

RNA Synthesis in a Cage—Structural Studies of Reovirus Polymerase $\lambda 3$

Yizhi Tao,¹ Diane L. Farsetta,^{2,5} Max L. Nibert,³ and Stephen C. Harrison^{1,4}

¹Howard Hughes Medical Institute
Department of Molecular and Cellular Biology
Harvard University
Cambridge, Massachusetts 02138

²Cellular and Molecular Biology Program
Institute for Molecular Virology
University of Wisconsin-Madison
Madison, Wisconsin 53706

³Department of Microbiology and Molecular
Genetics
Harvard Medical School
Boston, Massachusetts 02115

Summary

The reovirus polymerase and those of other dsRNA viruses function within the confines of a protein capsid to transcribe the tightly packed dsRNA genome segments. The crystal structure of the reovirus polymerase, $\lambda 3$, determined at 2.5 Å resolution, shows a fingers-palm-thumb core, similar to those of other viral polymerases, surrounded by major N- and C-terminal elaborations, which create a cage-like structure, with four channels leading to the catalytic site. This “caged” polymerase has allowed us to visualize the results of several rounds of RNA polymerization directly in the crystals. A 5′ cap binding site on the surface of $\lambda 3$ suggests a template retention mechanism by which attachment of the 5′ end of the plus-sense strand facilitates insertion of the 3′ end of the minus-sense strand into the template channel.

Introduction

Replicases and transcriptases have structural features that ensure specific initiation, adequate processivity, and high fidelity. Many polymerases must also reinitiate template-dependent synthesis accurately and rapidly. These characteristics demand complex spatial organization. Some simple DNA polymerases, such as the Klenow fragment of *E. coli* DNA polymerase I and its homologs (Doublet et al., 1998; Eom et al., 1996; Kiefer et al., 1998; Ollis et al., 1985), initiate on a polynucleotide primer and close in against the template:primer:nucleoside-triphosphate complex to enforce correct base pairing at the polymerization site by tight steric constraints. Enclosure of the region around the catalytic site also allows the enzyme to detect, by poor fit, a mismatch in the minor groove just upstream of the current polymerization position (Doublet et al., 1999). Processivity can be enhanced by a variety of ad hoc mechanisms—for example, the coopting by T7 DNA polymerase of a thio-

redoxin molecule (Doublet et al., 1998)—that provides a barrier to loss of template:primer. DNA replicases associate with a so-called sliding-clamp assembly (Stukenberg et al., 1991), which encircles a DNA strand when loaded by an ATP-dependent loading enzyme and which produces an extremely processive association. mRNA synthases, such as RNA polymerase II (Gnatt et al., 2001) and *E. coli* DNA-dependent RNA polymerase (Murakami et al., 2002), contain a central polymerase site surrounded by elaborations to guide template and product in and out of the catalytic cleft.

The $\lambda 3$ polymerase of reoviruses (Starnes and Joklik, 1993) and its homologs in other double-strand RNA viruses function within the confined interior of a viral core particle (Gillies et al., 1971; Shatkin and Sipe, 1968; Skehel and Joklik, 1969). For members of the *Reoviridae* family of dsRNA viruses, core assembly requires recruitment of one capped, plus-sense strand for each of the 10–12 genomic segments, along with a roughly equal number of RNA-dependent RNA polymerase complexes (Nibert and Schiff, 2001). Synthesis of minus-sense strands, full complements of the packaged plus strands, accompanies this process, catalyzed by the polymerase molecules (Figure 1A). The mature virion, to which viral outer-capsid proteins have also been added, contains a completely double-stranded genome. Upon reovirus infection of a host cell, the viral core (about 700 Å in diameter and containing about 23,500 bp of dsRNA) is probably released into the cytosol. There, the $\lambda 3$ polymerase molecules switch mode and catalyze fully conservative plus-strand synthesis, using each minus strand as a template (Figure 1A). Nascent plus-strand RNA is capped and methylated, by other core-associated enzymes, and exported into the cytosol. The likely enzyme for the first capping step, removal of the 5′- γ -phosphate, is $\mu 2$ (J. Kim and M.L.N., unpublished), one or two copies of which are probably associated with each $\lambda 3$ (Nibert and Schiff, 2001). The guanylyl and methyl transferase capping enzymes are domains of the core turret protein, $\lambda 2$ (Reinisch et al., 2000). Capped transcripts serve as mRNA, for translation of viral proteins, and also as genomic plus strands for packaging (Nibert and Schiff, 2001).

The structure of the reovirus core shows that the packaged $\lambda 3$ polymerase molecules lie just inside the $\lambda 1$ shell and close to a five-fold position, where an RNA exit pore leads to the $\lambda 2$ capping enzyme (Figure 1B; Dryden et al., 1993; Reinisch et al., 2000). Ten of the twelve five-fold exit positions are then associated with a polymerase, probably with a single genome segment “dedicated” to each; the remaining two exits may be vacant (or at least devoid of a genome segment). This organization suggests that during transcription the template (minus-sense) strand of the genomic RNA is pulled through the polymerase, which remains more or less in place, and that the transcript, as it emerges from the polymerase, is fed into the pore. We believe that the polymerase must stay tethered at the pore, because otherwise the 5′ end of the transcript would be unlikely to find its exit by free diffusion. To reinitiate, the 3′ end of the template

⁴Correspondence: harrison@crystal.harvard.edu

⁵Present address: ETAN Field Office, Social Justice Center, Madison, Wisconsin 53703.

strand must again find a polymerase—a second process unlikely to occur by free diffusion in the tightly packed interior of the particle (Figure 1B). Thus, there is probably a mechanism to hold the 3' end of the template close to the polymerase throughout the transcription cycle.

The atomic structure of the λ 3 polymerase, reported here from crystals of recombinant protein, could not be determined in the crystallographic analysis of the core because the single copy associated with each five-fold is averaged over all five equivalent orientations. The molecule, now seen at 2.5 Å resolution, is a “polymerase in a cage”—a catalytic module with fingers, palm, and thumb subdomains remarkably similar to those in other viral polymerases (Cheetham and Steitz, 2000), surrounded by elaborate additional elements derived from long N- and C-terminal extensions (Figures 1C and 2). Because crystal contacts are confined to the exterior of the cage, we can obtain multiple rounds of polymerization in the crystals and visualize directly the structures of various intermediates in the initiation and polymerization processes. There are channels in the cage corresponding to entry paths for template and for nucleoside triphosphates (NTPs) and to exit paths for minus-strand template or dsRNA product and for plus-strand transcript. A particularly striking result is the detection of an RNA cap binding site on the surface of the cage between the template entrance and exit channels. We propose that by binding the 5' end of the nontemplate strand, the enzyme ensures that it can reinitiate efficiently at the 3' end of the template, fulfilling the suggestion at the end of the preceding paragraph. Other viral polymerases may contain related sites, potential targets for inhibitors that do not affect cellular RNA polymerases.

Results

Preparation and Structure Determination

We expressed and purified λ 3 as described in Experimental Procedures, obtaining approximately 1 mg purified protein from 1 l of insect cell culture. The elution profile from a Superdex-200 size-exclusion column suggested a monomer. The protein crystallized in the absence of substrates in space group P2₁2₁2₁, with one molecule in the asymmetric unit. The structure was determined to 2.5 Å resolution by combining phases from multiple isomorphous replacement (MIR) and multiple wavelength anomalous dispersion (MAD) (Table 1). The excellent experimental map, improved by solvent flipping, allowed unambiguous placement of more than 95% of the residues. The final atomic model after crystallographic refinement contains 1256 residues out of 1267 in total (Figure 1C). The missing part includes a flexible loop between residues 957 and 964, one residue at the N terminus, and two at the C terminus. All complexes of λ 3 and substrates were prepared by soaking native crystals. Their structures were determined by molecular replacement, using the native λ 3 structure as a phasing model.

Overview of the Structure

The λ 3 polypeptide chain folds into a compact unit with a central cavity (Figure 2). Its overall diameter is about

65 Å. Comparison with other viral RNA polymerases (Ago et al., 1999; Bressanelli et al., 1999; Butcher et al., 2001; Hansen et al., 1997; Huang et al., 1998; Lesburg et al., 1999) shows that λ 3 can be divided into three domains: a central polymerase domain, which contains “fingers,” “palm,” and “thumb” subdomains, an N-terminal domain, which bridges fingers and thumb on one side of the catalytic cleft, and a C-terminal “bracelet” domain, which covers the catalytic cleft on the other side (Figure 2A). The N- and C-terminal domains create a cage that encloses the conserved catalytic region within a largely hollow center.

