Oleinick, N. L., Chiu, S., Ramakrishnan, N., & Xue, L. (1987) *Br. J. Cancer 55* (Suppl. VIII), 135-140.
Panyim, S., & Chalkley, R. (1969) *Arch. Biochem. Biophys. 130*, 337-346.
Richter, C., Park, J.-W., & Ames, B. N. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6465-6467.
Schneider, W. C. (1956) *Methods Enzymol. 111*, 680-684.
Smith, K. C. (1976) in *Aging, Carcinogenesis and Radiation Biology* (Smith, K. C., Ed.) pp 67-81, Plenum, New York.
Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., & Klenk, D. C. (1985) *Anal. Biochem. 150*, 76-85.
Steenken, S. (1989) *Chem. Rev. 89*, 503-520.
Teebor, G. W., Frenkel, K., & Goldstein, M. S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 318-321.
Teoule, R., & Cadet, J. (1978) in *Effects of Ionizing Radiation on DNA* (Hüttermann, J., Köhnlein, W., Teoule, R., & Bertinchamps, Eds.) pp 171-203, Springer Verlag, New York.
Vieira, A. J. S. C., & Steenken, S. (1987) *J. Am. Chem. Soc. 109*, 7441-7448.
von Sonntag, C. (1987) *The Chemical Basis of Radiation Biology*, Taylor & Francis, London.
Yamamoto, O. (1976) in *Aging, Carcinogenesis and Radiation Biology* (Smith, K. C., Ed.) pp 165-192, Plenum New York.

RNA Folding during Transcription by *Escherichia coli* RNA Polymerase Analyzed by RNA Self-Cleavage†

Joseph A. Monforte, Jason D. Kahn,‡ and John E. Hearst*
Department of Chemistry, University of California, Berkeley, and Laboratory of Chemical Biodynamics, Lawrence Berkeley Laboratories, Berkeley, California 94720

Received March 14, 1990; Revised Manuscript Received May 11, 1990

ABSTRACT: We have used a self-cleaving RNA molecule related to a subsequence of plant viroids (a "hammerhead") to study the length-dependent folding of RNA produced during transcription by *Escherichia coli* RNA polymerase. Transcript elongation is arrested at defined positions using chain-terminating ribonucleoside triphosphate analogues (3′-deoxyNTP's or 3′-*O*-methylNTP's). When the transcript can form the "hammerhead" structure it self-cleaves to give a truncated product. The experiment yields an RNA sequencing ladder which terminates at the length at which cleavage becomes possible; the sequencing ladder is compared to those generated by using a noncleaving transcript or by using [α-thio]ATP in place of ATP. We have shown that 15-18 nucleotides (nt) of RNA past the cleavage point must be synthesized before the transcript can self-cleave within a ternary complex, whereas RNA freed from the complex by heating can cleave with only 3 or more nt present beyond the cleavage point. There are sequence-dependent as well as length-dependent effects. The results suggest that 12 ± 1 nt are sequestered within the ternary complex and are consistent with the presence of a DNA-RNA hybrid within the transcription bubble, as proposed by others. The results indicate that the "hammerhead" structure does not disrupt the hybrid. It appears that the RNA beyond the hybrid is not restrained by interactions with the enzyme, since the last stem of the self-cleaving structure forms as soon as the RNA composing it emerges from the DNA-RNA hybrid. Self-cleaving of the transcript offers a simple structural probe for studying less well-characterized transcription complexes. The relevance of the results to models for transcription termination is discussed.

Transcription is the most important level of control of gene expression. The process of transcription can be divided into four stages (Chamberlin, 1974): promoter binding, initiation of transcription, elongation of the transcript, and termination. Mechanistic details of all of these processes are becoming available, and it is clear that cellular regulatory mechanisms act at all four stages. Our work concerns the structure and dynamics of the ternary polymerase·DNA·RNA complex during elongation.

The formation of structure in the nascent RNA chain is an essential step in pausing, antitermination, and termination of transcription by *Escherichia coli* RNA polymerase. These processes are believed to require the formation of an RNA hairpin near the 3′ end of the transcript. The prevailing model for "simple", ρ factor independent transcription termination and for pausing requires that an RNA hairpin disrupt the RNA-DNA hybrid helix within the transcription bubble (Farnham & Platt, 1982; von Hippel et al., 1984; Yager & von Hippel, 1987, 1990). RNA structure formation during transcription is required for attenuation of transcription in the *E. coli trp* operon (Landick & Yanofsky, 1987). The action of several transcription regulation factors such as NusA, proteins N and Q, and ρ is affected by RNA structure (Brennan et al., 1987; Faus et al., 1988; Landick & Yanofsky, 1987; Yang & Roberts, 1989). These proteins have been shown to affect pausing and termination of transcription, but in some cases, it is not clear whether they interact with the nascent transcript, the polymerase itself, or both. Specific folding patterns in the nascent RNA have been proposed to direct splice site selection (Eperon et al., 1988), to mediate Tat protein activation of HIV-1 gene expression (Berkhout et al., 1989), and to affect RNA-RNA interactions responsible for the control of plasmid ColE1 replication (Polisky et al.,

†This work was supported by NIEHS Training Grant 5-T32-ES07075-12 to J.D.K., by NIH Grant GM41911 to J.E.H., and by the Director, Office of Energy Research, Office of General Life Sciences, Structural Biology Division of the U.S. Department of Energy under Contract DE-AC03-76SF00098.
*To whom correspondence should be addressed at the Department of Chemistry, University of California, Berkeley.
‡Present address: Department of Chemistry, Yale University, New Haven, CT 06511.

