THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 29, Issue of July 22, pp. 26984–26991, 2005
Printed in U.S.A.

Acidic Residues in the Nucleotide-binding Site of the Bacteriophage T7 DNA Primase*

Received for publication, May 2, 2005, and in revised form, May 24, 2005
Published, JBC Papers in Press, May 25, 2005, DOI 10.1074/jbc.M504817200

Seung-Joo Lee and Charles C. Richardson‡

From the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

DNA primases catalyze the synthesis of oligoribonucleotides to initiate lagging strand DNA synthesis during DNA replication. Like other prokaryotic homologs, the primase domain of the gene 4 helicase-primase of bacteriophage T7 contains a zinc motif and a catalytic core. Upon recognition of the sequence, 5′-GTC-3′ by the zinc motif, the catalytic site condenses the cognate nucleotides to produce a primer. The TOPRIM domain in the catalytic site contains several charged residues presumably involved in catalysis. Each of eight acidic residues in this region was replaced with alanine, and the properties of the altered primases were examined. Six of the eight residues (Glu-157, Glu-159, Asp-161, Asp-207, Asp-209, and Asp-237) are essential in that altered gene 4 proteins containing these mutations cannot complement T7 phage lacking gene 4 for T7 growth. These six altered gene 4 proteins can neither synthesize primers *de novo* nor extend an oligoribonucleotide. Despite the inability to catalyze phosphodiester bond formation, the altered proteins recognize the sequence 5′-GTC-3′ in the template and deliver preformed primer to T7 DNA polymerase. The alterations in the TOPRIM domain result in the loss of binding affinity for ATP as measured by surface plasmon resonance assay together with ATP-agarose affinity chromatography.

DNA primase is an essential component of most replisomes in that it catalyzes the synthesis of oligoribonucleotides that serve as primers for the lagging strand DNA polymerase (1). In the bacteriophage T7 replication system, the DNA primase is encoded by gene 4 of the phage (2). The multifunctional gene 4 protein contains both a helicase and a primase domain located in the carboxyl- and amino-terminal halves, respectively, of the polypeptide (3). The gene 4 helicase-primase, together with the gene 5 DNA polymerase, the processivity factor *Escherichia coli* thioredoxin, and the gene 2.5 single-stranded DNA (ssDNA)[1]-binding protein, constitutes the T7 replisome (2). These four proteins mediate coordinated DNA synthesis in which leading and lagging strand DNA syntheses are coupled; both the leading and lagging strand DNA polymerases remain associated with the replisomes, and a replication loop containing a nascent Okazaki fragment is formed (4).

The primase domain of gene 4 protein catalyzes the synthesis of tetraribonucleotides in a template-mediated reaction (5). The T7 DNA primase initiates synthesis from a trinucleotide recognition site, 5′-GTC-3′. The first nucleoside, cytidine, is essential for recognition but is not copied into the product (6). The resulting dinucleotide, pppAC, is then extended to yield the functional tetraribonucleotide primers, pppACCC, pppACAC, and pppACCA (7). A comparative study of prokaryotic DNA primases identified six conserved motifs designated I–VI in Fig. 1*A* (8). A Cys_4 zinc-binding motif (motif I) located at the amino terminus plays a critical role in the recognition of the trinucleotide recognition site (9) and in the transfer of this site to the catalytic core (10), which consists of motifs II–VI. The entire primase domain is covalently attached to the carboxyl-terminal helicase domain via a linker region that is also important in the assembly of the gene 4 protein into the functional hexamer through which the DNA passes (11). The nucleotide sequences encoding the helicase and primase domains have each been cloned, and the resulting helicase and primase fragments have full helicase and primase activity, respectively (12, 13). However, the covalent association of helicase with primase bestows specific properties on the primase; the primase relies upon the helicase to increase its binding to ssDNA and to transport it along DNA in search of primase recognition sites that are extruded behind the helicase in close proximity to the amino-terminal primase domain.

A crystal structure of the T7 DNA primase domain reveals a distinct physical separation of the amino-terminal zinc motif from the catalytic core in that the two are linked via an extended 11-residue polypeptide (10). Such a flexible alignment of these two motifs could modulate primase activity and provide a mechanism by which primers are transferred to the lagging strand DNA polymerase. The flexibility of the linker is illustrated by the observation that the zinc motif of one gene 4 protein subunit within a hexamer can functionally interact with the catalytic core of an adjacent subunit (14). The RNA polymerase core can be further divided into two substructures. One subdomain that encompasses motifs II and III consists of an antiparallel four-stranded β-sheet. The other subdomain (motifs IV–VI) contains a TOPRIM fold also found in topoisomerases (15). A shallow cleft separates the amino-terminal subdomain of the RNA polymerase core from the TOPRIM fold. *E. coli* DnaG primase also has a cleft between these two subdomains (16, 17), and both clefts contain a number of basic residues. It has been proposed that the basic residues within this cleft electrostatically bind the single-strand DNA and that the polymerization of ribonucleotides occurs at a neighboring metal-binding center within the TOPRIM subdomain. We have described previously two lysines (Lys-122 and Lys-128) located in the basic cleft near motif III that are critical for the catalytic activity of T7 DNA primase (18).

* This work was supported by United States Public Health Services Grant GM 54397 and by U. S. Department of Energy Grant DE-FG02-96ER62251. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, MA 02115. Tel.: 617-432-1864; Fax: 617-432-3362; E-mail: ccr@hms.harvard.edu.

[1] The abbreviations used are: ssDNA, single-stranded DNA; DTT, dithiothreitol.

The TOPRIM folds found in prokaryotic DNA primases (motifs IV–VI) of both *E. coli* DnaG primase and T7 DNA primase are quite similar, both structurally and with regard to conserved residues (10, 16). Both have a number of acidic residues that cluster on the TOPRIM side of the cleft formed by the amino-terminal portion of RNA polymerase core (motifs II and III) and the TOPRIM subdomain. In the T7 DNA primase structure, five acidic residues create the acidic patch and bind two metal ions, presumably Mg^{2+} (8) (Fig. 1*B*). One of the metal-binding sites consists of a conserved triad of acidic residues (Glu-157, Asp-207, and Asp-209) similar to that found in DnaG primase as well as in the active site of RNA polymerases (19, 20). The second metal-binding site contains Asp-237 that coordinates another Mg^{2+}. *E. coli* DnaG primase also contains two metal-binding sites along with the same conserved amino acids.

In the present study we have investigated eight conserved acidic residues located in the TOPRIM fold based on sequence homology (8, 15). From the crystal structure of T7 DNA primase, four of these residues appear to be involved in the coordination of metal in the active site (10). In order to determine the role of these residues in primase function, we used *in vitro* mutagenesis to change each residue to alanine. Genetic and biochemical studies of the altered proteins show that six of these residues are indeed involved in the binding of nucleoside triphosphate.

