

## Interaction of Bacteriophage T7 Gene 4 Primase with Its Template Recognition Site\*

(Received for publication, June 9, 1999, and in revised form, September 3, 1999)

David N. Frick and Charles C. Richardson‡

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

**The primase fragment of the bacteriophage T7 63-kDa gene 4 helicase/primase protein contains the 271 N-terminal amino acid residues and lacks helicase activity. The primase fragment catalyzes the synthesis of oligoribonucleotides at rates similar to those catalyzed by the full-length protein in the presence of a 5-nucleotide DNA template containing a primase recognition site (5'-GGGTC-3', 5'-TGGTC-3', 5'-GTGTC-3', or 5'-TTGTC-3'). Although it is not copied into the oligoribonucleotides, the cytosine at the 3'-position is essential for synthesis and template binding. Two nucleotides flanking the 3'-end of the recognition site are required for tight DNA binding and rapid oligoribonucleotide synthesis. Nucleotides added to the 5'-end have no effect on the rate of oligoribonucleotide synthesis or the affinity of the primase for DNA. The binding of either ATP or CTP significantly increases the affinity of the primase for its DNA template. DNA lacking a primase recognition site does not inhibit oligoribonucleotide synthesis, suggesting that the primase binds DNA in a sequence-specific manner. The affinity of the primase for templates is weak, ranging from 10 to 150  $\mu\text{M}$ . The tight DNA binding (<1  $\mu\text{M}$ ) observed with the 63-kDa gene 4 protein occurs *via* interactions between DNA templates and the helicase domain.**

DNA polymerases require the presence of a 3'-hydroxyl terminated DNA or RNA primer to initiate polymerization. By far the most common mechanism for priming DNA synthesis involves the utilization of short oligoribonucleotides. Replicative DNA polymerases efficiently add deoxynucleotides to the 3'-end of RNA primers annealed to ssDNA.<sup>1</sup> In the cell, these RNA primers are synthesized by a specialized enzyme, DNA primase, a DNA-dependent RNA polymerase that catalyzes the synthesis of oligoribonucleotides at specific DNA sequences (1). The enzymatic analysis of oligoribonucleotide synthesis by DNA primases has been complicated primarily because DNA primases function as part of macromolecular complexes. These multiprotein complexes often catalyze additional reactions affecting NTP or DNA concentrations, confounding the examination of RNA synthesis alone. Moreover, when DNA primases

have been separated from these complexes, their RNA synthesis activities are often profoundly decreased (2, 3).

Gene 4 of *Escherichia coli* bacteriophage T7 encodes a multifunctional protein that provides both helicase and primase activities required for the replication of the phage DNA. The gene 4 product is a hexamer (4–6) comprised of two co-linear proteins translated from separate in-frame start sites (7). The smaller 56-kDa protein translocates 5'-to-3' on ssDNA and unwinds duplex DNA (helicase activity) in a reaction coupled to the hydrolysis of nucleoside triphosphates (8–10). The larger 63-kDa protein has all of the activities found in the 56-kDa protein and also acts as a primase by catalyzing the template-directed synthesis of short RNA polymers (11–13). T7 DNA polymerase (14, 15) uses these oligoribonucleotides as primers.

Although the presence of helicase and primase activities in the 63-kDa protein simplifies reconstruction of the T7 replisome (16, 17), the competitive effects of multiple activities have complicated characterization of the T7 DNA primase and helicase activities. For example, although the T7 helicase preferentially hydrolyzes dTTP as the energy source to unwind double-stranded DNA, the enzyme also hydrolyzes ATP (11). Hence, this activity would dramatically lower the concentration of ATP in primase assays, and therefore artificially low rates of primer synthesis would be recorded. The absence of primase activity in the 56-kDa gene 4 protein has enabled detailed kinetic studies to be conducted using the T7 helicase (10, 18, 19). In addition, a "helicase fragment" that contains only the C-terminal helicase domain of the T7 gene 4 protein has been purified and crystallized (20, 21).

When the start codon for the small form of the gene 4 protein is eliminated by substituting methionine-64 with leucine (22) or glycine (23), the mutant alleles support the growth of phage lacking gene 4. When purified, the 63-kDa gene 4 M64L (22) and M64G (23) proteins are indistinguishable from wild-type 63-kDa gene 4 protein. The 63-kDa gene 4 M64G allele was modified to code for a protein containing only the N-terminal 271 amino acids of the T7 gene 4 protein. This "primase fragment" lacks dTTPase and helicase activities but retains primase activity (24). The elimination of the DNA-binding site in the helicase domain of the 63-kDa gene 4 protein allows the direct analysis of protein-DNA interactions occurring at the active site for oligonucleotide synthesis. Comparisons of DNA binding of the primase fragment and the full-length protein reveal that DNA binds more tightly to the helicase domain than to the primase domain of the gene 4 protein, presenting two complications for the study of primer synthesis. First, the tight binding of the DNA to the helicase domain often results in the contamination of 63-kDa gene 4 protein preparations with trace amounts of DNA. Removal of DNA from the gene 4 protein is aggravated by the fact that this tightly bound DNA is protected from digestion by nucleases (25, 26). Second, when short, synthetic DNA oligonucleotides are used to analyze primer synthesis by the 63-kDa gene 4 protein, their binding to

\* This work was supported in part by United States Public Health Service Grant AI-06045 and by American Cancer Society Grant NP-1Z (to C. C. R.). 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.

‡ Supported by American Cancer Society Post-doctoral Fellowship PF-4452. 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@bcmp.med.harvard.edu.

<sup>1</sup> The abbreviations used are: ssDNA, single-stranded DNA; DTT, dithiothreitol.

the helicase DNA-binding site could make them inaccessible to the distal primase active site (27). Hence, the tight binding of DNA to the helicase domain masks the analysis of weaker interactions of DNA with the primase domain.

Here, we present an analysis of oligoribonucleotide synthesis by the T7 primase fragment to understand how the primase synthesizes oligoribonucleotides only at specific DNA sequences called "primase recognition sites." We first determine minimum template requirements for sequence-specific DNA binding by the primase fragment. We show that upon binding a recognition sequence, the primase synthesizes oligoribonucleotides in a template-directed manner. The deoxynucleotides at the 5'-end of the recognition sequence are used to dictate the sequence of the oligoribonucleotide, which occurs in the anti-parallel direction from 5' to 3'. A conserved dC at the 3'-end of the recognition sequence is essential for DNA binding but does not serve as a template. The roles of the NTP precursors in sequence-specific DNA binding by the primase fragment are also defined. These studies uncover important protein-DNA interactions previously masked by the tight binding between DNA and the helicase domain of T7 gene 4 protein. Once bound to a primase recognition site, the T7 primase rapidly synthesizes oligoribonucleotides from NTPs. A detailed analysis of the selective binding of NTPs by the primase and its subsequent oligoribonucleotide synthesis is presented in the accompanying paper (28).

#### EXPERIMENTAL PROCEDURES

**DNA and Enzymes**—Oligonucleotides were synthesized and purified by Integrated DNA Technologies (Corralville, IA). Their sequences are listed in Table II. Throughout this paper, primase recognition sites are underlined. The concentration of each oligonucleotide (mol of 3'-ends/liter) was calculated from  $A_{260}$  and its extinction coefficient.

The 56-kDa and 63-kDa gene 4<sub>G64</sub> proteins were purified by B. Beauchamp (Harvard Medical School) as described (29). The 63-kDa gene 4<sub>G64</sub> protein has a glycine substituted for methionine to eliminate the start site of the 56-kDa gene 4 protein but is indistinguishable from the wild-type protein (30). Therefore, here the 63-kDa gene 4<sub>G64</sub> will be referred to as "wild-type gene 4" for simplicity. The primase fragment contains residues 1–272 of the 63-kDa gene 4<sub>G64</sub> protein and was expressed and purified from colonies of strain Bl21(DE3) containing pETg4P as described (24). The helicase fragment contains residues 227–566 of the 63-kDa gene 4 protein and was expressed and purified by M. Sawaya (Harvard Medical School) as described for the gene 4D protein (20).

**Oligoribonucleotide Synthesis Assay**—Oligoribonucleotide synthesis reactions were measured essentially as described (31, 32) with the following modifications. Standard 10- $\mu$ l reactions contained 40 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 50 mM potassium glutamate, [ $\alpha$ -<sup>32</sup>P]CTP, ATP, and 50 nM 63-kDa gene 4 protein or primase fragment. All reactions were carried out in the presence of 10 mM free Mg<sup>2+</sup>. Assuming each mol of NTP binds one mol of Mg<sup>2+</sup> tightly, NTPs were premixed equal molar concentrations of MgCl<sub>2</sub> before use, and 10 mM additional MgCl<sub>2</sub> was added to each reaction. After incubation at 23 °C for 30 min, reactions were terminated by the addition of 10  $\mu$ l of sequencing dye (98% formamide, 10 mM EDTA, pH 8.0, 0.1% xylene cyanol FF, and 0.1% bromophenol blue). After heating to 95 °C for 5 min, the labeled products were separated by electrophoresis through 25% polyacrylamide gels containing 3 M urea. Radiolabeled products were measured using phosphor-image analysis with a Fuji BAS 1000 Bio-imaging analyzer. Under these conditions, the rates measured were linear with both time and enzyme concentration.

