechanics of Motor Proteins and the Cytoskeleton. Sunderland: Sinauer Associates, Inc., 2001.
63. Alberts B. The cell as a collection of protein machines:preparing the next generation of molecular biologists. Cell 1998; 92(3):291-294.

64. Ishijima A, Yanagida T. Single molecule nanobioscience. Trends in Biochem Sci 2001; 26(7):438-444.
65. White JH, Richardson CC. Gene 19 of bacteriophage T7. Overexpression, purification, and characterization of its product. J Biol Chem 1988; 263(5):2469-2476.
66. White JH, Richardson CC. Gene 18 protein of bacteriophage T7. Overproduction, purification, and characterization. J Biol Chem 1987; 262(18):8845-8850.
67. Hamada K, Fujisawa H, Minagawa T. Purification and properties of gene 18 product of bacteriophage T3. Virology 1984; 139(2):251-259.
68. Hamada K, Fujisawa H, Minagawa T. Overproduction and purification of the products of bacteriophage T3 genes 18 and 19, two genes involved in DNA packaging. Virology 1986; 151(1):110-118.
69. Morita M, Tasaka M, Fujisawa H. 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 1995; 245(5):635-644.
70. Morita M, Tasaka M, Fujisawa H. Analysis of functional domains of the packaging proteins of bacteriophage T3 by site-directed mutagenesis. J Mol Biol 1994; 235(1):248-259.
71. Kimura M, Fujisawa F. Dissection of functional domains of the packaging protein of bacteriophoage T3 by site-directed mutagenesis. Virology 1991; 180:709-715.
72. Serwer P. Models of bacteriophage DNA packaging motors. J Struct Biol 2003; 141:179-188.
73. Woods L, Catalano CE. Kinetic characterization of the GTPase activity of phage λ terminase: Evidence for communication between the two "NTPase" catalytic sites of the enzyme. Biochemistry 1999; 38:14624-14630.
74. Hwang Y, Catalano CE, Feiss M. Kinetic and mutational dissection of the two ATPase activities of terminase, the DNA packaging enzyme of bacteriophage Chi. Biochemistry 1996; 35:2796-2803.
75. Guo P, Peterson C, Anderson D. Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage ϕ29. J Mol Biol 1987; 197(2):229-236.
76. Duffy C, Feiss M. The large subunit of bacteriophage lambda's terminase plays a role in DNA translocation and packaging termination. J Mol Biol 2002; 316(3):547-561.
77. Serwer P. Internal proteins of bacteriophage T7. J Mol Biol 1976; 107:271-291.
78. Steven AC, Trus BL. The structure of bacteriophage T7. In: Harris JR, Horne RW, eds. Electron Microscopy of Proteins. Vol 5. London: Academic Press, 1986:1-35.
79. Cerritelli ME, Studier FW. Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. J Mol Biol 1996; 258(2):286-298.
80. Valpuesta JM, Sousa N, Barthelemy I et al. Structural analysis of the bacteriophage T3 head-to-tail connector. J Struct Biol 2000; 131(2):146-155.
81. Cerritelli ME, Studier FW. Purification and characterization of T7 head-tail connectors expressed from the cloned gene. J Mol Biol 1996; 258(2):299-307.
82. Steven AC, Trus BL, Maizel JV et al. Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. J Mol Biol 1988; 200(2):351-365.
83. Cerritelli ME, Trus BL, Smith CS et al. A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. J Mol Biol 2003; 327(1):1-6.
84. Moak M, Molineux IJ. Role of the Gp16 lytic transglycosylase motif in bacteriophage T7 virions at the initiation of infection. Mol Microbiol 2000; 37(2):345-355.
85. Serwer P. Observation of DNA by negative staining: Phage T7 DNA-capsid complexes. Ninth International Congress on Electron Microscopy. Vol II. Toronto: 1978:228-229.
86. Cerritelli ME, Cheng N, Rosenberg AH et al. Encapsidated conformation of bacteriophage T7 DNA. Cell 1997; 91(2):271-280.
87. Serwer P, Khan SA, Hayes SJ et al. The conformation of packaged bacteriophage T7 DNA: Informative images of negatively stained T7. J Struct Biol 1997; 120(1):32-43.
88. Stroud RM, Serwer P, Ross MJ. Assembly of bacteriophage T7. Dimensions of the bacteriophage and its capsids. Biophys J 1981; 36(3):743-757.
89. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21(2):319-331.

90. Overman SA, Aubrey KL, Reilly KE et al. Conformation and interactions of the packaged double-stranded DNA genome of bacteriophage T7. Biospectroscopy 1998; 4(5 Suppl):S47-S56.
91. Riemer SC, Bloomfield VA. Packaging of DNA in bacteriophage heads: Some considerations on energetics. Biopolymers 1978; 17(3):785-794.
92. Raman CS, Hayes SJ, Nall BT et al. Energy stored in the packaged DNA of bacteriophage T7. Biophys J 1993; 64:A12.
93. Smith DE, Tans SJ, Smith SB et al. The bacteriophage straight ϕ29 portal motor can package DNA against a large internal force. Nature 2001; 413(6857):748-752.
94. Cerritelli M, Conway JF, Cheng N et al. Molecular mechanisms in bacteriophage T7 procapsid assembly, maturation and DNA containment. In: Chiu W, Johnson JE, eds. Advances in Protein Chemistry. Vol. 64. Virus Structure. San Diego: Academic Press, 2003:301-323.
95. Serwer P, Pichler ME. Electrophoresis of bacteriophage T7 and T7 capsids in agarose gels. J Virol 1978; 28(3):917-928.
96. Dokland T. Scaffolding proteins and their role in viral assembly. Cell Mol Life Sci 1999; 56(7-8):580-603.
97. Bjornsti MA, Reilly BE, Anderson DL. In vitro assembly of the Bacillus subtilis bacteriophage ϕ29. Proc Natl Acad Sci USA 1981; 78(9):5861-5865.
98. Serwer P, Hayes SJ, Watson RH. The structure of a bacteriophage T7 procapsid and its in vivo conversion product probed by digestion with trypsin. Virology 1982; 122(2):392-401.
99. Masker WE, Serwer P. DNA packaging in vitro by an isolated bacteriophage T7 procapsid. J Virol1982; 43(3):113811-42.
100. Sun M, Louie D, Serwer P. Single-event analysis of the packaging of bacteriophage T7 DNA concatemers in vitro. Biophys J 1999; 77(3):1627-1637.
101. Serwer P, Watson RH, Hayes SJ et al. Comparison of the physical properties and assembly pathways of the related bacteriophages T7, T3 and ϕII. J Mol Biol 1983; 170(2):447-469.
102. Serwer P. A metrizamide-impermeable capsid in the DNA packaging pathway of bacteriophage T7. J Mol Biol 1980; 138(1):65-91.
103. Serwer P, Masker WE, Allen JL. Stability and in vitro DNA packaging of bacteriophages: effects of dextrans, sugars, and polyols. J Virol 1983; 45(2):665-671.
104. Gope R, Serwer P. Bacteriophage P22 in vitro DNA packaging monitored by agarose gel electrophoresis: rate of DNA entry into capsids. J Virol 1983; 47(1):96-105.
105. Koch AL. The bacterium's way for safe enlargement and division. Appl Environ Microbiol 2000; 66(9):3657-3663.
106. Minton AP. Implications of macromolecular crowding for protein assembly. Curr Opin Struct Biol 2000; 10(1):34-39.
107. Son M, Hayes SJ, Serwer P. Optimization of the in vitro packaging efficiency of bacteriophage T7 DNA: effects of neutral polymers. Gene 1989; 82(2):321-325.
108. Shibata H, Fujisawa H, Minagawa T. Early events in DNA packaging in a defined in vitro system of bacteriophage T3. Virology 1987; 159(2):250-258.
109. Shibata H, Fujisawa H, Minagawa T. Characterization of the bacteriophage T3 DNA packaging reaction in vitro in a defined system. J Mol Biol 1987; 196:845-851.
110. Evilevitch A, Lavelle L, Knobler CM et al. Osmotic pressure inhibition of DNA ejection from phage. Proc Natl Acad Sci USA 2003; 100(16)9292-9295.
111. Louie D, Serwer P. Quantification of the effect of excluded volume on double-stranded DNA. J Mol Biol 1994; 242(4):547-558.
112. Yanagida T, Kitamura K, Tanaka H et al. Single molecule analysis of the actomyosin motor. Curr Opin Cell Biol 2000; 12(1):20-25.
113. Sun M, Son M, Serwer P. Formation and cleavage of a DNA network during in vitro bacteriophage T7 DNA packaging: light microscopy of DNA metabolism. Biochemistry 1997; 36(42):13018-13026.
114. Huang S, Hayes SJ, Serwer P. Fluorescence microscopy of single viral capsids. J Struct Biol 2001; 135(3):270-280.

CHAPTER 5

DNA Packaging by Bacteriophage P22

Sherwood Casjens and Peter Weigele

Introduction

Bacteriophage P22 was isolated by Zinder and Lederberg[1] a half century ago, and was immediately put to work by *Salmonella* bacterial geneticists because of its unusual (at that time) DNA packaging properties. It was the first generalized transducing phage to be discovered – a small fraction (~2%[2]) of its virions carry a fragment of the host DNA instead of phage DNA, and this host DNA can be delivered into a host cell. Subsequently it has become one of the paradigm molecular genetic systems for understanding the lifecycles of temperate phages.[3,4] P22 virions contain partially circularly permuted (see below for details) and 3.8% terminally redundant dsDNA molecules about 43,400 bp long,[5] have very short tails[6] and infect smooth (O-antigen surface polysaccharide carrying) strains of *Salmonella typhimurium*. The many conditional lethal mutations isolated in P22 allowed the delineation of its assembly pathway over a quarter century ago.[7-10]

Building a Protective Shell for Viral DNA

Like other large dsDNA viruses, P22 assembles a protein procapsid and then inserts the DNA chromosome into this preformed container (Fig. 1). This strategy appears to be common to all dsDNA bacteriophages as well as the *Herpesviridae*. The genes required to build the procapsid and fill it with DNA map in a contiguous cluster on the P22 chromosome (Fig. 2). The four critical protein players in this process in P22 are analogous to those in other dsDNA viruses—coat, scaffold, portal and terminase (Table 1). The coat protein shells of P22 procapsids and virions have T=7 icosahedral symmetry (each procapsid contains 420 molecules of coat, less any that are replaced by portal), and three-dimensional reconstructions of both kinds of particles have been made from cryo-electron micrographs.[11-15] Procapsids contain about 250 molecules of scaffolding protein in the interior, all of which leave at or before the time of DNA packaging.[8,16] Scaffolding protein has essential roles in procapsid assembly that include directing coat protein to assemble properly, recruiting portal protein into procapsids, and perhaps excluding cytoplasmic proteins from the procapsid interior.[16-19] Twelve molecules of portal protein form a ring at a single capsid vertex to which tails will later attach.[20] DNA is thought to pass through this ring during packaging and injection. As in other large, well-studied viruses, a terminase which contains two types of subunits is responsible for DNA recognition and cleavage and perhaps provides the force required for DNA translocation through the portal complex. Sometime after scaffolding protein exit but around the time DNA packaging, the procapsid shell expands and the conformational change to the coat protein lattice results in a physically more robust capsid that is better able to withstand chemical insult.

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

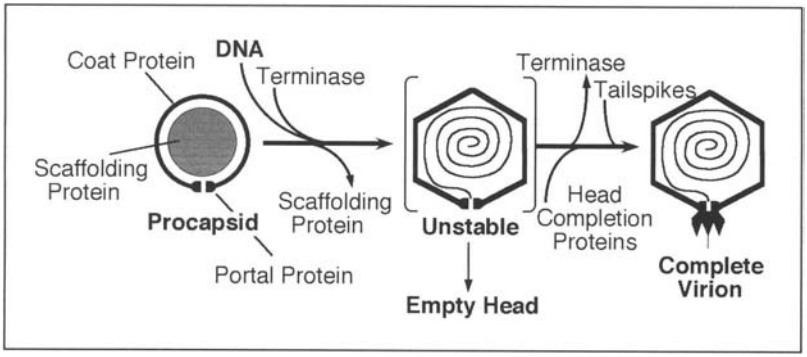


Figure 1. Bacteriophage P22 assembly pathway. The coat proteins of all three shells have T=7 icosahedral geometry. Four additional proteins (the products of genes *7*, *14*, *20* and *16*) which are not shown here are required for full virion infectivity, but are not required for the assembly of a particle that contains a stably packaged DNA molecule.

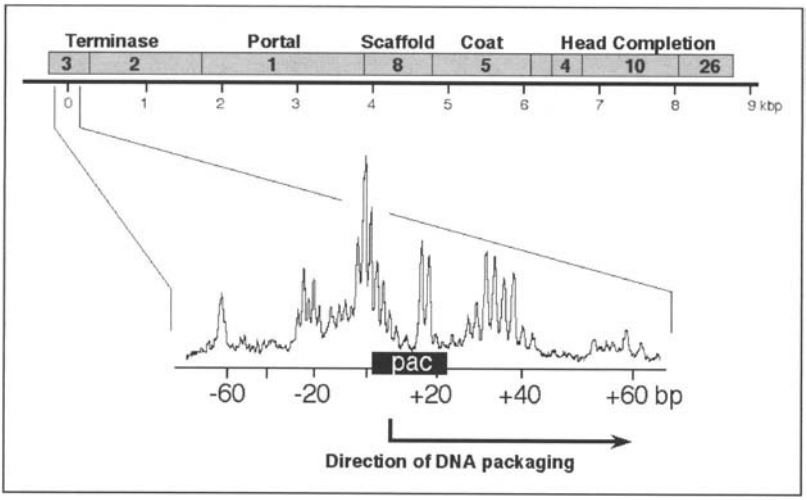


Figure 2. P22 DNA packaging genes and *pac* site.Gene numbers are indicated within the boxes denoting the genes above. Below is a densitometer scan of an electrophoresis gel of *Nru*I generated P22 packaging series left-end fragments. The peaks indicate the locations and relative frequencies at which ends are generated for initiation of DNA packaging series at the native *pac* site. The horizontal black box indicates the size and location of the *pac* site.[34] This figure is modified from Figure 4 of Casjens et al.[30]

DNA Packaging Strategy

The linear P22 virion DNA is circularized by homologous recombination upon infection and replication eventually produces daughter head-to-tail concatemers that are the substrate for the DNA packaging apparatus.[21] The circular permutation and terminal redundancy of the packaged DNA strongly suggests a "headful packaging" mechanism.[22] Headful packaging refers to a mechanism in which the interior volume of the capsid shell measures the length of the chromosome that is encapsulated (since it is filled to a particular DNA density) and a "headful nuclease" cleaves packaged DNA from the remainder of the concatemer only when the head is "full" of DNA (this dsDNA cut is called "headful cleavage"). A more careful study of P22

Table 1. Proteins required for bacteriophage P22 DNA packaging

Gene	Protein	Role in DNA Packaging
1	portal protein	forms the hole through which DNA enters the capsid; has an active role in DNA headful measurement and perhaps in DNA entry into the procapsid
2	large terminase subunit	required for DNA cleavage; detailed role not known; not present in virion
3	small terminase subunit	required for DNA cleavage; recognition of phage DNA for packaging; not present in virion
5	coat protein	forms the T=7 icosahedral phage head or capsid
8	scaffolding protein	required for proper coat assembly and for recruiting portal into procapsid; not present in the virion
4,10,26	head completion proteins	required for stable DNA packaging; may plug the portal hole after DNA has entered the capsid

virion DNA by Tye et al[23] and Jackson et al[24] showed elegantly that this is true, and in addition discovered that the circular permutation was not random, but that its DNA is "partially permuted". That is, DNA ends are not randomly positioned on the sequence, but are found largely in one region of the phage's genome sequence. The model proposed to explain partial permutation is that packaging initiates at a particular location called *pac*. A DNA end is thought to be created here by a DNA cleavage called the "series initiation cleavage" (Fig. 3), and the DNA end on one side of this cut is inserted into the procapsid so that packaging proceeds in only one direction from the initiation cleavage point. When DNA inside this procapsid reaches sufficient density, the headful nuclease is activated to cleave the DNA (the first headful cleavage of the packaging series), and the capsid full of DNA is released from the concatemer. A second DNA headful is then inserted into a new procapsid starting at the concatemer end created by the previous headful cleavage and ending with a new headful cleavge (the second headful cleavage of the series). Subsequent packaging events then follow sequentially in the same manner (Fig. 3). Such "processive" packaging series can vary from 2 to 12 packaging events long in phage P22, depending upon the infection conditions[24,25]—it is not known if the number of packaging events in a processive series is limited by concatemer length, or whether multiple series can initiate on a single concatemer molecule.

The DNA length in capsids is 43,400±750 bp, where the range represents the actual variation in length, not the uncertainty of the measurement.[5] This corresponds to a DNA density within the phage head of about 0.56 bp per cubic nm which is similar to DNA densities found in crystals of of short dsDNA molecules.[26] There is no evidence for any proteins bound to the DNA in the P22 virion, and there would be little room for them if they were present. The nucleotide sequence of the P22 genome is 41,725 bp long and so on average each DNA has an approximately 1800 bp direct repeat (terminal redundancy) at the ends. Homologous recombination between these repeats is responsible for circularization of the DNA in the next round of infection.

Initiation of DNA Packaging

The first step in P22 (or any virus) DNA packaging is recognition of DNA to be packaged. Current evidence points strongly toward the P22 gene 3 protein being responsible for this recognition. P22 is a generalized transducing phage in that about 1-2% of the time it

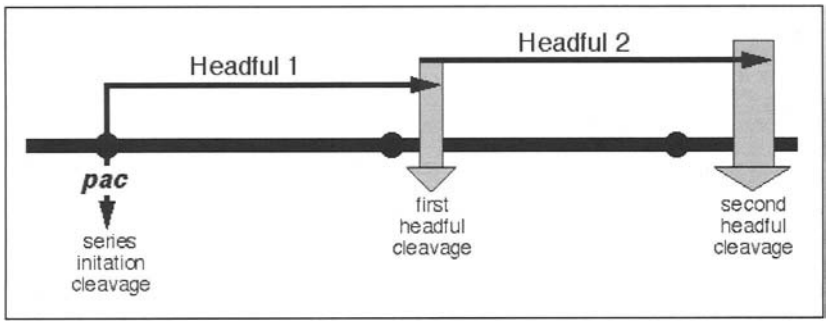


Figure 3. Processive packaging series. Concatemeric P22 DNA is represented by the thick horizontal line, and *pac* sites are represented by black circles. The vertical black line at the *pac* site indicates the location of series initiation cleavage for a processive packaging series, and horizontal arrows indicate the viral chromosomes cut from the concatemer in packaging events 1 and 2 of that series. A P22 packaging series starts at one *pac* site and packages DNA in one direction from that site. An average headful corresponds to 103.8% of the genome sequence (terminal redundancy not drawn to scale). The second packaging event in the series begins at the DNA end created by the headful cleavage of the previous event, and subsequent headfuls proceed in the same manner. Regions in which the headful cleavages take place are indicated by wide gray arrows. The inherent imprecision of the P22 headful measuring device is ±1.8% of the genome, and so the size of the headful cut region expands by this much with each subsequent headful in the series.

"mistakenly" initiates DNA packaging on the bacterial chromosome and encapsidates a phage chromosome-sized fragment of host DNA.[2] The resulting "virus" particle can deliver the host DNA it contains into another *Salmonella* cell as if it were a phage chromosome. Any such host DNA that is injected into such a cell can recombine with the resident chromosome to alter or "transduce" its genetic properties. Mutants originally isolated by Horst Schmieger as having higher than wild type frequencies of generalized transduction map in gene 3. Some of these do alter the target specificity of DNA recognition by gene 3 protein.[27-30] Such mutational changes in specificity are considered to be strong in vivo evidence of direct interaction between DNA and a putative DNA binding protein, since it is unlikely that altering one protein would cause it to, for example, change the specificity of a partner protein. Thus gene 3 protein is thought to be responsible for recognizing DNA to be packaged by P22. Gene 3 protein is found bound to gene 2 protein in infected cells.[31] Although its biochemical functions have not yet been successfully studied in the test tube, the gene 2/3 protein complex no doubt constitutes the P22 "terminase". The gene 2 and 3 proteins are not found in procapsids or in virions,[8] but these proteins are able to bind purified rings of portal protein.[32] Nonsense mutations in either of these genes completely block DNA packaging.[7] In addition, genes 3 and 2 (the latter encodes an ATPase motif, whereas the portal gene does not appear to encode such a motif) occupy positions on the chromosome expected for the recognition and DNA cleavage/ATPase subunits of terminase, respectively (The gene order in nearly all tailed-phages is small terminase - large terminase - portal - scaffold - coat; Fig. 2; see ref. 33).

The DNA-terminase complex formation and the DNA cleavage that initiates a packaging series appear not to be as simple as might have been imagined. Both terminase subunits are required for this in vivo cleavage of P22 DNA, however series initiation cleavage occurs in the absence of procapsids (it remains unknown why all *pac* sites are not cleaved in the absence of procapsids). The site on P22 DNA that is recognized by gene 3 protein, the *pac* site, has been characterized, and it is an asymmetric site about 22 bp long that lies inside gene 3.[29,34] The series initiation cleavage does not occur at a discrete site, but cleavages are scattered across an approximately 120 bp region (Fig. 2). The *pac* site itself lies near the center of this region. The

distribution of cuts within this region is not random, but cuts are (1) concentrated at 20 bp intervals, and (2) occur almost exclusively at 2 bp intervals throughout the region (Fig. 2). The locations of these cleavages are a property of the terminase alone, since they do not depend upon the presence of procapsids, are modified by mutations that alter the amino acid sequence of the gene 3 protein, but are not strongly affected by the nucleotide sequence outside the *pac* site itself.[30] The correct explanation for this complex pattern of cleavages is unknown, but it could be the result of a large repetitive protein complex being built on the DNA in which any of the units could perform the cleavage, or (2) movement of the terminase on the DNA after recognition of the *pac* site. P22 virion DNAs made under conditions in which the levels of terminase were lowered substantially have an unaltered *pac* series initiation DNA end pattern, suggesting that if model (1) is correct the complex does not enlarge along the DNA as more terminase is supplied. Recent findings with phage Sf6 suggest that (2) could be correct. Phage Sf6 is a close relative of P22 whose terminase is very different in sequence from that of P22. Sf6 appears to recognize a sequence within its gene 3 equivalent (its exact *pac* site has not yet been identified) but packaging series-initiating cleavages in the DNA are made over a nearly 2 kbp region.[35] The simplest model for this is a ± 1 kbp-diffusion of the terminase (nuclease) after recognition but before cleavage. A terminase complex large enough to span this length of DNA seems extremely unlikely, since only small amounts of terminase are made during infection. If sliding explains Sf6 series initiation, it could explain P22 as well.

After DNA is recognized by the gene 3 protein, a cleavage is made in the DNA and one of the ends created (the end to the right of the cleavage on the standard map) is inserted into the procapsid (Fig. 3). Analogy with other phage suggests that a gene 2 protein/gene 3 protein/DNA complex docks with the portal containing vertex of the procapsid, and that this interaction somehow inserts the DNA end into the portal hole, thus poising it for translocation into the capsid interior.

Filling the Capsid with DNA

As in other tailed phages, the motor that drives P22 DNA into the capsid is thought to be composed of the portal protein and terminase components. Because P22 is a headful packaging phage, its head filling device must have two parts—the force generating motor itself and a "headful sensing" device that controls the cleavage of DNA when the head is full. P22 heads contain DNA molecules that are on average 43,400 bp long, but their actual lengths vary from 42650 to 44150 bp.[5] The headful nuclease appears to have weak sequence specificity in vivo, since right end fragments of virion DNA are only created at a fraction of the bp in the headful cleavage region. Nonetheless, there are many potential cleavage sites in any region determined by the headful sensor to denote a headful of DNA (i.e., weak sequence specificity may determine the many exact cleavage sites that are observed within the general regions chosen for cleavage by the sensing device). Experiments with deleted P22 genomes,[5,36] mutants that package a longer DNA molecule than wildtype[37] and mutants with small T=4 capsids (Moore S and Prevelige P, personal communication) have demonstrated unambiguously that it is the size of the phage head that determines the length of DNA molecule encapsidated. Thus, this ±750 bp variation in chromosome length represents the ±1.8% imprecision of the headful-measuring device.

It is not known if the P22 DNA entry motor and headful sensor are separable functions or are two manifestations of the same device. Molecular genetic evidence suggests a more than simply passive role for the P22 portal protein in the DNA entry process. A mutant of P22 exists which encapsidates chromosomes that are 4.7% longer than wildtype. Although it is technically very difficult to directly prove that the mutant capsid interiors are not about 5% larger, measurements of internal DNA hydration showed that that the volume of water missing from these particles (compared to wildtype) is virtually identical to the volume of the "extra" DNA

present in the mutant virions.[37] It thus appears that the capsid interior volumes of the mutant and wildtype are the same, and the DNA length difference is the result of either a less sensitive headful-sensing device or of a motor that can push DNA in tighter or both. In either case, since this mutation alters a single amino acid in the portal, this finding shows that the portal ring is not a purely passive hole through which DNA is driven. In addition, experiments with procapsids that contain a mixture of mutant and wildtype portal proteins showed that the subunits of the portal ring can integrate information over the twelve subunits to determine where the headful cleavages occur, since a procapsid with about half wildtype and half mutant subunits packages a chromosome about 2.5% longer than wildtype, about halfway between procapsids with only wildtype or only mutant portal proteins.[26]

Although P22 DNA packaging is dependent upon ATP hydrolysis in crude extracts,[38] it has not been studied with purified components and the protein that hydrolyzes the ATP remains unknown. Nonetheless, it seems likely by analogy with other phages that the large terminase (gene 2 protein) subunit is the ATPase, and the P22 gene 2 protein does carry an ATPase amino acid sequence consensus.[39] Other circumstantial evidence suggests that P22 terminase may participate in DNA entry. Analysis of structures made by various mutants have shown that when gene 4, 10 or 26 protein is missing, DNA packaging occurs apparently normally, but the DNA is packaged unstably and falls out the particles before they can be isolated.[40,41] These three proteins are called "head completion protein" (see Fig. 1). The gene 4 and 10 proteins have been found only at the portal vertex of the mature virion,[42] and have been proposed to plug the portal hole so that DNA does not "leak" back out after packaging. Virion-like "empty heads" that have lost their DNA can be isolated from viruses containing mutations in genes 4, 10 or 26, and these empty heads have 5 to 10 molecules of both terminase subunits tightly bound to them[10] (Adams M, Casjens S, unpublished). Terminase subunits are also present in the "empty head" fraction of particles isolated from wildtype infections, suggesting that this is not an aberrant association seen only under the mutant conditions (Adams M, Casjens S, unpublished). Since it is required for packaging initiation and present in the particles after packaging is complete, it seems that terminase molecules are most likely also present and bound to the procapsid during the DNA entry process. It is quite possible, therefore, that they could participate directly in that process. Furthermore, the stoichometry of the bound terminase suggests that DNA packaging is carried out by a complex consisting of the procapsid and multiple terminase subunits. How the structure and organization of such a packaging complex might affect the packaging strategy is not known. Since we and others have failed to observe any ATPase activity by portal protein in any virus system, our current model for P22 DNA entry is thus that the DNA translocase motor consists of the portal ring and 5-15 molecules of each of the terminase subunits. We imagine, currently without proof (largely by analogy to other phage systems that are better studied in this regard), that the P22 large terminase subunit is both the ATPase that provides the packaging energy as well as the nuclease that acts during initiation of packaging series and headful DNA cleavage. The portal appears to be more than a passive hole, and is likely part of the headful-sensing device.

Termination of DNA Packaging

After the headful DNA cleavage releases the capsid containing the newly packaged chromosome from the concatemer, the three head completion proteins (the products of genes 4, 10 and 26) add to the particle and stabilize the packaged DNA (Fig. 1). After the head completion proteins add to the stucture, 18 molecules of one more protein, called "tailspike" protein, add to the portal vertex. Tailspike protein is responsible for binding virions to the outer surface of cells to initate the next round of infection.[3,4] Tailspike defective virions and wildtype virions contain no terminase subunits, but particles made by mutants defective in any of the three

head completion genes do contain them (above), suggesting that addition of the head completion proteins causes the release of terminase. The head completion proteins are known to add in the following obligate order (each in small numbers and to the portal vertex portion of the virion): gene 4 protein, gene 10 protein, and finally gene 26 protein.[41,42] Since 26 defective particles, which have bound 4 and 10 proteins, contain terminase, we have proposed that 26 protein addition is responsible for removing terminase from the particle to complete each DNA packaging event.[43]

It is not known what causes processive packaging series to end. Do packaging series terminate with an abortive attempt to package a too short DNA and simply run off the end of the DNA being packaged? Or might they collide with a replication apparatus? Or might the packaging machine have some way of only allowing a packaging event to start if enough DNA is available for a productive event? Many questions remain about headful packaging and the processive packaging series.

Summary

P22 virion assembly is one of the prototypic virus nucleic acid packaging systems. Its terminase, portal, scaffold, coat and head completion proteins have little sequence similarity to the analogous proteins of other well-studied dsDNA virus types, yet a perfect parallel exists between these P22 general functions and those of their analogs in other systems. This relationship suggests that either the large dsDNA viruses (perhaps including herpesviruses, and even iridoviruses and adenoviruses?[44,45]) have an extremely distant common ancestor from which all their DNA packaging machines are derived or that the procapsid/DNA packaging strategy represents an optimal solution arrived at more than once during the evolution of bacterial and eukaryotic viruses.

Acknowledgements

Research on P22 DNA packaging has been supported by the National Science Foundation (grant number MCB990526 to S.C.).

References

1. Zinder N, Lederberg J. Genetic exchange in Salmonella. J Bacteriol 1952; 64:679-699.
2. Ebel-Tsipis J, Botstein D, Fox MS. Generalized transduction by phage P22 in Salmonella typhimurium. I. Molecular origin of transducing DNA. J Mol Biol 1972; 71:433-448.
3. Poteete AR. Bacteriophage P22. In: Calendar R, ed. The Bacteriophages. New York: Plenum Press, 1988:647-682.
4. Susskind MM, Botstein D. Molecular genetics of bacteriophage P22. Microbiol Rev 1978; 42:385-413.
5. Casjens S, Hayden M. Analysis in vivo of the bacteriophage P22 headful nuclease. J Mol Biol 1988; 199:467-474.
6. Anderson T, Yamamoto N. Genomic masking and recombination between serologically unrelated phages P22 and P221. Virology 1961; 14:430-437.
7. Botstein D, Waddell CH, King J. Mechanism of head assembly and DNA encapsulation in Salmonella phage p22. I. Genes, proteins, structures and DNA maturation. J Mol Biol 1973; 80:669-695.
8. Casjens S, King J. P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly. J Supramol Struct 1974; 2:202-224.
9. King J, Lenk EV, Botstein D. Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. II. Morphogenetic pathway. J Mol Biol 1973; 80:697-731.
10. Poteete AR, King J. Functions of two new genes in Salmonella phage P22 assembly. Virology 1977; 76:725-739.
11. Casjens S. Molecular organization of the bacteriophage P22 coat protein shell. J Mol Biol 1979; 131:1-14.