The “right-hand” configuration of the polymerase domain closely resembles those of HIV-1 reverse transcriptase (RT) (Huang et al., 1998; Kohlstaedt et al., 1992) and the RNA-dependent RNA polymerases of poliovirus (Hansen et al., 1997), HCV (Ago et al., 1999; Bressanelli et al., 1999; Lesburg et al., 1999), and phage ϕ 6 (Butcher et al., 2001). The palm, which in all these enzymes also resembles the palm subdomain of the pol(I)-like DNA polymerases, is a four-stranded, antiparallel β sheet supported by three α helices. It contains the conserved acidic residues of the catalytic center (Figure 2A; Bruenn, 1991). The closest relatives of the λ 3 fingers subdomain (among those in known structures) are those of the polymerases from HCV and ϕ 6. Except for two insertions (397–420 and 617–642) in λ 3, secondary structural elements of the HCV and reovirus fingers subdomains can be aligned unambiguously in three dimensions. The thumb subdomain includes a β strand followed by three α helices. The strand forms a small sheet with a β -turn at the C terminus of the palm, a motif found in other polymerases (the “primer grip” of Jacobo-Molina et al., 1993). A β - α - α - α thumb is present in HIV-1 RT and poliovirus polymerase, but the lengths of the secondary structural elements and their spatial arrangements are rather different. The thumb of HCV polymerase is larger, but it also begins with a β - α - α - α structure.

Superposition onto the λ 3 polymerase domain of coordinates of a catalytic complex of HIV-RT shows that palm, fingers, and thumb can all be brought into register. Thus, like the HCV and ϕ 6 polymerases (Bressanelli et al., 2002; Butcher et al., 2001; Lesburg et al., 1999), λ 3 appears to have a “closed” structure, even in its resting state. A hairpin loop between the initial strand and the first helix in the thumb interacts with a loop at the tip of the fingers, creating a complete annulus around the catalytic site on the palm (Figure 2A).

The N-terminal domain of λ 3, residues 1–380, covers one side of the active site cleft. It anchors and augments the continuous surface between the fingers and thumb subdomains. The initial 40 residues of HCV polymerase occupy a similar position, but far more sparsely. Catalysis can occur in the λ 3 crystal without disturbing the lattice, suggesting that the fingertips move only slightly, if at all, during the catalytic cycle. Any significant shift in the position of the fingertips would lead to substantial domain displacements throughout the molecule.

The C-terminal domain of λ 3, residues 891–1267, is an annular structure (bracelet). Its shape and orientation relative to the core polymerase recall those of the sliding clamp of processive DNA polymerases (Shamoo and Steitz, 1999; Stukenberg et al., 1991), but it has an unrelated fold. Sliding clamps encircle DNA and act as teth-

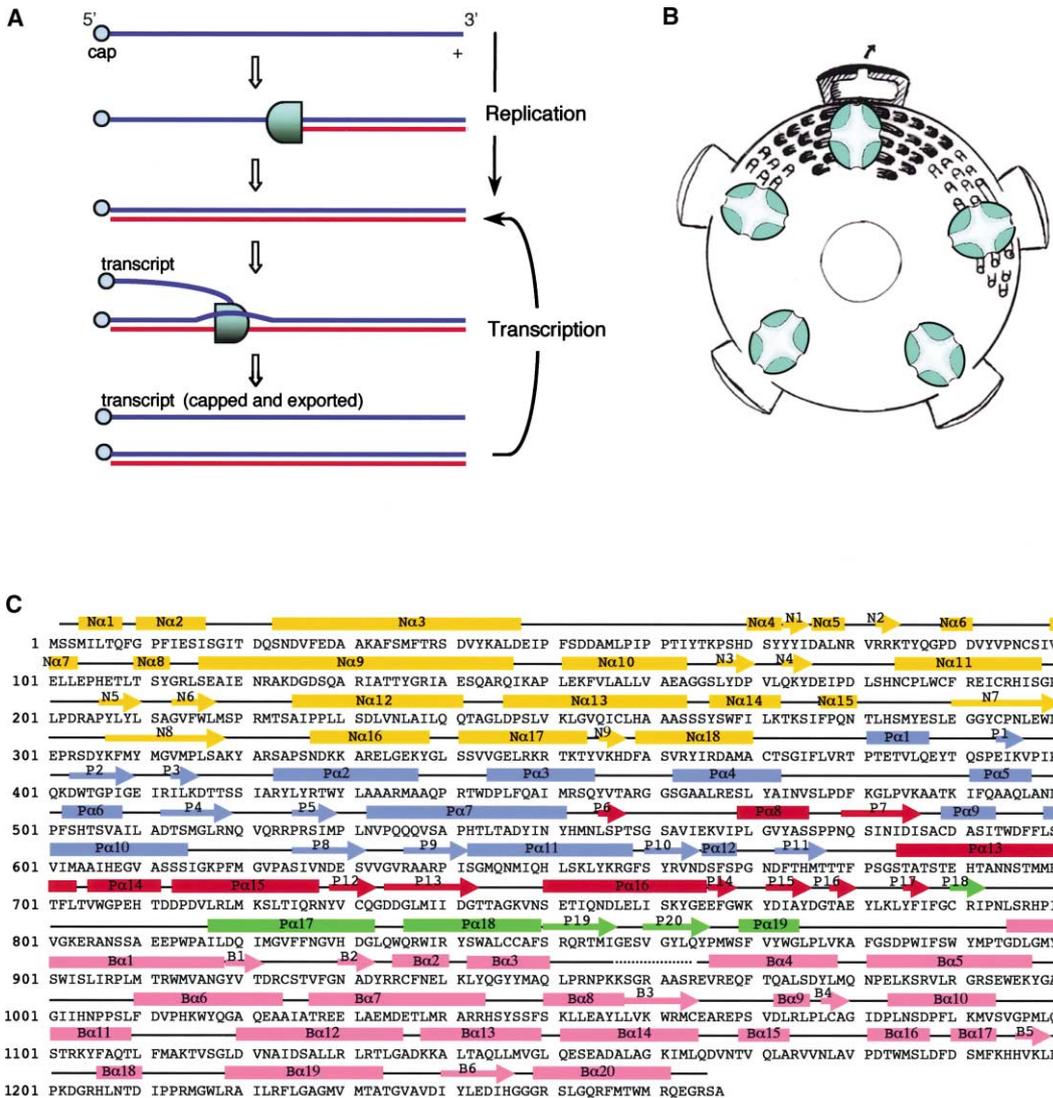


Figure 1. Functions and Primary Structure of the Reovirus Polymerase, $\lambda 3$

(A) RNA synthesis catalyzed by $\lambda 3$. The blue and red lines represent plus (+) and minus (–)-sense RNA, respectively. The polymerase complex is represented by the green, cup-shaped object.

(B) The position of $\lambda 3$ in the reovirus core, with dsRNA packaged in liquid-crystalline arrays. The cylindrical protrusions on the outer surface of the core represent pentamers of the viral capping enzyme. Nascent mRNA passes from the polymerase into these capping chambers.

(C) The amino acid sequence of $\lambda 3$ (type 3 Dearing). Above the sequence are secondary-structure assignments from the work reported here. α helices, β strands, and coils/turns are shown by rods, arrows, and black lines, respectively. Helices and strands are numbered (preceded by N, P α , or B to indicate domain). Disordered elements are shown by dotted lines. Secondary-structure elements are colored by domain as in Figure 2A.

ers to ensure polymerase processivity. The $\lambda 3$ bracelet is unlikely to open and close dynamically, like sliding clamps in DNA replication, because it folds into two tightly sealed circles. The opening has a diameter of about 20 Å and bears a net positive charge.

Each of the domains has extended contacts with the other two (see caption to Figure 2 for buried surfaces). The interfaces contain hydrophobic side chains and bridge the hydrophobic cores of the individual domains. At one point, just beneath the active site cleft in the palm, a short β sheet links the N- and C-terminal domains. These aspects of the structure are consistent with our conclusion that the domains do not move with respect to each other during catalysis.

Channels in the Cage

When represented by its molecular surface, $\lambda 3$ resembles a cube-like cage. The catalytic site, enclosed in the center, is accessible to solvent, substrate, or product through four channels of various sizes at the “front,” “left,” “rear,” and “bottom” (Figures 2B and 2C). The front channel is the bracelet opening just described. It is the only one of the four large enough to admit dsRNA. The left channel is at the interface of the polymerase and bracelet domains, and the rear channel is at the interface of the polymerase and N-terminal domains. All three domains border the bottom channel.

