Zinc-binding Domain of the Bacteriophage T7 DNA Primase Modulates Binding to the DNA Template

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:41483421

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Zinc-binding Domain of the Bacteriophage T7 DNA Primase Modulates Binding to the DNA Template

Seung-Joo Lee, Bin Zhu, Barak Akabayov, and Charles C. Richardson

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

Received for publication, August 27, 2012, and in revised form, September 25, 2012. Published, JBC Papers in Press, September 28, 2012, DOI 10.1074/jbc.M112.414151

Background: The zinc-binding domain of DNA primase has been implicated in template recognition.

Results: Residues around cysteines within the zinc-binding domain coordinate template binding but not sequence recognition.

Conclusion: The integrity and functional conformation of the zinc-binding domain are important for DNA primase function.

Significance: The zinc-binding domain of DNA primase is essential for template binding and primer synthesis.

The zinc-binding domain (ZBD) of prokaryotic DNA primases has been postulated to be crucial for recognition of specific sequences in the single-stranded DNA template. To determine the molecular basis for this role in recognition, we carried out homolog-scanning mutagenesis of the zinc-binding domain of DNA primase of bacteriophage T7 using a bacterial homolog from *Geobacillus stearothermophilus*. The ability of T7 DNA primase to catalyze template-directed oligoribonucleotide synthesis is eliminated by substitution of any five-amino acid residue-long segment within the ZBD. The most significant defect occurs upon substitution of a region (Pro-16 to Cys-20) spanning two cysteines that coordinate the zinc ion. The role of this region in primase function was further investigated by generating a protein library composed of multiple amino acid substitutions for Pro-16, Asp-18, and Asn-19 followed by genetic screening for functional proteins. Examination of proteins selected from the screening reveals no change in sequence-specific recognition. However, the more positively charged residues in the region facilitate DNA binding, leading to more efficient oligoribonucleotide synthesis on short templates. The results suggest that the zinc-binding mode alone is not responsible for sequence recognition, but rather its interaction with the RNA polymerase domain is critical for DNA binding and for sequence recognition. Consequently, any alteration in the ZBD that disturbs its conformation leads to loss of DNA-dependent oligoribonucleotide synthesis.

A problem encountered by the replisome is the simultaneous replication of antiparallel DNA strands by the unidirectional polymerizing activity of DNA polymerases. This dilemma is resolved by different modes of DNA synthesis for the leading and lagging strands. Unlike the continuous synthesis of the leading strand, the lagging strand is synthesized discontinuously in the form of Okazaki fragments. The Okazaki fragments reside within a replication loop that allows for the overall synthesis of both strands in the same direction with all events occurring within the replisome. Synthesis of the Okazaki fragments however necessitates a supply of oligoribonucleotides on the single-stranded lagging strand extruded by the helicase during leading strand synthesis. DNA primases are designated to fulfill this function by catalyzing the synthesis of oligoribonucleotides (1).

Unlike their eukaryotic counterparts, prokaryotic DNA primases (DnaG family) synthesize primers from specific sites on single-stranded DNA (ssDNA). Amino acid sequence alignment of DNA primases from bacteria and bacteriophages has identified six conserved motifs (2). The first motif at the N terminus of the primase is a zinc-binding domain (ZBD) that consists of four cysteines (Cys_{4}) or three cysteines and one histidine (Cys-His-Cys_{3}) that coordinate a zinc ion. The other five motifs reside in an RNA polymerase domain (RPD) located in the C-terminal portion of the primase where binding and the catalytic condensation of NTPs occur. Motifs 4–6 form a TOPRIM-fold that contains two magnesium ions that participate in a two metal-mediated phosphodiester formation reaction similar to most polymerases (3). Together with the TOPRIM-fold, motifs 2 and 3 are believed to create the template-binding site. Although the RPD is responsible for the template-directed coupling of the NTPs, the ZBD plays a critical role in the binding of the primase to specific trinucleotide sequences in the template defined as primase recognition sites (1). Each prokaryotic primase has its own recognition sequence for primer synthesis, distinct from others. The “cryptic” nucleotide at the 3’-side of the trinucleotide is not copied into RNA products, but its presence is essential for recognition of the sequence by the primase. NTPs complementary to the next nucleotides at the 5’-side of the template are then added to generate the dinucleotide, and it is further extended to the oligonucleotide of a functional length that can be used by the DNA polymerase as a primer. Primers synthesized *de novo* must be stabilized by the primase and then handed off to DNA polymerase for the initiation of DNA synthesis.

*This work was supported, in whole or in part, by National Institutes of Health Grant GM 54397 (to C. C. R.).

This article contains supplemental Figs. S1–S5 and Table 1.

1 To whom correspondence may be addressed: Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., C2-219, Boston, MA 02115. Tel.: 617-432-1864; Fax: 617-432-3362; E-mail: ccr@hms.harvard.edu.