12. Prasad BV, Prevelige PE, Marietta E et al. Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus. J Mol Biol 1993; 231:65-74.
13. Thuman-Commike PA, Greene B, Jakana J et al. Three-dimensional structure of scaffolding-containing phage p22 procapsids by electron cryo-microscopy. J Mol Biol 1996; 260:85-98.
14. Thuman-Commike PA, Greene B, Malinski JA et al. Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids. Biophys J 1999; 76:3267-3277.
15. Thuman-Commike PA, Greene B, Jakana J et al. Identification of additional coat-scaffolding interactions in a bacteriophage P22 mutant defective in maturation. J Virol 2000; 74:3871-3873.
16. King J, Casjens S. Catalytic head assembling protein in virus morphogenesis. Nature 1974; 251:112-119.
17. Bazinet C, King J. Initiation of P22 procapsid assembly in vivo. J Mol Biol 1988; 202:77-86.
18. Earnshaw W, Casjens S. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21:319-331.
19. Greene B, King J. Scaffolding mutants identifying domains required for P22 procapsid assembly and maturation. Virology 1996; 225:82-96.
20. Bazinet C, Benbasat J, King J et al. Purification and organization of the gene 1 portal protein required for phage P22 DNA packaging. Biochemistry 1988; 27:1849-1856.
21. Botstein D, Levine M. Intermediates in the synthesis of phage P22 DNA. Cold Spring Harb Symp Quant Biol 1968; 33:659-667.
22. Streisinger G, Enrich J, Stahl M. Chromosome structure in T4. III. Terminal redundancy and length determination. Proc Natl Acad Sci USA 1967; 57:292-295.
23. Tye BK, Botstein D. P22 morphogenesis. II: Mechanism of DNA encapsulation. J Supramol Struct 1974; 2:225-238.
24. Jackson EN, Jackson DA, Deans RJ. EcoRI analysis of bacteriophage P22 DNA packaging. J Mol Biol 1978; 118:365-388.
25. Adams MB, Hayden M, Casjens S. On the sequential packaging of bacteriophage P22 DNA. J Virol 1983; 46:673-677.
26. Casjens S, Wyckoff E, Hayden M et al. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. J Mol Biol 1992; 224:1055-1074.
27. Schmieger H. Phage P22-mutants with increased or decreased transduction abilities. Mol Gen Genet 1972; 119:75-88.
28. Schmieger H, Backhaus H. The origin of DNA in transducing particles in P22-mutants with increased transduction-frequencies (HT-mutants). Mol Gen Genet 1973; 120:181-190.
29. Casjens S, Huang WM, Hayden M et al. Initiation of bacteriophage P22 DNA packaging series. Analysis of a mutant that alters the DNA target specificity of the packaging apparatus. J Mol Biol 1987; 194:411-422.
30. Casjens S, Sampson L, Randall S et al. Molecular genetic analysis of bacteriophage P22 gene 3 product, a protein involved in the initiation of headful DNA packaging. J Mol Biol 1992; 227:1086-1099.
31. Poteete AR, Botstein D. Purification and properties of proteins essential to DNA encapsulation by phage P22. Virology 1979; 95:565-573.
32. Moore SD, Prevelidge P. Bacteriophage P22 portal vertex formation in vivo. J Mol Biol 2002; 315:975-994.
33. Casjens S, Hendrix R. Control mechanisms in dsDNA bacteriophage assembly. In: Calendar R, ed. The Bacteriophages. New York: Plenum Press, 1988:15-91.
34. Wu H, Sampson L, Parr R et al. The DNA site utilized by bacteriophage P22 for initiation of DNA packaging. Molec Micro 2002; 45:1631-1646.
35. Casjens S, Winn D, Gilcrease L et al. The chromosome of Shigella flexneri bacteriophage Sf6: Complete nucleotide sequence, genetic mosaicism, and DNA packaging. J Mol Biol 2004; 339:379-394.
36. Tye BK, Huberman JA, Botstein D. Nonrandom circular permutation of phage P22 DNA. J Mol Biol 1974; 85:501-528.

37. Casjens S, Wyckoff E, Hayden M et al. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. J Mol Biol 1992; 224:1055-74.
38. Poteete AR, Jarvik V, Botstein D. Encapsulation of phage P22 DNA in vitro. Virology 1979; 95:550-564.
39. Eppler K, Wyckoff E, Goates J et al. Nucleotide sequence of the bacteriophage P22 genes required for DNA packaging. Virology 1991; 183:519-538.
40. Lenk E, Casjens S, Weeks J et al. Intracellular visualization of precursor capsids in phage P22 mutant infected cells. Virology 1975; 68:182-199.
41. Strauss H, King J. Steps in the stabilization of newly packaged DNA during phage P22 morphogenesis. J Mol Biol 1984; 172:523-543.
42. Hartweig E, Bazinet C, King J. Injection apparatus of phage P22. Biophys J 1986; 49:24-27.
43. Casjens S. Bacteriophage P22 DNA packaging. In: Adolph K, ed. Chromosomes, Eukaryotic, Prokaryotic and Viral. Boca Raton: CRC Press, 1989:241-261.
44. Casjens S. Principles of virion structure, function and assembly. In: Chiu W, Burnett R, Garcea R, eds. Structural Biology of Viruses. Oxford University Press: Oxford, 1997:3-37.
45. Newcomb WW, Juhas RM, Thomsen DR et al. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. J Virol 2001; 75:10923-10932.

CHAPTER 6

Bacteriophage SPP1 DNA Packaging

Anja Dröge and Paulo Tavares

Introduction

SPP1 is a virulent double-stranded DNA (dsDNA) phage that infects the Gram-positive bacterium *Bacillus subtilis* strain 168. SPP1 belongs to the Siphoviridae family. The virion is composed of an icosahedral, isometric capsid (~60 nm diameter) and a long, flexible, noncontractile tail.[1] The phage head encloses the mature SPP1 chromosome, which is a linear double-stranded DNA molecule of ~45.9 kbp size.[2,3] Replication of the SPP1 genome produces DNA concatemers from which single chromosomes are cleaved and encapsidated into preformed procapsids. SPP1 packages its DNA by a headful packaging mechanism, a strategy first described for bacteriophage T4.[4] As for phages P22 or P1, in case of SPP1 DNA packaging starts by recognition and cleavage of a defined nucleotide sequence termed *pac*.[5,6] The concatemeric DNA is then translocated unidirectionally to the viral capsid interior. When the level of capsid DNA headfilling reaches a threshold amount (headful) a second endonucleolytic cleavage occurs. This second cleavage, the "headful cut", defines the size of the packaged DNA. In contrast to *pac* cleavage the sequence independent headful cut is imprecise. The size of mature SPP1 DNA therefore varies for about 6% around the average chromosome size of 45.9 kb.[7,8] Since the genome of SPP1 encompasses 44,007 bp,[9] the chromosomes generated by headful cut are terminally redundant (~4.1%). While the first round of packaging is initiated by a sequence specific cut at *pac*, the second and subsequent packaging events start from the DNA end generated by the headful cut of the previous packaging event. Therefore this strategy generates a population of phage chromosomes that are partially circularly permutated and terminally redundant (Fig. 1). Terminal redundancy of the packaged DNA molecule is essential for circularisation of the viral genome by homologous recombination after its injection in the host cell at the beginning of a new infection cycle.

The aim of this article is to review recent progress on the understanding of the molecular mechanisms used by phage SPP1 to package its DNA and to provide an integrated view of our current knowledge of this process. We identify also some of the central questions that are likely to be tackled in the near future. Reference to other virus systems is made only when of direct relevance for the comprehension of SPP1 DNA packaging.

Assembly of the SPP1 Procapsid, the Proteinaceous DNA Container

The procapsid (or prohead) of SPP1 consists of four proteins: the major capsid protein gp13, the scaffold protein gp11, the portal protein gp6 and a minor component gp7[10] (Table 1). Gp13 forms the head shell of the SPP1 procapsid. The protein composition of the SPP1 procapsid suggests that 415 copies of gp13 form a T=7 icosahedral head shell lattice.[11] The

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

Table 1. Properties of the SPP1 proteins that participate in capsid assembly and DNA packaging

SPP1 Protein	Monomer M_r (kDa)	Oligomeric State in Solution	Number of Copies in Virion	Assembly Intermediates	Essential Role	Function and Activities	References
Procapsid assembly							
gp6	57.3	13-mer	12	PC, CP, V	Yes	**Portal protein**, initiation of procapsid assembly, component of DNA translocation machine, headful sensor, connector assembly	3,11,19,20,22, 24,28,35
gp7	35.1	Monomer	2-3	PC, CP, V	No	Function unknown, absence reduces SPP1 burst 5-fold, binds gp6	10,11,14, unpublished results
gp11	23.5	14 to 16-mer	None	PC	Yes	**Scaffolding protein**, directs polymerisation of the major capsid protein gp13	10,11
gp13	35.4	Variable number	415	PC, CP, V	Yes	**Major capsid protein**	10,11
DNA packaging							
gp1	20.8	10-mer	None	–	Yes	**Terminase small subunit**, specific binding to *pac*, modulates endonuclease and ATPase activity of gp2	6,15-18
gp2	48.8	Monomer	None	–	Yes	**Terminase large subunit**, ATPase, non-specific endonuclease, interacts with gp1	6,17,18,21
Stabilisation of the packaged DNA							
gp12	6.6	Unknown	~70	CP, V	No	**Capsid decoration protein**	10, unpublished results
gp15	11.6	Unknown	12	CP, V	Yes	**Head completion protein**, prevents leakage of packaged DNA	10,22,35
gp16	12.5	Unknown	12	CP, V	Yes	**Head completion protein**, prevents leakage of packaged DNA, proteolysed during assembly	22,35

PC—procapsid; CP—DNA-filled capsid; V—virion.

interior of the structure is filled with 100-180 copies of gp11 that exit the procapsid at the beginning of DNA packaging. A complex of either 12 or 13 copies of the portal protein is present at a unique vertex of the icosahedron termed the portal vertex, the site of DNA entry during encapsidation of the phage chromosome. Furthermore the procapsid contains 2-3 copies of the minor capsid protein gp7.[11]

The role of the individual procapsid components for procapsid assembly was defined by the characterization of SPP1 mutants carrying conditional lethal mutations in the corresponding genes.[10] Expression of various combinations of the cloned procapsid genes in the bacterium *Escherichia coli* allowed establishing the network of protein-protein interactions during procapsid assembly and their role in procapsid shape determination.[11] The four procapsid proteins were shown to assemble into procapsids competent for DNA packaging in vitro.[11,12] During assembly of the head shell lattice the scaffold protein controls the polymerisation of gp13 into an icosahedral shell. Incorporation of the portal complex requires interaction with both the capsid and the scaffolding protein. A particular feature of SPP1 is that the portal protein regulates the size of the procapsid. Coproduction of gp11 and gp13 in *E. coli* in absence of the portal protein or during infections with SPP1 mutants deficient for portal protein production leads to formation of a mixed population of T=7 and smaller procapsids.[11] However, in presence of the portal protein formation of small sized procapsids is suppressed. The decision whether a T=7 or a smaller procapsid is assembled is made at an initial stage of procapsid assembly.[13] This observation indicates that the portal protein together with the coat protein and the scaffold is part of an initiation complex from which subsequently polymerisation of the head shell proceeds. Presence of the portal protein in the procapsid is an essential requirement for DNA packaging.

The role of the fourth procapsid protein in SPP1 assembly, gp7, is poorly understood. Gp7 strongly binds to gp6. This interaction recruits gp7 to the procapsid structure but it is disrupted at a late stage of SPP1 assembly leaving gp7 probably bound to the DNA packed inside the capsid.[14] Gp7 is not essential for the formation of biologically active procapsids but its absence leads to a 5 to 10-fold reduction in the biological activity of phages formed in DNA packaging reactions in vitro.[11]

Procapsid assembly: questions for future research

- How is the portal oligomer recruited to the initiation complex of procapsid assembly?
- What is the molecular nature of the initiation complex in procapsid assembly?
- How is the portal protein embedded in a nonmatching symmetry environment?
- What is the function of gp7 in SPP1 assembly?

Selective Recognition and Cleavage of SPP1 DNA by the Terminase Complex gp1-gp2

Packaging of SPP1 DNA into procapsids is initiated by *pac* cleavage of the concatemeric DNA precursor (Fig. 1). This endonucleolytic cleavage is performed by the terminase enzyme. The terminase of SPP1 is a hetero-oligomer of small (gp1) and large (gp2) subunits. The holoenzyme possesses several activities including recognition, binding and cleavage of the DNA substrate, and ATP-hydrolysis.

Gp1 is an oligomer in solution with a native mass of about 200 kDa. Electron micrographs of negatively stained gp1 molecules and analysis of the native protein by gel filtration suggested that ten molecules of gp1 form a ring-shaped particle.[15,16] Gp1 is composed of several functional domains.[15] Information about this domain organisation of gp1 was gained by

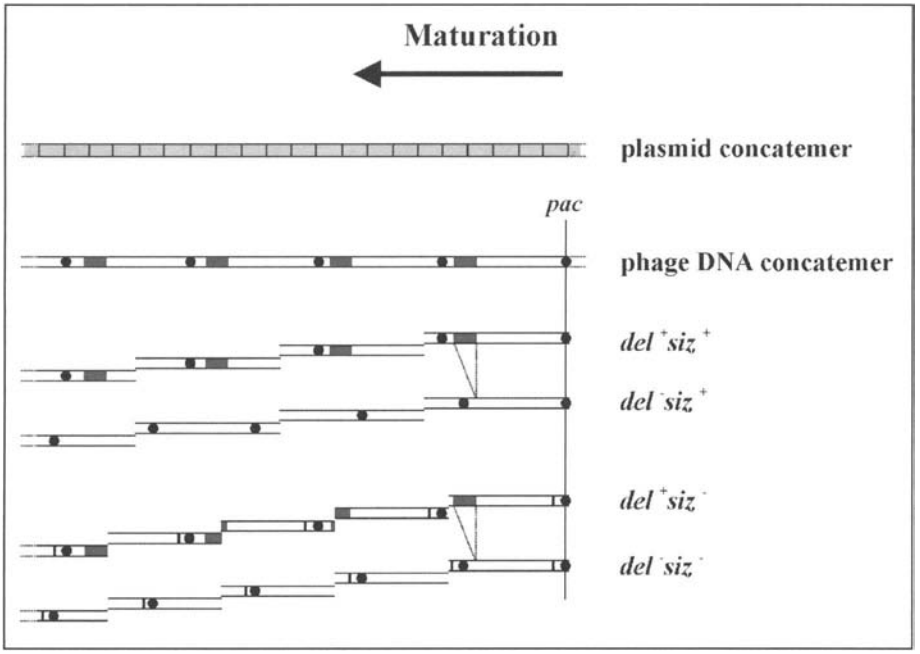


Figure 1. Processive sequential packaging of the SPP1 genome by the headful mechanism.[2,3] The substrate for packaging is a concatemer of the viral genome generated by replication of the SPP1 DNA shown as a white bar with black circles (*pac* sequence) and dark grey rectangles (non-essential region of the SPP1 genome). Headfuls of host DNA can be encapsidated at low frequency, as in case of plasmid concatemers (top bar), and delivered by the SPP1 particle to a host cell (transduction). SPP1 DNA packaging is unidirectional (top arrow). It initiates at *pac* and sequential encapsidation cycles proceed along the SPP1 concatemer. SPP1 wild type (*del*$^+$ *siz*$^+$) packages a 45.9 kbp headful that corresponds to a DNA molecule with 104 % the size of the genome (44,007 bp) having therefore a terminal redundancy of 4 %. A deletion in the phage genome (*del*$^-$ *siz*$^+$) is compensated by a proportional increase in the molecule terminal redundancy leading to packaged molecules with the same size as those found in wild type virions. Mutations *siz* (thin vertical bars) cause single amino acid substitutions in the SPP1 portal protein that lead to the encapsidation of a shorter viral chromosome. The chromosome size is not affected by the absence (*del*$^+$ *siz*$^-$) or presence (*del*$^-$ *siz*$^-$) of deletions in the SPP1 genome.

comparing the sequence of gp1 with the sequence of the corresponding functional analog of closely related phages (SF6 and ρ15) and functional analysis of chimeras between the SPP1 and SF6 proteins[15] (Fig. 2). The two proteins show 71 % identity that is clustered in three discrete regions (I, II and III). Segment I contains a putative NTP binding motif (B-type) and a DNA-binding motif (putative helix-turn-helix motif).[15,16] Furthermore a gp1-gp2 interaction region is located in segment I.[17] The region responsible for the interaction between gp1 subunits is located in segment II, the central part of the protein[17] (Fig. 2). Even though the segment II additionally contains a NTP hydrolysis motif, purified gp1 shows no ATPase activity in vitro.[15] No function could yet be assigned to segment III.

Gp1 alone binds specifically to SPP1 DNA at the packaging initiation sequence, *pac* (Fig. 3), but is unable to cleave it. The target *pac* sequence is asymmetric. DNAse I footprints have shown that gp1 binds to two sites within *pac* that were termed *pac*L, at the nonencapsidated left end, and *pac*R, at the encapsidated right end.[16] These two sites flank the site of cleavage *pac*C (Fig. 3). The intrinsically bent sequence *pac*L and *pac*R share no sequence homology and

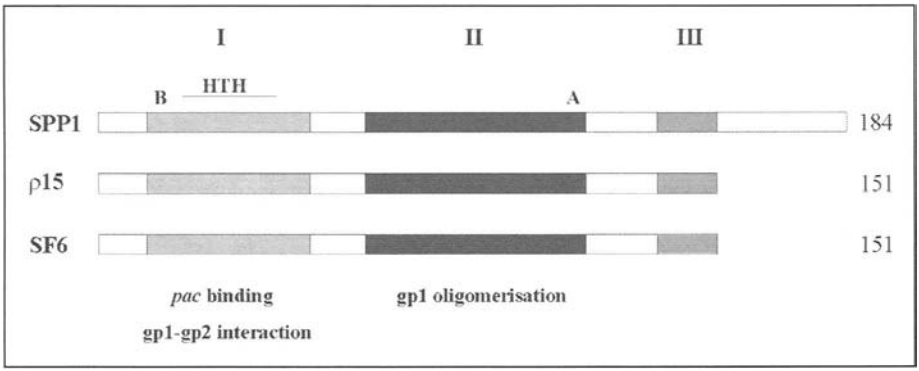


Figure 2. Comparison between the amino acid sequences of gp1 from the homologous phages SPP1, ρ15, and SF6. The number of amino acids of each gp1 form is shown on the right side. High amino acid sequence homology between the three proteins is found in three segments (I, II, and III).[15] These segments include a putative helix-turn-helix (HTH) DNA binding motif and the Walker NTP binding motifs A and B in the positions indicated on top of the SPP1 gp1 sequence.[15] The activities associated to each segment that are shown in the bottom part of the figure are derived from functional studies.[15,17]

form dissimilar complexes with gp1. In case of *pac*L the pattern of DNaseI cleavage indicates that *pac*L DNA wraps around the gp1 decamer. Based on the cooperative binding of the two gp1 decamers to *pac* DNA and based on footprinting data it was suggested that the gp1 decamers bound to *pac*L and *pac*R interact and thereby generate a DNA loop of 204 bp length that includes *pac*C[16] (Fig. 4). This model is further supported by the finding that purified gp1 promotes cyclization in vitro of a DNA fragment containing *pac*L.[16] The ensemble of the data showed that binding of gp1 to *pac* generates an asymmetric nucleoprotein complex.

The large subunit of the terminase, gp2, is a protein of 422 residues with a predicted molecular mass of 49 kDa. Analysis of gp2 has long been hampered because the protein could not be overproduced and purified in adequate amounts. However, this problem was recently overcome by affinity purification of hexa-histidine gp2 (H6-gp2), a gp2 derivative that fully complements the deficiency of SPP1 conditional lethal mutants in gene *2*.[6,17] H6-gp2 interacts with gp1, binds to dsDNA in a sequence independent manner and possesses a modest ATPase activity.[6] H6-gp2 introduces nicks and breaks in dsDNA with poor substrate specificity. Interestingly, in presence of distamycin H6-gp2 specifically nicks both DNA strands at the bona fide *pac*C sequence. Distamycin competes with gp1-*pac* complex formation and, like gp1, introduces local distortions into DNA. It was therefore suggested, that the DNA structure generated by gp1 promotes cleavage at *pac*C by gp2. Interaction with the two gp1 decamers was proposed to place two molecules of gp2 in close vicinity to the b sites of *pac*C (Figs. 3,4) within the *pac*-terminase nucleoprotein complex.[6] Upon cleavage at both b boxes the gp2 molecule bound at the b box proximal to *pac*L would unbind gp1 and degrade the nonencapsidated left DNA end (Fig. 4). The gp2 molecule bound to the b box proximal to *pac*R, however, remains in the packaging complex. Interaction of the gp2 monomer with both gp1 decamers would stimulate the ATPase activity of gp2 thereby energising ATP-driven DNA translocation.[6] The model is supported by the fact that raising the concentration of gp1 stimulates the ATPase activity of H6-gp2 but exerts a negative effect on the H6-gp2 endonuclease activity in vitro.[6] The terminase-DNA complex subsequently participates in DNA translocation to the procapsid interior through the portal protein pore.

In the model described above asymmetry of the *pac* site causes formation of an asymmetric DNA-terminase complex, thereby providing the structural basis for discrimination between

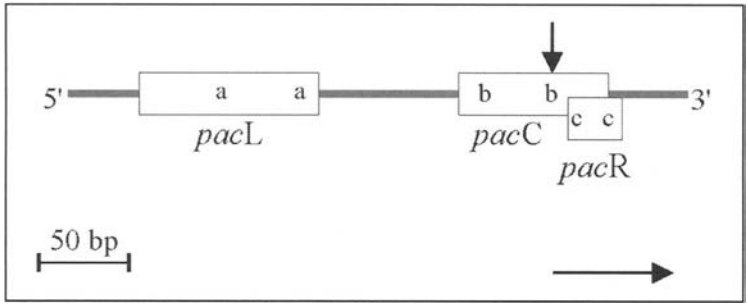


Figure 3. Bacteriophage SPP1 *pac* site. Gp1 binds to *pac*L (nonencapsidated DNA end, or left end) and *pac*R (encapsidated DNA end or right end) promoting 1 bp staggered nicking at the *pac*C site that is carried out by the gp2 endonuclease activity.[6,16] The predominant site of terminase cleavage which defines the first nucleotide of the SPP1 sequence is indicated by the vertical arrow. The position of directly repeated sequences (a, b, c) localised within the *pac*L, *pac*R and *pac*C boxes are shown. The direction of DNA encapsidation is indicated by the horizontal arrow in the bottom (note that the *pac* sequence is schematised here with an orientation opposite to the *pac* in Fig. 1). A model for the terminase-*pac* nucleoprotein complex was proposed[16] (Fig. 4).

the nonencapsidated left end and the encapsidated right end of the DNA concatemer.[6,16] This mechanism ensures the unidirectionality of DNA packaging.

Terminase-DNA interaction: questions for future research

- What is the exact mechanism ensuring sequence specific cleavage at *pac* by gp2?
- How is *pac* cleavage controlled and how is excessive *pac* cleavage throughout the DNA concatemer prevented?
- How does the terminase switch between the DNA cleavage and DNA translocation activities?

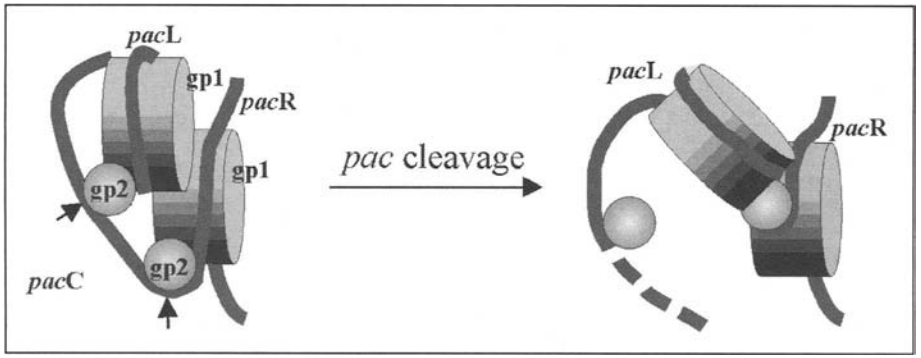


Figure 4. Model of the terminase-*pac* nucleoprotein complex before and after *pac* cleavage.[6,16] The gp1 cyclical decamers bound to *pac*L and *pac*R were proposed to position two molecules of gp2 to cut at the positions indicated by the arrows within the *pac*C sequence. After the endonucleolytic cleavage, one gp2 subunit would unbind gp1 and degrade the nonpackaged DNA end while the other DNA end is encapsidated in the viral procapsid (see text and ref. 6 for details). Organisation of the *pac* sequence is shown in Figure 3.

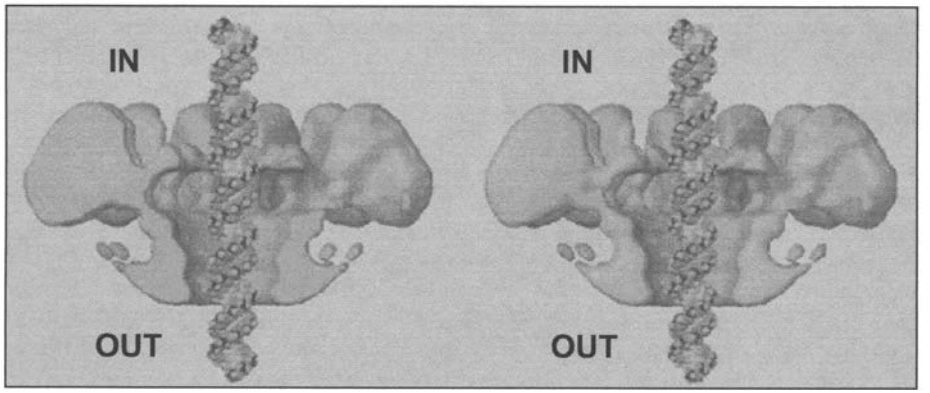


Figure 5. Structure of the SPP1 portal protein at 18 Å resolution. Stereo view of wild type gp6 oligomer cut-open along a vertical plane. A B-form DNA structure is modelled in the central channel to illustrate the hypothesis that gp6 encircles the double-helix during DNA packaging. The gp6 structure shown is from the isolated molecule, which is a 13-mer.[28] Even though the gp6 found in the assembled SPP1 virion is a 12-mer,[22,35] implying that the central gp6 channel is a few angstrom narrower, the potential regions of gp6-DNA interaction can be inferred from the figure. The wide region of the molecule faces the interior (IN) of the viral capsid[22,28,35] while the lower stem is exposed to the outside (OUT) or tail side of the capsid. Approximately three helical turns of DNA fit in the gp6 channel. Courtesy of Dr. E.V. Orlova.

Assembly of the DNA Packaging Machine, DNA Translocation, and Capsid Expansion

Specific recognition of the viral DNA and cleavage at *pac* is mediated by the terminase complex in the absence of other SPP1 proteins.[6,18] Productive SPP1 DNA encapsidation, however, requires the presence of procapsid structures containing an active portal protein.[11,19,20] This feature is common to every tailed phage system presently characterised. In analogy to other phage systems, in case of SPP1 the viral DNA-terminase complex is thought to dock at the portal vertex to assemble the machinery that pumps DNA to the procapsid interior.[21] The interaction activates the ATPase activity of the terminase large subunit gp2 in order to power DNA translocation to the procapsid interior.[21] The organisation of the DNA-terminase-procapsid ternary complex in SPP1 is not known. DNA packaging in vitro using crude cell extracts showed that neither the presence of the isolated portal protein nor of procapsids lacking the portal structure inhibits DNA packaging in a competitive fashion.[12] The terminase-DNA complex therefore has the capacity to discriminate between the pool of free portal protein, procapsids incompetent for DNA packaging and the portal protein embedded in the procapsid portal vertex. This distinction is essential to avoid nonproductive interactions that would lead to abortive packaging reactions. A possible structural basis for the specificity of the terminase-DNA complex with procapsid-embedded portal could be a structural transition of the portal protein. While isolated gp6 displays 13-fold symmetry, gp6 with a 12-fold symmetry is found in DNA-filled capsids. This change was proposed to occur during procapsid assembly.[22]

The working hypothesis for DNA encapsidation is that chemical energy generated by the ATPase activity of gp2[6,21] is converted to mechanical energy by the packaging apparatus powering DNA translocation against a steep concentration gradient. DNA movements to enter and exit the viral capsid are believed to occur through the portal protein central channel (Fig. 5). Single amino acid changes in gp6 were shown to block or reduce the efficiency of DNA translocation revealing the central role of the portal structure in this process.[19,20] The most recent models for DNA translocation are built on the seminal concept proposed by Roger

Hendrix that symmetry mismatches between components of the translocation machine provide the structural basis for a DNA rotary pumping action carried out by the portal protein.[23] A model that exploits the symmetry mismatch between the 10-fold helical symmetry of DNA and the symmetry of the portal protein was proposed.[24] The model was originally developed for a 13-fold symmetric portal protein but it applies also to a protein composed of 12 subunits. The portal protein was proposed to interact with the DNA phosphate backbone leading to the pumping step of 2 base-pairs through the portal channel. The step would consume 1 ATP molecule hydrolysed by the terminase. The portal protein rotates relative to the DNA and to the capsid in order to bring a new subunit in contact to the DNA molecule to exert the following pumping step. This strategy ensures a constant positioning of the portal protein subunits relative to the DNA and leads to translation of the double-helix to the capsid interior without significant rotations relative to the capsid.[24] Determination of the portal protein structure from bacteriophage ϕ29 provided enough detail to propose a more elaborated molecular mechanism for DNA translocation that exploits also the symmetry interactions between the portal protein and DNA[25] (Grimes and Anderson, this volume).

DNA packaging leads to a drastic conformational change of the SPP1 major capsid protein. The roundish procapsid with an outer diameter (d_o) of about 55 nm expands to a larger capsid structure (d_o ~ 66 nm) with a hexagonal outline.[11] It is not known if the trigger for expansion is a signal initiated at the portal vertex during DNA encapsidation or if it is due to the increasing pressure generated by the packed DNA on the capsid lattice inner surface (cf., refs. 20,26).

The DNA packaging machine and DNA translocation: questions for future research

- What is the structural organization of the DNA translocating machine?
- How is the chemical energy generated by the terminase ATP hydrolysis transduced to mechanical pumping of DNA?
- Does the pumping of DNA require rotation of the portal protein?
- How is DNA organized inside the capsid and what stabilizes its tight packing?

Termination of DNA Packaging: The Headful Cleavage

Encapsidation of the SPP1 chromosome is terminated by endonucleolytic cleavage of the viral DNA concatemer. The site of this sequence independent cleavage is determined by the mass of DNA encapsidated. The mechanism requires a sensor that measures the level of DNA headfilling and an effector headful nuclease. Mutations *siz* (for sizing) that affect the mass of DNA packaged were found exclusively in the portal protein coding gene[3] (Fig. 1), a feature also reported for bacteriophage P22.[27] The *siz* mutations lead to an undersizing of the mature viral chromosome that correlates with a reduced packing density of DNA inside the viral capsid.[3] The capsid size does not change in these mutants. The topology of the portal oligomer in the capsid structure is ideally suited for a sensor activity. The molecule wider region is exposed to the capsid interior while its stem faces the capsid outside where the terminase is believed to bind. The central channel of gp6 is delimited by a tentacles fringe in the wider area of the molecule (Fig. 5). The fringe is flexible and a significant difference on its structure was found when the wild type portal protein and a SizA sensor mutant portal protein were compared.[28] The pressure generated on this flexible sensor domain by the increasing amount of packaged DNA was proposed to be the trigger for headful cleavage.