The superposition of the catalytic complex of HIV-1 RT, described above, places the RT template:primer

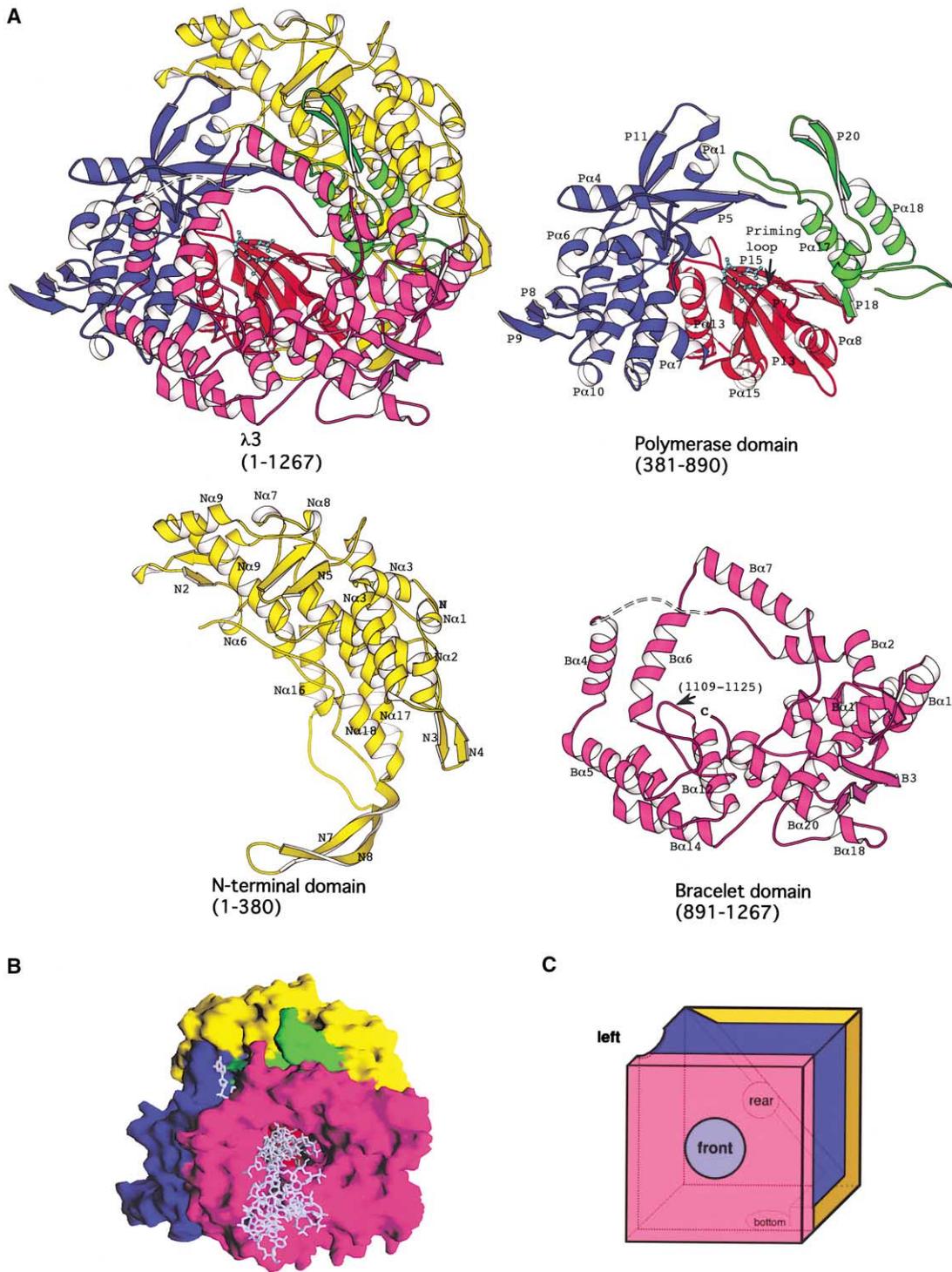


Figure 2. Crystal Structure of $\lambda 3$

(A) Ribbon diagram showing the structure of $\lambda 3$ and its three domains. The N- and C-terminal domains are in yellow and magenta, respectively. In the core polymerase domain, the fingers subdomain is in blue; the palm, in red; and the thumb, in green. The view corresponds to the conventional orientation (Joyce and Steitz, 1994). The three catalytically important aspartic acid residues in the palm are in cyan. Various secondary structural elements are labeled to establish correspondence with the sequence shown in Figure 1B. A disordered loop in the braceletlet, residues 956–965, is shown by a broken line.

(B) $\lambda 3$, in the same orientation as in (A), represented by its molecular surface, colored according to domain and subdomain (see above). A template:primer duplex modeled from the RT ternary complex structure (see Experimental Procedures) is in gray. Total interdomain buried surfaces are: N/Pol, 9200 Å²; Pol/B, 6900 Å²; N/B, 2100 Å².

(C) A simplified representation of $\lambda 3$ as a cubic cage, illustrating the presence of four channels leading to the enclosed active site.

Table 1. Data Collection Statistics for Native Structure Determination

Data Set	Native	EMTS		K ₂ PtCl ₄		K ₂ UO ₂ F ₅		Phenyl Mercury Nitrate	
		Peak	Backsoak	Peak	Backsoak	Peak	Backsoak	Peak	Backsoak
Preparation of heavy-atom derivative	-	0.1 mM, 2 hr + backsoak	0.2 mM, 4 hr + backsoak	0.5 mM, 4 hr no backsoak	1% of saturated solution, 5 hr + backsoak				
Wavelength (Å)	1.0000	1.0063	1.0715	1.0715	1.0060				
Number of reflections ^a	337,229 (52,430)	247,466 (38,398)	162,593 (30,588)	287,466 (34,800)	208,616 (35,236)	0.9844	207,894 (34,862)	200,234 (35,094)	
Completeness (%) ^b	97.9 (85.0)	93.2 (61.8)	97.4 (95.0)	98.5 (97.2)	92.1 (80.6)		92.0 (79.7)	92.1 (80.5)	
Resolution (Å)	30–2.5	30–2.8	30–3.0	30–2.9	30–2.9		30–2.9	30–2.9	
R _{merge} (%) ^{b,c}	6.9 (26.7)	5.4 (23.0)	7.6 (33.7)	6.5 (19.4)	7.6 (27.3)		8.9 (36.4)	8.1 (31.0)	
R _{iso} (%) ^{b,c} (15.0–3.0 Å)	-	13.6 (17.4)	20.0 (25.0)	22.0 (25.2)	19.5 (23.2)		19.3 (23.5)	18.9 (22.5)	
R _{ano} (%) ^{b,c}	-	2.5 (13.5)	4.2 (17.4)	-	4.0 (16.2)		4.5 (22.1)	3.8 (18.6)	
R _{all} (%) ^{b,c} (centric/acentric)	-	0.82/0.93 (MIR)	0.93/0.86 (MIR)	0.88/0.81 (MIR)	0.53/0.72 (MIR)		-	-	
Phasing power ^d (centric/acentric)	-	1.10/0.73 (MIR)	0.66/0.54 (MIR)	0.88/0.65 (MIR)	2.49/1.43 (MIR)		-	-	
Figure of merit	-	-	-	0.6915 (MIR + MAD)	-		-	-	
Refinement statistics									
R	30–2.5 Å								
R _{free}	21.3% for 49,367 reflections								
R _{ms} bond length and bond angles	24.5% for 2,767 reflections								
	0.008 Å, 1.470°								

^aNumbers in parentheses are the number of unique reflections.

^bNumbers in parentheses are for the highest-resolution bins.

$$R_{\text{merge}} = \frac{\sum_h \sum_l |I_{hl} - \langle I_h \rangle|}{\sum_h \sum_l I_{hl}}$$

$$R_{\text{iso}} = \frac{\sum_h |F_{Ph} - F_{P(h)}|}{\sum_h F_{Ph}}$$

$$R_{\text{ano}} = \frac{\sum_h |I_h^+ - I_h^-|}{\sum_h (I_h^+ + I_h^-)}$$

$$R_{\text{all}} = \frac{\sum_h |F_{Ph} - F_{P(h)}|}{\sum_h |F_{Ph} + F_{P(h)}|}$$

$$R_{\text{phasing}} = \frac{\sum_h (Caic) \sqrt{\sum_h |F_{Ph} - F_{P(h)}|}}{\sum_h |F_{Ph} + F_{P(h)}|}$$

^dPhasing power = $\text{rms}(F_h/E)$, where E is the lack of closure error.