EXPERIMENTAL PROCEDURES

Materials—Oligonucleotides were obtained from the Biopolymer Laboratory at Harvard Medical School. Restriction endonucleases, alkaline phosphatase, and Deep Vent® polymerase were purchased from New England Biolabs. T4 polynucleotide kinase, T4 DNA ligase, radiolabeled nucleotides, and high molecular weight protein markers were purchased from Amersham Biosciences. Agarose and β,γ-methylene-dTTP were from U. S. Biochemical Corp. Polyethyleneimine cellulose TLC plates were from J. T. Baker. NTP-agarose resins and biotin were purchased from Sigma. T7 DNA polymerase (T7 gene 5 protein-*E. coli* thioredoxin complex) and M13mp18 ssDNA were kindly provided by Donald Johnson (Harvard Medical School). Biotinylated NTP analogs were from Enzo Life Sciences Inc (Farmingdale, NY).

Site-directed Mutagenesis, Protein Overproduction, and Purification—Plasmids containing nucleotide changes that result in a single amino acid substitution were constructed as described previously (18). The entire gene 4 coding region was confirmed by DNA sequence analysis. After transformation of the plasmid into *E. coli* HMS 174(DE3), gene 4 proteins were overproduced by induction with 1 mM isopropyl 1-thio-β-D-galactopyranoside for 3 h at 37 °C. Gene 4 proteins were purified as described previously (18).

Phage Complementation Assay—*E. coli* DH 5α containing a plasmid that expresses T7 gene 4 under a T7 promoter (pET24-gp4) was grown to an A_{600} of 1. Serially diluted T7 phage stocks were mixed with an aliquot of the *E. coli* culture in 0.7% soft agar and poured onto LB plates. After an overnight incubation at 37 °C, the number of plaques that appeared on the plate were determined.

Biochemical Assays of Gene 4 Protein—Most of the assays used in this study have been described previously (11, 18). All assays (template-directed oligoribonucleotide synthesis, template-directed oligoribonucleotide extension, template-independent diribonucleotide synthesis, and RNA-primed DNA synthesis by DNA polymerase) used a reaction buffer containing 40 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, and 50 mM potassium glutamate plus the additional components described in each assay. All reactions were carried out at 37 °C for the indicated time.

Template-directed Oligoribonucleotide Synthesis—The ability of gene 4 protein to catalyze *de novo* synthesis of oligoribonucleotides was determined by measuring the incorporation of [α-^{32}P]CMP into oligoribonucleotides using a synthetic DNA template containing a primase recognition site. The reaction (10 μl) included 10 μM template DNA 5′-GGGTCAAA-3′, 0.1 mM each of ATP and [α-^{32}P]CTP (0.1 μCi), and the indicated amount of gene 4 protein. After incubation at 37 °C for 20 min, the reaction was terminated by the addition of 3 μl of sequencing dye and loaded onto a 25% denaturing polyacrylamide sequencing gel containing 3 M urea. Electrophoresis was carried out at 1800 V for 3 h, and the gel was dried for autoradiography. Radioactive oligoribonucleotide products were analyzed using a Fuji BAS 1000 Bioimaging analyzer.

Template-directed Oligoribonucleotide Extension—T7 gene 4 primase can also catalyze the extension of the diribonucleotide AC to the tri- and tetraribonucleotide provided it is annealed at a primase recognition site 5′-GTC-3′ (21). The reaction was as described above for a template-directed oligoribonucleotide synthesis assay except that various amounts (0.1, 0.2, 0.4, and 0.8 mM) of diribonucleotide AC were substituted for ATP. Gene 4 protein was present at 160 nM.

Template-independent Diribonucleotide Synthesis—Gene 4 primase can catalyze the synthesis of random diribonucleotide, albeit at a low rate, in the absence of DNA (22). Gene 4 protein (300 nM) was incubated with 0.1 mM [α-^{32}P]CTP (0.1 μCi) in the absence of template for 1 h at 37 °C. Reaction buffer contained Mn^{2+} instead of Mg^{2+}. Reaction products were dephosphorylated for 20 min at 37 °C and separated on a 25% denaturing polyacrylamide sequencing gel.

RNA-primed DNA Synthesis by T7 DNA Polymerase—The oligoribonucleotide synthesized by gene 4 protein at primase recognition sites on ssDNA can be transferred to DNA polymerase by the gene 4 protein to serve as primers for DNA synthesis. Such RNA-primed DNA synthesis was measured in an assay containing 20 nM M13 ssDNA, 0.3 mM each of dGTP, dCTP, dATP, and [α-^{32}P]dTTP (0.1 μCi), 100 nM gene 4 protein, 20 nM T7 DNA polymerase, and either 0.1 mM each of ATP and CTP or 20 μM preformed tetraribonucleotide ACCA. The reaction was incubated for 10 min at 37 °C and terminated by the addition of EDTA to a final concentration of 20 mM. The reaction products were spotted onto a DE81 membrane (Whatman), and the membrane was washed three times with 10 ml of 0.3 M ammonium formate, pH 8.0, to remove unincorporated radioactive nucleotide. The amount of radioactively synthesized DNA was determined by measuring the radioactive products retained on the membrane.

Assays for Activities of Helicase Domain—The oligomerization of gene 4 protein, the binding of gene 4 protein to ssDNA, ssDNA-dependent dTTP hydrolysis, and the unwinding of duplex DNA were measured as described previously (11).

Determination of Binding Ability of Gene 4 Protein to Nucleotide Resin—A small volume (0.4 ml) of pre-swollen nucleotide-coupled agarose resin was packed into a column and washed with buffer A (20 mM potassium phosphate, pH 6.8, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol). A protein pool containing gene 4 protein was prepared from phosphocellulose column chromatography as described previously (18), and $MgCl_2$ was added to a final concentration of 10 mM. After loading of the protein pool by gravity, the column was washed with 2 ml of buffer A containing 10 mM $MgCl_2$. Gene 4 protein was eluted using 2 ml of buffer E (20 mM potassium phosphate, pH 6.8, 20 mM EDTA, 1 mM DTT, and 10% glycerol). Each aliquot from the initial protein pool, flow-through, wash, and elution fraction was analyzed on 10% SDS-PAGE and visualized by staining with Coomassie Blue.

