Characterization of elongating T7 and SP6 RNA polymerases and their response to a roadblock generated by a site-specific DNA binding protein

Pamela A. Pavco* and Deborah A. Steege*
Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

Received June 6, 1991; Revised and Accepted August 2, 1991

ABSTRACT
As a means of generating homogeneous populations of elongation complexes with the RNA polymerases encoded by phages T7 and SP6, transcription has been carried out in vitro on templates associated with the Gin-111 mutant of EcoRI endonuclease. The Gin-111 protein, as a result of a single amino acid substitution at position 111, lacks cleavage function yet shows higher than wild-type affinity for the EcoRI recognition sequence GAATTC. On a series of linear and circular templates associated with Gin-111 protein, blockage of the phage RNA polymerase elongation complex is observed. The 3’ endpoint of the major blocked-length RNA species, just 3 bp upstream from the GAATTC, reveals an extremely close approach of polymerase’s leading edge to essential contacts between Gin-111 protein and its binding site. In contrast to E. coli RNA polymerase, which is blocked stably and quantitatively by Gin-111 protein (Pavco, P.A. and Steege, D.A. (1990) J. Biol. Chem. 265, 9960–9969), the phage polymerases show substantial levels of readthrough transcription beyond the protein block.

INTRODUCTION
The RNA polymerases specified by bacteriophages T7, T3 and SP6 are small, single-subunit enzymes of Mr ~100,000 (1). They are highly specific for their cognate promoters and terminators and are able to maintain rapid rates of transcription in vitro. The elongation rate for T7 RNA polymerase, for example, has been measured at ~230 nucleotides per second (2). The phage RNA polymerases are also very efficient at recycling in vitro. Each active enzyme can carry out multiple rounds of initiation, elongation and termination (3). As a result of these favorable characteristics, T7 and SP6 RNA polymerases are widely used for large-scale production of specific RNAs.

Although the relatively weak binding of the phage RNA polymerases to promoter sequences initially hampered biochemical analysis of individual steps in transcription, the nature of the complexes they form with the DNA template during initiation and immediately thereafter has been established. T7 RNA polymerase binds to its cognate promoters with association constants ≤ 10^7 M^-1 (3,4). From footprinting analysis (4–6), the complex with the promoter sequence appears to extend over about 19 bp. The Fe(II)-EDTA footprint extends from positions -17 to +2 and gives a pattern of protection suggesting that polymerase binds to regions primarily on one face of the DNA helix (6). Between positions -5 and -12, specific contacts inferred from methylation and ethylation interference data are important for polymerase binding (7). The region from -6 to +2 is accessible for nuclease cleavage, indicating that it may be transiently single-stranded (6,8,9). Also suggesting that the helix in this region is disrupted during initiation is the fact that removal of bases from either strand between positions -5 and +3 results in enhanced T7 polymerase binding (7). As is the case with other RNA polymerases, T7 RNA polymerase is subject to rounds of abortive cycling after initiation until RNA chain lengths of ~8 nucleotides have been attained (10).

Less is known about the phage RNA polymerase complex once it has moved fully into the elongation phase. The rapid RNA chain elongation rate and tendency not to pause on the template have made it difficult to isolate homogeneous populations of elongation complexes positioned at a specific site. In only two cases have T7 RNA polymerase elongation complexes been examined. By omitting UTP, Ikeda and Richardson (5) were able to use methidiumpropyl-EDTA-Fe(II) to study a ternary complex stalled at position +15; by arresting polymerase at a psoralen crosslink, Shi et al. (11) used DNase I to probe a complex arrested at +36. The sizes of the footprints obtained with these reagents, 24 and 20 bp, respectively, are not markedly different from that of the promoter complex. From this it has been suggested that no dramatic change in enzyme-DNA contacts occurs during the transition from initiation to elongation (6).

Previous work in our laboratory (12) has led to the development of an alternative method that permits study of homogeneous populations of elongation complexes. The basis of our approach is to block forward progress of transcribing RNA polymerase by means of a protein bound tightly to a specific site on the DNA template, and then to characterize the blocked ternary complex.

* To whom correspondence should be addressed
+ Present address: United States Biochemical Corporation, 26111 Miles Road, Cleveland, OH 44128, USA
The sequence-specific DNA binding proteins used are two EcoRI endonuclease mutants isolated by Modrich and co-workers (13,14). These proteins lack appreciable cleavage activity as a result of single amino acid substitutions for glutamic acid at position 111. The Gly-111 protein contains glycine at this position, and the Gln-111 protein, glutamine. Both proteins retain high affinity for the wild-type EcoRI recognition sequence GAATTC. The Gly-111 derivative shows near wild-type binding parameters. The Gln-111 derivative actually has a specific affinity for GAATTC 1000-2000 times greater than the wild-type enzyme due to its decreased dissociation rate. These EcoRI derivatives can therefore be used as high-affinity DNA binding proteins specific for the EcoRI recognition sequence, which is easily engineered at any position in a DNA template.