0006-2960/90/0429-7882$02.50/0

1990). Transient RNA structure could also affect RNA transport, assembly of protein–RNA complexes, or other aspects of RNA metabolism. Similar considerations may be important in protein folding during translation or membrane translocation (Rothman, 1989).

The order in which subsets of RNA sequence become available for folding may kinetically determine the final transcript structure, especially for large RNA molecules in which many conformations may be thermodynamically equivalent but once formed are too stable to interconvert. It has been shown experimentally that some RNA structures are present only transiently during RNA synthesis by Qβ replicase (Kramer & Mills, 1981). Possible examples of kinetically dependent structures important in vivo are those of long multiple intron mRNA's and the *E. coli* 16S ribosomal RNA. Ericson and Wollenzein (1989) have identified conformational differences between active and inactive forms of 16S RNA. The active conformation appears to be favored by kinetically driven structure formation, meaning that one of a large population of energetically similar conformations is formed preferentially during transcription. The active and inactive conformations are otherwise metastable and require heating to interconvert.

Knowledge of the detailed structure of newly formed RNA and of when it becomes available for intramolecular interactions could provide insight into the molecular mechanisms of processes such as termination and pausing, which require hairpin formation to occur within a few nucleotides of the site of polymerization. Studying the effects of antitermination factors on transient structure could help determine whether their effects are mediated through the transcript or through protein–protein interactions.

RNA structure formation during transcription antitermination has been studied by using *E. coli trp* operon mutations and RNase T_1 probing of paused transcription complexes. The transcript forms structure upstream from an RNA polymerase paused at the base of the 1:2 hairpin. The exact structure of the presumed RNA hairpin, how much of the hairpin forms, and whether it interacts with or disrupts the RNA–DNA hybrid are not known (Landick & Yanofsky, 1987).

Several groups have examined the contacts and accessibility of the nascent RNA chain. Hanna and Meares (1983) have used cleavable photoaffinity labeling techniques to show that the 5′ end of a growing RNA chain labels the DNA coding strand for 12–14 base pairs upstream from the point of synthesis on the T7 A1 promoter. The leading end of the transcript becomes accessible to DTT as transcript length increases from 12 to 14 nucleotides (nt) (Bernhard & Meares, 1986). RNA chains up to 94 nucleotides in length can contact the β and β′ subunits of polymerase. The existence of these RNA–DNA and RNA–protein interactions does not indicate that the RNA does not interact with itself, especially as photoaffinity labeling yields are typically low (Stackhouse & Meares, 1988). Additionally, the RNA–DNA contacts identified are those present in the initiation complex. The stability of these complexes is limited because of abortive initiation, and the σ subunit is often still bound. The RNA–DNA contacts may be significantly different in the elongation complex; photoaffinity labels which can be incorporated into the body of an RNA chain are now available to probe elongation complex interactions (Hanna et al., 1989).

Kumar and Krakow (1975) performed RNase digest experiments on *Azotobacter vinelandii* RNA polymerase transcription complexes. The amount of RNA protected from RNase T_1 digestion was found to be between 8 and 14 nucleotides. The photoaffinity and digestion data are consistent with the existence of an RNA–DNA hybrid of length 12 ± 2 base pairs within the transcription bubble, though direct evidence for RNA–DNA base pairing was not obtained.

Our experiments address the exact point at which an RNA transcript becomes free to form structure with itself, and whether this structure is capable of forming immediately adjacent to the RNA–DNA hybrid or interacting with the hybrid. To this end, we have developed a simple, nonintrusive assay for structure formation, using the RNA transcript itself as a probe. By using an RNA molecule which autocatalytically cleaves when it forms structure, we can identify exactly how far transcription must proceed before the RNA transcript is free enough to form the self-cleaving structure. This assay introduces no artifacts due to the presence of exogenous probes and accesses the elongation complex rather than an initiation complex. It also provides a simple method of assessing the effects of accessory proteins such as NusA on the transcript structure.

RNA capable of autocatalytic self-cleavage is common among certain viroids and virusoids and other RNA molecules [Forster & Symons, 1987; Prody et al., 1986; Uhlenbeck, 1987; reviewed in Keese and Symons (1987)]. This "hammerhead" structure requires the presence of 13 consensus sequence nucleotides, interacting to form an unknown tertiary structure linked by 3 helical stems. The cleavage reaction is site-specific and yields 2′,3′-cyclic phosphate and 5′-OH termini. Uhlenbeck (1987) has measured the rate of cleavage for "hammerheads" constructed from two separate molecules. The substrate RNA's have half-lives as short as 2 min, and unimolecular cleavage rates are likely to be even faster. The cleavage reaction provides an internal assay for RNA secondary and tertiary structure formation since only a fully formed "hammerhead" is capable of cleavage. At the salt concentrations used in transcription, the cleavage reaction is absolutely dependent on Mg^{2+} or other divalent cations, so cleavage can be rapidly and conveniently prevented by the addition of EDTA.

We find that 9–13 nt of RNA beyond the final stem of the "hammerhead" structure are required for cleavage, indicating that this length of RNA is restrained by interactions in the ternary complex. These results are consistent with the proposed length of the RNA–DNA hybrid, and they show that once the transcript is free of the hybrid its folding is not constrained by other interactions within the ternary complex. The self-cleaving transcript experiment is very simple and straightforward, and it is readily applicable to other polymerase systems including complexes modified by accessory transcription factors.

