Annu. Rev. Biochem. 2001. 70:181–208

REPLISOME-MEDIATED DNA REPLICATION

Stephen J. Benkovic, Ann M. Valentine, and Frank Salinas
Pennsylvania State University, Department of Chemistry, 414 Wartik Laboratory, University Park, Pennsylvania 16802; e-mail: sjb1@psu.edu; amv8@psu.edu; fgs4@psu.edu

Key Words leading strand, lagging strand, replication fork, polymerase, primosome

■ **Abstract** The elaborate process of genomic replication requires a large collection of proteins properly assembled at a DNA replication fork. Several decades of research on the bacterium *Escherichia coli* and its bacteriophages T4 and T7 have defined the roles of many proteins central to DNA replication. These three different prokaryotic replication systems use the same fundamental components for synthesis at a moving DNA replication fork even though the number and nature of some individual proteins are different and many lack extensive sequence homology. The components of the replication complex can be grouped into functional categories as follows: DNA polymerase, helix destabilizing protein, polymerase accessory factors, and primosome (DNA helicase and DNA primase activities). The replication of DNA derives from a multistep enzymatic pathway that features the assembly of accessory factors and polymerases into a functional holoenzyme; the separation of the double-stranded template DNA by helicase activity and its coupling to the primase synthesis of RNA primers to initiate Okazaki fragment synthesis; and the continuous and discontinuous synthesis of the leading and lagging daughter strands by the polymerases. This review summarizes and compares and contrasts for these three systems the types, timing, and mechanism of reactions and of protein-protein interactions required to initiate, control, and coordinate the synthesis of the leading and lagging strands at a DNA replication fork and comments on their generality.

CONTENTS

INTRODUCTION

The replication of duplex DNA by an assembly of enzymes and proteins that collectively constitute the replisome is an outstanding example of a finely crafted biological machine. Perhaps its greatest challenge is to synthesize simultaneously two strands of DNA with opposite polarity because the DNA polymerases act unidirectionally, extending a primer in a $5' \rightarrow 3'$ fashion. Consequently, leading strand synthesis is highly processive, whereas lagging strand synthesis is discontinuous, marked by the synthesis of 1- to 3-kilobase (kb) Okazaki fragments. The problem then is how to coordinate the movement of both polymerases; the ostensible answer is through the replisome.

This review explores the numbers and nature of the proteins composing the replisome of three species and is framed around questions of replisome function. How is the replisome assembled from its constitutive proteins? When accessory proteins such as clamp loaders and clamps are needed, how are they loaded onto single-stranded DNA to create the holoenzyme? What protein-protein interactions and DNA contacts are manifest? For the replication fork to move, single-stranded DNA must be extended by a helicase; does it set the rate of fork movement? For lagging strand synthesis, multiple initiation events by a primase are required; what controls their frequency and how is the short oligoribonucleotide primer captured by the polymerase before dissociation? Obviously, two polymerases or two holoenzymes are in operation; are they in contact directly or through other proteins? Do the lagging strand polymerase or associated proteins dissociate at the completion of each Okazaki fragment synthesis or are they recycled? What process(es) set the length of Okazaki fragments? What is the protein and DNA topology at the replication fork? We do not, in this review, consider how the primers are eventually excised and the fragments ligated to form a continuous lagging strand.

Early models connected the polymerization events at the leading and lagging strand by looping the lagging strand back through the replisome—the trombone model (1). A loop has remained a constant feature of subsequent models that differ primarily in the relative orientation of the two polymerases and contacts between other members of the replisome (2). Dimerization of the two polymerases is also anticipated in this model, and evidence for this particular interaction has been a common objective in many studies (3). Much of our present understanding of the replisome and its mode of action has come from investigations of the T4 and T7 bacteriophage and *E. coli* systems, and these will be the focus of our review. The

literature is voluminous, punctuated by papers with unexpected observations and a few with false leads. We hope our perforce limited selection conveys an accurate and adequate summation of the present state of understanding.

THE T7 REPLISOME

The Proteins

The advantage of the T7 replication system lies in its reconstitution by only four proteins: the T7 gene 5 protein (gp5), a DNA polymerase that forms a tight 1:1 complex ($K_d \approx 5$ nM) with *E. coli* thioredoxin (4, 5); the T7 gene 4 protein (gp4), a multifunctional enzyme that contains both helicase and primase activities (6, 7); and the T7 gene 2.5 protein (gp2.5), a single-stranded DNA binding protein (8, 9). Following literature precedent, we use T7 DNA polymerase to refer to the above complex with thioredoxin.

The 80,000-molecular-weight (80K) gp5 is a distributive enzyme in the absence of thioredoxin, incorporating 1–15 nucleotides before dissociating from the primer template; in the presence of thioredoxin, the enzyme is highly processive, incorporating thousands of nucleotides during a single pass. A crystal structure of the polymerase complexed to thioredoxin, a primer template, and a nucleoside triphosphate resembles that of other members of the polymerase I (Pol I) family that are shaped like a right hand in which the palm, fingers, and thumb form a DNA-binding groove (10, 11). Thioredoxin binds to an extended loop on the thumb, creating a flexible tether that could swing across the DNA-binding groove to encircle the primer-template exiting the polymerase; thus, it functions akin to the multimeric clamp proteins to facilitate processivity.

The 63K T7 helicase-primase protein is a member of the family of hexameric helicases with five conserved sequence motifs that includes *E. coli* DnaB (12). Its approximate shape and subunit arrangement were first revealed by electron microscopy to be a topologically closed hexamer that encircles one strand of DNA (13). The three-dimensional reconstruction from the images showed two stacked rings, one small and one large, capable of accommodating 25–30 nucleotides, which are associated with the primase and helicase activities. A shorter 56K protein that begins at a second initiation codon has only helicase activity (14–16). A crystal structure of a C-terminal helicase domain that includes all five of the conserved helicase motifs (residues 272–566) revealed a six-fold symmetric ring 120 Å in diameter with a central hole of 35 Å, closely matching that seen in the electron microscopic reconstruction (17a). A bound nucleotide (dTTP or ATP) is found at the subunit interface, providing insight into the origin of the cooperative binding found for assembly of the hexameric helicase from monomers and in its cooperative binding and hydrolysis of the ATP (18–20) that fuels its movement. Like the homologous RecA and F_1-ATPase protein, the oligomerization domain is adjacent to the amino terminus of the nucleoside triphosphate (NTP)-binding domain (21).

The 26K gene 2.5 protein binds specifically to single-stranded DNA ($K_d \approx 0.4\ \mu M$) with a site size of approximately seven nucleotides per monomer (8). In solution, it exists as a dimer. Genetic analysis revealed that gene 2.5 protein is essential for DNA replication in the host *E. coli*; phage DNA synthesis is reduced to a level of less than 1% for a T7 phage lacking this gene product relative to wild-type T7 phage (22). The gp2.5 protein interacts in a 1:1 stoichiometry with the DNA polymerase:thioredoxin complex ($K_d \approx 1\ \mu M$) as well as with both the 56K and 63K forms of gp4 (23).

A truncated gp2.5 protein that lacks the 21 C-terminal amino acids of the wild type retains its ability to bind to single-stranded DNA but no longer physically interacts with the polymerase. The truncated protein cannot substitute for wild-type gp2.5 in vivo, nor can wild-type *E. coli* single-strand binding (SSB) protein substitute for this mutant in vivo. A comparison of the amino acid sequences of T7 gene 2.5 protein, *E. coli* SSB protein, and T4 gene 32 protein implicates an amino-terminal domain as responsible for DNA binding in the SSB (24) and gp32 proteins (25).

Synthesis at the Leading Strand

Leading strand synthesis has been investigated using a topologically stable replication fork represented in Figure 1 (26). In the absence of gp4 protein, the T7 polymerase catalyzes no net strand displacement. In its presence, provided a single-stranded 5′ tail exists in the DNA duplex substrate to serve as an entry route for

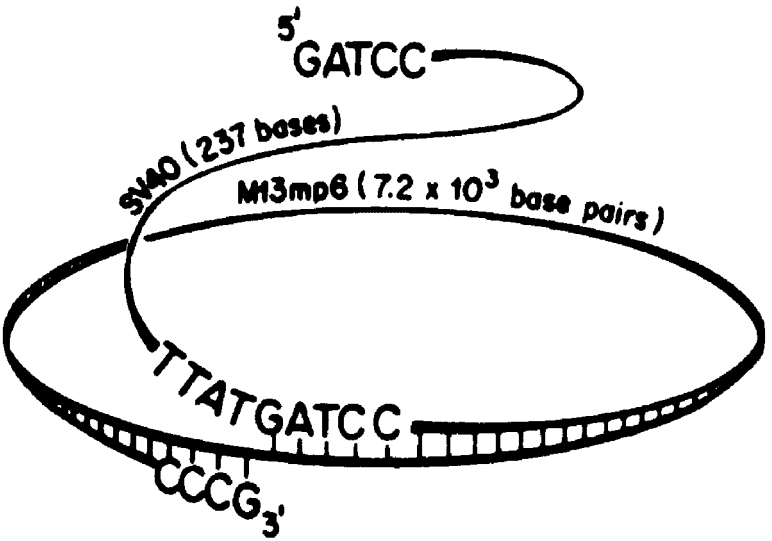


Figure 1 A preformed topologically stable replication fork derived from M13. (From Reference 26.)

the helicase, T7 polymerase catalyzes the polymerization of tens of thousands of nucleotides at a rate of approximately 300 nucleotides s^{-1}. The specificity of the gp4-gp5 interaction is underscored by the inability of gp4 to catalyze strand displacement synthesis by polymerases from *E. coli* and T4. A ternary complex consisting of single-stranded DNA, gp4, and gp5 proteins formed in the presence of a nucleotide analog (e.g. β,γ-methylene dTTP), and Mg^{2+} is stable to gel filtration and supports strand displacement synthesis (27).