When E. coli RNA polymerase is used in in vitro transcription on templates associated with Gln-111 protein (12), complete blockage of elongation is observed. Blockage is efficient with Gly-111 protein as well. Blockage by this means generates a population of ternary complexes positioned at a specific location on the DNA template. The 3' end of the nascent, blocked-length RNA is 14 nucleotides (nt) upstream of the GAATTC sequence and, based on exonuclease III footprinting data obtained for both Gln-111 protein alone and the blocked polymerase complex, is ~7 nt from the leading edge of the polymerase. The blocked ternary complexes remain stable over time in an active form. They are capable of either resuming elongation once the blocking protein is displaced from the DNA template by an increase in ionic strength, or interacting with E. coli termination factor rho and undergoing release.

It was of interest to determine if the blocking strategy described could be applied to study elongation complexes formed by the phage RNA polymerases. In this paper, binding of Gln-111 protein to DNA templates containing a unique EcoRI recognition sequence is examined under the standard conditions employed for transcription in vitro. Gln-111 protein is then tested for its ability to block elongating T7 and SP6 RNA polymerase. Using a series of templates in which the distance from the transcription start point to the EcoRI site is varied from 35-58 bp, the positions to which the 3' ends of the nascent, blocked-length RNAs map in the DNA template are determined.

**Enzymes and biochemicals**

Carrier-free [32P]orthophosphoric acid for preparation of [γ-32P]GTP (15) and [γ-32P]ATP for preparation of [5'-32P]pCp (16) were purchased from ICN; [α-32P]CTP was obtained from DuPont-NEN. E. coli DNA polymerase I Klenow fragment, T4 DNA ligase, alkaline phosphatase, mung bean nuclease, restriction enzymes, unlabeled ribonucleoside triphosphates (rNTPs) and RNase/DNase-free bovine serum albumin were purchased from Pharmacia LKB Biotechnology, Inc. RNA ligase was obtained from New England Biolabs. RNases T₁ and T₂ were purchased from Calbiochem.

T7 RNA polymerase was a gift of J. J. Dunn (Brookhaven National Laboratories). SP6 RNA polymerase was from Promega Corp. EcoRI endonuclease and Gln-111 protein (14) were provided by P. Modrich (Duke University). Concentrations for EcoRI endonuclease and Gln-111 protein are reported as dimer equivalents, assuming a dimer of identical subunits as the active DNA binding species (14).

**Transcription templates**

Plasmids in circular or linear form and specific restriction fragments were used as DNA templates for transcription. Plasmids pSP72, pSP73 and pGEM-2 (Promega Corp) bear both SP6 and T7 promoters and contain a single EcoRI site. Plasmid pGEM-2, obtained from T. Hsieh (Duke University), was used as the parent for a series of derivatives that vary from each other only by the distance between the T7 transcription start site and the EcoRI site. These plasmids were made by digesting pGEM-2 with HindIII alone to linearize the plasmid or with HindIII and SalI to remove a 15-bp fragment. The DNA was then treated with either the Klenow fragment of DNA polymerase I or mung bean nuclease, and the resultant blunt-ended plasmids were recircularized using T4 DNA ligase. Plasmids were introduced into strain DS70 (F- Δlac trpR λ+) by transformation (17) and purified by equilibrium centrifugation in CsCl-ethidium bromide gradients. Each template construct was verified by DNA sequencing (18,19). This analysis showed that the pGEM-2 isolate used, and therefore all subsequent derivatives, differed from the published sequence. Rather than GG at positions encoding nucleotides 9 and 10 in the T7 transcript, there is a single C. Consequently, transcript lengths reported here are 1 nt less than lengths predicted from the published sequence.