Combination of *siz* mutations in the portal protein leads to packaging of viral chromosomes significantly smaller than the SPP1 mature chromosome of single *siz* mutants.[29]

Additionally the efficiency of packaging is reduced in a number of these multiple gp6 mutants suggesting a cross-talk between some step in DNA encapsidation (e.g., DNA translocation) and the headful trigger/cleavage.[20,29] The data accumulated is compatible with a headful mechanism in which DNA translocation stops and renders the DNA molecule accessible to headful endonuclease attack. Stopping of DNA translocation might be caused by a conformational change of the headful sensor domain that physically blocks the DNA in the portal channel and/or by the incapacity of the DNA translocation machine to exert force for further DNA pumping to the capsid interior. The latter case implies a direct role of the portal protein in the mechanical DNA translocation action as discussed in the previous section. Arrest of DNA translocation would trigger headful cleavage carried out most likely by the terminase large subunit gp2. The process might be associated to a reduction in the gp1/gp2 ratio, a condition shown to de-repress the gp2 endonuclease activity and to inhibit its ATPase activity.[6,21] The studies described provide the first molecular clues to understand the mechanism of headful cleavage.

The headful packaging mechanism: questions for future research

- How does the portal protein measure the level of DNA headfilling inside the viral capsid?
- What is the signal that triggers headful cleavage and how is it transduced to the headful nuclease?
- How is the activity of the headful nuclease (terminase?) regulated?
- Are headful cleavage and binding of the head completion proteins coupled (see also next section)?

Stabilization of Packaged DNA and Control of DNA Release. Assembly and Structure of the SPP1 Connector

DNA is packaged at a high density inside the viral capsid of tailed bacteriophages. A direct correlation between the amount of encapsidated DNA and the sensitivity to chelating agents was demonstrated for a variety of phages including SPP1.[3,30] Divalent cations thus play a central role to stabilise the DNA-filled capsid most probably by neutralizing the negative charges of closely packed DNA phosphate backbones. They might also relieve repulsion between DNA and the capsid lattice interior in case the latter is negatively charged as shown for phage HK97.[26] The DNA packing generates an internal pressure that can lead to its spontaneous release from the phage capsid.[8,31-34]

After termination of DNA packaging the portal pore needs to be rapidly closed to prevent leakage of the viral chromosome—a process that must be reversed when the virus infects a host cell. In SPP1 this role is achieved by the head completion proteins gp15 and gp16 that bind sequentially to the portal vertex forming two rings of subunits stacked below the portal protein.[22,35] We call the complete structure connector (Fig. 6). A central channel that crosses the connector appears closed at the level of the gp16 ring.[22,35] In phages whose capsid was disrupted the connector-tail complex was shown to protect a short DNA fragment of 187 to 288 Å, a size that fits well the height of the connector. This DNA fragment is the extremity of the viral chromosome that is packaged last and is first to exit the virion when ejection is triggered.[8] The connector is a dynamic structure that binds the viral chromosome, serves as an interface for attachment of the helical tail, and controls DNA release from the virus. When the phage adsorption apparatus interacts with its bacterial receptor a signal is communicated along the tail structure to the connector region to trigger opening of the gp16 valve. One SPP1 chromosome extremity is fixed to this valve ready for polarised ejection from the viral capsid.

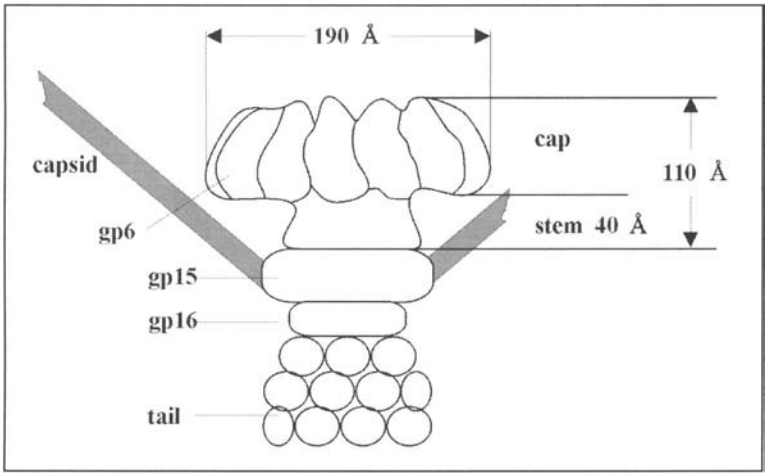


Figure 6. Organisation of the SPP1 head-to-tail interface.[22,35] The protein components of the connector structure are identified.

Processivity of DNA Packaging Events during SPP1 Infection

After termination of the first packaging cycle initiated at *pac* (initiation cycle) a second packaging event is initiated at the nonencapsidated concatemer extremity created by the headful cleavage and additional cycles of encapsidation follow (processive headfuls) (Fig. 1). Processivity of SPP1 DNA packaging is high reaching an average of 5-6 headfuls *per* concatemer. Some of the packaging series can even yield 12 headfuls or more.[8] The main factor that limits the series of encapsidation cycles during normal infection is host lysis.[8] Processive DNA packaging requires the generation of concatemers as long as 500 kbp within the time frame of SPP1 infection. It is possible that after the first cut at *pac* the packaging apparatus follows closely the replication machinery in a dynamic process that would limit the physical length of the concatemer and ensure that newly generated *pac* sites would be encapsidated before terminase attack.[8] Long packaging series do indeed absolutely require that the terminase leaves the *pac* sequences along the substrate concatemer being encapsidated intact. Even if only ~25% of all the *pac* sequences were cut randomly along the concatemer[18] the size of the substrate DNA would be considerably limited. Cleavage at *pac* sequences of concatemers where the initiation packaging cycle occurred must thus be avoided. The observation that DNA encapsidated in vitro when extracts of infected cells are mixed originates exclusively from the terminase donor extract suggests that this extract does not provide active free terminase to bind SPP1 DNA present in the extract lacking terminase.[12] The terminase-DNA complex is therefore the form competent to participate in the DNA packaging reaction in vitro while free terminase might be significantly unstable. Since the number of terminase small subunit decamers (gp1), responsible for recognition of *pac*, is auto-regulated[36] a limited number of terminases are likely to be loaded to the DNA substrate at early stages of the virus infection cycle. The terminase then remains stably associated to the same concatemeric molecule and processive DNA packaging predominates during the rest of the infection cycle. This strategy would ensure specificity of viral DNA encapsidation (terminase binding at *pac*) and furthermore maintain the integrity of the SPP1 DNA concatemer by avoiding excessive cleavage at *pac* sequences and thus ensure occurrence of long processive packaging series along a concatemer.[8] Similar observations were reported for bacteriophage P22.[37]

SPP1-Mediated Transduction: Packaging of Host DNA

Bacteriophage SPP1 mediates generalised transduction.[38] This process involves the transfer of nonviral genetic material between a donor cell and a receptor cell mediated by a virus. SPP1 transducing particles are indistinguishable from SPP1 virions with the exception that they contain a host DNA molecule with a size similar to that of the mature SPP1 chromosome.[39] In absence of homology between the host and SPP1 DNAs the frequency of transduction is very low: $10^{-7} - 10^{-8}$ for host chromosome genes[38] and $10^{-4} - 10^{-5}$ for plasmid DNA.[40] One SPP1 particle can transduce approximately 1 % of the *B. subtilis* chromosome.[38] Plasmid DNA, which is usually smaller than a phage chromosome unit-length, is encapsidated as a linear concatemer containing multiple copies of the plasmid in a head-to-tail arrangement (Fig. 1). SPP1 infection, as in the case of other phages, was shown to change the mode of plasmid replication to generate plasmid concatemers that are the substrate for DNA packaging.[41,42] Generalised transduction is attributed to a low frequency error of the terminase to recognise and initiate processive headful packaging at *pac*-related sequences in the host DNA or, in some cases, at DNA free ends.

The efficiency of transduction is significantly increased in presence of homology (>47 base pairs) between the host DNA molecule and the virus chromosome.[43,44] This facilitated transduction is attributed to homologous recombination between the two replicons yielding chimeric phage-host DNA molecules. After initiating encapsidation at a *pac* sequence of the phage genome the highly processive phage packaging machinery would then access host DNA through the region of recombination and package headfuls of it. Presence of *pac* in the host cell DNA leads to packaging initiation at this sequence and to unidirectional encapsidation of host DNA.[45]

Concluding Remarks

The phenomenology of bacteriophage SPP1 DNA packaging is well known and some of the molecular mechanisms involved are among the best understood in tailed phages systems. These include the terminase recognition and cleavage of its target sequence *pac*, the structure and function of the portal protein, the mechanism of headful sensor, and connector assembly. Other aspects of the DNA packaging process were not yet studied in detail like the terminase-procapsid interaction, the properties and assembly of the DNA translocating complex, or the mechanism of DNA translocation. A complete picture of the DNA packaging process requires knowledge of all these molecular mechanisms and their integration in the dynamics of the infected cell. An interdisciplinary approach combining genetics, biochemistry, structural biology and bacterial cell biology, on one side, and comparative analysis of the strategies used in different virus systems, on the other side, will undoubtedly be necessary to characterise the unifying mechanisms optimised by tailed phages and herpesviruses to encapsidate their viral genome.

Acknowledgments

We thank T.A. Trautner for critically reading the review and for his long-lasting engagement and support to SPP1 research. We thank E.V. Orlova for kindly providing (Fig. 5).

References

1. Riva S, Polsinelli M, Falaschi A. A new phage of Bacillus subtilis with infectious DNA having separable strands. J Mol Biol 1968; 35(2):347-356.
2. Morelli G, Fisseau C, Behrens B et al. The genome of B. subtilis phage SPP1: The topology of DNA molecules. Mol Gen Genet 1979; 168(2):153-161.
3. Tavares P, Santos MA, Lurz R et al. Identification of a gene in Bacillus subtilis bacteriophage SPP1 determining the amount of packaged DNA. J Mol Biol 1992; 225(1):81-92.

4. Streisinger G, Emrich J, Stahl MM. Chromosome structure in phage T4. III. Terminal redundancy and length determination. Proc Natl Acad Sci 1967; 57:292-295.
5. Deichelbohrer I, Messer W, Trautner TA. Genome of Bacillus subtilis bacteriophage SPP1: Structure and nucleotide sequence of pac, the origin of DNA packaging. J Virol 1982; 41:83-90.
6. Gual A, Camacho AG, Alonso JC. Functional analysis of the terminase large subunit, G2P, of Bacillus subtilis bacteriophage SPP1. J Biol Chem 2000; 275(45):35311-35319.
7. Humphreys GO, Trautner TA. Maturation of bacteriophage SPP1 DNA: Limited precision in the sizing of mature bacteriophage genomes. J Virol 1981; 37(2):832-835.
8. Tavares P, Lurz R, Stiege A et al. Sequential headful packaging and fate of the cleaved DNA ends in bacteriophage SPP1. J Mol Biol 1996; 264(5):954-967.
9. Alonso JC, Lüder G, Stiege AC et al. The complete nucleotide sequence and functional organization of Bacillus subtilis bacteriophage SPP1. Gene 1997; 204(1-2):201-212.
10. Becker B, de la Fuente N, Gassel M et al. Head morphogenesis genes of the Bacillus subtilis bacteriophage SPP1. J Mol Biol 1997; 268(5):822-839.
11. Dröge A, Santos MA, Stiege AC et al. Shape and DNA packaging activity of bacteriophage SPP1 procapsid: Protein components and interactions during assembly. J Mol Biol 2000; 296(1):117-132.
12. Dröge A, Tavares P. In vitro packaging of DNA of the Bacillus subtilis bacteriophage SPP1. J Mol Biol 2000; 296(1):103-115.
13. Thuman-Commike PA, Greene B, Malinski JA et al. Role of the scaffolding protein in P22 procapsid size determination suggested by T=4 and T=7 procapsid structures. Biophys J 1998; 74(1):559-568.
14. Stiege A, Isidro A, Dröge A et al. Specific and stoichiometric targetting of a DNA-binding protein to the SPP1 procapsid by interaction with the portal oligomer. Mol Microbiol 2003; 49:1201-1212.
15. Chai S, Kruft V, Alonso JC. Analysis of the Bacillus subtilis bacteriophages SPP1 and SF6 gene 1 product: A protein involved in the initiation of headful packaging. Virology 1994; 202(2):930-939.
16. Chai S, Lurz R, Alonso JC. 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 1995; 252(4):386-398.
17. Gual A, Alonso JC. Characterization of the small subunit of the terminase enzyme of the Bacillus subtilis bacteriophage SPP1. Virology 1998; 242(2):279-287.
18. Chai S, Bravo A, Lüder G et al. Molecular analysis of the Bacillus subtilis bacteriophage SPP1 region encompassing genes 1 to 6. The products of gene 1 and gene 2 are required for pac cleavage. J Mol Biol 1992; 224(1):87-102.
19. Isidro A, Santos MA, Henriques AO et al. The high-resolution functional map of bacteriophage SPP1 portal protein. Mol Microbiol 2004; 51:949-962.
20. Isidro A, Henriques AO, Tavares P. The portal protein plays essential roles at different steps of the SPP1 DNA packaging process. Virology 2004; 322:253-263.
21. Camacho AG, Gual A, Lurz R et al. Bacillus subtilis bacteriophage SPP1 DNA packaging motor requires terminase and portal proteins. J Biol Chem 2003; 278:23251-23259.
22. Lurz R, Orlova EV, Günther D et al. Structural organisation of the head-to-tail interface of a bacterial virus. J Mol Biol 2001; 310(5):1027-1037.
23. Hendrix RW. Symmetry mismatch and DNA packaging in large bacteriophages. Proc Natl Acad Sci USA 1978; 75(10):4779-4783.
24. Dube P, Tavares P, Lurz R et al. The portal protein of bacteriophage SPP1: A DNA pump with 13-fold symmetry. EMBO J 1993; 12(4):1303-1309.
25. Simpson AA, Tao Y, Leiman PG et al. Structure of the bacteriophage ϕ29 DNA packaging motor. Nature 2000; 408(6813):745-750.
26. Wikoff WR, Liljas L, Duda RL et al. Topologically linked protein rings in the bacteriophage HK97 capsid. Science 2000; 289(5487):2129-2133.
27. Casjens S, Wyckoff E, Hayden M et al. Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. J Mol Biol 1992; 224(4):1055-1074.
28. Orlova EV, Dube P, Beckmann E et al. Structure of the 13-fold symmetric portal protein of bacteriophage SPP1. Nat Struct Biol 1999; 6(9):842-846.
29. Tavares P, Dröge A, Lurz R et al. The SPP1 connection. FEMS Microbiol Rev 1995; 17:47-56.
30. Parkinson J, Huskey R. Deletion mutants of bacteriophage lambda. I. Isolation and initial characterization. J Mol Biol 1971; 56(2):369-384.

31. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21(2):319-331.
32. Smith DE, Tans SJ, Smith SB et al. The bacteriophage∞ϕ29 portal motor can package DNA against a large internal force. Nature 2001; 413(6857):748-752.
33. Kindt J, Tzlil S, Ben-Shaul A et al. DNA packaging and ejection forces in bacteriophage. Proc Natl Acad Sci USA 2001; 98(24):13671-13674.
34. Evilevitch A, Lavelle L, Knobler CM et al. Osmotic pressure inhibition of DNA ejection from phage. Proc Natl Acad Sci USA 2003; 100:9292-9295.
35. Orlova EV, Gowen B, Dröge A et al. Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. EMBO J 2003; 22:1255-1262.
36. Chai S, Szepan U, Alonso JC. Bacillus subtilis bacteriophage SPP1 terminase has a dual activity: It is required for the packaging initiation and represses its own synthesis. Gene 1997; 184(2):251-256.
37. Adams M, Hayden M, Casjens S. On the sequential packaging of bacteriophage P22 DNA. J Virol 1983; 46(2):673-677.
38. Yasbin RE, Young FE. Transduction in Bacillus subtilis by bacteriophage SPP1. J Virol 1974; 14(6):1343-1348.
39. de Lencastre H, Archer LJ. Characterization of bacteriophage SPP1 transducing particles. J Gen Microbiol 1980; 117(2):347-355.
40. Canosi U, Lüder G, Trautner TA. SPP1-mediated plasmid transduction. J Virol 1982; 44(2):431-436.
41. Bravo A, Alonso JC. The generation of concatemeric plasmid DNA in Bacillus subtilis as a consequence of bacteriophage SPP1 infection. Nucleic Acids Res 1990; 18(16):4651-4657.
42. Viret JF, Bravo A, Alonso JC. Recombination-dependent concatemeric plasmid replication. Microbiol Rev 1991; 55(4):675-683.
43. Deichelbohrer I, Alonso JC, Lüder G et al. Plasmid transduction by Bacillus subtilis bacteriophage SPP1: Effects of DNA homology between plasmid and bacteriophage. J Bacteriol 1985; 162(3):1238-1243.
44. Alonso JC, Lüder G, Trautner TA. Requirements for the formation of plasmid-transducing particles of Bacillus subtilis bacteriophage SPP1. EMBO J 1986; 5(13):3723-3728.
45. Bravo A, Alonso JC, Trautner TA. Functional analysis of the Bacillus subtilis bacteriophage SPP1 pac site. Nucleic Acids Res 1990; 18(10):2881-2886.

CHAPTER 7

The ø29 DNA Packaging Motor:

Seeking the Mechanism

Dwight Anderson and Shelley Grimes

Abstract

The *Bacillus subtilis* bacteriophage ø29 research team in Minneapolis has marveled at (and reveled in) the intricacies of ø29 assembly for more than 30 years. Here we highlight the current state of knowledge of ø29 DNA packaging. We describe the in vitro packaging system and focus on recent advances that address the mechanism of the packaging motor. Among advances, the head-tail connector has been visualized in proheads and the packaging motor resolved in partially packaged particles by electron microscopy, the structure of the connector has been solved by X-ray crystallography, and the force-velocity relationship of the motor has been established in single molecule studies. A challenge in the future is to determine the structure and interaction of motor components as well as the conformational changes in these components during energy transduction that define the mechanism of DNA translocation.

The compaction of ø29 DNA by more than 100-fold in length during packaging into the prohead is remarkable in that it overcomes the entropic, electrostatic and bending energies of DNA. The ø29 packaging motor is a multisubunit protein-RNA complex at the prohead portal vertex. The motor, driven by ATP hydrolysis, is force-generating and highly processive, and it opposes a strong internal force that builds up within the capsid as DNA is compressed. The motor can work against loads of 57 picoNewtons on average, making it one of the strongest molecular motors yet reported. We aim to identify and characterize the intermediates during the assembly and function of the motor and to determine the structure of each component of the motor at atomic resolution.

A brief overview of the ø29 DNA packaging system follows, and accordingly citation of published work is highly selective. For more detail, refer to a recent comprehensive review of ø29 DNA packaging.[1] Also, a comparison of all the bacteriophage DNA packaging systems under study has been addressed (Jardine and Anderson, in press).

Components of the ø29 DNA Packaging Motor

The DNA packaging motor consists of the prohead, which contains the dodecameric head-tail connector at its portal vertex, a multimer of prohead RNA (pRNA) that is attached to the connector, and a multimer of the ATPase gp16 (gene product 16) that binds to pRNA (Table 1). The active motor also includes the packaging substrate DNA-gp3.

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

Table 1. Components of the ø29 DNA packaging system

Component	Gene Product (Structure)	Molecular Weight[a]	Copies[b]	Mass (MDa)
Fiberless	gp7 (scaffold)	11,267	~147	~1.7
Prohead	gp8 (major capsid)	49,847	235	11.71
	gp10 (connector)	35,880	12	0.43
	pRNA (prohead RNA)	57,594	5-6	0.29 or 0.35
Packaging	gp3 (DNA terminal protein)	31,051	2	0.06
Proteins	gp16 (packaging ATPase)	38,967	5-6?	0.19 or 0.23
DNA with gp3				12.76 + 0.06 (19,285 bp)

[a] Protein molecular weights were inferred from the nucleotide sequence. [b] For a detailed description of copy number determinations and literature references, see Peterson C, Simon M, Hodges J et al. Composition and mass of the bacteriophage ø29 prohead and virion; J Struct Biol 2001; 135:18-25.

Proheads are produced in infection of the nonpermissive *Bacillus subtilis* host with the amber mutant *sus*16(300)-*sus*14(1241) that is defective for the packaging ATPase gp16 and that provides delayed lysis. The prohead is assembled from the head-tail connector (gp10), the scaffolding protein (gp7), the major capsid protein (gp8) and the head fibers (gp8.5) with the aid of an unidentified host protein.[2] The prohead is prolate, an extended T=3 icosahedron with dimensions of 54 x 42 nm and a shell thickness of only 1.6 nm (ref. 3) (Fig. 1). The shell contains 235 copies of the 49 kiloDalton (kDa) gp8, arranged as 11 pentamers at the vertices (the connector occupies the portal vertex) and 30 hexamers to expand the lattice. The 55 head fibers that project about 250Å from the quasi three-fold axes at the vertices are nonessential in both the prohead and virion. Occasionally most of the scaffolding protein is lost in cell lysis, yet the proheads retain full biological activity for in vitro DNA packaging (unpublished). Proheads can also be produced from products of the cloned genes 7, 8 and 10 in *Escherichia coli*, although the yield is about one-third that of infection in *B. subtilis*;[4] these proheads must be reconstituted with pRNA prior to DNA packaging. Each prohead can be active in DNA packaging.

The head-tail connector, the crux of the DNA packaging motor, is a dodecamer of the 36 kDa gp10. The structure of the connector is solved to 3.2Å resolution[5] and 2.1Å[6,7] (Fig. 2). It is 75Å in height, with three approximately cylindrical regions having external diameters of 138Å, 94Å and 66Å. The wide part is in the prohead, and the narrow part protrudes from the prohead and contains the RNA binding domain (Fig. 1A, B). The central part of the connector is comprised of three long alpha helices in each monomer that are arranged at an angle of about 40° relative to the central 12-fold axis. The outer surface of the connector is hydrophobic, which may facilitate connector rotation within the capsid.[8] The inner surface of the connector channel is highly negatively charged, which may ease DNA passage. Investigation of the role of the lysine residues at positions 200 and 209 that may interact with DNA[6] as well as the residues 229-246 in the wide region of the connector that are disordered in the crystal structure[5] may be fruitful. Each monomer has positively and negatively charged sides that may stabilize the structure. The X-ray structure of the connector fits well into the cryo-EM 3D reconstruction of the prohead (Fig. 3).

Prohead RNA is a 174-base transcript of the far left end of the genome and is essential for DNA packaging[9] (Fig. 4). A 120-base form of pRNA (domain I of Fig. 4), generated by

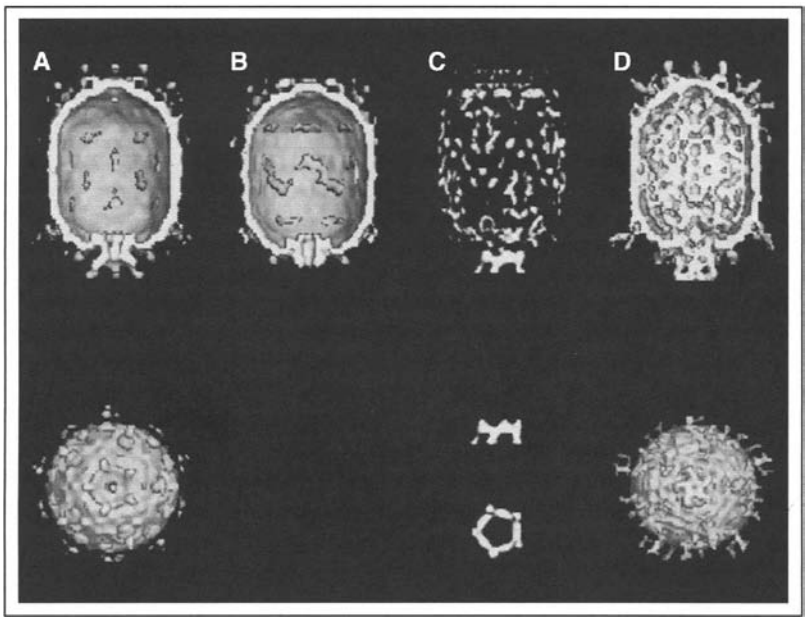


Figure 1. Cryo-EM 3D reconstructions. Top and bottom rows: A) prohead + 120-base pRNA; B) prohead treated with RNase; C) difference map between A and B; and D) partially packaged particle (with DNA in channel). End-on views, looking along the tail towards the head, are given for A and D. Orthogonal difference pRNA densities are shown below the difference map in C. Reprinted with permission from: Simpson AA, Tao Y, Leiman PG et al. Nature 2000; 408:745-750. © 1998 Macmillan Publishers Ltd.

the action of adventitious ribonucleases during prohead isolation, is fully active in DNA packaging and phage assembly.[10] The pRNA can be released from the prohead and reattached, with concomitant loss and restoration of biological activity, facilitating mutational analysis of RNA structure and function.[11-16] The number of copies of pRNA needed for DNA packaging is six by genetic methods,[17,18] and hexamers and dimers are the predominant forms of pRNA in solution by analytical ultracentrifugation.[17] The hexamer is formed by a novel intermolecular base pairing between identical monomers (helix G of Fig. 4), and dimers are intermediates in formation of the hexamer.[17-19] However, cryo-EM 3D reconstruction demonstrates five-fold symmetry of pRNA attached to the prohead,[5,20] or six-fold symmetry using different methods,[21] thus the symmetry of pRNA in the packaging motor is controversial. The electron density due to pRNA is visualized on the connector by cryo-EM as a ring around its protruding end, with a skirt of five dense spokes[5] (Fig. 1C). If the five-fold symmetric pRNA ring contacts the icosahedral capsid as proposed, then the capsid-pRNA-gp16 complex may act as a stator around the dodecameric connector (see model of the mechanism below).

The ATPase gp16 is produced by sucrose induction of *B. subtilis* (pSac16) that contains the cloned gene 16 (unpublished). The 39 kDa protein is the larger of the two ø29 DNA packaging proteins and contains ATP binding consensus sequences and the predicted secondary structure of ATPases.[22] The ATPase activity of gp16 is pRNA-dependent.[23] gp16 has an RNA recognition motif and binds to pRNA on the prohead to complete assembly of the packaging motor. gp16 is needed in multiple copies for DNA packaging in vitro,[24] and at least six copies of purified gp16 per prohead are added to achieve full packaging activity in the defined

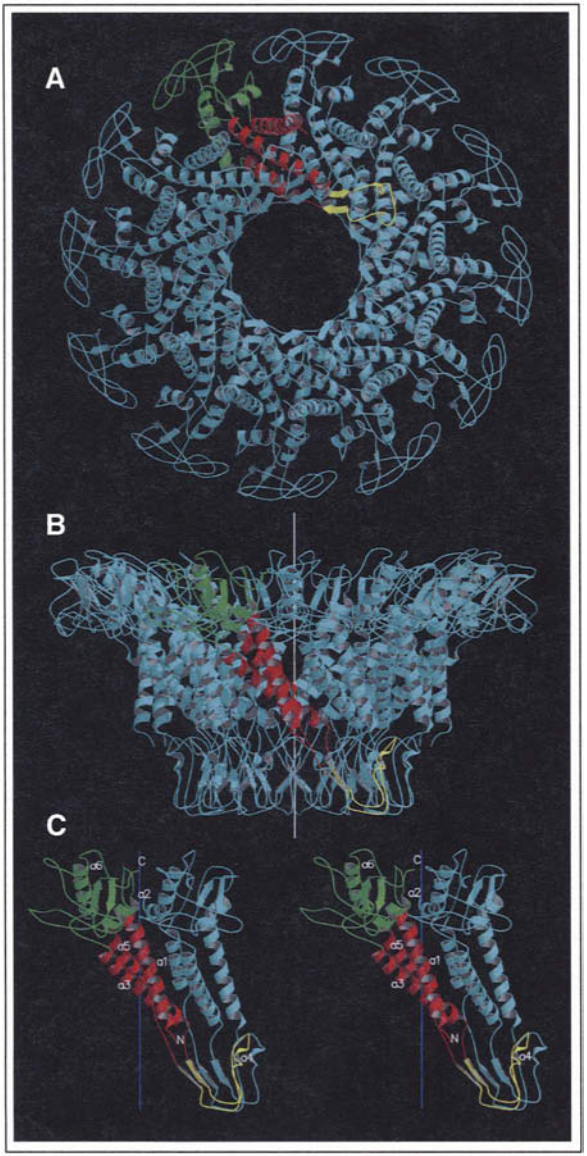


Figure 2. Connector structure ribbon diagrams. A) The dodecameric connector seen from the tail looking towards the head; B) a side view with the pRNA-binding site at the bottom, showing the conical structure of the connector and the helical twist of each subunit around the 12-fold axis (white); C) a stereo diagram of a pair of monomers. One monomer is colored red in the central domain, green in the wide-end domain that resides inside the capsid, and yellow at the narrow-end domain. The other monomer is colored blue. The ordered part of the polypeptide starts with helix α 1 on the outside of the connector, going towards the wide end (residues 61 to 128 and 247 to 286). Helix α 3 (residues 129 to 157) returns the chain to the narrow end (residues 158 to 202). The tip of the connector at the narrow end is formed by residues164 to 170 and 185 to 196. Helix α 5 (residues 208 to 226) returns the polypeptide to the wide end through the second disordered section. (Drawn with the program MOLSCRIPT and RASTER3D). Reprinted with permission from: Simpson AA, Tao Y, Leiman PG et al. Nature 2000; 408:745-750.

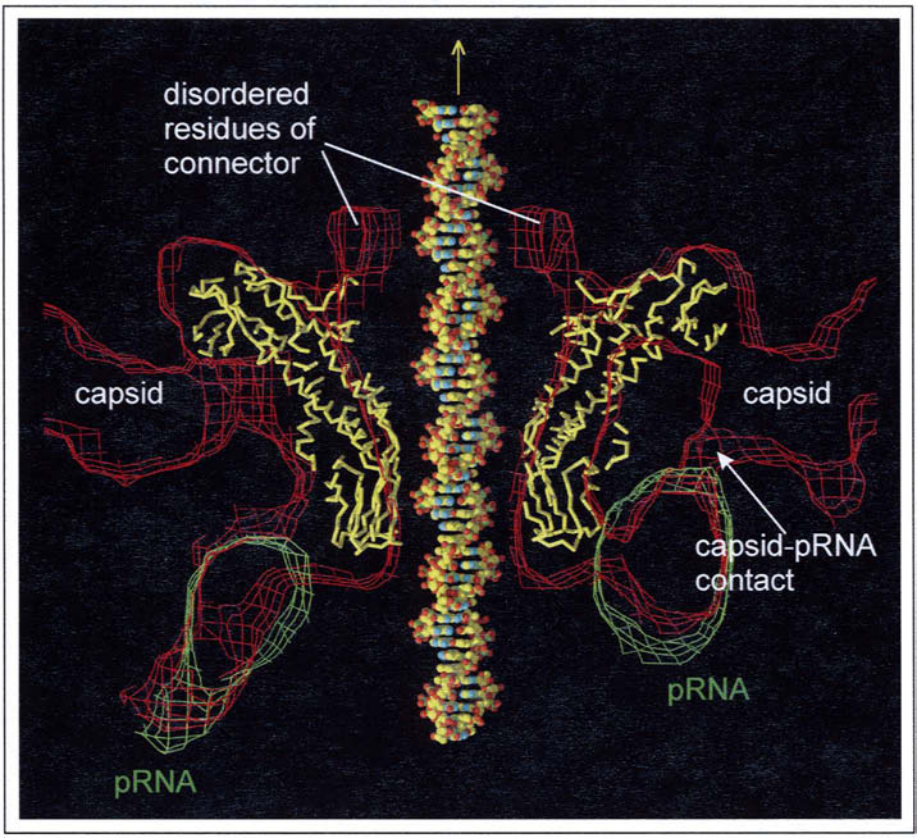


Figure 3. Prohead cryo-EM density fitted with atomic structure. Cross-section of the cryo-EM prohead density (red) fitted with the Cα backbone of the connector (yellow) and the cryo-EM pRNA difference map (green). Shown also is a DNA molecule placed through the central channel of the connector. The prohead capsid, one of the five contacts between the pRNA with the capsid, and the partially disordered residues 229 to 246 and 287 to 309 in the connector are indicated. (Drawn with the programs XTALVIEW and RASTER3D). Reprinted with permission from: Simpson AA, Tao Y, Leiman PG et al. Nature 2000; 408:745-750. ©1998 Macmillan Publishers Ltd.

system. However, the number of copies of gp16 in the motor is unknown. gp16 is purportedly visualized by cryo-EM attached to the pRNA spokes on the partially packaged particle[5] (Fig. 1D). pRNA and gp16 exit the filled head after packaging and are not components of the virion.[9,25]

The ø29 genome is 19.3 kilobasepairs (kbp) and has a single subunit of gp3 covalently attached at each 5' terminus (DNA-gp3). The 31 kDa gp3 is essential for both DNA replication and DNA packaging. gp3 primes DNA replication, which proceeds by a strand-displacement mechanism, producing unit-length genomes for packaging (for a review see ref. 26). gp3 is the smaller of the two packaging proteins commonly found in phage systems. It also confers structure on the DNA, as the gp3 ends of DNA-gp3 interact to yield circles, and gp3 loops back on the DNA to form lariats (see packaging events below).[27] DNA-gp3 is produced in restrictive infection with the mutant *sus*4(369)-*sus*8(22).