Materials and Methods

Purified *E. coli* RNA polymerase holoenzyme and NusA protein were generously provided by Prof. Michael Chamberlin (University of California, Berkeley). Ribonucleoside triphosphates, 3′-deoxyribonucleoside triphosphates (3′-dNTP's), 3′-*O*-methylnucleoside triphosphates (3′-*O*-MeNTP's), and ribonucleases were purchased from Pharmacia. Ribonuclease inhibitor (RNAsin, human placenta) was from Promega Biotec or Pharmacia. Adenosine 5′-(1-*O*-thiotriphosphate) ([α-thio]ATP) was obtained from New England Nuclear. Adenosine 5′-[γ-^{32}P]triphosphate and T4 DNA ligase were from Amersham. Cloned T4 polynucleotide kinase was supplied by New England Biolabs. Guanylyl(3′,5′)guanosine (GpG) and heparin were purchased from Sigma. Electrophoresis-grade acrylamide, bis(acrylamide), ammonium persulfate, and TEMED were obtained from Bio-Rad. Ultrapure

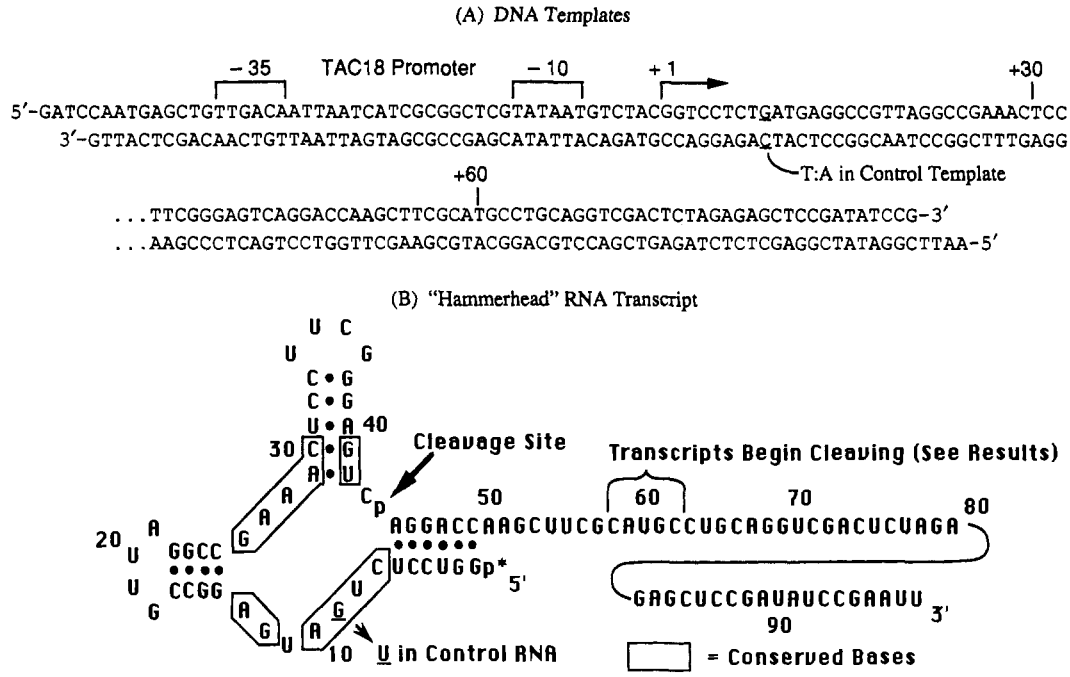

FIGURE 1: Template and transcript sequences. (A) Sequence of the synthetic double-stranded DNA template, containing a modified TAC 18 promoter followed by a sequence coding for the "hammerhead" structure. The "hammerhead" region is followed by a sequence of minimal complementarity to the "hammerhead" sequence. The control template contains a base pair change at +9 coding for a U instead of a G in the RNA product. This mutation destroys autolytic activity in the RNA. (B) The full-length RNA product. The "hammerhead", shown in its proposed secondary structure, cleaves autocatalytically between nucleotides C_{43} and A_{44} when the "hammerhead" structure is fully formed. The conserved bases required for activity, as determined phylogenetically, are enclosed in boxes. Any deviation from this sequence results in the loss of activity, as is the case for the control transcript. In addition to the conserved nucleotides, the three helical stems shown must form to promote self-cleavage. Constraints on the numbers and types of base pairs in these three stems are minimal though different sequences do exhibit a wide range of cleavage rates.

urea was from Schwarz/Mann Biotech. Diethyl pyrocarbonate treated water, silanized bullets (5% dichlorodimethylsilane/chloroform), and deionized formamide were used throughout.

Deoxyoligonucleotides (45–55 nt) were synthesized on an Applied Biosystems 381A automated synthesizer, purified by gel electrophoresis on 12% polyacrylamide gels, and ligated by standard methods (Maniatis et al., 1982) to give the transcription templates. The oligonucleotides were lightly labeled with ^{32}P for ease of purification. The ligated 145-nt products were isolated on 8% polyacrylamide denaturing gels and then electrophoresed on 8% nondenaturing gels to remove any excess single strand. The identities and double-stranded character of the templates were confirmed by restriction analysis. The templates have modified TAC 18 promoters (Mulligan et al., 1985) immediately followed by the self-cleaving transcript sequence, as shown in Figure 1.

