Annu. Rev. Biochem. 2001. 70:39–80

DNA PRIMASES

David N. Frick[1] and Charles C. Richardson[2]
[1]*Department of Biochemistry and Molecular Biology, New York Medical College, Valhalla, New York 10595; e-mail: DAVID_FRICK@NYMC.edu*
[2]*Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115; e-mail: ccr@hms.harvard.edu*

Key Words replication, RNA, oligoribonucleotides, polymerase, helicase

■ **Abstract** DNA primases are enzymes whose continual activity is required at the DNA replication fork. They catalyze the synthesis of short RNA molecules used as primers for DNA polymerases. Primers are synthesized from ribonucleoside triphosphates and are four to fifteen nucleotides long. Most DNA primases can be divided into two classes. The first class contains bacterial and bacteriophage enzymes found associated with replicative DNA helicases. These prokaryotic primases contain three distinct domains: an amino terminal domain with a zinc ribbon motif involved in binding template DNA, a middle RNA polymerase domain, and a carboxyl-terminal region that either is itself a DNA helicase or interacts with a DNA helicase. The second major primase class comprises heterodimeric eukaryotic primases that form a complex with DNA polymerase alpha and its accessory B subunit. The small eukaryotic primase subunit contains the active site for RNA synthesis, and its activity correlates with DNA replication during the cell cycle.

CONTENTS

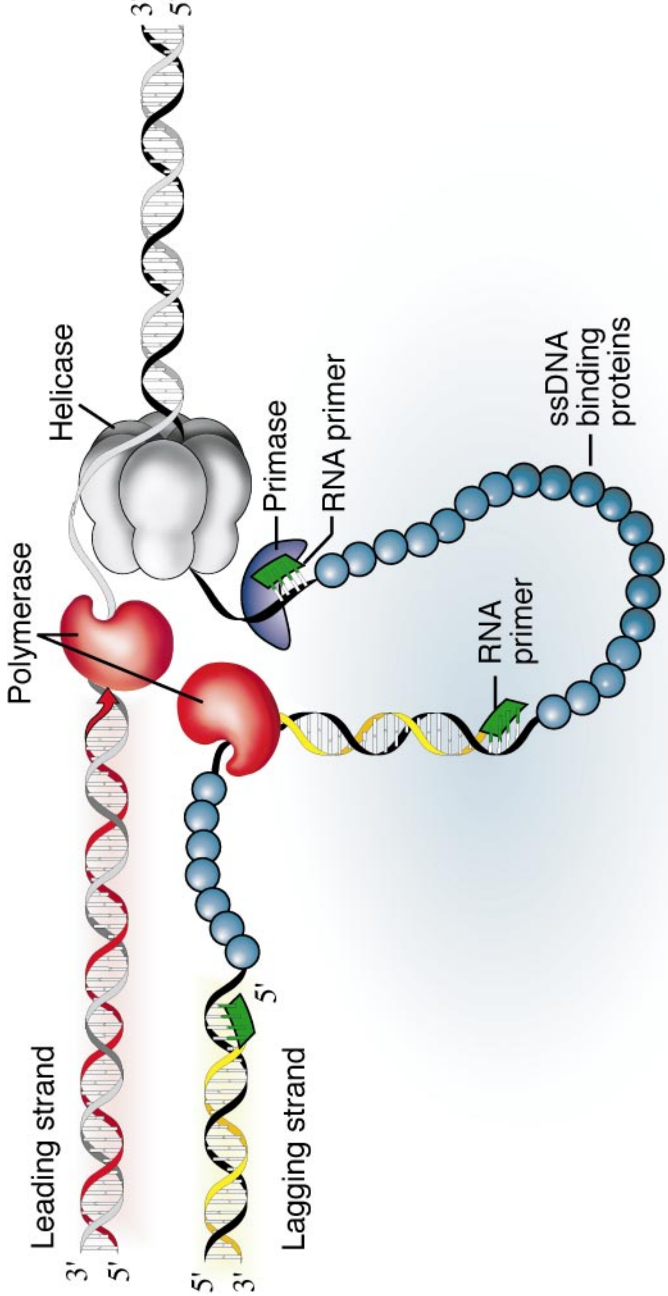
INTRODUCTION

On the basis of their structural model for DNA, Watson & Crick proposed in 1953 that each strand of the parental DNA molecule could serve as a template for the synthesis of two daughter molecules (1). During the subsequent decade, a succession of experiments attempted to elucidate the semiconservative mechanism of DNA replication whereby both new strands are synthesized in the same direction as the advancing replication fork (Figure 1). The simultaneous synthesis of both DNA templates is a remarkably intricate process mainly because the two strands have opposite polarities but DNA polymerase works in only one direction, by adding nucleotides to the 3′ end of a growing chain. A semidiscontinuous model of DNA replication was proposed to explain the unidirectional growth of both DNA strands. According to this model, one strand is synthesized continually and the other strand is synthesized in small units (Okazaki fragments) that are subsequently joined to yield a complete duplex DNA (2).

Semidiscontinuous DNA synthesis can explain the simultaneous replication of both strands in the same overall direction, but the process is complicated by the fact that the replication of the discontinuous strand requires many additional reactions. Besides the ligase needed to join the Okazaki fragments, an additional component is needed at the replication fork to provide primers on the single-stranded DNA (ssDNA) template. These primers are normally short RNA molecules, synthesized by an RNA polymerase, which, unlike known DNA polymerases, can initiate the synthesis of RNA chains de novo. Although early experiments suggested that any cellular RNA polymerase could provide primers (3), subsequent work has shown

→

Figure 1 Role of DNA primase in semi-discontinuous DNA replication. A DNA replication fork is shown progressing from *left* to *right* to copy duplex DNA (*black/gray*). Because the two strands of DNA are oriented in opposite polarities, only the leading strand (*red*) is synthesized continuously while the lagging strand (*yellow*) is synthesized discontinuously. If the two required DNA polymerases maintain contact to coordinate their activities, a looping of the lagging strand occurs. The DNA helicase (*gray*) separates the helix into ssDNA, which is stabilized by ssDNA binding proteins (*blue*). The primase (*purple*) synthesizes RNA primers (*green*) on the lagging strand that are transferred periodically to the DNA polymerase (*red*).


Leading strand
3'
5'
Lagging strand
5'
3'
5'
Polymerase
Helicase
Primase
RNA primer
RNA primer
ssDNA binding proteins
3'
5'

that DNA replication depends on a novel class of RNA-synthesizing enzymes distinct from the classical RNA polymerases. Early studies of T7 DNA replication implicated the phage gene 4 protein in priming DNA synthesis both in vivo (4,5) and in vitro (6,7). Initially called a priming protein (8), the T7 gene 4 protein later was designated a primase (9). The name has since been used to designate its *Escherichia coli* counterpart, the DnaG protein, and all other members of this enzyme class (10). Details of the role of DNA primase in the replication of the lagging strand are presented in Figure 1.

This review focuses on the structure and function of DNA primases and on their role in the initiation of Okazaki fragments. The coordination of primase activity with the numerous proteins and enzymes that compose the replication complex are also discussed. Three components of the replication complex in particular greatly influence primase activity: DNA polymerase, DNA helicase, and ssDNA binding protein. The importance of the interaction of primases with these proteins is highlighted by the fact that in virtually all systems, one or more of these proteins is physically associated with DNA primase or is essential for efficient primer synthesis. Besides permitting semidiscontinuous replication, primases play other significant roles in DNA synthesis and other cellular processes. For example, primases initiate leading strand synthesis at chromosomal origins, and they are involved in DNA repair (11). It should also be noted that DNA primases are not the sole means of providing primers for DNA polymerases. Other mechanisms for strand initiation include the use of DNA ends generated by recombination or repair, transcripts made by conventional RNA polymerases, or priming proteins attached to the ends of linear DNAs (12).

BASIC PROPERTIES OF PRIMASE GENES AND PROTEINS

With the recent explosion in genome sequencing, numerous proteins have been uncovered that possess significant similarity to known primases. Most primases can be classified into one of two large groups. One group contains the primases from bacteria and their phages, and the second comprises eukaryotic DNA primases. All primases share many properties, but the proteins in the two classes differ both in structure and in their relationship with other proteins in the replication complex. The prokaryotic primases are normally found associated with the replicative DNA helicase, whereas the eukaryotic primases are found as complexes with DNA polymerase α (Pol α). Several examples of each class of primases as well as several primases that do not fit into either of these two classes are discussed below. The basic properties of primases are summarized in Table 1.

Prokaryotic Primases (Helicase Associated)

As shown in Figure 2, the prokaryotic primases include proteins from both bacteria and their phages. The bacterial primases (Figure 2*A*) and the phage primases

TABLE 1 Biochemical properties of purified DNA primases

Source	Gene	Amino acids	*M*r (approx)	Primer length	Primer sequence	Recognition site[a]
Prokaryotic						
E. coli	dnaG	581	65,572	10–12	pppAG(N)$_{8-10}$	5′-CT$\underline{\text{G}}$-3′
T7	gene 4	566	62,655	4–5	pppAC(N)$_{2-3}$	5′-GT$\underline{\text{C}}$-3′
SP6	primase	661	74,096	4–10	pppGC(N)$_{2-8}$	5′-GC$\underline{\text{A}}$-3′
T4	gene 61	342	39,768	4–5	pppAC(N)$_{2-3}$	5′-GT$\underline{\text{T}}$-3′
P4	α	777	84,912	2–5	pppAG(N)$_{0-3}$	5′-CT-3′
Plasmid						
ColE2	rep	308	34,851	2–3	ppAGA	5′-TCT$\underline{\text{G}}$-3′
Viral						
Human HSV	UL52	1,058	114,341	10–12	pppGG(N) $_{8-10}$	5′-ACCCTCC C$\underline{\text{GA}}$-3′
Eukaryotic						
S. cerevisiae	PRI1	409	47,609	8–10	ppp(A/G)(N)$_{7-9}$	
	PRI2	528	62,263			
Drosophila melanogaster	PRI1	438	50,167	8–15	ppp(A/G)(N)$_{7-14}$	
	PRI2	533	61,390			
Mus musculus	PRI1	417	49,295	9–11	ppp(A/G)(N)$_{8-10}$	
	PRI2	505	58,409			
Homo sapiens	PRI1	420	49,902	11–14	ppp(A/G)(N)$_{8-13}$	
	PRI2	509	58,778			

[a] Cryptic nucleotides not copied into the primer are underlined. Eukaryotic primases have not been shown to use specific recognition sites.

(Figure 2*B*) are distantly related but share many biochemical properties, such as a close association with DNA helicase and several consensus signature sequences (13–15).

T7 Primase The protein to which primase activity was first ascribed is the product of gene 4 of bacteriophage T7 of *E. coli*. Gene 4 of phage T7 is essential for phage replication (16), and T7 mutants defective in gene 4 synthesize a small amount of DNA that hybridizes to only one strand (4, 5). A sensitive assay for primer synthesis by the gene 4 primase was developed using circular ssDNA templates and T7 DNA polymerase. In this assay, DNA synthesis is dependent on the presence of nucleoside triphosphates (NTPs) (6, 7) and the T7 gene 4 protein (9, 17). The T7 gene 4 primase requires only ATP and CTP to synthesize the primers attached to DNA, which are predominantly of the sequence pppACN$_1$N$_2$, where N$_1$ and N$_2$ are normally C or A (9, 17). G and U can also be incorporated into the third and fourth positions but at a much lower rate (18). In vivo, it is tetraribonucleotides that are predominantly attached to the 5′ termini of the Okazaki fragments (19), and their sequences are the same as those synthesized

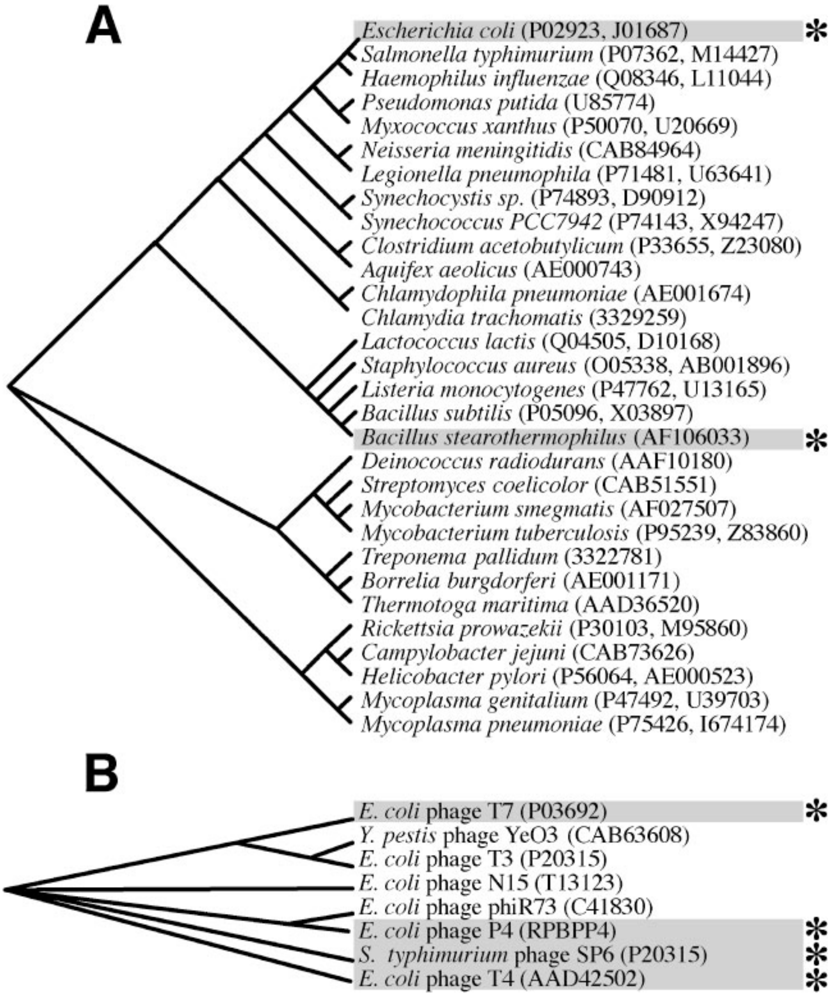


Figure 2 The prokaryotic DNA primases. (*A*) Bacterial encoding proteins significantly similar to the *E. coli* DnaG primase. (*B*) Bacteriophages with genes significantly similar to the T7 63-kDa gene 4 DNA primase. Genetic database accession numbers for the primase proteins are listed in parentheses. The amino acid sequences of putative DNA primases were aligned to generate the phylogenetic trees depicted. Alignments of most of the primases shown here are available on the World Wide Web courtesy of Mark Griep (http://chem-mgriep2.unl.edu/replic/primase.html). Primases discussed in detail in this review are marked (*).

in vitro (20,21). Each of the four deoxyribonucleotides is present at the RNA-DNA junction of Okazaki fragments (17,22,23). Oligoribonucleotide synthesis catalyzed by the T7 primase can also be measured directly by isolating the products synthesized in reactions not coupled to DNA synthesis (9). On large natural ssDNAs, a mixture of di- (pppAC), tri- (pppACC/A), and tetraribonucleotides (pppACCA/C and pppACAC/A) is synthesized. Similar products can also be detected in reactions using oligonucleotide templates of defined length, provided that appropriate template sequences, or primase recognition sites (see below), are present (24,25).

In addition to its role as a DNA primase, the gene 4 protein is also a DNA helicase. T7 gene 4 actually encodes two colinear proteins: a full-length 63-kilodalton (kDa) protein and a 56-kDa protein that is translated from an internal start codon located 189 bases from the 5′ end of the gene (26). The full-length 63-kDa gene 4 protein is both a primase and a helicase (27). The colinear 56-kDa gene 4 protein is a helicase but lacks primase activity (27). Only the 63-kDa gene 4 protein is necessary and sufficient for T7 growth, although a mixture of the two forms provides for the optimal rate of DNA synthesis in phage-infected cells (28,29).