2 The abbreviations used are: ssDNA, single-stranded DNA; ZBD, zinc-binding domain; RPD, RNA polymerase domain.
One of the well characterized prokaryotic primases is the primase-helicase encoded by gene 4 of bacteriophage T7. The C-terminal domain of the gene 4 protein is a member of superfamily 4 hexameric helicases. The helicase domain binds ssDNA and unwinds parental DNA at the expense of dTTP hydrolysis. The primase domain, located in the N-terminal half of the gene 4 protein, recognizes the sequence 5′-GTC-3′ in ssDNA. At this trinucleotide recognition site, the T7 DNA primase catalyzes the synthesis of the diribonucleotide pppAC and further extends it to functional tetraribonucleotides, and pppACCC, pppACCA, and pppACAC provided cognate sequences are available in the template (4, 5). The crystal structure of the T7 primase revealed a modular organization in which the ZBD and the RPD are connected by a flexible linker (6). The ZBD of T7 primase contains a four-stranded, antiparallel β-sheet (Fig. 1). Four cysteines, located at loop regions connecting β-strands, coordinate a zinc ion to form a structure typical of members of the zinc-ribbon family. The RPD of T7 primase contains an N-terminal subdomain consisting of four β-strands, coordinate a zinc ion during primer synthesis (9). Within the hexameric structure composed of full-length gene 4 primase-helicase, a ZBD from one subunit can interact with the RPD from the adjacent subunit to catalyze synthesis of oligoribonucleotides in a trans mode (10), suggesting a regulatory role of primase during DNA synthesis (11).

The importance of the ZBD in DNA-dependent oligoribonucleotide synthesis was initially noticed in studies with the 56-kDa gene 4 protein, a truncated form of the 63-kDa gene 4 protein, that arises from in-frame translation of the full-length gene 4 (12). The 56-kDa gene 4 protein lacks the N-terminal 63 amino acid residues corresponding to the ZBD and is not capable of catalyzing primer synthesis in a template-dependent manner (13). However, the 56-kDa gene 4 protein, like the full-length protein, does synthesize random diribonucleotide in the absence of template, albeit at a very slow rate. Replacement of a single cysteine in the ZBD with a serine results in loss of the zinc ion and abolishes template-directed synthesis (14). Furthermore, site-directed mutagenesis has identified amino acid residues critical to the recognition of 5′-GTC-3′ (15).

A persistent goal has been to elucidate the molecular mechanism by which the T7 primase recognizes a trinucleotide sequence in ssDNA. Prokaryotic DNA primases, each having a distinct ZBD, synthesize primers of different sequences. Consequently, it has been suggested that sequence specificity is dictated by interactions between amino acid residues in the ZBD and the trinucleotide sequence in DNA (1). Previously, we attempted to modify the specificity by either swapping the entire ZBD (16) or a portion of the zinc-spanning region with equivalent parts of homologs (17). These studies demonstrated that the integrity of ZBD is important in the recognition of specific sequences, but they failed to reveal the basis for the recognition.

Recent reports of the structures and recognition sequences of DNA primases have provided an approach to obtain additional information on the role of the ZBD in template recognition and binding (18–20). X-ray crystallographic structures show that the ZBDs of the T7 primase and of the primase of Geobacillus stearothermophilus (formerly known as Bacillus stearothermophilus) are similar in their overall structures (6, 21). With the exception of some misaligned regions at the N terminus of the G. stearothermophilus primase and the C terminus of the T7 primase, the two enzymes exhibit a similar folding in the zinc-spanning region (Fig. 1). Both primases have four antiparallel β-sheets forming two loops in which cysteines (or a histidine) coordinate a zinc ion. Despite these resemblances, T7 primase recognizes 5′-GTC-3′ in the template, whereas G. stearothermophilus primase recognizes either 5′-TTA-3′ or 5′-CTA-3′ (22). In an attempt to identify a region critical to the function of the ZBD, we have constructed altered T7 primases containing homologous portion of the G. stearothermophilus primase and examined their ability to catalyze template-directed synthesis of oligoribonucleotides and to deliver them to DNA polymerase. In addition, a collection of proteins consisting of a variety of amino acid substitutions in the vicinity of zinc-coordinating residues was generated and analyzed to examine their interaction with the template DNA. Our results lead us to conclude that the zinc-binding motif is not responsible for sequence recognition. However, its interaction with the RNA polymerase domain is critical for DNA binding and hence for sequence recognition.

**EXPERIMENTAL PROCEDURES**

**Materials—**Oligonucleotides were obtained from Integrated DNA Technology and Dharmacon. Plasmid DNA purification kits and nickel-nitrilotriacetic acid resin were from Qiagen. Restriction endonucleases, Deep Vent® polymerase, and T4 DNA ligase were purchased from New England Biolabs. Radio-labeled nucleotides were purchased from PerkinElmer Life Sciences. Sensor chips and reagents for surface plasmon resonance studies were from GE Healthcare.