Surface Plasmon Resonance Analysis—The binding of proteins to NTPs was examined using surface plasmon resonance on the Biacore instrument. Bio-17-ATP and Bio-11-CTP are analogs of ATP and CTP, respectively, which have a biotin attached to the base of the NTP through the indicated number of atoms. The biotinylated NTPs were immobilized on the surface of Sensor Chip SA through a streptavidin-biotin interaction by manually injecting 10 μl of 10 μM in a buffer (10 mM Tris-HCl, pH 7.5, 0.5 M NaCl, and 1 mM EDTA) at the flow rate of 20 μl/min. Proteins in a flow buffer (10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 50 mM potassium glutamate, and 1% glycerol) were injected for 1.25 min at the flow rate of 20 μl/min, and the binding signal was collected. After binding of protein, the chip was regenerated by repeatedly injecting 1 M NaCl for 1 min at a flow rate of 50 μl/min until the signal was restored to the level before injection of proteins. A control lane contained biotin immobilized on the streptavidin chip. The signal obtained when the proteins were flowed over the control lane was subtracted from the signal measured when proteins were flowed over the biotinylated NTP lane.

RESULTS

A sequence alignment of bacterial and bacteriophage DNA primases revealed six conserved motifs in the DnaG family of primases (Fig. 1*A*) (8). Motifs IV and V have homologies to motifs found in proteins that bind nucleotides such as ATPases, DNA, and RNA polymerases. An iterative profile search combined with structural modeling proposed that motifs IV–VI form a unique fold, the TOPRIM fold, also found in topoisomerases (15). Within these motifs, a number of acidic residues were considered candidates for metal-mediated binding of NTP and for catalysis (23). The crystal structures of T7 DNA primase and *E. coli* DnaG primase confirmed that motifs IV and V are

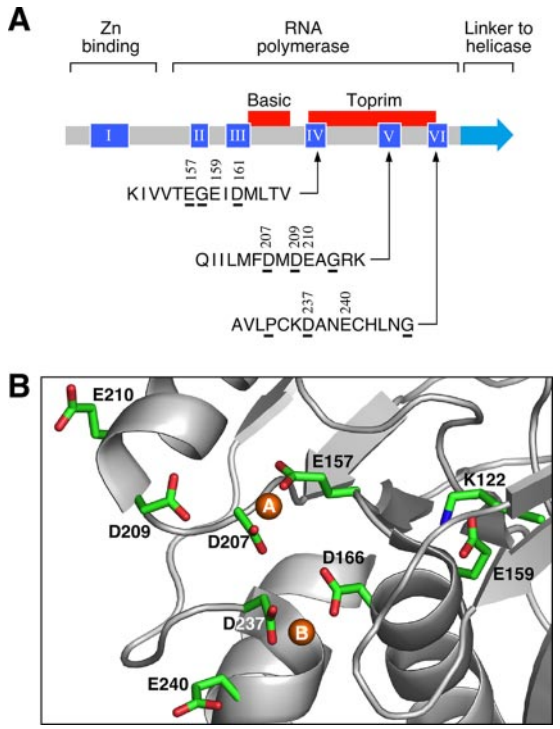

FIG. 1. **Primase domain of T7 gene 4 protein.** *A*, the schematic presentation shows the primase domain located in the amino-terminal half of the gene 4 protein. The *boxes*, labeled with r*oman numerals*, represent conserved motifs in the primase superfamily. The TOPRIM fold is composed of motifs IV–VI and the RNA polymerase-basic region is located between motifs III and IV. Based on sequence homology among prokaryotic DNA primases, conserved amino acid residues (*underlined*) were identified. The eight acidic residues investigated in this study are numbered. *B*, crystal structure of the catalytic subdomain is shown as a ribbon model (10). The eight acidic residues as well as Lys-122 are labeled, and their side chains are shown. Two magnesium ions are depicted by *spheres* with their electron density and labeled as *A* and *B*.

indeed involved in Mg^{2+}-mediated NTP binding, and several acidic residues were found to coordinate metal or to be in a position to do so (10, 16). In the present study, we have substituted alanine for each of the eight acidic residues in motifs IV–VI (Fig. 1*A*) and then examined the biochemical properties of each of the purified proteins.

Complementation of Phage Growth by Recombinant Gene 4

The ability of each of the genetically altered gene 4 proteins to support T7 phage growth was examined by using a complementation assay. In this assay, growth of phage T7Δ4 lacking gene 4 is dependent on gene 4 protein exogenously expressed from a plasmid harbored in the T7-infected *E. coli*. Defects in the function of gene 4 protein do not allow the phage to propagate, and no plaques are formed. Substitution at all 8 positions except for Glu-210 and Glu-240 resulted in loss of the ability of the protein to support growth of phage T7Δ4 (Table I). These results suggest that the six residues, Glu-157, Glu-159, Asp-161, Asp-207, Asp-209, and Asp-237, are each essential for gene 4 protein function. When the cells expressing each of the altered gene 4 proteins were infected with wild-type T7 phage, no significant difference between wild-type and altered gene 4 proteins was found. Thus, none of the altered proteins have a dominant negative effect on the function of the wild-type protein expressed by the phage.

Overproduction and Purification of Altered Gene 4 Proteins

In order to examine the biochemical properties of the altered gene 4 proteins, the proteins were purified to apparent homogeneity. The altered gene 4 proteins are designated gp4-E157A, E159A, D161A, D207A, D209A, E210A, D237A, and E240A. Although gp4-E210A and gp4-E240A could replace wild-type gene 4 protein *in vivo*, both of these altered proteins were also purified and characterized.

TABLE I
Ability of altered gene 4 proteins to complement growth of T7 phage

Gene 4 proteins containing the indicated amino acid substitution were expressed in *E. coli* DH5α. After infection with either T7Δ4 or wild-type (WT) T7 phage, the number of plaques were counted and normalized to the value obtained with wild-type gene 4 protein. Data were obtained from at least duplicated experiments.

Alteration in gene 4 protein	T7 phage	
	Δ4	WT[a]
WT	1.0[b]	1.0
E157A	$<10^{-5}$	0.67
E159A	$<10^{-5}$	1.7
D161A	$<10^{-5}$	1.7
D207A	$<10^{-5}$	1.0
D209A	$<10^{-5}$	3.3
E210A	2.9	2.7
D237A	$<10^{-5}$	0.83
E240A	2.0	2.7

[a] WT, wild type.
[b] Relative efficiency of plating.

Initially we attempted to purify all of the gene 4 proteins by a standard purification procedure that utilizes both adsorption to and elution from a phosphocellulose and an ATP-agarose resin (18). Whereas gp4-E210A and gp4-E240A demonstrated similar affinity for both resins, the other altered proteins displayed reduced affinity for both resins. For example, only 50% of gp4-D209A and gp4-D237A adsorbed to the phosphocellulose resin and less than 10% of gp4-E159A and gp4-D161A adsorbed. Gp4-E157A and gp4-D207A behaved similar to the wild-type protein on phosphocellulose column chromatography. More interestingly, all of these altered proteins except for gp4-E210A and gp4-E240A showed greatly reduced affinity for the ATP-agarose resin. As shown in Fig. 2, the majority of gp4-E157A and gp4-D207A applied to the column (*lane 1*) did not bind to the resin (*lane 2*), and consequently only a small percentage of the protein applied to the resin could be recovered (*lane 3*). This behavior is precisely opposite that observed with wild-type gene 4 protein irrespective of whether the proteins are monitored by staining with Coomassie Blue (Fig. 2*A*) or by a Western blot (Fig. 2*B*). These results suggest that the altered proteins are defective in their ability to bind nucleoside triphosphate, a property that we address below.