E. coli *Primase* Early studies on *E. coli* defective in the *dnaG* gene implicated its product in the initiation of synthesis of Okazaki fragments both in vivo and in vitro. The 60-kDa DnaG protein, initially purified using an in vitro complementation assay, was subsequently shown to catalyze the synthesis of small RNA primers (30). To synthesize oligoribonucleotides efficiently on ssDNA, *E. coli* primase requires either an origin coated with *E. coli* ssDNA binding protein or the replicative helicase of *E. coli* (the product of the *dnaB* gene). Consequently, the action of the *E. coli* primase has been studied in the context of two different systems. The first system uses the G4 phage DNA (10), a close relative of ϕX174, as a ssDNA template. On G4 ssDNA, synthesis is dependent on the presence of *E. coli* ssDNA binding protein, and it begins at the site where replication of the complementary strand begins in vivo, termed the G4 origin (G4ori$_c$) (31). The RNA product is a 26- to 29-nucleotide-long transcript of G4ori$_c$ (32). In the second system, the activity of the DnaG protein is measured in the presence of the *E. coli* DnaB helicase. This system has been designated the "general priming reaction" because DnaB permits the primase to synthesize RNA primers on most ssDNAs (33). In general priming reactions, the *dnaG* protein catalyzes the synthesis of oligoribonucleotides 10 to 60 nucleotides long (34,35), although most are 11 residues (35). In both systems, primers begin with a pppAG dinucleotide at the 5′ end. Genes encoding DnaG-like DNA primases have also been identified in many other organisms (Figure 2A).

Bacteriophage P4 Primase Bacteriophage P4 is a temperate phage of *E. coli* that can be maintained autonomously in the cell as a multicopy plasmid, and it encodes some of its own replication apparatus (36). Mutations in the bacteriophage P4 α

gene product do not support plasmid DNA synthesis (37). Characterization of a purified P4 α protein revealed that the 88-kDa protein contains RNA polymerase activity (38, 39). To synthesize primers, P4 α protein requires a DNA template, NTPs, and Mg^{2+} and K^+ ions for optimal activity (40–42). On ssDNA, the P4 α protein synthesizes two- to five-nucleotide-long primers that begin with pppAG (41). The P4 α protein, like the T7 gene 4 protein, is also a helicase. However, unlike the T7 helicase, which translocates in a $5'$ to $3'$ direction, the P4 α protein moves with a $3'$ to $5'$ directionality (41). P4 α protein possesses a third activity. The protein binds duplex DNA with specificity for the six repeats of the sequence $5'$-TGTTCACC-$3'$ that are found at the P4 origin of replication (41). The primase, helicase, and origin binding activities reside in three separate functional domains mapping to the N terminus, a middle region, and the C terminus of the polypeptide, respectively (43, 44).

Bacteriophage T4 Primase The requirement for NTPs in a T4 DNA replication system in vitro implicated RNA synthesis in the priming of DNA synthesis (45). Dissection of these reconstituted systems revealed that primer synthesis was dependent on the product of T4 gene 61 (gp61). T4 phage defective in gene 61 have a reduced rate of DNA synthesis (46) and accumulate ssDNA (47, 48). Like the other primases in this class, the gp61 primase requires a helicase for optimal activity, namely the T4 gene 41 protein (gp41) (49). The gp61 and gp41 have molecular masses of 44 kDa and 58 kDa, respectively. In the presence of gp41 helicase, T4 primase synthesizes mainly pentanucleotides that begin with pppAC. Pentaribonucleotides are the predominant primer species at the $5'$ ends of Okazaki fragments during T4 DNA replication (49–51), and the Okazaki fragments initiate with pppACNNN (50, 52). Because T4 gp61 cannot prime T4 DNA synthesis in the absence of gp41, the functional primase can be considered a complex between the gene 41 and gene 61 proteins. The two-protein complex requires the presence of Mg^{2+} and K^+ ions and ssDNA for the synthesis of oligoribonucleotides.

SP6 Primase Another novel prokaryotic primase was purified from *Salmonella typhimurium* bacteriophage SP6, a distant cousin of phage T7. Phage SP6 is morphologically similar to the *E. coli* phage T7, yet the viruses are so distantly related that their two proteins share only 22% identical amino acids. The SP6 primase requires GTP and CTP to synthesize oligoribonucleotides on ssDNA templates that begin primarily with the sequence pppGC. Sequence comparisons have suggested that, like the phage polymerases from T7 and P4, the SP6 primase protein may also contain a replicative helicase function, although none has yet been identified (53).

Eukaryotic Primases (Polymerase α Associated)

The eukaryotic primases (Figure 3) typically purify as a complex of four proteins: the 180-kDa DNA polymerase α (Pol α), a 70-kDa protein called the polymerase B subunit, and the two primase subunits of approximately 49 and 58 kDa,

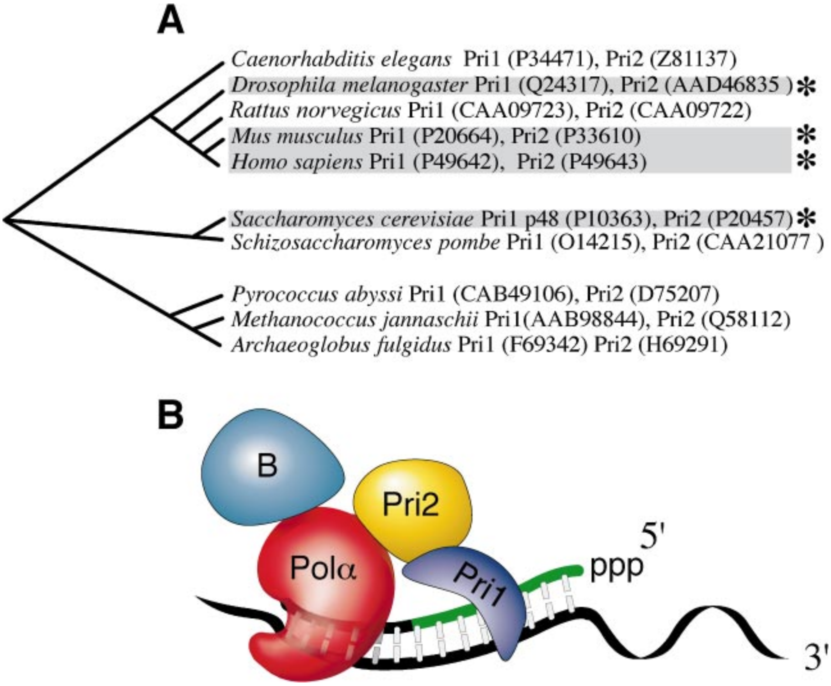


Figure 3 The eukaryotic family of DNA primases. (*A*) Organisms encoding proteins similar to both the large and small subunits of human DNA primase. Cladograms were generated from alignments of Pri1 proteins. Alignments are available on Mark Griep's web site (http://chem-mgriep2.unl.edu/replic/EukPri1AA.html). Primases discussed in this review are marked (*). (*B*) Cartoon depicting the organization of the eukaryotic Pol α/primase complex. The Pol α DNA polymerase (*red*) and its B subunit (*blue*) form a complex with the large primase subunit (*yellow*), encoded by the Pri2 gene, and the small primase subunit (*purple*), encoded by the Pri1 gene.

respectively. The smaller primase subunit contains the active site for oligoribonucleotide synthesis. The genes encoding all of the subunits of the Pol α/primase complex have been identified in humans, rats, mice, *Drosophila*, and the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, although the primase from *S. pombe* has not yet been purified or characterized. In addition, although the genes are not yet cloned, the primase from calf thymus has been purified and characterized. The recently sequenced genomes of several archaeons, particularly those from the family Euryarchaeota, contain regions homologous to genes encoding eukaryotic primases (Figure 3A).

Yeast Primase *S. cerevisiae* DNA primase copurifies with yeast DNA polymerase I (Pol I), the Pol α equivalent in yeast (54, 55). A monoclonal antibody specific for Pol I will coprecipitate this primase; the same antibody linked to a column can be used to separate the polymerase and primase subunits (55). Yeast primase synthesizes primers that are 8–10 nucleotides long, and yeast Pol I extends these primers at multiple sites on an M13 ssDNA template (56). Initiation begins only on templates containing polypyrimidines (57), whereas the first NTP incorporated at the 5′ end of the primer is a purine (58). The complex is composed of 180-, 86-, 62-, and 48-kDa subunits; the two smaller subunits compose the primase. The smallest, 48-kDa subunit (p48) alone catalyzes oligoribonucleotide synthesis. Genetic analyses confirm that both primase subunits are essential for yeast survival (59). The 62- and 48-kDa primase subunits physically interact with the 180-kDa polymerase subunit but not with the 86-kDa B subunit, which binds only to the 180-kDa polymerase (60). The p48 and p58 subunits of the yeast primase are encoded by two single-copy genes, PRI1 and PRI2, respectively. PRI1 is located on chromosome IX (61) and PRI2 is located on chromosome XI (62,63). Mutations in either PRI1 or PRI2 result in severe defects in cell growth and DNA synthesis that correlate with a high rate of mitotic recombination and spontaneous mutation (64). Other yeast mutants have been selected that are lethal to strains carrying primase defects. One of these mutants has been mapped to the MEC3 gene, which is part of the G_2 DNA damage checkpoint system, thus linking the primase with DNA repair (65).

Drosophila *Primase* From *Drosophila* extracts, DNA Pol α purifies as a complex with primase activity that has properties similar to the yeast primase discussed above (66) and can likewise be purified using immunoaffinity chromatography (67). The *Drosophila* Pol α/primase primes DNA synthesis on either poly(dT) or M13 ssDNA templates (66) by synthesizing RNA primers that are 8–15 nucleotides long (68). In the presence of deoxynucleoside triphosphates (dNTPs), Pol α extends the primers at multiple nonrandom sites (66,68). The Pol α/primase complex is composed of four subunits of 182, 73, 61, and 50 kDa (69). The 182-kDa subunit is the DNA polymerase, and the 61-kDa and 50-kDa subunits are required for efficient primer synthesis (69). The genes encoding the DNA polymerase subunit (70), its B subunit (71), and the large (72) and small (73) primase subunits have been cloned and sequenced. The recombinant purified small (50-kDa) subunit synthesizes oligoribonucleotides similar to those produced by the native complex but is more thermally labile than the intact complex (73).

Mouse Primase Mouse primase was first purified using monoclonal antibodies specific for human DNA Pol α (74,75). The heterodimeric primase is part of a complex containing the 180-kDa Pol α catalytic subunit and its 77-kDa B subunit (76). The primase subunits are 58 kDa and 49 kDa and together synthesize oligoribonucleotides 9–10 nucleotides long that prime DNA synthesis (77). The mouse primase requires NTPs and Mg^{2+} and synthesizes RNA on a variety of ssDNA templates (78,79). In the absence of DNA polymerase and dNTPs, mouse primase synthesizes multimeric oligoribonucleotides with a modal length of 9–10 nucleotides (80). The genes for the primase subunits have been cloned and expressed in *E. coli* and baculovirus (81). The properties of the recombinant proteins are similar to the native enzyme (82). The mouse primase 49-kDa and 58-kDa subunits map to chromosome 10 in the distal part of band D, and to region A5-B of chromosome 1, respectively (83).

Human Primase Like the eukaryotic primases described above, the human Pol α/primase (84, 85) can be purified from cell lysates using a monoclonal antibody (86). In humans, both polymerase subunits are phosphorylated on serine or threonine residues (87). The gene for the catalytic subunit of Pol α encodes a 165-kDa protein that has sequence similarity with the *E. coli* Pol I family of DNA polymerases (88). Due to its posttranslational modifications, human Pol α migrates on polyacrylamide gels as a protein with a higher apparent M_r of 180–195 kDa. The other three subunits of the human Pol α/primase complex have masses of 77, 59, and 50 kDa. The human primase complex synthesizes RNA primers 6–10 nucleotides long with ATP at the 5′ end. When dNTPs are present in the reaction, Pol α extends the primers to chains up to 3000 nucleotides long (86). The cDNAs for all four subunits have been cloned and expressed to produce an active recombinant complex (81, 89). The small primase subunit alone will synthesize oligoribonucleotides but is very unstable. The activity of the small primase protein can be stabilized by the addition of the large primase subunit or by the presence of divalent metal cations Mg^{2+} or Mn^{2+} during purification (89). The small primase subunit maps to the long arm of chromosome 12 (90), and the large subunit maps to two loci on chromosome 6 at bands 6p11.1–p12 (91).

Miscellaneous Primases

Not all known primases are homologous to proteins in either of the two classes outlined above. One set of these miscellaneous primases comprises viral proteins that are immunologically distinct from the eukaryotic primases. The members of another class are the products of plasmid *sog* genes. These "suppressors of *dnaG*" permit the growth of strains carrying temperature-sensitive alleles of *dnaG* at nonpermissive temperatures. Some, but not all, plasmid-encoded primases share homology with prokaryotic primases, but they also contain unique properties that distinguish them from other primases.

Transfer Protein Primases The first plasmid-encoded primases described were a series of proteins that are transported from donor to recipient *E. coli* cells during conjugation and are encoded by plasmids from the IncI and IncP compatibility groups (92). One of these transfer proteins is the product of the *sog* gene of the IncI plasmid ColIb-P9. As is the case with T7 gene 4, the *sog* gene encodes two polypeptides of 210 kDa (Sog_{210}) and 160 kDa (Sog_{160}) from separate in-frame translational start sites (93). The primase function is part of the amino terminus of the 210-kDa protein (93, 94). Two other transfer protein primases (called TraC proteins) are encoded by plasmids from the IncP group. One is the TraC protein of plasmid RP4 (IncPα), and another is the TraC protein of plasmid R751 (IncPβ) (95). Again, several forms of the TraC protein are encoded from different start sites. The RP4 *traC* gene encodes the 116-kDa TraC$_1$ protein and the 81-kDa TraC$_2$ protein (96). Plasmid RP751 *traC* encodes four proteins: TraC$_1$ (192 kDa), TraC$_2$ (152 kDa), TraC$_3$ (135 kDa), and TraC$_4$ (83 kDa). Even the smallest, TraC$_4$, is a functional primase (95), which suggests that in TraC proteins the primase domain likely resides in the C terminus. Segments of all these transfer proteins share sequences with the primase domain of the P4 primase (93), but only these small regions are similar to other prokaryotic primases. The other domains of these large proteins likely facilitate the transfer of genetic material from one bacterium to another.

ColE2 Rep Primase Another plasmid gene that encodes a novel primase is the *rep* gene of plasmid ColE2 (97). At 35 kDa, the Rep primase is smaller and lacks sequence homology with other primases. Remarkably, this small protein can prime DNA synthesis in the presence of plasmid DNA, *E. coli* DNA polymerase I, *E. coli* ssDNA binding protein, ATP, GTP, ADP, and dNTPs (98). The Rep protein primes DNA synthesis with ppAGA opposite the sequence 5′-TCTG-3′ in the plasmid origin. Synthesis requires ADP, ATP, and GTP, and products retain [β-^{32}P]ADP at their 5′ end, suggesting that RNA synthesis begins opposite the dT in the recognition site (99). Binding of the ColE2 Rep primase to the plasmid origin is so specific that it does not bind to the origin of the related plasmid ColE3 (100), the two origins differing by only 2 of 33 base pairs (97). Two regions near the C terminus of the Rep primase are likely involved in sequence-specific DNA binding (101). Unlike other known primases, the Rep primase facilitates DNA synthesis on a circular duplex DNA molecule.

Herpes Simplex Virus Primase Evidence that herpes simplex virus (HSV) encodes its own DNA primase was first found in HSV-infected HeLa cells (102). The HSV primase exists as a three-polypeptide complex consisting of the UL52, UL5, and UL8 gene products (103) that also contains 5′-3′ helicase activity (104). A subassembly of the UL5 (97-kDa) and UL52 (120-kDa) proteins is sufficient for helicase and primase activity (105, 106). The UL52 protein contains the catalytic center for primase activity (107), and the UL5 subunit is likely the helicase (108, 109). The herpes UL5/UL52 complex synthesizes 10- to 13-nucleotide-long primers on M13 ssDNA and unwinds DNA at a rate of two base pairs per second (110). In contrast with the host eukaryotic primase, HSV primase does not reinitiate from these primers to make primer multimers (110). Although it is not essential, the presence of the UL8 protein stimulates primer synthesis by the UL5 and UL52 components (111, 112), especially on templates coated with the herpes ssDNA binding protein, ICP8 (113). Similar genes have been identified in other viruses including equine HSV, human Epstein-Barr virus, human cytomegalovirus, and varicella-zoster virus. Alignments of these viral primases reveal amino acids likely to be critical for primer synthesis. One such alignment is available on the web site of Mark Griep at the University of Nebraska (http://chem-mgriep2.unl.edu/replic/VirPrAA.html).