All transcripts were initiated with [α-^{32}P]pGpG so that the 5′ end of the RNA was labeled. The dinucleotide GpG was kinased in the standard buffer (Maniatis et al., 1982) overnight at 4 °C using [γ-^{32}P]ATP. Typical reaction mixtures included 100 μCi of label (5000 Ci/mmol), 25 pmol of GpG, and 14 units of kinase in 15 μL. The kinase reaction was heat-killed for 3 min at 65 °C, and the buffer conditions were adjusted to those appropriate for RNA polymerase by adding 0.8 volume of water (or other reaction components) and 0.2 volume of a solution containing 1.2 M KCl, 10 mM $MgCl_2$, and 92 mM Tris base as described (Kahn & Hearst, 1989). Typically, this mixture was then diluted about 1.5-fold by the addition of RNA polymerase and DNA as below, and further diluted 2.5-fold into the final reaction mixture.

All transcription reactions were performed at 22 °C (room temperature), except as noted, in a buffer containing 40 mM Tris–acetate, pH 8.0, 6 mM $Mg(OAc)_2$, 20 mM KOAc, 6 mM β-ME, 0.05 mM EDTA, and 40 μg/mL acetylated BSA. *E. coli* RNA polymerase (typically 40 nM final concentration) and a 50% molar excess of DNA template were added to the [α-^{32}P]pGpG solution, and the binary complex was allowed to form for 8 min at 22 °C, at which time the polymerase–DNA complexes were aliquoted into Eppendorf tubes containing ribonuclease inhibitor (final concentration 2 units/μL), heparin (final concentration 25 μg/mL), and various nucleoside triphosphate and chain terminator mixtures. Final NTP and 3′-dNTP concentrations are given in the figure legends. We found that the 3′-dNTP terminators gave more reproducible sequencing patterns and less "crosstalk" between terminators than the 3′-*O*-methylNTP's; 3′-*O*-methylNTP's are also difficult to obtain commercially. Transcripts terminated with a 3′-dNMP migrate slightly faster than the corresponding transcript with a 3′-hydroxyl (see Figure 2), which is an aid to sequencing because it allowed us to distinguish paused transcripts and specifically terminated transcripts. The final reaction volume was 5 μL (a minimal volume was necessary for good resolution of RNA sequencing ladders). Reactions were stopped after 30–60 min (as indicated in the figure legends) by the addition of an equal volume of 100% formamide containing 20 mM EDTA and dyes. Analysis is described below.

In experiments where ternary complexes were disrupted and the released transcripts were allowed to self-cleave, reaction mixtures were boiled for 1.5 min, rapidly cooled to room temperature, incubated for the indicated length of time, quenched with formamide/EDTA, and analyzed as below.

Samples were concentrated in a Speed-Vac (Savant) to 5 μL, boiled for 1.5 min, cooled on ice, and loaded directly onto 0.5-mm-thick, 33 × 43 cm, 12% polyacrylamide [1:19 bis(acrylamide):acrylamide ratio] gels containing 8 M urea and 0.6× TBE (50 mM Tris, 50 mM borate, and 1 mM EDTA,

pH 8.3). The gels were run at 55 W (~2300 V, about 50 °C) until xylene cyanol was about 5 cm from the bottom of the gel. Minimum loading volume and a high electrophoresis temperature were required for optimum resolution and complete denaturation of the transcripts. The gels were dried carefully, and end-labeled RNA was detected by autoradiography at −70 °C using Kodak XAR-5 X-ray film and Cronex Lightning Plus enhancing screens (Du Pont); typical exposure times were 3–4 days.

To verify the identity of the cleavage product, enzymatic RNA sequencing was performed using standard methods as described (Kahn & Hearst, 1989; data not shown).

Results

The principal objective of this work was obtaining information on structure formation in nascent RNA during transcription, and the interactions of the nascent RNA with the rest of the polymerase·DNA·RNA ternary elongation complex. Structure formation of the nascent RNA is affected by steric interactions within the ternary complex and by the length of the RNA–DNA hybrid, and therefore transcript structure can report on the static and dynamic structure of the ternary complex. Our approach can be applied to examining molecular processes believed to require the formation of RNA structure during elongation, such as termination and antitermination.

The autocatalytic cleavage reaction of the "hammerhead" RNA requires a specific secondary and tertiary structure, and, therefore, the self-cleavage event provides a powerful assay for structure formation. In order for cleavage to occur in a nascent transcript, folding of the "hammerhead" sequence must be free of constraints imposed by the ternary complex. By determining the transcript length necessary for cleavage, we have measured the length of RNA between the site of polymerization and the region of unhindered structure formation. This internal RNA structure assay provides a simple way to probe the ternary complex, and use of the transcript itself as an RNA structure reporter avoids artifacts potentially introduced by exogenous chemical or enzymatic probes. The "hammerhead" structure is especially useful because of its small size and well-defined secondary structure.