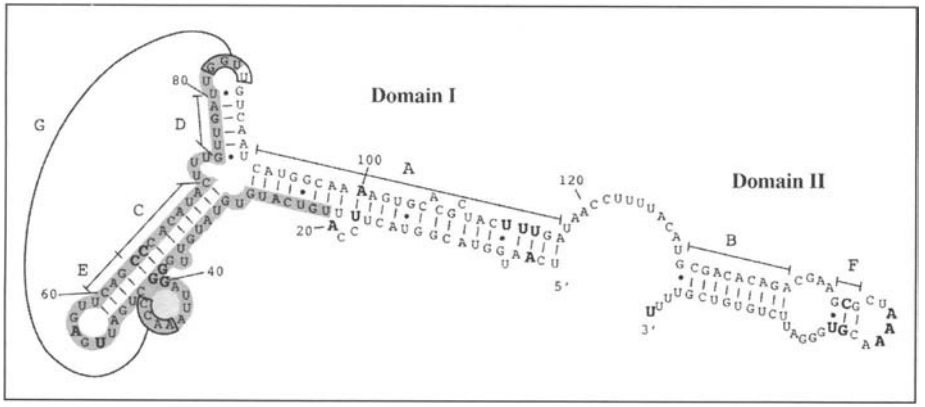


Figure 4. Secondary structure of the 174-base pRNA. The bold line shows the proposed tertiary interaction. Helices are designated A, B, C, D, E, F and G. The prohead binding domain is marked by shading. Reprinted with permission from: Grimes S, Jardine PJ, Anderson D. Adv Virus Res 2002; 58:255-294. © Harcourt Press, Inc.

The DNA Packaging Assay and Provisional Packaging Events

The packaging assay is based on protection of DNA in filled heads from DNase treatment.[28] Proheads, DNA-gp3 and gp16 are generally mixed in a ratio of 2:1:12 to provide an estimated two-fold excess of proheads and gp16, and the ternary complex forms during a 5 minute incubation at room temperature. ATP is added to initiate packaging, and incubation is continued for 10 minutes. DNase I is added to digest unpackaged DNA, and the protected packaged DNA is extracted and quantified by agarose gel electrophoresis (Fig. 5). Packaging efficiency approaches 100% (lanes 1 and 2). Packaging is oriented, as left-end restriction fragments are packaged preferentially[29] (lanes 4 and 5).

Provisional events of ø29 DNA packaging are illustrated in Figure 6. We begin with a description of the packaging substrate DNA-gp3, which has a maturation pathway for packaging that is independent of prohead assembly.[27] Supercoiling of the packaging substrate occurs by gp3-dependent (but sequence independent) formation of a lariat, binding of multiple copies of gp16 to the lariat loop junction, and supercoiling of the lariat loop by concerted gp3-gp16 action (Fig. 6A). The prohead then binds to the lariat loop of the supercoiled DNA-gp3-gp16 (not to the tail of the lariat), as observed by electron microscopy (Fig. 6B; unpublished). The supercoiled DNA-gp3-gp16 is shown wrapped around the connector, a projection of our finding that the prohead connector wraps supercoiled plasmid DNA, reducing the DNA contour length (unpublished). Thus the connector of the prohead behaves as free connectors, which wrap supercoiled plasmid DNA around the outside to reduce DNA contour length and restrain negative supercoils.[30] Linear plasmid DNA is not wrapped by free connectors or connectors in proheads. Efficient prohead connector-DNA binding may be due to prohead recognition of the helical pitch of the supercoiled DNA or affinity for DNA crossovers. Supercoiled DNA-gp3 is preferentially packaged over linear DNA-gp3.[27] Although both the left and right termini of DNA-gp3 are supercoiled in the presence of gp16, the prohead recognizes and packages the left end preferentially; gp3 may be posttranslationally modified, or the pRNA may hybridize transiently to its coding sequence near the left end.

Once DNA is wrapped around the connector, the prohead may slide on the DNA to the lariat loop junction, to position the DNA left end at the connector channel (Fig. 6C). Proheads are observed at the lariat loop junction by electron microscopy (unpublished). Then gp16 is

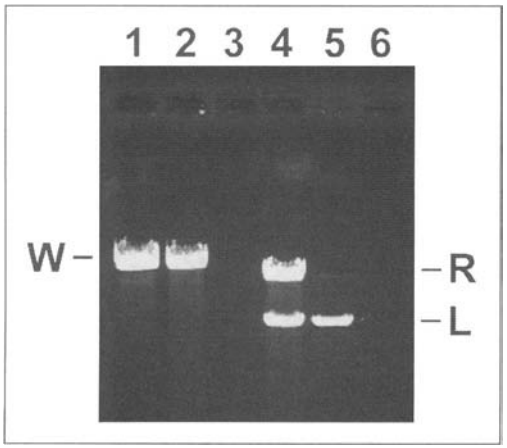


Figure 5. DNA packaging in the defined system. DNA-gp3, gp16, proheads and ATP are mixed and incubated for packaging. DNase I is then added to digest the unpackaged DNA, the DNase is inhibited, and the packaged DNA is extracted and quantified by agarose gel electrophoresis. Input DNA is the standard for quantification of packaging efficiency. ATP is omitted from the negative control lanes and shows that unpackaged DNA is sensitive to DNase I. The lanes contained: 1) whole-length DNA input; 2) packaged whole-length DNA; 3) negative control; 4) *Cla I* DNA input; 5) packaged *Cla I* DNA; 6) negative control. "W" designates whole-length; "R" and "L" the right and left ends. Reprinted with permission from: Grimes S, Jardine PJ, Anderson D. Adv Virus Res 2002; 58:255-294. © Harcourt Press, Inc.

hypothesized to bind to pRNA to complete the assembly of the packaging motor, and concurrently the left end of the DNA is freed and positioned for packaging (Fig. 6C and D; the pRNA and gp16 are not shown). It is not known if the copies of gp16 bound to DNA are sufficient for DNA packaging. However, it is clear that gp16 is needed in multiple copies,[24] and at least six copies of purified gp16 per prohead are added to achieve full packaging in vitro (unpublished). Hydrolysis of ATP by gp16-pRNA energizes sequential connector conformational change that results in DNA translocation (Fig. 6E), as described below in the model of the motor mechanism.

Several aspects of the ø29 packaging system and model are currently unique to ø29: RNA is an essential component of the packaging motor, the packaging substrate is supercoiled by the packaging proteins, and the supercoiled DNA is wrapped around free connectors and connectors in the prohead. We hope to determine if other phage DNA packaging systems share these features.

A Model of the Packaging Mechanism

The most striking structural characteristic of the connector is the 36 long alpha helices in the central region, inclined roughly obliquely to the DNA helix, that give the connector the appearance of a compressible spring. The idea that the connector can expand and contract is the basis of a ratchet mechanism model[5] (Fig. 7). There are three concentric symmetries of the active motor: the 10-fold symmetry of DNA in the connector channel, the 12-fold symmetry of the connector, and the five-fold symmetry of the capsid-pRNA-gp16. To start, the active ATPase I contacts connector monomer A, which in turn interacts with the DNA (Fig. 7A). The connector extends down the DNA by 2 bp, and concurrently the connector narrow end rotates counterclockwise by 12°; this aligns ATPase II with the connector monomer C and in turn monomer C with the DNA backbone (Fig. 7B). Then the connector contracts by 6.8Å to

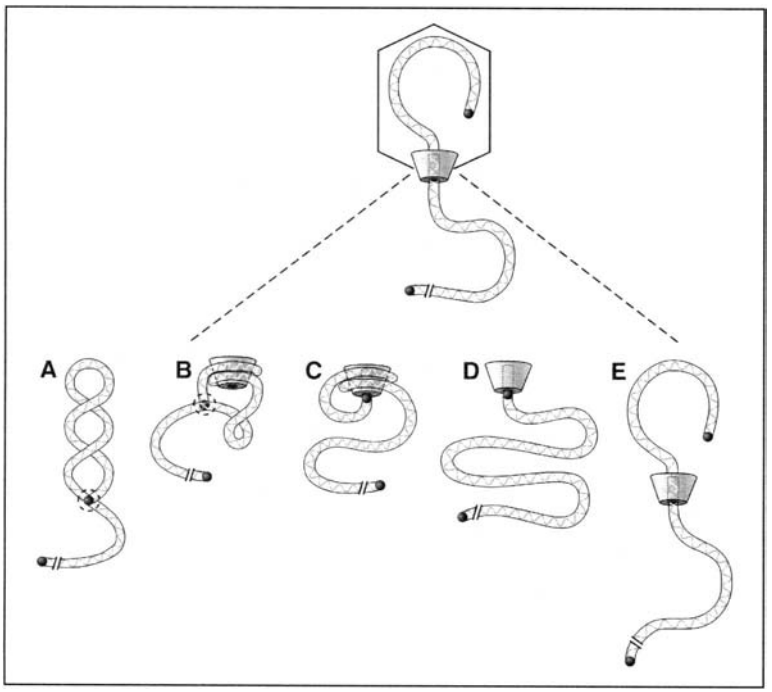


Figure 6. ø29 DNA-gp3 packaging model. A) gp3 covalently bound at the left end (gray dot) loops and forms a lariat by interaction with DNA, gp16 binds at the lariat loop junction (dotted circle), and gp3-gp16 produce a supercoiled packaging substrate (Grimes and Anderson, 1997); B) Supercoiled DNA wraps around the connector of the prohead; C) The left DNA end is freed and binds the connector; D) DNA is released from the outside of the connector; and E) DNA is packaged as the connector expands, contracts and rotates with the aid of ATP hydrolysis (see Fig. 7). The pRNA and gp16 components of the motor are not shown. Reprinted with permission from: Grimes S, Jardine PJ, Anderson D. Adv Virus Res 2002; 58:255-294. © Harcourt Press, Inc.

advance 2 bp of DNA into the prohead, and the wide end of the connector concurrently and passively rotates counterclockwise by 12° to return to the position for activation of ATPase II in the next cycle (Fig. 7C). Thus the rotation of the connector does not actively drive DNA translation, as a rotating nut advances a bolt,[31] but rather the rotation acts to index ordered firing of ATPases and thus coordinate sequential conformational changes in the connector to translocate DNA. Atomic force microscopy shows that the connector can be compressed reversibly by about one-third of its height,[32] lending support to our model.

The ø29 packaging motor, in translocating DNA in one direction through the connector channel, is different in kind from the other molecular motors under study. We have proposed that the ø29 motor is a ratchet mechanism that depends on passive rotation of the connector for sequential ATP-energized protein conformational change.[5] The F_1-ATPase[33] and the bacterial flagellar[34] rotary motors use sequential rotary conformational change to turn a spindle and are normally driven by a proton gradient. Although the F_1 motor uses ATP hydrolysis to rotate in one direction, the F_0 motor attached to the common shaft uses an electrochemical gradient to turn in the opposite direction. The linear motors of the myosin group that drive muscle contraction and move cargo along actin filaments, and motors of the kinesin group that move along microtubules are track motors that use conformational fluctuations to effect translation

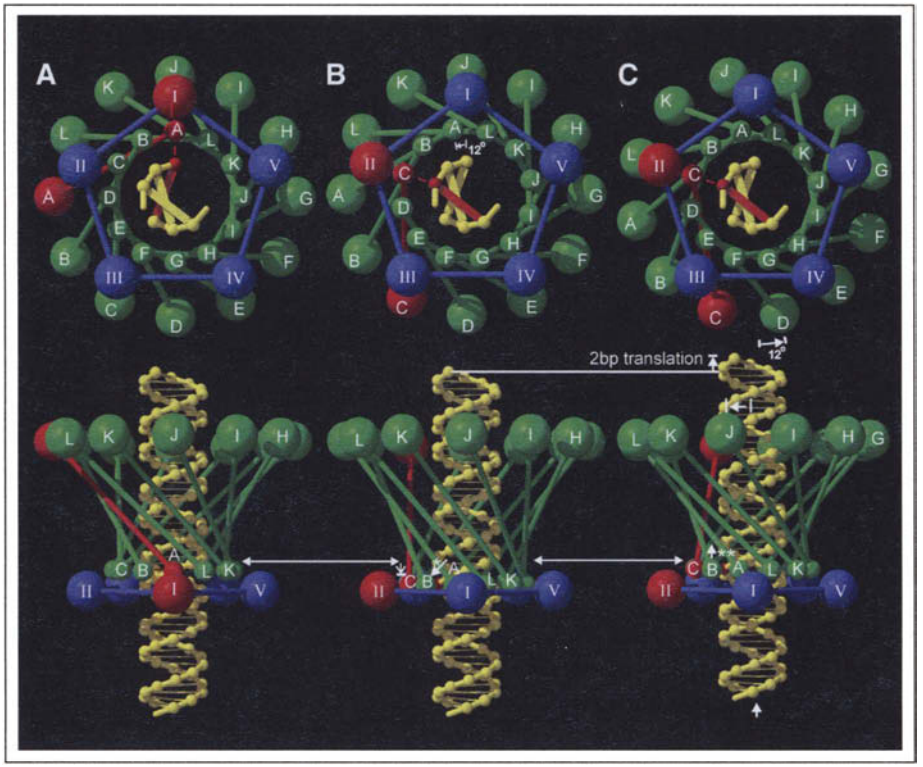


Figure 7. The DNA packaging mechanism. Shown is one cycle in the mechanism that rotates the connector and translates the DNA into the head. The view down the connector axis (top) is towards the head (as in (Fig. 2A), whereas the bottom row shows side views corresponding to that seen in Figure 2B. Eleven of the 12 subunits (A, B,...., L) of the connector are shown in green; the "active" monomer is shown in red. The connector is represented as a set of small spheres at the narrow end and a set of larger spheres at the wide end connected by a line representing the central helical region. The pRNA-ATPase complexes (I, II, III, IV, V), surrounding the narrow end, are shown by a set of four blue spheres and one red sphere. The DNA base aligned with the active connector monomer is also shown in red. In A) the active pRNA-ATPase I interacts with the adjacent connector monomer (A), which in turn contacts the aligned DNA base. In B) the narrow end of the connector has moved counterclockwise by 12° to place the narrow end of monomer C opposite ATPase II, the next ATPase to be fired, causing the connector to expand lengthwise by slightly changing the angle of the helices in the central domain (white arrow with asterisk). In C) the wide end of the connector has followed the narrow end, while the connector relaxes and contracts (white arrow with two asterisks), thus causing the DNA to be translated into the phage head. For the next cycle, ATPase II is activated, causing the connector to be rotated another 12°, and so forth. (Drawn with the program RASTER3D). Reprinted with permission from: Simpson AA, Tao Y, Leiman PG et al. Nature 2000; 408:745-750. © 1998 Macmillan Publishers Ltd.

in steps (for a review see ref. 35). In the putative ø29 ratchet mechanism, the spindle or bar (DNA) is translocated, and the connector serves as the pawl.

The ø29 Packaging Motor Is Processive and Powerful

In single molecule studies, optical tweezers pull on ø29 DNA as it is packaged and demonstrate that the packaging motor is force-generating.[36] Tethering a single packaging complex between two beads allows measurement of the packaging rates in real time and the forces

during packaging. DNA packaging is initiated and then stalled after about 15% packaging by the addition of the ATP analogue, γS-ATP. The partially packaged complexes are attached to a streptavidin-coated polystyrene bead via the biotinylated unpackaged end of the DNA and captured in the optical trap. This bead is brought into contact with a second bead, coated with anti-head antibodies, that is held by a pipette to form a stable tether (Fig. 8A, left). In the "constant force feedback" mode, the bead distance is adjusted to maintain a constant DNA tension of 5 picoNewtons (pN) (Fig. 8A, middle). Only when ATP is restored do the beads move closer together, as DNA is packaged (Fig. 8B). Although packaging is highly processive, in that packaging of 5 μm lengths of DNA are observed, pauses and slips occur. On average there are about three pauses per μm of DNA packaged, and pausing frequency increases with higher capsid fillings. During a pause the motor remains engaged with the DNA, despite the 5 pN of applied external force (Fig. 8B inset). During slips, an abrupt increase in tether length occurs, suggesting that the motor loses grip on the DNA, but then the packaging motor recovers and resumes packaging.

Figure 8C shows a plot of the packaging rate of a ø29-lambda hybrid DNA versus the percent of the genome packaged. Interestingly, the packaging rate is not constant. The packaging rate drops dramatically when about 50% of the DNA is packaged, from an initial rate of about 100 bp/second to zero when about 105% of the ø29 genome length is packaged. This rate decrease that is specific to the last half of packaging likely results from the build-up of internal force due to DNA confinement that opposes and slows the motor.

To investigate the build-up of internal force, the "no feedback" mode, where the positions of the trap and pipette are fixed (Fig. 8A, right), is used with a DNA substrate about 0.8x the length of ø29 DNA to obtain Force-Velocity (*F-V*) traces. As the DNA is reeled into the prohead during packaging, the bead is displaced in the laser trap and tension in the DNA molecule increases. Eventually the force reaches a high enough level to stall the motor. Shown are mean *F-V* curves at two points, when one-third and when two-thirds of the genome is packaged (Fig. 9). At one-third filling no internal force opposes the motor, as there is no velocity decrease, while at two-thirds filling there is an internal force since it takes about 14 pN less external force to stall the motor. Shifting of the two-thirds curve over 14 pN (in the direction of the arrows) shows that there is good overlap with the one-third curve. Thus the internal and external forces add together to act on the motor, and the inherent *F-V* relationship of the motor in the absence of any internal forces is obtained. The packaging rate decreases even for small forces, and it is estimated that a rate-limiting step produces a mechanical displacement of only 1Å, much less than the estimated movement of 2 bp per ATP hydrolyzed in bulk experiments.[22] When the packaging velocity decreases sharply at about 45 pN, a second step in the cycle of the motor associated with a larger movement becomes rate-limiting.[36]

The motor is one of the strongest molecular motors yet reported, working against loads of up to about 70 pN. The average stall force of 57 pN times the reported 0.68 nm (2 bp) moved per ATP gives a work done of 39 pN•nm, representing an energy conversion efficiency of about 30% of the estimated 120 pN•nm associated with ATP hydrolysis. The step size of the motor is less than 5 bp if the maximum force is 70 pN (120 pN•nm/70 pN = 1.7 nm).

By combining rate dependence on external force (Fig. 9) and on the fraction of DNA packaged (Fig. 8C), the build-up of internal force is calculated to reach about 50 pN when the whole genome is packaged. Dividing this force by the hexagonal cell surface area[37] gives an estimate of a pressure of about 6 megaPascals inside the phage head, assuming there is no significant dissipation of energy. This force builds only after about one-half of the DNA is packaged and may be due in part to rearrangements of DNA as it is progressively compressed. Total work done in packaging is estimated at 2×10^4 kT by integrating the curve of internal force (pN) versus percentage of genome packaged (not shown). If dissipation is not great, internal pressure may supply a part of the driving force for DNA injection into the host cell.

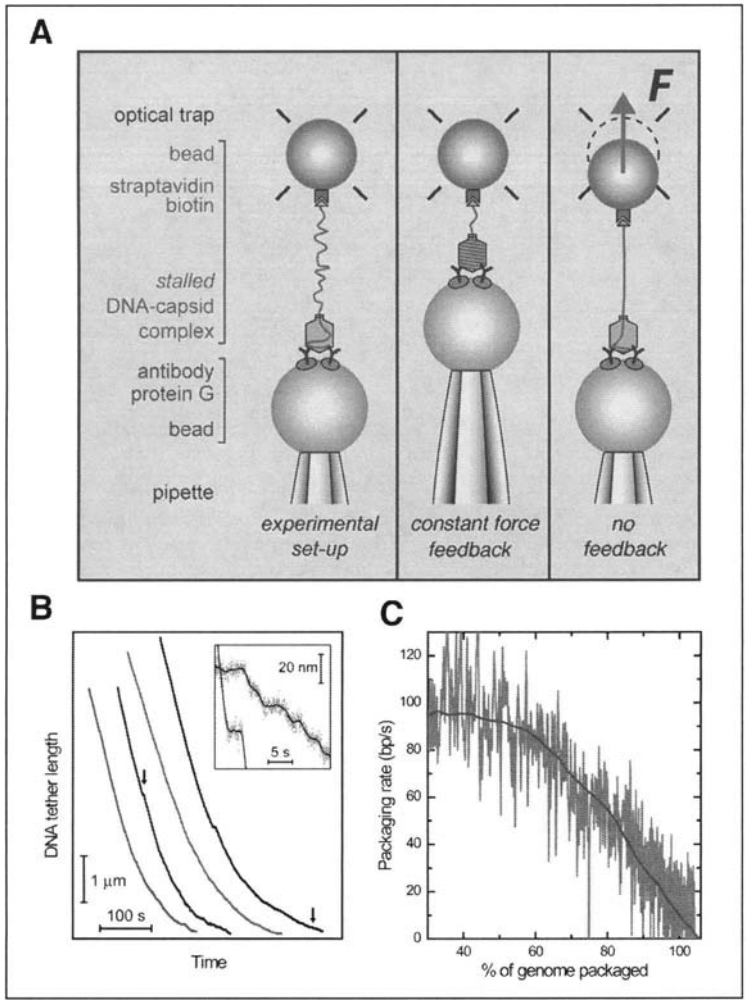


Figure 8. Single molecule DNA packaging in the laser tweezers. A) Diagrams showing: experimental set-up at the start of a measurement (left), constant force feedback mode (middle) and no feedback mode (right). A single ø29 packaging complex is tethered between two microspheres. Optical tweezers are used to trap one microsphere and measure the forces acting on it, while the other microsphere is held by a micropipette. To ensure measurement on only one complex, the density of complexes on the microsphere is adjusted so that only about one out of 5 to10 microspheres yields hookups. Such attachments break in one discrete step as the force is increased, indicating that only one DNA molecule carries the load. B) Plots of DNA tether length against time for four different complexes with a constant force of ~5 pN and using a 34.4 kb ø29-lambda DNA construct. Inset: increased detail of regions, indicated by arrows, showing pauses (curves have been shifted). The solid lines are a 100 point average of the raw data. C) Packaging rate vs. the amount of DNA packaged, relative to the original 19.3 kbp ø29 genome. Trace for a single complex is derived from rightmost trace in panel B). Rates were obtained by linear fitting in a 1.5 s sliding window. The thick center line is an average of eight such measurements. Large pauses (velocity drops >30 bp/s below local average) were removed, and the curves were horizontally shifted to account for differences in microsphere attachment points. The center line was smoothed using a 200 nm sliding window. The standard deviation for the ensemble of measurements varies from ~20 bp/second at the beginning down to ~10 bp/s at the end. Reprinted with permission from: Smith DE, Tans SJ, Smith SJ et al. Nature 2001; 413:748-752. © 1998 Macmillan Publishers Ltd.

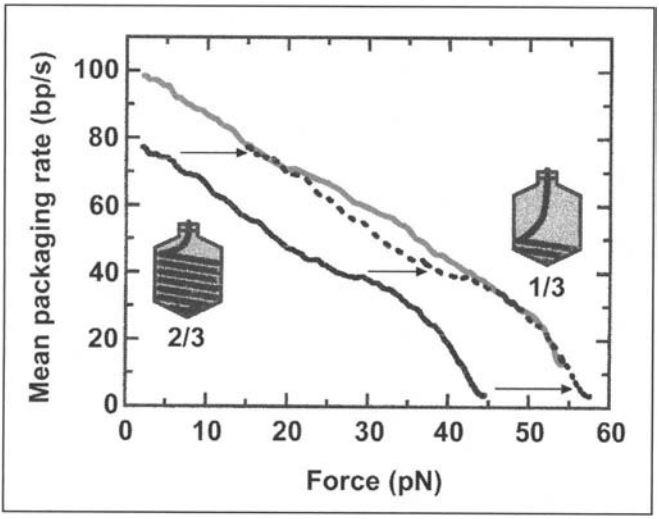


Figure 9. Force-Velocity (*F-V*) analysis. Plots of external force against velocity when about one-third (rightmost line) and about two-thirds (leftmost line) of the genome is packaged. Curves were obtained by averaging 14 and 8 individual traces, respectively. If, in the case of two-thirds of the genome being packaged, an additional 14 pN of internal force were acting on the motor (see text), the dashed line would show the expected behaviour. The rightmost line and the dashed lines would then represent the inherent (total) *F-V* curve for the motor. Reprinted with permission from: Smith DE, Tans SJ, Smith SJ et al. Nature 2001; 413:748-752. ©1998 Macmillan Publishers Ltd.

Ongoing Analysis of ø29 DNA Packaging: The Path to Enlightenment

The paradigm of analysis of any cell biological process at the molecular level, attributed to Tom Pollard and published in a review of the meeting *Proteins as Machines* in reference 38, is reproduced in Figure 10. The approach is to identify all components, describe all reaction intermediates, determine the rates of all transitions, and determine component structures at atomic resolution. Although arduous, all parts of the approach are vital to a complete understanding of the biological process. Multisubunit protein or protein-nucleic acid complexes dominate cell biochemical processes, for example DNA replication, targeting and transport of proteins, pre-mRNA splicing, or signal transduction. In machines of diverse cellular processes, the energy of nucleotide hydrolysis is coupled to conformational changes in protein subunits. The ø29 DNA packaging motor is such a multisubunit complex, with its connector, pRNA and gp16 components attached to the prohead and DNA-gp3 in its channel. We hope to determine the mechanism of the ø29 packaging motor by use of the approach outlined in Figure 10.

The inventory of components of the ø29 packaging motor is complete, but the order of interactions in motor assembly and the identity of packaging intermediates are not known with certainty. Considering motor assembly, it is likely that gp16 of the supercoiled DNA-gp3-gp16 complex binds to pRNA on the prohead to complete the assembly of the motor. The number of copies of gp16 and its structure on DNA-gp3, as well as the mechanism and extent of supercoiling of the DNA substrate, are to be determined. The structure of the packaging initiation complex, i.e., the prohead with DNA-gp3 wrapped around the connector in the absence of ATP, is to be determined by high resolution cryo-EM 3D reconstruction. The steps by which the supercoiled DNA is resolved to a linear form for packaging are unknown. Above all, the precise location and structure of pRNA and gp16 on the partially packaged

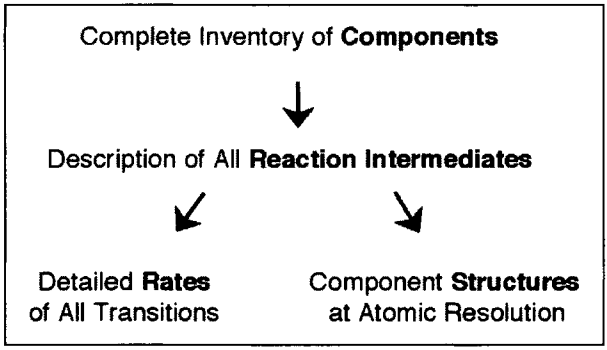


Figure 10. The path to enlightenment in dissection of complex biological processes. Reprinted with permssion from: Alberts B, Miake-Lye R. Cell 1992; 68:415-420. © 1999-2001 Elsevier Science.

particle (packaging intermediate) are to be determined by high resolution cryo-EM 3D reconstruction. As a complement to electron microscopy, the approach of obtaining atomic resolution structures of ø29 packaging motor components and substructures by X-ray crystallography is ongoing, and when possible these high resolution structures will be fitted into cryo-EM 3D reconstructions (Fig. 3).

Dynamic motor function is being investigated with a battery of methods to explore macromolecular interactions, reaction intermediates, and protein movement. Connector rotation is being investigated by polarization of single fluorophores. Packaging intermediates in ATP binding and hydrolysis are being isolated in an attempt to characterize the steps in chemo-mechanical energy transduction and to correlate these steps with protein conformational change. Electron paramagnetic resonance (EPR) and fluorescence resonant energy transfer (FRET) are being employed to detect and characterize conformational change in the connector, pRNA and gp16. Single molecule packaging in the laser tweezers is expected to reveal properties of the packaging motor that are inaccessible using packaging in bulk, such as precise measurement of the rate (and variations in rate) of packaging, the forces exerted on the DNA, effects of perturbation of local DNA structure, details of the catalytic cycle, and the step size of the motor.

Acknowledgements

We thank Charlene Peterson for preparation of the manuscript and Paul Jardine for comments. The work was supported by NIH grants to DA.

References

1. Grimes S, Jardine PJ, Anderson D. Bacteriophage ø29 DNA packaging. Adv Virus Res 2002; 58:255-294.
2. Rajagopal BS, Reilly BE, Anderson DL. Bacillus subtilis mutants defective in bacteriophage ø29 head assembly. J Bacteriol 1993; 175:2357-2362.
3. Tao Y, Olson NH, Xu W et al. Assembly of a tailed bacterial virus and its genome release studied in three dimensions. Cell 1998; 95:431-437.
4. Guo P, Erickson S, Xu W et al. Regulation of the phage ø29 prohead shape and size by the portal vertex. Virology 1991; 183:366-373.
5. Simpson AA, Tao Y, Leiman PG et al. Structure of the bacteriophage ø29 DNA packaging motor. Nature 2000; 408:745-750.
6. Guasch A, Pous J, Ibarra B et al. Detailed architecture of a DNA translocating machine: The high-resolution structure of the bacteriophage ø29 connector particle. J Mol Biol 2002; 315:663-676.