INTERACTIONS BETWEEN PRIMASES AND DNA TEMPLATES

Primase must first bind a DNA template before synthesizing RNA primers (25, 114). The eukaryotic Pol α/primase complexes bind DNA with a K_d of 0.1 to 1.0 μM and protect 9 nucleotides of the primer and 13 nucleotides of the template from nuclease digestion (114, 115). To enable frequent primer synthesis on the lagging strand, most primases will synthesize oligoribonucleotides complementary to virtually any ssDNA to which they are bound. Experiments using several prokaryotic and viral primases have shown, however, that certain DNA sequences support a dramatically higher rate of primer synthesis. These sequences, termed primase recognition sites, differ for each of the primases examined and may play an important role in the coordination of lagging strand synthesis. Despite not having a stringent requirement for a specific recognition site, many eukaryotic DNA primases likewise display sequence specificity.

Prokaryotic and Viral Primase Recognition Sites

With ϕX174 (22) or M13 ssDNA (23) as a template, T7 primase enables T7 DNA polymerase to initiate DNA synthesis primarily at 13 and 9 sites, respectively. Most of the T7 primase recognition sites contain the sequences 5′-GGGTC-3′, 5′-TGGTC-3′, or 5′-GTGTC-3′, which direct the synthesis of the primers pppACCC, pppACCA, and pppACAC, respectively. The identities of the template recognition sites have been confirmed in vivo and in vitro by mapping the exact locations of five primer RNA to DNA transition sites on a fragment from the T7 genome (116) and using short synthetic oligonucleotide templates (24, 25). All T7 primase recognition sites contain the same trinucleotide, 5′-GTC-3′. Primer synthesis begins opposite the 3′ dT, and the 3′ cytosine is cryptic, meaning that its complement is not incorporated into the oligoribonucleotide products. No RNA synthesis is supported by short synthetic templates lacking either the cryptic dC or the adjacent dT in the recognition site (25). Templates containing modified dC or dT bases likewise support a lower rate of oligoribonucleotide synthesis than sequences containing the preferred recognition site 5′-GTC-3′ (117, 118). In particular, templates lacking the N3 nitrogen of the cryptic dC or the N3 nitrogen of dT support little detectable synthesis, highlighting potential interactions between the primase and functional groups involved in Watson-Crick base-pair formation (118).

Other phage primases preferentially initiate synthesis at different characteristic primase recognition sites. The SP6 primase begins primers with pppGC preferentially at sites containing the sequence 5′-GCA-3′ (53). The T4 primase recognizes the site 5′-GTT-3′ or 5′-GCT-3′ depending on certain conditions. On M13 ssDNA, all initiation sites share the sequence 5′-GTT-3′ or 5′-GCT-3′ and give rise to the primers pppACNNN and pppGCNNN (119). Only the 5′-GTT-3′ site is used on T4 DNA that contains hydroxymethylcytosine (120). The first dT in the recognition site is cryptic. Synthesis by the P4 primase begins at the phage origin of replication, where the P4 α typically starts primers with pppAG at the two-base sequence 5′-CT-3′ (41).

The bacterial *E. coli* DnaG primase also recognizes a trinucleotide sequence in DNA. On the G4ori$_c$ template, the *E. coli* DnaG protein catalyzes the synthesis of 26- to 29-nucleotide-long oligoribonucleotides (121, 122) that begin with pppAG at the sequence 5′-CTG-3′ (123). G4ori$_c$ consists of three stem-loop structures that lie on the 5′ side of the trinucleotide 5′-CTG-3′ sequence and direct the binding of *E. coli* ssDNA binding protein to the DNA. None of the three hairpins normally present in G4ori$_c$ are required for DnaG protein to initiate primer synthesis, but the sequence 5′-CTG-3′ is required (124). Other synthetic ssDNA templates containing 5′-CTG-3′ but not any of the remaining G4ori$_c$ sequence likewise support DnaG-catalyzed primer synthesis (124). This trinucleotide sequence is also recognized by DnaG protein at the origins of related phages St-1, ϕK, and α3 (125). Although early models invoked secondary structures of G4ori$_c$ as the recognition site for the primase, later studies strongly support a model in which the hairpin loops help expose the primase recognition site in the presence of *E. coli* ssDNA binding protein (126). The *E. coli* primase recognition site 5′-CTG-3′ remains free of ssDNA binding protein (126) and can interact directly with the primase (127). A second primase molecule binds to a region flanking the 5′ end of the primase recognition site, which also remains free of ssDNA binding protein (128). Thus, although DnaG protein is a monomer in solution (30, 129), a dimer may be required to support replication at replication origins (130).

Clearly, prokaryotic primases preferentially synthesize primers at certain DNA sequences. However, different sequences are recognized by the various proteins: 5′-CTG-3′ by the *E. coli* DnaG protein, 5′-GTT-3′ (or 5′-GCT-3′) by the T4 gene 61 protein, 5′-GCA-3′ by the SP6 primase, and 5′-GTC-3′ by the T7 primase. In each case, a trinucleotide sequence, in which the 3′ residue is cryptic, is required for recognition. The biological role of these trinucleotide primase recognition sites remains unclear.

Although it is not evolutionarily related to the prokaryotic primases, the herpes simplex primase has also been shown to synthesize primers preferentially at specific sequences (112). On ϕX174 DNA, purified herpes primase predominantly synthesizes primers of the sequence 5′-GGGAGGGUAG-3′ beginning opposite the 3′ dC in the template sequence 5′-ACCCTCC$\underline{\text{C}}$GA-3′. Only oligonucleotides containing this recognition sequence will inhibit RNA primer DNA synthesis by the HSV primase/polymerase complex (131).

Sequence-Specific Interactions by Eukaryotic Primases

Eukaryotic primases require the presence of pyrimidines in the ssDNA template for activity and bind polypyrimidine templates more tightly than they do polypurine templates (132, 133). This common characteristic is likely due to the fact that all primers are initiated with purine NTPs (134). However, on templates containing a mixture of purines and pyrimidines, eukaryotic primases do not initiate synthesis randomly but prefer certain sites. The most convincing evidence of eukaryotic primase sequence specificity comes from experiments that mapped specific primase initiation sites on simian virus 40 (SV40) DNA to a region located near the 65-base minimal origin of replication (79, 134–137). On SV40 DNA, replication begins almost exclusively at template dTs and is spaced regularly within, and to the 3′ side of, the 65-base-pair minimal origin (134–137). When given a ssDNA template made from the strand encoding the SV40 early mRNA, mouse primase begins synthesis at only five sites, four of which are clustered within 10 nucleotides of each other. On the other strand encoding the late mRNA, primase synthesizes primers at six sites adjacent to the 65-nucleotide minimal origin (79). Deletion of only six nucleotides in this origin significantly affects initiation site usage (137).

Mouse primase initiation has also been studied using ssDNA templates derived from the minute virus of mice. On these DNAs, primer synthesis begins mainly at 17 sites that all share a template sequence (5′-CCA-3′ or 5′-CCC-3′) 2–14 nucleotides downstream from the start site for primer synthesis (138, 139). Mouse primase binds DNA with these trinucleotide sequences more tightly than it does DNA lacking these sequences and has the highest affinity for templates containing 5′-CCC-3′ approximately 10 nucleotides downstream from the initiation sites for primer synthesis (139). Similar pyrimidine trinucleotides have also been identified in initiation sites used by primases purified from simian cells (134) and HeLa (140) cells. Thus, like the prokaryotic primases, human and mouse primases recognize a trinucleotide. However, whereas prokaryotic primases initiate synthesis directly opposite to their recognition sequences, eukaryotic primases initiate primers opposite to nucleotides flanking the 3′ end of the recognition sequence.

In a study using a Pol α/complex isolated from calf thymus, Suzuki et al (141) found that only 9 out of 140 ssDNA templates screened support detectable primer synthesis. The active templates contained an unusually high percentage of pyrimidines. Substitutions of adenylates for the pyrimidines in the active templates affect both the rate of primer synthesis and the affinity of the primase for the template. A similar template specificity was observed using primase free of Pol α (142). In these studies, the initiation of primer synthesis seems to require a six-base pyrimidine-rich tract of DNA, bounded by a 3′ purine, but unlike the mouse primase initiation sites, the calf thymus sites share no commonly conserved sequence (141). The nonrandom primer initiation by calf thymus Pol α/primase has been confirmed using other, longer templates (143). Initiation site usage is dependent on the reaction conditions and appears more random at higher concentrations of NTPs (144) and when Mn^{2+} is added to the reaction (145).

PRIMER SYNTHESIS

Primases catalyze the synthesis of oligoribonucleotides in a minimum of five discrete steps: template binding, NTP binding, initiation, extension to a functional primer, and primer transfer to DNA polymerase. We have proposed a simple

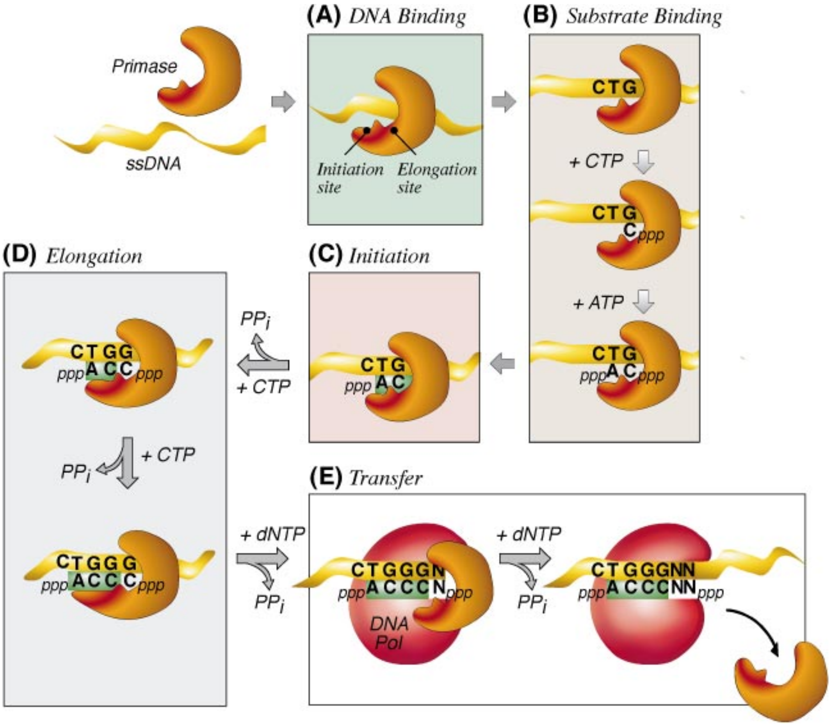


Figure 4 Steps in primer synthesis. This hypothetical mechanism is intended to depict primer synthesis by the T7 primase. Other primase mechanisms differ concerning NTPs used, primer sequence, primer length, and template recognition site. (*A*) DNA primase (*orange*) binds to ssDNA (*yellow*). (*B*) When the primase encounters an appropriate initiation site, two NTPs bind. The first NTP binds to the elongation site, eventually becoming the second nucleotide in the primer. The second NTP binds to the initiation site, and is incorporated at the 5′ end of the primer. (*C*) Primer synthesis is initiated by the formation of a dinucleotide and inorganic pyrophosphate from the two bound NTPs. (*D*) The growing oligoribonucleotide is transferred to the initiation site while additional NTPs bind to the elongation site; nucleotides are incorporated at the 3′ end of the primer. (*E*) Primer RNA is transferred to the replicative DNA polymerase that adds deoxynucleotides derived from dNTPs to their 3′ ends.

mechanism that involves two NTP binding sites on the primase protein (146). In this mechanism (Figure 4), the site at which the NTP is to be incorporated at the 5′ end of the primer is referred to as the initiation site. The second site, which binds the NTP added to the 3′ end of the primer, is referred to as the elongation site. At each elongation step of primer synthesis, the product ($n + 1$) oligonucleotide must be transferred to the initiation site so that another NTP may bind to the elongation site. The length of the oligonucleotide that can bind in the initiation site may limit the final length of the oligoribonucleotide synthesized by the primase.

Substrate Binding and Primer Initiation

After binding a DNA template (Figure 4*A*), primase must next bind two NTP substrates (Figure 4*B*) to catalyze the formation of a dinucleotide and inorganic pyrophosphate (Figure 4*C*). The NTP incorporated at the 5′ end of the primer is usually a purine, ATP or GTP, and retains its 5′ triphosphate moiety after dinucleotide synthesis. Most primases show little preference regarding the triphosphate end of the NTP incorporated at the 5′ end of the primer. NTP analogs with modified 5′ phosphate groups (146–148) and even nucleotides cross-linked to proteins via their 5′ phosphates are incorporated at the 5′ end of primers (149, 150).

In the phage T7 system, primers usually begin with pppAC. However, in the absence of DNA, T7 primase synthesizes other dinucleotides at a low rate in a reaction that is enhanced greatly when Mn^{2+} is present (151, 152). In these template-independent reactions, CMP predominates at the 3′ end of the dinucleotide, but each of the four canonical NTPs is incorporated at the 5′ end, suggesting that during initiation, CTP binds first and helps to locate primase recognition sites in DNA (153). Indeed, the affinity of T7 primase for DNA templates increases when primase is bound to either ATP or CTP (25). The K_m for CTP in the reaction catalyzed by T7 primase (0.8 mM) is higher than the cellular concentration of CTP, suggesting that the concentration of CTP could contribute to control of the rate of primer synthesis (146). Under optimal conditions, T7 primase can initiate primers at a rate of four dinucleotides per second (25), a rate sufficient to facilitate T7 DNA synthesis (154).

The kinetics of primer synthesis catalyzed by *E. coli* primase have been examined using a 23-nucleotide DNA template, which contains the primase recognition site 5′-GTC-3, the minimal template required for full-length (10- to 12-nucleotide) primers (155). In this system, under steady-state conditions, mainly full-length products are synthesized, and intermediates do not accumulate, suggesting that the rate-limiting step occurs before the synthesis of the dinucleotide. The rate of primer synthesis by *E. coli* primase on a 23-nucleotide template is quite slow at 0.00089 sec^{-1} (155). However, when the *E. coli* DnaB helicase is present, the rate of primer synthesis increases 15-fold (156).

Reported K_m values for NTPs in reactions catalyzed by eukaryotic primase range from 8 to 175 μM (78, 115, 133, 157, 158). The rate-limiting step during the synthesis of eukaryotic primers occurs before or during dinucleotide synthesis. With the calf thymus primase, after the formation of an E•DNA•NTP•NTP complex, dinucleotides are synthesized at a rate of 0.003 s^{-1} and are then rapidly extended to full-length primers (159).

As shown in Figure 4*B*, the NTPs may bind the primase in a particular order, with the first NTP binding to the primase at the elongation site. The best evidence supporting this binding order comes from studies using eukaryotic primases. Because eukaryotic primases do not initiate solely at specific primase recognition sites, they may incorporate either ATP or GTP at the 5′ end of the primer. By identifying the 5′-terminal nucleotide in primers synthesized at various NTP concentrations, Sheaff & Kuchta have demonstrated that the first NTP that binds calf thymus primase becomes the second nucleotide in the primer and determines where on the template synthesis begins (159).

Primer Elongation

During strand elongation, the growing oligoribonucleotide is transferred to the initiation NTP binding site to make room for another NTP to bind to the elongation site and subsequently attach to the 3′ hydroxyl of the oligoribonucleotide (Figure 4*D*). The length of oligoribonucleotides that are synthesized varies among the different primases and conceivably could be regulated by the length of the oligoribonucleotide that can be accommodated at the initiation NTP binding site of the enzyme. Oligoribonucleotides normally used as primers by DNA polymerase are designated full-length or unit-length primers. Unit-length primers are synthesized processively (80, 115, 146, 160) but are often not the only products observed in vitro. Shorter "abortive primers" accumulate with all primases, and under certain conditions, eukaryotic primases synthesize primer multimers by adding nucleotides to the ends of unit-length primers. Primer multimers are normally the length of multiple unit-length primers. For example, if the primase synthesizes primers 10 nucleotides long, then primer multimers would be 20, 30, or 40 nucleotides long.