In view of the altered behavior of these proteins on phosphocellulose and ATP-agarose, it was necessary to design an alternative purification procedure. Basically, chromatography on ATP-agarose resins was replaced by chromatography on DEAE-Sepharose, a purification step that has been used previously (18). Consistent with the ability of gp4-E210A and gp4-E240A to support T7Δ4 growth, these two proteins behaved identically to wild-type gene 4 protein on both phosphocellulose and ATP-agarose, and hence these two proteins were purified following the normal procedure for wild-type protein. The purity of all of the gene 4 proteins, regardless of the purification procedure, was greater than 90% as measured by SDS-PAGE and staining with Coomassie Blue.

Biochemical Assays

Catalytic Activity of Primase—The single amino acid substitution introduced into the TOPRIM region of primase domain could potentially affect its catalytic function. In the template-directed *de novo* synthesis of primers, a zinc motif in T7 primase first recognizes a specific sequence (5′-GTC-3′) on the ssDNA template. Subsequently, the catalytic RNA polymerase

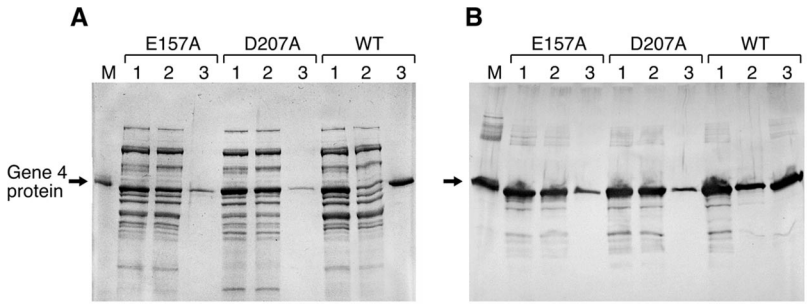

FIG. 2. **Binding of gene 4 proteins to ATP-agarose.** Two representatives of the altered gene 4 proteins, gp4-E157A and gp4-D207A, were compared with wild-type (*WT*) protein for their ability to bind to ATP-agarose. The phosphocellulose fraction containing gene 4 protein (*lane 1*) was applied to an ATP-agarose affinity column, and the flow-through fraction (*lane 2*) and elution fraction (*lane 3*) were analyzed on a 10% SDS-polyacrylamide gel with a marker of gene 4 protein (*M*). Elution from the ATP-agarose column was accomplished by the addition of a buffer containing 20 mM EDTA. Proteins on gels were visualized by either Coomassie staining (*A*) or Western blot by using polyclonal antibody against gene 4 protein (*B*).

domain polymerizes nucleotide cognate to the template. T7 primase is also able to extend a preformed ribonucleotide annealed to the primase recognition sequence in a template-dependent manner. In addition to template-directed synthesis, the catalytic center can also catalyze the synthesis of random diribonucleotides in the absence of template. We have examined these three catalytic activities of the altered gene 4 proteins and compared them to the activities of the wild-type gene 4 protein.

In the first assay shown in Fig. 3*A*, a short ssDNA template (5′-GGGTCAAA-3′) containing the primase recognition site was incubated with various amounts of gene 4 protein in the presence of ATP and [α-^{32}P]CTP. Reaction products were analyzed on a sequencing gel, and the amount of oligoribonucleotides synthesized was determined by measuring radioactivity incorporated into the major product, pppACCC. The results showed that both gp4-E210A and gp4-E240A catalyzed synthesis of oligoribonucleotides as efficiently as the wild-type gene 4 protein, whereas none of the other altered proteins were able to catalyze the synthesis of oligoribonucleotides.

In the second assay (Fig. 3*B*), the ability of the proteins to extend a preformed diribonucleotide was measured. In this assay, ATP was replaced with the diribonucleotide, 5′-AC-3′, and the incorporation of [α-^{32}P] CTP was measured as in the previous assay. None of the altered proteins except for gp4-E210A and gp4-E240A could incorporate CTP.

In the third assay, the gene 4 proteins were incubated with [α-^{32}P]CTP in the absence of DNA template (Fig. 3*C*). Only gp4-E210A, gp4-E240A and wild-type protein were capable of synthesizing the diribonucleotide 5′-CC-3′.

Taken together, all three assays show that residues Glu-157, Glu-159, Asp-161, Asp-207, Asp-209, and Asp-237 are critical to the catalytic function of T7 primase, whereas residues Glu-210 and Glu-240 are not. Acidic residues in nucleolytic enzymes are known to be involved in Mg^{2+}-mediated binding of NTP. By assuming that the loss of catalytic activity by the amino acid substitutions described here is due to disruption of the metal ion-mediated event, we attempted to recover the enzymatic activity of the altered proteins by supplying excess Mg^{2+}. The addition of 10-fold more Mg^{2+} than that present in the standard reaction assay did not result in any measurable *de novo* synthesis of oligoribonucleotides (data not shown).

Reconstitution of Primase Activity—For template-directed primer synthesis, both the zinc motif and the catalytic center of gene 4 protein are required. We have shown previously that a defect in either of these motifs can be complemented by a functioning motif in an adjacent subunit within a hexamer of the gene 4 protein (14). For example, a hetero-complex between a gene 4 protein lacking the zinc motif (56-kDa gene 4 protein) and a gene 4 protein with an inactive catalytic site (gp4-K122A)