**In vitro transcription**

Transcription reactions (final vol 15 μl) using T7 RNA polymerase were carried out essentially as described for E. coli RNA polymerase (12), except that the transcription buffer was 20 mM sodium phosphate, pH 7.5, 5 mM dithiothreitol, 5 mM NaCl, 50 μg/ml bovine serum albumin. Where indicated, SP6 RNA polymerase was used in place of T7 RNA polymerase with no additional changes. Transcription reactions contained 10–500 μM each rNTP and 10 μCi [α-32P]CTP. The DNA template (4 nM) was preincubated for 5 min at 37°C in transcription buffer prior to the addition of Gln-111 protein to a final concentration of 4–80 nM. Gln-111 protein was diluted immediately prior to use with ice-cold 20 mM potassium phosphate, pH 7.4, 0.1 M NaCl, 0.5 mM dithiothreitol, 0.2 mM EDTA, 10% glycerol, 50 μg/ml bovine serum albumin. After a 10-min incubation to allow binding of Gln-111 protein, T7 RNA polymerase (diluted in 20 mM sodium phosphate, pH 7.5, 0.1 M dithiothreitol, 0.1 mM EDTA, 10% glycerol) was added to ~40 nM. Transcription was initiated 10 min later by the addition of MgCl₂ to 5 mM and stopped after 10 min by the addition of 2 μl 0.2 M EDTA and an equal vol (17 μl) of buffered formamide containing xylene cyanol FF (20). The samples were heated to 90°C for 3 min, chilled quickly, and analyzed on sequencing gels containing 8% polyacrylamide gel (20). Gels were exposed to Kodak XAR-5 film at -70°C in the presence of a Dupont Lightening Plus intensifying screen. Quantitation of autoradiograms was accomplished using the Photometrics Star I CCD camera in conjunction with the Image digital processing program.

To generate RNA size markers, full-length transcripts labeled at the 5' end using [γ-32P]ATP as the radioactive rNTP were prepared in a standard transcription reaction scaled up 10-fold. Following isolation from a denaturing polyacrylamide gel, the end-labeled transcript was subjected to partial digestion with RNase T₁ under denaturing conditions (21). An RNA size ladder was generated by alkaline hydrolysis of the end-labeled RNA (22).
Analysis of protein-DNA complexes

Binding of Gln-111 protein to a 32P-labeled template containing a single EcoRI site was analyzed by nondenaturing polyacrylamide gel electrophoresis (22,23) as described previously (12), except that transcription buffers and diluents were as given above. The 32P-labeled DNA template used was the 231-bp Ndel/HpaI fragment of pSP73. It was prepared by treating pSP73 DNA digested with HpaI with T4 polynucleotide kinase in the presence of [γ-32P]ATP, digesting the 32P-labeled DNA with Ndel and purifying the appropriate 32P-labeled fragment from a polyacrylamide gel. The standard protocol for in vitro transcription was followed, except that T7 RNA polymerase dilution buffer was substituted for T7 RNA polymerase. Ten min after the addition of MgCl2, 5 μl of 20% glycerol, 10 mM Tris-HCl, pH 7.6, 1 mM EDTA was added with gentle mixing. A portion of the sample was immediately applied to a 5% polyacrylamide gel (12).

Challenge with EcoRI endonuclease as an assay of EcoRI site occupation by Gln-111 protein

The 32P-labeled DNA used in challenge experiments was the 179-bp PvuII/Rsal fragment from pGEM-2. It was prepared by digestion of pGEM-2 DNA with PvuII, treatment with T4 polynucleotide kinase in the presence of [γ-32P]ATP to label the PvuII end, and final digestion with Rsal. The resulting DNA was extracted with phenol and precipitated twice from ethanol. Free DNA or protein-DNA complexes prepared by incubating the DNA with a 20-fold excess of Gln-111 protein were applied to a nondenaturing polyacrylamide gel. Each lane contained 4 nM DNA, the amount of DNA present in a standard transcription reaction. Regions of the gel containing the free DNA or the 1:1 DNA-protein complex were located by autoradiography and excised. Gel slices from appropriate lanes were combined and placed in buffer (30 μl per slice) that approximated standard T7 transcription conditions: 23 mM sodium phosphate, pH 7.5, 5.4 mM dithiothreitol, 12.5 mM NaCl, 540 μM each rNTP, 14 μM EDTA, 0.7% glycerol, 93 μg/ml bovine serum albumin. During a 30-min incubation at 37°C, 10–20% of the DNA or protein-DNA complex eluted from the polyacrylamide. To constitute each reaction for the challenge assay, 15 μl of the eluate was supplemented with 1 μl [α-32P]CTP (10 μCi), and 1 μl T7 RNA polymerase was added to a final concentration of 200 nM. Samples were incubated for 2 min at 37°C, and transcription was initiated by the addition of MgCl2 to 5 mM. After 2 min, wild-type EcoRI endonuclease was added in a vol of 1 μl to a final concentration of 20 nM. In control reactions, the appropriate dilution buffers replaced T7 RNA polymerase, Gln-111 protein or EcoRI endonuclease. The dilution buffer for Gln-111 protein was also used for EcoRI. Portions (8 μl) of each sample were removed after a further 5 or 20 min and mixed with 2 μl 0.2 M EDTA. Following ethanol precipitation in the presence of 10 μg tRNA, the samples were resuspended in 10 μl of buffered formamide containing xylene cyanol FF (20) and analyzed on 8% sequencing gels.