The size of unit-length primers varies depending on each primase and the needs of its cognate replicative polymerase. Phage primases synthesize the shortest primers, normally fewer than 5 nucleotides. Most other primases synthesize primers slightly longer than 10 nucleotides (Table 1). During steady-state RNA synthesis catalyzed by primase, the quantities of the various products do not change over time, indicating a processive reaction (115, 146, 160, 161). Significant amounts of nucleotides are incorporated into abortive products, but the relative amount of shorter products can be minimized in conditions favoring the addition of the next nucleotide (146, 160). The physiological significance of abortive products remains elusive, but primases are able to rebind and elongate oligonucleotides, suggesting that abortive primers could serve as intermediates in primer synthesis (147).

Primer multimers are synthesized in reactions containing purified eukaryotic primases free of associated polymerase activity, or in reactions containing the Pol α/primase complex but free of dNTP substrates. For example, in the absence of Pol α, mouse primase continues to synthesize primers in modal increments of the original 9- to 11-base RNA primer up to 50 bases in length (78, 80). Yeast primase likewise rebinds and extends unit-length primers to synthesize multimeric primers during repeated cycles of processive oligoribonucleotide synthesis (161).

Multimeric primers are synthesized more frequently at low temperatures, which would favor the re-annealing of unit-length primers (160).

Several DNA primases select and incorporate NTPs with low fidelity. The *E. coli* primase can incorporate deoxynucleotides as well as ribonucleotides into primers and will add ribonucleotides to the 3′ hydroxyl of a DNA chain (34, 156, 162–164). The biological significance of these in vitro observations is not known, nor is it known if such mixed primers occur in vivo. Because of its ability to incorporate deoxyribonucleotides into primers, the DnaG protein can also incorporate dideoxyribonucleotides into primer chains, thus preventing their further extension (164). Besides lacking the ability to discriminate between NTPs and dNTPs, *E. coli* primase also lacks fidelity in the selection of the appropriate nucleotide base (128, 165). *E. coli* primase forms non–Watson-Crick base pairs with a rate approaching 1 in every 10 nucleotides polymerized (155). The *E. coli* primase shares low fidelity with eukaryotic primases, notably *Drosophila* primase, which incorporates deoxynucleotides and frequently synthesizes primers not complementary to the template (166). Low fidelity is not, however, a universal feature of all primases. For example, T7 primase differentiates between ribo and deoxy sugars and will misincorporate nucleotides at a detectable rate only when solely the incorrect nucleotide is present in a reaction (146).

Transfer of Primers to DNA Polymerase

To fulfill their primary function, DNA primases must not only initiate and elongate oligoribonucleotides but also transfer primers to DNA polymerase (Figure 4*E*). The communication between primase and DNA polymerase is most apparent when primer synthesis is examined in the presence and absence of DNA polymerase. In eukaryotic systems, where the primase and polymerase form a tight complex, the two activities are clearly interconnected. In the presence of ssDNA template, NTPs, and dNTPs, eukaryotic primase transfers primers directly to DNA Pol α without dissociating from the template or from the enzyme complex (167).

Whereas in eukaryotic systems primases and polymerases form tight complexes, in the prokaryotic replication fork, DNA polymerases and primases are more transiently related. The primase repeatedly associates and dissociates from the replication fork while catalyzing multiple rounds of primer synthesis (168–170). In *E. coli*, the DNA polymerase III (Pol III) enzyme complex influences primer synthesis, probably by competing with the primase for the primer 3′ end. In the presence of Pol III, the most abundant product is 8 nucleotides long instead of 11, and longer oligoribonucleotides are not synthesized (35, 162). In a complete system containing primase, Pol III, and dNTPs, the primers at the ends of Okazaki fragments have been found to be 1 to 9 nucleotides long in one study (171) and 9 to 14 in another (35). Yuzhakov et al (172) have proposed an elegant mechanism involving a three-point switch to explain the transfer of the primer from the *E. coli* primase to Pol III during DNA replication. They have demonstrated that Pol III, with its sliding-clamp processivity subunit β, cannot assemble onto primed DNA while the primase is still bound. To load onto a primed template, the Pol III χ subunit, part of the clamp-loading γ complex, makes contact with ssDNA binding protein, which coats the template. Because DnaG primase also contacts ssDNA binding protein to assemble properly on DNA, Yuzhakov et al have proposed that the χ subunit may act to displace the primase from the replication fork. Thus, ssDNA binding protein facilitates primer transfer by switching its contact between primase and the Pol III χ subunit (172).

Inhibitors

Inhibitors of DNA primases are valuable compounds not only for the elucidation of biochemical pathways but also as lead compounds to design drugs that block proliferative, bacterial, or viral diseases. Most of the compounds known to inhibit primases are nucleotide analogs. For example, *E. coli* primase will use 2′,3′-dideoxynucleoside 5′-triphosphates (ddNTPs) as substrates, and when incorporated, these compounds act as chain terminators (124). Another nucleotide analog, 2′-deoxy-2′-azidocytidine, potently inhibits DnaG and is incorporated into the primer, thus severely limiting primer extension (173). The T7 primase will not use either dNTPs or ddNTPs, but it will incorporate 3′-dCMP as a chain terminator, albeit at a rate 1/100 that of CMP (146).

The triphosphate derivatives of the anticancer drug cytarabine (cytosine arabinoside, araC) and the antiviral drug vidarabine (adenosine arabinoside, araA) both inhibit DNA primases (133, 158, 174, 175). AraATP is a potent inhibitor of calf thymus DNA primase with a K_i of approximately 2 μM (158) and is preferred by the enzyme over ATP, 2′-dATP, 3′-dATP, or ddATP (174). With the human eukaryotic primases, araCTP (1-β-D-arabinofuranosylcytosine 5′-triphosphate) and araATP (9-β-D-arabinofuranosyladenosine 5′-triphosphate) compete with GTP, with K_i values of 122 and 128 μM, respectively (133).

The potency of araCTP and araATP in inhibiting human primase is surpassed by halogenated nucleotide analogs, including FaraATP (9-β-D-arabinofuranosyl-2-fluoroadenosine 5′-triphosphate), and the even more potent inhibitors FMAC-TP [1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-methylcytosine 5′-triphosphate], FIAC-TP [1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine 5′-triphosphate], and 2′-F-araTTP [1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)thymine 5′-triphosphate] (133). FaraATP is a noncompetitive inhibitor of primer synthesis with a K_i 25-fold lower than the K_m of ATP (176). Eukaryotic primases will preferentially incorporate FaraATP into RNA primers 30-fold more efficiently than they do ATP. Once incorporated, FaraATP acts as a chain terminator. As a result, in the presence of FaraATP, large amounts of short abortive oligoribonucleotide are produced that are not extended to full-length primers (177). In addition, (dichloroanilino)purines and -pyrimidines selected to inhibit the herpes simplex virus type 1 helicase-primase have been shown to slow the replication of HSV1 in infected HeLa cells (178).

Other compounds that are not derived from nucleosides also inhibit primer synthesis. The DNA trypanocidal drug suramin [hexasodium sym-bis(*m*-aminobenzoyl-*m*-amino-*p*-methylbenzoyl-1-naphthylamino-4,6,8-trisulfonate)carbamide] inhibits the action of the Pol α/primase complex from human cells. Suramin inhibits both the primase, by competing with GTP, and Pol α, by competing with the template/primer (179). Another interesting primase inhibitor is sphingosine, a biological molecule involved in apoptosis, which inhibits primase but not DNA polymerases α or β (180). Sphingosine not only inhibits in vitro RNA-primed DNA synthesis but also slows the growth of the human leukemic cell line Hl-60 (181).

Roles of Helicases in Primer Synthesis

Helicases unwind and separate DNA using energy derived from the hydrolysis of NTPs. The effects of helicases on primer synthesis are most easily studied in systems where two functions are encoded by separate genes. In systems where the two functions are fused into a single protein, functional primase domains first must be isolated and compared with the full-length proteins. Such a separation has been carried out with both the T7 and P4 primases. A 253-amino-acid N-terminal fragment of the phage P4 α protein retains the ability to prime DNA synthesis (43). Likewise, a 271-amino-acid N-terminal fragment of the T7 primase retains the ability to synthesize oligoribonucleotides although it lacks helicase activity (182). These studies reveal three main roles for the helicase in primer synthesis. The helicase appears to tether the primase to DNA, transport the primase to primase recognition sites, and coordinate the actions of primase and DNA polymerase.

E. coli *DnaB Helicase* McMacken et al (183) first noted that although DnaG is required to synthesize the initial primer in the conversion of ϕX174 ssDNA to double-stranded DNA (dsDNA), DnaB is required to initiate and sustain the formation of multiple primers. These studies led to the hypothesis that the DnaB protein acts as a mobile promoter to transport DnaG to initiation sites (183). Because DnaB and DnaG bind together to form a physical complex (184, 185), the movement of the helicase in a $5'$ to $3'$ direction on ssDNA (186) can transport the primase to initiation sites. The addition of DnaB to primase reactions also allows DnaG to prime DNA synthesis on almost any ssDNA template (33). The importance of helicase movement in primer synthesis is highlighted by experiments using DnaB mutants that are unable to translocate. For example, one DnaB mutant in which Ala replaces Lys-236 (K236A) assembles into hexamers and binds DNA containing an origin of replication but will not hydrolyze ATP, the fuel for helicase translocation. When combined with the other components of the DNA replication apparatus, DnaB K236A will not support primer synthesis or Pol III assembly (187).

Although one role of the DnaB helicase clearly is to provide mobility for the DnaG primase to reach primase recognition sites, in the presence of helicase a fundamental change in primer synthesis seems to occur that is unrelated to translocation. On certain ssDNA templates, such as poly dT, the *E. coli* primase alone is relatively inert, but in the presence of the DnaB protein and ATP, the DnaG primase catalyzes the synthesis of oligoribonucleotides (33). In addition, DnaG primase will synthesize oligoribonucleotides on short templates lacking the primase recognition site 5′-CAG-3′ when attached to the DnaB helicase. Moreover, only in the presence of DnaB will the *E. coli* primase initiate oligoribonucleotides with GTP at the 5′ end. ATP is normally used (156). Yoda & Okazaki (188) have proposed that the association of DnaG with DnaB loosens primase specificity such that it can initiate RNA synthesis at the broader sequence 5′-PyPyPu-3′. This change in specificity may enable the primase to function differently on either the leading or lagging strand (189). Because the leading strand is synthesized continually, it needs to be primed only at the origin of replication. However, the lagging strand must be primed continually to begin each new Okazaki fragment, and the ability of the primase to work at more DNA sequences would allow frequent priming.

T4 Gene 41 Helicase The T4 gene 41 protein (gp41) is a helicase that uses NTP hydrolysis to move through duplex DNA with a 5′ to 3′ polarity. When complexed with the T4 gp61 primase, the gp41 helicase influences the rate of primer synthesis and both the composition and length of RNA primers. On a ϕX174 ssDNA template, the major products of gp61 alone are the dinucleotides pppAC and pppGC, whereas in the presence of gp41 the products are mainly pentamers of the sequence pppACNNN (120, 190, 191). Only a complex of the gp41 and gp61 has full catalytic activity in both primase and helicase assays (192–194). Protein cross-linking and gel-retardation assays reveal that the gp41 primase binds DNA as a monomer at an appropriate primase recognition site around which the helicase can assemble as a hexamer (195, 196).

T7 Gene 4 Helicase Domain Unlike the two primases described above, the T7 primase is covalently attached to the replicative T7 gene 4 helicase (Figure 5). In the presence of DNA and dTTP, the T7 gene 4 protein forms a hexamer (197, 198) that, when viewed in the electron microscope, appears as bilobal subunits surrounding DNA (199). The two lobes likely correspond to the helicase and primase domains. The larger lobe oriented at the 3′ end of the ssDNA likely corresponds to the helicase because the helicase moves 5′ to 3′ to separate the DNA to provide the primase with a ssDNA template (Figure 5*A*).

The first evidence that T7 helicase translocation was not necessary for primer synthesis was presented in early studies that showed that the T7 primase synthesizes primers in the presence of the dTTP analog β,γ-methylene-dTTP (200). In the presence of β,γ-methylene-dTTP, the gene 4 protein forms hexamers and binds

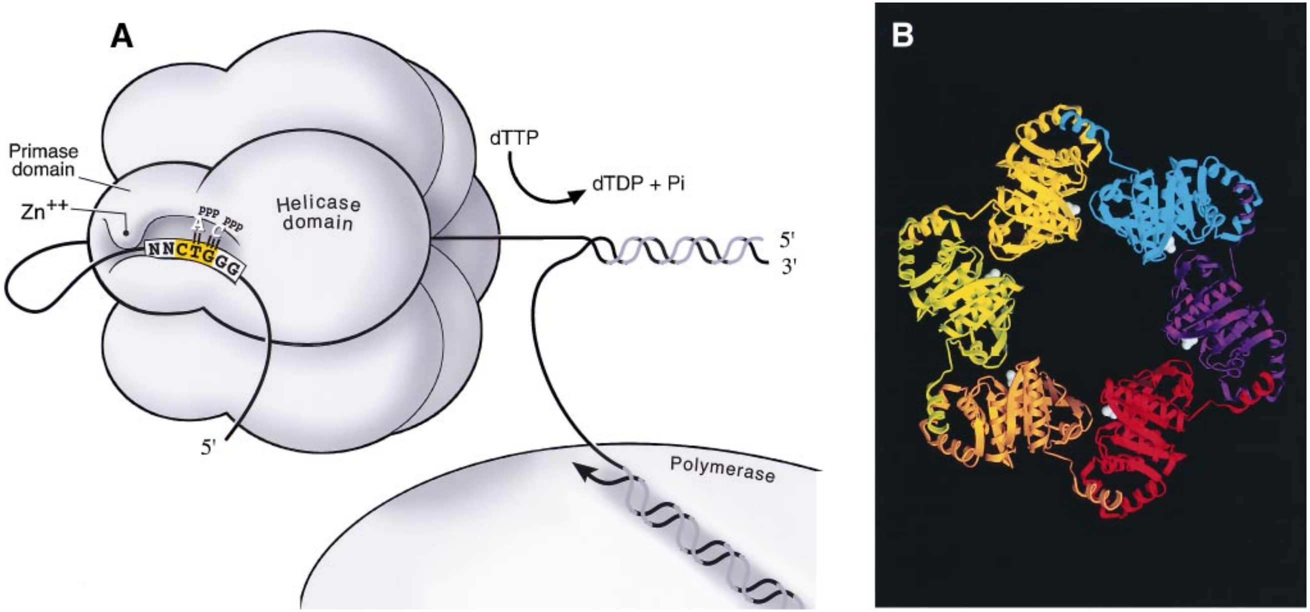


Figure 5 Role of the C-terminal helicase domain of the T7 primase in primer synthesis. (*A*) The hexameric T7 gene 4 protein complex is composed of subunits containing two large functional domains, an N-terminal primase domain (182) and a C-terminal helicase domain (209). The hexamer surrounds ssDNA on which it translocates in a 5′ to 3′ direction to separate the double helix. The primase synthesizes primers on the strand to which the helicase is bound but in the opposite direction, by reading the template in a 3′ to 5′ direction. The leading strand DNA polymerase continually synthesizes DNA using the other ssDNA strand. (*B*) X-ray crystal structure of the helicase domain of the hexameric T7 gene 4 protein. The dTTP fuel for helicase action (*gray*) is bound at the interface of the protein subunits. (*B* from Reference 210, courtesy of Tom Ellenberger and Michael Sawaya.)

DNA tightly but cannot translocate (201, 202). Consequently, the primase does not dissociate from the template to which it is bound. The primase does, however, synthesize primers if the bound DNA contains an appropriate primase recognition site and can function on primase recognition sites on DNA strands to which it is not bound tightly (203). The primase and helicase DNA-binding sites are physically sufficiently close, however, so that DNA lacking a primase recognition site bound to the helicase will block the interaction with other DNAs and inhibit oligoribonucleotide synthesis (25, 117).