**Gel Shift Assay**—A gel shift assay was used to measure the binding of DNA to the gene 4 proteins and peptide fragments. The 5' terminus of the oligonucleotide 5'-GGGTC<sub>A10</sub>-3' was labeled using [ $\gamma$ -<sup>32</sup>P]ATP and T4 polynucleotide kinase (New England Biolabs). Each binding reaction contained 40 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 50 mM potassium glutamate, 0.3 mg/ml protein, 50  $\mu$ M [5'-<sup>32</sup>P]GGGTC<sub>A10</sub>-3', and 1 mM  $\beta$ , $\gamma$ -methylene dTTP. After incubation for 5 min at 23 °C, the free DNA was separated from protein-DNA complexes by electrophoresis through a 10% polyacrylamide gel using a Mini-PROTEAN II Electrophoresis system (Bio-Rad). In this system, the unbound DNA migrates with the loading dye front. Radiolabeled

DNA was visualized by autoradiography, and proteins were located by staining with Coomassie Blue.

#### RESULTS

In reactions requiring ATP, CTP, and a ssDNA template, the T7 gene 4 protein synthesizes short oligoribonucleotides for use as primers by T7 DNA polymerase (14). The oligoribonucleotides used as primers are mainly tetramers of the sequence pppACCA and pppACCC (33, 34) and pppACAC (35, 36). For example, with  $\Phi$ X174 DNA as a template, primers are synthesized at 13 major primase recognition sites that each share the sequence 5'-GGGTC-3' or 5'-TGGTC-3' (12). The gene 4 primase also synthesizes primers with the sequences 5'-(A/C)G-GTC-3' and 5'-NTGTC-3' on natural ssDNA templates, albeit less frequently (12, 37). Each of the primase recognition sites contains a core trinucleotide sequence 5'-GTC-3', which is sufficient to support the synthesis of the dinucleotide pppAC (31). Oligonucleotide synthesis begins with this dinucleotide at the dT in the recognition sequence, and all primers start with pppA at the 5' terminus (14, 34). The cytosine in the recognition sequence, although essential, is not copied into the primer and hence is designated as "cryptic" (12, 38). Studies using synthetic DNA templates have suggested that in addition to the primase recognition site, the T7 primase requires flanking DNA sequences to catalyze optimal rates of oligoribonucleotide synthesis (27, 31). It remains unclear whether these template requirements stem from direct interactions of the ssDNA with the primase active site or from interaction of ssDNA with the helicase domain of the T7 gene 4 protein. Therefore, the template requirements of the T7 primase were re-examined using the primase fragment of the T7 gene 4 protein. This fragment lacks helicase but retains primase activity (24), and thus reveals only protein-DNA interactions that are essential for oligoribonucleotide synthesis.

**Minimum Template Requirements for Oligoribonucleotide Synthesis Catalyzed by the T7 Gene 4 Primase Fragment**—Fig. 1 shows the products of oligonucleotide synthesis reactions containing ATP, [ $\alpha$ -<sup>32</sup>P]CTP, the primase fragment, and various concentrations of several different DNA templates. The reaction products were separated from [ $\alpha$ -<sup>32</sup>P]CTP on a polyacrylamide gel. The first three templates shown, 5'-GGGTC-3' (lanes 1–5), 5'-TGGTC-3' (lanes 6–10), and 5'-GTGTC-3' (lane 11–15), are composed of the sequences at which the majority of primers are synthesized by the gene 4 primase on  $\Phi$ X174 (12) or M13 DNA (39). The fourth template, 5'-TTGTC-3' (lanes 16–20), is the sequence of a secondary site found on these templates. Pentanucleotides are the minimum DNA template required for the synthesis of tetranucleotides by the T7 primase because, in addition to the four 5'-nucleotides encoding the oligoribonucleotide, they contain the essential 3' cryptic C. In the presence of these pentanucleotide templates, the primase fragment catalyzes the synthesis of di-, tri-, and tetranucleotides, and the rate of oligoribonucleotide synthesis is dependent on the template concentration (Fig. 1).

Like the primase fragment, the full-length protein is capable of synthesizing oligoribonucleotides given a 5-nucleotide-long DNA template containing a primase recognition site but only at high DNA concentrations (>50  $\mu$ M) in the absence of dTTP. These data contrast with earlier reports (27, 31). This may be due to the fact that the previous studies used the 63-kDa gene 4 protein and a lower concentration (1  $\mu$ M) of DNA template or that a low rate of synthesis was out of the range of detection (27, 31). Another possibility is that at low DNA concentrations, the majority of these short DNA templates are bound tightly at the helicase active site and are inaccessible to the primase active site (27). The previous assays were carried out in the presence of dTTP (31) or its nonhydrolyzable analogue  $\beta$ , $\gamma$ -

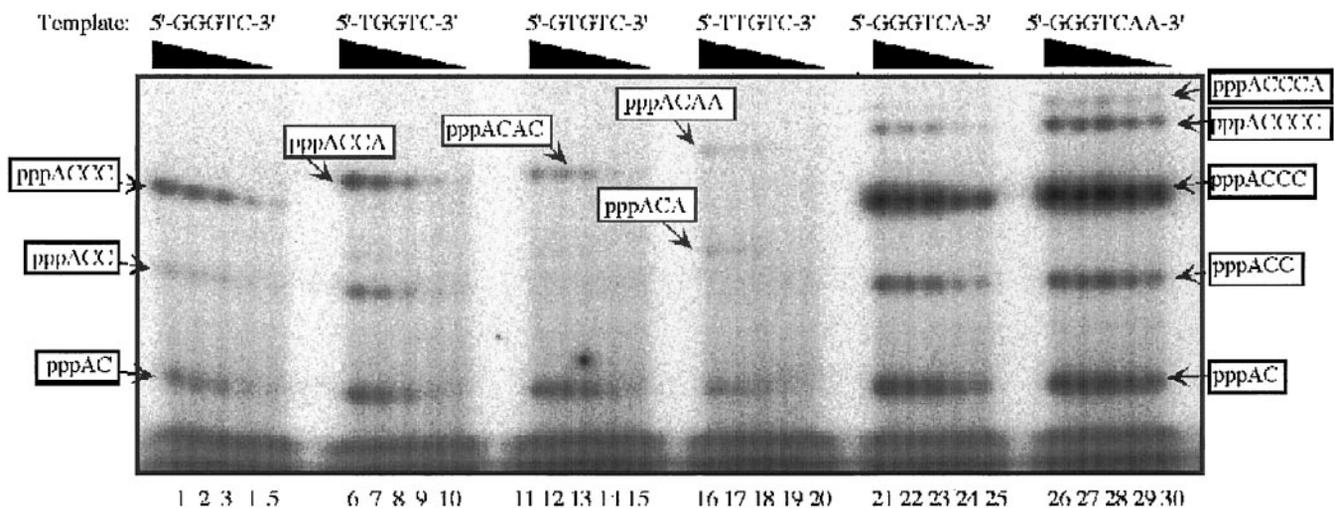


FIG. 1. **Products of oligoribonucleotide reactions catalyzed by the T7 primase fragment.** Oligoribonucleotide synthesis reactions were performed at 23 °C for 30 min in the presence of 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol), 1 mM ATP, DNA template as indicated, and 50 nM primase fragment. The reaction products were separated on a 25% polyacrylamide gel containing 3 M urea and visualized using autoradiography. Reactions shown in lanes 1–5 contained 230, 150, 77, 23, and 7.7  $\mu$ M of the template 5'-GGGTC-3', respectively. Reactions shown in lanes 6–10 contained 270, 180, 90, 27, and 9  $\mu$ M of the template 5'-TGGTC-3', respectively. Reactions shown in lanes 11–15 contained 180, 120, 59, 18, and 5.9  $\mu$ M of the template 5'-GTGTC-3', respectively. Reactions shown in lanes 16–20 contained 180, 120, 64, 18, and 6.4  $\mu$ M of the template 5'-TTGTC-3', respectively. Reactions shown in lanes 21–25 contained 180, 120, 60, 18, and 6.0  $\mu$ M of the template 5'-GGGTCA-3', respectively. Reactions shown in lanes 26–30 contained 170, 120, 58, 17, and 5.8  $\mu$ M of the template 5'-GGGTCAA-3', respectively.

methylene-dTTP (27), both of which dramatically increase the affinity of the T7 helicase for DNA (25).