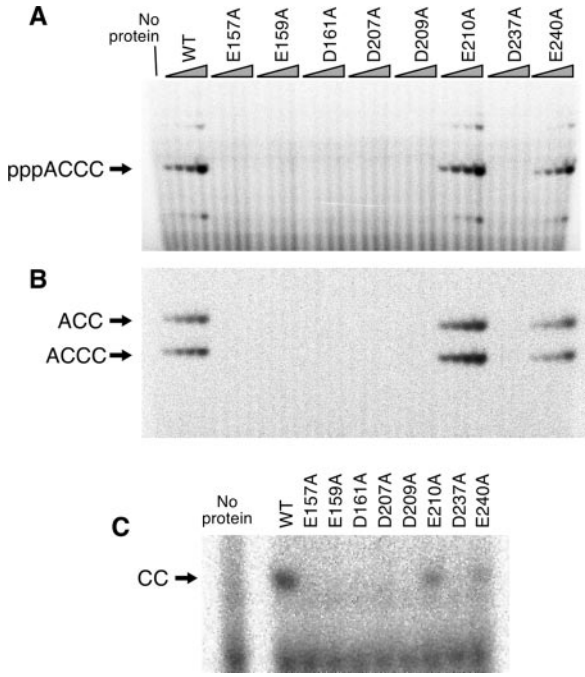

FIG. 3. **Catalytic activity of gene 4 primase.** *A*, template-directed oligoribonucleotide synthesis. Increasing amounts (40, 80, 160, and 320 nM) of wild-type (*WT*) and altered gene 4 proteins were incubated with ATP and [α-^{32}P]CTP in the presence of template DNA 5′-GGGTCAAA-3′ for 20 min at 37 °C. The reaction products were analyzed on 25% denaturing polyacrylamide gel. The major product, 5′-pppACCC-3′, is indicated to the *left* of the gel. *B*, template-directed oligoribonucleotide extension. Reactions similar to the one described in *A* were carried out except that the diribonucleotide 5′-AC-3′ replaced ATP. The concentration of 5′-AC-3′ was varied from 0.1, 0.2, and 0.4 to 0.8 mM. The reaction products, 5′-ACC-3′ and 5′-ACCC-3′, are shown to the *left* of the gel. *C*, template-independent diribonucleotide synthesis. Gene 4 proteins were incubated with [α-^{32}P]CTP in the absence of template for 1 h at 37 °C, and the resulting products were dephosphorylated and separated on 25% denaturing polyacrylamide. The reaction product, 5′-CC-3′, is shown to the *left* of the gel.

catalyzes template-directed oligoribonucleotide synthesis, whereas neither subunit alone can (14). In order to examine such a reconstitution of primase activity, we prepared various combinations of gene 4 protein mixtures with the 56-kDa gene 4 protein, gp4-K122A, and the six altered primase-inactive gene 4 proteins identified from this study. After mixing the two proteins, the ability of the gene 4 protein mixture to synthesize template-directed oligoribonucleotides was determined by measuring the incorporation of [α-^{32}P]CTP into oligoribonucleotides (Fig. 4). Similar to gp4-K122A, all the primase-defective proteins could catalyze the synthesis of oligoribonucleotides in

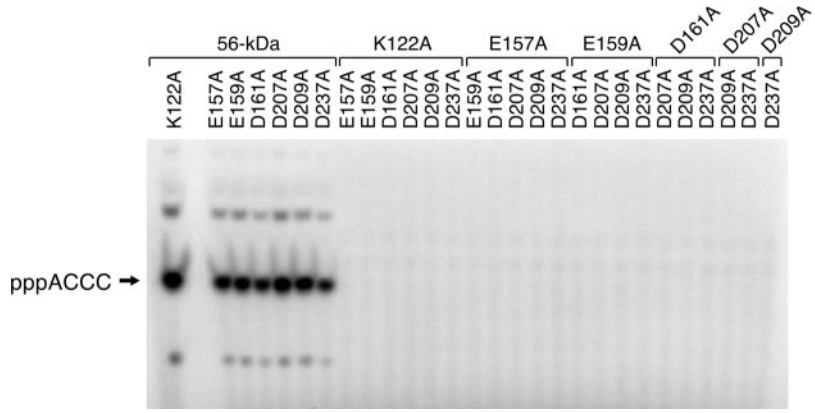

FIG. 4. **Reconstitution of primase activity by gene 4 proteins.** The ability of primase-inactive gene 4 proteins to reconstitute one another within a hexamer was measured. Pairs of the indicated gene 4 proteins were mixed and preincubated at room temperature for 30 min prior to the addition of ATP, [α-^{32}P]CTP, and dTTP in the presence of a template 5′-GGGTCA$_{10}$-3′. After incubation for 20 min at 37 °C, the reaction products were analyzed on 25% denaturing polyacrylamide gel. The major product 5′-pppACCC-3′ is indicated to the *left* of the gel. The 56-kDa gene 4 protein lacks the zinc motif of the primase domain.

the presence of the 56-kDa protein. However, all combinations of inactive primases did not show any activity for oligoribonucleotides synthesis. The ability to reconstitute primase activity with the 56-kDa gene 4 protein demonstrates that the zinc motif of the altered proteins can function properly.

RNA-primed DNA Synthesis by T7 DNA Polymerase—The ultimate role of DNA primase during DNA replication is to supply oligoribonucleotides for use as a primer by DNA polymerase to initiate lagging strand synthesis. It has been shown previously that a gene 4 protein defective in primer synthesis due to an inability to catalyze phosphodiester bonds can nonetheless successfully deliver a preformed tetraribonucleotide annealed at a primase recognition site to DNA polymerase to initiate DNA synthesis (18). In order to examine this functional aspect of the altered proteins, RNA-primed DNA synthesis by DNA polymerase was carried out by using two different assays.

In the first assay, ATP and CTP were provided as precursors for *de novo* synthesis of oligoribonucleotides. Consistent with the previous results, under these conditions wild-type, gp4-E210A, and gp4-E240A were enabled by T7 DNA polymerase to initiate DNA synthesis on ssDNA (Fig. 5*A*). None of the other altered proteins could enhance T7 DNA polymerase activity because they are unable to synthesize oligoribonucleotides. However, when the preformed oligoribonucleotide 5′-ACCA-3′ replaced ATP and CTP, all the altered gene 4 proteins could deliver the primer to DNA polymerase to initiate DNA synthesis (Fig. 5*B*). Thus, despite the inability of the altered proteins to synthesize primers, the defective gene 4 proteins can recognize the preformed oligoribonucleotides and their recognition sequence and then deliver the primer to T7 DNA polymerase.

Affinity for NTP—Although the defect in oligoribonucleotides synthesis created by the replacement of the six acidic residues in the TOPRIM fold could arise by any one of several mechanisms, several lines of evidence suggest that the alterations affect nucleotide binding. As discussed earlier, all of the altered proteins defective in primer synthesis had very low affinity for the ATP-agarose resin normally used for purification of the gene 4 protein, necessitating the replacement of this resin by another. Furthermore, acidic residues in other proteins having TOPRIM folds are thought to bind ATP, and the crystal structures of T7 gene 4 primase domain and *E. coli* DnaG primase actually show interactions of these residues with a divalent cation, presumably Mg^{2+} (10, 16). We have examined this likely possibility by comparing the binding of wild-type gene 4 proteins and the genetically altered proteins to nucleotides.

In preliminary experiments, we examined the binding of wild-type gene 4 protein to nucleotides bound to agarose resins, and the results are shown in Table II. Gene 4 protein binds strongly to ATP linked to agarose through either C-8 or N-6 of the base. However, when ATP is linked via its 2′- and 3′-ribose hydroxyls to the resin, no binding of the protein could be detected. These results are in agreement with those reported earlier for gene 4 protein (24). Gene 4 protein showed no affinity for either AMP or ADP (Table II).