**Mutagenesis, Overproduction, and Purification of Recombinant Proteins—**Alterations at selective positions of the ZBD were introduced into plasmids expressing either the entire T7 gene 4 (pET24-gp4 or pET28-gp4) or the T7 DNA primase coding region (pET24-gP4) following procedures described previously (7). A DNA fragment containing alterations in gene 4 was generated by PCR and incubated with NdeI and BstBI. The insert DNA purified on an agarose gel was ligated into the vector pET24a (+) treated with the same restriction enzymes. After the insert DNA was ligated, the plasmid was transformed into Escherichia coli BL21 (DE3). Proteins were overproduced by inducing cells with 1 mM isopropyl β-D-1-thiogalactopyranoside at 37 °C for 3 h after the culture reached A600 of around 1. A full-length gene 4 protein containing no His tag was purified as described previously (7) using DEAE-anion exchange column chromatography instead of ATP-agarose affinity chromatography. The ATP-agarose affinity chromatography was omitted to avoid use of EDTA, a potential chelator for zinc ion during the protein elution step. Pro-
Zinc-binding Domain of T7 DNA Primase

...proteins containing a His tag at the N terminus were purified using nickel-nitriilotriacetic acid affinity chromatography followed by DEAE-anion exchange column chromatography (23). All purified proteins were over 90% pure as judged on SDS-polyacrylamide gel.

Complementation of Phage Growth—E. coli DH5α (0.3 ml) harboring a plasmid that expresses T7 gene 4 was cultured until an A600 of around 1 and mixed with serially diluted T7Δ4 phage lacking gene 4 (0.1 ml) in 3 ml of soft agar. The mixture was plated on the top of LB plate, and the plaques that emerged after incubation at 37 °C for 5 h were counted. Additional deletions around gene 4 in T7Δ4 (T7 nucleotide 11,481–13,621) prevent homologous recombination between the phage and the gene 4-expressing plasmid.

Template-directed Oligoribonucleotide Synthesis—Reaction mixture contained 40 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 50 mM potassium glutamate, 0.1 mM each of dATP and [α-32P]CTP (0.1 μCi/μl), indicated oligomer DNA containing a primase recognition sequence (5'-GGGTC-3' or 5'-TGGTC-3') or 10 mM M13 ssDNA, and the indicated amount of gene 4 protein or primase. After incubation at 37 °C for 30 min, the reaction was terminated by the addition of 2 μl of sequencing dye, and reaction products were separated on a 25% denaturing polyacrylamide sequencing gel containing 3 M urea. The gel was dried for autoradiography, and the amount of CMP incorporated into the major product, the tetraribonucleotide, was rated into the major product, the tetraribonucleotide, was determined using a Fuji BAS 1000 Bioimaging analyzer.

DNA-binding Domain of T7 DNA Primase—Reaction mixture contained 40 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 50 mM potassium glutamate, 0.1 mM each of ATP and [α-32P]CTP (0.1 μCi/μl), indicated oligomer DNA containing a primase recognition sequence (5'-GGGTC-3' or 5'-TGGTC-3') or 10 mM M13 ssDNA, and the indicated amount of gene 4 protein or primase. After incubation at 37 °C for 30 min, the reaction was terminated by the addition of 2 μl of sequencing dye, and reaction products were separated on a 25% denaturing polyacrylamide sequencing gel containing 3 M urea. The gel was dried for autoradiography, and the amount of CMP incorporated into the major product, the tetraribonucleotide, was determined using a Fuji BAS 1000 Bioimaging analyzer.

Measurement of DNA Binding Affinity by Surface Plasmon Resonance—The binding affinity of T7 primase to ssDNA was determined using surface plasmon resonance as described previously (25). Single-stranded 25-mer DNA biotinylated at the 5' end was immobilized on a streptavidin-coated sensor chip. Binding of protein was detected by flowing the indicated primase (10 μM) over the chip in a binding buffer (10 mM Tris-HCl, pH 7.5, 50 mM potassium glutamate, 10 mM MgCl₂, and 1 mM DTT) at a flow rate of 40 μl/min. The surface of the chip was regenerated by injecting 100 μl of 1 M NaCl and 50 mM NaOH at a flow rate of 100 μl/min to remove bound proteins. Sensing was determined by subtracting the nonspecific signal from a control flow cell in which biotin was immobilized instead of biotinylated DNA.

RESULTS

Homolog-scanning Mutagenesis of ZBD of T7 DNA Primase—The structures of the ZBDs of T7 DNA primase and G. stearothermophilus primase are similar especially near the zinc atom as seen in their superimposed structures (Fig. 1, A and B). An alignment of amino acid sequences shows that their secondary structures are also conserved (Fig. 1C). Yet the two primases recognize different trinucleotide sequences for primer synthe-
Proteins exhibit a 102- to 107-fold decrease in their ability to support growth of T7 phage lacking gene 4. Plaque sizes obtained with the wild-type gene 4 protein. Data were from at least duplicate experiments.