7. Guasch A, Pous J, Ibarra B et al. (Note to ref. 6) Detailed architecture of a DNA translocating machine: The high-resolution structure of the bacteriophage ø29 connector particle. J Mol Biol 2002; 321:379-380.
8. Simpson AA, Leiman PG, Tao Y et al. Structure determination of the head-tail connector of bacteriophage ø29. Acta Cryst 2001; D57:1260-1269.
9. Guo P, Erickson S, Anderson DL. A small viral RNA is required for in vitro packaging of bacteriophage ø29 DNA. Science 1987; 236:690-694.
10. Wichitwechkarn J, Bailey S, Bodley JW et al. Prohead RNA of bacteriophage ø29: Size, stoichiometry, and biological activity. Nucl Acids Res 1989; 17:3459-3468.
11. Guo P, Bailey S, Bodley JW et al. Characterization of the small RNA of the bacteriophage ø29 DNA packaging machine. Nucl Acids Res 1987; 15:7081-7090.
12. Wichitwechkarn J, Johnson D, Anderson D. Mutant prohead RNAs in the in vitro packaging of bacteriophage ø29 DNA-gp3. J Mol Biol 1992; 223:991-998.
13. Reid RJD, Bodley JW, Anderson D. Identification of bacteriophage ø29 prohead RNA domains necessary for in vitro DNA-gp3 packaging. J Biol Chem 1994; 269:9084-9089.
14. Reid RJD, Zhang F, Benson S et al. Probing the structure of bacteriophage ø29 prohead RNA with specific mutations. J Biol Chem 1994; 269:18656-18661.
15. Zhang C, Lee CS, Guo P. The proximate 5' and 3' ends of the 120-base viral RNA (pRNA) are crucial for the packaging of bacteriophage ø29. Virol 1994; 201:77-85.
16. Zhang C, Tellinghuisen T, Guo P. Use of circular permutation to assess six bulges and four loops of DNA-packaging pRNA of bacteriophage ø29. RNA 1997; 3:315-323.
17. Zhang F, Lemieux S, Wu X et al. Function of hexameric RNA in packaging of bacteriophage ø29 DNA in vitro. Mol Cell 1998; 2:141-147.
18. Guo P, Zhang C, Chen C et al. Inter-RNA interaction of phage ø29 pRNA to form a hexameric complex for viral DNA transportation. Mol Cell 1998; 2:149-155.
19. Chen C, Sheng S, Shao Z et al. A dimer as a building block in assembling RNA. J Biol Chem 2000; 275:17510-17516.
20. Morais MC, Tao Y, Olson NH et al. Cryo-EM image reconstruction of symmetry mismatches in bacteriophage ø29. J Struct Biol 2001; 135:38-46.
21. Ibarra B, Caston JR, Llorca O et al. Topology of the components of the DNA packaging machinery in the phage ø29 prohead. J Mol Biol 2000; 298:807-815.
22. Guo P, Peterson C, Anderson D. Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage ø29. J Mol Biol 1987; 197:229-236.
23. Grimes S, Anderson D. RNA dependence of the bacteriophage ø29 DNA packaging ATPase. J Mol Biol 1990; 215:559-566.
24. Bjornsti MA, Reilly BE, Anderson DL. Bacteriophage ø29 proteins required for in vitro DNA-gp3 packaging. J Virol 1984; 50:766-772.
25. Chen C, Guo P. Sequential action of six virus-encoded DNA-packaging RNAs during phage ø29 genomic DNA translocation. J Virol 1997; 71:3864-3871.
26. Meijer WJJ, Horcajadas JA, Salas M. ø29 Family of Phages. Microbiol and Mol Biol Rev 2001; 65:261-287.
27. Grimes S, Anderson D. The bacteriophage ø29 packaging proteins supercoil the DNA ends. J Mol Biol 1997; 266:901-914.
28. Grimes S, Anderson D. In vitro packaging of bacteriophage ø29 DNA restriction fragments and the role of the terminal protein gp3. J Mol Biol 1989; 209:91-100.
29. Bjornsti MA, Reilly BE, Anderson DL. Morphogenesis of bacteriophage ø29 of Bacillus subtilis: Oriented and quantized in vitro packaging of DNA-gp3. J Virol 1983; 45:383-396.
30. Turnquist S, Simon M, Egelman E et al. Supercoiled DNA wraps around the bacteriophage ø29 head-tail connector. Proc Natl Acad Sci USA 1992; 89:10479-10483.
31. Hendrix RW. Symmetry mismatch and DNA packaging in large bacteriophages. Proc Natl Acad Sci USA 1978; 75:4779-4783.
32. Muller DJ, Engel A, Carrascosa JL et al. The bacteriophage ø29 head-tail connector imaged at high resolution with the atomic force microscope in buffer solution. EMBO J 1997; 16:2547-2553.
33. Kinosita Jr K, Yasuda R, Noji H et al. F_1-ATPase: A rotary motor made of a single molecule. Cell 1998; 93:21-24.

34. Silverman MR, Simon MI. Flagellar rotation and the mechanism of bacterial motility. Nature 1974; 249:73-74.
35. Vale RD, Milligan RA. The way things move: Looking under the hood of molecular motor proteins. Science 2000; 288:88-95.
36. Smith DE, Tans SJ, Smith SJ et al. The bacteriophage ø29 portal motor can package DNA against a large internal force. Nature 2001; 413:748-752.
37. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell 1980; 21:319-331.
38. Alberts B, Miaki-Lye R. Unscrambling the puzzle of biological machines: The importance of the details. Cell 1992; 68:415-420.

CHAPTER 8

Encapsidation of the Segmented Double-Stranded RNA Genome of Bacteriophage ϕ6

Minna M. Poranen, Markus J. Pirttimaa and Dennis H. Bamford

Abstract

Bacteriophage ϕ6 has a segmented double-stranded RNA genome that is incorporated into a preformed capsid during viral assembly. The three viral genomic segments are packaged as single-stranded precursors, which are later replicated into the mature double-stranded genome inside the capsid by the viral polymerase. The packaging efficiency of ϕ6 is high; virtually all particles released from ϕ6-infected cells are infectious and carry one copy of each of the genome segments. This feature makes ϕ6 an appropriate model to analyze the principles of packaging of a multi-segmented genome. In vitro analyses have revealed that the packaging of ϕ6 involves sequential uptake of the three genome segments. The initiation of RNA replication is a checkpoint for correct genome packaging and is dependent on a specific interaction between the viral RNA and the preassembled polymerase complex, the procapsid. During the maturation the polymerase complex undergoes structural changes which lead to the expansion of the particle. The packaging NTPase of ϕ6 is a ring-like hexamer that is located at each of the five-fold vertices in the procapsid. However, it appears that only one vertex is adequate to carry out efficient packaging, while the others are required when the newly synthesized message RNAs exits from the capsid. ϕ6 packaging and replication shares features with bacterial double-stranded DNA viruses as well as with eukaryotic double-stranded RNA viruses.

Introduction

Double-Stranded RNA Virus ϕ6

ϕ6 is an enveloped double-stranded (ds)RNA virus of a gram-negative bacterium *Pseudomonas syringae*[1] and the type-organism of the family *Cystoviridae* (http://www.ncbi.nlm.nih.gov/ICTV/). For a long time ϕ6 was the only isolate of its kind. However, recently several new phages with a segmented dsRNA genome have been isolated.[2] Some of these are close relatives of ϕ6 with high sequence identity, while others are more distantly related and have retained only the general genomic organization and the overall virion structure common in the *Cystoviridae* family.

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

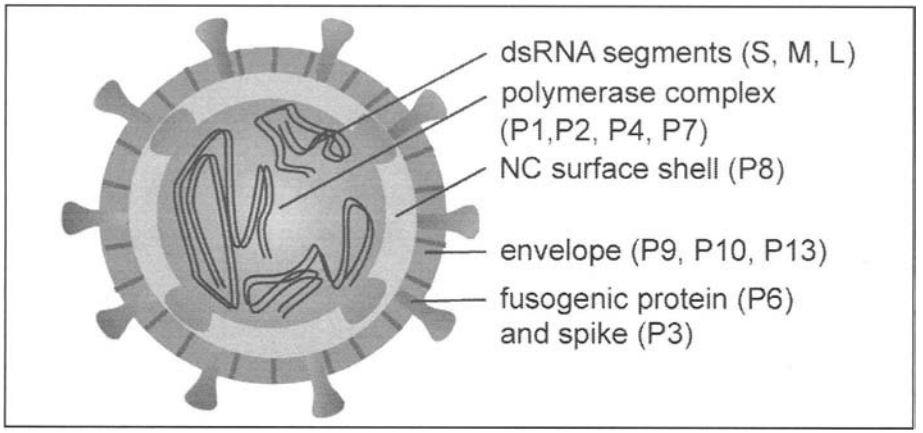


Figure 1. Schematic presentation of the ϕ6 virion. The virion is composed of three concentric layers. Four proteins, P1, P2, P4 and P7, comprise the polymerase complex, which is covered by lattice of protein P8 forming the nucleocapsid surface shell. The turret-like extrusions made of protein P4 at the five-fold vertexes of the polymerase complex extend to the P8 layer. The envelope contains four viral encoded integral membrane proteins (P6, P9, P10, P13), one of which (fusogenic protein P6) anchors the receptor-binding spike P3. S, M and L are the dsRNA genomic segments.

ϕ6 shares features in common with the dsRNA viruses of the family *Reoviridae*. The genome is segmented and the virion is composed of three structural layers (Fig. 1). The structural and functional analogues found between ϕ6 and the members of *Reoviridae* have been taken as an evidence of common evolutional origin.[3,4] The similarity is most evident within the innermost protein capsid; the outer layers can vary both in composition and structure, also within each viral family. The conserved inner layer is a multi-functional molecular machine, which can perform genome encapsidation, replication, and transcription, while the outer layers facilitate interactions with the host cell during the particle internalization.[5]

Some of the functions of the inner capsid have been well documented for the members of the *Reoviridae* and the high resolution structure has been determined for the core particles of bluetongue virus and reovirus.[6,7] However, the mechanisms of genome encapsidation is relatively poorly understood,[8] and ϕ6 has so far been the only model to study the packaging and the consecutive replication of a segmented genome of a dsRNA virus.

Structure of the ϕ6 Virion and Life-Cycle

The spherical ϕ6 virion has three structural layers, which can be sequentially removed.[9,10] The outermost layer is a lipid-protein envelope and the nucleocapsid (NC) surface shell covers the polymerase complex (Fig. 1; Table 1). The envelope surrounding the proteinaceous NC contains five viral-encoded proteins and phospholipids originating from the host plasma membrane.[9,11] The NC surface is composed of a single viral protein (P8) which is arranged into an incomplete T=13 lattice in which the regions immediately surrounding the five-fold vertices are occupied by proteins extending from the underlying polymerase complex.[9,12] A similar arrangement has also been described for the intermediate layer of reovirus.[13] The polymerase complex of ϕ6 is formed by 120 copies of protein P1 arranged as 60 dimers on the icosahedral lattice.[12] Such an organization with two molecules of the same protein in two different conformations in an icosahedral asymmetric unit is so far only discovered in the core particles of the dsRNA viruses and is referred as the T=2 structure.[6]

Table 1. Nomenclature

Particle	Composition
Polymerase Complex	General name for the protein complex that carries out the RNA packaging and replication Forms the innermost structural layer of the virion Composed of proteins: P1- major capsid protein P2- RNA polymerase P4- multi-functional NTPase P7- small particle stabilizing protein
Procapsid (PC)	Empty polymerase complex
Core	Polymerase complex enclosing the dsRNA genome
Nucleocapsid (NC)	Core particle encapsidated in the protein P8 shell P8 forms the middle structural layer of the virion
Virion	Nucleocapsid encapsidated by lipid bilayer envelope Envelope forms the outermost structural layer of the virion Envelope contains proteins: P9, P10, and P13- integral membrane proteins P6- fusogenic protein, anchors P3 spike P3- the receptor-binding spike

As a dsRNA virus, ϕ6 needs to deliver not only its genome but also the viral replication machinery (the polymerase complex) into the host cell cytoplasm to initiate infection. During the complex entry process the structural layers of the ϕ6 virion are sequentially removed and the transcriptionally active polymerase complex is delivered into the cytoplasm.[14-16] The single-stranded (ss)RNA molecules produced early in the infection serve as templates for protein synthesis and are also packaged into the newly synthesized virions.

The first assembly intermediate detected during the ϕ6 infection cycle is an empty polymerase complex, also referred as the procapsid (PC).[17] This particle packages the single-stranded genomic precursor molecules and replicates them into the double-stranded form inside the capsid. The dsRNA-filled polymerase complex directs the late transcription and matures to a progeny virion by addition of the NC surface shell and the viral envelope.[18,19]

Genome Organization

The three genomic segments of ϕ6 are named according to their size; small, medium, and large (Fig. 2).[20] Each virion contains only one copy of each segment.[21] Abbreviations S, M and L are used for the dsRNA form of the genome, while lowercase letters, s, m and l, indicate the single-stranded, plus-sense RNA precursor molecules (Fig. 3). The dsRNA genome resides inside the polymerase complex throughout the viral life-cycle and the cell cytoplasm is exposed only to the single-stranded messenger (m)RNA molecules. The genes are clustered into functional groups to the three genome segments and the coding regions are flanked by distinct noncoding areas at both termini of each segment (Fig. 2).[20]

In Vitro Studies

The current understanding of the ϕ6 assembly pathway and life-cycle is largely based on in vitro studies. Immediately after the isolation of ϕ6, the virion associated transcriptase activity was detected[22] which led to in vitro transcription studies with purified core particles and

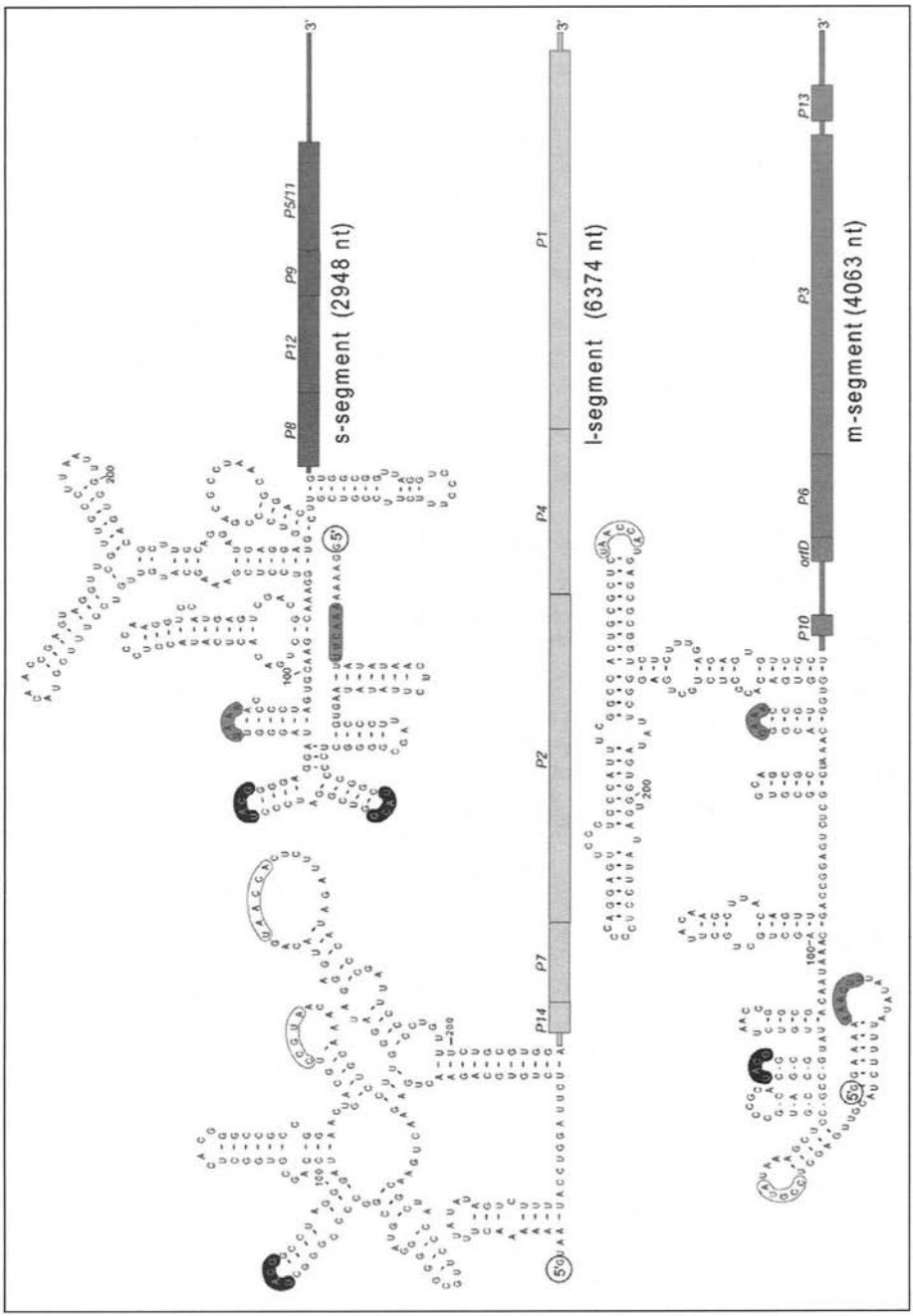


Figure 2. Genome organization. The three plus-sense segments are designated as s, m, and l. The protein coding regions are shown by wide bars and the noncoding areas by narrow bars. The secondary structure elements of the 5'-end pac-regions[35] are depicted enlarged. Conserved loop structure common to the three segments is shown with black background. Some single-stranded regions common between s and m, or m and l are indicated using a similar background.

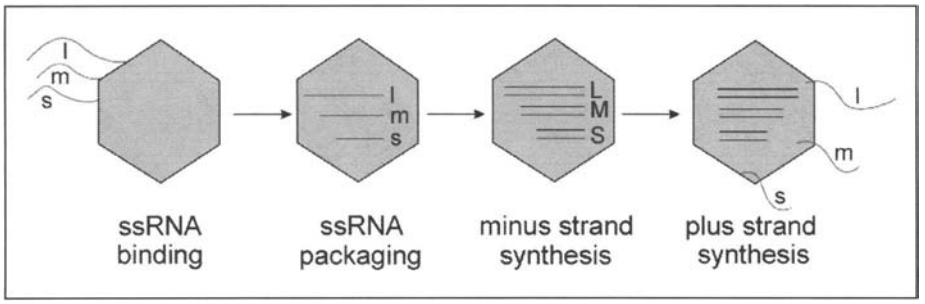


Figure 3. The in vitro replication cycle of ϕ6. The empty polymerase complex, procapsid (PC) binds the single-stranded genomic precursor molecules (s, m, l) on the outer surface of the particle. If ATP (or any NTP) is available, the PC packages the ssRNA molecules into the capsid. The polymerase then replicates the packaged segments into dsRNA form (S, M, L) inside the particle. These dsRNA-containing particles can start to transcribe the viral ssRNA molecules (s, m, l), which exit from the particle.

NCs.[23,24] A breakthrough for the genome packaging and replication analysis came a few decades later when isolated PCs, produced in *Escherichia coli* cells,[25] were shown to be capable of carrying out the whole ϕ6 RNA replication cycle in vitro conditions (see Fig. 3).[26] Since then the recombinant PCs have been an invaluable tool for the analysis of ssRNA packaging, replication, and transcription activities. Later an NC-spheroplast infection system was set up which allowed direct measurement of the biological functionality of viral NCs.[27] Furthermore, the NC surface shell protein was purified and was shown to reassemble on isolated dsRNA-filled polymerase complexes to yield infectious NCs.[10] These important tools were used to demonstrate that recombinant PCs could be turned to infectious particles.[28] This confirmed that the packaging of the tripartite genome in vitro conditions was a true intermediate reaction in the assembly pathway. Recently, the in vitro studies have extended to include the assembly of the empty PCs, and nowadays it is possible to carry out the entire assembly pathway from purified proteins and genomic ssRNA precursors to infectious NCs under defined in vitro conditions.[29] These in vitro systems have opened avenues to gain detailed information on assembly, genome packaging, replication, and transcription. Here we will focus on processes, which lead from the empty PC to the dsRNA-filled particle. Transcriptional activity and particle assembly are discussed when these processes are intimately connected to the topic of this article.

Components of the ϕ6 RNA Packaging and Replication System

The ϕ6 in vitro RNA packaging and replication system contains isolated empty polymerase complexes (procapsids, PCs) and single-stranded copies of the viral genome segments in a defined reaction mixture (Fig. 3). The packaging or the replication is neither dependent on any viral encoded nonstructural protein, nor on any host components.

Single-Stranded Genomic Precursor Molecules

The virus specific plus-sense ssRNAs for packaging reaction can be obtained from an NC transcription reaction,[23] or by using T7 RNA polymerase to transcribe the cDNA copies of the genome segments cloned under the T7 promoter.[28,30] The latter system allows the alteration of the single-stranded packaging substrates, which has been important for the localization of the genomic regulatory elements.

The noncoding regions at the 5'- and 3'-ends of the genome segments contain *cis*-acting signals for the packaging and replication of the genome (Fig. 2). The 5'-terminus was shown to carry the packaging signal, referred as the pac-site. It extends from the end of the segment 273,

280, and 209 nucleotides (nt) in s, m, and l segments, respectively.[31-35] Except the 5'-terminal homologous 18 nt, these regions have only very limited sequence similarity. However, several common structural elements can be found (Fig. 2). Some of the loop structures are common to two of the segments and one is found in all three segments. The pac-site of the s segment contains two copies of this conserved loop structure.[35] At the 3'-end of the segments there is a longer (75 nt) region of partial identity, which is predicted to fold into a transfer RNA-like secondary structure.[36]

Procapsid (PC)

The ϕ6 PC is composed of four protein species, P1, P2, P4, and P7 (Table 1). In addition, it is possible to produce and isolate PC particles missing any of the proteins P2, P4, P7, or both the proteins P2 and P7.[25,26,37] These incomplete particles together with the purified recombinant proteins[29,38-40] have been essential for the assigning of specific functions to the component proteins.

Capsid Forming Protein P1

Protein P1 is the major component of the PC and forms the peculiar T=2 capsid structure.[12,41,42] Although P1 has mainly a structural role in the polymerase complex it is likely to be involved in the initial binding and recognition of the genomic ssRNAs[43] as well as in the organization of the dsRNA genome inside the capsid. The other polymerase complex proteins P2, P4, and P7 can associate independently to the P1 shell.[37,38,44]

The Polymerase P2

A consensus sequence of known RNA-dependent RNA polymerases has been found in protein P2.[45] Particles missing P2 are unable to replicate RNA, but have ssRNA packaging activity.[46] Purified P2 can polymerize RNA also in the absence of the other PC proteins. The biochemically determined copy number for P2 is ~14 molecules per particle,[21] and P2 is proposed to reside underneath the five-fold vertices leading to actually 12 copies per particle.

A Multi-Functional NTPase P4

Protein P4 contains a consensus sequence common in NTP binding proteins.[47] In the presence of NDP or NTP and divalent cations, Mg^{2+} and Ca^{2+}, purified P4 assembles into ring-like homo-hexamers. The hexameric form of P4 is an active NTPase[39,48] hydrolyzing NTPs, dNTPs as well as ddNTPs to corresponding nucleoside diphosphates. The P4 NTPase specificity spectrum correlates with that observed for the RNA encapsidation reaction (see below) and the NTPase activity of P4 is stimulated by ssRNA.[39,48] Therefore it has been postulated that P4 is the NTPase providing energy for the genome translocation reaction. This is supported by the recent isolation and analysis of P4 null particles; Particles devoid of P4 were totally inert in ssRNA packaging.[49] P4 has also a role during the transcription when the newly synthesized mRNA molecules are extruded from the particle.[49] Furthermore, although P4 null particles can be isolated from the cell, under defined in vitro conditions the particle assembly is dependent on hexameric P4.[29]

A Small Particle Stabilizing Protein P7

P7 is a minor component of the polymerase complex. Purified P7 forms stable dimers in solution which are elongated in shape.[29,38] There is no known enzymatic activity for P7, but the analyses of particles missing P7 have revealed that P7 is needed for stable ssRNA packaging as well as for normal transcription.[38,46] In addition, P7 accelerates the in vitro assembly process.[29] The estimated copy number for P7 is 60 (or 30 dimers).[38] The exact location of P7 in

PC is not known, but it has been suggested to be located near the two-fold axes of the PC so that the elongated P7 dimer could stabilize the contacts between two P1 dimers.[29]

Procapsid to Core Transition

Cryo-electron microscopy based image reconstructions have revealed a considerably different conformation for the empty polymerase complex (depicted in Fig. 5A) compared to that of the virion-derived dsRNA-filled core particle (depicted in Fig. 5F) despite the same protein composition. The empty particle is smaller (~46 nm versus ~50 nm) and more angular. It also appeared to have inwards oriented "cups" at each of the five-fold facets instead of the round appearance of the core particle.[12] The difference between the dsRNA-filled and the compressed empty polymerase complex suggests that the particle maturation involves considerable conformational changes resulting also in the expansion of the particle. Whether the expansions take place during the genome packaging, minus-strand synthesis, or prior the activation of transcription is currently unknown.

In addition to the compressed dodecahedral and the rounded form of the polymerase complex, an intermediate conformation has been described. This conformation, referred as to an expanded PC (Fig. 5E), appears as a dodecahedron, with shallow depressions at the five-fold vertices. Such conformation can be observed in the preparations of empty PCs, and are also formed when the virion-derived polymerase complexes lose the dsRNA genome.[12] If this conformation represents a true assembly intermediate, it would suggest that the expansion occurs at least in two stages.

The ϕ6 in Vitro ssRNA Packaging and Replication System

The ϕ6 in vitro ssRNA packaging and replication system (Fig. 3) allows to dissect the ssRNA binding, ssRNA packaging, combined ssRNA packaging and replication (i.e., minus-strand synthesis reaction), and combined packaging, replication, and transcription activities (i.e., plus-strand synthesis reaction) by modifying the reaction conditions.[30,46,50-52]

The translocation of the ssRNA into the closed compartment requires energy in the form of hydrolysable nucleoside triphosphates, which can be either rNTPs, dNTPs, or ddNTPs.[50,51] Uncleavable nucleotide triphosphate analogs can not support the RNA packaging, although ssRNA binding to the empty particles can be observed in these conditions.[46] The packaged ssRNA and the replicated dsRNA are resistant to RNase-treatment and cosediment with the particle.[50,53]

The packaging is specific for ϕ6 plus-strands and this specificity is determined by the 5'-terminal pac-sites (Fig. 2).[31-33] The packaging of ssRNA segments containing an internal artificial hairpin structure is blocked so that the 3'-terminus is sensitive to RNase treatment.[54] This suggests that the packaging is initiated from the 5'-end of the segments. The synthesis of the complementary minus-strand occurs inside the particle and can initiate only after the 3'-end is accessible to the polymerase.[40,50,53]

The analyses of ϕ6 packaging using genomic ssRNA segments of reduced size have shown that the PCs of ϕ6 can uptake more than three ssRNA molecules to compensate the reduction in the total amount of nucleic acid to be packaged. The additional ssRNA appears to be always of the same class as the truncated or deleted segment.[55]

Sequential Packaging

ϕ6, like other dsRNA viruses with segmented genome, needs to have a mechanism to ensure that one copy of each of the individual genome segment will be packaged into the virus particle. Studies on ϕ6 have shed light to the mechanism how the tripartite genome may be incorporated into the capsid efficiently and accurately.

Apparently the polymerase complex has the ability to recognize the three viral single-stranded genomic precursor molecules in segment-specific manner. The segment-specific binding-sites are likely to be exposed on the outer surface of the particle,[34,46] so that the selection and recognition of the ssRNA molecules to be packaged could be accomplished already prior to the energy consuming packaging process. The binding affinities of the three genome segments differ so that the s segment has the highest affinity, followed by m and then by l.[46] Although particles can also bind ssRNA molecules that do not contain the viral pac-site, such unspecific binding does not lead to packaging.[26,46] Therefore we suggest that only specific binding can trigger on the packaging machinery or direct the end of the RNA molecule in a correct entering through P4 hexamer.

The PC can package each of the three ϕ6 plus-stands independently of each other into the PC, and the packaging efficiencies of the segments follow the binding-affinity order, s>m>l.[30,51] When the reaction contains more than one segment, the segments interfere with each other's packaging. The presence of s stimulates the packaging of m and similar relationship has been observed between m and l.[33,51] Evidence for this packaging order between the segments has also been observed in vivo experiments; particles containing either only the s segment, both the s and m segments, or all the three segments (but no other combinations of the segments) have been isolated from ϕ6 infected cells.[53] These results suggest that the three segments are packaged in a preferred order so that the s segment is packaged first followed by m and then by l.

It is not known how the sequential packaging is achieved, but this phenomenon is likely to be crucial for the correct and accurate encapsidation. The serial order of dependence between the segments could be achieved by conformational switches, so that the packaging of a segment creates the high affinity binding-site for the next segment to be packaged.[33,34,56] Alternatively, the RNA molecules could interact with each other and these RNA-RNA interactions could then interfere with the binding or the packaging process.

The model of conformational switches claims that binding-sites for the m and l segments would not be exposed on the surface of the naive PC, and that m could be packaged only after s and l only after s and m. However, as stated above, all the three genomic segments can be packaged into the PC independently of each other—albeit with different efficiencies. In addition it is known that the empty PC is not a rigid structure, but gentle heat-treatment (at 22°C) can induce conformational changes in the particle.[12] We suggest that the empty PC can spontaneously adopt the different conformations in which either s, m, or l binding-sites are exposed. At elevated temperatures, like during the packaging reaction at 30°C, these spontaneous changes are likely to occur. The highest packaging efficiency for s, could reflect the fact that the conformation which are capable of packaging s are the majority, while the forms having high affinity for m or l represent the minority. The idea of slow spontaneous conformational change in the PC during packaging reaction towards particles ready for packaging l segment would also explain the relatively slow packaging kinetics measured for l segment. The conformation which exposes binding-site(s) for s are probably favored in vivo since the s is preferentially packaged not only in vitro but also in vivo. Presumably, the switches during the packaging do not involve high-energy transitions as they are likely to be spontaneous and reversible (see below).

Mutations That Change the Sequential Packaging

Mutations in P1 (Arg_{14} to Gly) and P4 (Ser_{250} to Gln) affect the sequential packaging cascade so that s is not packaged or is packaged poorly without considerable effect on m and l packaging. This phenotype of the P1 capsid mutant has been taken as an evidence that P1 contains the primary recognition sites for genomic ssRNA segments.[43] Interestingly, the inability to package s by the P4 NTPase mutant particles appears only after the particles have been

frozen and thawed. Fresh P4 mutant particles function like wild-type PCs in packaging and there is no detectable difference between the fresh and frozen particles in sedimentation behavior or protein composition.[49] We suggest that the mutations in P4 and P1 affect the particle so that it easily undergoes the first conformational switch in the packaging cascade leading to the phenotype which expose binding-sites for m. Both of the mutant particles, with this phenotype, have also a tendency to lose the majority of protein P4 during purification. Our recent results show that the in vitro assembly of wild-type P4 on isolated mutant particles deficient of P4 rescues the packaging of segment s. This indicates that the proposed spontaneous conformational switch in the segment-specific binding-site may be reversible.

Model of Packaging Induced Particle Expansion

Qiao et al[34] and Mindich[57] have suggested a model that connects the sequential packaging of the three genomic ssRNA segments, the expansion of the particle and the headful packaging phenomenon. In this model it is envisaged that the empty unexpanded particle has binding-sites only for the segment s. After the packaging of s, the particle slightly expands, the binding-sites for s disappear, and binding-sites for m are exposed. Upon packaging of m, the particle further expands and binding-sites for l segment appear allowing the recognition and packaging of the l segment. It was proposed that the expected expansions of the PC and the changes in packaging specificity would not occur unless headful of ssRNA is translocated into the particle.