Synthesis at the Lagging Strand

The discontinuous replication of the lagging strand requires RNA-primed DNA synthesis; the necessary primers are generated by the primase activity of gp4. The in vivo and in vitro primer sequences are pppA (C) $(N)_{2\text{-}3}$ with the predominant DNA recognition sequence being 3′-CTGGG-5′, 3′-CTGTG-5′ or 3′-CTGGT-5 (6, 28). The 3′ cytosine residue is cryptic, required for recognition but not copied into primer. The unique 7K N terminus of the 63K gp4 protein contains a Cys_4 metal binding motif and binds one equivalent of Zn that contributes to recognition of the template priming sequence. Cys to Ser mutants of the Zn binding motif have a lower Zn content and lead to loss of template-directed but not template-independent synthesis of oligoribonucleotides (29). Both chimeric primases whose Zn motifs were derived from those found in the *E. coli* and T4 primases hydrolyze dTTP at their helicase site but are drastically impaired in catalyzing template-dependent priming. These chimeric proteins, however, do not use the recognition sequence of any parent, which implicates other regions of the gene 4 protein besides the zinc-binding domain as recognition site determinants (30). The role of the C-terminal amino acids is revealed by the behavior of a truncated gene 4 protein containing only the N-terminal 271 amino acids; it is dimeric, devoid of helicase and associated dTTPase activities, but catalyzes the template-directed synthesis of di-, tri- and tetranucleotides at rates similar to wild-type protein (31). Neither this truncated protein nor one in which the C-terminal 17 amino acids have been deleted can interact with the polymerase to couple primer synthesis with DNA synthesis (32).

The 56K T7 helicase interacts asymmetrically with both single-stranded tails of forked DNA, binding about 25 nucleotides of the 5′ strand through the hole in the enzyme. Optimal DNA unwinding rates require contact between the 3′ strand and a contact region estimated to be on the outer surface of the hexamer approximately 10–15 nucleotides from the fork junction, a value estimated by varying the length of the 3′ tail (33). With an optimal substrate, the T7 DNA helicase can move unidirectionally 5′ to 3′ on the lagging strand template fueled by NTP hydrolysis at a rate approaching 260 base pairs (bp) per second (260 bp s^{-1}, at 30°C), very close to that noted earlier for leading strand synthesis. The mechanism of unwinding is probably via exclusion of the 3′ tail from the central hole (34).

Lagging strand DNA synthesis has been reconstituted on ϕX174 single-stranded DNA in the presence of T7 DNA polymerase and the T7 helicase-primase (35).

Lagging strand synthesis has also been demonstrated on the preformed replication fork derived from M13 (Figure 1) and in contrast to leading strand synthesis is inhibited by the addition of a challenger DNA trap, as well as by dilution. The Okazaki fragments that are 2–6 kb in length are sensitive to the level of gp4 protein but not the T7 DNA polymerase (36). Inclusion of gp2.5, single-strand binding protein, increases the frequency and efficiency of primase initiation some 10-fold at gp2.5 levels that require gp4 to bind to protein-coated single-stranded DNA.

A key insight into the geometry of access to the primase recognition site was obtained by fixing gp4 to single-stranded DNA in the presence of β,γ-methylene dTTP. Under these conditions, the protein does not translocate, nor does it readily dissociate to be captured by challenger DNA. Nevertheless, the primase can function to catalyze primer formation at distal recognition sites on the same or other single-stranded DNA molecules (37), which suggests that the primase active site faces away from the DNA threading through the helicase (Figure 2). This would foster a direct hand-off of the primer to the lagging strand polymerase. The affinity of gp4 for DNA arises in part from within the helicase domain (39) that requires 5–10 nucleotide flanking sequences. Importantly, the synthesis of the primer proceeds at a rate of $\sim$1 s^{-1}, sufficient to satisfy the time requirements for Okazaki fragment synthesis (1–2 s) (30, 40).

Coordinated Synthesis of Leading and Lagging Strands at a Replication Fork

A key step that relates replication fork movement to lagging strand synthesis is the priming of the Okazaki fragments and its effect on helicase travel. On minimal replication forks consisting of two synthetic oligonucleotides annealed to form a 30-bp duplex and two single-stranded tails of 22 and 40 nucleotides at the 3′ and 5′ ends respectively, the helicase activity in terms of fork movement of gene 4 protein is inhibited when the primase activity of gp4 is activated at a primase recognition site. In the absence of primase recognition sites, the addition of NTPs has no effect (41).

With a preformed replication fork similar to that in Figure 1, the presence of gp4 and T7 DNA polymerase plus the complement of NTPs and deoxynucleoside triphosphates (dNTPs) leads to leading strand products greater than 40 kb and Okazaki fragments of 0.5–6 kb. An increase in lagging strand DNA synthesis caused by increased levels of priming nucleotides results in a corresponding decrease in the amount of leading strand synthesis and a drop in the rate from 320 to 250 nucleotides s^{-1}, which is consistent with the two activities being coupled and the priming rate influencing replication fork movement. The synthesis of both strands is resistant to 20-fold dilution, which is consistent with the recycling of both the initially bound helicase-primase and polymerase proteins. (The contrary previous result was attributed to the presence of the 56K gp4.) Confirmatory evidence for coordinated synthesis is that the processivity of leading and lagging strand synthesis is resistant to template challenge experiments, and that

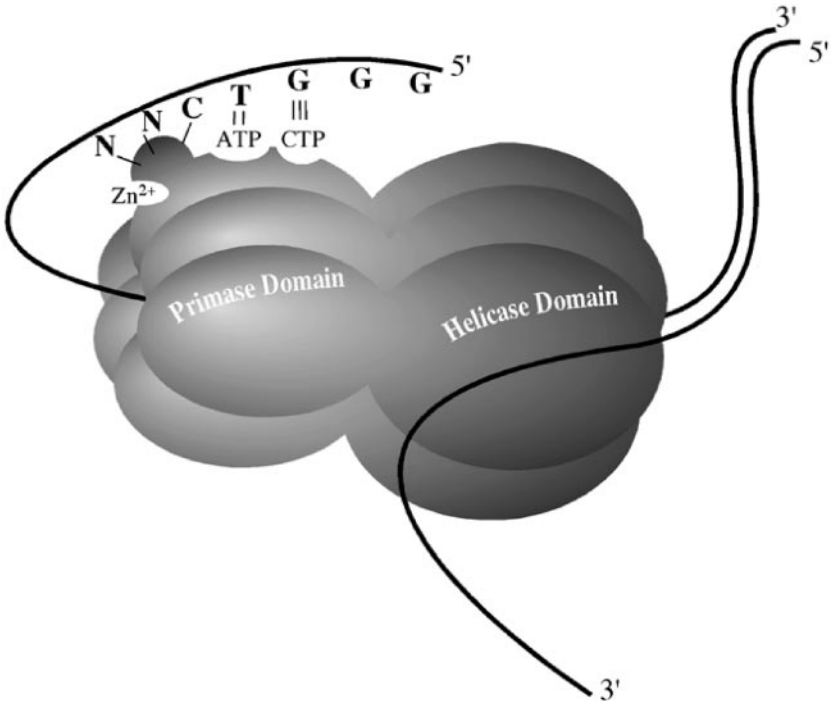


Figure 2 Diagram showing DNA-protein interactions between T7 gene 4 protein and the primase recognition site. When viewed in the electron microscope, the T7 protein forms a hexamer of bilobal subunits that surrounds single-stranded DNA (13). Interactions between single-stranded DNA and the helicase domain likely occur on the inner and outer surface of the hexamer (33, 38). The primase domain is shown interacting with the primase recognition site via contacts between bound nucleoside triphosphates and the Cys_4 zinc ribbon. The zinc-binding domain and the ATP- and CTP-binding sites are shown on the outside of the hexamer for clarity. In the absence of a high-resolution structure of the gene 4 protein, attempts to localize the protein domains involved in the helicase or primase activities are speculative. (From Reference 39.)

the average length of Okazaki fragments remains constant at around 2 kb at all levels of the polymerase. Although dimeric forms of T7 polymerase have not been demonstrated, the data require two such enzymes at the replication fork, possibly coordinated through the primase and helicase domains of the gene 4 protein.

A more quantitative analysis of events at a functioning replisome was obtained by examining a functioning replisome on a replication fork composed of a small, circular duplex DNA molecule (70 bp) bearing a 5′ single-stranded tail of 40 nucleotides (nt) (Figure 3) (42). The lagging strand contains two primase recognition sites, and because of the sequence of the construct, lagging strand synthesis can be identified and measured by the incorporation of dCMP and leading stranding

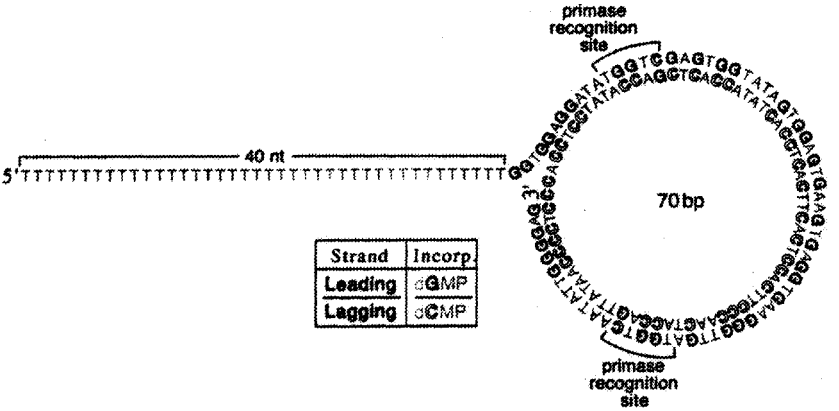


Figure 3 A minicircle with a replication fork. It consists of a double-stranded DNA circle bearing a 5′ single-stranded DNA tail to which the gp4 can bind. Two recognition sites (5′-TGGTC-3′) at which the gp4 primase can catalyze the synthesis of RNA primers are present on the lagging strand. One strand of the minicircle contains 33 cytosine residues but only 2 guanine residues, whereas the other contains 33 guanine residues and 2 cytosine residues. Consequently, leading and lagging strand synthesis can be identified and quantitated by measuring the incorporation of dGMP and dCMP, respectively. (From Reference 42.)

synthesis by dGMP. In addition, its small size provides for a molar excess of forks over the replication proteins, typically minicircle:gp4:DNA polymerase in a ratio of 10:1:8. Sufficient gp2.5 was also present to coat any single-stranded DNA formed during the reaction. Direct monitoring of the rate of leading and lagging synthesis revealed that both proceeded at the same rate provided gp2.5 was present. In the absence of gp2.5 or with the deletion mutant gp2.5 Δ21C, which lacks the acidic carboxyl tail, the rate of lagging strand synthesis is fourfold less than that of leading strand synthesis. Addition of a trace amount of dideoxy CTP (ddCTP) inhibits both lagging and leading strand synthesis concurrently. In short, the system is coupled.

The question of whether the polymerase responsible for Okazaki fragment synthesis is recycled was addressed by using T526F T7 DNA polymerase. This form of the enzyme discriminates against incorporation of dideoxy nucleotides. Initiating replisome synthesis with the T526F form of the polymerase in the presence of wild-type enzyme and ddGTP or ddCTP does not terminate either leading or lagging strand synthesis, which is consistent with recycling of the lagging strand polymerase and the high processivity of both polymerases.