<table>
<thead>
<tr>
<th>Gene 4 protein</th>
<th>Efficiency of plating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>1</td>
</tr>
<tr>
<td>H1</td>
<td>0.03</td>
</tr>
<tr>
<td>H2</td>
<td>2 × 10⁻⁷</td>
</tr>
<tr>
<td>H3</td>
<td>0.03</td>
</tr>
<tr>
<td>H4</td>
<td>0.06</td>
</tr>
<tr>
<td>H5</td>
<td>1 × 10⁻⁴</td>
</tr>
</tbody>
</table>

**TABLE 1**

Ability of gene 4 hybrid proteins to complement growth of T7 phage lacking gene 4

The T7 phage lacking gene 4 (T7Δ4) was plated on E. coli DH5α containing a plasmid expressing each of the recombinant genes 4 described in Fig. 1 (H1, H2, H3, H4 and H5). The plaques were counted and complementation ability of gene 4 protein was expressed as relative plating efficiency normalized against the value obtained with the wild-type gene 4 protein. Data were from at least duplicate experiments.

FIGURE 2. Template-directed oligoribonucleotide synthesis by hybrid gene 4 primases. A, gene 4 protein (200 nm) was incubated with 36 μM 15-mer ssDNA containing a primase recognition sequence 5′-GGGTC-3′ and 0.1 mM each of ATP and [α-³²P]CTP (0.1 μCi/μl) in a reaction buffer (40 mM Tris-HCl, pH 7.5, 10 mM DTT, 10 mM MgCl₂, and 50 mM potassium glutamate). After incubation for 30 min at 37°C, reaction products were separated on a 25% denaturing polyacrylamide gel and detected by autoradiography. Major reaction products were indicated at the left side of the gel. B, gene 4 protein (200 nm) was incubated with 10 nm M13 ssDNA, 0.1 mM each of all four NTPs, and [α-³²P]CTP (0.1 μCi/μl) in the buffer as described in A. dTTP was supplied at a final concentration of 0.5 mM to facilitate translocation of gene 4 protein on the ssDNA template.

FIGURE 3. DNA synthesis by hybrid gene 4 primases.

A, gene 4 protein (200 nm) was incubated with 30 μM ssDNA containing a primase recognition sequence 5′-GGGTC-3′ and 0.1 mM each of ATP and [α-³²P]CTP (0.1 μCi/μl) in a reaction buffer (40 mM Tris-HCl, pH 7.5, 10 mM DTT, 10 mM MgCl₂, and 50 mM potassium glutamate). After incubation for 30 min at 37°C, reaction products were separated on a 25% denaturing polyacrylamide gel and detected by autoradiography. Major reaction products were indicated at the left side of the gel. B, gene 4 protein (200 nm) was incubated with 10 nm M13 ssDNA, 0.1 mM each of all four NTPs, and [α-³²P]CTP (0.1 μCi/μl) in the buffer as described in A. dTTP was supplied at a final concentration of 0.5 mM to facilitate translocation of gene 4 protein on the ssDNA template.

**Zinc-binding Domain of T7 DNA Primase**

Complementation of T7Δ4 Growth—The effect of the alterations on the in vivo function of gene 4 protein was examined by a phage growth complementation assay. In this assay, the viability of T7 phage lacking the entire gene 4 (T7Δ4) depends on the expression of gene 4 from a plasmid harbored in E. coli. As shown in Table 1, compared with wild-type protein, all hybrid proteins exhibit a 10² to 10⁷-fold decrease in their ability to support growth of T7Δ4. Plaque sizes obtained with the hybrid proteins were also significantly smaller than those observed with the wild-type gene 4 protein (data not shown). In particular, alterations of a segment containing Cys-17 and Cys-20 that coordinate the zinc (H2) and a segment neighboring another zinc-coordinating cysteine, Cys-36 (H5), result in a more severe defect than the other alterations (H1, H3, and H4).

Oligoribonucleotide Synthesis Activity—To determine defects that lead to the loss of in vivo function, the hybrid proteins were purified and examined in biochemical assays. In all assays, the 56-kDa gene 4 protein lacking ZBD was included for comparison. At first, the ability of gene 4 protein to catalyzate template-directed oligoribonucleotide synthesis was examined using a short ssDNA template containing the sequence 5′-GGGTC-3′. All hybrid proteins as well as the 56-kDa gene 4 protein fail to synthesize oligoribonucleotide (Fig. 2A). Because alterations of the ZBD could change specificity of the primase recognition site, we used M13 ssDNA that should contain all possible recognition sequences for primer synthesis. Again, none of the hybrid gene 4 proteins synthesize any oligoribonucleotide on M13 ssDNA (Fig. 2B). The multiple products observed with wild-type protein reflect the use of minor recognition sites within the M13 DNA sequence. Use of radiolabeled NTPs other than CTP did also not yield any oligoribonucleotides synthesized by the hybrid proteins (data not shown).