Based on the data just presented for wild-type gene 4 protein binding to ATP, we examined the ability of each of the eight altered gene 4 proteins to bind to an ATP affinity column where the ATP is linked to the resin via C-8 of the base. In contrast to wild-type gene 4 protein, none of the altered gene 4 proteins defective in primase activity could bind to the column (Table II and Fig. 2). However, gp4-E210A and gp4-E240A bound as well as did wild-type gene 4 protein.

We have also examined the binding of wild-type and the altered gene 4 proteins to ATP by using surface plasmon resonance. In these experiments, biotinylated ATP or CTP was immobilized on the surface of a Biacore chip via a biotin-streptavidin interaction. Each of the gene 4 proteins were then flowed over the NTP linked to the chip, and the response units (RU) were measured. It was observed that gene 4 protein nonspecifically bound to streptavidin on the surface of the chip. To eliminate the contribution of this nonspecific binding, proteins were flowed over a chip where only biotin was immobilized to the streptavidin chip instead of the biotinylated NTP. Specific binding of gene 4 protein to immobilized NTP was expressed by the difference between a binding signal from an experimental lane and the one from the control lane. Although wild-type gene 4 protein displayed a rapid association with the bound ATP, neither gp4-E157A, gp4-D207A, nor gp4-D237A, all defective in primase activity, displayed any measurable binding (Fig. 6). We also examined gp4-K122A, which we had shown previously to be defective in catalyzing phosphodiester bond formation, and we found that it also did not bind to ATP (Fig. 6). When CTP, another preferred nucleotide bound by T7 primase, was immobilized to the chip, no binding signal to the wild-type protein was detected (data not shown).

Helicase Domain Activity—Although primase and helicase domains are separated in T7 gene 4 protein (Fig. 1*A*), the activity of one domain is often affected by the other domain. For instance, the efficiency of oligoribonucleotides synthesis on the long template is dependent on not only the primase domain but also the helicase domain (11). The enhancement of primase activity results from the ability of the helicase domain to transport the primase domain to recognition site. In order to determine whether the alterations in TOPRIM also affect the helicase domain, we examined the helicase activity of the altered proteins. As summarized in Table III, all activities of the altered proteins (oligomerization, DNA binding, ssDNA-dependent dTTP hydrolysis, and DNA unwinding) were similar to those of the wild-type protein. Therefore, we conclude that the alterations are confined within the primase domain.

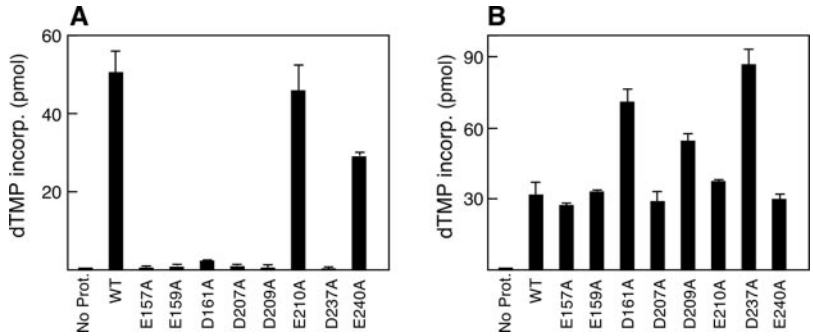

FIG. 5. **RNA-primed DNA synthesis by T7 DNA polymerase.** DNA synthesis on an M13 ssDNA template catalyzed by T7 DNA polymerase requires primers annealed to the template. The primer was either synthesized *de novo* from ATP and CTP by gene 4 protein (*A*) or supplied as 5′-ACCA-3′ (*B*). Reaction mixtures contained dGTP, dCTP, dATP, [α-^{32}P]dTTP, M13 ssDNA, T7 DNA polymerase, and 100 nM of the indicated gene 4 protein. After incubation for 5 min at 37 °C, the products of the reaction were spotted on DE81 membrane, and unincorporated dNTP was removed by extensively washing with ammonium formate. The amount of DNA synthesis was determined by measuring the amount of radioactive dTMP incorporated into DNA. *WT,* wild type; *No prot*, no protein.

TABLE II
Ability of gene 4 protein to bind nucleotide-agarose resins

A variety of nucleotides attached to agarose resin was examined for binding affinity of gene 4 proteins using chromatography on a small column. The binding (+) or absence of binding (−) to the gene 4 protein is indicated.

Nucleotide	Position of attachment	Length of linker	Binding of proteins		
			WT	E210A, E240A	E157A, E159A, D161A, D207A, D209A, D237A
AMP	N-6 of base	8	−		
ADP	C-8 of base	8	−		
ATP	C-8 of base	9	+	+	−
ATP	N-6 of base	11	+		
ATP	Hydroxyls of ribose	22	−		

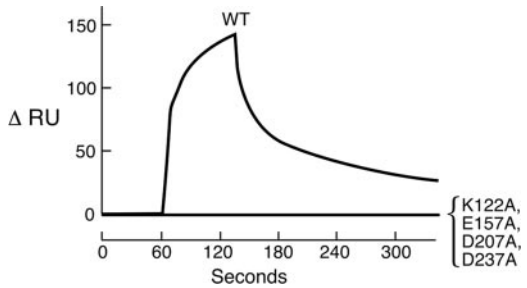

FIG. 6. **NTP binding affinity detected by surface plasmon resonance.** ATP was immobilized on the surface of chip via biotin-streptavidin interaction, and gene 4 protein was passed over the chip in the presence of Mg^{2+}. Binding affinity of protein to immobilized ATP is represented by the difference in binding signal between immobilized biotinylated ATP (experimental lane) and immobilized biotin (control lane) as shown on the *y* axis. The wild-type (*WT*) gene 4 protein and four primase-inactive gene 4 proteins (0.3 μM each) were injected for 1.25 min at the flow rate of 20 μl/min. *RU*, response units.

DISCUSSION

DNA primases are an essential component of the replisomes found in most replication systems. They have a dual role in that they must catalyze the template-directed synthesis of oligoribonucleotides, usually of a well defined length, and then interact with their cognate DNA polymerase to deliver the oligonucleotide for use as a primer to initiate DNA synthesis. In the first step of primer synthesis, the prokaryotic DNA primases recognize a trinucleotide sequence, on single-stranded DNA, catalyze the synthesis of a dinucleotide, and then extend it to a length that can be used as a primer by the respective DNA polymerase (1). Thus, their synthesis of RNA and recognition of a DNA sequence from which synthesis is initiated are shared with RNA polymerases that recognize, albeit much longer DNA sequences, promoters and catalyze the synthesis of RNA. The fact that the recognition sequence for DNA primases resides on single-stranded DNA serves as another distinguishing characteristic of the primases. Unlike classical RNA polymerases, the primases have low processivity of nucleotide polymerization and relatively low fidelity (1).