From the data in Fig. 1, we can conclude that the synthesis of oligoribonucleotides by T7 primase requires DNA and the rate of synthesis is dependent on template concentration. No synthesis occurs in the absence of added template. Thus, the gene 4 protein does not synthesize oligoribonucleotides in the absence of template and does not use these oligonucleotides to locate primase recognition sites on DNA. Moreover, the incorporation of ribonucleotides is template-directed as confirmed by the fact that different oligoribonucleotides are synthesized in the presence of different templates. Oligoribonucleotides containing AMP or CMP have different mobilities in the 25% acrylamide gels used to separate the products of the primase reactions. Oligoribonucleotides containing AMP migrate slightly more slowly than the same length oligoribonucleotides containing CMP. Hence, the trinucleotide pppACA migrates more slowly than the trinucleotide pppACC. The trinucleotide pppACC is synthesized in reactions containing the templates 5'-GGGTC-3', 5'-TGGTC-3', whereas the trinucleotide pppACA is synthesized in reactions containing 5'-GTGTC-3' and 5'-TTGTC-3'. Likewise, the tetranucleotide pppACAA migrates more slowly than pppACAC, which is equivalent to pppACCA in molecular weight. The tetranucleotide having the lowest molecular weight, pppACCC, migrates the most rapidly. The tetranucleotide pppACCC is synthesized in the presence of the recognition sequence 5'-GGGTC-3'; pppACCA is synthesized at 5'-TGGTC-3'; pppACAC is synthesized at 5'-GTGTC-3'; and pppACAA is synthesized at 5'-TTGTC-3'.

**Effect of Nucleotides Flanking the 3'-End of the Primase Recognition Site**—Oligoribonucleotide synthesis catalyzed by the primase fragment in the presence of pentanucleotide templates (Fig. 1, lanes 1–20) shows that a template comprised of only a primase recognition site is necessary and sufficient to support the oligoribonucleotide synthesis. The DNA flanking these primase recognition sites may also affect the rate of oligonucleotide synthesis. Like the cryptic C, these nucleotides need not be used to direct oligoribonucleotide synthesis but may instead interact with the enzyme to stabilize the DNA in the active site. Previously, it was observed that DNA templates

with less than 15 nucleotides flanking the 3'-end of the primase recognition site less efficiently support oligoribonucleotide synthesis by the full-length 63-kDa gene 4 protein (37). However, because these earlier studies were all carried out in the presence of dTTP, the energy source for helicase activity, the requirement for flanking sequences may stem from interactions between the DNA template and the helicase domain of the gene 4 protein. It was important then to re-examine the primase template requirement using the primase fragment, which lacks helicase activity.

After screening a variety of DNA template (see Table II), a dramatic stimulation of primase fragment catalyzed oligoribonucleotide synthesis was observed when nucleotides were added to the 3'-end of the primase recognition site. This effect is illustrated by the final sets of reactions shown in Fig. 1 (lanes 21–30). Lanes 21–25 show the products of reactions containing the template 5'-GGGTCA-3', and lanes 26–30 shows the products of reactions containing the template 5'-GGGTCAA-3'. The same di-, tri-, and tetranucleotide products are produced in reactions containing 5'-GGGTCA-3' and 5'-GGGTCAA-3' as are produced with the sequence 5'-GGGTC-3' (Fig. 1, lanes 1–5). This result demonstrates that, like the cryptic C, these nucleotides are not copied into the oligoribonucleotide primer. All oligoribonucleotide products begin opposite the dT recognition sequence with the dinucleotide pppAC.

In addition to enhancing the rate of primer synthesis, the presence of the 3'-nucleotides allows for the synthesis of a pentanucleotide. The template DNA does not strictly encode this additional nucleotide, as evidenced by the fact that both pppACCCC and pppACCCA are synthesized. However, more pentanucleotide pppACCC is synthesized, suggesting that this phenomenon may be due to the slippage of the primase along the template strand. If such slippage occurs, then the incoming CTP would base pair with the template dG. This pentamer synthesis may be similar to the effect seen when the 56-kDa gene 4 protein is added to 63-kDa protein to stimulate the synthesis of "pseudo-templated" pppACCCC pentamers at the recognition sequence 5'-GGGTC-3' (31).

**Kinetic Constants Describing the Binding of the T7 Primase to Template DNA**—The fact that the presence of two nucleo-

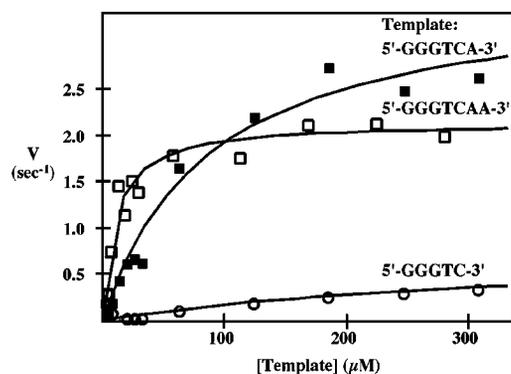


FIG. 2. Effect of nucleotides flanking the 3'-end primase recognition site on the rate of oligonucleotide synthesis. Oligonucleotide synthesis reactions were performed at 23 °C for 30 min in the presence of 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 2 mM ATP, 2 mM [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol), 50 nM primase fragment, and DNA templates ranging from 0.6 to 310  $\mu$ M. Rates of incorporation of [<sup>32</sup>P]CMP into oligoribonucleotides (pmol of CMP/min/pmol of enzyme) were calculated by measuring the amount of each product after separation on a 25% polyacrylamide gel containing 3 M urea using a BAS Fuji BAS 1000 Bio-imaging analyzer. The sequence of the DNA templates were 5'-GGGTC-3' (○), 5'-GGGTCA-3' (■), and 5'-GGGTCAA (□). Curves are calculated using Equation 1, and the values for each template are listed in Table I.

tides flanking the 3'-end of the primase recognition site enhances the rate of primer synthesis suggests that the role of the additional nucleotides is to stabilize the template in the primase active site. To test this hypothesis, we tried to estimate the  $K_d$  of the enzyme for DNA by comparing the rates of oligoribonucleotide synthesis in the presence of different concentrations of DNA templates. The amount of each reaction product was measured by phosphorimage analysis taking into account the number of labeled nucleotides in each oligoribonucleotide. Such an analysis of the data in Fig. 1 reveals that the rate of oligoribonucleotide synthesis is dependent on the concentration of template and that the relative amounts of each product depend on the sequence of the template. For example, the majority of the nucleotides (68%) are incorporated into tetramers in the presence of templates containing the primase recognition sites 5'-GGGTC-3' (lanes 1–5 and 21–30), whereas in the presence of other templates such as 5'-TTGTC-3' (lanes 16–20) the primase fragment incorporates 70% of the nucleotides into dinucleotides. Because tetramers, but not dimers, function as primers in DNA synthesis (40), this observation could explain why secondary primase recognition sites such as 5'-TTGTC-3' are not used frequently to prime DNA synthesis on natural DNA templates (12). A detailed analysis of the amounts and ratios of the various oligoribonucleotide products is presented in the accompanying paper (28). In this study, the rates of primer synthesis will be reported simply as the total amounts of nucleotides incorporated into all of the products per unit time. The rates reflect the fact that one phosphodiester bond is formed per dimer, two bonds are formed per trinucleotide, and three bonds are formed per tetranucleotide product. All rates in this study were obtained under conditions where the rate of product formation was linear with both time and enzyme concentration, and less than 20% of [ $\alpha$ -<sup>32</sup>P]CTP was consumed.

The data in Fig. 1 were insufficient to determine kinetic constants because too few template concentrations were analyzed. Consequently, the rate of oligoribonucleotide synthesis was measured at many additional concentrations of the three most interesting templates: 5'-GGGTC-3', 5'-GGGTCA-3', and 5'-GGGTCAA-3' (Fig. 2). The rate of the primase reaction is clearly dependent on the concentration of DNA and was fit to

TABLE I

Kinetic constants for oligoribonucleotide synthesis by the T7 primase fragment

Rates of nucleotide incorporation into oligoribonucleotides were calculated for reactions containing 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol), 1 mM ATP, and 50 nM primase fragment at 15 concentrations (ranging from 0.6 to 310  $\mu$ M) of the indicated DNA templates. Products of reactions were separated on a 25% polyacrylamide gel containing 3 M urea, and their amounts were measured using a BAS Fuji BAS 1000 Bio-imaging analyzer. Data were fit to Equation 1 using the Macintosh computer program Enzyme Kinetics (Trinity Software) by nonlinear least squares analysis as shown in Fig. 2. Primase recognition sites are underlined.