In order to recycle the lagging strand polymerase, one anticipates that the polymerase at the 3′ end of the completed Okazaki fragment is juxtaposed to the new priming site. This positioning requires the looping out of an extent of single-stranded DNA connecting these two sites (42, 43). Electron microscopy of the minicircle replication system revealed the presence of duplex replication loops at

the end of linear duplex molecules for ~50% of the replicating molecules. The loops ranged in length from 0.1 to 2.1 kb and were associated with a replisome. Further analysis revealed that many of the loops contained stretches of single-stranded DNA separated by a duplex segment. These apparently represent the in-progress synthesis of an Okazaki fragment because the sum of the duplex and one single-stranded region is ~3 kb, the average length of an Okazaki fragment. In the absence of gp2.5, there are many more single-stranded regions. The role of gp2.5 may be multifold: tethering the DNA to the helicase, recruiting the lagging strand holoenzyme to the primer site, slowing the helicase to trigger primase priming through its accumulation on the helicase-generated single strand, and mediating loop formation. All are aspects of a speculative model proposed for coordinated synthesis by the T7 replisome (42).

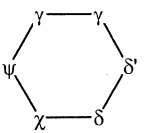
THE *E. COLI* REPLISOME

Protein Components

Compared to the T7 replisome, that of *E. coli* is more complex in terms of the numbers of proteins required. The polymerase III holoenzyme alone consists of 10 unique subunits, whose composition has been reviewed (44–48). The holoenzyme contains an isolable core polymerase ($\alpha\varepsilon\theta$) derived from three polypeptides: α, the polymerase catalytic subunit; ε, the 3′-5′ exonuclease; and θ, an accessory unit that binds to ε, stimulating its editing function. A second, stable subassembly adds τ protein, derived from the *dnaX* gene, which bridges two α units to form a dimeric core $(\alpha\varepsilon\theta)_2\tau_2$. The carboxyl terminus of τ appears to be necessary for binding to α as well as to the DnaB helicase (49, 50). A third subassembly is the γ complex composed of five proteins: $\gamma_4\delta\delta'\chi\psi$ in the stoichiometry shown. If the core has as presumed a two-fold axis of symmetry, binding of γ complex to core would introduce asymmetry.

The interaction between these various proteins and their enzymatic activities are listed in Table 1. The γ protein, also derived from the *dnaX* gene, is a truncated version of τ lacking the latter's carboxyl-terminal sequence. Both γ and τ proteins possess DNA-dependent ATPase activity; however, the γ complex rather than the corresponding τ complex (τ substituted for γ) acts as a clamp loader to load β, a clamp protein, onto DNA through the δ subunit (51, 52). The combination of these three subassemblies in an ordered fashion leads to the Pol III polymerase with the composition $(\alpha\varepsilon\theta)_2\tau_2\gamma_2\delta\delta'\chi\psi$, a polymerase that lacks high processivity in the absence of β protein (53). The addition of two β units creates the Pol III holoenzyme with an overall molecular weight of 900,000. The clamp protein, β, is a dimer that slides freely on duplex DNA in both directions. As anticipated from the binding properties of β, X-ray crystallography revealed that β is indeed in the shape of a ring with a central cavity of ~35 Å, large enough to accommodate duplex DNA, and a width of approximately one turn of DNA. The cavity is presumably insulated from contact with DNA by a sheath of water molecules, which permits its

TABLE 1 Composition of DNA polymerase III holoenzyme[a]

Subunit	Mass (kilodaltons)	Function	Contacts in subassemblies[b]
α	130	DNA polymerase	α–ε–θ
ε	27.5	3′, 5′ exonuclease	
θ	8.6	Stimulates exonuclease	
τ	71	Stimulates helicase, dimerizes core	α–ε–θ θεα–τ–τ–αεθ γ–γ
γ	47.5	ATP-requiring clamp loader	
δ	38.7	Accessory protein, binds to β	
δ′	36.9	Accessory protein, stimulates ATPase of γ	
χ	16.6	Accessory protein, binds SSB	
ψ	15.2	Accessory protein, with χ increases affinity for γ	
β	40.6	Clamp protein,	

[a]Adapted from References 48, 50a.

[b]K_d values for contacts are generally <100 nM.

facile diffusion along the DNA duplex and dissociation at its ends. Topologically similar sliding clamps have been characterized by X-ray crystallography, including the proteins gp45 (T4) and proliferating cell nuclear antigens (PCNAs) (yeast, humans) (54–56).

DnaB, 314K, is the primary replicative DNA helicase in *E. coli* essential for replication (57). The protein as visualized by electron microscopy exists as a stable ring-shaped hexamer that can adopt in the presence of nucleoside di- or triphosphates two different conformational states (C-3 or C-6 symmetry) as a consequence of asymmetric dimers within these rings (58–60). The conformational flexibility is also manifest kinetically in the sequential isomerization of fluorescently tagged DnaB upon binding nucleotide cofactors in which the first three cofactors add independently with high affinity and the remaining three bind with negative cooperativity (61). Fluorescent energy transfer experiments indicate that single-stranded DNA binds to DnaB predominantly through a single subunit, and as with the T7 gp4, it passes through the cross-channel of the hexamer (62). The helicase, in the presence of nonhydrolyzable ATP analogs, preferentially binds to the 5′ area of forked substrates (63) consistent with its 5′ → 3′ directionality when unwinding duplex DNA powered through ATP hydrolysis (7; reviewed in 64).

The *E. coli* 65K DNA primase, DnaG protein, initiates RNA synthesis at a 5′-CTG-3′ sequence within a single-stranded DNA template to produce primers of 8–12 nucleotides for lagging strand synthesis (65, 66). A crystal structure of the RNA polymerase domain (36K core region, residues 111–433), which retains low polymerase activity, reveals a cashew-shaped molecule characterized by a shallow, wedge-shaped cleft on the concave side of the protein measuring ~9 Å at one end and 20 Å at the other. The single-stranded DNA is imagined to thread $3' \rightarrow 5'$ through the narrower opening, with the synthesis of the RNA:DNA duplex extending into the wide, shallow depression. Whether this primase binds the single-stranded DNA as it passes through the channel of DnaB or requires it to loop back as in Figure 2 for the T7 gp4 helicase-primase is presently not definitively established (67).

Interaction of DnaG with DnaB is mediated by the carboxy-terminal 16 amino acids (68), resulting in $>10^3$-fold activation of the primase activity and targeting that action to be mainly at the replication fork (69, 70). The interaction with helicase also appears to broaden the primase recognition sequence to favor 5′-CAG-3′ (71). This binding of DnaG/DnaB is critical for the recruitment of the primase to the replication fork. DnaG protein tightly binds one Zn, which is ligated by three cysteinyl sulfhydryls and one histidine nitrogen near the N terminus (72, 73). Apoprimase, however, synthesizes primers in a sequence-specific manner, which eliminates a role for Zn in catalysis (73) and ascribes to it a structural one, possibly to prevent disulfide formation. Primer synthesis by DnaG outside the context of a replication fork, however, is too slow to prime the Okazaki fragment synthesis observed with the replisome.

The single-strand binding protein (SSB) from *E. coli* (reviewed in 74) forms a tetrameric structure that binds single-stranded DNA, the χ subunit of Pol III holoenzyme, and DnaG, and is essential for coupled leading and lagging strand synthesis (75). In addition to DnaB, DnaG, and SSB, primosome assembly requires five other protein components, the products of the *dnaC*, *dnaT* (protein i), *priA* (protein n′), *priB* (protein n), and *priC* (protein n′) genes (76). The stepwise assembly of these seven proteins provides the primosome unit required for Okazaki fragment synthesis (76, 77), with DnaC acting like the T4 gp59 to recruit the helicase to the fork (see T4 section below) (78).

Leading Strand Synthesis

The core polymerase $(\alpha\varepsilon\theta)_2\tau_2$ synthesizes DNA at a rate of approximately 20 nucleotides s^{-1} with a processivity in the tens of nucleotides (79). When Pol III holoenzyme is associated with β clamp protein, the overall rate increases to ~750 nucleotides s^{-1} and to a processivity in excess of 50,000 nucleotides (77, 80). The processivity is totally dependent on the presence of β clamps.

The assembly of the clamp on DNA by the γ complex follows a sequential process: (*a*) binding of ATP to the γ complex changes its conformation, exposing the δ subunit to bind β (81); (*b*) binding of the γ complex to β protein opens the

β clamp prior to ATP hydrolysis (the interaction with δ is sufficient to open the β clamp); (*c*) hydrolysis of one equivalent of ATP by the DNA $\cdot$ γ complex $\cdot$ β assembly acts either to release the β clamp to close it on the DNA or to position the duplex within the interior of the β clamp; and (*d*) hydrolysis of the second ATP equivalent causes dissociation of the γ complex from DNA, leaving β behind. These events are inferred from rapid kinetic studies and the interception of different conformational states; direct measurements of the similar ring opening by the gp45 protein in the T4 replisome discussed below permits a further mechanistic distinction (82). Nevertheless, the overall rate of clamp loading exceeds that required by the time frame of the Okazaki fragment synthesis (see below).

Both the δ subunit of the γ complex and the α subunit of the core inhibit the phosphorylation of a kinase recognition sequence, engineered near the C terminus of β, indicating in conjunction with protease mapping and site-specific mutagenesis that both bind β on the same face with overlapping sites (83). The affinity of β for γ complex is some 30-fold greater than for the core; however, in the presence of primed M13 mp18 (a derivative of M13 with a multiple cloning site) this affinity is reversed, with β preferring the core. These changes create a molecular switch during replisome action. The interaction of β with γ complex initially drives β loading onto duplex DNA followed by its replacement by core polymerase. The binding of the latter blocks the ability of the γ complex to unload β, which permits DNA synthesis to be finished and then β to be recycled via removal by the γ complex. There is a structural parallel to other systems that use other ringlike clamp proteins, gp45 and PCNA (56), whose assembly is also an ATP-driven process involving clamp loaders that bind to the same face of the clamp protein as their corresponding polymerases (84).

Coordinated Synthesis of Leading and Lagging Strands

Leading and lagging strand synthesis has been reconstituted at a tailed-form II DNA template as shown in Figure 4 in the presence of the Pol III holoenzyme, SSB, and the primosome, producing multigenome-length (−) strand DNA and multiple cycles of discontinuous (+) strand DNA (85). In the absence of primase, only leading strand synthesis was observed. The omission of DnaB, DnaC, and DnaT virtually eliminated rolling circle synthesis; the absence of protein n′ and n markedly reduced synthesis; and the presence of protein n′′ stimulated DNA synthesis. The reconstituted replication fork moves at a rate of 600–800 nucleotides s^{-1} (30°C), producing leading strands of 150–500 kb in length and Okazaki fragments of 1–2 kb, which requires the latter to be initiated every 1–2 s. Alternatively, rolling circle synthesis could be reconstituted from Pol III holoenzymes, DnaB, DnaC, and primase, although the number of active replication forks (DnaB forks) was reduced threefold (85).