RESULTS

Binding of Gln-111 protein to the DNA template

The ability of the Gln-111 derivative of EcoRI endonuclease to bind DNA in vitro under standard T7 transcription conditions was assessed by a mobility shift assay on nondenaturing polyacrylamide gels. Protein-DNA complexes were formed by incubating varied amounts of Gln-111 protein with a constant amount of 32P-end-labeled template containing a single EcoRI recognition sequence. As shown in Fig. 1, an increase in the molar ratio of Gln-111 protein to DNA results in an increase in the amount of a distinct complex migrating at the position of the DNA template bound by a single dimer of Gln-111 protein (1:1). At a 20-fold molar excess of protein to DNA, virtually all of the DNA template is associated with Gln-111 protein. As indicated by the presence of DNA associated with more than one Gln-111 protein dimer, some nonspecific binding to the template DNA is evident as well.

Although quantitative binding of the DNA template is approached at a 20-fold molar excess of Gln-111 protein, the low ionic strength conditions used are not optimal for Gln-111 protein. It is sensitive to inactivation when diluted into buffers of low ionic strength and exhibits more nonspecific binding under these conditions. Our previous studies of polymerase blockage in the E. coli transcription system were carried out at higher ionic strength, and quantitative binding was achieved at the much lower molar ratio of Gln-111 protein to DNA of 7.5:1 (12). In our experiments with T7 RNA polymerase, however, attempts to optimize Gln-111 protein binding conditions by increasing the ionic strength of the reaction buffer resulted in decreased transcription. Consequently, the lower ionic strength conditions were used so as not to compromise transcription.

In vitro transcription on templates associated with Gln-111 protein

To determine what effect Gln-111 protein bound to the DNA template has on elongation by T7 RNA polymerase, transcription was carried out in the presence of [α-32P]CTP and a 20-fold molar excess of Gln-111 protein. In this and all subsequent experiments, the transcription reactions permitted multiple initiation events on the DNA template, because there did not prove to be an effective way to limit T7 transcription to a single round that was compatible with the presence of Gln-111 protein. Fig. 2A shows the data for transcription on a circular template at rNTP concentrations of 500 μM. Transcription initiated at the
T7 promoter, but blocked by protein bound to the EcoRI site 59 bp downstream, will result in truncated RNAs reflecting the position of the bound protein. In the absence of Gln-111 protein, the major products are the high molecular weight RNAs visualized at the top of the lane (FL). In the presence of Gln-111 protein, shorter RNAs appear (BL). The amount of the shorter RNAs increases with increasing concentration of Gln-111 protein. By reference to RNA size standards, the most abundant short RNA species was determined to be 56 nt long, 3 nt less than the distance from the T7 transcription start to the EcoRI site. RNAs 1 and 2 nt longer are also visible. These transcripts are all of an appropriate, although maximal, length to represent the products of T7 RNA polymerase unable to elongate past the EcoRI site due to the presence of bound protein. Fig. 2B shows the products of transcription on the same template at a 50-fold lower rNTP concentration (10 μM). Under these conditions, the amount of the blocked-length species is increased, and there is less [α-32P]CTP incorporated into full-length products. This apparent decrease in readthrough transcription may not represent increased blockage efficiency, however, since the relative amounts of full-length RNA made at 500 μM and 10 μM each rNTP in the absence of Gln-111 protein (lanes 0, Fig. 2A and 2B) suggest that overall transcription may also be reduced at the lower, possibly limiting, rNTP concentration. While blockage efficiencies cannot be determined from the multiple-round transcriptions on circular templates, it is clear that even with levels

![Figure 2](image-url)

**Figure 2.** Analysis of [α-32P]CTP-labeled RNA products generated in vitro by transcription on a circular template in the presence of Gln-111 protein at 500 μM (A) and 10 μM (B) each rNTP. The template used is that described in the text and in subsequent figures as p(+4). The products of 10-min reactions were analyzed on 8% sequencing gels. Numbers at the top of each lane indicate the molar ratio of Gln-111 protein dimers to DNA template present. RNA size standards were prepared from an end-labeled 151-nt runoff RNA generated from the linear form of the same template by alkaline hydrolysis (N) or partial digestion with RNase T1 (G). The sequence of the region of interest is on the right, with the nucleotides corresponding to the EcoRI site in the template emphasized by shading. Nucleotides to which the 3' end of the blocked-length transcripts (BL) map are indicated, the largest arrow corresponding to the major blocked-length product. Full-length transcription products (FL) near the top of the gel are indicated.

of Gln-111 protein that give quantitative binding of the EcoRI site, substantial amounts of readthrough transcription beyond the protein block continue to occur.