The loss of ability to catalyze the template-directed synthesis of oligoribonucleotides could result from defective function of either the ZBD or RPD (10). However, all the hybrid DNA primases, like the 56-kDa gene 4 protein, catalyze the template-independent synthesis of random diribonucleotides (supplemental Fig. S1). Therefore, the inability to mediate template-directed oligoribonucleotide synthesis appears to arise from a defect in the ZBD.

RNA-primed DNA Synthesis—Oligoribonucleotides synthesized by DNA primase are delivered to DNA polymerase to initiate lagging strand synthesis. This function of the primase was examined in an RNA-primed DNA synthesis assay under two different conditions. In the first experiment, a mixture of all four NTPs was supplied as precursors for oligoribonucleotides. Wild-type gene 4 protein supports DNA synthesis by T7 DNA polymerase on M13 ssDNA by providing oligoribonucleotides that serve as primers to initiate DNA synthesis (Fig. 3A). Not surprisingly, in view of their inability to synthesize template-directed oligoribonucleotide synthesis, no DNA synthesis was observed with any of the hybrid gene 4 proteins (Fig. 3A). The lack of DNA synthesis confirms the absence of oligoribonucleo-
that are unable to deliver preformed oligoribonucleotides to T7 DNA polymerase have normal amounts of zinc. This finding suggests that the modifications in these hybrid primases do not completely disrupt conformation of the ZBD, yet they destroy the ability of the ZBD to properly interact with the template and RPD.

**Activity of Helicase Domain**—The helicase domain of gene 4 protein is located in the C-terminal half of the polypeptide. The helicase domain assembles on ssDNA as a hexamer and uses the energy derived from the hydrolysis of dTTP to translocate unidirectionally on the ssDNA. All the hybrid gene 4 proteins catalyze the ssDNA-dependent hydrolysis of dTTP in a concentration-dependent manner, confirming that the helicase domain is not affected by alterations in the ZBD (supplemental Fig. S2).

**Construction of Library and Selection of Functional Gene 4 Proteins**—The results from the homolog-scanning analysis described above show that all alterations of the ZBD as small as five amino acid residues lead to a defect in template-directed synthesis and delivery of oligoribonucleotide. The complementation analysis of the hybrid primases and the RNA-primed DNA synthesis assays indicate that alterations in the H2 segment spanning Pro-16 to Cys-20 result in the most severe defect in primer function relative to the other hybrid primases. The two cysteine residues, Cys-17 and Cys-20, in this portion of the ZBD are involved in coordination of a zinc ion and are strictly conserved in prokaryotic primases. Most bacteria have the same amino acid sequence, Cys-Pro-Phe-His, in this segment, whereas bacteriophages have various residues between the two cysteines, Cys-Xaa-Xaa-Cys (2). Therefore, this portion of the ZBD might be one of the most important regions for primer function, including template-directed oligoribonucleotide synthesis and its delivery to DNA polymerase.

To investigate the role of this region in primer function in more detail, we created a plasmid library that consists of random amino acid substitutions at the three variable positions (residue numbers 16, 18, and 19) by introducing mixed codons. *E. coli* cells containing the plasmid library were infected with a T7 phage lacking a portion of gene 4 coding the first 80 amino acids of the gene 4 protein (T7Δ4-1). Homologous recombination between the plasmid and T7Δ4-1 generated T7 phages that have various amino acid compositions at positions 16, 18, and 19 in gene 4 protein. T7 phages that acquire a functional gene 4 from recombination are viable and can be selectively isolated in the absence of gene 4 expression from a plasmid. When an *E. coli* pool containing the plasmid library (4 × 10⁵ cells) was infected with T7Δ4-1 at a multiplicity of infection of 0.05, 42 plaque plaques emerged. In a control in which the same number of bacterial cells containing wild-type gene 4-expressing plasmid was used, 100 plaques were observed, reflecting a recombination efficiency of 0.5%. Amino acid sequence analysis of recombiant phages reveals that 10 of 42 plages encode wild-type gene 4 protein (Table 2). The other 32 sequences were aligned based on characteristics of side chain of the residue at position 18 because this position appears to play the most critical role in the function of ZBD as discussed later.

**Analysis of Altered Gene 4 Primases Containing Substitution at Positions 16, 18, and 19**—Among the functional gene 4 proteins selected from the homologous recombination described...
above, we chose six representative proteins based on the amino acid at each position and purified them for biochemical analysis. Most of the altered proteins differ from one another by a single residue to facilitate analysis of the role. We also included gp4-C20H in which histidine is substituted for cysteine at position 20 to examine its effect on primase; most bacterial primases contain a histidine at this position.