Template sequence	$K_{DNA}$ $\mu$ M	$V_{max}$ $s^{-1}$	$V_{max}/K_{DNA}$ $s^{-1} mM^{-1}$
5'-GGGTC-3'	150 $\pm$ 80	0.8 $\pm$ 0.2	5 $\pm$ 3
5'- <u>TTGTC</u> -3'	87 $\pm$ 30	3.1 $\pm$ 0.5	36 $\pm$ 12
5'- <u>GTGTCAA</u> -3'	10 $\pm$ 1.5	2.1 $\pm$ 0.2	210 $\pm$ 8

the concentration of DNA template to yield a  $V_{max}$  and  $K_m$  for DNA ( $K_{DNA}$ ) using the following equation.

$$v = V_{max}[DNA]/K_{DNA} + [DNA] \quad (\text{Eq. 1})$$

For each template, nonlinear regression analysis was used to fit the data to Equation 1 using initial estimates of  $K_{DNA}$  and  $V_{max}$  obtained from a double-reciprocal plot of  $1/v$  versus  $1/DNA$  (41). The calculated curve fits produced using the computer program Enzyme Kinetics (Trinity Software) are shown along with the data in Fig. 2. A catalytic efficiency for the enzyme in the presence of each of the DNA template was also calculated, and the resulting kinetic constants are summarized in Table I. If the  $K_{DNA}$  approximates the equilibrium dissociation constant for the enzyme and DNA, then the data in Table I suggest that the two template 5'-GGGTC-3' and 5'-GGGTCA-3' bind the enzyme with similar affinities. However, the addition of a single 3'-nucleotide increases the maximum rate of oligoribonucleotide synthesis 3.8-fold. The template containing a second 3'-nucleotide binds the enzyme dramatically more tightly with a  $K_{DNA}$  of 10  $\mu$ M.

**Effect of Template Length on Primase Catalyzed Oligoribonucleotide Synthesis**—The data in Table I demonstrate that the shortest template required for rapid oligoribonucleotide synthesis and tight DNA binding is a 7-mer. To further define the template requirements for the primase, we have compared the rate of oligoribonucleotide synthesis at saturating levels of the template 5'-GGGTCAA-3' with rate of synthesis in the presence of the same concentration of other DNA templates. The results are presented in Table II. In the presence of oligonucleotides containing only a primase recognition site (oligos 2–5 in Table II), the rate of primer synthesis is 5–32-fold lower than the synthesis in the presence of 5'-GGGTCAA-3'. Synthesis in the presence of the primase recognition sites alone closely parallels the observed frequency of use of primase recognition site on natural DNAs such as  $\Phi$ X174 DNA (12) or M13 DNA (37). The primase recognition site that supports the most synthesis is 5'-GGGTC-3' followed by 5'-TTGTC-3'.

Oligos 6–12 in Table II, each of which contains the primase recognition site 5'-GGGTC-3', were designed to examine the effect of flanking DNA on oligoribonucleotide synthesis by the primase fragment. The presence of a 5'-flanking sequence (5'-CCCCGGGTC-3') had no effect on the rate of oligoribonucleotide synthesis. However, the same primase recognition site with flanking nucleotides on either side (5'-CCCCGGGTCAA-3') supports a 5-fold higher rate. DNA templates containing the recognition sequence 5'-GGGTC-3' and different length flanking sequences at the 3'-end of the recognition site reveal that only a single nucleotide flanking the 3'-end of

TABLE II  
Relative rates oligoribonucleotide synthesis by the T7 primase fragment on various DNA templates

Rates of nucleotide incorporation into RNA primers were measured for reactions containing 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM ATP, 1 mM GTP, 1 mM UTP, 1 mM [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol), 200  $\mu$ M template ( $\mu$ mol 3'-ends/liter), and 50 nM primase fragment at the indicated concentrations of DNA template. Products of reactions were separated on a 25% polyacrylamide gel containing 3 M urea, and their amounts were measured using a BAS Fuji BAS 1000 Bio-imaging analyzer. For comparison, the rates of nucleotide incorporation were divided by the rate of synthesis observed in the presence of the template 5'-GGGTCAA-3' (2 pmol of nucleotides incorporated/s/pmol enzyme). Primase recognition sites are underlined.

Oligo	Sequence	<i>v</i>
		%
1	5'-GGGTCAA-3'	100
2	5'-GGGTC-3'	20
3	5'-TGGTC-3'	10
4	5'-GTGTC-3'	6.4
5	5'-TTGTC-3'	3.1
6	5'-CCCCGGGTC-3'	9.6
7	5'-CCCCGGGTCAAAAA-3'	106
8	5'-GGGTCA-3'	125
9	5'-GGGTCAA-3'	100
10	5'-GGGTCAAAAA-3'	108
11	5'-GGGTCAAAAA-3'	83
12	5'-GGGTCAAAAAAAA-3'	130
13	5'-TGGTCAA-3'	100
14	5'-TGGTCGG-3'	75
15	5'-TGGTCTT-3'	88
16	5'-TGGTCCC-3'	72
17	5'-GGGTGGGGGG-3'	<0.1
18	5'-GGGTAAAAAA-3'	<0.1
19	5'-GGGTTTTTTTT-3'	<0.1
20	5'-GGGTCCCCC-3'	60
21	5'-TGGGCAA-3'	<0.1
22	5'-TGGACAA-3'	<0.1
23	5'-TGGCCAA-3'	<0.1
24	5'-TGATCAA-3'	2.2
25	5'-TGTCAA-3'	4.3
26	5'-TGCTCAA-3'	2.2
27	5'-CAACTGGGGTGGTCTGGTGG-3'	<0.1

the recognition sequence (5'-GGGTCA-3') is sufficient to support the maximum rate of oligoribonucleotide synthesis. Oligonucleotide templates containing 2, 3, 4, 5, or 10 nucleotides flanking the 3'-end of the recognition sequence supported similar rates of primer synthesis.

The role of the base moiety of the 3'-flanking nucleotides was examined using oligos 13–16 (Table II). These four templates contain the primase recognition site 5'-TGGTC-3' flanked on the 3'-end by one of the 4 deoxynucleotides. Each of these templates support similar activity, although flanking dTs and dAs support slightly more synthesis than dGs and dCs.

**Oligoribonucleotide Synthesis on Templates Containing Modified Primase Recognition Sites**—All known T7 primase recognition sites begin with the core trinucleotide sequence 5'-GTC-3'. Oligos 17–26 (Table II) were designed to test the requirement of the primase fragment for these three nucleotides. For example, template 5'-GGGTAAAAAA-3' (oligo 18) is identical to the template 5'-GGGTCAAAAA-3' (oligo 11) except that the cryptic C is replaced with a dA. Even at 200 mM, templates in which the cryptic C was substituted with a dA, dT, or dG residue do not support any detectable oligoribonucleotide synthesis by the primase fragment (Table II). In templates 21–23, the dT in 5'-GTC-3' is replaced with another nucleotide, and in templates 24–27 the dG in 5'-GTC-3' is substituted. Only when the dG is altered is any synthesis detected, albeit at low rate. Templates lacking the dC or the dT did not support detectable synthesis reflecting at least a 1000-fold decrease in rate. We conclude that the cryptic C and the first T are critical

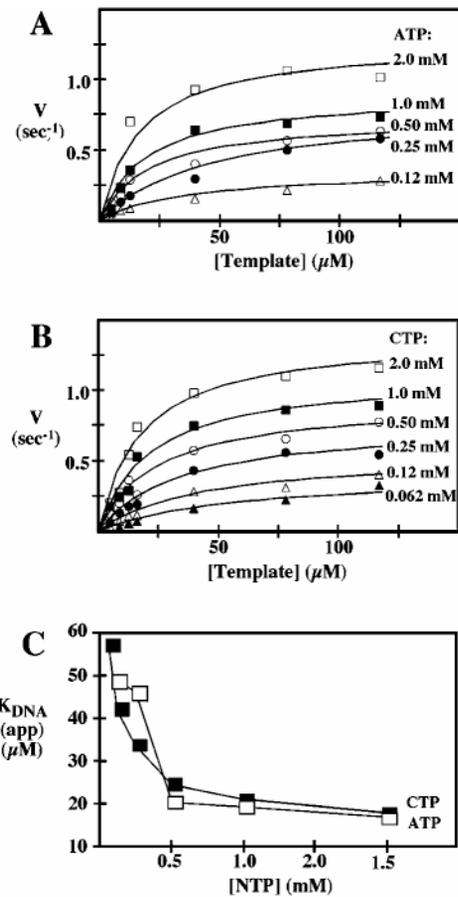


FIG. 3. Effects of ATP and CTP on the affinity of the primase fragment for a DNA template. Oligonucleotide synthesis reactions were performed at 23 °C for 30 min in the presence of 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 50 nM primase fragment, and the indicated concentrations of ATP and [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol) and the DNA template 5'-GGGTCA<sub>10</sub>-3'. A, plot shows velocity of oligoribonucleotide synthesis versus DNA at 2 mM CTP and 0.125 mM ( $\Delta$ ), 0.25 mM ( $\bullet$ ), 0.5 mM ( $\circ$ ), 1 mM ( $\blacksquare$ ), and 2 mM ( $\square$ ) ATP. B, plot shows velocity versus DNA at 2 mM ATP and 0.0625 mM ( $\blacktriangle$ ), 0.125 mM ( $\triangle$ ), 0.25 mM ( $\bullet$ ), 0.5 mM ( $\circ$ ), 1 mM ( $\blacksquare$ ), and 2 mM ( $\square$ ) CTP. Data in A and B were fit to Equation 1 by nonlinear least squares analysis to yield apparent  $K_{DNA}$  and  $V_{max}$  values. C, apparent  $K_{DNA}$  values from A were plotted versus ATP concentration ( $\square$ ), and apparent  $K_{DNA}$  values from B are plotted versus CTP concentration ( $\blacksquare$ ).

for template binding and/or oligoribonucleotide synthesis.