**Analysis of EcoRI site occupation by Gln-111 protein during transcription by challenge with EcoRI endonuclease**

The next step was to determine whether readthrough transcription was occurring because Gln-111 protein was being displaced from its binding site or because Gln-111 protein was somehow being bypassed by the transcription apparatus. Since the presence of bound Gln-111 protein would be expected to protect the single GAATTC sequence in the template from digestion by wild-type EcoRI endonuclease, the two possibilities were differentiated by determining whether or not the GAATTC becomes susceptible to EcoRI cleavage during transcription. For these experiments it proved important to use a DNA template associated with a single dimer of Gln-111 protein. Accordingly, free DNA and the 1:1 complex with Gln-111 protein were excised from the appropriate regions of nondenaturing gels and eluted into a buffer containing the necessary components for transcription. The eluted samples were directly supplemented with [α-32P]CTP and incubated in the absence or presence of T7 RNA polymerase. Two min after transcription was initiated with 5 mM MgCl₂, appropriate samples were challenged with wild-type EcoRI endonuclease.

The results in Fig. 3 show that the gel-purified free DNA (left panel) is digested to completion in a 5-min incubation with EcoRI, both in the absence and presence of T7 RNA polymerase. All of the DNA once present at the position of the full-length template (DNA-FL) now appears at the position of the EcoRI-cleaved DNA (DNA-RI). When T7 RNA polymerase is present, the additional bands representing runoff RNAs from the full-length (FL) and EcoRI-cleaved DNA (RI) show that transcription is occurring over time on these templates. The isolated 1:1 complex

![Figure 3](image-url)

**Figure 3.** The effect of transcription on EcoRI site occupation by Gln-111 protein as assayed by challenge with EcoRI endonuclease. [32P]-labeled template DNA, as the free DNA and the 1:1 complex with Gln-111 protein, was excised from a nondenaturing gel and eluted into a buffer appropriate for transcription. Samples of the eluate were directly supplemented with [α-32P]CTP and brought to 37°C. Following addition of T7 RNA polymerase to 200 nM and incubation at 37°C for 2 min, MgCl₂ was added to 5 mM to initiate transcription. EcoRI endonuclease was added to 20 nM 2 min later, and the incubation was continued for 5 or 20 min. In control reactions, appropriate dilution buffers replaced T7 RNA polymerase, Gln-111 protein and EcoRI endonuclease. The labeled DNAs and RNA products were electrophoresed on an 8% sequencing gel. The full-length template (DNA-FL) was a 179-bp PvuII/Rsal fragment of pGEM-2 labeled at the PvuII end. Cleavage at the single EcoRI site results in a 91-bp labeled DNA fragment (DNA-RI). The full-length RNA (FL), blocked-length RNA (BL) and the product of runoff transcription on the template cleaved at the EcoRI site (RI) are indicated.
of template and Gln-111 protein (right panel), as expected, remains only partially digested by EcoRI in the absence of T7 RNA polymerase, and some of this cleavage presumably reflects complexes disrupted during their isolation. After a 5-min incubation with EcoRI, the amount of radioactivity at the position of the full-length DNA indicates that 85% of the template remains protected from cleavage due to the presence of bound Gln-111 protein. After a 20-min incubation, the same amount remains in the position of the full-length DNA. The band at the position of the EcoRI-cleaved DNA suggests that additional cleavage has occurred, but its overlap with a more diffuse spot of radioactivity running slightly faster overestimates the band's intensity. By contrast, if the 1:1 complex is incubated with EcoRI under identical conditions, but in the presence of T7 RNA polymerase, the amount of template at the position of the full-length DNA indicates that only 55% remains undigested after a 5-min incubation with EcoRI. This decreases further to 35% after a 20-min incubation. In the position of the EcoRI-cleaved DNA, corresponding increases are observed. Proof that transcription is occurring on these templates is again provided by the presence of full-length and blocked-length RNAs (FL and BL), as well as the RNA generated by runoff transcription at the newly-cut EcoRI site (RI). In the absence of EcoRI, the same amount of total RNA is made by the end of the 5-min incubation from the 1:1 complex as from free DNA, with 26% comprising the blocked-length species. The total amount of blocked- and full-length RNA drops to 57% in a 5-min challenge with EcoRI, in accord with the loss of the full-length template. Because full-length templates are being cleaved by EcoRI, little further increase in full-length RNA occurs over time. Consequently, after a 20-min challenge, the blocked-length species comprises 44% of the blocked- plus full-length RNA. During the same period of time, multiple rounds of transcription occur on the EcoRI-cleaved DNA.