Table 2. Data Collection Statistics for Enzyme-Substrate Complexes

Experiment	Initiation Complex	Elongation Complex #1	Elongation Complex #2	Cap Analog Complex
Experimental condition	0.5 mM RNA oligo AUUAGC 1 mM 3'-deoxy rGTP 1 mM 3'-deoxy rCTP 1.5 mM Mn ²⁺ Soak time > 6 hr	0.5 mM RNA oligo AUUAGC 1 mM rGTP 1 mM 3'-deoxy rCTP 1 mM 3'-deoxy rUTP 1.5 mM Mn ²⁺ Soak time ≈ 45 min	0.5 mM RNA oligo UAGCCCC 1 mM rGTP 1 mM 3'-deoxy rCTP 1.5 mM Mn ²⁺ Soak time ≈ 45 min	0.5 mM cap analog m ⁷ GpppG Soak time ≈ 2 hr
Unit cell dimensions	a = 71.06, b = 85.14, c = 250.25 Å α = β = γ = 90°	a = 70.72, b = 85.21, c = 248.47 Å α = β = γ = 90°	a = 70.88, b = 85.01, c = 249.29 Å α = β = γ = 90°	a = 70.89, b = 84.94, c = 249.25 Å α = β = γ = 90°
Resolution (Å)	30–2.8	30–2.8	30–2.5	30–2.5
Completeness (%) ^a	79.4 (75.6)	86.4 (82.1)	89.7 (75.3)	93.3 (80.8)
Redundancy	2–3	5–6	3–4	5–6
R _{merge} ^a	12.0 (28.3)	13.9 (52.2)	8.7 (48.2)	8.8 (40.4)
Refinement Statistics				
R (%)	19.3	19.9	22.0	22.0
R _{free} (%)	26.1	26.0	27.5	25.2
Rms bond length (Å)	0.007	0.007	0.007	0.009
Rms bond angles (°)	1.395	1.394	1.489	1.719

^aThe numbers in parentheses refer to the highest-resolution shell.

DNA duplex (the “upstream” end of the template) in the front channel of λ3 and the 5' overhang of the template (the “downstream” end) in the left channel. The latter is lined with basic residues, which should facilitate template entry. Moreover, the incoming nucleotide and the two divalent metal ions in the superposed RT structure (and in the catalytic complexes of λ3 described below) lie near the three conserved aspartate residues in the λ3 palm. The modeled incoming nucleotide is near the inner mouth of the rear channel, and nucleotides needed for RNA polymerization are likely to diffuse through this opening into the active site. A well-defined channel for nucleotide diffusion has also been noted in the ϕ6 polymerase (Butcher et al., 2001) and in DNA-dependent RNA polymerases (Gnatt et al., 2001; Murakami et al., 2002).

The bottom channel in λ3 is a structural feature not yet observed in other polymerases. We propose that when the polymerase is operating in transcriptional mode, the transcript exits the active site through this bottom channel. We discuss this assignment further in relation to the structures of catalytic complexes, described below.

Initiation Complexes

The structure of λ3 in its apo form suggests that only internal adjustments might be required to accommodate substrates and that catalytic complexes might form in the crystals. As expected, initiation complexes were obtained by soaking crystals of λ3 in solutions containing oligoribonucleotides, 3'-deoxy rNTPs, and Mn²⁺ (see Table 2 and Experimental Procedures). The absence of a 3'-OH group on the ribonucleotide prevented attack on the triphosphate of a nucleotide in the incoming (substrate) position by a nucleotide in the priming position. The complex was thus stalled at the initiation stage. The sequences of the oligoribonucleotides corresponded to the termini of reovirus genome segment L3, which encodes the major inner-capsid protein, λ1. Replication

and transcription are initiated at the 3' ends of the plus- and minus-sense genomic strands, respectively. The three oligonucleotides we used are shown in Figure 3A, in the register with respect to the incoming nucleotide seen in the crystals.

Template Conformation

Difference maps of crystals soaked as described all showed strong RNA electron density in the left channel. The 3' end enters the active site, making base pairs with bound rNTPs (Figure 3A). Assignment of RNA sequence was based on electron density, which was clear enough to distinguish pyrimidines and purines. The phosphate backbone of the template follows a path similar to that taken by the template overhang in HIV-1 RT (Huang et al., 1998) and T7 DNA and RNA polymerase catalytic complexes (Cheetham and Steitz, 1999; Doublet et al., 1998). The ribose and base of the templating nucleotide (position +1) stack tightly under the side chains of P530 and I528, forcing the downstream template to bend away from the catalytic pocket (Figures 3B and 3C).

Initiation Sequence Specificity

The observed positions of the oligonucleotides depend on their base sequences (Figure 3A). A template G is favored at position +1. It provides a carbonyl for a hydrogen bond with R518 and an –NH for a hydrogen bond with the side chain of S682 (Figure 3C). U at this position can also interact favorably with R518, but A and C cannot. Given the conserved 3' terminal sequences of reovirus RNAs (plus, UCAUC-3'; minus, UAGC-3'), the preference for G or U at +1 promotes synthesis of full-length RNAs of either sense. The ϕ6 RNA polymerase appears to have a different mechanism for ensuring that the initiation of RNA synthesis occurs at the very end of its template. Its C-terminal domain blocks the product tunnel and helps to position the 3' end of template in the correct initiation register (Butcher et al., 2001). It presumably then moves away to allow elongation.

rNTP Sites

The overall active site conformation and the positioning of the incoming and priming nucleoside triphosphates

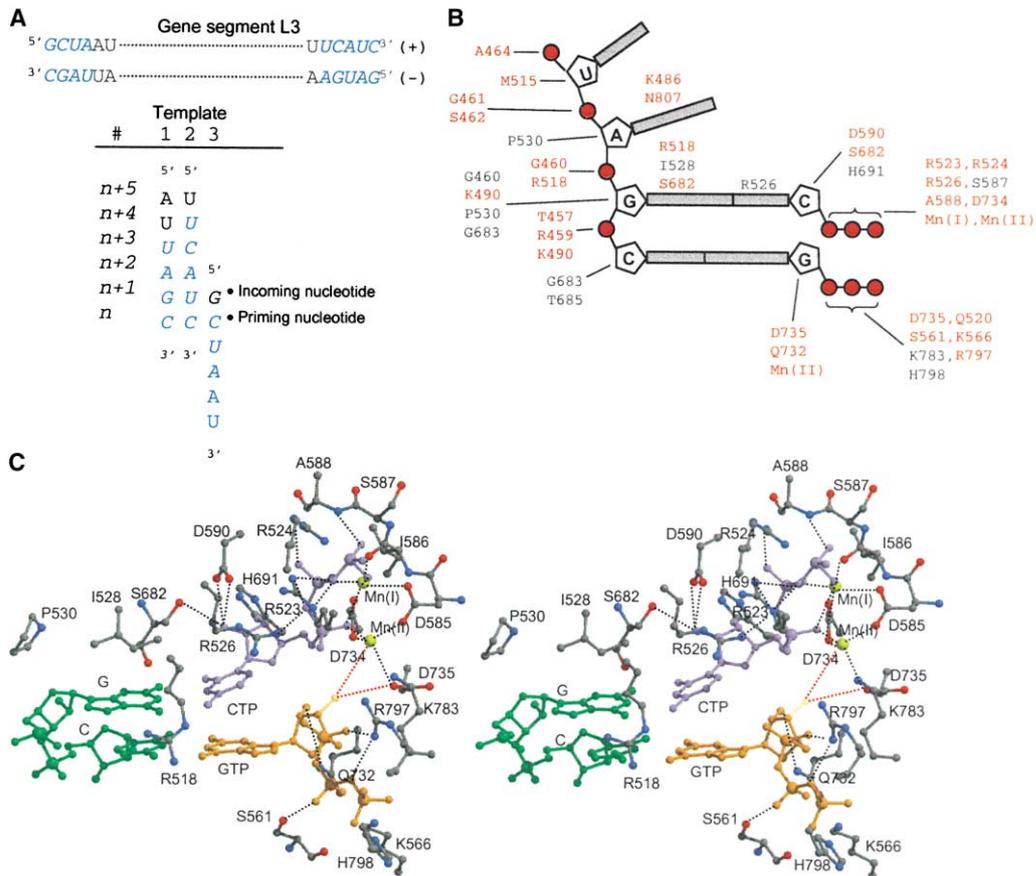


Figure 3. Structure of the $\lambda 3$ Initiation Complex

(A) Three oligonucleotides and their spatial register (as seen in soaked crystals) when bound in the $\lambda 3$ active site. These oligonucleotides are terminal sequences in gene segment L3. Motifs conserved among different gene segments are shown in blue italics.

(B) Interactions between $\lambda 3$ and its substrates. Residues making polar interactions with substrates are highlighted in red.