**Complementation of T7Δ4 Growth and Zinc Content**—Examination of their ability to complement growth of T7Δ4 phage lacking gene 4 confirms that the altered proteins behave similar to wild-type protein except for gp4-H2-KHR and gp4-C20H. gp4-H2-KHR with Lys, His, and Arg at positions 16, 18, and 19, respectively, exhibits a 5-fold reduction in the complementation assay (Table 2). gp4-C20H with histidine replacing cysteine at position 20 is ~30-fold less efficient than wild-type protein in the complementation assay. All of the altered proteins have a zinc content equivalent to wild-type protein except for gp4-C20H, which has essentially no zinc (supplemental Table S1). Taken together with the phage complementation data, it appears that histidine cannot replace the cysteine at position 20 in the T7 primase. This finding is relevant to the defect observed with the hybrid primase gp4-H2.

**Oligoribonucleotide Synthesis Activity**—Because the altered primases were selected from a positive screening for functional gene 4 proteins, they must have at least a minimal primase activity. The ability of the primases to synthesize oligoribonucleotides was examined using a 60-mer ssDNA containing a primase recognition site (5'-TGGTC-3'). As expected, all the altered primases catalyze the synthesis of ACCA (Fig. 4A). However, when oligoribonucleotide synthesis is measured in an assay with varying amounts of template, some proteins were observed to synthesize oligoribonucleotides more efficiently than does the wild-type protein. For example, gp4-H2-KHR, RSG, and SSG have apparent binding constants 10–30-fold lower than the wild-type (Fig. 4, B and C). The other proteins exhibit similar or slightly higher values than wild-type protein. A comparison of the amino acid sequence of the altered proteins reveals an approximate correlation between the net charge of the residues in this segment and the binding constant. The more positively charged residues present in the segment, the better binding to the ssDNA template and the more oligoribonucleotides produced at low concentrations of template. For example, gp4-H2-KHR containing the most basic residues has the highest DNA binding affinity, whereas gp4-H2-ENE containing the most acidic residues shows the lowest affinity (Fig. 4C).

**Sequence Recognition**—To determine whether alterations in the H2 segment lead to any modification in sequence-specific recognition, we carried out oligoribonucleotide synthesis using M13 ssDNA as examined earlier for the hybrid primases (Fig. 5). No significant difference in sequence recognition by any of

### Table 2

<table>
<thead>
<tr>
<th>Property at position 18</th>
<th>Occurrence (42 in total)</th>
<th>Selected for analysis</th>
<th>T7Δ4 plating efficiency*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negatively charged</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Pro</td>
<td>10</td>
<td>WT 1</td>
</tr>
<tr>
<td>Glu</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td>Ala</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td>Arg</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td>Thr</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td>Lys</td>
<td>Lys</td>
<td>1</td>
<td>Lys 1</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>1</td>
<td>Gly 1</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>1</td>
<td>Ser 1</td>
</tr>
<tr>
<td>Thr</td>
<td>Thr</td>
<td>1</td>
<td>Thr 1</td>
</tr>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td><strong>Polar/no charge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Glu</td>
<td>1</td>
<td>H2-ENE 1</td>
</tr>
<tr>
<td>Lys</td>
<td>Lys</td>
<td>1</td>
<td>Lys 1</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>1</td>
<td>Gly 1</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>1</td>
<td>Ser 1</td>
</tr>
<tr>
<td>Thr</td>
<td>Thr</td>
<td>1</td>
<td>Thr 1</td>
</tr>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>1</td>
<td>Asp 1</td>
</tr>
<tr>
<td><strong>Aliphatic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td>1</td>
<td>Asn 1</td>
</tr>
<tr>
<td>Glu</td>
<td>Glu</td>
<td>1</td>
<td>Glu 1</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>2</td>
<td>Ala 2</td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td>1</td>
<td>Val 1</td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>1</td>
<td>Ser 1</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>1</td>
<td>Gly 1</td>
</tr>
<tr>
<td>Met</td>
<td>Met</td>
<td>1</td>
<td>Met 1</td>
</tr>
<tr>
<td>Asn</td>
<td>Asn</td>
<td>1</td>
<td>Asn 1</td>
</tr>
<tr>
<td><strong>Positively charged</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Lys</td>
<td>2</td>
<td>H2-KHR 0.18</td>
</tr>
<tr>
<td>Glu</td>
<td>His</td>
<td>1</td>
<td>C20H 0.03</td>
</tr>
</tbody>
</table>

*The relative efficiency of plating was normalized against the value obtained with the wild-type gene 4 protein.
the altered primase was found as compared with the wild-type primase.