We interpret the increased cleavage of the 1:1 complex during transcription as evidence that elongating T7 RNA polymerase is capable of displacing Gln-111 protein from the EcoRI recognition sequence and thus clearing the template for readthrough transcription. The data of Figs. 2 and 3, however, show that the presence of Gln-111 protein on the DNA template does block the T7 RNA polymerase elongation complex to some degree and leads to production of discrete blocked-length RNAs.

In vitro transcription in the presence of Gln-111 protein on templates of different lengths

Present evidence has not established whether T7 RNA polymerase movement during elongation is such that the leading edge of the polymerase remains a constant physical distance from the site within the enzyme at which nucleotides are being added to the RNA chain. If this distance is maintained by incremental movement of the polymerase after each nucleotide is added, then regardless of chain length, the 3' end of a nascent RNA in a ternary complex blocked by Gln-111 protein will always be located at the same position relative to the EcoRI recognition sequence. However, if polymerase movement is such that this is not the case, e.g. several nucleotides are added before the polymerase shifts forward on the template, blockage by Gln-111 protein at different positions from the transcription start might trap ternary complexes in distinct stages of elongation, with the site of nucleotide addition at different distances from the GAATTC.

With this as the rationale, a series of templates was constructed from plasmid pGEM-2 in which the distance from the T7 transcription start and the EcoRI site was varied over 23 bp. Within the series, the distances range from 4 bp longer to 19 bp shorter than that present in the parent plasmid. The 5' terminal sequences of the transcripts predicted for these templates are shown in Fig. 4; plasmid names denote the difference in template length relative to pGEM-2, which is designated p(0). Absolute lengths from the transcription start site to the GAATTC are 58, 54, 50, 43, and 35 bp. Even the shortest of these is of sufficient length to ensure that T7 RNA polymerase has moved fully into the elongation mode (5,6,10,11) before it encounters the blocking protein. In all templates, the promoter region and the 25 bp immediately preceding the EcoRI recognition sequence were left unchanged so that transcription levels would be the same from

![Figure 5. Analysis of the blocked-length RNA products generated from circular (A) and linear (B) DNA templates of varied length. The [α-32P]CTP-labeled products of in vitro transcription in the presence of a 20-fold molar excess of Gln-111 protein on the templates of Fig. 4 were analyzed on 8% sequencing gels. Alkaline hydrolysis products (N) separate the sets. Within each set, the individual lanes represent transcription in the presence of a high (left lane) or low (right lane) rNTP concentration. These were, for the circular template series, 500 and 10 μM each rNTP, and for the linear series, 500 and 100 μM each rNTP. Included in the outermost lanes of each panel are sequencing lanes generated with RNase T1. (G) using 5'32P-labeled RNA prepared by transcription on linearized p(4) DNA. RNA lengths are indicated to the right of each panel. The full-length runoff RNAs obtained from the linearized templates resolved as two bands on sequencing gels containing either 8.3 M urea (20) or 7 M urea and 10 M formamide. Since the sequencing lanes generated from the p(+4) RNA show that it has a unique sequence, the two bands probably represent 3' end heterogeneity (10,24-26).](image-url)
one template to another and an identical sequence context would be present in the region where polymerase encounters bound Gln-111 protein.

The products of transcription on these templates in the presence of a 20-fold molar excess of Gln-111 protein are shown in Fig. 5. The lengths of the blocked-length RNAs were determined by comparison to RNA size standards. The result was that for all templates in the series, whether in circular or linear form (Figs. 5A and 5B, respectively), the 3' ends of the blocked-length RNAs mapped an identical distance from the GAATTTC. Regardless of actual transcript length, the major RNA 3' endpoint reflects transcription up to and including the template position 3 bp upstream of the EcoRI site. RNAs 1 and 2 nt longer are present in addition to the major blocked-length species.

The transcripts produced from each template at high and low rNTP concentrations are also compared in Fig. 5. As noted previously, limiting the rNTPs results in a modest increase in the amount of the blocked-length species. For the linear templates (Fig. 5B), the amount at high rNTP concentrations ranges from 27% of the total transcripts for p(−19) to 42% for p(0). The amount at low rNTP concentrations ranges from 42% to 48% of the total transcripts. These values cannot be taken to represent the fraction of polymerases blocked by bound Gln-111 protein because the transcription reactions permit recycling of polymerases. However, they indicate that even though conditions for quantitative binding of the templates by Gln-111 protein have been used, on no template does elongation remain completely blocked.