Another approach to examining the role of helicase activity in primer synthesis has been to construct mutants encoding amino acid substitutions in the gene 4 helicase domain and then examine the effect of the substitutions on primase activity. The first T7 helicase mutants shown to affect primase activity contained substitutions in conserved helicase motif 1A, a P-loop NTP-binding site (13, 204, 205). One of these mutants, in which Gly-317 is changed to Val and Lys-318 is changed to Met, lacks the ability to hydrolyze NTPs and inhibits the helicase activities of the 56-kDa and 63-kDa wild-type gene 4 proteins. The G317V, K318M double mutant retains the ability to synthesize oligoribonucleotides but primes DNA synthesis on ϕX174 ssDNA only with the aid of wild-type 56-kDa gene 4 helicase protein (205). The more conservative mutation, changing Lys-318 to Ala, also abolishes detectable dTTPase and helicase activities (206). Unlike the double mutant, the K318A primase retains the ability to bind ssDNA in the presence of dTTP and assembles into hexamers. Consequently, the K318A mutant possesses more primase activity than the double mutant and can prime DNA synthesis without the aid of the 56-kDa gene 4 protein (206). Several other noteworthy mutations lie in the conserved helicase motif 4 (13) and indicate that this region is involved in DNA binding (198, 207).

C-terminal fragments of the gene 4 protein containing the residues required for helicase activity (208, 209) have been crystallized and their structures have been determined (210, 211). The structure of the gene 4 helicase domain (Figure 5*B*) is similar to that of other proteins that use the energy derived from NTP hydrolysis, such as other helicases, F_1ATPase, and the DNA strand exchange factor RecA. The helicase structures reveal that the amino acid side chains in helicase motif 1A are in intimate contact with the bound NTP, including K318, which contacts a nonbridging oxygen of the β-phosphate. Motif 4 is located on the inside of the hexamer, supporting the notion that DNA is bound on the inside of the ring structure (211). The smaller helicase fragment (210) lacks a region linking the primase and helicase domains and does not form a hexamer (209) but crystallizes as a helical filament that, when viewed down its axis, resembles the gene 4 hexameric rings seen in electron micrographs (199). The larger helicase fragment forms an asymmetric hexamer that binds a nonhydrolyzable ATP analog at four of the six possible NTP binding sites (211). The structures provide significant insights into a possible mechanism of DNA unwinding involving nucleotide hydrolysis (210, 211). Neither of these two gene 4 structures, however, contain any residues thought to be involved in primer synthesis.

Eukaryotic Helicases The interactions between eukaryotic primases and replicative helicases have been demonstrated most clearly in studies examining viral replication in human cells. The large tumor antigen (TAg) of simian virus 40 (SV40) is a hexameric helicase (212) that moves in a 3′ to 5′ direction on ssDNA. In the presence of ATP, two hexamers of TAg initiate SV40 replication by binding to the minimal origin of replication (213). The presence of TAg stimulates primer synthesis by Pol α/primase (214), and the proteins permit the synthesis of Okazaki fragments on SV40 duplex DNA templates coated with the human ssDNA binding protein, replication protein A (RP-A) (215). Initiation of SV40 DNA replication by Pol α/primase in the presence of the TAg helicase occurs on the lagging strand (213, 216) but not on the leading strand (217). Physical interactions between TAg and the Pol α/primase complex have been demonstrated. These interactions occur between TAg and specific domains of the 180-kDa polymerase α subunit (218, 219) and the polymerase B subunit (220). There is no evidence for a direct physical interaction between TAg and either primase subunit.

Matsumoto et al (221) have screened several host helicases for their ability to replace SV40 TAg. They found that DNA helicase II isolated from mouse FM3A cells stimulates DNA synthesis in a reaction dependent on RP-A, mouse DNA Pol α/primase complex, and *E. coli* DNA gyrase. Unlike TAg, DNA helicase II is a 5′ to 3′ helicase, suggesting that, like DnaB and other prokaryotic helicases, the protein may load onto the lagging strand DNA template to coordinate semidiscontinuous DNA replication. Purified DNA helicase B stimulates the synthesis of oligoribonucleotides by mouse DNA primase on M13 ssDNA and influences the length of the products by inhibiting the production of longer RNA primers (222). There may be additional replicative helicases. One possibility is DNA Pol α accessory protein, which purifies with and stimulates eukaryotic primase and has been shown to react with antibodies raised against TAg (223).

Roles of Single-Stranded DNA Binding Proteins in Primer Synthesis

Proteins that coat ssDNA affect so many reactions in DNA metabolism that it is often difficult to assess their relevance to specific events in vivo. The DNA primases are no exception, and physical interactions as well as dramatic enzymatic effects are seen when ssDNA binding proteins are added to DNA primases in vitro.

E. coli *ssDNA Binding Protein* *E. coli* ssDNA binding protein helps coordinate genome replication and in addition guides the assembly of the primase at origins of replication (128). G4ori$_c$ is made of three stem-loop structures that act to direct the binding of *E. coli* ssDNA binding protein, which subsequently directs the primase to DNA. Primer synthesis is restricted to a specific site, a hairpin that resists coating by ssDNA binding protein (126–128). The fact that ssDNA binding protein stimulates the rate of oligoribonucleotide synthesis on such templates suggests that some amino acid residues in ssDNA binding protein could participate in primase recognition site binding or in primer synthesis. The ssDNA binding protein is so intimately involved with the primase that it can be cross-linked to a modified ATP bound to the catalytic site of DnaG primase (149). Once attached, this analog can be extended to an oligoribonucleotide (150). The labeled section of *E. coli* ssDNA binding protein has been mapped to the region between Met48 and Trp88 (150), which suggests that this region may make up part of the active site for primer synthesis. Once primers are synthesized, ssDNA binding protein switches its contact from primase to the χ subunit to help transfer primers to DNA Pol III (172). Whereas in the absence of DnaB helicase, ssDNA binding protein aids primer synthesis, in the presence of DnaB, ssDNA binding protein actually appears to inhibit the primase. In the absence of ssDNA binding protein, the DnaG and DnaB proteins synthesize primers on ssDNA templates (34). However, when ssDNA is precoated with ssDNA binding protein, a DnaB/DnaG/Pol III complex will synthesize little or no RNA-primed DNA (33), most likely because various other proteins are required to load DnaB onto DNA coated with ssDNA binding protein (224).

T7 Gene 2.5 Protein In the phage T7 system, the ssDNA binding protein is encoded by phage gene 2.5, an essential gene for T7 DNA replication (225). T7 gene 4 primase binds to gene 2.5 protein–coated ssDNA in the absence of NTPs (168). A direct physical interaction between the T7 gene 4 protein and the C-terminal domain of the T7 gene 2.5 protein has been demonstrated by affinity chromatography (226). A genetically altered gene 2.5 protein lacking its acidic carboxyl-terminal domain does not bind to gene 4 protein, although it still binds to ssDNA (227). On ssDNA templates, the T7 gene 2.5 protein increases the frequency of initiation of DNA synthesis by more than 10-fold (168). This stimulation arises both from an increase in the rate of primer synthesis and from more efficient primer elongation by T7 DNA polymerase (24, 168).

T4 Gene 32 Protein Similar interactions occur between T4 gp61 primase and the T4 ssDNA binding protein, T4 gp32. The gp61 and gp32 proteins sediment together as a complex in the absence of DNA (228). Like the gene 2.5 protein of T7, gp32 has a negatively charged domain at its carboxyl terminus that is essential for interaction with the gp61 primase (228, 229). As in the *E. coli* system, when gp32 is present at concentrations sufficient to coat all of the ssDNA in a reaction, primer synthesis is essentially abolished (191, 230). However, this inhibition is relieved when the gene 44/62 and 45 accessory proteins are present (230). The effect of gp32 on primer synthesis may play a role in the coordination of primer synthesis with primer elongation by the T4 DNA polymerase. In the absence of gp32, only a small percentage of the oligoribonucleotides synthesized are used as primers, whereas in the presence of the gene 32 protein, essentially all primers are successfully transferred to T4 DNA polymerase (191).

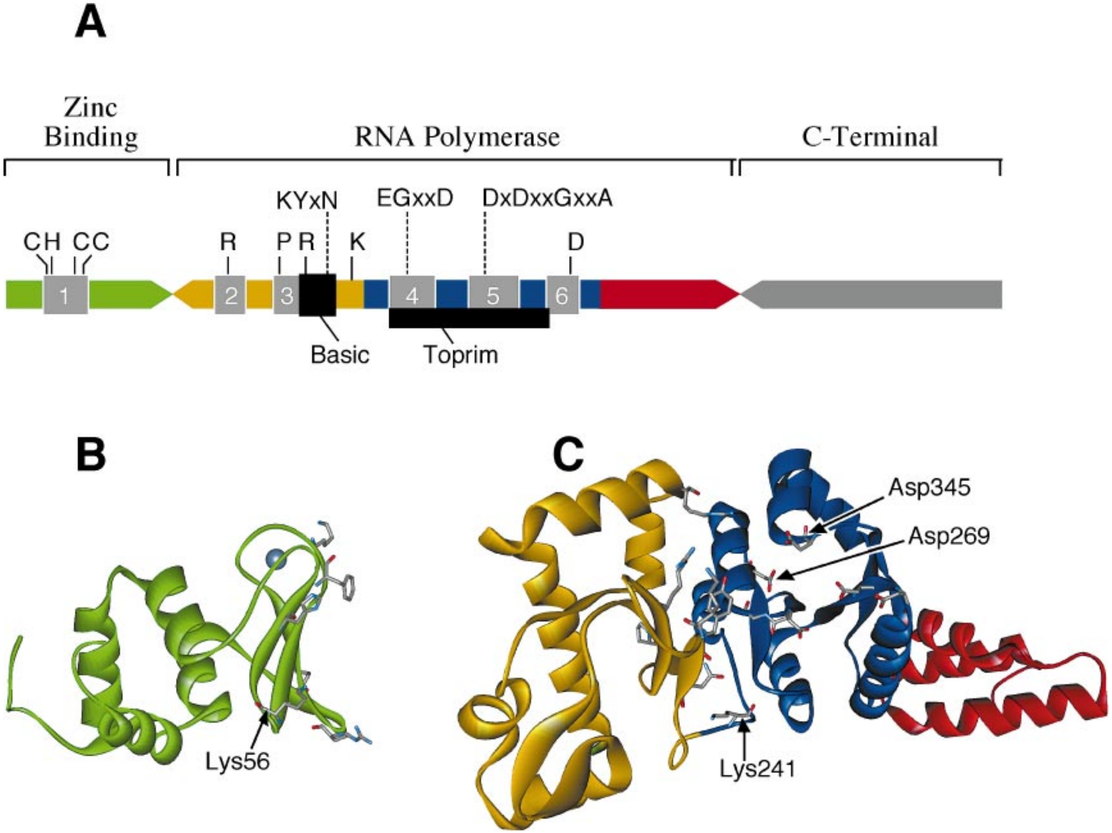
Eukaryotic ssDNA Binding Proteins Whereas the above bacterial and phage ssDNA binding proteins are relatively small, with masses typically between 20 and 30 kDa, their eukaryotic counterparts are larger and more complex. Eukaryotic ssDNA binding protein is usually referred to as replication protein A (RP-A), although in the past it has also been called replication factor A (RF-A) and when isolated from human cells, human ssDNA binding protein (HSSB). RP-A is a multifunctional heterotrimeric protein with subunits of 70 kDa (RP-A_1), 32 kDa (RP-A_2), and 11–14 kDa (RP-A_3). Primer synthesis as well as the synthesis of Okazaki fragments on the SV40 duplex DNA template requires the presence of RP-A in addition to TAg and the Pol α/primase complex (215). Pol α/primase will not synthesize DNA on natural DNA templates coated with RP-A unless SV40 TAg helicase is added to the reaction. The specific nature of the RP-A primase interaction is highlighted by the fact that only human RP-A, not yeast RP-A, stimulates human Pol α/primase in the presence of TAg on natural DNA templates (231). Physical interactions have been demonstrated, using enzyme-linked immunosorbent assay and other immuno techniques, between the human Pol α/primase complex, the SV40 TAg helicase, and human RP-A. RP-A associates both with the Pol α/primase complex and with the primase subunits alone (232). The 70-kDa subunit of RP-A plays an essential role in promoting the formation of the complex of SV40 TAg and the catalytic subunit of human polymerase α (220). The physical interaction of RP-A with Pol α/primase complex occurs via this 70-kDa subunit (232).

STRUCTURE OF PROKARYOTIC PRIMASES

The bacterial and phage primases have similar functional characteristics, and their primary amino acid sequences are derived from a common ancestor (13, 15). Proteolytic digestion of *E. coli* and T7 primase has shown that the prokaryotic primase proteins are composed of three regions (207, 223, 234). One fragment contains the N-terminal zinc binding region, the second contains the oligoribonucleotide synthesis site, and the third contains either a helicase or a region for interaction with a helicase. All prokaryotic primases share the N-terminal zinc binding region and the RNA polymerase domain but differ in their C terminus. In the primases that do not contain a helicase function, the C-terminal domain appears responsible for interaction with the replicative helicase. In the T7 and P4 primases, the C-terminal domain itself is the helicase. Below and in Figure 6, information on the structure and function of these three regions is summarized.

Zinc Binding Domain

Bernstein & Richardson (27) found that the 56-kDa form of the gene 4 protein lacks both the amino terminus of the 63-kDa gene 4 protein and the ability to synthesize primers. An analysis of the sequence of the missing N-terminal amino


A
Zinc Binding
RNA Polymerase
C-Terminal
KYxN
EGxxD
DxDxxGxxA
CH
CC
R
P R
K
D
1
2
3
4
5
6
Basic
Toprim
B
Lys56
C
Asp345
Asp269
Lys241

acids revealed the presence of a potential zinc motif that is also present in other primases from both prokaryotic and eukaryotic sources (13, 151, 235). The 63-kDa gene 4 protein contains one gram-atom of zinc per mole of protein whereas the 56-kDa gene 4 protein is devoid of zinc (151). All DNA primases contain a metal binding site, which is composed of four conserved Cys or His residues that could potentially coordinate zinc. In prokaryotic primases, the metal binding site is in the N terminus of the polypeptide. In contrast, metal binding sites are located in the C terminus of primases from viral systems, and centrally in the catalytic (small) subunits of eukaryotic primases. Substitution of serine for any one of the four conserved cysteine residues in the zinc motif of the T7 primase abolishes primer synthesis, but the mutant proteins retain some ability to synthesize a low level of random dinucleotides, which indicates that the zinc motif is not involved in phosphodiester bond formation (151). In the P4 primase, two of the four cysteines postulated to bind zinc have been substituted with glycine. Neither mutant G4 α protein can prime DNA synthesis (43). Like the T7 primase, purified preparations of the DnaG protein of *E. coli* contain one gram-atom of zinc per mole of protein (236). X-ray absorption spectroscopy indicates that the *E. coli* primase-bound zinc ion is coordinated by three sulfur atoms and one nitrogen atom at distances of 2.36 Å and 2.15 Å respectively (237). Although *E. coli* primase retains full activity when zinc is removed with *p*-(hydroxymercuri)-benzenesulfonate, the cysteine residues in the apoenzyme are more sensitive to modifying agents, and the enzyme is less stable in the absence of zinc (238).

Sequence-specific DNA binding by zinc motifs has been extensively studied in a number of systems, and a similar role has been proposed for the primase zinc motifs. So far there is only indirect evidence that zinc finger domains of prokaryotic primases are located near template DNA. In the T4 system, DNA protects sites in the N terminus of T4 primase that are sensitive to tryptic cleavage, and when a kinase recognition site is attached to the N terminus of gp61, the rate of protein phosphorylation is significantly lower in the presence of DNA (196). There is convincing evidence, however, that the zinc motif of the T7 primase directly

←

Figure 6 Structure of the prokaryotic DnaG-like DNA primase. (*A*) Schematic diagram of the domain structure and the conserved residues in DnaG-like primase. The six consensus sequences identified by Ilina et al (13) are *numbered and shaded in gray*. The charged basic region conserved in RNA polymerases (14) and the toprim domain (15) are in *black*. (*B*) Structure of the zinc ribbon domain of the DnaG protein from *B. stearothermophilus* [Protein Data Bank accession code 1D0Q (240)]; the zinc metal ion is shown in *gray*. Side chains of conserved residues that may interact with the ssDNA template are shown (240). His-33 of the T7 gene 4 protein, which helps locate primase recognition sites in DNA (153), aligns near Lys-46. (*C*) Structure of the RNA polymerase fragment of the *E. coli* DnaG primase [Protein Data Bank accession code 1DDE (249)]; it has three subdomains (*yellow, blue, red*). Conserved amino acid side chains are shown. Asp-269 and Asp-345 likely coordinate the required divalent metal cations (247). Lys-241 is critical for dinucleotide extension (245).

contacts DNA. Using site-directed mutagenesis and chimeric proteins, the roles of almost all of the amino acids in this region of the T7 primase have been investigated (152, 153, 239). Two notable T7 mutants contain substitutions for either Asp-31 or His-33 in the zinc binding region. T7 primase with substitutions for either of these residues no longer synthesizes oligoribonucleotides predominantly at the 5′-GTC-3′ recognition sequence. When His-33 is changed to Ala, the protein preferentially synthesizes oligoribonucleotides at sequences containing a cryptic purine instead of a cryptic cytosine (153).