(C) Stereo view of the $\lambda 3$ active site. Template strand, green; 3'-deoxy rCTP at incoming position, purple; 3'-deoxy rGTP at priming position, orange; Mn^{2+} ions, gold spheres; assigned hydrogen bonds and metal ligand interactions, dashed lines. The missing 3'-OH is shown in a lighter shade, and the related hydrogen bonds, in red. In initiation and catalytic complexes, the α -phosphorous of the incoming nucleotide tends to be about 0.5–1 Å farther from the 3'-carbon of the priming nucleotide than expected in a productive complex. The absence of the 3'-OH (a ligand of one of the Mg^{2+} ions) may yield a slightly relaxed structure.

are very similar to those in other polymerase complexes. The 3'-OH group (missing here) of the priming rNTP would be placed for in-line attack on the α phosphate of the incoming rNTP. The latter pairs with the template base at position +1 and its triphosphate is coordinated by two divalent cations, three basic residues (R523, R524, and R526) from the NTP binding loop at the tip of the fingers subdomain, and the main chain –NH group of A588. The 2'-OH of the incoming rNTP forms hydrogen bonds with D590 and S682; the 2'-OH of the nucleotide at the primer position forms a hydrogen bond with Q732. These contacts presumably confer the observed specificity of $\lambda 3$ for ribonucleotides.

A loop containing residues 558–565 supports the priming rNTP. This “priming loop” is effectively an insertion within a strand that is found in the palms of all other polymerases. The insertion is unique (among known structures) to $\lambda 3$.

Conformational Change in $\lambda 3$

There is no major domain movement in $\lambda 3$ when it binds oligoribonucleotides and rNTPs (Figure 4A). Noticeable

conformational changes include side chain rearrangements in the NTP binding loop at the fingertips and a slight expansion of the template channel. Superposition of the apo and initiating forms of the enzyme gives an rms deviation of 0.66 Å for main chain and side chain atoms. One of the two loops that separates the front and template channels, residues 956–965, is disordered in the apo-enzyme but becomes ordered in the complex.

RNA Polymerization in the Crystal

Synthesis of dsRNA

Crystalline $\lambda 3$ catalyzes phosphodiester bond formation. Our strategy was to incubate the crystals with suitable ssRNA template, divalent cation Mn^{2+} , and rNTPs (see Table 2 and Experimental Procedures) and to force the polymerization reaction to terminate at a specific position by incorporation of a 3'-deoxy rNTP. In one crystal, a dinucleotide product 5'-GC-3' was synthesized using 5'-AUUAGC-3' as template. In a second, 5'-GGGGG-3' was made with the template 5'-UAGCCCC-3'. The products in both cases were dsRNA (Figure 4),

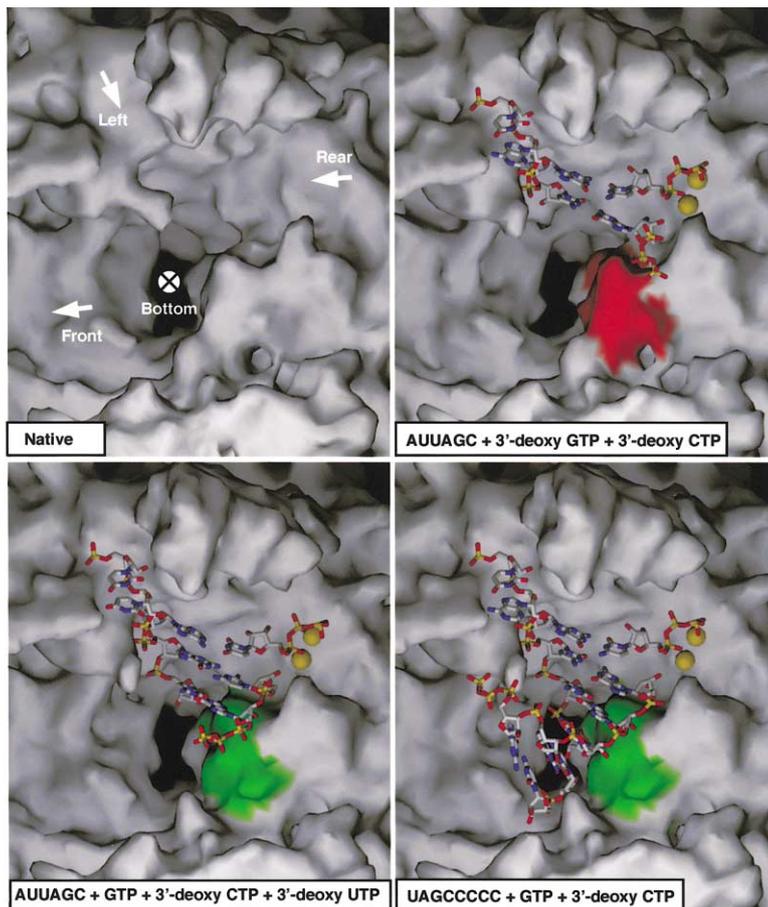


Figure 4. RNA Synthesis by $\lambda 3$

The active site of $\lambda 3$ is shown, along with experimental conditions, at four progressive stages—free $\lambda 3$, an initiation complex, after synthesis of a 2 nucleotide product, and after synthesis of a 5 nucleotide product. $\lambda 3$ is rendered as a molecular surface; RNA and nucleotides, as atomic models; metal ions, as yellow spheres. To reveal the active site, the N-terminal domain, the thumb, and residues 517–528, 942–975, and 1014–1046 have been removed. The four channels are indicated on the image of free $\lambda 3$. Upon formation of the first phosphodiester bond, the loop (residues 658–665) that supports the priming nucleotide changes its conformation (red \rightarrow green), thus avoiding collision with the dsRNA duplex.

which projected into the front channel. In the second crystal, we anticipated seeing 5'-GGGGGC-3' as product, because 3'-deoxy CTP was used in the soak. In fact, there is probably some heterogeneity of product, because occupancy of the product strand decreases toward its 5' end. Reasonable occupancy for the 5'-U of the template strand dictated our choice of primary product in model building. Note that the template in this second crystal did not contain the preferred initiation sequence described above. The polymerase contacts product RNA along its minor groove, forming both hydrogen bonds and van der Waals contacts with RNA backbone (Figure 3B) and van der Waals contacts with bases (except for two hydrogen bonds to a templating G at +1; see previous section).

Abortive Initiation and Strand Separation

The short double-strand products we obtained may correspond to initial products of either the replicative or the transcriptional mode of reovirus RNA synthesis. The loop that supports the priming rNTP has retracted by about 3 Å toward the palm with respect to its position in the apo-enzyme and in the initiation complexes, and it now fits snugly into the minor groove of the product duplex (Figure 4C). Failure of this loop to shift would block elongation and lead to abortive initiation. Indeed, abortive transcripts of reovirus are predominantly 2–4 bases long (Farsetta et al., 2000), suggesting that the priming loop can provide a kinetic barrier for the transition from initiation to elongation. Modeling further extension of the template:product shows that unless the du-

plex were to bend, the product strand would probably collide with a loop (1109–1125) from the C-terminal bracelet domain. This structure might help separate template and product in the transcriptional mode and to direct the transcript toward the bottom channel. Alternatively, separation might depend on partial constriction of the front channel, perhaps by a shift in the flexible loop between residues 957 and 964, or (as suggested by Butcher et al., 2001) by binding of $\mu 2$, the second internal protein in the reovirus core. In transcriptional RNA polymerases from prokaryotes and eukaryotes, a so-called rudder (a projection from the Rpb1 or β' subunit) is believed to split template and transcript as polymerization propels the base-paired product against it (Gnatt et al., 2001 [#119]; Murakami et al., 2002 [#121]; but see Kuznedelov et al., 2002 for evidence to the contrary).