**Interaction with Long DNA Templates**—In the above experiments, a trend was observed in the amount of oligoribonucleotides synthesized by the altered proteins. Altered primases that show better binding to the short DNA template compared with wild-type primase (Fig. 4) synthesize less oligoribonucleotides on the longer M13 template. One might expect the primase fragment containing the substitution of ENE (primase-H2-ENE) produces fewer oligoribonucleotides than does the wild-type primase fragment. The results confirm that alterations in template binding seen with full-length proteins result from changes within the primase domain.

In the above experiments, the activity of the primases is inferred to be a result of the binding of the primase to its recognition sequence, and the affinity for the primases was calculated based on the kinetic results. DNA binding affinity of the primases was also examined directly using surface plasmon resonance (Fig. 7). In this assay, a 25-mer ssDNA template either containing (5’-TGTTG-3’) or lacking (5’-TGTTG-3’) a primase recognition site was immobilized on a chip through biotin-streptavidin interaction, and the primase was flowed over the chip (25). The sequence 5’-TGTTG-3’ lacks the essential cryptic cytosine; the cytosine is replaced with guanosine. As reported previously (25), the wild-type primase binds to a tem-

**FIGURE 4. Oligoribonucleotide synthesis by gene 4 primases selected from functional recombinant phage library.** A, gel analysis of oligoribonucleotides synthesized on a 60-mer template. The experiment is similar to the one described in Fig. 2. Gene 4 protein (200 nM) was incubated with 10 nM M13 ssDNA, 0.1 mM each of all four NTPs, and [α-32P]CTP. Major reaction products are indicated at the left side of the gel. B, template-directed synthesis of oligoribonucleotides. Reactions similar to those described in Fig. 2. dTTP was supplied at a final concentration of 0.5 mM. Major products, pppACCA and pppACAC, are indicated at the left side of the gel. C, apparent binding of gene 4 protein to the 60-mer template was calculated from the data obtained in B. $V_{max}$ and $K_{DNA}$ values were calculated by the nonlinear regression method.
plate containing a primase recognition site slightly better than a template lacking the recognition sequence in the absence of NTP, although the binding to either is minimal. However, sequence-specific recognition of a template containing the primase recognition site is observed in the presence of ATP and CTP that are required for oligoribonucleotide synthesis (Fig. 7A). The primase fragment containing the alteration of KHR differs in two respects to wild-type primase fragment (Fig. 7B). First, primase-H2-KHR exhibits significantly increased binding to the template containing the recognition sequence (about 700 response units) even in the absence of ATP and CTP. Second, the binding is less dependent on the presence of the recognition sequence. The altered primase binds almost equally to templates that either contain or lack the recognition sequence. Regardless of the sequence in the template, the presence of ATP and CTP further enhances binding. In contrast, primase-H2-ENE does not bind to ssDNA in a sequence-dependent manner, and its binding is not affected by the presence of ATP and CTP (Fig. 7C). Despite the nonspecific binding to the template lacking a primase recognition site, no primer synthesis by the altered proteins was observed without a primase recognition sequence (data not shown).

**DISCUSSION**

Zinc-binding motifs play diverse roles in biological events from gene regulation to protein folding (30). Many zinc-binding motifs fulfill their functions through binding to specific sequences of DNA. Most DNA primases contain zinc-binding motifs. A small catalytic subunit of the two-component primases in archaea and eukaryotes contains a zinc knuckle consisting of Cys-His-Cys2 or Cys2-His2 (31). Despite structural information from several primases, the precise role of the zinc-binding motifs in primer synthesis remains unclear (32). Different from the primases in archaea and eukaryotes, primases encoded by bacteria and bacteriophages catalyze the template-directed synthesis of oligoribonucleotides from specific sequences in the DNA. Their modular structures reveal a ZBD composed of Cys-His-Cys2 or Cys4 separated from the RPD (6,
Zinc-binding Domain of T7 DNA Primase

33). Because the RPD alone exhibits a catalytic activity independent of template sequence, it has been speculated that the ZBD of the primase is responsible for the sequence-specific recognition of the template (13). Interestingly, each of the prokaryotic primases that share overall structural similarity recognizes distinct sequences in the template. In this report, we have investigated the molecular basis of template recognition by modifying the ZBD of T7 DNA primase through homolog-scanning mutagenesis.

Role of Zinc-binding Domain in Sequence Recognition—Most substitutions (H1, H3, and H4) examined in the homolog-scanning analysis result in loss of the zinc and thus disrupt the integrity of the ZBD. These DNA primases in which the ZBD has been distorted behave similar to the 56-kDa gene 4 protein lacking the entire ZBD. It appears that the ZBD has a unique conformation that is stabilized through elaborate interactions of the residues, particularly those that retain the zinc. A similar disruption of the ZBD was previously observed by the single amino acid replacement of His-33 with alanine (28). However, replacement of a region coordinating the zinc ion with a homologous region (H2 and H5) does not result in loss of the metal ion. This observation suggests that this portion of the primase has evolved to maintain the metal and thus structural integrity of ZBD. However, maintaining the zinc ion is not sufficient for sequence recognition by the primase. Altered proteins containing these substitutions fail to catalyze template-directed synthesis of oligoribonucleotides or to deliver preformed oligoribonucleotides to DNA polymerase. These results indicate that a suitable conformation of the ZBD is required for interactions with the DNA template and RPD for proper function.