In view of their ability to catalyze template-directed synthesis of oligoribonucleotides, it is not surprising that the DNA primases and polymerases share catalytic mechanisms and structural features. For example, two aspartic acid residues in a motif, D*X*D, are often conserved not only throughout many polymerases but also in DNA primases (1, 25). Those acidic residues form an acidic patch in the active sites of T7 DNA primase and *E. coli* DnaG primases and are thought to bind two divalent metal ions (10, 16, 26), reminiscent of the two metal ion-mediated mechanism of nucleotide condensation proposed for all polymerases by Steitz *et al.* (27). However, despite these similarities, the primases do differ significantly from the polymerases. The RNA polymerase domain of both T7 primase and DnaG primase contains a TOPRIM fold found in a variety of DNA topoisomerases but not polymerases (15). More interestingly, the DNA primases have a zinc-binding domain that, in the case of phage T7 primase, consists of a four-stranded antiparallel β-sheet flanked by a carboxyl-terminal α-helix (10). The zinc in T7 primase is bound by two pairs of cysteines (28) and, like the zinc-binding domain of DnaG primase, is located at the amino terminus of the protein, tethered to the catalytic core of the primase domain by a highly flexible linker (10). This zinc-binding domain is involved in nucleotide sequence recognition as well as in the transfer of the primer to the polymerase (9, 10).

The conserved nature of the acidic residues in the TOPRIM fold and the involvement of several of these in the coordination of two metal ions in the crystal structures of *E. coli* DnaG primase and T7 DNA polymerase (10, 16) make a strong case that they play a critical role in nucleotide binding and/or catalysis. Indeed, a number of biochemical studies have implicated motifs IV, V, and VI in the coordination of nucleoside triphosphate substrates and divalent metal cations necessary for catalysis (29–31). In the case of T7 DNA primase, one metal ion (*A* in Fig. 1*B*) is coordinated with Glu-157 and Asp-207 and possibly with the side chain of Asp-209, although the latter in the crystal structure appears too distant to make direct contact (10). Nonetheless, these three acidic residues are conserved in all proteins containing the TOPRIM fold and constitute the acidic triad found in the catalytic site of RNA polymerase (19, 20). The second metal (*B* in Fig. 1*B*) site clearly has a contact with Asp-237, the equivalent residue in DnaG that was detected by Fe^{2+} affinity (31). Asp-161 is also a member of the

TABLE III
Activities of the helicase domain of gene 4 proteins
Results were obtained from at least duplicated assays as described under "Experimental Procedures."

	Wild type	E157A	E159A	D161A	D207A	D209A	E210A	D237A	E240A
Oligomerization	+	+	+	+	+	+	+	+	+
DNA binding (nM)[a]	25 ± 2	30 ± 3	36 ± 5	39 ± 3	32 ± 4	35 ± 5	25 ± 4	29 ± 4	20 ± 3
dTTP hydrolysis (pmol)[b]	0.9 ± 0.3	0.8 ± 0.3	1.0 ± 0.2	1.7 ± 0.3	0.7 ± 0.2	1.4 ± 0.1	1.0 ± 0.1	2.0 ± 0.1	1.0 ± 0.2
DNA unwinding (fmol)[c]	2.3 ± 0.3	2.9 ± 0.3	2.5 ± 0.3	2.4 ± 0.3	2.2 ± 0.2	2.3 ± 0.04	2.0 ± 0.1	2.5 ± 0.3	2.5 ± 0.5

[a] Dissociation constant for hexameric gene 4 protein to DNA.
[b] Amount of dTTP hydrolyzed by 1 nM gene 4 protein (monomer) from initial 500 pmol in the presence of 8 nM M13 ssDNA.
[c] Amount of DNA substrate unwound by 1 nM gene 4 protein (monomer).

cluster of five invariant acidic residues on the TOPRIM side of the cleft that constitutes the RNA polymerase domain but is somewhat remote for direct involvement with either of the two metals in the crystal structure.

In the present study, we have determined the importance of eight of the acidic residues located in motifs IV–VI that compose the TOPRIM domain. Each of the eight acidic residues was individually substituted with alanine, and the altered protein was examined for its ability to support the growth of T7 phage lacking gene 4. Each of the altered proteins was then purified, and its biochemical properties were examined. Of the eight acidic residues, only two, Glu-210 and Glu-240, do not appear to be important for primase activity *in vivo* or *in vitro*. Although Glu-210 and Glu-240 are proximate to the critical acidic residues, Asp-209 and Asp-237, both gp4-E210A and gp-E240A support T7Δ4 growth, and no biochemical defects were detected in the battery of primase assays. This nonessential role is supported by the crystal structure of the primase domain; both Glu-210 and Glu-240 are exposed to solvent, and the side chain of Glu-210 is oriented opposite to the metal (10) (Fig. 1*B*). The side chain of Glu-240 does face the metal but at a distance of 6 Å.

We were reassured to find that four of the six acidic residues are essential for primase activity and for phage growth, either from inspection of the crystal structure of T7 primase or from homologous residues in other primases, in major roles in the catalytic center. The two residues (Glu-157 and Asp-207) observed to coordinate one of the two metal ions (*A* in Fig. 1*B*), in the structure were both found to be essential as well as Asp-209 whose side chain is in a position to complete the acidic triad. Asp-237, which coordinates the second metal ion (*B* in Fig. 1*B*), is likewise essential. Of the other two essential acidic residues, Glu-159 and Asp-161, the latter is a part of the cluster that forms the acidic patch that coordinates the second metal ion. Glu-159, on the other hand, is notably distant from the metal ions to be involved in coordination, and most interestingly, it is not one of the highly conserved residues among DNA primases. More likely, it is a critical determinant of the binding pocket for nucleotide via its interactions with other residues.

We find that all of the amino acid changes in the TOPRIM fold which affect primase activity also drastically reduce the affinity of gene 4 protein to bind to an ATP-affinity column. We have routinely used an ATP-affinity column to purify T7 gene 4 protein because the wild-type protein binds with high affinity in the presence of Mg^{2+}, necessitating elution with EDTA (18). Our initial attempts to purify the altered proteins described in this study revealed the defect in ATP binding, requiring the substitution of a DEAE-Sepharose chromatography step. The inability of the primase-defective gene 4 proteins to bind to ATP was confirmed using surface plasmon resonance to demonstrate their lack of binding to ATP immobilized to the surface of a Biacore sensor chip. Clearly, the essential acidic residues described in this study affect the ability of ATP to bind to the catalytic site of the primase, a defect either mediated through the inability to bind the essential metal ion or via the absence of side chains that interact directly with the nucleotide.