Packaging Densities

In hope to get understanding when the changes in the polymerase complex structure might (or have to) occur during the packaging and replication process we estimated the volumes of the different intermediate particles (Table 2; see also Fig. 5). We used an assumption that the segments are packaged in order s, m, l, and calculated the packaging densities for single- and double-stranded forms of the genome. It appeared, as expected, that the interior of the rounded polymerase complex (Fig. 5F; a conformation found from mature NCs and core particles) was large enough to accommodate the entire dsRNA genome. The RNA density observed was about 70% of that determined for dsDNA in crystals, which is considered as the highest packaging density. The packaging density of the expanded PC (Fig. 5E) reaches the upper density limit when the whole dsRNA genome is within the particle. This suggests that the expansion of the particle into the rounded form should occur prior to the initiation of the transcription. The highly active process of mRNA synthesis within the particle would probable be impossible if there was no space for RNA movements. The empty compressed particle (Fig. 5A) was far too small to include either the dsRNA or the ssRNA form of the genome. However, it appeared that the empty particle has space for s and m segments (Fig. 5D), but not for the l segment. This conclusion would hold even in the case that there would be ± 30% error in the calculated volume for the empty particle. Thus the compressed PC has to expand latest during the packaging of the l segment (Fig. 5F). Headful conditions are not achieved with s alone.

Initiation of Minus-Strand Synthesis—A Checkpoint for Genome Packaging

The packaging of the ssRNA genomic precursors is followed by the minus-strand synthesis inside the particle to form the dsRNA genome.[26,53] It is considered that the RNA replication does not take place simultaneously with the ssRNA uptake, it is rather turned on after completion of the packaging.

A question arises, what could be the signal (i) caused by the ssRNA genome packaging and (ii) sensed by the PC, that would lead to the activation of polymerase functions. Two

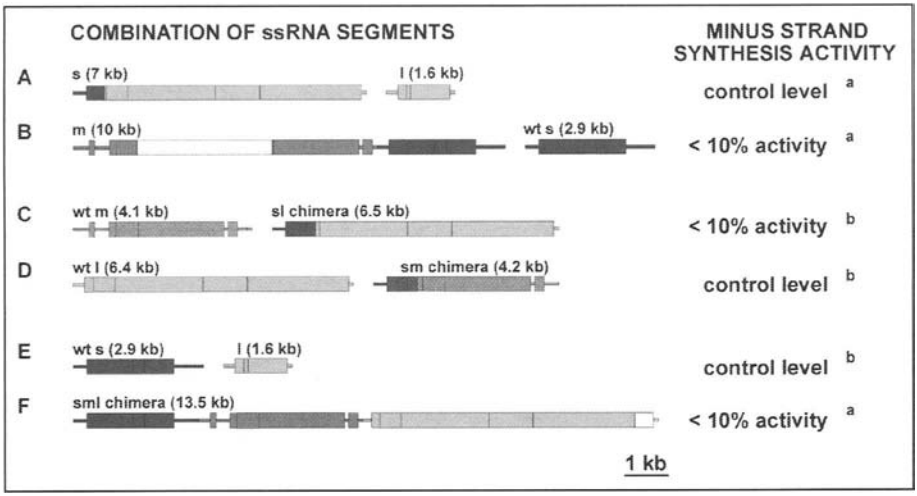


Figure 4. Minus strand synthesis initiation control. The minus strand synthesis activation has been tested using different combinations of modified and wild-type (wt) ssRNA segments. The ssRNA segments in each reaction (A- F) are presented schematically to illustrate the sizes and contents of the segments. Protein coding regions are shown by wide bars and the noncoding areas by narrow ones. The different grey levels indicate the origin of the RNA: light grey, L; middle grey, M; and dark grey, S segment. The white boxes indicate sequences that are not ϕ6-specific (originate from the *E. coli* lacZ gene). Control level refers to a reaction containing wild-type s, m, and l segments ([a]Results from Qiao et al[34]; [b]results from Poranen and Bamford[58]).

different models have been presented for the regulation the ϕ6 RNA replication. Qiao et al[34] proposed that ϕ6 uses a headful sensing mechanism so that initiation of minus-strand synthesis would simply be dependent on the amount of ssRNA packaged into the capsid. We have suggested[56,58] that the switch from packaging to minus-strand synthesis is caused by specific RNA-protein interactions.

To test the models, modified ssRNA segments were constructed and the activation of minus-strand synthesis was analyzed using different combinations of full-length genome segments together with the modified ones.[34,56,58] The sizes and the specificity classes of the modified segments varied so that it was possible to test the hypotheses.

A prediction of the model of Qiao et al[34] is, (i) that the activation of minus-strand synthesis should be equal in two reactions in which the total mass of ssRNA packaged is equal, and (ii) the minus-strand synthesis should only occur when the amount of ssRNA packaged matches to that of the full ϕ6 genome [2.9 kb(s) + 4.1 kb(m) + 6.4 kb(l) = 13.4 kb].

When a molecule equal in size to the sum of segments s and m (7 kb), with the pac-sequence from s, was packaged together with deleted l segment, Δl (1.6 kb), a control level minus-strand synthesis was obtained (Fig. 4A), while when a molecule equal in size to the sum of segments m and l (10 kb) with the pac-sequence from m was packaged together with the s segment (2.9 kb) only a about 10% of normal activation of RNA synthesis was observed (Fig. 4B).[34] The amount of RNA packaged into the particles in these two reactions is closely the same as the Δl segment is packaged in multiple copies to compensate the reduction in size.[55] Similarly, the sum of the lengths of the ssRNA molecules matches in reactions carried out with the normal m (4.1 kb) together with the chimeric segment of s and l (6.5 kb) (Fig. 4C), and with the normal l segment (6.4 kb) together with the chimera of s and m (4.2 kb) (Fig. 4D). In this case a control level ssRNA synthesis was only detected in reactions containing the latter combination of ssRNA

Table 2. Estimated inside volumes for ϕ6 assembly intermediates and nucleic acid packaging densities within the particles

Particle[a] Nucleic Acid	Particle Volume nm³	Nucleic Acid nt	Nucleic Acid Density nt/nm³
Rounded pol. c.	45 000[b]		
S+M+L		26 770	0.6
Expanded pol. c.	30 000[c]		
S+M+L		26 770	0.9
Compressed pol. c.	11 000[d]		
s		2 948	0.3
s+m		7 011	0.6
s+m+l		13 385	(1.2)[e]
S+M+L		26 770	(2.4)[e]
Bluetongue virus[f]	53 000		
dsRNA genome		38 438	0.7
Bacteriophage λ[g]	102 000		
dsDNA genome		97 000	0.9
dsDNA in crystals[g]			0.9

[a] pol. c. = polymerase complex; [b] Calculated assuming that the interior of the particle is a sphere. The radius (22 nm) was estimated from the three dimensional reconstructions of electron micrographs of NCs and PCs.[12,60] [c] Calculated assuming that the interior of the particle is a dodecahedron. The value for a side of a dodecahedron (a=15.7 nm) was calculated from the radius (r) according to a Råde and Westergren.[68] The radius was obtained as in note b. [d] Calculated as in note[c] and assuming that the inward oriented depressions at the five-folds occupy twelve times a volume of a five-cornered pyramid, which height is 11 nm. Calculations were base on formulas presented in Råde and Westergren.[68] [e] Values exceed those measured for dsDNA crystals. [f] Gouet et al;[69] [g] Casjens.[70]

molecules.[58] These two examples are against the prediction: (i) Furthermore, minus-strand synthesis was strongly activated in reactions where the sum of the lengths of the segments was only 4.5 kb (reaction with normal s and Δl) (Fig. 4E).[58] In this case the shorter than normal l segment could be packaged in multiple copies, but as shown by Mindich et al[55] the packaging of multiple copies of reduced size segments is not crucial for the activation of minus-strand synthesis. Moreover, a single ssRNA molecule having the size of total ϕ6 genome with the pac-site of the s (Fig. 4F) activated the RNA synthesis only weakly (less than 10% of the control).[34] These two examples are against prediction. (ii) It appears that other factors than the amount of packaged RNA are important to turn on the minus-strand synthesis.

It follows that specific signaling occurring during packaging might be involved in activating the replication. Systematic analyses using different combinations of truncated, deleted and modified segments allowed us to localize such sequence to the 5'-end of the l segment.[56,58] (The results described above are not in conflict with this observation). The replication always correlated with the presence of the l segment's pac-site irrespectively of the length of the segment containing it (full-length l, l with large internal deletion, chimeric segment with the 5'-terminal fragment from l, or 5'-end pac-site fragment of l). This suggests that the signal needed to activate the minus-strand synthesis is a specific interaction between the sequence or structure in the l segment's pac-region and a protein component in the particle. As the l segment is the last to be packaged it is rational that the l segment carries the information for replication initiation. The fact that the 5'-terminal fragment of l is a sufficient signal to turn on the RNA replication indicates that the minus-strand synthesis on s and m segments can begin

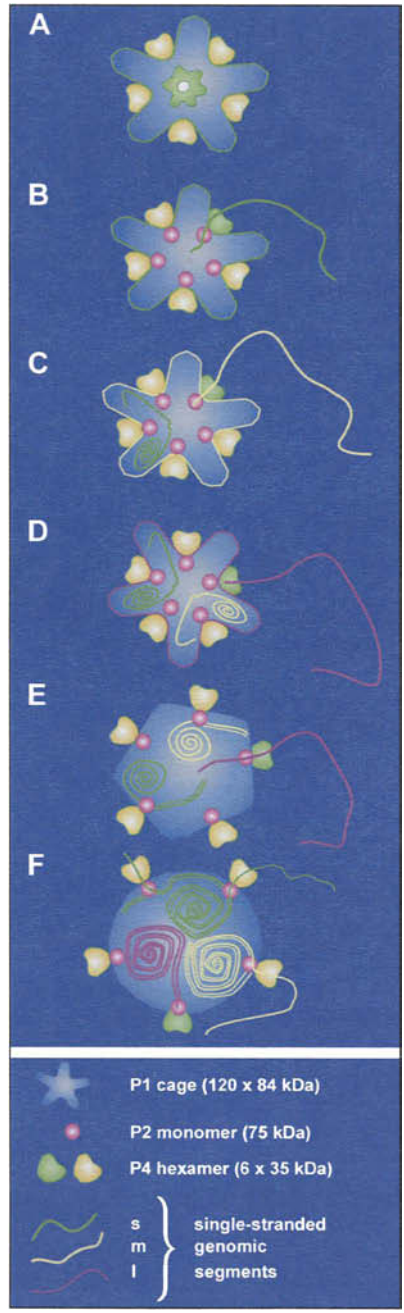


Figure 5. Model of ϕ6 RNA packaging and replication cascade. A) The empty procapsid of ϕ6 is composed of four proteins. P1 forms the structural skeleton of the particle, which is stabilized by P7 (not shown). The enzymatic components, the packaging NTPase P4 and the polymerase P2, are located at or near the five-fold axes. P2 is most likely located inside the capsid, while P4 comprises the turret-like extrusions. This location of hexameric P4 creates a symmetry mismatch into the site of ssRNA packaging. One of the five-folds differs from the others in its physical properties so that P4 is more tightly bound to this vertex (special vertex, colored in green) than the others (colored in yellow). The special vertex is active during the packaging of the single-stranded genomic segments (B-E), while the transcription is dependent on the P4 hexamers at the other vertices (F). During the genome packaging and replication the particle undergoes major conformational changes. Three structures have been described for the polymerase complex: the empty compressed particle (A), the expanded dodecahedral particle (E) and the mature particle with rounded appearance (F). In addition it is suggested that the particle undergoes conformational switches which change the specificity of the RNA binding-site on the outer surface of the particle (B-D; indicated on green, yellow and purple outlines). The packaging of the segments is initiated from the 5'-end, which contains a segment-specific pac-site. The empty particle is preferentially in a conformation that has high affinity binding-site for s segment (B). The binding or the packaging of the s segment induces a conformational change in the ssRNA binding-site so that the particle preferentially binds m segment (C) and the m is packaged. The m segment then induces switch to conformation that has high affinity for l segment (D). The internal volume of the compressed empty particle is too limited to pack the whole ssRNA genome and the particle has to expand latest during the packaging of the l segment. The 5'-end of the l segment carries a signal that is needed for the initiation of minus strand synthesis (on s and m, and later on l). Therefore we suggest that the 5'-end of the l segment induces an expansion of the particle, which allows the packaging of the full l segment and also switches on the replication. After completion of the minus strand synthesis on the l segment the particle can initiate transcription.[58] At this stage the dsRNA density within the expanded dodecahedral particle reaches its upper limits and we propose that the particle expands to the rounded conformation (F). In addition to the polymerase P2, the transcription is dependent on P4. The 11 putative transcriptional vertices (colored in yellow) are probably all active during transcription as multiple transcriptional forks has been detected in s and m segments at a given time.[67] The transcription occurs via a semi-conservative strand-displacement mechanism, and it is likely that the produced ssRNA molecules exit via the five-fold vertices.[67] It is not known whether the special packaging vertex contains a polymerase molecule, but the P4 of the special vertex can not compensate the absence of the P4 hexamers of the other vertices during transcription.

prior to the complete packaging of the l segment. This would be consistent with the observed sequential appearance of s, m, and l minus-strands in ϕ6 infected cells.[59]

Interestingly, the calculated packaging densities (Table 2) for different conformations of the polymerase complex indicate that the particle has to expand in order to be able to package the full-length l segment after the packaging of s and m. Thus, either the packaging of the s and m has to induce a conformational change in the particle to allow the packaging of the l (as suggested in the model of Qiao et al[34]) or the l segment 5'-end carries a signal which would induce a conformational change in the particle. The observation that the l segment 5'-end is needed for the activation of minus-strand synthesis could also be connected to such a conformational change: The polymerase (P2) active site could be buried in the compressed particle and only becomes available after transition from the compressed to the expanded conformation, or the 3'-end of the segments would not be accessible to the polymerase until the conformation is changed. (Interestingly, the transfer RNA-like conserved region at the 3'-end of the segments is crucial for the efficient initiation of minus-strand synthesis within the capsid[31] although is dispensable when replication is carried out with an isolated polymerase.[40] Based on these observations it has been proposed that 3'-end secondary structure might be needed for the correct localization of the ssRNA within the particle).[40]

We suggest that the transition of the polymerase complex from the compressed to the expanded form is a critical checkpoint during the particle maturation and is induced by the 5'-end of the l segment. The expansion allows the packaging of full l segment, and switches on the replication of the genome (Fig. 5E). The fact that the virion-derived polymerase complex collapses to the conformation of expanded PC (Fig. 5E) when it loses the genome, not to the form of the compressed naive PC,[12] suggest that the transition from the compressed to expanded PC is not a reversible change. The idea of particle expansion during the packaging of l does not exclude conformational change occurring already prior to the packaging of l. Conformational switches in the specific ssRNA binding-site in the PC are likely to regulate the changes in the packaging specificity from s to m and from m to l (Fig. 5B-D). As these changes can be spontaneous and reversible, and there is no requirement for change in the volume of the particle prior to packaging of the l (Table 2), we suggest that the switches in segment-specific binding-site do not involve major structural changes in the particle (Fig. 5B-D).

Symmetry Mismatch in the Procapsid of ϕ6

Cryo-electron microscope based image-reconstructions of the ϕ6 polymerase complex have depicted outward protruding turrets or satellite-like structures at each of the five-fold vertex of the icosahedral particle.[12,60] Images of empty particles composed of proteins P1, P2, and P7, however, lack these structures suggesting that the turrets are composed mainly of protein P4. The three-dimensional reconstructions of negatively stained purified P4 matched to the satellite structures supporting the idea that a P4 hexamer occupy the five-fold vertex positions in the PC.[60] This imply that there is a symmetry mismatch between the hexameric P4 and its location at the five-fold vertex.

A Special Packaging Vertex

The structural approaches have indicated that ϕ6 PCs are symmetrical particles having one putative packaging NTPase (P4 hexamer) at each vertex. Earlier studies have also suggested that the packaging of more than one segment can take place simultaneously, and this was taken as evidence that there could be more than one entry portal in the PC.[54]

Recent biochemical and genetic analysis have, however, suggested that one of the 12 vertexes in the ϕ6 PC may be physically and functionally different.[49] A single point mutation in P4 (Ser_{250} to Gln) leads to the formation of PC-like particles with reduced amounts of P4.

Similar particles can also be obtained by treating the wild-type PCs with detergent n-octyl-β-D-glucopyranoside. Both of these P4 deficient particles contain approximately 10% of the normal P4 amount. These results are based on in vitro recombinant systems. Similar observations have also been made earlier, during the very first replication analysis. It was noticed that the P4 content of particles isolated from ϕ6 infected cells was reduced 11-fold when the samples were treated with sarkosyl.[53] These observations suggest that there are at least two different sites for P4 molecules in the particle which differ in their binding affinity. The fraction of P4 that stays in the particle could comprise one P4 hexamer suggesting that the PC has one special vertex.

Although P4 null particles are totally inert in packaging, the particles with lowered amount of P4 were still able to package the three viral single-stranded genomic segments. Interestingly, the efficiency and the kinetics of the packaging were at least as good as with wild-type PCs proposing a similar mode of packaging.[49] Therefore we conclude that one P4 containing vertex is necessary and sufficient to carry out ϕ6 genome packaging, and propose that ϕ6 has only one packaging vertex.

If one of the vertexes is assigned for packaging, what is the role of the other P4 hexamers at the 11 remaining vertices? The packaged ssRNA in P4 deficient particles is efficiently replicated into dsRNA showing that the polymerase is fully functional. However, the transcription is totally blocked.[49] Likewise, the in vivo isolated sarkosyl treated particles were shown to be active in replication but inert in transcription.[53] It follows that the normal transcription is dependent on P4 and that the special vertex can not compensate the absence of the other P4 hexamers in transcription. Our recent results show that the transcriptional defect can be rescued by the addition of P4 on the P4 deficient particles, and the acquired activity correlates linearly to the amount of P4 incorporated.

How is the special vertex is determined in the apparently symmetrical PC of ϕ6? The nucleation of ϕ6 PC assembly is dependent on P4[29] and we propose that the P4 multimer that initiates the assembly of the particle would determine the location of the special vertex. This P4 could be more tightly incorporated into the particle.

Comparison to Tailed dsDNA Bacteriophages and Eukaryotic dsRNA Viruses

The packaging system of ϕ6 has certain similarities to the well-studied packaging systems of tailed dsDNA bacterial viruses. In both cases the genome is translocated into a preformed capsid (procapsid or prohead). The packaging is dependent on the hydrolysis of the high-energy bond in ATP (or any NTP in the case of ϕ6). There are global conformational changes in procapsid during the maturation, and these can be induced by elevated temperature, change in pH or genome packaging.[61,62]

The symmetry mismatch at the five-fold vertex of ϕ6 procapsid as well as the ring-shaped structure of P4 NTPase suggest that P4 might have a similar role in the RNA translocation as that of the portal proteins and terminases of dsDNA phages. For dsDNA phages it has been proposed that the symmetry mismatch might assist the rotation of the packaging machinery during the genome packaging[63] and this may apply also for ϕ6. As mentioned above ϕ6 P4 is also needed when the ssRNA molecules are extruded from the capsid.

The procapsid of the tailed dsDNA viruses contains one special vertex where the packaging machinery is located and through which the genome is both packaged and delivered. The ϕ6 procapsid appears symmetrical and contains altogether 12 putative packaging machineries, one in each vertex. However, one of the vertices differs from the others and, like in

dsDNA phages, this proposed special vertex was shown to function during genome packaging.[49] The exit of the single-stranded transcripts produced inside the ϕ6 capsid occurs via the other vertices, the transcriptional vertices. The dsDNA phages establish the unique vertex by nucleating their capsid assembly around the portal. The special vertex in ϕ6 procapsid could also be determined during the assembly as the multimeric P4 is needed for the nucleation of PC assembly.

On the other hand, certain characteristics seem to be more divergent between ϕ6 and the bacterial dsDNA viruses. First of all the ϕ6 genome is packaged in a single-stranded precursor form that is only later replicated into the dsRNA genome inside the capsid. Obviously the folded structure characteristics to ssRNA molecules has to be melted prior or during the packaging. The genome of ϕ6 is segmented, which brings additional complexity to the selection of the nucleic acid molecules to be packaged. The templates for packaging are genome length molecules, not concatameric nucleic acid, as is typical for the tailed dsDNA phages. There is no parallel for the scaffolding protein in the ϕ6 system, rather all the components active during the packaging are also found in the mature virion.

The eukaryotic *Reoviridae* and ϕ6 share similar genome replication strategy in which the single-stranded genomic precursor molecules are replicated to double-stranded form within the viral capsid. The five-fold vertices in the ϕ6 polymerase complex appear to be the centers of activity during genome packaging and replication. Likewise, transcription and mRNA capping enzymes are located at the icosahedral vertices of rotavirus, bluetongue virus and reovirus. In the case of rotavirus and bluetongue virus these enzymes are placed within the core, while in the reovirus the capping enzymes form turret-like structures on the surface of the inner capsid (which appear very similar to the structures made of P4 of ϕ6). Electron microscopic observations of transcribing rotavirus cores have depicted mRNA molecules extruding at, or close to, the five-fold axes of the core,[64] and this most likely applies also for ϕ6.

Unlike ϕ6, the encapsidation and replication processes of the members of the *Reoviridae* are dependent on several viral encoded nonstructural proteins and occur in cytoplasmic inclusions. Interestingly, one of the nonstructural proteins (NS2 of bluetongue virus, NSP2 of rotavirus, and σNS of reovirus) is an NTPase, which could be involved in genome packaging.[8] The NSP2 of rotavirus forms homo-multimers and has helix destabilization activity.[65,66]

Conclusions and Future Directions

Our current view of the ϕ6 packaging and replication cascade is summarized in Figure 5. Even though considerable progress has been made in developing and analyzing the ϕ6 packaging system, our understanding of the packaging at mechanistic level is still relatively poor. The structural changes during the packaging process are still highly hypothetical. However, the estimations of the interior volumes of the intermediate particles gave some indications when the transitions might occur (Table 2; Fig. 5). The compromized stability and homogeneity of the particles and the spontaneous structural changes have hampered the structural analyses. The specific RNA-protein interactions during the packaging and replication and the changes in the packaging specificity are interesting areas in which mutational analysis of the secondary structure elements at the pac-sites will probably give new insight. We are actively studying the single packaging vertex and the role of P4 NTPase in transcription. It has appeared that hexameric P4 has helicase activity.[71] We hope that the high-resolution structural information on the packaging components and possible intermediate particles will be available in the future. This together with the biochemical and genetic evidence would greatly increase our understanding of the mechanism of the ϕ6 RNA translocation events.

References

1. Vidaver AK, Koski RK, Van Etten JL. Bacteriophage φ6: A lipid-containing virus of Pseudomonas phaseolicola. J Virol 1973; 11:799-805.
2. Mindich L, Qiao X, Qiao J et al. Isolation of additional bacteriophages with genomes of segmented double-stranded RNA. J Bacteriol 1999; 181(15):4505-8.
3. Hendrix RW. Evolution: The long evolutionary reach of viruses. Curr Biol 1999; 9(24):R914-7.
4. Bamford DH. Do viruses form lineages across different domains of life? Res Microbiol 2003; 154(4):231-6.
5. Bamford DH. Virus structures: Those magnificent molecular machines. Curr Biol 2000; 10(15):R558-61.
6. Grimes JM, Burroughs JN, Gouet P et al. The atomic structure of the bluetongue virus core. Nature 1998; 395(6701):470-8.
7. Reinisch KM, Nibert ML, Harrison SC. Structure of the reovirus core at 3.6 Å resolution. Nature 2000; 404(6781):960-7.
8. Patton JT, Spencer E. Genome replication and packaging of segmented double-stranded RNA viruses. Virology 2000; 277(2):217-25.
9. Van Etten JV, Lane L, Gonzalez C et al. Comparative properties of bacteriophage φ6 and φ6 nucleocapsid. J Virol 1976; 18(2):652-8.
10. Olkkonen VM, Ojala PM, Bamford DH. Generation of infectious nucleocapsids by in vitro assembly of the shell protein on to the polymerase complex of the dsRNA bacteriophage φ6. J Mol Biol 1991; 218(3):569-81.
11. Laurinavicius S, Käkelä R, Bamford DH et al. The origin of phospholipids of the enveloped bacteriophage φ6. Virology 2004; in press.
12. Butcher SJ, Dokland T, Ojala PM et al. Intermediates in the assembly pathway of the double-stranded RNA virus φ6. EMBO J 1997; 16(14):4477-87.
13. Metcalf P, Cyrklaff M, Adrian M. The three-dimensional structure of reovirus obtained by cryo-electron microscopy. EMBO J 1991; 10(11):3129-36.
14. Bamford DH, Romantschuk M, Somerharju PJ. Membrane fusion in prokaryotes: Bacteriophage φ6 membrane fuses with the Pseudomonas syringae outer membrane. EMBO J 1987; 6(5):1467-73.
15. Romantschuk M, Olkkonen VM, Bamford DH. The nucleocapsid of bacteriophage φ6 penetrates the host cytoplasmic membrane. EMBO J 1988; 7(6):1821-9.
16. Poranen MM, Daugelavicius R, Ojala PM et al. A novel virus-host cell membrane interaction. Membrane voltage-dependent endocytic-like entry of bacteriophage φ6 nucleocapsid. J Cell Biol 1999; 147(3):671-82.
17. Mindich L, Davidoff-Abelson R. The characterization of a 120 S particle formed during φ6 infection. Virology 1980; 103(2):386-91.
18. Bamford DH, Mindich L. Electron microscopy of cells infected with nonsense mutants of bacteriophage φ6. Virology 1980; 107(1):222-8.
19. Ewen ME, Revel HR. In vitro replication and transcription of the segmented double-stranded RNA bacteriophage φ6. Virology 1988; 165(2):489-98.
20. Mindich L, Bamford DH. Lipid-containing bacteriophages. In: Calendar R, ed. The Bacteriophages. New York: Plenum Publishing Corporation, 1988:475-519.
21. Day LA, Mindich L. The molecular weight of bacteriophage φ6 and its nucleocapsid. Virology 1980; 103(2):376-85.
22. Van Etten JL, Vidaver AK, Koski RK et al. RNA polymerase activity associated with bacteriophage φ6. J Virol 1973; 12(3):464-71.
23. Emori Y, Iba H, Okada Y. Transcriptional regulation of three double-stranded RNA segments of bacteriophage φ6 in vitro. J Virol 1983; 46(1):196-203.
24. Ojala PM, Bamford DH. In vitro transcription of the double-stranded RNA bacteriophage φ6 is influenced by purine NTPs and calcium. Virology 1995; 207(2):400-8.
25. Gottlieb P, Strassman J, Bamford DH et al. Production of a polyhedral particle in Escherichia coli from a cDNA copy of the large genomic segment of bacteriophage φ6. J Virol 1988; 62(1):181-7.
26. Gottlieb P, Strassman J, Qiao XY et al. In vitro replication, packaging, and transcription of the segmented double-stranded RNA genome of bacteriophage φ6: Studies with procapsids assembled from plasmid-encoded proteins. J Bacteriol 1990; 172(10):5774-82.

27. Ojalsa PM, Romantschuk M, Bamford DH. Purified ϕ6 nucleocapsids are capable of productive infection of host cells with partially disrupted outer membranes. Virology 1990; 178(2):364-72.
28. Olkkonen VM, Gottlieb P, Strassman J et al. In vitro assembly of infectious nucleocapsids of bacteriophage ϕ6: Formation of a recombinant double-stranded RNA virus. Proc Natl Acad Sci USA 1990; 87(23):9173-7.
29. Poranen MM, Paatero AO, Tuma R et al. Self-assembly of a viral molecular machine from purified protein and RNA constituents. Mol Cell 2001; 7(4):845-54.
30. Gottlieb P, Strassman J, Qiao X et al. In vitro packaging and replication of individual genomic segments of bacteriophage ϕ6 RNA. J Virol 1992; 66(5):2611-16.
31. Frilander M, Gottlieb P, Strassman J et al. Dependence of minus-strand synthesis on complete genomic packaging in the double-stranded RNA bacteriophage ϕ6. J Virol 1992; 66(8):5013-7.
32. Gottlieb P, Qiao X, Strassman J et al. Identification of the packaging regions within the genomic RNA segments of bacteriophage ϕ6. Virology 1994; 200(1):42-7.
33. Qiao X, Casini G, Qiao J et al. In vitro packaging of individual genomic segments of bacteriophage ϕ6 RNA: Serial dependence relationships. J Virol 1995; 69(5):2926-31.
34. Qiao X, Qiao J, Mindich L. Stoichiometric packaging of the three genomic segments of double-stranded RNA bacteriophage ϕ6. Proc Natl Acad Sci USA 1997; 94(8):4074-9.
35. Pirttimaa MJ, Bamford DH. RNA secondary structures of the bacteriophage ϕ6 packaging regions. RNA 2000; 6(6):880-9.
36. Mindich L, Qiao X, Onodera S et al. Heterologous recombination in the double-stranded RNA bacteriophage ϕ6. J Virol 1992; 66(5):2605-10.
37. Paatero AO, Mindich L, Bamford DH. Mutational analysis of the role of nucleoside triphosphatase P4 in the assembly of the RNA polymerase complex of bacteriophage ϕ6. J Virol 1998; 72(12):10058-65.
38. Juuti JT, Bamford DH. Protein P7 of phage ϕ6 RNA polymerase complex, acquiring of RNA packaging activity by in vitro assembly of the purified protein onto deficient particles. J Mol Biol 1997; 266(5):891-900.
39. Juuti JT, Bamford DH, Tuma R et al. Structure and NTPase activity of the RNA-translocating protein (P4) of bacteriophage ϕ6. J Mol Biol 1998; 279(2):347-59.
40. Makeyev EV, Bamford DH. Replicase activity of purified recombinant protein P2 of double-stranded RNA bacteriophage ϕ6. EMBO J 2000; 19(1):124-33.
41. Ktistakis NT, Lang D. The dodecahedral framework of the bacteriophage ϕ6 nucleocapsid is composed of protein P1. J Virol 1987; 61(8):2621-3.
42. Olkkonen VM, Bamford DH. The nucleocapsid of the lipid-containing double-stranded RNA bacteriophage ϕ6 contains a protein skeleton consisting of a single polypeptide species. J Virol 1987; 61(8):2362-7.
43. Onodera S, Qiao X, Qiao J et al. Isolation of a mutant that changes genomic packaging specificity in ϕ6. Virology 1998; 252(2):438-42.
44. Casini G, Qiao X, Mindich L. Reconstitution of active replicase in procapsids of the segmented dsRNA bacteriophage ϕ6. Virology 1994; 204(1):251-3.
45. Koonin EV, Gorbalenya AE, Chumakov KM. Tentative identification of RNA-dependent RNA polymerases of dsRNA viruses and their relationship to positive-strand RNA viral polymerases. FEBS Lett 1989; 252(1-2):42-6.
46. Juuti JT, Bamford DH. RNA binding, packaging and polymerase activities of the different incomplete polymerase complex particles of dsRNA bacteriophage ϕ6. J Mol Biol 1995; 249(3):545-54.
47. Mindich L, Nemhauser I, Gottlieb P et al. Nucleotide sequence of the large double-stranded RNA segment of bacteriophage ϕ6: Genes specifying the viral replicase and transcriptase. J Virol 1988; 62(4):1180-5.
48. Paatero AO, Syvaoja JE, Bamford DH. Double-stranded RNA bacteriophage ϕ6 protein P4 is an unspecific nucleoside triphosphatase activated by calcium ions. J Virol 1995; 69(11):6729-34.
49. Pirttimaa M, Paatero AO, Frilander M et al. Nonspecific nucleoside triphosphatase P4 of double-stranded RNA bacteriophage ϕ6 is required for single-stranded RNA packaging and transcription. J Virol 2002; 76(20):10122-7.
50. Gottlieb P, Strassman J, Frucht A et al. In vitro packaging of the bacteriophage ϕ6 ssRNA genomic precursors. Virology 1991; 181(2):589-94.