The zinc binding region of the DnaG homolog from *Bacillus stearothermophilus* has recently been crystallized and its structure determined at 1.7-Å resolution (Figure 6*B*) (240). Unlike other zinc binding motifs, the primase region that coordinates zinc lacks α helices, thus classifying the primase motifs as members of the zinc ribbon subfamily of zinc binding motifs. The structure is similar to other zinc ribbon proteins, including transcription factor TFIIS (241) and yeast RNA polymerase II subunit 9 (242), and also resembles the T4 gene 32 protein (243), which binds ssDNA, and RP-A (244). Pan & Wigley (240) propose that DNA binds the primase on the surface of a β-sheet containing several residues conserved in other prokaryotic DNA primases that are in topologically equivalent positions in ssDNA binding proteins (240). One of these residues, Lys-56, aligns near His-33 in the T7 gene 4 protein (13); His-33 is an essential residue for determining primase sequence specificity (153).

RNA Synthesis Domain

In addition to the conservation of the zinc motif among prokaryotic primases, several other conserved sequence motifs (Figure 6*A*) are clustered in the central region (13). The conserved motifs are numbered from the N terminus to the C terminus, with the N-terminal zinc ribbon motif as motif 1. No functional or structural role has been attributed to primase motif 2, but many residues within motifs 3–6 appear to be critical for NTP binding and oligoribonucleotide synthesis.

The comparison of the sequence of the DnaG protein from *Haemophilus influenzae* with other primase sequences (14) helped to define further the conserved region near primase motif 3. This charged, basic region of primases is shared with a motif present in several large subunits of both prokaryotic and eukaryotic RNA polymerases, suggesting that this region could play a role in RNA synthesis. Labeling experiments with the *E. coli* DnaG primase have confirmed this prediction. Mustaev & Godson (149) used ATP derivatives to affinity-label the protein in the presence of G4ori$_c$ and ssDNA binding protein. The resulting primase was labeled in three regions and retained the ability to synthesize dinucleotides. Only one of the labeled residues, His-43, is within the zinc motif; the others, Lys-528 and either Lys-211, Lys-229, or Lys-241, are elsewhere in the protein. Lys-528 is near the C terminus in an area that is not conserved in other primases. Further mutagenesis of the DnaG protein has revealed that Lys-241, but not Lys-211 or Lys-229, is an essential part of the catalytic center (245). A mutant primase with Arg replacing Lys-241 synthesizes dinucleotides at a rate similar to that of the wild-type protein but few if any longer products (245).

Each of motifs 4, 5, and 6 has been implicated in the coordination of the NTP substrates and divalent metal cation(s) necessary for catalysis. The importance of the Glu residues in motif 4 in primer synthesis was first predicted by Strack et al (93), who recognized that the six-amino-acid sequence Glu-Gly-Tyr-Ala-Thr-Ala is shared between the P4 primase and the plasmid-encoded transfer protein primases. They showed that the conservative change of Glu-214 to Gln abolishes the P4 α protein's primase activity (93). From further sequence comparisons, Aravind et al (15) predicted that this Glu residue is the general base needed to facilitate nucleotide polymerization and that a D×D sequence in motif 5 acts to coordinate Mg^{2+} (15). The D×D in motif 5 is also present in the UL52 subunit of herpes primase, and the replacement of either of these aspartates with alanines greatly diminishes herpes primase activity (107, 246). Godson et al (247) have used Fe^{2+} ion cleavage and mutagenesis to locate the residues involved in the metal ion binding site of the *E. coli* primase. One of the important aspartates (D269) is in the Glu-Gly-Tyr-Met-Asp sequence in motif 4 and is only four amino acids removed from the conserved Glu discussed above. This Asp is conserved in all prokaryotic primases except P4 primase, which has a threonine in the corresponding position. The other metal ion–coordinating aspartates are in a D×D sequence located in motif 6. This Asp-Pro-Asp sequence is conserved in all DnaG-like bacterial primases (248), but only the first aspartate in this sequence is present in the phage primases.

The X-ray crystal structure of the RNA polymerase domain of the *E. coli* primase has been determined independently by two groups (249, 250) and is shown in Figure 6*C*. The structure of the RNA polymerase region of DnaG is unlike that of other DNA and RNA polymerases. Instead of a structure resembling a right hand, the primase is a shallow, cashew-shaped molecule composed of three subdomains. The N-terminal subdomain contains primase motifs 2 and 3, and the center subdomain contains motifs 4–6. As predicted by sequence comparisons, the center region resembles a domain, present in topoisomerases and similar metal-binding phosphotransfer proteins, that has been designated the toprim fold (15). Motifs 4 and 5 contain conserved negatively charged residues preceded by runs of hydrophobic residues that form a compact β/α fold. The residues that are strictly conserved in all bacterial DnaG proteins cluster around the central valley of the protein, which suggests that this region forms the active site (249, 250).

C-Terminal Domain

The C-terminal domains of the prokaryotic primases are not evolutionarily conserved. The only primases with similar amino acid sequences in this region are those possessing a helicase function, namely the P4 and T7 primases. The structure of the C-terminal domain (Figure 5*B*) of the T7 gene 4 protein has been determined by X-ray crystallography (210, 211). In addition to providing helicase function, acidic residues in the C terminus of the T7 primase have been implicated in binding the T7 DNA polymerase (251). Direct evidence for a physical interaction between the T7 primase and polymerase has been obtained with gel filtration, native polyacrylamide gel electrophoresis, and surface plasmon resonance techniques. Such complexes can be isolated free of unbound protein and have been shown to mediate RNA-primed DNA synthesis in the presence of NTPs and dNTPs (23). Deletion of the 17 C-terminal residues of the gene 4 protein abolishes this interaction, but the mutant protein retains primase and helicase activities (252).

The helicase portion of the P4 α protein likely adopts a RecA-like fold similar to the T7 helicase, but unlike the T7 protein, the C-terminal region of the P4 primase contains an additional origin binding function. Thus, the P4 α protein has a primase domain with a zinc ribbon at the N terminus, and a C-terminal region that contains both helicase and DNA binding activities (44). Amino acid substitutions in the middle section of the P4 α protein abolish ATPase and DNA unwinding ability, whereas peptides containing only the C-terminal 142 residues retain the ability to bind DNA (44). Alteration of the C-terminal third of the protein affects both the helicase and ATPase activities of purified proteins, suggesting that P4 helicase and DNA binding activities are related (44).

Less is known about the C-terminal domains of the prokaryotic primases lacking helicase activity. In the bacterial primases and in certain phage primases, the only known function of this domain is to interact with helicases or with other proteins at the replication fork. The physical interaction between the *E. coli* DnaG primase and DnaB helicase has been mapped to the C-terminal domain of DnaG (170, 184, 233, 253). When the C terminus of DnaG is removed by proteolysis, the protein retains the ability to synthesize oligoribonucleotides but is inactive in the helicase-dependent priming assay (170). A peptide fragment containing only the C-terminal DnaG domain blocks the stimulatory effect of the DnaB helicase and increases the size of Okazaki fragments (170). A DnaG protein containing a single amino acid substitution in the C terminus (Q576A) directs the synthesis of dramatically longer Okazaki fragments (254), binds DnaB with a significantly lower affinity than does the wild-type protein (184), and supports the initiation of bidirectional replication less efficiently than the native protein (255). Other mutants that map to the C terminus of *E. coli* DnaG may also influence the coordination of the primase in replication. These include *parB* and the slow stop *dnaG* mutation *dnaG*2903, in which the chromosomes partition abnormally during cell division (170, 256, 257). Whereas the weak or transient interaction between *E. coli* DnaG and DnaB proteins makes their binding difficult to measure (184), the homologous proteins from *B. stearothermophilus* have been clearly demonstrated to form a stable complex that can be isolated by gel filtration chromatography. As in the *E. coli* system (170, 184, 223), the C-terminal domain of *B. stearothermophilus* DnaG is necessary and sufficient to form a complex and to stimulate the enzymatic activities of the DnaB helicase (185).

STRUCTURE OF EUKARYOTIC PRIMASES

Far less is known about the structure of eukaryotic primases than of prokaryotic primases, in part because the larger, more complex eukaryotic genes are less readily expressed in recombinant systems. As detailed above, eukaryotic primases exist as part of a four-protein complex. Such a typical complex has an apparent mass of more than 300 kDa. For example, the Pol α/primase complex from yeast has a M_r of 344 kDa, a sedimentation coefficient of 8.9S, and a Stokes radius of 81.1 Å (258). The largest subunit, typically 165–180 kDa, contains the active site for DNA synthesis, and the smallest, typically about 48 kDa, synthesizes the RNA primer. Of the two other subunits, the larger, 70-kDa subunit likely regulates polymerase activity whereas the other, 58-kDa protein aids the primase.

The two smallest subunits of the four-protein Pol α/primase complex are generally regarded as the primase. Because the small primase subunit contains the active site for primer synthesis, it is often referred to as Pri1. The large primase subunit, which may coordinate primase and polymerase action, is called Pri2 (Figure 3*B*). Antibodies raised against the small subunit, but not those binding to the large subunit, potently inhibit primer synthesis. GTP and ATP are more efficiently cross-linked to the small subunit (259,260). However, when the active site of the yeast primase is labeled using an autocatalytic ATP derivative, both large and small subunits are modified, suggesting that both are involved in primer synthesis (261). With the mouse primase, ATP can be cross-linked to both subunits, although the NTP affinity of the small subunit is higher than that of the large subunit. In this species, the small subunit, rather than the large subunit, adsorbs to ssDNA-cellulose (82). Similar experiments with the human primase suggest that, although ATP selectively binds to the small subunit under certain conditions, synthesis actually occurs in a region where the two subunits contact each other (262). The small subunit of human primase alone will extend RNA primers annealed to DNA templates, but the larger subunit is required for primer initiation (263).

Comparisons of the amino acid sequences of the small as well as the large subunits of eukaryotic primases reveal five different conserved motifs (77,81,263, 264). All of the conserved charged residues in the small subunit of mouse primase have been changed to alanine and the mutant enzymes have been analyzed. Almost all conserved charged residues appear necessary for efficient primer synthesis, in particular E105, D109, and D111 in motif IV (265). The region surrounding D109 and D111 in motif IV shares homology with the D × D motif found in motif 5 of the prokaryotic primases, the HSV primase, and other polymerases from the Pol X family (246). Two arginines in motif IV could aid NTP binding because, when they are changed to alanine, the K_m of the mouse primase for GTP increases significantly (265). Motif IV in the small subunit also contains four cysteines that could form a zinc ribbon like the ones seen in prokaryotic primase (151,265). Motif V of the small subunits of eukaryotic primase also shares homology with the DNA polymerase β (Pol β) and the Pol X family of polymerases (266) centered on another essential aspartate of the mouse primase (265). Near this aspartate is a highly conserved arginine that in Pol β makes contact with the phosphate backbone of the primer strand (267). When the corresponding arginine (R304) of human primase is substituted with alanine, the enzyme more readily initiates primers with guanosine (rather than GTP), which suggests that the arginine side chain contacts the phosphates of the 5′ terminal NTP (266). A region around conserved motif III of the large subunit shares sequences with DNA Pol β. In Pol β this region is part of the 8-kDa domain that is involved in DNA binding, which suggests a similar role for the primase large subunit (266).

REGULATION OF PRIMASE ACTIVITY

Because all cells are not continually dividing, primase activity must be highly regulated. Most of the evidence for primase regulation comes from studies carried out with eukaryotic primases. Generally, primase activity seems to correlate with cellular DNA replication during S phase (268). Such regulation could be extremely important, as shown by the link between amplification of the primase gene PRIM1 and osteosarcoma (269). It is to be hoped that understanding the role that primase undoubtedly plays in cell division and proliferation will lead to improved cancer therapies.

There are several potential explanations for the observed variations in primase activity in cells. At the level of transcription, it is unclear whether the cellular amounts of primase mRNAs vary with cell cycle (264, 270). In mouse FM3A cells, mRNA levels for the p180 and p70 subunits vary throughout the cell cycle, with both transcripts increasing two- to threefold during mid–S phase (264). The genes encoding human Pol α and the two primase subunits are down-regulated during cell differentiation (271). The transcripts encoding the yeast primase subunits have been shown to peak late in G_1 together with other genes for DNA synthesis (62, 272).

Another possible mechanism for primase regulation could involve posttranslational modifications of the components of the primase complex. In yeast (273), mice (264), and humans (274), the polymerase B subunit is known to be phosphorylated in a cell-cycle–dependent manner. In humans, both the 180-kDa and 86-kDa proteins are phosphorylated, and two candidate kinases have been identified. They are two cyclin-dependent kinases, cyclin E/cdk2 and cyclin A/cdk2, that help regulate the transition from G_1 to S phase. Purified cyclin A/cdk2 phosphorylates both Pol α and the B subunit, resulting in a Pol α/primase with a reduced ability to initiate DNA synthesis. The second kinase, cyclin E/cdk2, primarily phosphorylates p180 and yields a more active Pol α/primase (275). Further biochemical studies of the actions of these and other kinases may reveal why primase activity, and hence DNA replication, is at its maximum during G_1/S phase and at its nadir during G_2/M.

A third possible reason for the observed variability in primase activity is that Pol α/primase is physically sequestered from the rest of the replication complex.

In eukaryotic cells, DNA replication is thought to take place around the nuclear matrix. A considerable amount of human primase is bound to the nuclear matrix (276,277), and the amount bound to the nuclear matrix varies throughout the cell cycle (278). In addition, a nuclear localization signal has been found on the large subunit of the mouse primase, indicating that the large subunit might facilitate the transport of primase to the nucleus (279). A similar role for the mouse polymerase B subunit has been proposed (280).

Other cellular proteins undoubtedly influence primer synthesis. For example, primase regulation may be linked to the phosphorylation of histone H1, which in its dephosphorylated state inhibits primer synthesis in a dose-dependent manner. The phosphorylation of histone H1 relieves some of this inhibition (281,282). Another proposed regulator of human primase activity is the retinoblastoma protein, which normally exists in both phosphorylated and dephosphorylated states in the cell. The phosphorylated form of retinoblastoma protein binds to the Pol α/primase complex and stimulates primase activity significantly more than does the dephosphorylated form (283). Besides the helicases and ssDNA binding proteins discussed above, numerous cellular factors have been identified and purified that stimulate the activity of the Pol α/primase complex (85, 284–287). It is likely that these stimulatory factors represent other important components of the replication fork.

FUTURE DIRECTIONS

Primases from a number of sources have been purified and extensively characterized. The recently elucidated structures of fragments of DnaG proteins have provided the first glimpses of primases at an atomic level. The functions of the various domains of the prokaryotic primases have been mapped using the plethora of available mutant primases with altered biochemical functions. However, the lack of structures containing templates or substrates makes speculation regarding the molecular details of the primase reaction difficult. Structural studies with eukaryotic primases are likewise anxiously anticipated. When compared to our detailed understanding of other proteins at the replication fork, such as DNA polymerase and helicase, we still know relatively little about DNA primase. As the critical first step in genomic replication, primase action clearly must be investigated further.

ACKNOWLEDGMENTS

Grants from the United States Public Health Service (AI-06045) and the American Cancer Society (NP-1Z) support work in the laboratory of CCR. A grant from the American Cancer Society (PF-4452) supported DNF. We thank Drs. Ingrid Richardson, Lisa Rezende, and Karyn Frick for critically reading the manuscript and Steve Moskowitz for help with figure preparation.