A surprising result has been the finding that the gene 4 protein binds to the ATP-affinity resin via the ATP-binding site of the primase domain. We had assumed previously that the binding of gene 4 protein to this resin was occurring through binding of the ATP at the nucleoside triphosphate site of the helicase domain, which binds and hydrolyzes both dTTP and ATP (32). The binding of ATP to the primase domain is relatively weak with a K_m of 0.32 mM as measured kinetically (33). Precisely why the nucleotide-binding site of the primase has such a high affinity for the ATP-agarose resin remains unclear.

As is the case for most polymerases, T7 DNA primase contains essential acidic residues on the floor of the active site. In the general mechanism of two metal ion-mediated catalysis by polymerases, one metal ion promotes the deprotonation of the 3′-hydroxyl of the primer strand, and the other facilitates the formation of the pentacovalent transition state of the α-phosphate of incoming (d)NTP and the departure of the PP_i-leaving group (34). Three conserved acidic residues (Glu-157, Asp-207, and Asp-209) clustered in one magnesium ion in T7 primase would play such a role as aspartates in the catalytic triad of an archaeal primase interacting with phosphates of UTP (35). From their proximity from the other metal ion, Asp-161 and Asp-237 are also considered to coordinate with the other metal ion. Therefore, the five metal ion-coordinated resides appear to interact with the first nucleotide ATP in oligoribonucleotide synthesis, as determined by ATP-affinity chromatography and surface plasmon resonance assay. However, Glu-159, from this work, and Lys-122, from the earlier study, are distant from any metal ion in the crystal structure, yet both are critical to binding to ATP. More interestingly, those residues are noticeably close to one another in the structure, suggesting they have a common basis for the loss of ATP binding upon replacement with alanine (Fig. 1*B*).

Based on the accumulated evidence derived from structural, biochemical, and mutagenesis studies, a picture of the catalytic site of the DNA primase is slowly emerging. This study demonstrates that ATP strongly binds to the catalytic site of primase in the absence of template, suggesting that the ATP-binding site might be pre-filled *in vivo*. We showed previously that T7 primase also selectively binds CTP in the absence of DNA (9). The preferential binding of the cognate NTPs to the primase recognition site (5′-GTC-3′) prior to DNA binding apparently provides advantage for the template-dependent oligoribonucleotide synthesis. In one scenario, the binding of the NTPs in the catalytic site induces significant conformational change, which facilitates binding of template DNA. Alternatively, the presence of NTPs bound to T7 DNA primase could provide via Watson-Crick base pairing a mechanism for binding to the trinucleotide recognition sites 5′-GTC-3′.

Acknowledgment—We thank Samir Hamdan for useful discussions and help during surface plasmon resonance analysis and preparation of Fig. 1*B*.

REFERENCES

1. Frick, D. N., and Richardson, C. C. (2001) *Annu. Rev. Biochem.* **70,** 39–80
2. Richardson, C. C. (1983) *Cell* **33,** 315–317
3. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166,** 477–535
4. Lee, J., Chastain, P. D., II, Griffith, J. D., and Richardson, C. C. (2002) *J. Mol. Biol.* **316,** 19–34
5. Tabor, S., and Richardson, C. C. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 205–209
6. Mendelman, L. V., and Richardson, C. C. (1991) *J. Biol. Chem.* **266,** 23240–23250
7. Frick, D. N., and Richardson, C. C. (1999) *J. Biol. Chem.* **274,** 35889–35898
8. Ilyina, T. V., Gorbalenya, A. E., and Koonin, E. V. (1992) *J. Mol. Evol.* **34,** 351–357
9. Kusakabe, T., Hine, A. V., Hyberts, S. G., and Richardson, C. C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4295–4300
10. Kato, M., Ito, T., Wagner, G., Richardson, C. C., and Ellenberger, T. (2003) *Mol. Cell* **11,** 1349–1360
11. Lee, S. J., and Richardson, C. C. (2004) *J. Biol. Chem.* **279,** 23384–23393
12. Frick, D. N., Baradaran, K., and Richardson, C. C. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 7957–7962
13. Guo, S., Tabor, S., and Richardson, C. C. (1999) *J. Biol. Chem.* **274,** 30303–30309
14. Lee, S. J., and Richardson, C. C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12703–12708
15. Aravind, L., Leipe, D. D., and Koonin, E. V. (1998) *Nucleic Acids Res.* **26,** 4205–4213
16. Keck, J. L., Roche, D. D., Lynch, A. S., and Berger, J. M. (2000) *Science* **287,** 2482–2486
17. Podobnik, M., McInerney, P., O'Donnell, M., and Kuriyan, J. (2000) *J. Mol. Biol.* **300,** 353–362
18. Lee, S. J., and Richardson, C. C. (2001) *J. Biol. Chem.* **276,** 49419–49426
19. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., and Kornberg, R. D. (2001) *Science* **292,** 1876–1882
20. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) *Nature* **417,** 712–719
21. Kusakabe, T., and Richardson, C. C. (1997) *J. Biol. Chem.* **272,** 12446–12453
22. Bernstein, J. A., and Richardson, C. C. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 396–400
23. Aravind, L., Mazumder, R., Vasudevan, S., and Koonin, E. V. (2002) *Curr. Opin. Struct. Biol.* **12,** 392–399
24. Bernstein, J. A., and Richardson, C. C. (1989) *J. Biol. Chem.* **264,** 13066–13073
25. Arezi, B., and Kuchta, R. D. (2000) *Trends Biochem. Sci.* **25,** 572–576
26. Augustin, M. A., Huber, R., and Kaiser, J. T. (2001) *Nat. Struct. Biol.* **8,** 57–61
27. Steitz, T. A., Smerdon, S. J., Jager, J., and Joyce, C. M. (1994) *Science* **266,** 2022–2025
28. Mendelman, L. V., Beauchamp, B. B., and Richardson, C. C. (1994) *EMBO J.* **13,** 3909–3916
29. Strack, B., Lessl, M., Calendar, R., and Lanka, E. (1992) *J. Biol. Chem.* **267,** 13062–13072
30. Klinedinst, D. K., and Challberg, M. D. (1994) *J. Virol.* **68,** 3693–3701
31. Godson, G. N., Schoenich, J., Sun, W., and Mustaev, A. A. (2000) *Biochemistry* **39,** 332–339
32. Matson, S. W., and Richardson, C. C. (1983) *J. Biol. Chem.* **258,** 14009–14016
33. Frick, D. N., Kumar, S., and Richardson, C. C. (1999) *J. Biol. Chem.* **274,** 35899–35907
34. Joyce, C. M., and Steitz, T. A. (1995) *J. Bacteriol.* **177,** 6321–6329
35. Ito, N., Nureki, O., Shirouzu, M., Yokoyama, S., and Hanaoka, F. (2003) *Genes Cells* **8,** 913–923