These experiments require that the time scales for transcription elongation, RNA structure formation, and RNA self-cleavage be much shorter than the time required for the experiment, and disassembly of the ternary complex must be slower than all of these processes. Elongation by one base takes 20–30 ms (von Hippel et al., 1984). The rate of structure formation is not well-known: the relaxation time for equilibration of an RNA pseudoknot and a hairpin has been measured to be 0.1–1 s for a process with an activation enthalpy of 42 kcal/mol (Puglisi et al., 1990), but the rate of processes with a lower activation energy would be expected to be much more rapid, perhaps in the millisecond range. The self-cleavage reaction rate establishes a lower limit for the rate of RNA structure formation, and the half-life for the isolated "hammerhead" used here is less than 30 s (see below). The ternary complex is stable for hours (Levin et al., 1987), and evidence presented below confirms that the ternary complex is stable for the time of the experiment. Therefore, a 30–60-min experiment provides a clear picture of the extent of the cleavage reaction at each position and the influence of the ternary complex on cleavage, and it allows us to observe cleavage events occurring with a wide range of rate constants. Varying NTP concentrations (giving slower or faster elongation) and incubation times (more or less time for cleavage or disassembly) did not substantially affect the observed set of transcripts capable of cleavage. The elongation step time may be faster than some RNA structure formation processes, but once transcription is arrested by the incorporation of a chain terminator, the RNA structure has ample time to equilibrate during the experiment; when transcription is not paused or stalled, there may be situations in which the relative rates of elongation and folding may affect the kinetically accessible RNA structures.

Transcription reactions were performed in the presence of chain-terminating nucleoside triphosphate analogues, usually 3′-deoxynucleoside triphosphates. A sequencing ladder of stalled complexes was created by using the four nucleotide analogues in separate reactions. A fifth reaction using only normal triphosphates was performed to give a ladder of natural pauses and to provide a control for sequencing. Transcripts initiated at a single site from a single round of transcription were generated by using a template with the strong TAC 18 promoter and by including heparin with the triphosphates to prevent reinitiation. Only RNA molecules which had been initiated with [α-^{32}P]pGpG were detected; because the labeling is at the 5′ end of the initial transcript, only the 5′ portion of the cleaved RNA molecules is visible in these experiments. Transcription reactions were quenched with a mixture of formamide and EDTA in order to simultaneously disrupt the ternary complex and prevent further "hammerhead" cleavage by denaturing the RNA and sequestering all divalent cations.

Two different templates were used: one coding for the "hammerhead" sequence and one control containing a single base change within the "hammerhead" consensus sequence (Figure 1). The control template produces a transcript which does not cleave, giving a standard sequencing pattern of stalled complexes for the full length of the DNA template (Figure 2, lanes 1–5). The transcript sequence was designed to minimize any alternate conformations which might compete with the "hammerhead" structure and inhibit cleavage. The sequence was derived from a combination of the known "hammerhead" sequences and suggestions from Olke Uhlenbeck. The entire transcript (99 nt) was analyzed by using the RNASS RNA folding program (Williams & Tinoco, 1986) in order to ensure that sequences in the long 3′ end of the transcript would not interfere with formation of the "hammerhead" structure. An undisturbed "hammerhead" was predicted for the minimum free energy structure and for all other structures within 3.5 kcal/mol of the minimum free energy (−40.5 kcal/mol at 37 °C). Visual examination of the other structures suggested that alternate conformations would be unlikely to form during transcription. The "final stem" of the "hammerhead", referred to below, is nucleotides 1–6:44–49; it is the last part of the "hammerhead" structure to form during transcription.

Figure 2 shows a typical autoradiogram of sequencing reactions performed on the cleaving and noncleaving templates, using either ATP or [α-thio]ATP in the reactions on the cleaving template. The pattern of stalled complex transcripts from the "hammerhead" template is identical with the pattern from the control template for the first 42 nucleotides, after which the pattern changes as a result of transcript self-cleavage. A strong band corresponding to the cleavage product appears at position C_{43} (the identity of the cleavage product was confirmed by enzymatic sequencing; data not shown). The patterns are again similar between A_{44} and G_{57}, indicating that transcripts from the "hammerhead" template do not cleave in this length range. (G_{57} indicates the transcript of length 57, whose last base is G.) Beginning at position C_{58}, there is a diminution of band intensity in the "hammerhead" lanes compared to the control template, and beyond position G_{61},