In order to determine which of the replisome proteins remain continuously associated with the replication fork, the dependence of Okazaki fragment length on the concentration of a particular protein was measured. For a distributively acting

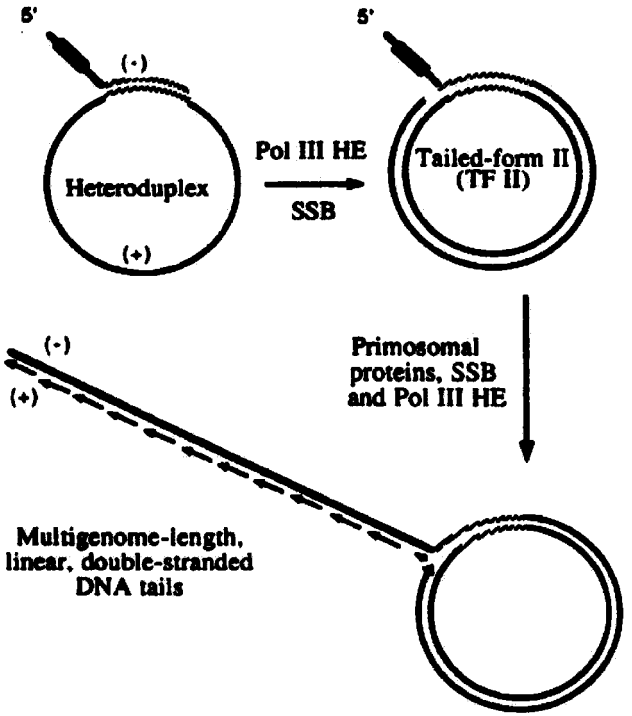


Figure 4 A rolling circle type of DNA replication system. Pol III HE is a polymerase III holoenzyme. (From Reference 85.)

protein, the length of the Okazaki fragment should vary inversely with protein concentration. All the primosomal proteins with the exception of the primase acted processively. Okazaki fragment synthesis on DnaB forks was sensitive to primase dilution, which suggests that primase reenters the replication fork complex via interactions with the helicase, DnaB. The β clamp protein also behaved distributively and catalytically, i.e. it was reused, presumably through its unloading by the γ complex described earlier.

Further evidence that the primase acts distributively was furnished by the observation that increases in NTP concentration decreased Okazaki fragment size owing to stimulation of the rate of primase association with the replication fork (86). Changes in dNTP levels that increase competition with NTP binding to the primase also affected Okazaki fragment length, but not the rate of fork movement. On the other hand, the behavior of β upon dilution to likewise increase Okazaki fragment length was traced to a decrease in the efficiency of initiation of DNA synthesis from primer termini but not priming frequency. Collectively these observations are consistent with the hypothesis that events at the new primer termini are

independent of the action of the DNA polymerase complex bound to the penultimate Okazaki fragment. Because the frequency of primer initiation is unchanged, the lagging strand polymerase either displaces unused primers or bypasses them (87).

Primase in the presence of the helicase on single-stranded ϕX174 DNA demonstrates an inherent size-determining mechanism (88) that sets the length of most primers at around 12–14 nucleotides, which is consistent with the active-site cavity in the X-ray structure with less abundant populations between 30 and 70 nucleotides. A similar distribution was observed with the primosome on the tailed-form II DNA template (Figure 4). Presence of either Pol III holoenzyme or Pol III core alters the primer population to be predominantly nine or fewer nucleotides in length, which supports a specific interaction with primase. On an active replication fork, primer lengths are unchanged even when the efficiency of their extension is reduced threefold by reducing β protein. These data suggest that protein interactions with clamp and clamp loader proteins occur subsequent to binding of primase to the Pol III core.

Guided by the changes in fragment size in response to alterations in reaction parameters, a cycle of Okazaki fragment synthesis has been proposed: (*a*) Primase binds to the helicase at the replication fork; (*b*) primase initiates primer synthesis on the lagging strand; (*c*) a primase-core interaction is established that limits the length of primer synthesis and exposes the 3′ end of the primer; (*d*) a β clamp from solution is loaded by the γ complex subunit onto the primer template junction; (*e*) the Pol III holoenzyme transits from the previously completed Okazaki fragment to the new primer terminus; and (*f*) primase releases and polymerase synthesizes the new fragment.

The question of whether other proteins remain within the replisome during repeated cycles of Okazaki fragment synthesis was answered by reconstitution of the replisome from its individual subassemblies: Pol III core, γ complex, and β. The various protein units were subject to dilution during lagging strand synthesis on the tailed-form II fork in the presence of SSB. Both the γ complex and the core polymerase of the lagging strand remained associated during multiple cycles (89, 90). The physical coupling required for the coordination of leading and lagging strand synthesis is attributed to the τ subunit, which bridges the two active core assemblies (91). The τ subunit, which also physically associates with the DnaB helicase, acts to increase the unwinding rate of the helicase alone, so that the fork moves at the speed of the polymerase (91).

The signal that sets the cycle of Okazaki fragment synthesis in motion then is step *a* in the cycle, the entry of the primase into the replisome proteins at the fork (92). Convincing evidence in support of this conjecture was derived from a series of mutant primases that retain their capacity to synthesize primers as well as to interact with the Pol III holoenzyme (93). These mutations in the C-terminal eight amino acids of primase are involved in interaction with the helicase. Changing the concentration of a mutant primase led to changes in Okazaki fragment size that correlated with the affinity of the primase for the helicase. Neither the efficiency

of primer utilization nor the rate of replication fork movement was affected. Collectively, the cumulative data argue against models in which the lagging strand holoenzyme triggers synthesis of a new primer.

A model for how the proteins at the replication fork are spatially coordinated to satisfy the detected contacts and the kinetic properties of the replisome is provided in Figure 5. The leading and lagging strand polymerases are oriented in a parallel manner. The helicase, DnaB, and the γ complex contact the τ subunit, providing the latter with easy access to the 3′ terminus of the primer and an incoming β clamp. This close proximity is also necessary for the χ subunit (not shown) of the γ complex to displace primase from its RNA product for recycling (94). The χ subunit is also the main contact between the holoenzyme and SSB (95) and, in another example of a molecular switch, its higher affinity for SSB may lead to the release of the primase from its contact to SSB within the replisome. The γ complex then can load β to sequester the primer and form

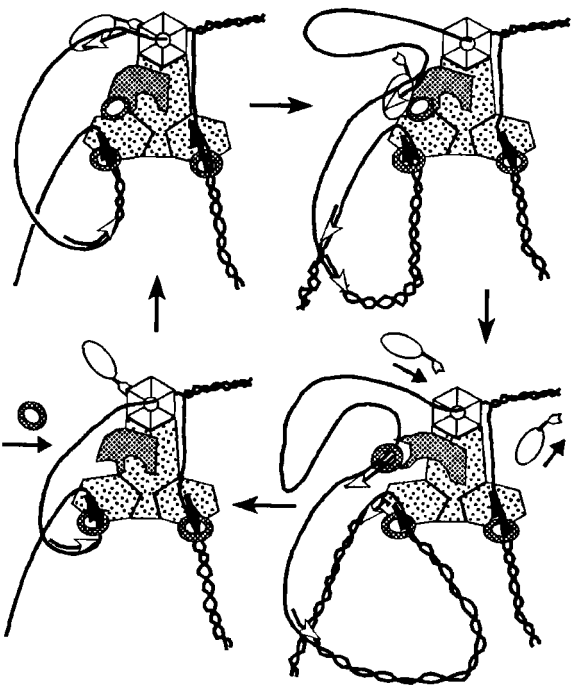


Figure 5 A model of a proposed replication cycle that depicts how protein-protein interactions between primase and the holoenzyme could localize the new primer to the fork at all times. Clear hexagon, DnaB; ring shape, β; spotted hexagons, core; clawlike irregular shape, γ complex; spotted irregular shape, τ dimer bridging the leading and lagging strand polymerases; tailed oval shape, primase; arrows, primers. For clarity, SSB has been omitted. The drawing is not to scale. (From Reference 93.)

the preinitiation complex needed to commence synthesis of the next Okazaki fragment.

THE T4 REPLISOME

Proteins

The eight protein components of the T4 replication system have been reviewed recently (2, 96–98), including the RNase H and ligase enzymes needed to seal the nick between adjacent Okazaki fragments.

The 110K T4 DNA polymerase, the product of phage gene 43 (gp43), contains both 5′-3′ polymerase and 3′-5′ exonuclease activities and functions both as the leading and lagging strand polymerase. The polymerase from phage RB69, a phylogenetic variant of the T4 family, has ∼64% identity of amino acid sequence with its T4 homolog, and its crystal structure exhibits the anatomical analogy (thumb, fingers, palm) found in the B family of DNA polymerases (99–101). In the absence of an accessory protein, the polymerase is primarily distributive, adding nucleotides at a rate of ∼500 bases s^{-1} (102, 103).

Three accessory proteins, gp45, gp44, and gp62, increase the processivity of the polymerase on primed single-stranded or forked duplex templates (104, 105). The gp44 and gp62 proteins (35K and 21K as subunits) copurify as a tight 44/62 complex with a 4:1 stoichiometry (106) and function as a clamp loader in an ATP-fueled process (107, 108). The protein gp45 (74.4K) was revealed by its crystal structure as a circular clamplike molecule capable of encircling DNA (109); it has an internal diameter of 35 Å and a ring thickness of 25 Å, a result anticipated by earlier studies (110). Consequently, gp45, *E. coli* β protein, and yeast PCNA share a common structural topology and function despite low sequence homology (55, 56).

The T4 primosome consists of the helicase, gp41, and the primase, gp61. The 54K helicase exists as a dimer but assembles into a ring-shaped hexamer of asymmetric dimers upon binding GTP, ATP, or ATPγS (111, 112), reminiscent of the *E. coli* and T7 helicases. The 40K primase, gp61, forms a weak complex with helicase, but in the presence of single-stranded DNA it participates in a stable ternary complex (half-life of $>$20 min) composed of a six-helicase:one-primase subunit stoichiometry (113). There is no evidence whether the primase is centrally located or shuttles from one subunit to another as the primosome functions. Mapping of the helicase-primase interaction in the presence of single-stranded DNA indicated local conformational changes within the primase that may expose the recognition site of the primase for DNA template sequences (114).