**SP6 RNA polymerase is blocked at the same distance from template-bound Gln-111 protein**

Transcription experiments with SP6 RNA polymerase verified that it is blocked in a similar way by Gln-111 protein. A 20-fold excess of Gln-111 protein was used, and transcription was carried out on a template bearing an SP6 promoter and a single EcoRI site (pSP72). As was found for T7 RNA polymerase, the products of SP6 transcription included short RNAs that did not appear in the absence of Gln-111 protein. The 3' endpoint of the major blocked-length RNA generated by SP6 RNA polymerase was determined by comparison to the known lengths of two blocked-length T7 RNAs, as shown in Fig. 6. The length of the major blocked-length RNA generated by SP6 RNA polymerase (left lane) was determined to be 64 nt. As is the case for the corresponding species made by T7 RNA polymerase (right lanes), this 3' endpoint is 3 bp upstream of the EcoRI site in the SP6 template.

**3' end analysis of the blocked-length RNAs**

On all templates tested, the major blocked-length RNA species produced by T7 RNA polymerase in the presence of Gln-111 protein is consistently accompanied by small amounts of species that are 1–2 nt longer. Since T7 RNA polymerase has been reported to incorporate extra nucleotides not encoded by the template onto the 3' ends of runoff RNAs (24–26), it was of interest to determine whether the longer species contain random nucleotides or nucleotides encoded by the template. For these experiments transcription was carried out with unlabeled rNTPs in the presence of a 20-fold molar excess of Gln-111 protein on template p(−19). The 3' ends of the RNA products were labeled using 5'32P-pCp and T4 RNA ligase (16).

Sufficient quantities of the major blocked-length RNA and the species 1 nt longer were obtained to isolate the labeled RNAs individually from a sequencing gel and digest them to completion with RNase T2. Thin layer chromatographic analysis (27) of the labeled 3' nucleoside monophosphates obtained from the major species indicated that its 3' terminal nucleotide is C. This corresponds exactly to the 3' terminal nucleotide predicted from the template sequence by aligning the blocked-length RNA to size standards. The 3' end of the transcript 1 nt longer is U, again corresponding to the template sequence. These results confirm that the majority of ternary complexes are blocked by Gln-111 protein such that the last nucleotide incorporated into the transcript corresponds to the template sequence 3 bp upstream of the EcoRI.

![Figure 6](image-url)  
**Figure 6. 3' endpoint mapping of SP6 and T7 blocked-length RNAs.** The SP6 RNAs generated on pSP72 in the presence of a 20-fold molar excess of Gln-111 protein are shown in the leftmost lane. On this template, the distance from the SP6 transcription start site to the EcoRI site is 67 bp. Shown on the right as size standards are the corresponding blocked-length RNAs generated by transcription with T7 RNA polymerase on pSP73 and p(+4); these are 72 and 56 nt long, respectively. Generation of the alkaline hydrolysis ladder (N) was as in Fig. 2. RNA lengths are indicated on the right.
A minority of the elongation complexes can elongate 1–2 nt farther, but no evidence of random incorporation onto the transcript end by T7 RNA polymerase is seen.

DISCUSSION

In this paper we have shown that Gln-111 protein bound to the EcoRI recognition sequence on linear and circular DNA templates functions as a barrier to elongation for the RNA polymerases specified by phages T7 and SP6. However, even under conditions where Gln-111 protein is bound quantitatively to the template, substantial amounts of readthrough transcription continue to occur. As judged from the results of challenging with the wild-type EcoRI endonuclease, the readthrough transcription reflects displacement of the blocking protein from the EcoRI site. For all templates tested, the blocked-length RNAs generated in response to bound Gln-111 protein have 3' ends 3 bp upstream of the GAATTCC sequence, with minor amounts of species containing 1–2 additional template-encoded nucleotides present as well. The relative positions of these ends from the GAATTCC are the same for a template series varying the distance transcribed to the EcoRI site. From this result we suggest that by this method of looking from the nucleotide addition site to the leading edge contacting a blocking protein, we do not see evidence for nonuniform movement of the phage polymerases during elongation.