Interaction with the $\lambda 1$ Capsid Shell and with $\mu 2$

Each $\lambda 3$ polymerase must lie close to one of the five-fold pores in the $\lambda 1$ shell so that mRNA molecules can pass smoothly into the turret chambers for capping and export. Inward-projecting density features in image reconstructions of empty cores suggest that the polymerase is present near the five-folds even in the absence of RNA (Dryden et al., 1998). If $\mu 2$ is the 5'-terminal phosphatase that prepares the mRNA terminus for capping, then the leading end of a transcript must pass across it before entering the capping chamber, and the two proteins may be jointly tethered. The exit channel for transcript and the $\mu 2$ catalytic site are likely to face

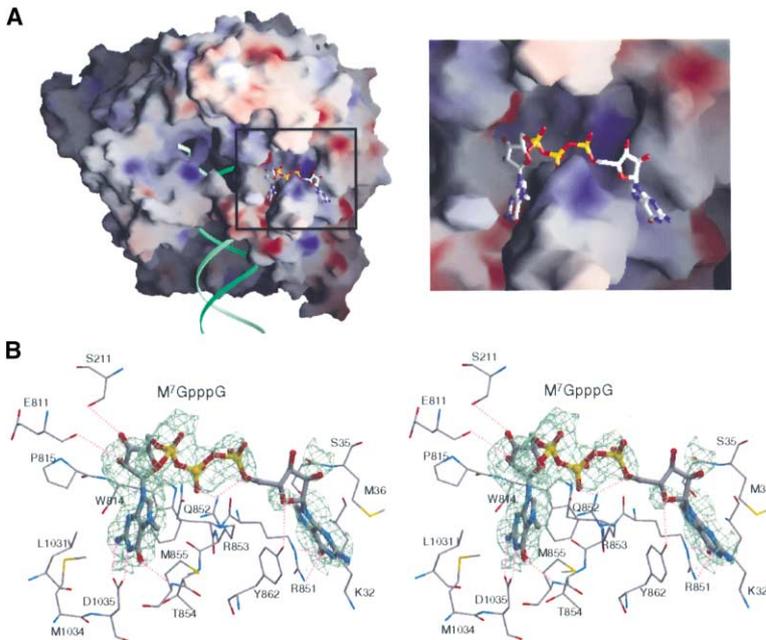


Figure 5. Cap Binding Site on $\lambda 3$

(A) Location. $\lambda 3$ is rendered as a molecular surface colored according to electrostatic potential, with positive in blue and negative in red. On the right is an enlarged view of the pocket. The phosphate backbone of a modeled template:primer duplex is shown by green ribbons.

(B) Stereo diagram of the cap binding pocket. The view is taken in the same orientation as in (A). Hydrogen bonds are represented by pink dotted lines. Electron density corresponding to the cap analog is shown in green at a level of 0.8σ .

the five-fold pore, leaving the left, front, and rear channels open to the required molecular traffic.

Cap Recognition

Detection, during the course of the experiments described in the previous section, of a GTP binding site on the polymerase surface led us to soak crystals with an analog of the N⁷-methyl-GpppG mRNA cap (Figure 5). Difference maps showed strong features for the N⁷-methyl G (G1) and its associated sugar; the triphosphate linkage and the second guanosine were not as well ordered. A clear “bump” on the better-ordered purine base, corresponding to the N⁷-CH₃, allowed us to distinguish the two halves of the ligand. Moreover, D1035 forms hydrogen bonds with G1 (the positively charged, methylated purine), whereas Arg851 contacts G2. Acidic residues are usually important in coordinating the positively charged, N⁷-methyl guanines in cap binding proteins (Hu et al., 1999; Marcotrigiano et al., 1997). Both bases are sandwiched between hydrophobic, largely aliphatic side chains (Figure 5B). The $\lambda 3$ cap binding site differs from those in vaccinia virus VP39 and eIF4, which provide stacking interactions with aromatic side chains to stabilize the methylated base and which do not appear to have a pocket for the unmethylated guanine (G2).

Soaking experiments also show that the conserved tetra-nucleotide 5'-GCUA-3' (Figure 3A) does not bind efficiently to $\lambda 3$, other than at the active site. Thus, the cap, rather than the 5' sequence, appears to be the primary element by which $\lambda 3$ grasps the 5' end of the positive strand. A few positively charged surface patches on the front face of $\lambda 3$ could provide a weak binding site for the conserved 5' nucleotides, perhaps contributing to selectivity in packaging. In the crystals, cap binding activity and the enzymatic activity of $\lambda 3$ are mutually independent.

Discussion

Relation to Other Polymerases

Because one can synthesize RNA in situ in native $\lambda 3$ crystals, this enzyme is an excellent model for characterizing intermediate steps in nucleotide incorporation. Unlike HIV-RT and the T7 DNA and RNA polymerases, $\lambda 3$ binds template and substrate with only localized rearrangements. Moreover, template translocation and nucleotide incorporation can be accomplished with little structural perturbation. The accuracy of RNA synthesis by $\lambda 3$ is about one error for every 10^4 bases, comparable to the fidelity of other polymerases without proofreading activity (Echols and Goodman, 1991).

The structural characteristics of the catalytic complex also resemble closely those of other polymerase catalytic complexes for which structures are known (Bressanelli et al., 2002; Butcher et al., 2001; Doublet et al., 1999; Huang et al., 1998). These conserved features include the conformation and ligation of the incoming nucleotide, the conformation of the template at the -2 , -1 , and $+1$ positions, the way in which the templating base is “jammed” against a segment of the fingers subdomain, and the way in which other interactions constrain the geometry of template-primer and template-incoming nucleotide base pairs. In the case of polymerases, such as T7 DNA polymerases and HIV-RT, that close in upon their substrates, it has not been clear whether they need to open and close completely for each catalytic step (Doublet et al., 1999). Our results, showing multistep polymerization within crystals of $\lambda 3$, suggest they do not need to do so.

Spatial Organization in Reovirus Replication and Transcription

The architecture of $\lambda 3$ polymerase has evidently evolved to ensure spatial and temporal coordination in the crowded interior of the reovirus core. The catalytic core

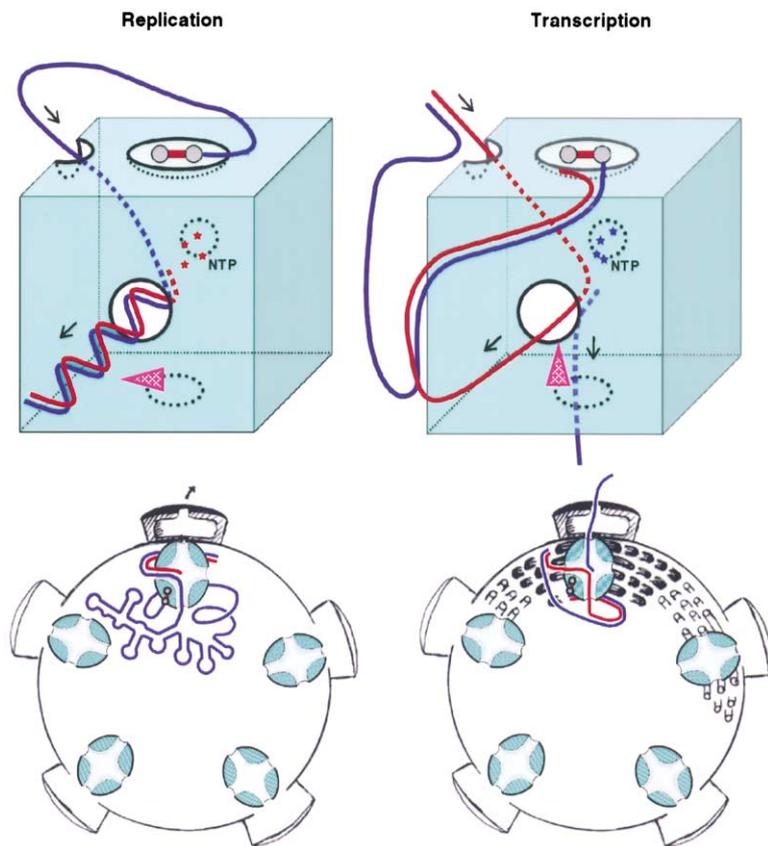


Figure 6. Model for Spatial Organization of Polynucleotide Synthesis

Left, replication; right, transcription. $\lambda 3$ is represented by a green cube. Blue lines and red lines correspond to (+) and (-) RNA, respectively. The saucer-like feature on the top of the $\lambda 3$ cube represents the cap binding site; the dumbbell-like object represents the 5' cap. The template RNA exits the active site through the front channel. The nascent strand has two possible routes: either through the front channel during replication or through the bottom channel during transcription. The pink arrow symbolizes a presumptive separator structure that splits the nascent transcript from its template.

lies at the center of a cage with four defined channels. There is an external site for binding a 5'-m⁷G RNA cap to tether this end to the polymerase. We can identify candidate structures for specifying the exit route of the RNA product.

Cap binding may be significant for two distinct stages in the replication cycle: RNA packaging into assembling cores and mRNA synthesis after viral entry into a host. In the packaging step, $\lambda 3$ could recognize capped viral RNA segments by conserved terminal elements: 5'-N⁷-MeG-GCUA and UCAUC-3'. Binding of the cap at one position on the polymerase surface and insertion of the 3' end into the template channel nearby could ensure association of a single polymerase molecule with each segment of viral plus-strand RNA. We suggest that higher-order structures might then form between different RNA segments to select a full set of ten distinct mRNAs for packaging. Interactions between the $\lambda 3$ polymerase molecules and the interior of the assembling $\lambda 1$ shell could promote efficient incorporation of the viral genome. Such a process would be different from the RNA packaging mechanism of bacteriophage $\phi 6$, in which plus-sense RNAs are inserted into preassembled procapsids (Mindich, 1999).