A fourth accessory, the 26K monomeric protein gp59, forms complexes with the gp41 helicase (1:1 stoichiometry based on subunits) through the C terminus of gp41 (115) and with single-strand binding protein, gp32, through its acidic C-terminal domain (residues 254–301) (116–119). This protein binds single-stranded or duplexed DNA with either a 5′ or 3′ single-strand extension, but it

has a higher affinity for forked DNA (115). A high-resolution X-ray structure has revealed a novel α-helical two-domain fold with sites that could simultaneously bind gp32, forked DNA, and gp41 in order to mediate the rapid assembly of the helicase on gp32-coated DNA (115) as proposed (120).

The 33K protein product of gene 32 is the T4 single-stranded DNA binding protein and is analogous in function to those of T7 and *E. coli*. It is essential for replication in vivo (121) and in vitro (122). gp32 binds cooperatively to single-stranded DNA and may cluster through translocation along the strand (123). The affinity of gp32 for single-stranded DNA is due in part to the presence of a tightly cysteinyl bound Zn as well as the side chains of aromatic amino acids in the core region (residues 22–253) (124, 125), as substantiated by a crystal structure of the gp32·DNA complex (126). In addition to its interaction with gp59, gp32 associates with the polymerase and the primase, presumably through the carboxyl terminus (127, 128), underscoring some of its plethora of interactions within the replisome.

Leading Strand Replication

Processive DNA synthesis requires the formation of a holoenzyme complex derived from the T4 DNA polymerase, gp43, and the accessory proteins, gp44/62 and gp45. The kinetics of holoenzyme formation has been defined by rapid-quench and stopped-flow fluorescent kinetics on the replication fork shown in Figure 6 (129, 130). The assembly follows an ordered process in which the clamp trimer

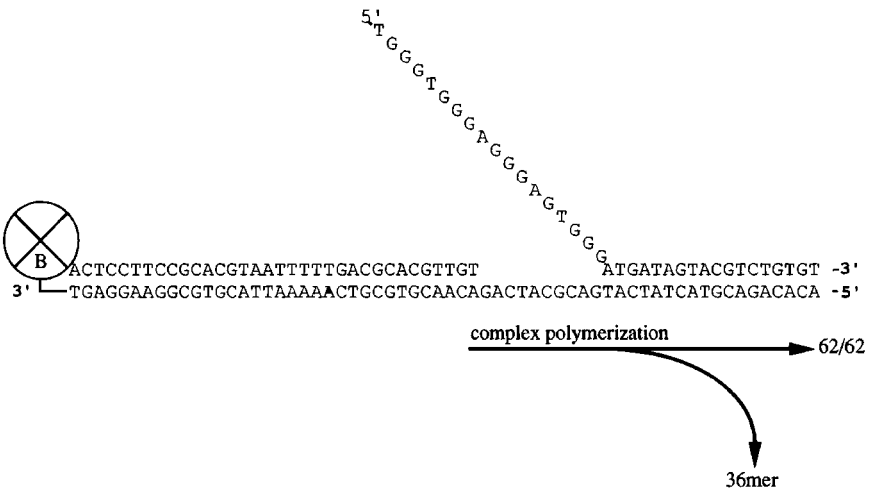


Figure 6 Sequence of the biotinylated fork primer/template, Bio34/62/36. DNA synthesis by the complex results in extension of the primer (34mer) to a 62mer, displacing the fork strand (36mer) in the process. Primer extension by the polymerase alone terminates after the 10-base-pair gap is filled. B is the biotin derivative incorporated into the template strand. The quartered circle represents the bound streptavidin tetramer. (From Reference 130.)

interacts stoichiometrically with the clamp loader, which leads to the hydrolysis of two of the four equivalents of ATP bound to gp44, followed by the binding of the clamp · clamp-loader complex to DNA with the subsequent hydrolysis of the remaining two equivalents of ATP (131). At limiting concentrations the clamp loader acts catalytically to assemble the clamp on the DNA, although its presence at equivalent levels to the clamp and polymerase proteins is required to form a totally productive holoenzyme. Subsequently, the clamp loader also functions as a chaperone for the polymerase to ensure proper holoenzyme formation but does not remain as a member of the holoenzyme (132–135).

Fluorescent tagging of the clamp molecule revealed that the formation of the gp44/62·gp45 complex is associated with a conformation change triggered by ATP hydrolysis (136, 137) that brings gp62 into contact with gp45 (138), reminiscent of the exposure of the δ subunit in the *E. coli* clamp loader during a similar process. Cross-linking of gp45 mutants in which a single cysteine was introduced at various loci and then derivatized showed that the binding of both the clamp loader and the polymerase were to the same face (C-terminal) of the clamp protein, gp45 (139, 140). The clamp loader does not act as a protein kinase by phosphorylating its target protein, but functions in analogy to other molecular motors such as kinesin and myosin (141). In this context, a mutant gp45 protein in which disulfide cross-links have been introduced that link the monomer-monomer interface at all three interfaces hyperstimulates the ATPase activity of the clamp loader consonant with a futile ATP ring-opening cycle (142). However, mutant clamp proteins with at least one nonlinked interface do support processive synthesis (143).

The kinetic events associated with assembly of the holoenzyme have been charted by stopped-flow fluorescence-resonance energy transfer using a gp45 protein whose subunit interface had a donor and acceptor pair on opposite subunits. The overall process is depicted in Figure 7. Steady-state measurements of energy transfer combined with the behavior of the clamp protein upon ultracentrifugation were consistent with an equilibrating solution structure for gp45 in which two of the subunit interfaces were closed and the third open to $\sim$40 Å (143). Beginning with that species, the distance across the open interface is increased to greater than 45 Å and then decreased to 30 Å as it closes onto DNA during a 10-step assembly process that mirrors earlier results (144). This process, and the locus of polymerase/clamp binding, has been described in detail (145–148). Hydrolysis of two equivalents of ATP by the gp44/62·gp45 complex powers the further opening of the clamp protein; binding to DNA spontaneously closes this distance. Hydrolysis of the remaining two equivalents of ATP is not manifest by a further change in this distance; this step is tentatively assigned to driving conformational changes in gp44/62 associated with its role to chaperone gp43 to the holoenzyme. Binding of the polymerase does not close the distance to that anticipated for a fully closed clamp protein, in accord with a separation imposed by the binding of the C terminus of the polymerase between two of the subunit interfaces of gp45 (145, 146).

Several steps in this sequence have rate constants ranging from 0.3 to 6.0 s^{-1} and involve ATP hydrolysis events. There are both similarities and differences relative

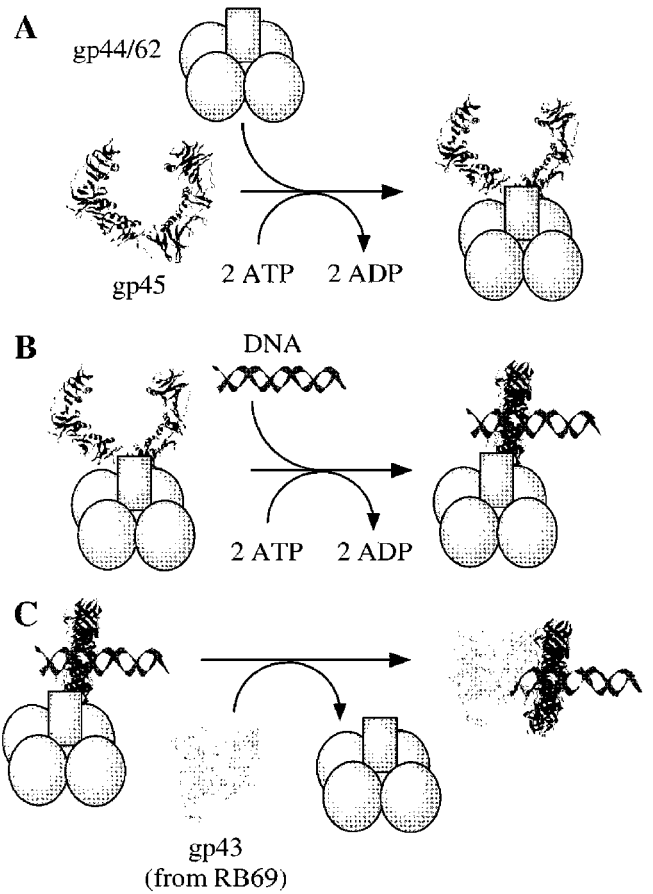


Figure 7 Proposed holoenzyme assembly model. (*A*) Interaction of gp45 and gp44/62 in the presence of ATP. Two molecules are hydrolyzed in a process that opens the gp45 ring. (*B*) Subsequent interaction of DNA. Two additional molecules of ATP are hydrolyzed in a process that closes the gp45 ring. (*C*) Subsequent interaction of gp43 and dissociation of gp44/62. The final holoenzyme complex is formed in a process that closes the gp45 ring further. (From Reference 144.)

to the assembly mechanism of β clamp loading onto the Pol III holoenzyme: A notable similarity is that the clamp in the clamp · clamp-loader · DNA complex for each species is not completely closed. An obvious difference is that ATP hydrolysis by the clamp loader is not required for formation of the β clamp·γ complex·DNA complex; ATPγS is an acceptable substitute. Whereas the disassembly of the T4 holoenzyme from duplex DNA proceeds through the dissociation of the gp45 subunits, that of *E. coli* requires active participation of the clamp-loader·γ complex. The rate for dissociation of the gp45 from the holoenzyme matches that of

the polymerase, which indicates that both depart at a rate of $\sim$0.01 s^{-1}, some one-hundred-fold slower than the polymerase alone, thereby accounting for the processivity of the holoenzyme (149, 150).

Coordinated Synthesis of Leading and Lagging Strands

It is instructive to examine the properties of the primosome within the context of an active replication fork. The gp41 helicase is a $5' \rightarrow 3'$ helicase that requires single-stranded regions on the $5'$ and $3'$ sides of the replication fork (such as one derived from single-stranded M13 annealed with a partially complementing fragment) for activity (151, 152). DNA unwinding mediated by gp41 is processive and requires the hydrolysis of ribo/deoxy-ATP or -GTP; it proceeds at a rate of approximately 400 nucleotides s^{-1} in the presence of active leading/lagging strand synthesis (153). Alone, the helicase activity was found to be stimulated by the primase; this interaction may in part be responsible for increasing its processivity at a replication fork (151). It is probable that, as with the T7 and *E. coli* helicases, the lagging strand passes through the center of the hexamer, since a substrate with a streptavidin-biotin block in the lagging strand inhibits gp41 activity (117).