**Effect of NTP Concentration on the Affinity of T7 Primase for Its Recognition Sequence**—Unlike the other classical RNA polymerases, prokaryotic primases catalyze the synthesis of oligonucleotides only in the presence of specific primase recognition sites. This specificity may be due to the fact that the primase uses the bound nucleoside triphosphates to locate the recognition sequences on DNA through base pair formation with the template (42). If so, then the affinity of the T7 primase for oligonucleotide templates containing a single recognition sequence should be affected by the concentration of nucleoside triphosphates. To test this hypothesis, the primase fragment was titrated with a DNA template (5'-GGGTCA<sub>10</sub>-3') at different concentrations of ATP but with a fixed amount of CTP, and the velocity of primer synthesis was measured. Using this template, ATP will be incorporated only at the first position in the product pppACCC. Fig. 3A shows a plot of the velocity of CMP incorporation into oligoribonucleotides at six different template concentrations ranging from 3.9 to 116  $\mu$ M and five different concentrations of ATP ranging from 0.125 to 2 mM.  $K_{DNA}$  values were determined by nonlinear least squares analysis and are clearly dependent on the concentration of ATP as

shown by the secondary plot in Fig. 3C. The values obtained for  $K_{DNA}$  range from 17 to 47  $\mu\text{M}$ , suggesting that the formation of an enzyme-ATP complex significantly increases the affinity of the enzyme for a recognition sequence.

To investigate whether or not the binding of CTP likewise affects the binding of the primase to the DNA template, the velocity of primer synthesis was measured at various concentrations of template and CTP. Again, the template DNA used was 5'-GGGTCA<sub>10</sub>-3'. The concentrations of DNA and CTP were varied at a fixed concentration of ATP (2 mM). Fig. 3B shows a plot of the velocity of CMP incorporation into oligoribonucleotides at six different DNA concentrations ranging from 7.7 to 116  $\mu\text{M}$  and six different concentrations of CTP ranging from 62 to 2 mM. Again, the apparent  $K_{DNA}$  decreases with CTP concentration (Fig. 3C). This result suggests that the formation of the enzyme-CTP complex affects the affinity of the enzyme for the template by at least 3-fold.

**Inhibition of Primase Activity of the Full-length Gene 4 Protein by DNA Lacking a Primase Recognition Site**—To determine whether the interaction of DNA with the primase fragment is sequence-specific, the rate of oligonucleotide synthesis was next determined in the presence of a DNA oligonucleotide containing a single primase recognition site (5'-CCCCGGGT-CAAAA-3') and in the presence of a second DNA oligonucleotide lacking a primase recognition site. If the primase binds to the second template with an affinity equal to that of the DNA containing the primase recognition site, then the second oligonucleotide should be a competitive inhibitor of the reaction. The sequence of the competitor DNA was 5'-CAACTGGGGT-GGTGCTGGTGG-3' (Table II, oligo 27). In this experiment, the concentration of the DNA containing the primase recognition site was held constant at 10  $\mu\text{M}$ . Rates of oligonucleotide synthesis were determined in the absence and in the presence of 12 to 210  $\mu\text{M}$  of the competitor DNA. Either the primase fragment or the 63-kDa gene 4 protein was included in the reactions. In control reactions (not shown), no oligonucleotide synthesis was observed in the presence of 210  $\mu\text{M}$  of the competitor DNA in reactions containing the primase fragment of the 63-kDa gene 4 protein. These reactions were repeated with the 63-kDa gene 4 protein in the presence of either 1 mM dTTP or 1 mM  $\beta,\gamma$ -methylene-dTTP. Both of these nucleoside triphosphates dramatically increase the affinity of the 63-kDa gene 4 protein (25, 43) for DNA but have no effect on the affinity of the primase fragment for DNA (24). As will be shown below, the NTP-dependent ssDNA binding of the gene 4 protein occurs primarily via interaction between DNA and the helicase domain. Hence, this series of reactions should reveal any effect of DNA bound tightly to the helicase domain on the rate of oligoribonucleotide synthesis.

In the absence of competitor DNA, the primase fragment supports 3-fold more oligoribonucleotide synthesis than does an equal molar concentration of the full-length protein. The rate of oligoribonucleotide synthesis by the primase fragment is 60 pmol NMP incorporated/min/pmol primase fragment, whereas the rate of synthesis supported by the 63-kDa gene 4 protein is 20 pmol of NMP incorporated/min/pmol protein monomer. Therefore to compare the relative activities of the two enzymes, all rates of oligonucleotide synthesis were normalized to the rate of oligonucleotide synthesis in the absence of competitor for either the primase fragment or the 63-kDa gene 4 protein (Fig. 4). At these high concentrations of template DNA, dTTP and  $\beta,\gamma$ -methylene-dTTP have no effect on the rate of primer synthesis by the 63-kDa gene 4 protein in the absence of competitor DNA.

No inhibition of oligoribonucleotide synthesis by DNA lacking a primase recognition site is seen in reactions catalyzed by

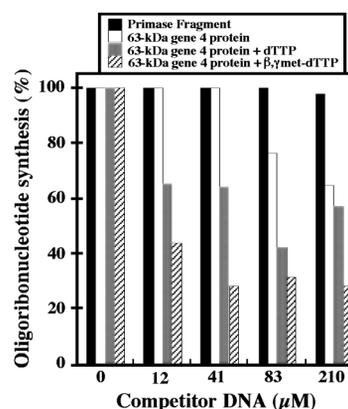


FIG. 4. Inhibition of primase activity of the wild-type gene 4 protein by DNA lacking a primase recognition site. Oligoribonucleotide synthesis reactions were performed at 23 °C for 30 min in the presence of 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM  $\text{MgCl}_2$ , 10 mM DTT, 1 mM [ $\alpha$ - $^{32}\text{P}$ ]CTP (100 Ci/mol), 50 nM primase, and 10  $\mu\text{M}$  DNA template (5'-CCCCGGGTCAAAA-3'). For comparison, the activities of the primase fragment or the 63-kDa gene 4 protein were normalized to their respective activities in the absence of competitor DNA (100%). The bar graph reports the percentage of activity remaining in the absence of competitor DNA and in the presence 12, 41, 83, and 210  $\mu\text{M}$  competitor DNA. The black bars represent reactions containing the primase fragment. The white bars represent data from reactions that contained the 63-kDa gene 4 protein. The gray bars are the results of reactions containing the 63-kDa gene 4 protein with 1 mM dTTP, and the hatched bars report data from reactions containing the 63-kDa gene 4 protein in the presence of  $\beta,\gamma$ -methylene-dTTP.

the primase fragment (Fig. 4, black bars). This result suggests that if the primase fragment binds to the competitor DNA sequence, then it does so very weakly. At the highest concentration of competitor DNA used, the competitor DNA was in 20-fold excess over the DNA with the primase recognition site, yet no less RNA synthesis was detected in these reactions compared with reactions in the absence of competitor DNA. If the primase fragment interacts with the competitor DNA, the  $K_i$  would be well in excess of 200  $\mu\text{M}$ .

The activity of the 63-kDa gene 4 protein (Fig. 4, white bars) is weakly inhibited by the competitor DNA. At the highest concentration of competitor DNA used, more than 50% of the activity of 63-kDa gene 4 protein remains. Hence, the  $K_i$  would again be in excess of 200  $\mu\text{M}$ , although evidently the 63-kDa protein binds the competitor DNA more tightly than the primase fragment. When either dTTP (Fig. 4, gray bars) or  $\beta,\gamma$ -methylene-dTTP (hatched bars) is added to the reaction, the competitor DNA becomes a potent inhibitor of primer synthesis catalyzed by the 63-kDa gene 4 protein. In reactions that included dTTP, 50% inhibition was achieved at concentrations at or above 82  $\mu\text{M}$ , and in reactions containing  $\beta,\gamma$ -methylene-dTTP, 50% inhibition was likewise achieved at concentrations above 12  $\mu\text{M}$ . Such potent inhibition has been described before for the 63-kDa gene 4 protein in reactions containing dTTP. For example, Mendelman *et al.* (38) showed that oligoribonucleotide synthesis on a template containing a primase recognition site is inhibited competitively by a template containing a dT residue replacing the cryptic C primase recognition site (3'-(N<sub>10</sub>)TTGGT(N<sub>5</sub>)-5') with an apparent  $K_i$  of 1.30  $\mu\text{M}$ .