Data from a previous structural study along with data from an alanine-scanning mutagenesis enabled the identification of residues critical for primase function on one face of the ZBD (34). This face of the ZBD might interact with RPD directly or through template DNA. However, no structural change from Asp-31 and His-33 upon the addition of DNA was detected. These residues were previously shown to be critical for sequence specificity, suggesting that more than a single residue in the ZBD determines the specificity (15, 34).

Requirement of Amino Acid Residue in the H2 Region for DNA Binding—The small number of recombinant phages arising from the screen for functional gene 4 limits our interpretation for amino acid requirement in the H2 region of the ZBD of T7 primase. Nonetheless, a direct comparison of the altered proteins does provide insight into preference at specific positions. For example, gp4-H2-SSG containing serine at position 18 (underlined) exhibits higher binding to DNA than does gp4-H2-SDG harboring aspartic acid at the same position. Similarly, a comparison of kinetic parameters between gp4-H2-KHR and gp4-H2-KER indicates positively charged histidine is better than negatively charged glutamic acid for template binding. The equivalent activity of gp4-H2-SDG and wild-type primase, both of which contain the same aspartic acid at this position, implicates this position as an important factor in DNA binding. Therefore, a trend emerges where a positively charged residue at position 18 leads to better binding with DNA.

In contrast, the residue at position 16 (underlined) does not make a significant difference in DNA binding as observed with gp4-H2-SSG and gp4-H2-RSG. The preference for a positively charged residue at position 18 can be further extended to the overall charge arising from residues at positions 16, 18, and 19 in the H2 region. Alignment of primases reveals a positive correlation between the number of positively charged residues in this region and DNA binding affinity. gp4-H2-KHR with positively charged residues at the three positions exhibits much higher DNA binding affinity than does gp4-H2-ENE containing the mostly negatively charged residues. Although structural analysis of the primase-template interaction will ultimately resolve the precise contacts necessary for DNA binding, it is tempting to propose that this region of ZBD is in proximity to the template as augmented in gp4-H2-KHR. The same pattern of oligoribonucleotide synthesis by proteins with alterations in this region also suggests that this region is not responsible for recognition of specific sequences in the template. Although we observed predominantly changes in the charged (47%) and aliphatic (50%) residues at positions 16 and 19, respectively, no significant tendency was found, implying that it could be biased by the small number of phages in the library.

Single substitution of histidine for cysteine at position 20, a distinctive difference in the ZBDs of bacteria and bacteriophages, results in the loss of the zinc in the T7 primase. This result is surprising because histidine was anticipated to coordinate the metal ion within a flexible loop as in the bacterial ZBD. The incompatibility of histidine suggests that the architecture of the T7 ZBD restricts the position for cysteine.

Implication on Interactions between ZBD and RPD—Several lines of data indicate that the altered ZBDs interact with the RPD in a manner different from that observed with the wild-type ZBD. First, both primase-H2-KHR and primase-H2-ENE do not bind to DNA in a sequence-specific manner as measured by surface plasmon resonance. Because neither the ZBD nor the RPD alone can support DNA binding (25), the changes in DNA binding must result from interactions between the ZBD and RPD. Second, the requirement for ATP and CTP binding in RPD for efficient binding is altered by changes in segment H2 of the ZBD. The binding of primase-H2-KHR to its recognition sequence is enhanced by the presence of ATP and CTP, but primase-H2-ENE is not. Third, the ZBD containing H2-KHR alone does not form a stable structure. Analysis of ZBD-H2-KHR by circular dichroism spectroscopy reveals a disordered conformation as indicated by the disappearance of characteristic peaks shown in the wild-type, and the altered ZBD forms a dimer in solution (supplemental Fig. S5). Indeed, no zinc was detected by inductively coupled plasma-mass spectrometry. The absence of zinc raises the possibility that ZBD-H2-KHR alone is not stable due to the lack of critical contacts with the RPD. These data suggest that changes in DNA binding by alterations in the ZBD are mediated through interactions with the RPD.

Thus, it seems that a more likely approach to change sequence specificity would be the introduction of amino acid changes in the RPD as well as the ZBD. A comparison of several bacterial primases whose template specificities were recently identified supports this argument. G. stearothermophilus primases and the primase of Staphylococcus aureus synthesize primers from 5′-(C/T)AT-3′, whereas the primase of Aquifex
Zinc-binding Domain of T7 DNA Primase

initiation and extension by T7 DNA primase. EMBO J. 25, 2199–2208
Zinc-binding Domain of T7 DNA Primase