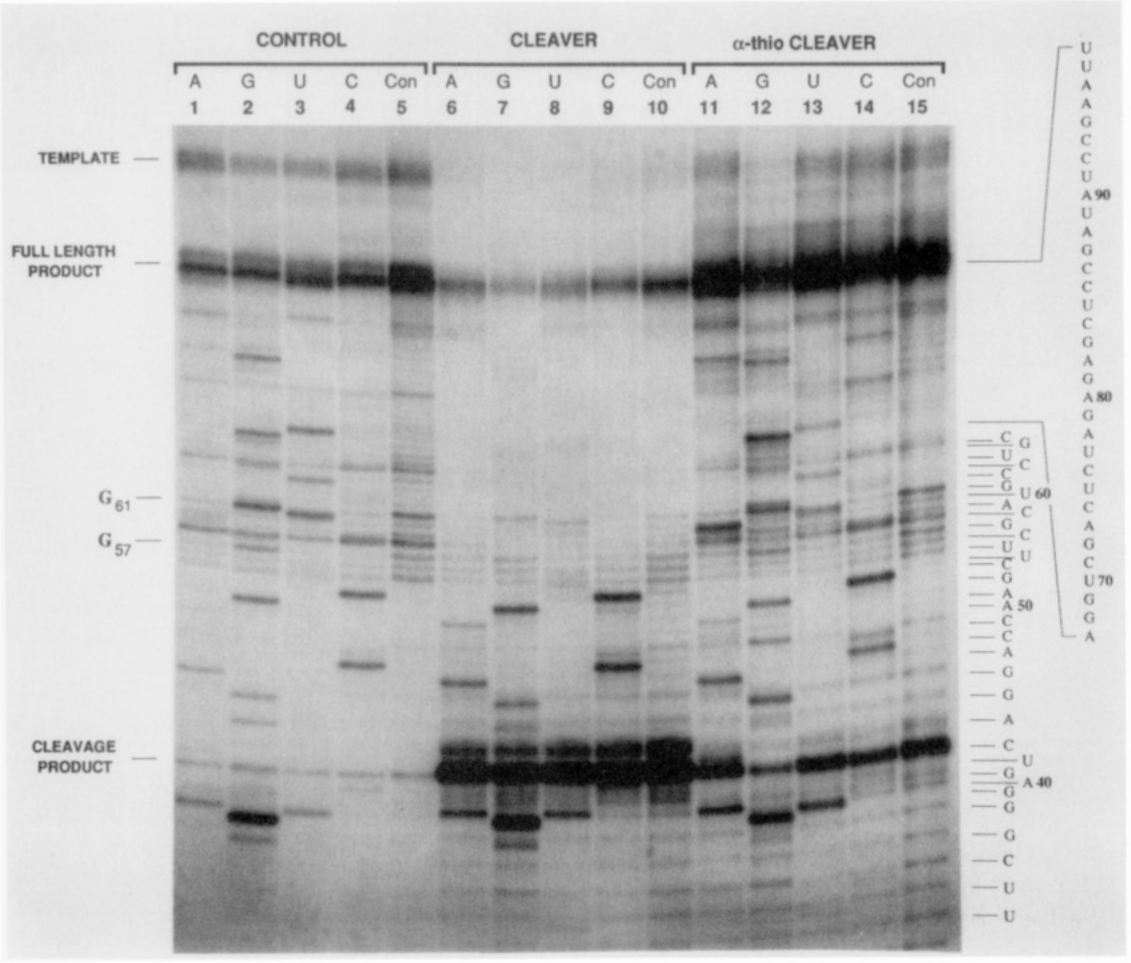

FIGURE 2: Self-cleavage of the nascent transcript. Comparison of RNA sequencing ladders from transcription of the control and "hammerhead" templates shows self-cleavage of the transcript. Lanes 1–5 show transcription on the control template in the presence of the four different chain terminators and a low rNTP concentration control. Lanes 6–10 are equivalent to lanes 1–5 except that the "hammerhead" template coding for the self-cleaving transcript is used; the cleaved product is the strong band at C_{43}. Lanes 11–15 also use the "hammerhead" template, but ATP is replaced by [α-thio]ATP in the rNTP mix. The presence of a phosphorothioate at the cleavage site significantly reduces the rate of cleavage, giving a mixture of cleaved and uncleaved products. DNA templates were lightly labeled during their preparation. The $G_9 \rightarrow U$ base change is not seen in sequencing because it is off the bottom of this gel. The reaction was performed at 22 °C for 60 min. The triphosphate concentrations used in each lane are as follows: Control and "hammerhead" experiments (lanes 1–10). A lanes: 3′-dATP, 100 μM; ATP, 80 μM; GTP, UTP, and CTP, 450 μM. G lanes: 3′-dGTP, 100 μM; GTP, 50 μM; ATP, UTP, and CTP, 450 μM. U lanes: 3′-dUTP, 100 μM; UTP, 30 μM; ATP, GTP, and CTP, 450 μM. C lanes: 3′-dCTP, 100 μM; CTP, 50 μM; ATP, GTP, and UTP, 450 μM. Control lanes: ATP, GTP, UTP, and CTP, 50 μM. [α-Thio]ATP experiment (lanes 11–15). A lane: 3′-dATP, 100 μM; [α-thio]ATP, 250 μM; GTP, UTP, and CTP, 450 μM. G lane: 3′-dGTP, 100 μM; GTP, 50 μM; [α-thio]ATP, UTP, and CTP, 450 μM. U lane: 3′-dUTP, 100 μM; UTP, 50 μM; [α-thio]ATP, GTP, and CTP, 450 μM. C lane: 3′-dCTP, 100 μM; CTP, 80 μM; [α-thio]ATP, GTP, and UTP, 450 μM. Control lane: [α-thio]ATP, 250 μM; GTP, 50 μM; UTP, 50 μM; and CTP, 80 μM.

the band intensity is dramatically reduced, indicating essentially 100% cleavage of transcripts in these stalled complexes. Thus, we find that only the transcripts which extend nine nucleotides beyond the end of the final "hammerhead" stem show any indication of cleavage, and nearly complete cleavage occurs when 13 or more nucleotides past the final "hammerhead" stem have been synthesized. This suggests that a maximum of 13 nucleotides of the transcript are unavailable for RNA–RNA base pair formation, beyond which the transcript is free to form structure.

The position of the "cutoff" for cleavage is highly reproducible. Experiments in which triphosphate or terminator concentrations were varied over a wide range (NTP's 20–400 μM; 3′-*O*-MeNTP's 50–1000 μM) gave the same cutoff position even though the overall distribution of transcript lengths was dramatically changed (data not shown). The invariance of the cutoff position indicates that differences in the fraction of polymerase that reaches a given template position are not responsible for the observed banding patterns. Note that some runoff product is visible in the "hammerhead" template lanes, and experiments described below show this is due to the presence of transcript trapped in a noncleaving conformation.