LITERATURE CITED

1. Watson JD, Crick FHC. 1953. *Nature* 171:737–38
2. Ogawa T, Okazaki T. 1980. *Annu. Rev. Biochem.* 49:421–57
3. Brutlag D, Schekman R, Kornberg A. 1971. *Proc. Natl. Acad. Sci. USA* 68:2826–29
4. Wolfsen J, Dressler D. 1972. *Proc. Natl. Acad. Sci. USA* 69:2682–86
5. Strätling W, Knippers R. 1973. *Nature* 245:195–97
6. Hinkle DC, Richardson CC. 1975. *J. Biol. Chem.* 250:5523–29
7. Scherzinger E, Litfin F. 1974. *Eur. J. Biochem.* 46:179–87
8. Scherzinger E, Seiffert D. 1975. *Mol. Gen. Genet.* 141:213–32
9. Scherzinger E, Lanka E, Morelli G, Seiffert D, Yuki A. 1977. *Eur. J. Biochem.* 72:543–58
10. Rowen L, Kornberg A. 1978. *J. Biol. Chem.* 253:758–64
11. Foiani M, Lucchini G, Plevani P. 1997. *Trends Biochem. Sci.* 22:424–27
12. Salas M. 1991. *Annu. Rev. Biochem.* 60:39–71
13. Ilyina TV, Gorbalenya AE, Koonin EV. 1992. *J. Mol. Evol.* 34:351–57
14. Versalovic J, Lupski JR. 1993. *Gene* 136:281–86
15. Aravind L, Leipe DD, Koonin EV. 1998. *Nucleic Acids Res.* 26:4205–13
16. Studier FW. 1969. *Virology* 39:562–74
17. Romano LJ, Richardson CC. 1979. *J. Biol. Chem.* 254:10476–82
18. Kusakabe T, Richardson CC. 1997. *J. Biol. Chem.* 272:5943–51
19. Shinozaki K, Okazaki T. 1977. *Mol. Gen. Genet.* 154:263–67
20. Seki T, Okazaki T. 1979. *Nucleic Acids Res.* 7:1603–19
21. Ogawa T, Okazaki T. 1979. *Nucleic Acids Res.* 7:1621–33
22. Tabor S, Richardson CC. 1981. *Proc. Natl. Acad. Sci. USA* 78:205–9
23. Nakai H, Richardson CC. 1986. *J. Biol. Chem.* 261:15217–24
24. Mendelman LV, Richardson CC. 1991. *J. Biol. Chem.* 266:23240–50
25. Frick DN, Richardson CC. 1999. *J. Biol. Chem.* 274:35889–98
26. Dunn JJ, Studier FW. 1983. *J. Mol. Biol.* 166:477–535
27. Bernstein JA, Richardson CC. 1988. *Proc. Natl. Acad. Sci. USA* 85:396–400
28. Mendelman LV, Notarnicola SM, Richardson CC. 1992. *Proc. Natl. Acad. Sci. USA* 89:638–42
29. Rosenberg AH, Patel SS, Johnson KA, Studier FW. 1992. *J. Biol. Chem.* 267:15005–12
30. Bouché J-P, Zechel K, Kornberg A. 1975. *J. Biol. Chem.* 250:5995–6001
31. Hourcade D, Dressler D. 1978. *Proc. Natl. Acad. Sci. USA* 75:1652–56
32. Bouché J-P, Rowen L, Kornberg A. 1978. *J. Biol. Chem.* 253:765–69
33. Arai K, Kornberg A. 1979. *Proc. Natl. Acad. Sci. USA* 76:4308–12
34. Arai K, Kornberg A. 1981. *J. Biol. Chem.* 256:5267–72
35. Zechner EL, Wu CA, Marians KJ. 1992. *J. Biol. Chem.* 267:4054–63
36. Lindqvist BH, Six EW. 1971. *Virology* 43:1–7
37. Gibbs W, Goldstein RN, Wiener R, Lindqvist B, Calendar R. 1973. *Virology* 53:24–39
38. Barrett KJ, Gibbs W, Calendar R. 1972. *Proc. Natl. Acad. Sci. USA* 69:2986–90
39. Barrett KJ, Blinkova A, Arnold G. 1983. *J. Virol.* 48:157–69
40. Krevolin MD, Calendar R. 1985. *J. Mol. Biol.* 182:509–17
41. Ziegelin G, Scherzinger E, Lurz R, Lanka E. 1993. *EMBO J.* 12:3703–8

42. Orejas RD, Ziegelin G, Lurz R, Lanka E. 1994. *Nucleic Acids Res*. 22:2065–70
43. Ziegelin G, Linderoth NA, Calendar R, Lanka E. 1995. *J. Bacteriol*. 177:4333–41
44. Ziegelin G, Calendar R, Lurz R, Lanka E. 1997. *J. Bacteriol*. 179:4087–95
45. Morris CF, Sinha NK, Alberts BM. 1975. *Proc. Natl. Acad. Sci. USA* 72:4800–4
46. Mufti S, Bernstein H. 1974. *J. Virol*. 14: 860–71
47. Hamlett N, Berger H. 1975. *Virology* 63:539–67
48. Broker TR. 1973. *J. Mol. Biol*. 81:1–16
49. Nossal NG. 1980. *J. Biol. Chem*. 255:2176–82
50. Kurosawa Y, Okazaki T. 1979. *J. Mol. Biol*. 135:841–61
51. Liu C-C, Alberts BM. 1980. *Proc. Natl. Acad. Sci. USA* 77:5698–702
52. Okazaki T, Kurosawa Y, Ogawa T, Seki T, Shinozaki K, et al. 1978. *Cold Spring Harbor Symp. Quant. Biol*. 43:203–19
53. Tseng TY, Frick DN, Richardson CC. 2000. *Biochemistry* 39:1643–54
54. Singh H, Dumas LB. 1984. *J. Biol. Chem*. 259:7936–40
55. Plevani P, Badaracco G, Augl C, Chang LMS. 1984. *J. Biol. Chem*. 259:7532–39
56. Badaracco G, Bianchi M, Valsasnini P, Magni G, Plevani P. 1985. *EMBO J*. 4: 1313–17
57. Badaracco G, Valsasnini P, Foiani M, Benfante R, Lucchini G, Plevani P. 1986. *Eur. J. Biochem*. 161:435–40
58. Roth Y-F. 1987. *Eur. J. Biochem*. 165:473–81
59. Francesconi S, Longhese MP, Piseri A, Santocanale C, Lucchini G, Plevani P. 1991. *Proc. Natl. Acad. Sci. USA* 88:3877–81
60. Brooke RG, Dumas LB. 1991. *J. Biol. Chem*. 266:10093–98
61. Lucchini G, Francesconi S, Fioani M, Badaracco G, Plevani P. 1987. *EMBO J*. 6:737–42
62. Foiani M, Santocanale C, Plevani P, Lucchini G. 1989. *Mol. Cell. Biol*. 9:3081–87
63. Purnelle B, Tettelin H, Van Dyck L, Skala J, Goffeau A. 1993. *Yeast* 9:1379–84
64. Longhese MP, Jovine L, Plevani P, Lucchini G. 1993. *Genetics* 133:183–91
65. Longhese MP, Fraschini R, Plevani P, Lucchini G. 1996. *Mol. Cell. Biol*. 16:3235–44
66. Conaway RC, Lehman IR. 1982. *Proc. Natl. Acad. Sci. USA* 79:2523–27
67. Kuroda K, Kagiyama-Takahashi R, Shinomiya T. 1990. *J. Biochem*. 108:926–33
68. Conaway RC, Lehman IR. 1982. *Proc. Natl. Acad. Sci. USA* 79:4585–88
69. Kaguni LS, Rossignol JM, Conaway RC, Banks GR, Lehman IR. 1983. *J. Biol. Chem*. 258:9037–39
70. Melov S, Vaughan H, Cotterill S. 1992. *J. Cell Sci*. 102:847–56
71. Cotterill S, Lehman IR, McLachlan P. 1992. *Nucleic Acids Res*. 20:4325–30
72. Huikeshoven H, Cotterill S. 1999. *Biochim. Biophys. Acta* 1445:359–62
73. Bakkenist CJ, Cotterill S. 1994. *J. Biol. Chem*. 269:26759–66
74. Tanaka S, Hu SZ, Wang TS, Korn D. 1982. *J. Biol. Chem*. 257:8386–90
75. Yagura T, Kozu T, Saneyoshi M, Hiraga S, Nagano H. 1983. *J. Biol. Chem*. 258:13070–75
76. Yagura T, Kozu T, Seno T, Tanaka S. 1987. *Biochemistry* 26:7749–54
77. Prussak CE, Almazan MT, Tseng BY. 1989. *J. Biol. Chem*. 264:4957–63
78. Tseng BY, Ahlem CN. 1983. *J. Biol. Chem*. 258:9845–49
79. Tseng BY, Prussak CE. 1989. *Nucleic Acids Res*. 17:1953–63
80. Suzuki M, Enomoto T, Masutani C, Hanakoa F, Yamada M, Ui M. 1989. *J. Biol. Chem*. 264:10065–71
81. Stadlbauer F, Brueckner A, Rehfuess C, Eckerskorn C, Lottspeich F, et al. 1994. *Eur. J. Biochem*. 222:781–93
82. Copeland WC, Wang TS-F. 1993. *J. Biol. Chem*. 268:26179–89
83. Adler DA, Tseng BY, Wang TS-F, Disteche CM. 1991. *Genomics* 9:642–46

84. Tseng BY, Ahlem CN. 1982. *J. Biol. Chem.* 257:7280–83
85. Yagura T, Kozu T, Seno T, Saneyoshi M, Hiraga S, Nagano H. 1983. *J. Biol. Chem.* 258:13070–75
86. Wang TS, Hu S, Korn D. 1984. *J. Biol. Chem.* 259:1854–65
87. Wong SW, Paborsky LR, Fisher PA, Wang TS, Korn D. 1986. *J. Biol. Chem.* 261:7958–68
88. Wang TS, Wong SW, Korn D. 1989. *FASEB J.* 3:14–21
89. Schneider A, Smith RW, Kautz AR, Weisshart K, Grosse F, Nasheuer HP. 1998. *J. Biol. Chem.* 273:21608–15
90. Cloutier S, Hamel H, Champagne M, Yotov WV. 1997. *Genomics* 43:398–401
91. Shiratori A, Okumura K, Nogami M, Taguchi H, Onozaki T, et al. 1995. *Genomics* 28:350–53
92. Lanka E, Scherzinger E, Gunther E, Schuster H. 1979. *Proc. Natl. Acad. Sci. USA* 76:3632–36
93. Strack B, Lessl M, Calendar R, Lanka E. 1992. *J. Biol. Chem.* 267:13062–72
94. Wilkins BM, Boulmois GJ, Lanka E. 1981. *Nature* 290:217–21
95. Miele L, Strack B, Kruft V, Lanka E. 1991. *DNA Seq.* 2:145–62
96. Lanka E, Lurz R, Kroger M, Furste JP. 1984. *Mol. Gen. Genet.* 194:65–72
97. Yasueda H, Horii T, Itoh T. 1989. *Mol. Gen. Genet.* 215:209–16
98. Takechi S, Itoh T. 1995. *Nucleic Acids Res.* 23:4196–201
99. Takechi S, Matsui H, Itoh T. 1995. *EMBO J.* 14:5141–47
100. Kido M, Yasueda H, Itoh T. 1991. *Nucleic Acids Res.* 19:2875–80
101. Shinohara M, Itoh T. 1996. *J. Mol. Biol.* 257:290–300
102. Holmes AM, Wietstock SM, Ruyechan WT. 1988. *J. Virol.* 62:1038–45
103. Crute JJ, Tsurumi T, Zhu L, Weller SK, Olivio PD, et al. 1989. *Proc. Natl. Acad. Sci. USA* 86:2186–89
104. Crute JJ, Mocarski ES, Lehman IR. 1988. *Nucleic Acids Res.* 16:6585–96
105. Calder JM, Stow ND. 1990. *Nucleic Acids Res.* 18:3573–78
106. Dodson MS, Lehman IR. 1991. *Proc. Natl. Acad. Sci. USA* 88:1105–9
107. Klinedinst DK, Challberg MD. 1994. *J. Virol.* 68:3693–701
108. Crute JJ, Bruckner RC, Dodson MS, Lehman IR. 1991. *J. Biol. Chem.* 266:21252–56
109. Zhu L, Weller SK. 1992. *J. Virol.* 66:469–79
110. Crute JJ, Lehman IR. 1991. *J. Biol. Chem.* 266:4484–88
111. Sherman G, Gottlieb J, Challberg MD. 1992. *J. Virol.* 66:4884–92
112. Tenney DJ, Hurlburt WW, Micheletti PA, Bifano M, Hamatake RK. 1994. *J. Biol. Chem.* 269:5030–35
113. Le Gac NT, Villani G, Hoffmann JS, Boehmer PE. 1996. *J. Biol. Chem.* 271:21645–51
114. Thompson HC, Sheaff RJ, Kuchta RD. 1995. *Nucleic Acids Res.* 23:4109–15
115. Anarbaev RO, Vladimirova OV, Lavrik OI. 1995. *Eur. J. Biochem.* 228:60–67
116. Fujiyama A, Kohara Y, Okazaki T. 1981. *Proc. Natl. Acad. Sci. USA* 78:903–7
117. Mendelman LV, Kuimelis RG, McLaughlin LW, Richardson CC. 1995. *Biochemistry* 34:10187–93
118. Searls T, Chen DL, Lan T, McLaughlin LW. 2000. *Biochemistry* 39:4375–82
119. Cha T-A, Alberts BM. 1986. *J. Biol. Chem.* 261:7001–10
120. Nossal NG, Hinton DM. 1987. *J. Biol. Chem.* 262:10879–85
121. Sims J, Dressler D. 1978. *Proc. Natl. Acad. Sci. USA* 75:3094–98
122. Fiddes JC, Barrell BG, Godson GN. 1978. *Proc. Natl. Acad. Sci. USA* 75:1081–85
123. Hiasa H, Sakai H, Tanaka K, Honda Y, Komano T, Godson GN. 1989. *Gene* 84:9–16
124. Swart JR, Griep MA. 1993. *J. Biol. Chem.* 268:12970–76

125. Sims J, Capon D, Dressler D. 1979. *J. Biol. Chem.* 254:12615–28
126. Sun W, Godson GN. 1993. *J. Biol. Chem.* 268:8026–39
127. Sun W, Godson GN. 1996. *J. Bacteriol.* 178:6701–5
128. Sun W, Godson GN. 1998. *J. Mol. Biol.* 276:689–703
129. Wickner SH, Wright M, Hurwitz J. 1973. *Proc. Natl. Acad. Sci. USA* 70:1613–18
130. Stayton MM, Kornberg A. 1983. *J. Biol. Chem.* 258:13205–15
131. Tenney DJ, Sheaffer AK, Hurlburt WW, Bifano M, Hamatake RK. 1995. *J. Biol. Chem.* 270:9129–36
132. Holmes AM, Cheriathundam E, Bollum FJ, Chang LMS. 1985. *J. Biol. Chem.* 260:10840–46
133. Parker WB, Cheng Y-C. 1987. *Mol. Pharmacol.* 31:146–51
134. Yamaguchi M, Hendrickson EA, DePamphilis ML. 1985. *Mol. Cell Biol.* 5:1170–83
135. Hay RT, DePamphilis ML. 1982. *Cell* 28:767–79
136. Hay RT, Hendrickson EA, DePamphilis ML. 1984. *J. Mol. Biol.* 175:131–57
137. Tseng BY, Ahlem CN. 1984. *Proc. Natl. Acad. Sci. USA* 81:2342–46
138. Faust EA, Nagy R, Davey SK. 1985. *Proc. Natl. Acad. Sci. USA* 82:4023–27
139. Davey SK, Faust EA. 1990. *J. Biol. Chem.* 265:3611–14
140. Vishwanatha JK, Yamaguchi M, DePamphilis ML, Baril EF. 1986. *Nucleic Acids Res.* 14:7305–23
141. Suzuki M, Savoysky E, Izuta S, Tatebe M, Okajima T, Yoshida S. 1993. *Biochemistry* 32:12782–92
142. Suzuki M, Izuta S, Savoysky E, Sakurai T, Simbulan C, et al. 1993. *Biochem. Mol. Biol. Int.* 29:645–52
143. Harrington C, Perrino FW. 1995. *Nucleic Acids Res.* 23:1003–9
144. Kirk BW, Harrington C, Perrino FW, Kuchta RD. 1997. *Biochemistry* 36:6725–31
145. Kirk BW, Kuchta RD. 1999. *Biochemistry* 38:10126–34
146. Frick DN, Kumar S, Richardson CC. 1999. *J. Biol. Chem.* 274:35899–907
147. Kusakabe T, Richardson CC. 1997. *J. Biol. Chem.* 272:12446–53
148. Matsuo H, Moriguchi T, Takagi T, Kusakabe T, Buratowski S, et al. 2000. *J. Am. Chem. Soc.* 122:2417–21
149. Mustaev AA, Godson GN. 1995. *J. Biol. Chem.* 270:15711–18
150. Godson GN, Mustaev AA, Sun W. 1998. *Biochemistry* 37:3810–17
151. Mendelman LV, Beauchamp BB, Richardson CC. 1994. *EMBO J.* 13:3909–16
152. Kusakabe T, Richardson CC. 1996. *J. Biol. Chem.* 271:19563–70
153. Kusakabe T, Hine AV, Hyberts SG, Richardson CC. 1999. *Proc. Natl. Acad. Sci. USA* 96:4295–300
154. Lechner RL, Richardson CC. 1983. *J. Biol. Chem.* 258:11185–96
155. Swart JR, Griep MA. 1995. *Biochemistry* 34:16097–106
156. Johnson SK, Bhattacharyya S, Griep MA. 2000. *Biochemistry* 39:736–44
157. Brooks M, Dumas LB. 1989. *J. Biol. Chem.* 264:3602–10
158. Kuchta RD, Willhelm L. 1991. *Biochemistry* 30:797–803
159. Sheaff RJ, Kuchta RD. 1993. *Biochemistry* 32:3027–37
160. Kuchta RD, Reid B, Chang LMS. 1990. *J. Biol. Chem.* 265:16158–65
161. Singh H, Brooke RG, Pausch MH, Williams GT, Trainor C, Dumas LB. 1986. *J. Biol. Chem.* 261:8564–69
162. McMacken R, Kornberg A. 1978. *J. Biol. Chem.* 253:3313–19
163. Sun W, Godson GN. 1998. *J. Biol. Chem.* 273:16358–65
164. Rowen L, Kornberg A. 1978. *J. Biol. Chem.* 253:770–74
165. Benz EW Jr, Reinberg D, Vicuna R, Hurwitz J. 1980. *J. Biol. Chem.* 255:1096–106