The gp61 primase alone can synthesize in the presence of a ϕX174 template the dimers pppApC and pppGpC, but only when associated with the helicase does it produce the biologically relevant oligonucleotides. These are primarily pentaribonucleotides of the sequence pppApC $(pN)_3$, which are needed to initiate lagging strand synthesis (154, 155). A mutant form of gp41 lacking approximately 20 amino acids from the C terminus is still effective in association with gp61 as a primosome in priming as well as in DNA synthesis on circular single-stranded DNA, but not in the presence of gp32. This finding recalls the general importance of the C terminus of the helicase in the T7 and *E. coli* systems and the unique, partly mysterious role of single-strand binding protein in the replisome (156).

The combination of the seven proteins (excluding gp59) that are required for holoenzyme and primosome function, and for coverage of exposed single-stranded DNA, is sufficient for replication fork movement on a primed single-stranded circular M13:DNA template at physiological rates (250–300 bases s^{-1}) with the accompanying generation of Okazaki fragments ($\sim$1200 bases). This system closely approximates the behavior of the replisome in the cell (157, 158). No dissociation of the lagging strand polymerase was detected by dilution experiments in which the Okazaki fragment length remained unchanged. The efficiency of primer synthesis on the lagging strand is greatly influenced by the presence of gp32, which blocks wasteful primer synthesis (159). Single-strand binding protein is, however, not recycled during DNA synthesis; its capture by a poly RNA trap arrests the progression of the replication fork (160).

Signaling of Okazaki fragment synthesis in the T4 system has been attributed to the completion of the previous Okazaki fragment, thus permitting the bypass of primosome recognition sequences without primer synthesis. Because the length

of Okazaki fragments proved sensitive to the level of ribonucleoside triphosphates (rNTPs) during active synthesis, a template mechanism which predicts that the length of a successive Okazaki fragment is set by the length of its predecessor was ruled out (160). Instead, size variation was accommodated by a timing mechanism in which the lagging strand holoenzyme pauses before releasing from its DNA template and recycling (160). This type of signaling also requires that the rate of the lagging strand holoenzyme be greater than that on the leading strand. Recently, however, measurements of the rate of leading and lagging strand synthesis at a replication fork showed the two rates to be equal (161). Furthermore, there is no detection of a discrete lag in the dissociation of the holoenzyme from duplexed DNA upon encountering either a 5′ triphosphate or nonphosphorylated 5′ end at the site of the previous Okazaki fragment (162, 163). The rate of this dissociation, 0.1–1 s^{-1}, approaches that required for completing Okazaki fragment synthesis within the time frame of 2–3 s.

To characterize quantitatively coordinated leading and lagging strand synthesis, a minicircle with an annealed replication fork similar to that invoked with the T7 replication proteins (see Figure 3) was used to assemble and trace the movement of the T4 replisome (161). The rates of leading and lagging strand synthesis were identical, which suggests that the two processes are tightly coupled, consistent with the recycling of the polymerase and primase components (164). In contrast, the gross rate of primer appearance, in the case of earlier studies on either the *E. coli* or T4 primase alone (156), was too slow to keep pace with the rate of Okazaki fragment synthesis, which suggests the presence of an unproductive primase complex or misassembly of the replisome owing to the experimental conditions. The appearance of unused tetra- and pentameric RNA primers even in the presence of gp32 may indicate sporadic priming by the primase without a specific signal. Given the rate of replication fork movement of $\sim$400 s^{-1}, a primase turnover number of 1 s^{-1} (as observed for the T7 helicase-primase) would be able to provide only 2–3 primers within that interval, even if all recognition sequences were primed.

The presence of gp32 is absolutely critical for lagging strand synthesis: Either its absence or its replacement with gp32-A, in which the C terminus is truncated, effectively abolishes lagging strand synthesis. gp32-A no longer binds either the primase or polymerase but inhibits their activities (127, 165), probably through retention of its DNA binding affinity. The coupling of leading and lagging strand synthesis is further supported by termination of lagging strand synthesis by the addition of dideoxy CTP, which specifically stops lagging strand synthesis and concomitantly aborts synthesis on the leading strand (Figure 3).

It will be of considerable interest to monitor the progress of the replication fork on an abbreviated time scale. The observation of a delay or pause in both leading and lagging strand synthesis coincident with the completion of an Okazaki fragment would allow dissection of the holoenzyme recycling process, which derives from a combination of factors: release of the holoenzyme from the lagging strand, reclamping of the clamp protein, and priming by the primase. Two

of these processes, release and reclamping, have rate constants (noted above) commensurate with a time frame of 1–3 s; it is likely that measurement of the rate of primer synthesis in the context of a replication fork may yield a similar value.

It would, therefore, appear that the great majority of the T4 replisome proteins recycle at the replication fork. A key polymerase-polymerase interaction has been traced to the amino acid region 401–600, which includes a long coiled-coil extension of a finger domain located next to that of the central core by alignment of gp43 with the RB69 polymerase (100). In addition to the protein-protein interactions involving gp32 already noted, gp32 showed affinity on chromatography for all the T4 replication proteins excepting the helicase (166). These contacts are also manifest kinetically in the reduction by gp32 of the primase rate of priming on single-stranded DNA, an inhibition relieved by the presence of the clamp and clamp loader proteins (156). These experiments, done without knowledge of gp59, might be interpreted as the displacement of gp32 by gp45 and gp44/62 (acting as surrogates for gp59), permitting binding of the primosome to its recognition sequence. However, a tryptic fragment of gp41 missing 17–20 amino acids from the C terminus does not promote priming on gp32-covered DNA even in the presence of the above accessory proteins. Consequently, gp45·gp44/62 (as well as the polymerase) may interact with this region of the helicase (96). The caveat to remember is that the priming rate in these experiments is well below that required in the context of a replication fork.

Two models of the T4 replisome at a replication fork have been created using these data as well as analogy when reasonable (Figure 8). The two holoenzymes can be oriented in a parallel or antiparallel orientation (or some intermediate geometry). The lagging strand DNA is threaded through the hexameric helicase with the primase active site either facing away from or toward its recognition site. The former requires a loop in the lagging strand DNA. Expansion of this loop occurs as a result of the forward travel of the helicase and synthesis of the Okazaki fragment. A small second loop may temporarily appear behind the helicase during the short pause between synthesis of successive Okazaki fragments. The primase may hand off the pentameric ribonucleotide through a sequence of clamp loader–promoted removal of gp32 and its replacement with the clamp protein, which obviates dissociation of the short RNA primer (analogous to the *E. coli* molecular switch), or the primase may pass the ribonucleotide directly to the lagging strand polymerase. Coverage of the emerging single-stranded DNA prevents excessive priming by the primase. A simple explanation for the dependency of Okazaki fragment length on the level of rNTP would be a change in the frequency of this priming event independent of the completion of the synthesis of the previous Okazaki fragment. There is no clamp loader nor gp59 shown as part of the replisome. It is possible, though not demonstrated, that gp45 recycles with the polymerase and closes spontaneously from a partially opened conformation in order to bypass gp32, as noted in the clamp loading sequence described earlier. Alternatively, gp45 is recruited from the large solution pool (165).

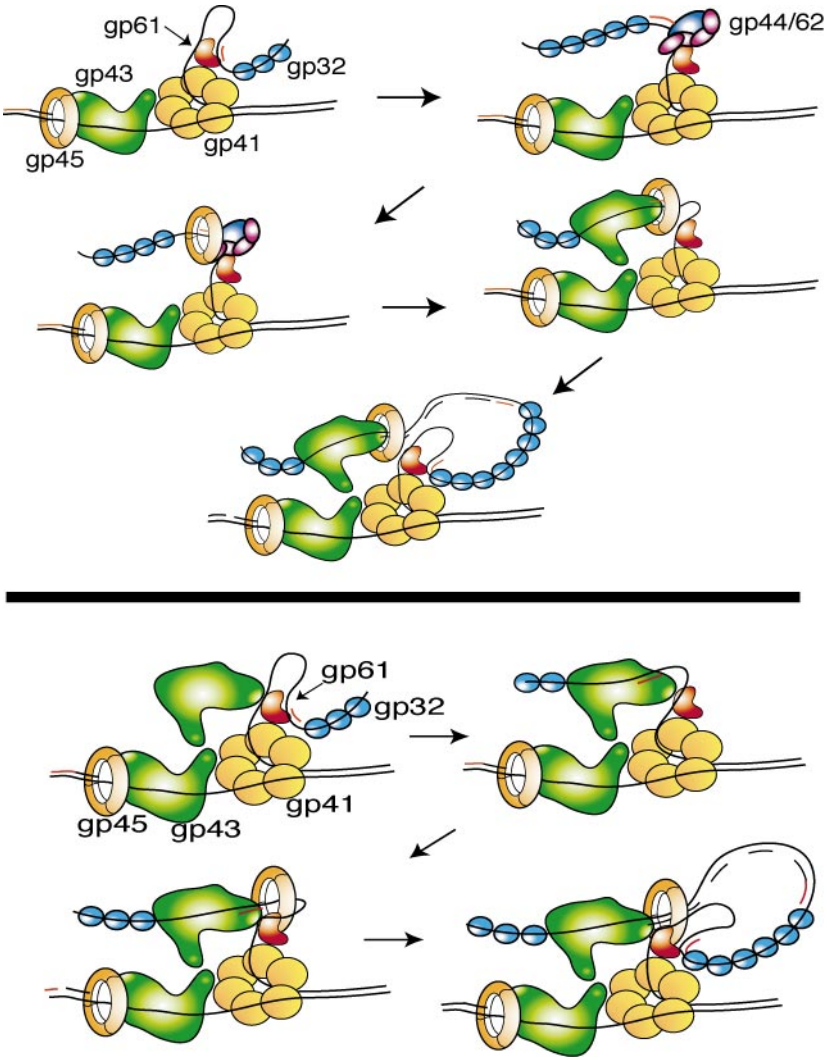


Figure 8 Two possible models for primer hand-off during coordinated synthesis of leading and lagging DNA strands in bacteriophage T4. In both, the two gp43 polymerase enzymes are antiparallel; they may instead be parallel or in some intermediate disposition. (*Top*) The newly synthesized primer is first handed off to the clamp loader, gp44/62. The primer is then threaded through the clamp protein and onto the lagging strand polymerase, which extends it, looping out the lagging strand behind. (*Bottom*) The primer is handed off directly from primase to polymerase. This exchange is followed by loading of the clamp and lagging strand synthesis.