The use of [α-thio]ATP instead of ATP in the transcription reaction results in the incorporation of a phosphorothioate at the cleavage site (and other A sites), which reduces the rate of cleavage but does not substantially perturb the "hammerhead" (Buzayan et al., 1988). Because of the reduced cleavage rate, relative to the experiment time, the phosphorothioate-containing transcripts exhibit normal elongation above G_{57} along with cleavage product. These experiments show that the disappearance of transcripts longer than G_{57}, in the cleaving template transcriptions, is due to a cleavage event rather than to enhanced pausing or other structural effects on the ternary complex. The single base change in the control transcript might potentially prevent "hammerhead" formation, allowing for normal elongation. But the fact that the phosphorothioate-containing "hammerhead" structure does not affect elongation shows that the lack of higher molecular weight transcripts is not an artifact caused simply by the

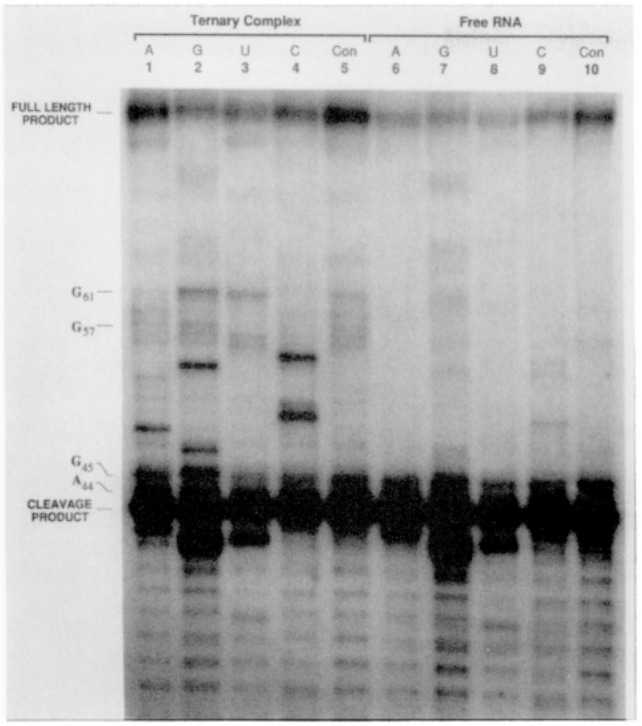

FIGURE 3: Cleavage of released transcripts. Comparison of transcript cleavage before and after disruption of the ternary complex. Transcription was performed using the "hammerhead" template at 22 °C for 30 min, at which time the reaction volumes were split. In lanes 1–5, the reactions were immediately quenched with EDTA/formamide as usual. In lanes 6–10, the reactions were heated to 90 °C for 2 min to disrupt the complex and release the transcripts, quick-cooled to 22 °C, incubated for 30 min, and then quenched. This procedure allows free transcripts of sufficient length to self-cleave. Only transcripts with two or fewer nucleotides past the cleavage site are unable to self-cleave after release, as indicated by the persistence of transcripts ending at A_{44} and G_{45}. Except as noted herein, the reaction conditions are the same as those for lanes 6–10 in Figure 2.

formation of the self-cleaving structure.

The results above demonstrate that formation of the "hammerhead" structure and self-cleavage of the transcript are inhibited for RNA molecules within the ternary complex up to 61 nucleotides in length (corresponding to molecules including up to 12 nucleotides past the end of the final "hammerhead" stem). The control experiment of Figure 3 shows that the integrity of the ternary complex is essential for the prevention of cleavage. When the complexes are disrupted by heating, the RNA transcript is released, and the figure shows that molecules long enough to form three or more of the base pairs in the final "hammerhead" stem self-cleave, as indicated by the disappearance of all but two bands above C_{43}. The significance of the minimum 3 base pair requirement in the final "hammerhead" stem for self-cleavage is discussed below. Separate experiments show that cleavage is complete within 30 s after release (data not shown). These results also indicate that the ternary complex is stable over the course of the experiment, since released RNA would rapidly self-cleave.

Figure 4 shows the temperature dependence of the self-cleavage reaction. As the temperature is increased from 4 to 50 °C, the abundance of transcripts of lengths from 58 to 61 nt decreases until virtually all have cleaved at 50 °C. This decrease could be due to enhanced displacement of the upstream end of the RNA–DNA hybrid by the final RNA–RNA stem of the "hammerhead", giving a large population of molecules with fully formed 6 base pair final stems. However, for reasons given under Discussion, we believe instead that the enhanced cleavage is due to an increased rate of cleavage for molecules with abbreviated 3 base pair final stems. (The experiment of Figure 3 shows that only these 3 base pairs are required for cleavage.) If this is the case, it again suggests that a minimum of 12 nucleotides, the distance from the third base pair of the stem to C_{58}, are unavailable for base pair formation. The stability of G_{57} at 50 °C shows that this distance does not vary with temperature.

Figure 4 also shows that the 99-nt runoff product and bands at several pause sites between 61 and 99 nt (which are visible in the control template reactions in Figure 2 as well) disappear as the reaction temperature is increased, demonstrating that these bands are probably due to a noncleaving alternate conformation. At elevated temperature, interconversion between cleaving and noncleaving conformations is more rapid than at lower temperatures, giving more efficient cleavage. Noncleaving alternate conformations have been proposed as an explanation for the failure of some "hammerhead" sequences to cleave (Uhlenbeck, 1987). At 4 °C, several intermediates of lengths between 62 and 99 nt are observable; at this temperature, the cleavage rate even within the active conformation may be slow.