We conclude that DNA lacking a primase recognition sequence blocks primer synthesis only when bound tightly to the helicase domain of the gene 4 protein. In the presence of  $\beta,\gamma$ -methylene-dTTP, the gene 4 hexamer is essentially locked on DNA and is unable to translocate (27). It is therefore not surprising that the DNA lacking primase recognition sites most effectively inhibits oligoribonucleotide synthesis in the presence of  $\beta,\gamma$ -methylene-dTTP. These data confirm a critical role of the helicase domain of the gene 4 protein in the regulating

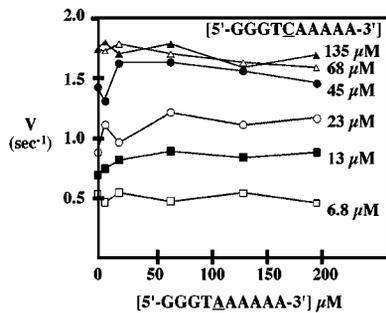


FIG. 5. Effect of a DNA template lacking a cryptic cytosine in the primase recognition site on oligoribonucleotide synthesis. Oligoribonucleotide synthesis reactions were performed at 23 °C for 30 min in the presence of 40 mM Tris-Cl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM [ $\alpha$ -<sup>32</sup>P]CTP (100 Ci/mol), 100 nM primase, at six concentrations of the active DNA template (5'-GGGTCAAAAA-3') and five concentrations of the competitor DNA (5'-GGGTAAAAA-3'). The plot reports the velocity of reaction in terms of pmol of nt incorporated into RNA/min/pmol of primase fragment.

contacts between the primase active site and primase recognition sequences.

*The Cryptic C in the Primase Recognition Site Is Essential for DNA Binding to the Primase Fragment*—We showed earlier (Table II) that the cryptic C in the primase recognition site is essential for oligoribonucleotide synthesis by the primase fragment. It is possible that the enzyme does not bind a template lacking the cryptic C or that the cryptic C plays a more general role in orienting the NTP substrates. To differentiate between these two possibilities, we measured primer synthesis in the presence of various concentrations of two 10-nucleotide oligonucleotides (5'-GGGTCAAAAA-3' and 5'-GGGTAAAAA-3') identical in every aspect except for the presence of the cryptic C. If the cryptic C plays a role in template binding, then the inactive template should not be an inhibitor of oligoribonucleotide synthesis catalyzed in the presence of the active template. The results are shown in Fig. 5. As before, the rate of the reaction was dependent on the concentration of the DNA template containing the primase recognition site. No inhibition of the reaction is seen even under conditions in which the competitor DNA is 28-fold in excess over the proper template containing the cryptic C. In addition, the rates of dimer, trimer, and tetramer formation were also insensitive to the concentration of competitor DNA, revealing that neither the overall rate nor the processivity of the reaction affected titration with competitor DNA. These data highlight the importance of this non-coding nucleotide in the binding of the DNA template to the primase and point to an essential contact between the primase and the cryptic C.

*DNA Binds More Tightly to the Helicase Domain than to the Primase Domain of the Gene 4 Protein*—The  $K_{\text{DNA}}$  values describing the binding of DNA to the primase fragment determined above are 50–100-fold higher than the  $K_{\text{DNA}}$  value of the 63-kDa gene 4 protein (38). Previously, we showed that the 63-kDa gene 4 protein was much more active in the presence of low concentrations of DNA than the primase fragment and that this difference requires the presence of dTTP. In the absence of dTTP the primase fragment and the 63-kDa gene 4 proteins have nearly identical activities at all concentrations of DNA. Because the binding of the 63-kDa gene 4 protein to ssDNA is markedly enhanced in the presence of dTTP, we speculated that the DNA binding results from the interaction of the DNA with the helicase active site (24). To test this prediction, we have analyzed the binding of [ $\gamma$ -<sup>32</sup>P]GGGTCA<sub>10</sub> to the 63-kDa gene 4 protein, the 56-kDa gene 4 protein, the primase fragment, and a helicase fragment. The helicase fragment contains amino acid residues 227–566 of the 63-kDa gene 4 protein,

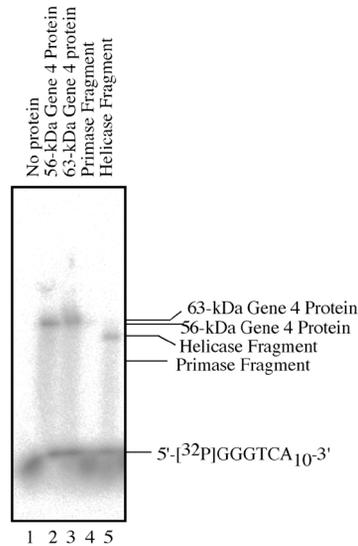


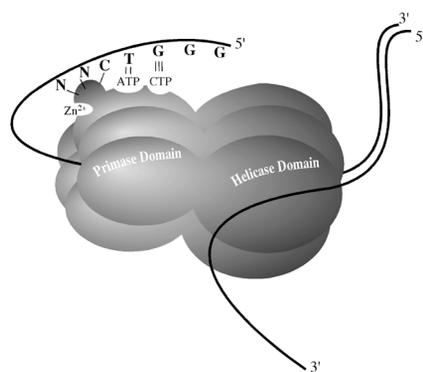
FIG. 6. Gel shift analysis of DNA binding to the gene 4 helicase/primase protein, the primase fragment, and the helicase fragment. Each binding reaction contained 0.3 mg/ml protein, 50  $\mu$ M [ $\gamma$ -<sup>32</sup>P]5'-GGGTCA<sub>10</sub>-3', and 1 mM  $\beta$ , $\gamma$ -methylene dTTP. After incubation for 5 min at 23 °C, the unbound DNA was separated from protein-DNA complexes by electrophoresis through a 10% polyacrylamide gel. Lane 1 (control) contains no protein. Lane 2 contains the 56-kDa gene 4 protein. Lane 3 contains the 63-kDa gene 4 protein. Lane 4 contains the primase fragment (residues 1–272 of the 63-kDa gene 4 protein). Lane 5 contains the helicase fragment (residues 241–566 of the 63-kDa gene 4 protein). Radiolabeled DNA was visualized by the autoradiogram shown. Proteins were located by staining with Coomassie Blue, and their locations are indicated at the left of the autoradiogram.

which includes all residues predicted to be involved in helicase activity (20, 44).

Each of these four proteins was mixed separately with 50  $\mu$ M [ $\gamma$ -<sup>32</sup>P]GGGTCA<sub>10</sub> in the presence of 1 mM  $\beta$ , $\gamma$ -methylene dTTP, a nonhydrolyzable analogue of dTTP that enhances the ssDNA binding of the gene 4 protein (43). Each binding reaction contained 3  $\mu$ g of protein in a final volume of 10  $\mu$ l. After incubation for 5 min at 23 °C, the free DNA was separated from protein-DNA complexes by electrophoresis through a 10% polyacrylamide gel. In this system, the free DNA migrates essentially with the loading dye front. An autoradiogram of the gel is shown in Fig. 6. The results clearly show that the DNA is bound tightly to the 63-kDa gene 4 protein, the 56-kDa gene 4 protein, and the helicase fragment. However, the weak binding of DNA to the primase fragment is not detectable in this assay. We conclude that the ssDNA binds preferentially to the helicase domain of the 63-kDa gene 4 protein.

#### DISCUSSION

The interaction of DNA with the primase fragment of the T7 gene 4 protein described in the present study yields several new insights that lead to a better understanding of the mechanism of action of DNA primases. The elimination of the helicase domain of the T7 gene 4 protein has allowed the first direct studies of the interaction of DNA with the primase active site. This analysis of DNA-primase interactions shows first that the primase requires only a 5-nucleotide DNA sequence containing a primase recognition site to catalyze the template-directed synthesis of oligoribonucleotides from ATP and CTP. Second, two nucleotides flanking the 3'-end of the recognition sequence act to stabilize the DNA in the primase recognition site. Third, the binding of the nucleoside triphosphates at the primase active site influences the affinity of the primase for the DNA template. Fourth, the primase binds DNA in a sequence-specific manner. Finally, the high affinity interaction previously observed between the full-length protein and DNA occurs



**FIG. 7. Diagram showing DNA-protein interactions between T7 gene 4 protein and the primase recognition site.** When viewed in the electron microscope, the T7 gene forms a hexamer of bi-lobal subunits that surrounds ssDNA (6). Interactions between ssDNA and the helicase domain likely occur on the inner and outer surface of the hexamer (18, 49). The primase domain is shown interacting with the primase recognition site *via* contacts between bound nucleoside triphosphates and the Cys<sup>4</sup> zinc finger. The Cys<sup>4</sup> zinc-binding domain and ATP- and CTP-binding sites are shown on the outside of the hexamer for clarity. In the absence of a high resolution structure of the gene 4 protein, attempts to localize the protein domains involved in helicase or primase activities are speculative.

mainly through the interaction of DNA with the helicase active site. The analysis of the binding of DNA to the primase fragment thus provides information critical to the comprehension of how the T7 primase selectively synthesizes oligoribonucleotide primers only at primase recognition sites of specific defined DNA sequences. These interactions are diagrammed schematically in Fig. 7.