51. Frilander M, Bamford DH. In vitro packaging of the single-stranded RNA genomic precursors of the segmented double-stranded RNA bacteriophage ϕ6: The three segments modulate each other's packaging efficiency. J Mol Biol 1995; 246(3):418-28.
52. van Dijk AA, Frilander M, Bamford DH. Differentiation between minus- and plus-strand synthesis: Polymerase activity of dsRNA bacteriophage ϕ6 in an in vitro packaging and replication system. Virology 1995; 211(1):320-3.
53. Ewen ME, Revel HR. RNA-protein complexes responsible for replication and transcription of the double-stranded RNA bacteriophage ϕ6. Virology 1990; 178(2):509-19.
54. Qiao X, Qiao J, Mindich L. Interference with bacteriophage ϕ6 genomic RNA packaging by hairpin structures. J Virol 1995; 69(9):5502-5.
55. Mindich L, Qiao X, Qiao J. Packaging of multiple copies of reduced-size genomic segments by bacteriophage ϕ6. Virology 1995; 212(1):213-7.
56. Frilander M, Poranen M, Bamford DH. The large genome segment of dsRNA bacteriophage ϕ6 is the key regulator in the in vitro minus and plus strand synthesis. RNA 1995; 1(5):510-8.
57. Mindich L. Precise packaging of the three genomic segments of the double-stranded RNA bacteriophage ϕ6. Microbiol Mol Biol Rev 1999; 63(1):149-60.
58. Poranen MM, Bamford DH. Packaging and replication regulation revealed by chimeric genome segments of double-stranded RNA bacteriophage ϕ6. RNA 1999; 5(3):446-54.
59. Pagratis N, Revel HR. Minus-strand RNA synthesis by the segmented double-stranded RNA bacteriophage ϕ6 requires continuous protein synthesis. Virology 1990; 177(1):281-8.
60. de Haas F, Paatero AO, Mindich L et al. A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage ϕ6. J Mol Biol 1999; 294(2):357-72.
61. King J, Chiu W. The procapsid-to-capsid transition in double-stranded DNA bacteriophages. In: Chiu W, Burnett RM, Garcea RL, eds. Structural biology of viruses. New York: Oxford University Press, 1997:288-311.
62. Lata R, Conway JF, Cheng N et al. Maturation dynamics of a viral capsid: Visualization of transitional intermediate states. Cell 2000; 100(2):253-63.
63. Hendrix RW. Symmetry mismatch and DNA packaging in large bacteriophages. Proc Natl Acad Sci USA 1978; 75(10):4779-83.
64. Lawton JA, Estes MK, Prasad BV. Three-dimensional visualization of mRNA release from actively transcribing rotavirus particles. Nat Struct Biol 1997; 4(2):118-21.
65. Taraporewala Z, Chen D, Patton JT. Multimers formed by the rotavirus nonstructural protein NSP2 bind to RNA and have nucleoside triphosphatase activity. J Virol 1999; 73(12):9934-43.
66. Taraporewala ZF, Patton JT. Identification and characterization of the helix-destabilizing activity of rotavirus nonstructural protein NSP2. J Virol 2001; 75(10):4519-27.
67. Usala SJ, Brownstein BH, Haselkorn R. Displacement of parental RNA strands during in vitro transcription by bacteriophage ϕ6 nucleocapsids. Cell 1980; 19(4):855-62.
68. Råde L, Westergren B. Mathematics handbook for Science and Engineering. 3rd ed. Lund: Studentlitteratur, 1995.
69. Gouet P, Diprose JM, Grimes JM et al. The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. Cell 1999; 97(4):481-90.
70. Casjens S. Principles of virion structure, function, and assembly. In: Chiu W, Burnett RM, Garcea RL, eds. Structural biology of viruses. New York: Oxford University Press, 1997:3-37.
71. Kainov DE, Pirttimaa M, Tuma R et al. RNA packaging device of double-stranded RNA bacteriophages, possibly as simple as hexamer of P4 protein. J Biol Chem 2003; 278(48):48084-48091.

CHAPTER 9

Cleavage and Packaging of Herpes Simplex Virus 1 DNA

Herpesvirus Assembly

Joel D. Baines and Sandra K. Weller

Abstract

Herpes simplex virus DNA accumulates in the nuclei of infected cells as long concatemers. The packaging machinery recognizes signals within the concatemers, cleaves the DNA to generate unit length monomers, and inserts the cleaved genomes into preformed capsids. The goals of this work are to review the current understanding of these processes and to provide working models based on this information.

Introduction

The family *Herpesviridae* encompasses many human and animal pathogens that can cause a wide variety of medically and economically important diseases in a broad range of hosts. Of the more than 100 herpesviruses isolated to date, eight can cause significant disease in humans, especially in immunocompromised individuals. Understanding DNA packaging of herpesviruses has important practical implications because the process is a potential target for new antiviral strategies. In fact, resistance to two classes of novel antiviral drugs map to genes known to be essential for DNA packaging.[1-3] Since the processes of DNA cleavage and packaging are likely to be highly conserved, agents targeting these processes may enjoy broad efficacy against many different herpesviruses. It is also anticipated that such antiviral drugs would be extremely safe given the unique protein machinery that mediates DNA cleavage and packaging.

While considering the material in this chapter, we anticipate that the reader will be struck by the remarkable similarities between the assembly and packaging mechanisms of herpesviruses and those of the more extensively studied DNA bacteriophages. Indeed, the reactions share a number of important features including (i) formation of a procapsid intermediate consisting of a capsid shell initially supported by an internal scaffold, (ii) replacement of the internal scaffold with viral DNA, (iii) insertion of DNA through a unique vertex (the portal vertex of the procapsid), and (iv) derivation of unit length molecules by endonucleolytic cleavage of complex DNA concatemers. By analogy to bacteriophage systems, it is also presumed that herpesviruses encode a multicomponent enzyme, the terminase that links the DNA to the capsid, cleaves the DNA precisely into genomic lengths, and drives the cleaved genomes into the capsid through the hydrolysis of ATP. Although identification of the terminase has been hampered by the lack of an in vitro packaging system, genetic and limited biochemical data

Viral Genome Packaging Machines: Genetics, Structure, and Mechanism, edited by Carlos Enrique Catalano.

have begun to identify likely terminase subunits and elucidate their respective contributions to the packaging reactions.

Reviews of HSV DNA replication and capsid assembly have been published and will not be explored in detail here except as they relate to DNA packaging.[4-6] Rather, the goals of this chapter are to first, illustrate and summarize what is known about herpesvirus DNA cleavage and packaging and second, suggest models consistent with current data that will serve as the basis for future experimental investigations.

Arrangement of the HSV Genome and Packaging Sequences

The HSV-1 genome contains two covalently linked segments of nonrepeated DNA. These unique segments are flanked by inverted repeats and have been designated unique long (U_L, approximately 128 Kbp) and unique short (U_S, approximately 25.0 Kbp)[7] (Fig. 1). The two segments can invert relative to one another, allowing 4 isomers to accumulate in equimolar amounts in infected cells.[8,9] The open reading frames in the unique regions of a prototypical isomeric arrangement are sequentially numbered from left to right.[10]

The inverted repeats flanking U_L have been designated ab and b'a', whereas repeats flanking U_S are designated a'c' and ca[11,12] (Fig. 1). The a sequences are necessary and sufficient for proper cleavage and packaging of viral genomes.[13-17] In packaged virion DNA, the a sequences at genomic termini are in the same orientation whereas the a sequences at the L/S junction are in opposite orientation (indicated a'). More than one a sequence can be present at the L terminus (indicated a_n), whereas the S terminus invariably contains a single a sequence.[12] Sequences within the a sequences of various HSV strains can be arranged as *DR1-Uc-DR4*$_m$*-DR2*$_n$*-Ub-DR1* where m and n designate variable numbers of repeats (Fig. 1). The sizes of these components vary from strain to strain but are believed to be consistent within strains. As an example, the F strain of HSV-1 contains a *Uc* of 58 bp, a *Ub* of 64 bp, and *DR1*, *DR4* and *DR2* sequences of 20, 37, and 12 bp, respectively. Thus, the a sequence length can vary considerably due to variable numbers and types of repeat units.[18,19] Single a sequences are flanked by copies of *DR1* whereas junctions containing more than one a sequence share a single intervening *DR1*.[13,18,19]

Most of the discussion in this review will concern highly conserved sequences within *Uc* and *Ub* that have been designated *Pac 2* and *Pac 1*, respectively[20,21] (see Fig. 1). Sequences

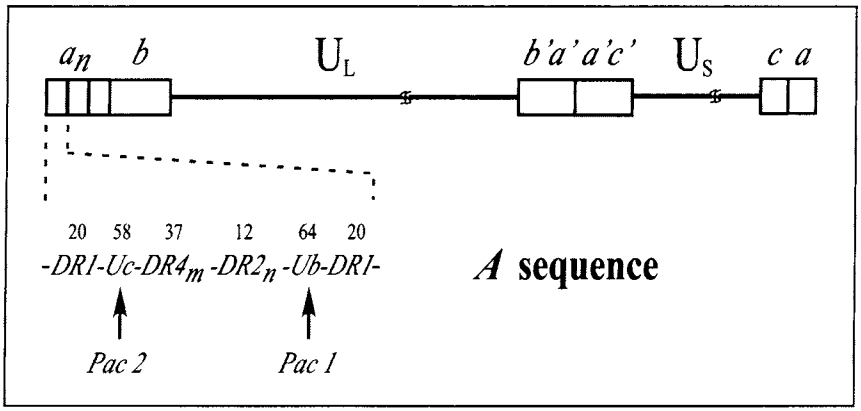


Figure 1. Schematic diagram of HSV-1 DNA. Unique long (U_L) and unique short (U_S) components are flanked by inverted repeats (boxes). The a sequences at genomic ends contain sequences necessary for packaging DNA. They are reiterated at the internal L/S junction in opposite orientation (designated by a'). Detail of the a sequence is also shown. The lengths of the a sequence components are indicated in bp.

containing a *DR1, Pac 2, Pac 1,* and at least some repeated units, are sufficient for the generation of genomic termini.[20,22-24] *Pac 1* sequences consist of a 3-7 bp A or T rich sequence flanked by a poly C sequence. *Pac 2* sequences contain a 5-10 bp A-rich region near a CGCGGCG motif. In murine cytomegalovirus (MCMV), the poly C of *Pac 1,* and the A-rich and CGCGGCG motifs of *Pac 2* have been shown to be essential for production of genomic termini found in extracellular virions.[25]

Concatemeric DNA cleavage occurs within *DR1* sequences leaving a trunctated *DR1* of 18 bp plus a single-base 3' overhang. Circularization of the genome early in the next round of infection would be expected to generate a single complete a sequence.[13,18,26] Although the simplest model would suggest that both ends could be derived from a single staggered cleavage event, evidence supports the contention that there are two separate cleavages.[23] The first cleavage event is believed to initiate packaging, and U_L DNA is threaded into the capsid first, followed by the U_L/U_S junction, and U_S (see Fig. 2). The second cleavage at the end of U_S is then believed to release the packaged genome from the concatemer. This model is consistent with the experimental observation that free ends on the concatemer contain U_L sequences but not U_S sequences,[27-29] and predicts that the concatemeric end consists of a truncated *DR1* of 18 bp plus a single-base 3' overhang, followed by *Pac 2, Pac 1,* and U_L sequences (see Fig. 2).[30] It follows that additional rounds of packaging could potentially initiate at the U_L terminus at the free end of concatemeric DNA.[31]

Capsid Maturation

During the herpesvirus cleavage/packaging reaction, DNA is inserted into preformed capsids. Cleaved viral genomes are not detected in cells infected with viruses that fail to assemble capsids, suggesting that capsids contain essential parts of the cleavage/packaging machinery.[32] There are four morphologic types of capsids that can be distinguished by electron microscopic examination of thin sections of herpesvirus-infected cells, and they are designated procapsids and capsid types A, B, and C.[33,34] The capsid shell of procapsids is morphologically distinct from all other capsid types and appears porous and roughly spherical by electron microscopy.[33,35] Procapsids contain internal scaffold proteins. Capsid types A, B and C are more angularized icosahedrons of approximately 120 nm in diameter. Type A capsids consist of only the icosahedral shell, type B capsids contain the angular shell and a roughly spherical internal scaffold, and type C capsids contain DNA and lack the scaffold.[34,36-38] The fact that type C capsids contain only the capsid shell and DNA indicates that the scaffold is lost during the packaging reaction (see Fig. 4).

The internal scaffold is comprised of the scaffold protein encoded by $U_L26.5$ and the viral protease encoded by U_L26.[39-43] The U_L26 and $U_L26.5$ proteins are encoded by overlapping transcripts and the encoded proteins share identical C-termini (see Fig. 3). The amount of $U_L26.5$ gene product far exceeds that of U_L26, presumably because of the strong $U_L26.5$ promoter that lies within the U_L26 open reading frame. The protease domain resides in the N-terminus of U_L26, and is absent from the $U_L26.5$-encoded scaffold. The unprocessed forms of U_L26 and $U_L26.5$ proteins are initially incorporated into capsids (Fig. 4).[44,45] The C-termini of the U_L26 and $U_L26.5$ gene products can interact with the major capsid protein, VP5, and such associations are believed to play critical roles in the assembly of the procapsid.[46-49]

In one widely accepted model of capsid maturation,[33] the procapsid is the precursor of all capsid types. Upon activation, the protease cleaves both (i) itself to release protease (VP24), scaffold (VP21), and a 25 amino acid C-terminal peptide, and (ii) the C-terminus of $U_L26.5$ protein to release the major scaffold protein (VP22a) and the 25 amino acid peptide[42,43,50] (Fig. 3). The protease domain (VP24) remains within the capsid while the C termini of the U_L26 and $U_L26.5$ scaffold proteins (VP21 and VP22a, respectively) are extruded. Presumably,

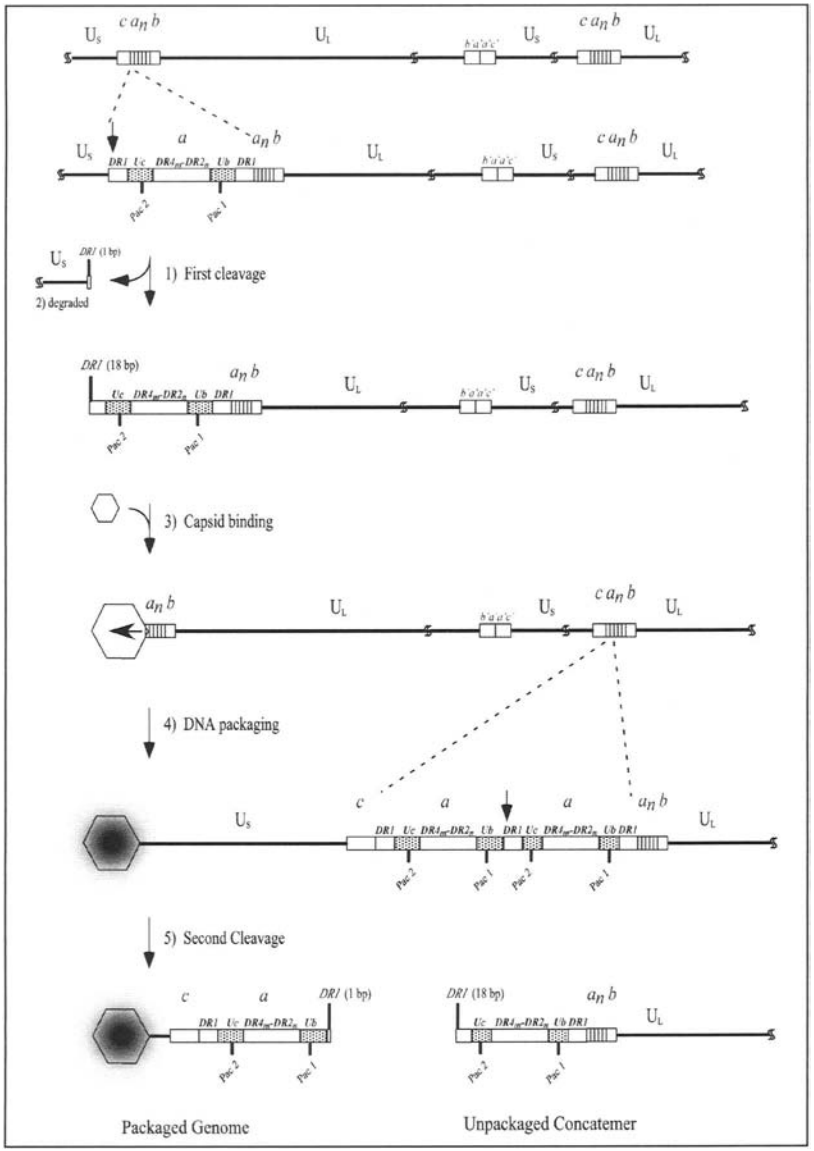


Figure 2. Proposed events in cleavage and packaging of HSV-1 DNA. Schematic diagram of concatemeric DNA, and a detailed view of the a sequence are shown in the first two lines. The first cleavage occurs within the *DR1* immediately adjacent to *Pac 2*, such that the DNA that is to be packaged contains *DR1* with 18 of its 20 base pairs and a single base 3' overhang (not indicated, step 1). The other (U_S) end, terminated by a single bp of *DR1* plus a single base 3' overhang (not indicated), is presumably degraded (step 2). After association of the cleaved DNA end with the capsid portal (step 3), U_L is packaged, followed by U_S (step 4). The second cleavage is diagrammed in the last 2 lines as occurring within the *DR1* shared between two a sequences (step 5), although the presence of 2 tandem a sequences is not necessary for successful cleavage/packaging. This second cleavage is dependent on *Pac 1* sequences, and occurs such that the DNA to be packaged is terminated by a single base pair of *DR1*. As a result, the terminus of the unpackaged concatemeric DNA (bottom right) is identical to the end to be packaging after the first DNA cleavage (step 1), and should therefore be able to initiate a subsequent round of packaging.

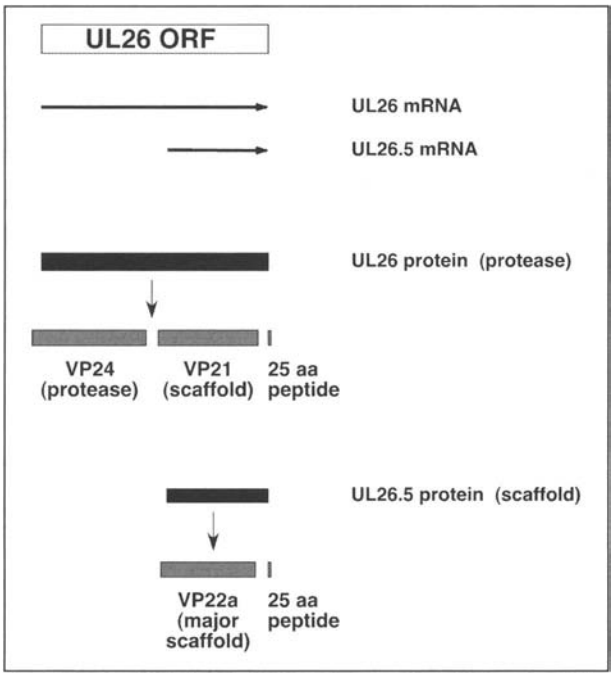


Figure 3. The U_L26 gene and the U_L26 and $U_L26.5$ mRNA transcripts and protein products are depicted at the top. Primary translation products are represented by black boxes. Proteolysis products are depicted in gray. The proteolytic products are VP24, which contains the protease domain, two scaffold proteins (VP21 and VP22a) and a 25 amino acid terminal peptide.

the capsid shell is thereby released from constraints that preclude maturation into the more angularized form of capsids. Supporting this proposition are observations that the timing of protease activation correlates with the morphological transformation of the procapsid to the mature angularized form.[51-53]

There are 3 possible outcomes of the proposed maturation pathway, but only one that leads to productive nucleocapsid assembly. (i) Type A capsids arise when DNA is not inserted or retained (i.e., packaging is incomplete or aborted) but the internal scaffold is expelled. (ii) Type B capsids result when DNA is not inserted and the internal scaffold is trapped within the angularized capsid shell, or (iii) C capsids arise when the packaging machinery is properly engaged, genomic DNA is inserted, the internal scaffold is expelled, and the capsid shell undergoes the conformational change, sealing the DNA inside. This model is supported by the observations that procapsids spontaneously form in vitro upon mixture of capsid components produced in heterologous expression systems,[54] and that such procapsids transform into type B capsids over time.[33]

The model is appealing because it incorporates the procapsid into the assembly pathway and explains the origin of all capsid types from a single reaction. Because types A and B capsids are dead ends in this model, the remarkable implication is that the packaging reaction must be tightly coordinated with capsid maturation to favor production of type C capsids over types A and B capsids. It might therefore be envisioned that in order to prevent premature angularization and sealing of the capsid, the viral protease packaged within the internal capsid shell must somehow be restrained from cleavage of the internal scaffold until the precise time that DNA and the packaging machinery engage the portal vertex.

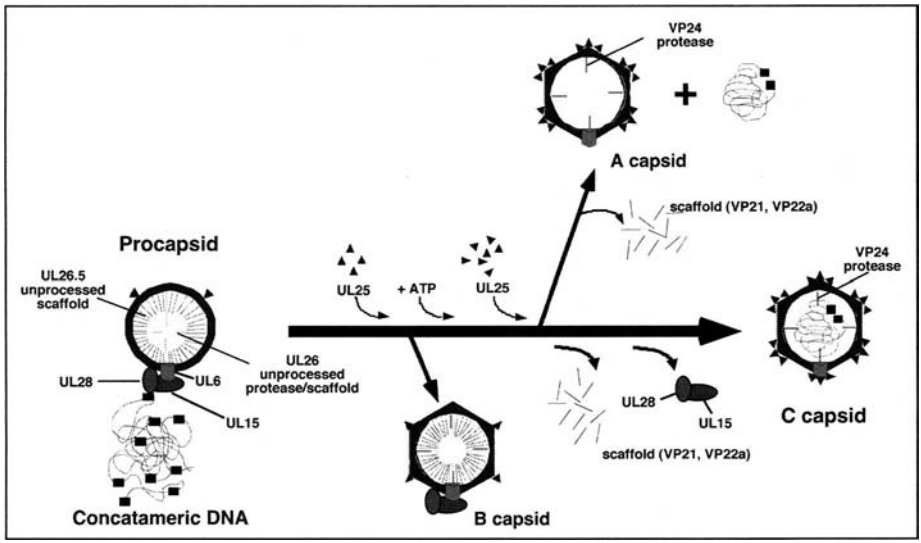


Figure 4. Model of HSV-1 DNA cleavage and packaging and capsid maturation. Based on published reports on capsid assembly, scaffold, and DNA packaging proteins associated with wild-type and mutant virus capsids, we propose a model for the maturation of capsids in herpesvirus infected cells (described in the text). The unprocessed form of U_L26 protease/scaffold protein is depicted in blue, and the $U_L26.5$ scaffold is depicted in red. The U_L6 protein is depicted as a gray cylinder and the putative two subunit terminase composed of U_L15 and U_L28 is depicted as green and purple ovals. Note, U_L33 has recently been shown to associate with U_L15 and U_L28 and may be a subunit of the teminase. U_L25 is depicted as a green triangle. Note that although U_L25 is drawn as being located on the outside of the capsid, it is possible that some or all of U_L25 is actually located inside the capsid. This figure is adapted from the model presented in Dr. Amy Sheaffer's thesis (University of Connecticut Health Center). A color version of this figure is available online at http://www.Eurekah.com.

Components of the Packaging Machinery

The U_L6, U_L15, U_L17, U_L25, U_L28, U_L32, and U_L33 genes are not needed for production of type B capsids but are required for packaging of viral DNA.[55-67] Unlike U_L6, U_L15, U_L17, U_L28, U_L32, and U_L33 that are required for DNA cleavage, U_L25 performs a later function and may be required for either optimal cleavage of full length genomes, or stabilization of capsids containing viral DNA (see below).[31,62] The proteins encoded by these cleavage/packaging genes can be categorized based on their respective roles of portal vertex (U_L6), DNA retention (U_L25), terminase (U_L15, U_L28, and possibly U_L33), or proteins that mediate transport of nucleocapsids to engage other components of the packaging machinery (U_L17 and U_L32). Each of these will be treated separately.

The Portal Vertex (U_L6)

It has long been recognized that absence of a functional HSV-1 U_L6 gene precludes cleavage of concatemeric viral DNA.[44,56,65,68,69] A high degree of homology exists between U_L6 homologs encoded by herpesvirus family members indicating that the encoded proteins likely play similar roles in the life cycles of all herpesviruses. pU_L6 is a component of procapsids, types A, B and C capsids, and virions.[53,56] Thus, the capsid association of pU_L6 is unaffected by scaffold loss, capsid angularization or DNA packaging, and strongly suggests that pU_L6 is an integral component of the capsid shell.

The first indication that pU_L6 might encode a docking site for the putative terminase complex came from the observation that pU_L15 failed to associate with type B capsids from cells infected with a U_L6 null virus.[70,71] By analogy with phage systems, a portal protein would be expected to be incorporated into the procapsid and form a unique vertex through which DNA can be packaged into capsids. In phage systems, the portal protein is not essential for capsid assembly, but when present, is thought to initiate assembly of the phage head.[72] Similarly, HSV-1 capsids can form in the absence of pU_L6 in cells infected with a U_L6 deletion mutant and in extracts of insect cells containing individual capsid proteins.[56,73]

Recent studies using immuno-electron microscopy with intact capsids have demonstrated that epitopes present in the C-terminus of pU_L6 are available for reaction with antibody at the external surface of a pentameric capsid vertex.[74] In general, only one such vertex was labeled whereas capsids from cells infected with a U_L6 null mutant did not exhibit significant immunolabeling. In the same study, estimates from immunoblots probed with pU_L6-specific antibody indicated approximately 12-17 pU_L6 molecules per capsid, consistent with the notion that the portal might consist of 12 molecules of pU_L6, much like analogous structures found in phage heads (see chapters 6 and 7 of this text). On the other hand, a recent report estimated that there were approximately 44 pU_L6 molecules per capsid.[75] Although this apparent discrepancy will have to be studied further, the immunoelectron microscopy experiment has provided convincing evidence that HSV-1 capsids contain a unique pU_L6-containing vertex. This is the first demonstration of a unique capsid vertex in any mammalian virus.

To further test the idea that pU_L6 is a portal protein, the properties of purified pU_L6 in solution were examined.[74] If pU_L6 actually forms a portal, it would be expected to have a structure consistent with a channel through which DNA could be translocated into the capsid. Purified soluble pU_L6 expressed in recombinant baculovirus infected insect cells was examined by electron microscopy and revealed a uniform population of rings with internal diameter of approximately 5 nm and external diameter 16 nm. The masses of individual pU_L6 rings were measured by scanning transmission electron microscopy and found to correspond to an oligomeric state of 12. By comparison, bacteriophage portal complexes have also been shown to consist of rings with an oligomeric state of 12 or 13.[72,76] The X-ray structure of the ϕ-29 portal complex has recently been solved and models based on the structure suggest that the 12 subunit portal ring turns within the capsid driving ATP-dependent translocation of DNA into the capsid[77] (Chapter 7 of this volume). The overall size and structure of HSV-1 pU_L6 resembles the portal complexes of DNA bacteriophages. For instance the outside diameters of phage connectors has been reported in the range of 14.5 nm - 17.5 nm, whereas the inside diameter is in the 2.5 nm - 4.5 nm range.[72,76]

Lateral views of the pU_L6-containing ring-like structures suggest a goblet-like shape with a globular shape at one end, and a stem and base at the other. Such a structure might explain previous failures to observe a unique vertex in HSV-1 capsids if the head portion of the goblet is lodged deep within the capsid shell, and the relatively smaller masses of stem and base are on the inside, obscured from detection by cryoelectron microscopy. The apparent goblet shape of the pU_L6 rings may also be consistent with the presence of three domains: a capsid interaction domain which may consist of the wide portion of the goblet, a long stem which may correspond to the DNA transport channel, and a base which may correspond to the exposed domain. The exposed domain may correspond to the region of pU_L6 that interacts with the terminase.

U_L25 Protein

U_L25 encodes a minor capsid component and it has been estimated that approximately 40 copies are present within each type B capsid.[62,75] There appears to be an inverse correlation between the amount of pU_L25 and the levels of scaffold protein processing within capsids.

Thus, very little pU_L25 is present on procapsids containing an unprocessed scaffold, more is seen on type B capsids, and the largest amounts of pU_L25 are seen on type C capsids that lack scaffold proteins.[53] The pseudorabies virus homologue of pU_L25 was also present at higher levels in types A- and C-capsids relative to that in type B-capsids.[78] It is therefore possible that scaffold loss may enable high levels of pU_L25 to associate with capsids.

U_L25 protein performs functions that are likely to be critical for late stages of DNA cleavage and packaging. Initial studies focused on two *ts* mutants with lesions in the U_L25 gene. One mutant (ts1204) was defective in initial steps of infection, whereas both ts1204 and another mutant (ts1208) failed to produce DNA-containing type C capsids at the nonpermissive temperature.[61] Characterization of a U_L25 null mutant indicated that pU_L25 was not required for DNA synthesis but virtually all of the DNA remained DNase sensitive, indicating that little was successfully packaged.[31,62] Unlike the case in all other cleavage/packaging mutants analyzed to date, genomic L-termini are correctly generated in the absence of pU_L25, but considerably fewer S-termini are detectable.[31] This observation and the observation that type A capsids accumulate to abnormally high levels in cells infected with the U_L25 null mutant suggest that in the absence of pU_L25, the packaging reaction is often aborted between the first and second DNA cleavage events. pU_L25 is less critical for the second DNA cleavage event if the genome to be packaged is shorter than normal.[31] Taken together, the results suggest that pU_L25 performs a function critical for packaging full length genomes, and in the absence of the protein, genomic DNA with a single free end is often expelled before the second cleavage event can occur. Thus, pU_L25 is critical for retention of DNA within the capsid when the potential energy for DNA exit is high, such as is presumed to be the case when large amounts of DNA are tightly packed within the capsid.[79]

How pU_L25 performs such a function is not known. It might be envisioned that pU_L25 comprises part of the packaging motor and contributes to the torque generating capacity thereof. Thus, in the absence of pU_L25, the motor cannot overcome the propensity of the DNA to exit the capsid, and the DNA slips out. Alternatively, and consistent with the addition of U_L25 protein after commencement of packaging, pU_L25 may act to anchor the DNA within the capsid. Thus, as DNA is inserted, pU_L25 is added to stabilize the packaged DNA, neutralizing the propensity of the DNA to exit.[62] Another possibility is that pU_L25 acts to trap packaged DNA within the capsid either by altering the conformation of DNA containing capsids to stabilize them, or to physically plug hole(s) through which DNA might escape. Such functions are not unprecedented inasmuch as some bacteriophage proteins serve to stabilize newly packaged DNA within capsids (reviewed in ref. 80). For example, pU_L25 may be analogous to the lambda phage protein gpW, which is required for the stabilization of DNA within the phage head.[81]

The Putative Terminase (U_L15, U_L28, U_L33)

It is logical to presume that herpesviruses encode a multisubunit functional homolog of bacteriophage terminases that specifically recognizes genomic ends within the DNA, links the DNA to the capsid, and mediates the packaging of DNA through the hydrolysis of ATP. Although the absence of an in vitro packaging system in HSV precludes definitive proof, a complex of the U_L15 and U_L28 gene products is the most likely to fulfill the role of terminase at the present time. (The U_L28 protein was originally referred to as ICP18.5[82]). If the U_L15 and U_L28 proteins were subunits of the HSV terminase it would be expected that they would interact with one another, transiently associate with capsids, hydrolyze ATP, and specifically bind and cleave *Pac* DNA sequences. Although evidence is indirect in some cases, most of these activities have been associated with the products of U_L15 and U_L28 and their homologs in other herpesviruses.