REFLECTIONS

Several conclusions, some reinforced by experimental facts and others more speculative, relate the three replisomes discussed. All three catalyze leading and lagging strand synthesis at a replication fork with nearly the same overall rate (300–400 bases s^{-1}) of fork movement; key individual proteins such as the helicases and polymerases are nearly matched in their turnover numbers. The two polymerases coordinate their syntheses, acting processively on both strands. The rates ($>1\ s^{-1}$) of clamp loading, of lagging strand polymerase release, and priming (at least for the T7 system) are sufficient to occur within the 1- to 3-s time frame required for Okazaki synthesis with minimal effect on fork movement. The primase can shuttle in/out of the replisome—in for T7 and maybe T4, in and out for *E. coli*. Given the high rate of polymerase synthesis, properly reconstituted replication forks will have few wasted primers even if the primase is not the recipient of a start signal. Clearly, the length of the Okazaki fragments can be made to vary by changing the levels or activities of various proteins or concentration of the nucleotides in this case.

The protein-protein contacts, the molecular cement that binds the replisome together, are only coming into view as more structures become known. There will be a bewildering number of conformational states, as befits a dynamic biological machine, so structural elucidation alone will not be the last word. Clearly, the C termini of single-strand binding proteins, polymerase, helicase, etc in all these systems serve as tie-rods in holding the structure together and orchestrating coupled movements. Undoubtedly, the lagging strand DNA is looped, but all its contact points have not been elucidated. We can be fairly sure that many of these statements will be valid for eukaryotic systems as well (167).

ACKNOWLEDGMENT

We thank Faoud T Ishmael for his preparation of Figure 8.

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

1. Alberts BM, Morris C, Mace D, Sinha N, Bittner M. 1975. In *DNA Synthesis and Regulation*, ed. M Goulian, PC Hanawalt, pp. 241–69. Menlo Park, CA: Benjamin
2. Nossal NG. 1994. In *Molecular Biology of Bacteriophage*, ed. JD Karam, pp. 43–55. Washington, DC: Am. Soc. Microbiol.
3. Arai N, Kornberg A. 1981. *J. Biol. Chem.* 256:5294–98
4. Huber HE, Tabor S, Richardson CC. 1987. *J. Biol. Chem.* 262:16224–32
5. Tabor S, Huber HE, Richardson CC. 1987. *J. Biol. Chem.* 262:16212–23
6. Tabor S, Richardson CC. 1981. *Proc. Natl. Acad. Sci. USA* 78:205–9
7. Kolodner R, Richardson CC. 1977. *Proc. Natl. Acad. Sci. USA* 74:1525–29
8. Kim YT, Tabor S, Bortner C, Griffith

JD, Richardson CC. 1992. *J. Biol. Chem.* 267:15022–31
9. Kim YT, Tabor S, Churchich JE, Richardson CC. 1992. *J. Biol. Chem.* 267:15032–40
10. Steitz TA. 1993. *Curr. Opin. Struct. Biol.* 3:31–38
11. Doublie S, Tabor S, Long AM, Richardson CC, Ellenberger T. 1998. *Nature* 391:251–58
12. Ilyina TV, Gorabalenya AE, Koonin EV. 1992. *J. Mol. Evol.* 34:351–57
13. Egelman EH, Yu X, Wild R, Hingorani MM, Patel SS. 1995. *Proc. Natl. Acad. Sci. USA* 92:3869–73
14. Dunn JJ, Studier FW. 1983. *J. Mol. Biol.* 166:477–75
15. Bernstein JA, Richardson CC. 1988. *Proc. Natl. Acad. Sci. USA* 85:396–400
16. Bernstein JA, Richardson CC. 1988. *J. Biol. Chem.* 263:14891–99
17. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB. 2000. *Cell* 101:589–600
17a. Sawaya MR, Guo S, Tabor S, Richardson CC, Ellenberger T. 1999. *Cell* 99:167–77
18. Hingorani MM, Patel SS. 1996. *Biochemistry* 35:2218–28
19. Patel SS, Hingorani MM. 1993. *J. Biol. Chem.* 268:10668–75
20. Notarnicola SM, Park K, Griffith JD, Richardson CC. 1995. *J. Biol. Chem.* 270:20215–24
21. Guo S, Tabor S, Richardson CC. 1999. *J. Biol. Chem.* 274:30303–9
22. Kim YT, Richardson CC. 1993. *Proc. Natl. Acad. Sci. USA* 90:10173–77
23. Kim YT, Richardson CC. 1994. *J. Biol. Chem.* 269:5270–78
24. Williams KR, Spicer EK, LoPresti MB, Guggenheimer RA, Chase JW. 1983. *J. Biol. Chem.* 258:3346–55
25. Lonberg N, Kowalczykowski S, Paul L, von Hippel PH. 1981. *J. Mol. Biol.* 145:123–38
26. Lechner RL, Richardson CC. 1983. *J. Biol. Chem.* 258:11185–96
27. Nakai H, Richardson CC. 1986. *J. Biol. Chem.* 261:15208–16
28. Fujiyama A, Kohara Y, Okazaki T. 1981. *Proc. Natl. Acad. Sci. USA* 78:903–7
29. Mendelman LV, Beauchamp BB, Richardson CC. 1994. *EMBO J.* 13:3909–16
30. Kusakabe T, Richardson CC. 1996. *J. Biol. Chem.* 271:19563–70
31. Frick DN, Baradaran K, Richardson CC. 1998. *Proc. Natl. Acad. Sci. USA* 95:7957–62
32. Notarnicola SM, Mulcahy HL, Lee J, Richardson CC. 1997. *J. Biol. Chem.* 272:18425–33
33. Ahnert P, Patel SS. 1997. *J. Biol. Chem.* 272:32267–73
34. Lohman TM. 1993. *J. Biol. Chem.* 268:2269–72
35. Engler MJ, Richardson CC. 1983. *J. Biol. Chem.* 258:11197–205
36. Nakai H, Richardson CC. 1988. *J. Biol. Chem.* 263:9818–30
37. Kusakabe T, Baradaran K, Lee J, Richardson CC. 1998. *EMBO J.* 17:1542–52
38. Hacker KJ, Johnson KA. 1997. *Biochemistry* 36:14080–87
39. Frick DN, Richardson CC. 1999. *J. Biol. Chem.* 274:35889–98
40. Frick DN, Kumar S, Richardson CC. 1999. *J. Biol. Chem.* 274:35899–907
41. Debyser Z, Tabor S, Richardson CC. 1994. *Cell* 77:157–66
42. Lee J, Chastain PD, Kusakabe T, Griffith JD, Richardson CC. 1998. *Mol. Cell. Biol.* 1:1001–10
43. Park K, Debyser Z, Tabor S, Richardson CC, Griffith JD. 1998. *J. Biol. Chem.* 273:5260–70
44. Herendeen DR, Kelly TJ. 1996. *Cell* 84:5–8
45. O'Donnell M. 1992. *BioEssays* 14:105–11
46. O'Donnell M. 1994. *Ann. NY Acad. Sci.* 726:144–53
47. McHenry CS. 1991. *J. Biol. Chem.* 266:19127–30
48. Kelman Z, O'Donnell M. 1995. *Annu. Rev. Biochem.* 64:171–200

49. Dallmann HG, Kim S, Pritchard AE, Marians KJ, McHenry CS. 2000. *J. Biol. Chem.* 275:15512–19
50. Studwell-Vaughn PS, O'Donnell M. 1991. *J. Biol. Chem.* 266:19833–41

50a. Onrust R, Finkelstein J, Naktinis V, Turner J, Fang L, O'Donnell M. 1995. *J. Biol. Chem.* 270:13348–57

51. Olson MW, Dallmann HG, McHenry CS. 1995. *J. Biol. Chem.* 270:29570–77
52. Xiao H, Naktinis V, O'Donnell M. 1995. *J. Biol. Chem.* 270:13378–83
53. Stukenberg PT, O'Donnell M. 1995. *J. Biol. Chem.* 270:13384–91
54. Hingorani MM, O'Donnell M. 2000. *Curr. Biol.* 10:R25–29
55. Kong X-P, Onrust R, O'Donnell M, Kuriyan J. 1992. *Cell* 69:425–37
56. Krishna TS, Kong XP, Gary S, Burgers PM, Kuriyan J. 1994. *Cell* 79:1233–43
57. Egelman EH. 1998. *J. Struct. Biol.* 124:123–28
58. Bujalowski W, Klonowska MM, Jezewska MJ. 1994. *J. Biol. Chem.* 269:31350–58
59. Yu X, Jezewska MJ, Bujalowski W, Egelman EH. 1996. *J. Mol. Biol.* 259:7–14
60. San Martin MC, Stamford NPJ, Dammerova N, Dixon NE, Carazo JM. 1995. *J. Struct. Biol.* 114:167–76
61. Bujalowski W, Jezewska MJ. 2000. *Biochemistry* 39:2106–22
62. Jezewska MJ, Rajendran S, Bujalowski W. 1998. *Biochemistry* 37:3116–36
63. Matson SW, Kaiser-Rogers KA. 1990. *Annu. Rev. Biochem.* 59:289–329
64. Jezewska MJ, Rajendran S, Bujalowski W. 1997. *Biochemistry* 36:10320–26
65. Yoda K, Yasuda H, Jiang XW, Okazaki T. 1988. *Nucleic Acids Res.* 16:6531–46
66. Kitani T, Yoda K, Ogawa T, Okazaki T. 1985. *J. Mol. Biol.* 184:45–52
67. Keck JL, Roche DD, Lynch AS, Berger JM. 2000. *Science* 287:2482–86
68. Tougu K, Marians KJ. 1996. *J. Biol. Chem.* 271:21391–97
69. LeBowitz JH, McMacken R. 1986. *J. Biol. Chem.* 261:4738–48
70. Johnson SK, Bahattacharyya S, Griep MA. 2000. *Biochemistry* 39:736–44
71. Bahattacharyya S, Griep MA. 2000. *Biochemistry* 39:745–52
72. Griep MA, Adkins BJ, Hromas D, Johnson S, Miller J. 1997. *Biochemistry* 36:544–53
73. Powers L, Griep MA. 1999. *Biochemistry* 38:7413–20
74. Lohman TM, Ferrari ME. 1994. *Annu. Rev. Biochem.* 63:527–70
75. Meyer R, Glassberg J, Kornberg A. 1979. *Proc. Natl. Acad. Sci. USA* 76:1702–5
76. Arai K-I, Low R, Kobori J, Shlomai J, Kornberg A. 1981. *J. Biol. Chem.* 256: 5273–80
77. Mok M, Marians KJ. 1987. *J. Biol. Chem.* 262:16644–54
78. Baker TA, Wickner SH. 1992. *Annu. Rev. Genet.* 26:447–77
79. Fay PJ, Johanson KO, McHenry CS, Bambara RA. 1981. *J. Biol. Chem.* 256:976–83
80. Studwell-Vaughn PS, O'Donnell M. 1990. *J. Biol. Chem.* 265:1171–78
81. Naktinis V, Onrust R, Fang L, O'Donnell M. 1995. *J. Biol. Chem.* 270:13358–65
82. Hingorani MM, Bloom LB, Goodman MF, O'Brien WE, O'Donnell M. 1999. *EMBO J.* 18:5131–44
83. Naktinis V, Turner J, O'Donnell M. 1996. *Cell* 84:137–45
84. Kelman Z, O'Donnell M. 1995. *Nucleic Acids Res.* 23:3613–20
85. Wu CA, Zechner EL, Marians KJ. 1992. *J. Biol. Chem.* 267:4030–44
86. Zechner EL, Wu CA, Marians KJ. 1992. *J. Biol. Chem.* 267:4045–53
87. O'Donnell M, Kornberg A. 1985. *J. Biol. Chem.* 260:12884–89
88. Zechner EL, Wu CA, Marians KJ. 1992. *J. Biol. Chem.* 267:4054–63
89. Kim S, Dallman NHG, McHenry CS, Marians KJ. 1996. *J. Biol. Chem.* 271:21406–12
90. Wu CA, Zechner EL, Hughes AJ, Franden MA, McHenry CS, Marians KJ. 1992. *J. Biol. Chem.* 267:4064–73