We had previously shown that the primase fragment of the 63-kDa gene 4 protein retains DNA-dependent oligoribonucleotide synthesis activity even in the absence of helicase or dTTPase activity (24). This finding was somewhat surprising because most primases need to interact with their cognate helicase proteins to achieve maximum activity. For example, in *E. coli*, the *dnaG*-encoded DNA primase and the *dnaB*-encoded helicase form a complex (45). In phage T4, the gp41 protein (helicase) combines in a 6:1 complex with the T4 gp61 protein (primase) when bound to DNA and ATP (46–48). Phage T7 has devised an elegant coupling of helicase and primase actions by combining both activities into a single protein, the 63-kDa gene 4 protein. This association, as well as the observation that in the absence of dTTP the primase activity of the 63-kDa gene 4 protein is dramatically reduced on M13 DNA, led to the proposal that helicase translocation was required for efficient primer synthesis (31). One clear role of helicase domain in promoting more efficient oligoribonucleotide synthesis by the gene 4 protein is to tether the protein tightly to DNA to bring the gene 4 primase in contact with specific primase recognition sites (24). The inability of the primase fragment to bind DNA lacking a primase recognition site (Fig. 4) highlights this critical role of the helicase domain.

The relationship between the helicase and primase domains of the gene 4 protein is diagrammed schematically in Fig. 7. Six subunits of the full-length 63-kDa gene 4 protein are shown because the active form of the gene 4 protein is likely a hexamer based on size exclusion chromatography, native gel electrophoresis, and chemical cross-linking (4, 5). Helicase fragments lacking residues 241–272, a region linking the helicase with the primase domain, fail to form stable hexamers, suggesting that these residues are critical for hexamer formation (21). When viewed in the electron microscope, gene 4 oligomers appear to form ring-like structures that surround ssDNA (6). The bi-lobal shape of each subunit is intended to indicate two large functional domains and is supported by the shape of the

gene 4 subunits observed in electron micrographs (6). The T7 helicase requires a forked template to initiate strand displacement (9), suggesting interactions between both DNA strands. Current models predict that the hexameric helicases likely bind one strand bound tightly in the inside of the ring complex and excludes the other strand, which makes contact with the outside of the ring (18, 49). The direction of helicase translocation on ssDNA is based on the observation that primase recognition sites are utilized in 5'-to-3' direction on ssDNA templates (12). The primase active site is shown on the exterior of the hexameric gene 4 for clarity and to convey the separation of the helicase and primase active sites. This proposed location is an attempt to explain the fact that when the gene 4 protein is tightly bound to DNA *via* the helicase domain, the distal primase active site is still accessible to other DNA templates (27).

The model in Fig. 7 is likely an oversimplification of the shape of the gene 4 protein and the resulting interactions between the gene 4 protein and DNA. Two lines of evidence point to a more complexly intimate relationship between the helicase and primase active sites. First, DNA bound tightly to the helicase domain blocks oligoribonucleotide synthesis at the primase active site on a second DNA template (Fig. 4), suggesting that a single DNA-binding site is required for helicase and primase activities or two overlapping DNA-binding sites. Secondly, studies using the full-length gene 4 protein show a requirement for 5–10-nucleotide flanking sequences on either side of the primase recognition site (27, 31). The primase fragment requires only two 3'-flanking nucleotides for maximum oligoribonucleotide synthesis (Table I). Taken together, these data suggest that primase DNA-binding region and the helicase DNA-binding region are sufficiently close that a 21-mer DNA template may be bound simultaneously in both sites. Moreover, in the absence of translocation, the template DNA must be long enough to loop between the helicase and primase DNA-binding sites, or the primase recognition site may be chelated in the helicase domain (27).

Although the helicase domain of the T7 gene 4 protein binds much more tightly to DNA than the primase domain (Fig. 6), it does so in a sequence-independent manner (25, 43). In contrast, the primase fragment only binds DNA containing a primase recognition site (Figs. 4 and 5). The fact that the primase fragment synthesizes oligoribonucleotides only at specific recognition sites shows that no residues in the helicase domain (amino acids 273–566) are necessary for sequence-specific DNA binding. In all previous studies using the full-length gene 4 protein, interactions between DNA and the primase domain of the gene 4 protein were masked because of tight binding of the helicase domain to DNA. The highly active primase fragment has thus been invaluable in determining how the primase selectively synthesizes oligonucleotides only at specific primase recognition sites. The data presented here (Figs. 4 and 5) indicate that DNA lacking a primase recognition site binds less tightly to the primase active site than DNA containing a primase recognition site. An alternative explanation would be that DNA lacking a primase recognition site binds equally well, and upon binding the functional groups required to catalyze phosphodiester bond formation are improperly aligned. The former explanation seems much more likely because the latter would predict inhibition by DNA lacking a primase recognition site.

We initially intended to investigate the interaction of DNA with the primase fragment by gel shift assays or by nitrocellulose filter binding assays (25). However, such techniques were far too insensitive to detect the relatively weak binding of the primase fragment to DNA (Fig. 6). Because oligoribonucleotide synthesis is dependent on the presence of a template DNA

containing a primase recognition site (Fig. 1), we used a kinetic assay to analyze the binding of DNA to the primase fragment. The shortest DNA sequence that supported DNA-dependent oligoribonucleotide synthesis contained only a 5-nucleotide primase recognition site. The primase fragment catalyzed the synthesis of oligonucleotides in the presence of 5-nucleotide DNA templates containing a primase recognition site, and the sequence of oligonucleotides synthesized is directed by the sequence of the DNA recognition site (Fig. 1). A DNA sequence flanking the 5'-end of the recognition sequence does not enhance the efficiency of primer oligoribonucleotide synthesis. One might expect the 5'-flanking sequence length to affect oligoribonucleotide synthesis because 63-kDa gene 4 protein uses primase recognition sites in a 5'-to-3' direction on ssDNA (12). In stark contrast, the addition of only one or two nucleotides flanking the 3'-end of the recognition sequence greatly enhances both the binding of the primase fragment to DNA and the efficiency of oligonucleotide synthesis. Sequence-specific DNA binding could be potentially achieved through base pair formation of the DNA template with bound nucleoside triphosphate substrates. Indeed, the formation of the enzyme-ATP and formation of the enzyme-CTP complex increases the affinity of the primase for DNA (Fig. 3). These effects could result from the formation of base pairs between the enzyme bound NTPs and the template DNA (Fig. 7).

Also shown in the model depicted in Fig. 7 are the interactions between the primase domain and the cryptic C nucleotide in the primase recognition site and the two nucleotides flanking the 3'-end of the recognition sequence. The presence of two 3'-flanking nucleotides greatly increases the efficiency of primase recognition site usage (Fig. 1 and Table I). Furthermore, templates in which the cryptic C is replaced by another nucleotide support less than 0.1% of the oligoribonucleotide synthesis as templates containing a cryptic C of the primase fragment. These direct experiments reveal that the requirement for the cryptic C is even more stringent than that previously thought (42). Hence, the requirement for the cryptic C was re-examined using the 63-kDa gene 4 protein in the presence of dTTP. As before (42), templates in which the cryptic C was replaced with G, A, or T supported 12, 15, and 2% oligoribonucleotide synthesis, respectively. However, when dTTP, which promotes ssDNA binding by the helicase (43), was omitted from the reaction, the 63-kDa gene 4 protein had a template specificity similar to that of the primase fragment. In the absence of dTTP, templates lacking the cryptic C support no detectable oligonucleotide synthesis by either the primase fragment or the full-length gene 4 protein. Thus, the full contribution of the cryptic C to primase template binding is revealed only in the absence of DNA binding to the helicase domain of the gene 4 protein. An interaction of the primase with the cryptic C could enable the primase to bind DNA in a sequence-specific manner. Such a role is supported by the fact that an oligonucleotide lacking the cryptic C fails to inhibit oligoribonucleotide synthesis (Fig. 5).