166. Cotterill S, Chui G, Lehman IR. 1987. *J. Biol. Chem.* 262:16105–8
167. Sheaff RJ, Kuchta RD, Ilsley D. 1994. *Biochemistry* 33:2247–54
168. Nakai H, Richardson CC. 1988. *J. Biol. Chem.* 263:9818–30
169. Wu CA, Zechner EL, Marians KJ. 1992. *J. Biol. Chem.* 267:4030–44
170. Tougu K, Peng H, Marians KJ. 1994. *J. Biol. Chem.* 269:4675–82
171. Ogawa T, Arai K, Okazaki T. 1983. *J. Biol. Chem.* 258:13353–58
172. Yuzhakov A, Kelman Z, O'Donnell M. 1999. *Cell* 96:153–63
173. Reichard P, Rowen L, Eliasson R, Hobbs J, Eckstein F. 1978. *J. Biol. Chem.* 253: 7011
174. Kuchta RD, Ilsley D, Kravig KD, Schubert S, Harris B. 1992. *Biochemistry* 31:4720–28
175. Harrington C, Perrino FW. 1995. *J. Biol. Chem.* 270:26664–69
176. Catapano CV, Chandler KB, Fernandes DJ. 1991. *Cancer Res.* 51:1829–35
177. Catapano CV, Perrino FW, Fernandes DJ. 1993. *J. Biol. Chem.* 268:7179–85
178. Crute JJ, Lehman IR, Gambino J, Yang TF, Medveczky P, et al. 1995. *J. Med. Chem.* 38:1820–25
179. Ono K, Nakane H, Fukushima M. 1985. *Nucleic Acids Symp. Ser.* 16:249–52
180. Simbulan CM, Tamiya-Koizumi K, Suzuki M, Shoji M, Taki T, Yoshida S. 1994. *Biochemistry* 33:9007–12
181. Tamiya-Koizumi K, Murate T, Suzuki M, Simbulan CM, Nakagawa M, Takemura M, et al. 1997. *Biochem. Mol. Biol. Int.* 41:1179–89
182. Frick DN, Baradaran K, Richardson CC. 1998. *Proc. Natl. Acad. Sci. USA* 95:7957–62
183. McMacken R, Ueda K, Kornberg A. 1977. *Proc. Natl. Acad. Sci. USA* 74:4190–94
184. Lu YB, Ratnakar PV, Mohanty BK, Bastia D. 1996. *Proc. Natl. Acad. Sci. USA* 93:12902–7
185. Bird LE, Pan H, Soultanas P, Wigley DB. 2000. *Biochemistry* 39:171–82
186. LeBowitz JH, McMacken KJ. 1986. *J. Biol. Chem.* 261:4738–48
187. Fang L, Davey MJ, O'Donnell M. 1999. *Mol. Cell* 4:541–53
188. Yoda K, Okazaki T. 1991. *Mol. Gen. Genet.* 227:1–8
189. Bhattacharyya S, Griep MA. 2000. *Biochemistry* 39:745–52
190. Hinton DM, Nossal NG. 1987. *J. Biol. Chem.* 262:10873–78
191. Cha T-A, Alberts BM. 1990. *Biochemistry* 29:1791–98
192. Venkatesan M, Silver LL, Nossal NG. 1982. *J. Biol. Chem.* 257:12426–34
193. Hinton DM, Silver LL, Nossal NG. 1985. *J. Biol. Chem.* 260:12851–57
194. Richardson RW, Nossal NG. 1989. *J. Biol. Chem.* 264:4725–31
195. Dong F, von Hippel PH. 1996. *J. Biol. Chem.* 271:19625–31
196. Jing DH, Dong F, Latham GJ, von Hippel PH. 1999. *J. Biol. Chem.* 274:27287–98
197. Patel SS, Hingorani MM. 1993. *J. Biol. Chem.* 268:10668–75
198. Notarnicola SM, Park K, Griffith JD, Richardson CC. 1995. *J. Biol. Chem.* 270:20215–24
199. Egelman EH, Yu X, Wild R, Hingorani MM, Patel SS. 1995. *Proc. Natl. Acad. Sci. USA* 92:3869–73
200. Scherzinger E, Lanka E, Hillenbrand G. 1977. *Nucleic Acids Res.* 4:4151–63
201. Matson SW, Richardson CC. 1985. *J. Biol. Chem.* 260:2281–87
202. Hingorani MM, Patel SS. 1996. *Biochemistry* 35:2218–28
203. Kusakabe T, Baradaran K, Lee J, Richardson CC. 1998. *EMBO J.* 17:1542–52
204. Notarnicola SM, Richardson CC. 1993. *J. Biol. Chem.* 268:27198–207
205. Mendelman LV, Notarnicola SM, Richardson CC. 1993. *J. Biol. Chem.* 268:27208–13
206. Patel SS, Hingorani MM, Ng WM. 1994. *Biochemistry* 33:7857–68

207. Washington MT, Rosenberg AH, Griffin K, Studier FW, Patel SS. 1996. *J. Biol. Chem.* 271:26825–34
208. Bird LE, Hakansson K, Pan H, Wigley DB. 1997. *Nucleic Acids Res.* 25:2620–26
209. Guo S, Tabor S, Richardson CC. 1999. *J. Biol. Chem.* 274:30303–9
210. Sawaya MR, Guo S, Tabor S, Richardson CC, Ellenberger T. 1999. *Cell* 99:167–77
211. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB. 2000. *Cell* 101:589–600
212. Stahl H, Droge P, Knippers R. 1986. *EMBO J.* 5:1939–44
213. Fanning E, Knippers R. 1992. *Annu. Rev. Biochem.* 61:55–85
214. Collins KL, Kelly TJ. 1991. *Mol. Cell Biol.* 11:2108–15
215. Matsumoto T, Eki T, Hurwitz J. 1990. *Proc. Natl. Acad. Sci. USA* 87:9712–16
216. Tsurimoto T, Melendy T, Stillman B. 1990. *Nature* 346:534–39
217. Murakami Y, Eki T, Hurwitz J. 1992. *Proc. Natl. Acad. Sci. USA* 89:952–56
218. Dornreiter I, Hoss A, Arthur AK, Fanning E. 1990. *EMBO J.* 9:3329–36
219. Dornreiter I, Copeland WC, Wang T. 1993. *Mol. Cell. Biol.* 13:809–20
220. Collins KL, Russo AAR, Tseng BY, Kelly TJ. 1993. *EMBO J.* 12:4555–66
221. Matsumoto K, Seki M, Masutani C, Tada S, Enomoto T, Ishimi Y. 1995. *Biochemistry* 34:7913–22
222. Saitoh A, Tada S, Katada T, Enomoto T. 1995. *Nucleic Acids Res.* 23:2014–18
223. Crouch E, Miller S, Wilson V, Busbee D. 1997. *Mutat. Res.* 374:109–23
224. Marians KJ. 2000. *Trends Biochem. Sci.* 25:185–89
225. Kim YT, Richardson CC. 1993. *Proc. Natl. Acad. Sci. USA* 90:10173–77
226. Kim YT, Tabor S, Churchich JE, Richardson CC. 1992. *J. Biol. Chem.* 267:15032–40
227. Kim YT, Richardson CC. 1994. *J. Biol. Chem.* 269:5270–78
228. Burke RL, Munn M, Barry J, Alberts BM. 1985. *J. Biol. Chem.* 260:1711–22
229. Burke RL, Alberts BM, Hosoda J. 1980. *J. Biol. Chem.* 255:11484–93
230. Richardson CC, Nossal NG. 1989. *J. Biol. Chem.* 264:4732–40
231. Melendy T, Stillman B. 1993. *J. Biol. Chem.* 268:3389–95
232. Dornreiter I, Erdile LF, Gilbert IU, von Winker D, Kelly TJ, Fanning E. 1992. *EMBO J.* 11:769–76
233. Sun W, Tormo J, Steitz TA, Godson GN. 1994. *Proc. Natl. Acad. Sci. USA* 91:11462–66
234. Yong Y, Romano LJ. 1995. *J. Biol. Chem.* 270:24509–17
235. Toh H. 1986. *FEBS Lett.* 194:245–48
236. Stamford NPJ, Lilley PE, Dixon NE. 1992. *Biochim. Biophys. Acta* 1132:17–25
237. Powers L, Griep MA. 1999. *Biochemistry* 38:7413–20
238. Griep MA, Lokey ER. 1996. *Biochemistry* 35:8260–67
239. Hine AV, Richardson CC. 1994. *Proc. Natl. Acad. Sci. USA* 91:12327–31
240. Pan H, Wigley DB. 2000. *Struct. Fold. Des.* 8:231–39
241. Qian X, Gozani SN, Yoon HS, Jeon CJ, Agarwal K, Weiss MA. 1993. *Biochemistry* 32:9944–59
242. Wang B, Jones DN, Kaine BP, Weiss MA. 1998. *Structure* 6:555–69
243. Shamoo Y, Friedman AM, Parsons MR, Konigsberg WH, Steitz TA. 1995. *Nature* 376:362–66
244. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L. 1997. *Nature* 385:176–81
245. Sun W, Schoneich J, Godson GN. 1999. *J. Bacteriol.* 181:3761–67
246. Dracheva S, Koonin EV, Crute JJ. 1995. *J. Biol. Chem.* 270:14148–53
247. Godson GN, Schoenich J, Sun W, Mustaev AA. 2000. *Biochemistry* 39:332–39
248. Szafranski P, Smith CL, Cantor CR. 1997. *Biochim. Biophys. Acta* 1352:243–48

249. Keck JL, Roche DD, Lynch AS, Berger JM. 2000. *Science* 287:2482–86
250. Podobnik M, McInerney P, O'Donnell M, Kuriyan J. 2000. *J. Mol. Biol.* 300:353–62
251. Nakai H, Richardson CC. 1986. *J. Biol. Chem.* 261:15208–16
252. Notarnicola SM, Mulcahy HL, Lee J, Richardson CC. 1997. *J. Biol. Chem.* 272:18425–33
253. Tougu K, Marians KJ. 1996. *J. Biol. Chem.* 271:21391–97
254. Tougu K, Marians KJ. 1996. *J. Biol. Chem.* 271:21398–405
255. Hiasa H, Marians KJ. 1999. *J. Biol. Chem.* 274:27244–48
256. Grompe M, Versalovic J, Koleuth T, Lupski JR. 1991. *J. Bacteriol.* 173:1268–78
257. Versalovic J, Lupski JR. 1997. *Microbiology* 143:585–94
258. Goulian M, Heard CJ. 1989. *J. Biol. Chem.* 264:19407–15
259. Nasheuer H-P, Grosse F. 1988. *J. Biol. Chem.* 263:8981–88
260. Foiani M, Lindner AJ, Hartmann GR, Lucchini G, Plevani P. 1989. *J. Biol. Chem.* 264:2189–94
261. Plevani P, Foiani M, Falconi MM, Pizzagalli A, Santocanale C, et al. 1988. *Biochim. Biophys. Acta* 951:268–73
262. Zakharova OD, Podust VN, Mustaev AA, Anarbaev RO, Lavrik OI. 1995. *Biochimie* 77:699–702
263. Copeland WC. 1997. *Protein Ex. Pur.* 9:1–9
264. Miyazawa H, Izumi M, Tada S, Takada R, Masutani M, et al. 1993. *J. Biol. Chem.* 268:8111–22
265. Copeland WC, Tan XH. 1995. *J. Biol. Chem.* 270:3905–13
266. Kirk BW, Kuchta RD. 1999. *Biochemistry* 38:7727–36
267. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. 1994. *Science* 264:1891–903
268. Vishwanatha JK, Coughlin SA, Wesolowski-Owen M, Baril EF. 1986. *J. Biol. Chem.* 261:6619–28
269. Yotov WV, Hamel H, Rivard GE, Champagne MA, Russo PA, et al. 1999. *Genes Chromosomes Cancer* 26:62–69
270. Tseng BY, Prussak CE, Almazan MT. 1989. *Mol. Cell. Biol.* 9:1940–45
271. Moore AL, Wang TS. 1994. *Cell Growth Differ.* 5:485–94
272. Johnston LH, White JH, Johnson AL, Lucchini G, Plevani P. 1990. *Mol. Gen. Genet.* 221:44–48
273. Foiani M, Liberi G, Lucchini G, Plevani P. 1995. *Mol. Cell. Biol.* 15:883–91
274. Nasheuer HP, Moore A, Wahl AF, Wang TS. 1991. *J. Biol. Chem.* 266:7893–903
275. Voitenleitner C, Rehfuess C, Hilmes M, O'Rear L, Liao PC, et al. 1999. *Mol. Cell Biol.* 19:646–56
276. Wood SH, Collins JM. 1986. *J. Biol. Chem.* 261:7119–22
277. Martelli AM, Capitani S, Neri LM. 1998. *J. Cell. Biochem.* 71:11–20
278. Collins JM, Chu AK. 1987. *Biochemistry* 26:5600–7
279. Mizuno T, Okamoto T, Yokoi M, Izumi M, Kobayashi A, et al. 1996. *J. Cell Sci.* 109:2627–36
280. Mizuno T, Ito N, Yokoi M, Kobayashi A, Tamai K, et al. 1998. *Mol. Cell Biol.* 18:3552–62
281. Takada S, Magira T, Yamamura M. 1989. *Biochem. Biophys. Res. Commun.* 160:711–14
282. Takada S, Nakagawa A, Yamada K, Endo I, Yamamura M. 1994. *Biochem. Mol. Biol. Int.* 34:935–41
283. Takemura M, Kitagawa T, Izuta S, Wasa J, Takai A, et al. 1997. *Oncogene* 15:2483–92
284. Yagura T, Kozu T, Seno T. 1983. *Nucleic Acids Res.* 11:6369–80
285. Orlando P, Geremia R, Frusciante C, Grippo P. 1989. *Cell Differ. Dev.* 27:129–36
286. Goulian M, Heard CJ, Grimm SL. 1990. *J. Biol. Chem.* 265:13221–30
287. Goulian M, Heard CJ. 1990. *J. Biol. Chem.* 265:13231–39

Annual Review of Biochemistry
Volume 70, 2001

CONTENTS