How closely the phage RNA polymerases approach bound Gln-111 protein becomes evident in Fig. 7, which shows blocked-length RNA 3' endpoints and information pertaining to interactions of T7 RNA polymerase and Gln-111 protein with DNA. Fig. 7A shows the blocked-length RNA 3' endpoints for T7 and SP6 polymerase and for comparison, the corresponding 3' endpoint for blocked E. coli elongation complexes 14 bp upstream from the GAATTCC (12). Shown for Gln-111 protein are the boundaries on each strand defined by exonuclease III digestion (12) and the Fe(II)-EDTA footprint (P. Pavco, unpublished results). Included also are data for chemical probes from studies of the wild-type EcoRI endonuclease in the absence of MgCl2 (28). Fig. 7B summarizes footprinting data for T7 promoter and elongation complexes, superimposed on our template sequence by appropriate alignment on the major blocked-length RNA 3' end. These include the Fe(II)-EDTA footprint of promoter-bound polymerase (6), the DNase footprint of an elongation complex containing a 15-nt RNA (11) and the methidiumpropyl-EDTA-Fe(II) footprint of an elongation complex containing a 36-nt RNA (5). If the Fe(II)-EDTA footprint is taken simply as visualizing a region of minimum size protected by polymerase, it is clear that polymerase's leading edge moves well into the region Gln-111 protein protects from exonuclease III digestion. It overlaps nucleotides in the EcoRI recognition sequence whose phosphates are thought to be in direct contact with bound EcoRI, but does not encroach upon bases whose N7 nitrogens are implicated in binding (28). Thus the minor blocked-length RNAs 1–2 nt longer that we observe may reflect the final limit to which the elongation complex can work into Gln-111 protein contacts with the EcoRI recognition sequence before Gln-111 protein is no longer stably bound and polymerase can resume full-length transcription. In any case, the proximity of the blocked-length RNA 3' endpoints to the binding site for Gln-111 protein suggests that the site at which nucleotides are added to the growing RNA chain lies very close to the leading edge of T7 and SP6 RNA polymerase.

On the one hand, the capacity of Gln-111 protein to only modestly block the phage RNA polymerase elongation complex was surprising in view of Gln-111 protein's high specific affinity for and long lifetime on its binding site (14). The characteristics of blockage observed here are markedly different than seen with E. coli RNA polymerase in our earlier studies (12). Gln-111 protein blocks E. coli ternary complexes completely, in a form that is stable over the 1-hr time periods tested and that gives little evidence of readthrough transcription. The blocked complexes are fully capable of resuming elongation, but only when Gln-111 protein is removed from the template by auxiliary means, e.g. via an increase in ionic strength. The Gly-111 EcoRI derivative also blocks E. coli ternary complexes efficiently, but Gln-111 protein, with its far slower dissociation rate, is the more effective block. On the other hand, the phage RNA polymerase elongation complexes, although sensitive to blockage by E. coli RNA polymerase-rifampicin open complexes (29), have been shown to displace or bypass a number of other proteins and protein complexes. These include lac repressor (30,31), the Xenopus SS gene transcription complex (32), mammalian RNA polymerase III (33) and nucleosomes (33,34). Several specific properties of the phage elongation complex are likely to be the basis for its different response to DNA-bound proteins. One is its very rapid elongation rate and tendency not to pause; the elongation rate is nearly 5-fold higher than that of E. coli ternary complexes. Consistent with this, we observe higher amounts of the blocked-length RNAs when transcription is carried out at low, and possibly rate-limiting, rNTP concentrations. A second feature is the enzyme's small size and ability to pry directly into essential contacts between Gln-111 protein and its recognition sequence. By comparison, in the blocked complex formed with the E. coli ternary complex, the distance between the blocked-length RNA 3' endpoint and the GAATTCC is much larger (14 bp). Together, these properties give T7 RNA polymerase a kinetic advantage in continuing to add nucleotides and move forward on the template when Gln-111 protein shifts at all on its binding site.

In the more general context of studies on prokaryotic and eukaryotic DNA binding proteins as regulators of elongation, the results presented here add to growing evidence suggesting that the outcome of an elongational block is determined by the properties of both the elongation complex and the blocking protein. Present evidence now indicates that the result ranges from transient or partial displacement of DNA-bound protein to a complete block that leads to termination and release (12,32–42). Although many of the properties of blocked elongation complexes that determine the ultimate outcome remain to be established, there is as yet no indication that blockage by the EcoRI derivatives reflects anything but DNA-bound protein functioning as a simple physical roadblock. These proteins were selected for study only on the basis of their ability to bind tightly to a unique sequence. No specific interaction with any RNA polymerase is known or expected. Thus far Gln-111 protein has been shown to block three RNA polymerases, albeit to different degrees. It will be very interesting to determine if Gln-111 protein will also prove a useful reagent to study elongation by mitochondrial and eukaryotic RNA polymerases.

ACKNOWLEDGEMENTS

We are indebted to P. Modrich and D. Wright for Gln-111 protein and EcoRI endonuclease, J. Dunn for T7 RNA polymerase, A. Perrotta and M. Been for assistance with certain of the
RNA analyses, and J.-C. Hsieh for help with the digital imaging analysis. G. Vergara provided photographic assistance. We thank P. Modrich, J. Dunn, J. Coleman, W. McAllister and M. Chamberlin for valuable discussions. This work was supported by grant GM33349 from the National Institute of General Medical Sciences.

REFERENCES