91. Kim S, Dallman HG, McHenry CS, Marians KJ. 1996. *Cell* 84:643–50
92. Wu CA, Zechner EL, Reems JA, McHenry CS, Marians KJ. 1992. *J. Biol. Chem.* 267:4074–83
93. Tougu K, Marians KJ. 1996. *J. Biol. Chem.* 271:21398–405
94. Yuzhakov A, Kelman Z, O'Donnell M. 1999. *Cell* 96:153–63
95. Kelman Z, Yuzhakov A, Andjelkovic J, O'Donnell M. 1998. *EMBO J.* 17:2436–49
96. Nossal NG. 1992. *FASEB J.* 6:871–78
97. Alberts BM. 1990. In *Mechanisms and Enzymology of DNA Replication*, ed. FYH Wu, CW Wu, pp. 223–37. New York: Raven
98. Young MC, Reddy MK, von Hippel PH. 1992. *Biochemistry* 31:8675–90
99. Wang J, Yu P, Lin T-C, Konigsberg WH, Steitz TA. 1996. *Biochemistry* 35:8110–19
100. Wang J, Sattar AK, Wang CC, Karam JD, Konigsberg WH, Steitz TA. 1997. *Cell* 89:1087–99
101. Karam JD, Konigsberg WH. 2000. *Prog. Nucleic Acid Res.* 64:65–96
102. Fairfield FR, Newport JW, Dolejsi MK, von Hippel PH. 1983. *J. Biomol. Struct. Dyn.* 1:715–27
103. Capson TL, Peliska JA, Kaboord BF, Frey MW, Lively C, et al. 1992. *Biochemistry* 31:10984–94
104. Huang C-C, Hearst JE, Alberts BM. 1981. *J. Biol. Chem.* 256:4087–94
105. Roth AC, Nossal NG, Englund PT. 1982. *J. Biol. Chem.* 257:1267–73
106. Jarvis TC, Paul LS, von Hippel PH. 1989. *J. Biol. Chem.* 264:12709–16
107. Mace DC, Alberts BM. 1984. *J. Mol. Biol.* 177:279–93
108. Jarvis TC, Paul LS, Hockensmith JW, von Hippel PH. 1989. *J. Biol. Chem.* 264:12717–29
109. Vinkemeier U, Moarefi I, Darnell JE, Kuriyan J. 1998. *Science* 279:1048–52
110. Gogol EP, Young MC, Kubasek WL, Jarvis TC, von Hippel PH. 1992. *J. Mol. Biol.* 224:395–412
111. Dong F, Gogl EP, von Hippel PH. 1995. *J. Biol. Chem.* 270:7462–73
112. Liu CC, Alberts BM. 1981. *J. Biol. Chem.* 256:2813–20
113. Dong F, von Hippel PH. 1996. *J. Biol. Chem.* 271:19625–31
114. Jing DH, Dong F, Latham GJ, von Hippel PH. 1999. *J. Biol. Chem.* 274:27287–98
115. Jones CE, Mueser TC, Nossal NG. 2000. *J. Biol. Chem.* 275:27145–54
116. Morrical SW, Beernink HTH, Dash A, Hempstead K. 1996. *J. Biol. Chem.* 271:20198–207
117. Raney KD, Carver TE, Benkovic SJ. 1996. *J. Biol. Chem.* 271:14074–81
118. Barry JE, Alberts BM. 1994. *J. Biol. Chem.* 269:33049–62
119. Yonesaki T. 1994. *J. Biol. Chem.* 269:1284–89
120. Morrical SW, Hempstead K, Morrical MD. 1994. *J. Biol. Chem.* 269:33069–81
121. Curtis MJ, Alberts BM. 1976. *J. Mol. Biol.* 102:793–816
122. Huberman JA, Kornberg A, Alberts BM. 1971. *J. Mol. Biol.* 62:39–52
123. Lohman TM, Kowalczykowski S. 1981. *J. Mol. Biol.* 152:67–109
124. Giedroc DP, Keating KM, Williams KR, Konigsberg WH, Coleman JE. 1986. *Proc. Natl. Acad. Sci. USA* 83:8452–56
125. Shamoo Y, Williams KR, Konigsberg WH. 1988. *Proteins* 4:1–6
126. Shamoo Y, Friedman AM, Parsons MR, Konigsberg WH, Steitz TA. 1995. *Nature* 376:362–66
127. Burke RL, Alberts BM, Hosoda J. 1980. *J. Biol. Chem.* 255:11484–93
128. Formosa T, Burke RL, Alberts BM. 1983. *Proc. Natl. Acad. Sci. USA* 80:2442–46
129. Kaboord BF, Benkovic SJ. 1995. *Curr. Biol.* 5:149–57
130. Kaboord BF, Benkovic SJ. 1996. *Biochemistry* 35:1084–92
131. Sexton DJ, Kaboord BF, Berdis AJ,

Carver TE, Benkovic SJ. 1998. *Biochemistry* 37:7749–56
132. Berdis AJ, Benkovic SJ. 1996. *Biochemistry* 35:9253–65
133. Capson TL, Benkovic SJ, Nossal NG. 1991. *Cell* 65:249–58
134. Munn MM, Alberts BM. 1991. *J. Biol. Chem.* 266:20034–44
135. Reddy MK, Weitzel SE, von Hippel PH. 1993. *Proc. Natl. Acad. Sci. USA* 90:3211–15
136. Sexton DJ, Carver TE, Berdis AJ, Benkovic SJ. 1996. *J. Biol. Chem.* 271:28045–51
137. Latham GJ, Pietroni P, Dong F, Young MC, von Hippel PH. 1996. *J. Mol. Biol.* 264:426–39
138. Pietroni P, Young MC, Latham GJ, von Hippel PH. 1997. *J. Biol. Chem.* 272:31666–76
139. Latham GJ, Bacheller DJ, Pietroni P, von Hippel PH. 1997. *J. Mol. Biol.* 272:31677–84
140. Latham GJ, Bacheller DJ, Pietroni P, von Hippel PH. 1997. *J. Biol. Chem.* 272:31685–92
141. Berdis AJ, Benkovic SJ. 1997. *Biochemistry* 36:2733–43
142. Latham GJ, Dong F, Pietroni P, Dozono JM, Bacheller DJ, von Hippel PH. 1999. *Proc. Natl. Acad. Sci. USA* 96:12448–53
143. Alley SC, Shier VK, Abel-Santos E, Sexton DJ, Soumillion P, Benkovic SJ. 1999. *Biochemistry* 38:7696–709
144. Alley SC, Abel-Santos E, Benkovic SJ. 2000. *Biochemistry* 39:3076–90
145. Berdis AJ, Soumillion P, Benkovic SJ. 1996. *Proc. Natl. Acad. Sci. USA* 93:12822–27
146. Alley SC, Jones AD, Soumillion P, Benkovic SJ. 1999. *J. Biol. Chem.* 274:24485–89
147. Shamoo Y, Steitz TA. 1999. *Cell* 99:155–66
148. Alley SC, Trakselis MA, Mayer MU, Ishmael FT, Jones AD, Benkovic SJ. 2000. *J. Biol. Chem.* Submitted
149. Soumillion P, Sexton DJ, Benkovic SJ. 1998. *Biochemistry* 37:1819–27
150. Sexton DJ, Berdis AJ, Benkovic SJ. 1997. *Curr. Opin. Chem. Biol.* 1:316–22
151. Venkatesan M, Silver LL, Nossal NG. 1982. *J. Biol. Chem.* 257:12426–34
152. Richardson RW, Nossal NG. 1989. *J. Biol. Chem.* 264:4725–31
153. Schrock RD, Alberts B. 1996. *J. Biol. Chem.* 271:16678–82
154. Nossal NG, Hinton DM. 1987. *J. Biol. Chem.* 262:10879–85
155. Hinton DM, Nossal NG. 1987. *J. Biol. Chem.* 262:10873–78
156. Richardson RW, Nossal NG. 1989. *J. Biol. Chem.* 254:4732–39
157. Alberts BM, Barry J, Bedinger BP, Formosa T, Jongeneel CV, Kruezer KN. 1983. *Cold Spring Harbor Symp. Quant. Biol.* 47:655–68
158. Cha T-A, Alberts BM. 1989. *J. Biol. Chem.* 264:12220–25
159. Cha TA, Alberts BM. 1990. *Biochemistry* 29:1791–98
160. Selick HE, Barry J, Cha T-A, Munn M, Nakanishi M, et al. 1987. In *DNA Replication and Recombination*, ed. R McMaken, TJ Kelly, pp.183–214. New York: Liss
161. Salinas F, Benkovic SJ. 2000. *Proc. Natl. Acad. Sci. USA* 97:7196–201
162. Hacker KJ, Alberts BM. 1994. *J. Biol. Chem.* 269:24221–28
163. Carver TE Jr, Sexton DJ, Benkovic SJ. 1997. *Biochemistry* 36:14409–17
164. Richardson RW, Ellis RL, Nossal NG. 1990. In *Molecular Mechanisms in DNA Replication and Recombination*, ed. CC Richardson, IR Lehman, pp. 247–89. New York: Liss
165. Burke RL, Munn M, Barry J, Alberts BM. 1985. *J. Biol. Chem.* 260:1711–22
166. Cha T-A, Alberts BM. 1988. *Cancer Cells* 6:1–10
167. Baker TA, Bell SP. 1998. *Cell* 92:295–305