An obvious candidate region of the 63-kDa gene 4 protein that may contact the primase recognition site in the template DNA is the Cys<sup>4</sup> zinc finger motif. The zinc finger is located in the N terminus of the protein between amino acid residues 17 and 39. This region is included within the 63 amino acids missing from the 56-kDa gene 4 protein (7), which lacks primase activity (50). Substitution of a single cysteine in this motif with serine (C36S) abolishes DNA directed oligoribonucleotide synthesis by the T7 primase (51). In addition, further site-directed mutagenesis of residues in this motif, in particular His<sup>33</sup> and Asp<sup>31</sup>, confirm the importance of the zinc finger in primase activity and points out a possible interaction between

the Cys<sup>4</sup> motif and the cryptic C in the recognition sites (42). It should be noted that this motif alone does not appear to confer all the sequence specificity of the primase for primase recognition sites based on the characterization of chimeric primases (52, 53). Hence, in Fig. 7 the Cys<sup>4</sup> motif is not shown making contacts with the entire primase recognition site.

The relatively weak DNA binding of the primase active site to DNA fits with current models of how the T7 primase functions in the T7 DNA replication complex (17). The apparent  $V_{\max}$  of the primase fragment catalyzed synthesis corresponds to a rate of nucleotide incorporation of 2.1 nucleotides/s/monomer. This rate exceeds the rate previously measured for oligonucleotide synthesis on M13 DNA by the full-length 63-kDa gene 4 protein of 0.8/s/hexamer (0.13/s/monomer) (37). The 16-fold faster rate measured under optimal conditions for the primase fragment is more than sufficient to account for the initiation of new Okazaki fragments of an average length of 3000 base pairs (16, 17, 54) assuming a rate of DNA synthesis of 300 nucleotides/s (55). In fact, the length of the Okazaki fragments synthesized suggests that the primase only uses a fraction of the available primase recognition sites. Tight DNA binding by the primase active site would actually inhibit the progress of the replisome by causing the complex to pause at each and every primase recognition site. The weak DNA binding by the primase domain explains the distributive nature of the primase reaction at the replication fork. Moreover, the tight DNA binding by the helicase explains how the helicase permits processive synthesis of the leading strand (39). Weak DNA binding by the primase would also allow rapid synthesis of DNA. Only when the replisome pauses would primer synthesis occur at specific primase recognition sites. Pausing may occur because of physical constraints on the replisome that occur during the coupled synthesis of leading and strands of DNA. Rapid oligonucleotide synthesis and weak DNA binding by the primase domain of the gene 4 protein would thus enable the efficient synthesis of primers only when needed to initiate new Okazaki fragments.

*Acknowledgments*—We are grateful to Benjamin B. Beauchamp and Michael R. Sawaya for providing purified proteins and to Shenyuan Guo and Kajal Chowdhury for critically reading the manuscript.

#### REFERENCES

- Griep, M. A. (1995) *Indian J. Biochem. Biophys.* **32**, 171–178
- Swart, J. R., and Griep, M. A. (1995) *Biochemistry* **34**, 16097–16106
- Sheaff, R. J., and Kuchta, R. D. (1993) *Biochemistry* **32**, 3027–3037
- Patel, S. S., and Hingorani, M. M. (1993) *J. Biol. Chem.* **268**, 10668–10675
- Notarnicola, S. M., Park, K., Griffith, J. D., and Richardson, C. C. (1995) *J. Biol. Chem.* **270**, 20215–20224
- Egelman, H. H., Yu, X., Wild, R., Hingorani, M. M., and Patel, S. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3869–3873
- Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477–535
- Kolodner, R., and Richardson, C. C. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1525–1529
- Matson, S. W., Tabor, S., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 14017–14024
- Bernstein, J. A., and Richardson, C. C. (1988) *J. Biol. Chem.* **263**, 14891–14899
- Matson, S. W., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 14009–14016
- Tabor, S., and Richardson, C. C. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 205–209
- Bernstein, J. A., and Richardson, C. C. (1989) *J. Biol. Chem.* **264**, 13066–13073
- Scherzinger, E., Lanka, E., and Hillenbrand, G. (1977) *Nucleic Acids Res.* **4**, 4151–4163
- Nakai, H., and Richardson, C. C. (1986) *J. Biol. Chem.* **261**, 15208–15216
- Debyser, Z., Tabor, S., and Richardson, C. C. (1994) *Cell* **77**, 157–166
- Lee, J., Chastain, P. D., Kusakabe, T., Griffith, J. D., and Richardson, C. C. (1998) *Mol. Cell* **1**, 1001–1010
- Hacker, K. J., and Johnson, K. A. (1997) *Biochemistry* **36**, 14080–14087
- Washington, M. T., and Patel, S. S. (1998) *J. Biol. Chem.* **273**, 7880–7887
- Bird, L. E., Hakansson, K., Pan, H., and Wigley, D. B. (1997) *Nucleic Acids Res.* **25**, 2620–2626
- Guo, S., Tabor, S., and Richardson, C. C. (1999) *J. Biol. Chem.*, **274**, 30303–30309
- Rosenberg, A. H., Patel, S. S., Johnson, K. A., and Studier, F. W. (1992) *J. Biol. Chem.* **267**, 15005–15012
- Mendelman, L. V., Notarnicola, S. M., and Richardson, C. C. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10638–10642
- Frick, D. N., Baradaran, K., and Richardson, C. C. (1998) *Proc. Natl. Acad. Sci.*

- U. S. A. **95**, 7957–7962
25. Hingorani, M. M., and Patel, S. S. (1993) *Biochemistry* **32**, 12478–12487
26. Yong, Y., and Romano, L. J. (1995) *J. Biol. Chem.* **270**, 24509–24517
27. Kusakabe, T., Baradaran, K., Lee, J., and Richardson, C. C. (1998) *EMBO J.* **17**, 1542–1552
28. Frick, D. N., Kumar, S., and Richardson, C. C. (1999) *J. Biol. Chem.* **274**, 35899–35907
29. Notarnicola, S. M., and Richardson, C. C. (1993) *J. Biol. Chem.* **268**, 27198–27207
30. Mendelman, L. V., Notarnicola, S. M., and Richardson, C. C. (1993) *J. Biol. Chem.* **268**, 27208–27213
31. Mendelman, L. V., and Richardson, C. C. (1991) *J. Biol. Chem.* **266**, 23240–23250
32. Patel, S. S., Hingorani, M. M., and Ng, W. M. (1994) *Biochemistry* **33**, 7857–7868
33. Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D., and Yuki, A. (1977) *Eur. J. Biochem.* **72**, 543–558
34. Romano, L. J., and Richardson, C. C. (1979) *J. Biol. Chem.* **254**, 10483–10489
35. Nakai, H., and Richardson, C. C. (1986) *J. Biol. Chem.* **261**, 15217–15224
36. Nakai, H., and Richardson, C. C. (1988) *J. Biol. Chem.* **263**, 9831–9839
37. Kusakabe, T., and Richardson, C. C. (1997) *J. Biol. Chem.* **272**, 5943–5951
38. Mendelman, L. V., Kuimelis, R. G., McLaughlin, L. W., and Richardson, C. C. (1995) *Biochemistry* **34**, 10187–10193
39. Nakai, H., and Richardson, C. C. (1988) *J. Biol. Chem.* **263**, 9818–9830
40. Kusakabe, T., and Richardson, C. C. (1997) *J. Biol. Chem.* **272**, 12446–12453
41. Lineweaver, H., and Burk, D. (1934) *J. Am. Chem. Soc.* **56**, 658–666
42. Kusakabe, T., Hine, A. V., Hyberts, S. G., and Richardson, C. C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4295–4300
43. Matsun, S. W., and Richardson, C. C. (1985) *J. Biol. Chem.* **260**, 2281–2287
44. Ilyina, T. V., Gorabalenya, A. E., and Koonin, E. V. (1992) *J. Mol. Evol.* **34**, 351–357
45. Arai, K., and Kornberg, A. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4308–4312
46. Liu, C. C., and Alberts, B. M. (1981) *J. Biol. Chem.* **256**, 2821–2829
47. Hinton, D. M., and Nossal, N. G. (1987) *J. Biol. Chem.* **262**, 10873–10878
48. Dong, F., and von Hippel, P. H. (1996) *J. Biol. Chem.* **271**, 19625–19631
49. Ahnert, P., and Patel, S. S. (1997) *J. Biol. Chem.* **272**, 32267–32273
50. Bernstein, J. A., and Richardson, C. C. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 396–400
51. Mendelman, L. V., Beauchamp, B. B., and Richardson, C. C. (1994) *EMBO J.* **13**, 3909–3916
52. Hine, A. V., and Richardson, C. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12327–12331
53. Kusakabe, T., and Richardson, C. (1996) *J. Biol. Chem.* **271**, 19563–19570
54. Park, K., Debyser, Z., Tabor, S., Richardson, C. C., and Griffith, J. D. (1998) *J. Biol. Chem.* **273**, 5260–5270
55. Lechner, R. L., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 11185–11196