In assembling cores, RNA synthesis probably proceeds as visualized in our crystallographic studies (Figure 6). Hydrogen bonds between protein side chains and the second template base (Figures 3B and 3C) prevent overshoot before initiation. The dsRNA product exits through the front channel (Figure 6). There are five equivalent interactions with the capsid shell at each five-fold

pore, and it is possible that the polymerase rotates from one to the next during RNA synthesis. Counterrotation of $\lambda 3$ as dsRNA product emerges could facilitate spooling of individual genome segments (Figures 1 and 6). Local genomic RNA secondary structure, and any higher-order RNA structures joining different genomic strands, will be disrupted as the genomic RNA feeds into the active site. Moreover, when RNA synthesis approaches the 5' end of the template, the terminus will be forced to leave the cap site and enter the template channel. As it emerges through the front channel, the cap can be recaptured by the cap binding site, which is adjacent to the front-channel exit.

After infection and activation of the core by stripping of the outer-shell proteins, transcription can proceed using the negative strand as a template (Figure 6). Attachment of the 5' end of the positive strand to the cap site will facilitate insertion of the free 3' end of the negative strand into the template channel. Moreover, after completion of each round of mRNA synthesis, the 3' end of the negative strand will easily reinsert and reinitiate. Were the 5' end of the positive strand not anchored firmly to the polymerase, the 3' end of the template would "get lost" within the coils of dsRNA in the interior of the core, greatly decreasing the efficiency of reinitiation. Moreover, a cap binding site prevents initiation from the wrong strand. Similar mechanisms might be adopted by other RNA viruses in order to synthesize predominantly plus- or minus-sense RNA molecules during viral replication (Ahola et al., 2000).

Both the core capsid shell protein $\lambda 1$ (Bisaillon et al.,

1997) and the internal core protein μ 2 (Noble and Nibert, 1997) have been proposed to have helicase activity. Cap binding by λ 3 may eliminate the need for any helicase to initiate duplex melting by keeping the terminal base pairs separated. Our structure also suggests that separation of template and transcript occurs within the polymerase cage, so that this function likewise appears to reside within λ 3. It could involve a triggered switch to a strand-separating conformation, but we have not yet identified the likely trigger.

The special structural features of λ 3 illustrate that functional requirements for some degree of long-range order—e.g., finding the 3' end of the template or packaging a full genome efficiently within the cavity of the core—can be satisfied by appropriately organized, *local* machinery. Moreover, there may be additional surface features of λ 3 not fully revealed by the present structures that make other contributions to long-range organization. In particular, there may be aspects of the λ 3 cage that help guide the plus-strand (nontemplate) side of the transcription bubble, and there is probably a docking surface for μ 2. Indeed, working out the role of μ 2 is clearly the next step in analyzing this elegantly simple replication/transcription complex.

Experimental Procedures

Expression and Purification

Sf21 cells in suspension culture at 1×10^6 /ml were infected at MOI of 2 with baculovirus harboring the L3 gene under control of the polyhedrin promoter. The cells were harvested 60 hr later, pelleted, washed with cold PBS, resuspended at 7.5×10^6 cells/ml in 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol, and homogenized. The membrane fraction was removed by low-speed centrifugation. Nearly two-thirds of the total λ 3 could be recovered from the supernatant by precipitation with ammonium sulfate [10%–20% (w/v)]. The protein was dissolved in fractionation buffer (50 mM Na-HEPES [pH 7.8], 50 mM NaCl, 1 mM EDTA, 10% glycerol, 2 mM 2-mercaptoethanol), applied to heparin-Sepharose, and eluted with a gradient to 500 mM NaCl. The λ 3 peak was collected, concentrated to \sim 2 ml, and applied to a 60 ml Superdex-200 column (Amersham Pharmacia Biotech) in 100 mM NaCl. Fractions containing λ 3 were pooled and dialyzed against fractionation buffer. In a final step, the protein was applied to a Mono-S column (Amersham Pharmacia Biotech), from which it eluted at a NaCl concentration of 150 mM. The protein was about 95% pure, as judged by Coomassie blue staining after SDS-PAGE.

Crystallization and Data Collection

Crystals of λ 3 in space group $P2_12_12_1$, $a = 70.9$, $b = 85.2$, $c = 248.6$ Å, were obtained by hanging-drop vapor diffusion. The best crystals grew at 22°C when 1.2 μ l droplets of the recombinant protein (8.7 mg/ml in 50 mM Na-HEPES [pH 7.8], 10% glycerol, 150 mM NaCl, 10 mM DTT, 0.02% NaN_3) were mixed with 1 μ l well solution containing 7.5% PEG8000, 10% glycerol, 0.1 M Na-HEPES (pH 7.8). With microseeding, thin plate-like crystals appeared after 1 day and grew to full size ($\sim 10 \times 400 \times 400$ μm^3) in about a week. Each data set was collected from a single frozen crystal that had been transferred to stabilizing solution (7.5% PEG8000, 25% glycerol, 50 mM Na-HEPES [pH 7.8]) immediately before flash freezing in a boiling nitrogen stream at 100°K. Heavy atom derivatives were prepared by soaking crystals in stabilizing solution containing heavy atom compounds at the concentrations in Table 1. Crystals derivatized in Hg and Pt compounds were back-soaked for at least 2 hr. For the EMTS and K_2PtCl_6 derivatives, the wavelength was adjusted to maximize f'' , determined from an extended X-ray absorption fine structure (EXAFS) scan. For the PMN derivative, complete MAD data sets were collected at three different wavelengths (Table 1). All data were recorded on Quantum Q4 CCD area detectors and processed

using DENZO and SCALEPACK (Table 1; Otwinowski and Minor, 1997).

Structure Determination

We combined multiple isomorphous replacement (MIR) phases from three different derivatives and phases from multiple-wavelength anomalous dispersion (MAD) from one derivative. We used an anomalous difference Patterson map calculated with data collected at the f''_{max} wavelength to identify the two major heavy-atom sites of the PMN derivative. Three PMN minor sites and the heavy-atom sites for the Pt, UO_2 , and EMTS derivatives were located in Fourier difference maps. Ethyl-Hg from EMTS bound at the two major sites of the PMN derivative. Heavy-atom sites were refined using MLPHARE (Table 1; CCP4, 1994). Anomalous signals were included in both refinement and phasing for all data sets except for UO_2 . Heavy-atom positions of the correct hand gave positive anomalous occupancies. For the PMN derivative, data at the f'_{min} wavelength were used as a reference against which dispersive differences from the other two data sets were phased. Phases from the two sources, MIR and MAD, were combined with the program SIGMAA (CCP4, 1994) and were improved and extended using DM (Cowtan and Main, 1998) with solvent flipping. The resulting map, calculated to 2.5 Å resolution, allowed >95% of the polypeptide chain to be modeled unambiguously using the program O (Jones et al., 1991). For crystallographic refinement we used CNS (Brunger et al., 1998) with idealized amino acid parameters (Engh and Huber, 1991). Six percent of the reflections were omitted during the refinement and used to calculate R_{free} (Brunger, 1992).

Individual atomic temperature factors were refined after several rounds of positional refinement and manual correction. A total of 600 water molecules were placed automatically in CNS in two iterations. The overwhelming majority of amino acids were associated with well-defined electron density. The exceptions were the surface regions 388–392, 471–483, 559–566, and 630–639, where B factors were high. The final protein model was analyzed with the program PROCHECK (Laskowski et al., 1993). Over 89% of the residues were in the most favored regions of a Ramachandran plot, and none were in disallowed regions.

Soaking Experiments

Initiation Complex

Crystals of λ 3 were transferred to stabilizing solution containing 0.5 mM RNA oligonucleotide, 1 mM 3'-deoxy rGTP, 1 mM 3'-deoxy rCTP, and 1.5 mM Mn^{2+} and incubated for 6 hr (Table 2). To inhibit traces of RNase activity, the stabilizing solution (prior to adding nucleotides) was pretreated with *RNase*secure (Ambion, Inc.). The structure of native λ 3 was used as a phasing model. After initial rigid body refinement, R and R_{free} dropped to 0.289 and 0.319, respectively. The subsequent map showed new electron density in both the template channel and at the active site. An RNA oligonucleotide, rNTPs, and Mn^{2+} ions were built using the program O (Jones et al., 1991). Their positions were very close to those in a model derived from the RT ternary complex (Huang et al., 1998), by aligning equivalent residues 580–591, 728–740, and 687–702 from λ 3 with structurally equivalent residues 105–116, 179–191, and 156–171 from the catalytic subunit of RT. The structure of the λ 3 initiation complex was refined using CNS (Table 2; Brunger et al., 1998).