Terminase Subunits

Indirect evidence that the U_L28 and U_L15 proteins interact has come from the study of antiviral drugs such as 2-bromo-5,6,-dichloro-1-β-D-ribofuranosyl benzimidazole (BDCRB) that are directed against the pU_L15 and pU_L28 homologs of human cytomegalovirus (HCMV).[2,3] Specifically, the observation that mutations in both the pU_L28 and pU_L15 homologs promote increased resistance to BDCRB suggests that the proteins interact in some fashion. Other evidence supporting an interaction include the observations that (i) small amounts of the HSV-1 U_L15 and U_L28 proteins can be partially copurified from lysates of infected cells as a 1:1 heterodimer by ion exchange and DNA affinity chromatography,[83](ii) under some circumstances, the U_L28 protein requires the U_L15 protein for import into the nucleus,[83,84] and (iii) the proteins can be coimmunoprecipitated from lysates of HSV infected cells and insect cells containing pU_L15 and pU_L28.[83-85]

Recently, it has been shown that the U_L33 gene product of HSV also coimmunoprecipitates with pU_L15 and pU_L28, suggesting that pU_L33 interacts transiently with pU_L15 and pU_L28, or may act as a third subunit of the terminase.[85] Additional studies will be necessary to determine the role of pU_L33 in DNA cleavage and packaging.

DNA Binding Activity

Purified U_L28 protein can bind *Pac 1* sequences in vitro with a Kd of approximately 5 nM.[86] The aberrant electrophoretic migration of the bound substrates suggest that U_L28 protein does not recognize Watson-Crick base paired duplex DNA, but requires the presence of single stranded regions for optimal binding. Such regions are available for binding in various DNA oligomers bearing *Pac 1* sequences, suggesting that *Pac 1* has a propensity to form such structures. The U_L28 protein can also bind structured DNA induced by the *Pac 1* of HCMV with only slightly lower affinity. Indirect evidence that recognition of *Pac 1* structured DNA is relevant in vivo comes from the observations that (i) point mutations within the C-rich element of *Pac 1* that eliminate cleavage/packaging of HCMV DNA preclude formation of the structures recognized by HSV U_L28 protein and U_L28 protein/DNA interaction,[86] (ii) *DR2* sequences adjacent to *Pac 1* can adopt novel nonB DNA structures with deformed regions containing single strands,[87] implying that such regions may exert torsional stresses that expose single strands within *Pac 1*, and (iii) the *Pac* sequences of human herpesvirus 6 contain areas of S1 nuclease hypersensitivity suggesting the presence of single stranded regions.[88] The HCMV homolog of pU_L28 (HCMV pU_L56) was reported to bind the *pac* motif and may have specific nuclease activity; however, these studies were not confirmed with purified protein.[89]

The recent observation that *Pac 1* sequences are required for correct generation of S-termini indicate that *Pac 1* is specifically required for the second DNA cleavage event.[90] Thus, it is possible that high affinity binding of *Pac 1* DNA by pU_L28 occurs late in the packaging reaction when torsional stress causes the extrusion of single stranded regions of *Pac 1*. Presumably, this would normally focus the second DNA cleavage within *DR1* to release the packaged genomic DNA from the concatemer. Such a model might explain the longstanding observation that several a sequences in correct orientation are not cleaved until a certain amount of DNA is packaged,[15,91] i.e., when a certain amount of torsional stress is applied to the packaged DNA to extrude the single strands recognized by U_L28 protein.

ATPase Activity

U_L15 was initially suggested to encode part of the terminase based on sequence similarities with the T4 terminase.[10,92] The sequence similarities are especially strong in the region of the ATP binding motif, and genetic studies have shown that the putative ATP binding site of pU_L15 is essential for its function.[93] ATP hydrolysis is also required for packaging DNA in vivo.[94]

Transient Capsid Association

The U_L15 and U_L28 gene products are associated with procapsids and type B capsids.[53,70,71,95] Lesser amounts of pU_L15 and pU_L28 associate with type C capsids suggesting that the proteins are excluded from the capsid once DNA is packaged (see Fig. 4). The location of pU_L15 and pU_L28 within capsids is not known and the mechanism of release from capsids is not understood. It is possible that the proteins engage DNA, bind to the portal vertex, and once DNA is packaged, release the packaged end of the DNA and disengage from the capsid. However, it is equally plausible that pU_L15 and pU_L28 lie within the capsid in association with the internal scaffold and are expelled along with scaffold proteins during capsid angularization.

Epitopes encoded by the first 35 amino acids of pU_L15 are proteolytically removed around the time of the first DNA cleavage event.[96] Analysis of this series of amino acids suggests the presence of a highly charged alpha helix with a propensity to form a coiled-coil that might mediate interaction with other proteins.[10] Removal of such a sequence could release the protein from capsids or other proteins and thus provide an elegant method to regulate terminase activity. Release of pU_L15 also requires a function encoded by U_L25 inasmuch as pU_L15 associates with capsids of the U_L25 null mutant.[71,96] The removal of pU_L15 from capsids may also require insertion of viral DNA into the capsid and concurrent expulsion of the inner scaffold.

Intranuclear Transport Proteins (U_L17, U_L32)

During infection with Herpes simplex virus 1, gene expression and DNA replication occur within globular intranuclear compartments termed replication compartments.[99] At least four cleavage and packaging proteins, U_L6, U_L15, U_L32 and U_L33, accumulate in replication compartments at relatively early times after infection (6-8 hours).[93,100-102] Moreover, capsids have also been observed to colocalize completely within replication compartments at these times.[101] These results have led to the suggestion that DNA packaging occurs in replication compartments; however, the mechanism by which capsids are directed to these compartments is not understood. Capsid transport is apparently a regulated function inasmuch as aggregates of procapsids accumulate in cells infected with viruses containing temperature sensitive mutations within the viral protease.[44] Upon shift to the permissive temperature, the procapsids mature and the resulting type A, B and C capsids redistribute away from regions containing capsid aggregates and accumulate more diffusely in replication compartments.[51] Recent evidence has suggested that the U_L17 and U_L32 gene products are necessary for proper targeting of capsids to replication compartments.[101,103]

The virion tegument is delimited by the inside of the viral envelope and the outer surface of nucleocapsids. The U_L17 proteins are readily detected in immunoblots of the capsid/tegument fraction of virions, and within light particles that consist of enveloped tegument proteins but lack capsid proteins.[60] These observations suggest that the U_L17 proteins are associated with the virion tegument but it is also possible that the proteins associate with nucleocapsids.[104] In cells infected with wild type viruses, most type B capsids accumulate in a diffuse pattern within replication compartments.[105] In contrast, capsids accumulate exclusively within intranuclear aggregates separate from replication compartments in cells infected with a virus lacking U_L17.[60,103] These observations suggest that U_L17 is involved in normal pathways that serve either to (i) directly transport capsids through the nucleus or (ii) produce capsids competent for intranuclear transport.

A capsid transport function has also been proposed for the U_L32 gene of HSV-1.[101] In the absence of U_L32, capsids do not accumulate in replication compartments, nor in punctate regions as in the U_L17 mutant, but in a perinuclear region abutting the nuclear membrane. The U_L32 protein appears to accumulate in both the nucleus and cytoplasm, suggesting that it may play roles in addition to that of viral DNA cleavage/packaging.[101,106]

Table 1. Proteins associated with one or more capsid forms

Gene	Protein	Present In Capsids				Role in HSV Assembly
		A	B	C	Procapsids	
U_L6	pU_L6	+	+	+	+	Portal vertex protein
U_L15	pU_L15	+	+	-	+	Putative terminase subunit
U_L25	pU_L25	+	+	+	+/-	DNA stabilization
U_L28	pU_L28	-	+	-	+	Putative terminase subunit
U_L26	VP24	+	+	+	+	Protease
U_L26	VP21	-	+	-	+	Scaffold
$U_L26.5$	VP22a	-	+	-	+	Scaffold
U_L18	VP23	+	+	+	+	Triplex
U_L19	VP5	+	+	+	+	Major capsid protein
U_L38	VP19c	+	+	+	+	Triplex
U_L35	VP26	?	+	+	?	Hexon tips

References for assignments are provided in the text.

Model of DNA Cleavage and Packaging

A model for the maturation of capsids in herpesvirus-infected cells that is consistent with the current literature is depicted in Figures 2 and 4. The procapsid containing unprocessed scaffold proteins is depicted as the precursor to the angular type A, B, and C capsids found within cells infected with wild type viruses (Fig. 4). As described above, procapsids contain the unprocessed versions of the pU_L26 and $pU_L26.5$ gene products. It is believed that immature capsids are delivered via pU_L17 and pU_L32-mediated transport to intranuclear replication compartments (not shown).

The putative terminase consisting of pU_L15 and pU_L28 is proposed to bind viral DNA via the pU_L28 subunit, and dock with procapsids primarily through an interaction with pU_L6. A model of the DNA cleavage events is shown in Figure 2. The DNA concatemer is scanned by the capsid-associated terminase for the first *Pac 2* in proper orientation. The DNA is cleaved at a fixed distance from *Pac 2*; normally this occurs within the upstream *DR1* leaving a truncated element of 18 bp plus a single-base 3' overhang. In some genomes, this *Pac 2* is immediately followed by additional a sequences. The now generated L-terminus with 18 bp of *DR1* followed by the rest of the a sequence is inserted into the capsid. The cleaved U_S end not bound to the procapsid is presumably degraded (Fig. 2).

It is believed that protease activation and cleavage of the scaffold proteins triggers the DNA cleavage and packaging reactions and, in turn, the angularization of the procapsid to the mature capsid. The products of the protease reaction are VP24 (protease domain), two scaffold proteins (VP21 and VP22a) and a 25 aa C-terminal peptide that serves to anchor the unprocessed forms of pU_L26 and $pU_L26.5$ to the capsid (see Figs. 3,4). VP24 remains in the capsid and VP21 and VP22a are extruded (Fig. 4). The fate of the 25 amino acid peptide is unknown. Concurrently, DNA is threaded into the capsid through a portal vertex comprised of a pU_L6 dodecamer.[74] Energy for DNA insertion is provided partly through hydrolysis of ATP mediated by pU_L15 within the capsid-associated terminase. U_L DNA is inserted which is followed by the L-S junction and U_S.

The addition of U_L25 protein to the maturing capsid is proposed to occur in response and proportion to the release of the internal scaffold proteins (ref. 52). Although Figure 4 depicts

pU_L25 on the surface of the capsid, it is possible that some or all of pU_L25 is located inside the capsid. Whether pU_L25 acts to anchor DNA within the capsid or to stabilize DNA-containing capsids is currently unknown.

Near the end of packaging, single strands of *Pac 1* in the correct orientation in *Ub* are extruded and recognized by pU_L28 within the procapsid-associated pU_L15/pU_L28 complex, and a second cleavage occurs leaving the packaged U_S terminus with a single a sequence followed by a single bp of *DR1* (see Fig. 2). The protein or protein complex that actually cleaves the DNA is unknown. The second cleavage event and possibly functions of pU_L25 help mediate release of the putative terminase from the capsid. An efficient use of resources would include initiation of further rounds of packaging as the truncated *DR1* on the free U_L end of the concatemer engages newly delivered procapsids.

This model proposes several testable hypotheses. It is anticipated that investigation of these will not only reveal the workings of a most extraordinarily complex and fascinating molecular machine, but also provide novel leads for future developments in antiviral therapy.

Acknowledgements

We would like to thank the past and current members of our laboratories, and especially Amy Sheaffer for the use of the capsid maturation diagram. These studies were supported by public health service grants GM50740 and AI37549 from the National Institutes of Health.

References

1. van Zijl M, Fairhurst J, Jones TR et al. Novel class of thiourea compounds that inhibit herpes simplex virus type 1 DNA cleavage and encapsidation: Resistance maps to the UL6 Gene. J Virol 2000; 74:9054-9061.
2. Krosky PM, Underwood MR, Turk SR et al. Resistance of human cytomegalovirus to benzimidazole ribonucleosides maps to two open reading frames: UL89 and UL56. J Virol 1998; 72:4721-4728.
3. Underwood MR, Harvey RJ, Stanat SC et al. Inhibition of a human cytomegalovirus DNA maturation by a benzimidizole ribonucleoside is mediated through the UL89 gene product. J Virol 1998; 72:717-725.
4. Weller SK. Herpes simplex virus DNA replication and genome maturation. In: Cooper GM, Temin BS, Temin R, eds. Implications of the DNA provirus: Howard Temin's Scientific Legacy. Washington, DC: ASM Press,1995:189-213.
5. Lehman IR, Boehmer PE. Replication of herpes simplex virus DNA. J Biol Chem 1999; 274(40):28059-28062.
6. Homa FL, Brown JC. Capsid assembly and DNA packaging in herpes simplex virus. Reviews in Medical Virology 1997; 7:107-122.
7. Sheldrick P, Berthelot N. Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harbor Symp Quant Biol 1974; 39:667-678.
8. Hayward GS, Jacob RJ, Wadsworth SC et al. Anatomy of herpes simplex virus DNA: Evidence for four populations of molecules that differ in the relative orientations of their long and short segments. Proc Natl Acad Sci USA 1975; 72:4243-4247.
9. Delius H, Clements JB. A partial denaturation map of herpes simplex virus type 1 DNA: Evidence for inversions of the unique DNA regions. J Gen Virol 1976; 33:125-133.
10. McGeoch DJ, Dalrymple MA, Davison AJ et al. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J Gen Virol 1988; 69:1531-1574.
11. Wadsworth S, Jacob RJ, Roizman B. Anatomy of herpes simplex virus DNA. II. Size, composition and arrangement of inverted terminal repetitions. J Virol 1975; 15:1487-1497.
12. Wagner MJ, Summers WC. Structures of the joint region and the termini of the DNA of herpes simplex virus type 1. J Virol 1978; 27:374-387.
13. Mocarski ES, Roizman B. Structure and role of the herpes simplex virus DNA termini in inversion, circularization and generation of virion DNA. Cell 1982; 31:89-97.

14. Stow ND, McMonagle EC. Characterization of the TRs/IRs origin of DNA replication of herpes simplex virus type 1. Virology 1983; 130:427-438.
15. Spaete RR, Frenkel N. The herpes simplex virus amplicon: A new eucaryotic defective-virus cloning-amplifying vector. Cell 1982; 30:295-304.
16. Stow ND, McMonagle EC, Davison AJ. Fragments from both termini of the herpes simplex virus type 1 genome contain signals required for the encapsidation of viral DNA. Nucleic Acids Res 1983; 11:8205-8220.
17. Vlazny DA, Kwong A, Frenkel N. Site-specific cleavage/packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full length viral DNA. Proc Natl Acad Sci USA 1982; 79:1423-1427.
18. Davison AJ, Wilkie NM. Nucleotide sequence of the joint between the L and S segments of herpes simplex virus types 1 and 2. J Gen Virol 1981; 55:315-331.
19. Mocarski ES, Roizman B. Site-specific inversion sequence of the herpes simplex virus genome: Domain and structural features. Proc Natl Acad Sci USA 1981; 78:7047-7051.
20. Deiss LP, Chou J, Frenkel N. Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA. J Virol 1986; 59:605-618.
21. Nasseri M, Mocarski ES. The cleavage recognition signal is contained within sequences surrounding an a-a junction in herpes simplex virus DNA. Virology 1988; 167:25-30.
22. Deiss LP, Frenkel N. Herpes simplex virus amplicon: Cleavage of concatemeric DNA is linked to packaging and involves amplification of the terminally reiterated a sequence. J Virol 1986; 57:933-941.
23. Varmuza SL, Smiley JR. Signals for site-specific cleavage of HSV DNA: Maturation involves two separate cleavage events at sites distal to the recognition sequences. Cell 1985; 41:793-802.
24. Mocarski ES, Deiss LP, Frenkel N. The nucleotide sequence and structural features of a novel U_S-a junction present in a defective herpes simplex virus genome. J Virol 1985; 55:140-146.
25. McVoy MA, Nixon DA, Adler SP et al. Sequences within the herpesvirus-conserved pac1 and pac2 motifs are required for cleavage and packaging of the murine cytomegalovirus genome. J Virol 1998; 72:48-56.
26. Poffenberger KL, Roizman B. A noninverting genome of a viable herpes simplex virus 1: Presence of head-to-tail linkages in packaged genomes and requirements for circularization after infection. J Virol 1985; 53:587-595.
27. Severini A, Morgan AR, Tovell DR et al. Study of the structure of replicative intermediates of HSV-1 DNA by pulsed-field gel electrophoresis. Virology 1994; 200(2):428-435.
28. Zhang X, Efstathiou S, Simmons A. Identification of novel herpes simplex virus replicative intermediates by field inversion gel electrophoresis: Implications for viral DNA amplification strategies. Virology 1994; 202:530-539.
29. Martinez R, Sarisky RT, Weber PC et al. Herpes simplex virus type 1 alkaline nuclease is required for efficient processing of viral DNA replication intermediates. J Virol 1996; 70(4):2075-2085.
30. McVoy MA, Nixon DE, Hur JK et al. The ends on herpesvirus DNA replicative concatemers contain pac2 cis cleavage/packaging elements and their formation is controlled by terminal cis sequences. J Virol 2000; 74(3):1587-1592.
31. Stow ND. Packaging of genomic and amplicon dna by the herpes simplex virus type 1 ul25-null mutant kul25ns. J Virol 2001; 75(22):10755-10765.
32. Desai P, DeLuca NA, Glorioso JC et al. Mutations in herpes simplex virus type 1 genes encoding VP5 and VP23 abrogate capsid formation and cleavage of replicated DNA. J Virol 1993; 67:1357-1364.
33. Newcomb WW, Homa FL, Thomsen DR et al. Assembly of the herpes simplex virus capsid: Characterization of intermediates observed during cell-free capsid formation. J Mol Biol 1996; 263:432-446.
34. Gibson W, Roizman B. Proteins specified by herpes simplex virus VIII. Characterization and composition of multiple capsid forms of subtypes 1 and 2. J Virol 1972; 10:1044-1052.
35. Trus BL, Booy FP, Newcomb WW et al. The herpes simplex virus procapsid: Structure, conformational changes upon maturation, and roles of the triplex proteins VP19C and VP23 in assembly. J Mol Biol 1996; 263:447-462.

36. Booy FP, Newcomb WW, Trus BL et al. Liquid crystalline, phage-like packaging of encapsidated DNA in herpes simplex virus. Cell 1991; 64:1007-1015.
37. Schrag JD, Prasad BV, Rixon FJ et al. Three-dimensional structure of the HSV1 nucleocapsid. Cell 1989; 56:651-660.
38. Zhou ZH, Prasad BV, Jakana J et al. Protein subunit structures in herpes simplex virus A-capsid determined from 400 kV spot-scan electron cryomicroscopy. J Mol Biol 1994; 242:456-469.
39. Liu F, Roizman B. The promoter, transcriptional unit, and coding sequences of herpes simplex virus 1 family 35 proteins are contained within and in frame with the UL26 open reading frame. J Virol 1991; 65:206-212.
40. Liu F, Roizman B. The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. J Virol 1991; 65:5149-5156.
41. Davison MD, Rixon FJ, Davison AJ. Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes simplex type 1. J Gen Virol 1992; 73:2709-2713.
42. Person S, Laquerre S, Desai P et al. Herpes simplex virus type 1 capsid protein, VP21, originates within the UL26 open reading frame. J Gen Virol 1993; 74:2269-2273.
43. Rixon FJ, Cross Am, Addison C et al. The products of herpes simplex virus type 1 gene UL26 which are involved in DNA packaging are strongly associated with empty but not with full capsids. J Gen Virol 1988; 69:2879-2891.
44. Preston VG, Coates AM, Rixon FJ. Identification and characterization of a herpes simplex virus gene product required for encapsidation of viral DNA. J Virol 1983; 45:1056-1064.
45. Thomsen DR, Newcomb WW, Brown JC et al. Assembly of the herpes simplex virus capsid: Requirement for the carboxy-terminal twenty-five amino acids of the proteins encoded by the UL26 and UL26.5 genes. J Virol 1995; 69:3690-3703.
46. Hong Z, Beaudet-Miller M, Durkin J et al. Identification of a minimal hydrophobic domain in the herpes simplex virus type 1 scaffolding protein which is required for interaction with the major capsid protein. J Virol 1996; 70(1):533-540.
47. Holland LE, Anderson KP, Stringer JR et al. Isolation and localization of herpes simplex virus type 1 mRNA abundant before viral DNA synthesis. J Virol 1979; 31:447-462.
48. Desai P, Person S. Second site mutations in the N-terminus of the major capsid protein (VP5) overcome a block at the maturation cleavage site of the capsid scaffold proteins of herpes simplex virus type 1. Virology 1999; 261(2):357-366.
49. Warner SC, Desai P, Person S. Second-site mutations encoding residues 34 and 78 of the major capsid protein (VP5) of herpes simplex virus type 1 are important for overcoming a blocked maturation cleavage site of the capsid scaffold proteins. Virology 2000; 278(1):217-226.
50. Weinheimer SP, McCann III PJ, O'Boyle II DR et al. Autoproteolysis of herpes simplex virus type 1 protease releases an active catalytic domain found in intermediate capsid particles. J Virol 1993; 67:5813-5822.
51. Wilson DW, Church GA. Study of herpes simplex virus maturation during a synchronous wave of assembly. J Virol 1997; 71:3603-3612.
52. Sheaffer AK, Newcomb WW, Brown JC et al. Evidence for controlled incorporation of herpes simplex virus type 1 UL26 protease into capsids. J Virol 2000; 74(15):6838-6848.
53. Sheaffer AK, Newcomb WW, Gao M et al. Herpes simplex virus DNA cleavage and packaging proteins associate with the procapsid prior to its maturation. J Virol 2001; 75(2):687-698.
54. Newcomb WW, Homa FL, Thomsen DR et al. Cell-free assembly of the herpes simplex virus capsid. J Virol 1994; 68:6059-6063.
55. Weller SK, Carmichael EP, Aschman DP et al. Genetic and phenotypic characterization of mutants in four essential genes that map to the left half of HSV-1 U_L DNA. Virology 1987; 161:198-210.
56. Patel AH, Rixon FJ, Cunningham C et al. Isolation and characterization of herpes simplex virus type 1 mutants defective in the UL6 gene. Virology 1996; 217:111-123.
57. Poon APW, Roizman B. Characterization of a temperature-sensitive mutant of the U_L15 open reading frame of herpes simplex virus 1. J Virol 1993; 67:4497-4503.
58. Baines JD, Cunningham C, Nalwanga D et al. The U_L15 gene of herpes simplex virus type 1 contains within its second exon a novel open reading frame that is translated in frame with the U_L15 gene product. J Virol 1997; 71:2666-2673.

59. Yu D, Shaeffer AK, Tenny DJ et al. Characterization of ICP6:Lac Z insertion mutants of the UL15 gene of herpes simplex virus type 1 reveals the translation of two proteins. J Virol 1997; 71:2656-2665.
60. Salmon B, Cunningham C, Davison AJ et al. The herpes simplex virus 1 U_L17 gene encodes virion tegument proteins that are required for cleavage and packaging of viral DNA. J Virol 1998; 72:3779-3788.
61. Addison C, Rixon FJ, Palfreyman JW et al. Characterisation of a herpes simplex virus type 1 mutant which has a temperature-sensitive defect in penetration of cells and assembly of capsids. Virology 1984; 138:246-259.
62. McNab AR, Desai P, Person S et al. The product of the herpes simplex virus type 1 UL25 gene is required for encapsidation but not for cleavage of replicated DNA. J Virol 1998; 72:1060-1070.
63. Tengelsen LA, Pedersen NE, Shaver PR et al. Herpes simplex virus type 1 DNA cleavage and capsidation require the product of the UL28 gene: Isolation and characterization of two UL28 deletion mutants. J Virol 1993; 67:3470-3480.
64. Addison C, Rixon FJ, Preston VG. Herpes simplex virus type 1 UL28 gene product is important for the formation of mature capsids. J Gen Virol 1990; 71:2377-2384.
65. Sherman G, Bachenheimer SL. Characterization of intranuclear capsids made by ts morphogenetic mutants of HSV-1. Virology 1988; 163:471-480.
66. Weller SK, Aschman DP, Sacks WR et al. Genetic analysis of temperature-sensitive mutants of HSV-1: The combined use of complementation and physical mapping for cistron assignment. Virology 1983; 130:290-305.
67. Al-Kobaisi MF, Rixon FJ, McDougall I et al. The herpes simplex virus UL33 gene product is required for the assembly of full capsids. Virology 1991; 180:380-388.
68. Sherman G, Bachenheimer SL. DNA processing in temperature-sensitive morphogenetic mutants of HSV-1. Virology 1987; 158:427-430.
69. Lamberti C, Weller SK. The herpes simplex virus type 1 UL6 protein is essential for cleavage and packaging but not for genomic inversion. Virology 1996; 226:403-407.
70. Salmon B, Baines JD. Herpes simplex virus DNA cleavage and packaging: Association of multiple forms of U_L15-encoded proteins with B capsids requires at least the U_L6, U_L17 and U_L28 genes. J Virol 1998; 72:3045-3050.
71. Yu D, Weller SK. Herpes simplex virus type 1 cleavage and packaging proteins UL15 and UL28 are associated with B but not C capsids during packaging. J Virol 1998; 72:7428-7439.
72. Bazinet C, King J. The DNA translocating vertex of dsDNA bacteriophage. Annu Rev Microbiol 1985; 39:109-129.
73. Thomsen DR, Roof LL, Homa FL. Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. J Virol 1994; 68:2442-2457.
74. Newcomb WW, Juhas RM, Thomsen DR et al. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. J Virol 2001; 75(22):10923-10932.
75. Ogasawara M, Suzutani T, Yoshida I et al. Role of the UL25 gene product in packaging DNA into the herpes simplex virus capsid: Location of UL25 product in the capsid and demonstration that it binds DNA. J Virol 2001; 75:1427-1436.
76. Valpuesta JM, Fernandez JJ, Carazo JM et al. The three-dimensional structure of a DNA translocating machine at 10 A resolution. Structure Fold Des 1999; 7(3):289-296.
77. Simpson AA, Tao Y, Leiman PG et al. Structure of the bacteriophage phi29 DNA packaging motor. Nature 2000; 408(6813):745-750.
78. Kaelin K, Dezelee S, Masse MJ et al. The UL25 protein of pseudorabies virus associates with capsids and localizes to the nucleus and to microtubules. J Virol 2000; 74(1):474-482.
79. Smith DE, Tans SJ, Smith SB et al. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. Nature 2001; 413(6857):748-752.
80. Black LW. DNA packaging in dsDNA bacteriophages. Annu Rev Microbiol 1989; 43:267-292.
81. Perucchetti R, Parris W, Becker A et al. Late stages in bacteriophage lambda head morphogenesis: In vitro studies on the action of the bacteriophage lambda D-gene and W-gene products. Virology 1988; 165(1):103-114.

82. Pellett PE, Jenkins FJ, Ackermann M et al. Transcription initiation and nucleotide sequence of a herpes simplex virus 1 gene conserved in the Epstein-Barr virus genome and reported to affect the transport of viral glycoproteins. J Virol 1986; 60:1134-1140.
83. Koslowski KM, Shaver PR, Casey JTI et al. Physical and functional interactions between the herpes simplex virus UL15 and UL28 DA cleavage and packaging proteins. J Virol 1999; 73:1704-1707.
84. Abbotts AP, Preston VG, Hughes M et al. Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. J Gen Virol 2000; 81 Pt 12:2999-3009.
85. Beard PM, Taus NS, Baines JD. The DNA cleavage and packaging proteins encoded by genes U_L28, U_L15,and U_L33 of herpes simplex virus 1 form a complex in infected cells. J Virol 2002; 76:4785-4791.
86. Adelman K, Salmon B, Baines JD. Herpes simplex DNA packaging sequences adopt novel structures that are specifically recognized by a component of the cleavage and packaging machinery. Proc Natl Acad Sci USA 2001; 98:3086-3091.
87. Wohlrab F, McLean MJ, Wells RD. The segment inversion site of herpes simplex virus type 1 adopts a novel DNA structure. J Biol Chem 1987; 262:6407-6416.
88. Deng H, Dewhurst S. Functional identification and analysis of cis-acting sequences which mediate genome cleavage and packaging in human herpesvirus 6. J Virol 1998; 72:320-329.
89. Bogner E, Radsak K, Stinski MF. The gene product of human cytomegalovirus open reading frame UL56 binds the pac motif and has specific nuclease activity. J Virol 1998; 72:2259-2264.
90. Hodge PD, Stow ND. Effects of mutations within the herpes simplex virus type 1 DNA encapsidation signal on packaging efficiency. J Virol 2001; 75(19):8977-8986.
91. Vlazny DA, Frenkel N. Replication of herpes simplex virus DNA: Location of replication recognition signals within defective virus genomes. Proc Natl Acad Sci USA 1981; 78:742-746.
92. Davison AJ. Channel catfish virus: A new type of herpesvirus. Virology 1992; 186:9-14.
93. Yu D, Weller SK. Genetic analysis of the UL 15 gene locus for the putative terminase of herpes simplex virus type 1. Virology 1998; 243:32-44.
94. Chi JHI, Wilson DW. ATP-dependent localization of the herpes simplex virus capsid protein VP26 to sites of procapsid maturation. J Virol 2000; 74:1468-1475.
95. Taus NS, Baines JD. Herpes simplex virus DNA cleavage and packaging: The U_L28 gene product is a minor component of B capsids. Virology 1998; 252:443-449.
96. Salmon B, Nalwanga D, Fan Y et al. Proteolytic cleavage of the amino-terminus of the U_L15 gene product of herpes simplex virus 1 U_L15 protein is coupled with maturation of viral DNA into unit length genomes. J Virol 1999; 73:8338-8348.
97. Catalano C, Feiss M. Virus DNA packaging: The strategy used by phage??. Molecular Microbiology 1995; 16:1075-1086.
98. Catalano CE, Tomka MA. Role of gpFI protein in DNA packaging by bacteriophage lambda. Biochemistry 1995; 34(31):10036-10042.
99. Quinlan MP, Chen LB, Knipe DM. The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. Cell 1984; 36:857-868.
100. Reynolds AE, Fan Y, Baines JD. Characterization of the U_L33 gene product of herpes simplex virus 1. Virology 2000; 266:310-318.
101. Lamberti C, Weller SK. The herpes simplex virus type 1 cleavage/packaging protein, UL32, is involved in efficient localization of capsids to replication compartments. J Virol 1998; 72:2463-2473.
102. Ward PL, Ogle WO, Roizman B. Assemblons: Nuclear structures defined by aggregation of immature capsids and some tegument proteins of herpes simplex virus 1. J Virol 1996; 70:4623-4631.
103. Taus NS, Salmon B, Baines JD. The herpes simplex virus 1 U_L17 gene is required for localization of capsids and major and minor capsid proteins to intranuclear sites where viral DNA is cleaved and packaged. Virology 1998; 252:115-125.
104. Goshima F. Herpes simplex virus UL17 protein is associated with B capsids and colocalizes with ICP35 and VP5 in infected cells. Arch Virol 2000; 145:417-426.
105. Roizman B, Sears AE. Herpes simplex viruses and their replication. In: Fields BN, Knipe DM, eds. Virology. New York: Raven Press Ltd, 1990:1795-1841.
106. Chang YE, Poon APW, Roizman B. Properties of the protein encoded by the U_L32 open reading frame of herpes simplex virus 1. J Virol 1996; 70:3938-3946.

Index

G

H

I

M

N

O

P

R

S

T

U

V

W