Elongation Complexes

For *in situ* RNA synthesis, two RNA oligonucleotides, 5'-AUUAGC-3' and 5'-UAGCCCCC-3', were used in two independent experiments. Crystals were incubated as described (Table 2). The λ 3 initiation complex was used for starting phases. Rigid body refinement brought the initial R/R_{free} to 0.231/0.279 and 0.243/0.279 for structures complexed with 5'-AUUAGC-3' and 5'-UAGCCCCC-3', respectively. Nascent dsRNA helix was observed in ($F_o - F_c$) map in both cases. For the crystal soaked in 5'-AUUAGC-3', the density showed a dinucleotide 5'-GC-3'; for the crystal soaked in 5'-UAGCCCCC-3', the product was 5'-GGGGG-3'. The electron density for both the nascent and template RNA became progressively weaker as one moved away from the active site. The triphosphate group at the 5'-end of the nascent chain was not visible.

λ 3 Complexed with Cap Analog

Crystals were soaked for 2 hr in stabilizing solution containing the 0.5 mM ribo m⁷G cap analog, m⁷GpppG (Promega, Co.). Native λ 3

coordinates were used to phase the structure. The structure was refined using the program CNS, with modified ligand parameter files calculated from the refined structure of GpppG complexed with guanylyltransferase (Hakansson and Wigley, 1998).

Preparation of Figures

Figures were prepared with Molscrip (Kraulis, 1991), Grasp (Nicholls et al., 1991), and Spock (J.A. Christopher, <http://quorum.tamu.edu/spock>).

Acknowledgments

We thank Karin Reinisch, Yakov Korkhin, Andrea Csejtei, Lei Jin, Piotr Sliz, Susanne Liemann, Min Lei, Brian McClain, Kimberly Mahan, Qing Fan, and other members of the laboratories of Stephen C. Harrison and Don C. Wiley. This work was supported by NIH grant CA13202 to S.C.H. and NIH grants AI-39533 and AI-47904 to M.L.N. S.C.H. is an Investigator in the Howard Hughes Medical Institute.

Received: July 26, 2002

Revised: September 25, 2002

References

- Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K., and Miyano, M. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure Fold. Des.* 7, 1417–1426.
- Ahola, T., den Boon, J.A., and Ahlquist, P. (2000). Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment, negative-strand RNA synthesis, and viral RNA capping. *J. Virol.* 74, 8803–8811.
- Bisaillon, M., Bergeron, J., and Lemay, G. (1997). Characterization of the nucleoside triphosphate phosphohydrolase and helicase activities of the reovirus lambda1 protein. *J. Biol. Chem.* 272, 18298–18303.
- Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R.L., Mathieu, M., De Francesco, R., and Rey, F.A. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* 96, 13034–13039.
- Bressanelli, S., Tomei, L., Rey, F.A., and De Francesco, R. (2002). Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J. Virol.* 76, 3482–3492.
- Bruenn, J.A. (1991). Relationships among the positive strand and double-strand RNA viruses as viewed through their RNA-dependent RNA polymerases. *Nucleic Acids Res.* 19, 217–226.
- Brunger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355, 472–474.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J., Kuzewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* 54, 905–921.
- Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H., and Stuart, D.I. (2001). A mechanism for initiating RNA-dependent RNA polymerization. *Nature* 410, 235–240.
- CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for X-ray crystallography. *Acta Crystallogr. D* 50, 760–763.
- Cheetham, G.M., and Steitz, T.A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science* 286, 2305–2309.
- Cheetham, G.M., and Steitz, T.A. (2000). Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases. *Curr. Opin. Struct. Biol.* 10, 117–123.
- Cowtan, K., and Main, P. (1998). Miscellaneous algorithms for density modification. *Acta Crystallogr. D Biol. Crystallogr.* 54, 487–493.
- Double, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 391, 251–258.
- Double, S., Sawaya, M.R., and Ellenberger, T. (1999). An open and closed case for all polymerases. *Struct. Fold. Des.* 7, R31–R35.
- Dryden, K.A., Wang, G.W., Yeager, M., Nibert, M.L., Coombs, K.M., Furlong, D.B., Fields, B.N., and Baker, T.S. (1993). Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. *J. Cell Biol.* 122, 1023–1041.
- Dryden, K.A., Farsetta, D.L., Wang, G., Keegan, J.M., Fields, B.N., Baker, T.S., and Nibert, M.L. (1998). Internal structures containing transcriptase-related proteins in top component particles of mammalian orthoreovirus. *Virology* 245, 33–46.
- Echols, H., and Goodman, M.F. (1991). Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.* 60, 477–511.
- Engh, R.A., and Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* 47, 392–400.
- Eom, S.H., Wang, J., and Steitz, T.A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. *Nature* 382, 278–281.
- Farsetta, D.L., Chandran, K., and Nibert, M.L. (2000). Transcriptional activities of reovirus RNA polymerase in reconstituted cores. Initiation and elongation are regulated by separate mechanisms. *J. Biol. Chem.* 275, 39693–39701.
- Gillies, S., Bullivant, S., and Bellamy, A.R. (1971). Viral RNA polymerases: electron microscopy of reovirus reaction cores. *Science* 174, 694–696.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876–1882.
- Hakansson, K., and Wigley, D.B. (1998). Structure of a complex between a cap analogue and mRNA guanylyl transferase demonstrates the structural chemistry of RNA capping. *Proc. Natl. Acad. Sci. USA* 95, 1505–1510.
- Hansen, J.L., Long, A.M., and Schultz, S.C. (1997). Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* 5, 1109–1122.
- Hu, G., Gershon, P.D., Hodel, A.E., and Quioco, F.A. (1999). mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. *Proc. Natl. Acad. Sci. USA* 96, 7149–7154.
- Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* 282, 1669–1675.
- Jacobo-Molina, A., Ding, J., Nanni, R.G., Clark, A.D., Jr., Lu, X., Tantillo, C., Williams, R.L., Kamer, G., Ferris, A.L., Clark, P., et al. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA* 90, 6320–6324.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47, 110–119.
- Joyce, C.M., and Steitz, T.A. (1994). Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* 63, 777–822.
- Kiefer, J.R., Mao, C., Braman, J.C., and Beese, L.S. (1998). Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* 391, 304–307.
- Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 256, 1783–1790.
- Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946–950.
- Kuznedelov, K., Korzheva, N., Mustaev, A., and Severinov, K. (2002). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stabil-

ity and is not involved in the maintenance of RNA-DNA hybrid length. *EMBO J.* 21, 1369–1378.

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

Lesburg, C.A., Cable, M.B., Ferrari, E., Hong, Z., Mannarino, A.F., and Weber, P.C. (1999). Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Biol.* 6, 937–943.

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89, 951–961.

Mindich, L. (1999). Precise packaging of the three genomic segments of the double-stranded-RNA bacteriophage phi6. *Microbiol. Mol. Biol. Rev.* 63, 149–160.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290.

Nibert, M.L., and Schiff, L.A. (2001). Reoviruses and their replication. In *Fields Virology*, D.M. Knipe and P.M. Howley, eds. (New York: Lippincott-Raven), pp. 1679–1728.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281–296.

Noble, S., and Nibert, M.L. (1997). Core protein mu2 is a second determinant of nucleoside triphosphatase activities by reovirus cores. *J. Virol.* 71, 7728–7735.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 313, 762–766.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*, C.W. Carter and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

Reinisch, K.M., Nibert, M.L., and Harrison, S.C. (2000). Structure of the reovirus core at 3.6 Å resolution. *Nature* 404, 960–967.

Shamoo, Y., and Steitz, T.A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* 99, 155–166.

Shatkin, A.J., and Sipe, J.D. (1968). RNA polymerase activity in purified reoviruses. *Proc. Natl. Acad. Sci. USA* 61, 1462–1469.

Skehel, J.J., and Joklik, W.K. (1969). Studies on the in vitro transcription of reovirus RNA catalyzed by reovirus cores. *Virology* 39, 822–831.

Starnes, M.C., and Joklik, W.K. (1993). Reovirus protein lambda 3 is a poly(C)-dependent poly(G) polymerase. *Virology* 193, 356–366.

Stukenberg, P.T., Studwell-Vaughan, P.S., and O'Donnell, M. (1991). Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* 266, 11328–11334.

Accession Numbers

Atomic coordinates for the structures described here have been deposited in the Protein Data Bank with following accession numbers: 1MUK (unliganded λ 3), 1MWH (λ 3 plus cap), 1N1H (initiation complex), 1N38 (short elongation complex), and 1N35 (long elongation complex).