The data above show that self-cleavage of the transcript occurs only after 12–13 nucleotides have been synthesized beyond the "hammerhead". When the transcript is elongated to C_{58}, the "hammerhead" sequence becomes capable of forming the minimal 3 base pair final stem, allowing cleavage to occur. Partial cleavage at 22 °C and complete cleavage at higher temperatures are observed. The results indicate that the separation between the site of polymerization and the 3 base pair final stem is at least 12 nucleotides. As the transcript is elongated to C_{62} the "hammerhead" sequence can form the complete 6 base pair final stem, giving complete cleavage at 22 °C and higher temperatures. These data determine a maximum separation of 13 nucleotides between the site of polymerization and the end of a complete 6 base pair final stem whose structure formation appears to be unhindered.

Landick and Yanofsky (1987) and Faus et al. (1988) have suggested that the transcription factor NusA affects transcription by interacting with the transcript as well as the polymerase. We have performed transcription reactions similar to those above with the "hammerhead" template in the presence and absence of NusA (data not shown). The sequencing patterns were identical under the two sets of conditions, but the lack of an observable effect does not preclude the existence of NusA/RNA interactions. NusA may not interact with this particular transcript, the interaction may not disrupt the self-cleaving structure, or the time scale of the experiment may be too slow to see any effect if NusA binding is in a relatively fast equilibrium.

Discussion

Our results clearly show that in the RNA·DNA·polymerase ternary complex the formation of intramolecular structure in the nascent RNA transcript is inhibited for the first 12–13 nucleotides upstream of the site of polymerization, whereas beyond 13 nucleotides RNA structure can form freely. The assay for RNA structure formation is the self-cleavage of the "hammerhead" transcript sequence as its final stem forms upon emerging from the ternary complex. This approach has advantages over intrusive structural probes such as RNases in that no artifacts are introduced due to steric interactions between the probe and the polymerase, and the functional assay for structure gives more global information than chemical protection experiments.

The self-cleavage experiments we have described are straightforward to perform, though they require practice and

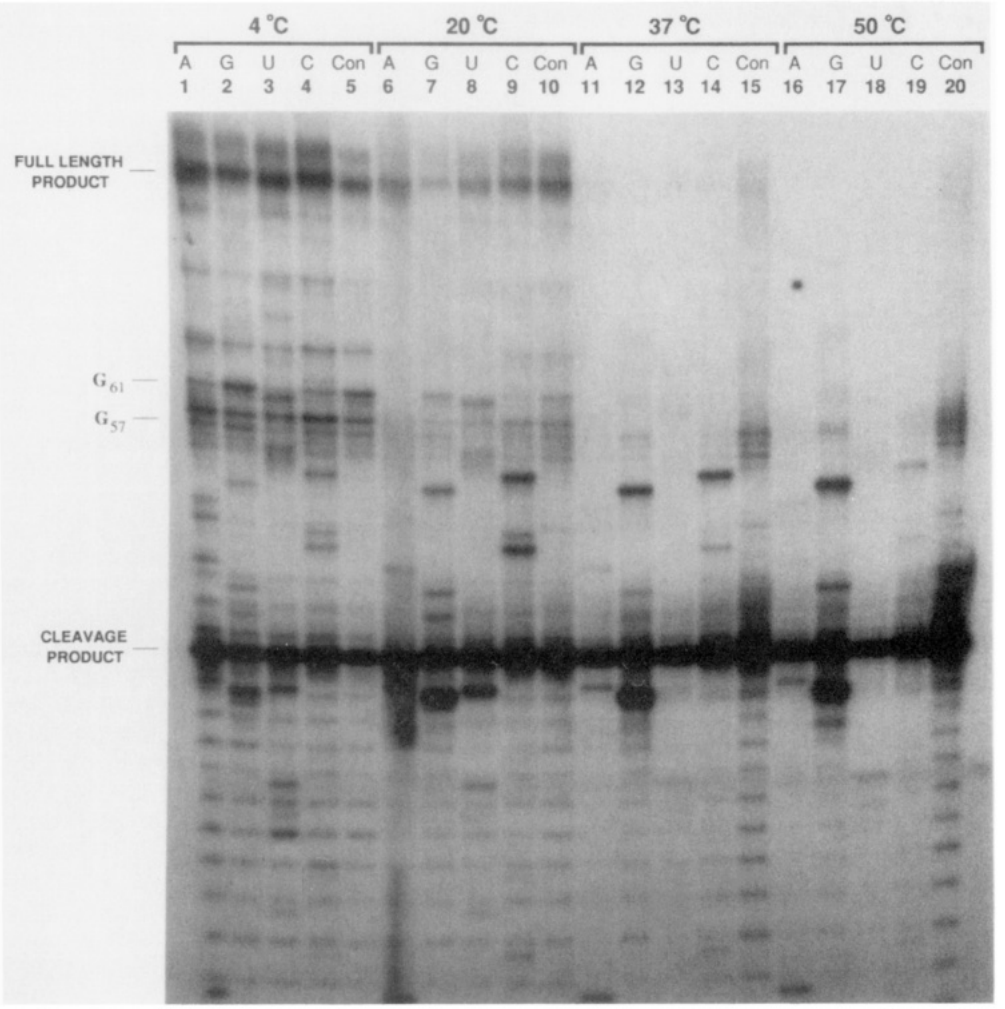

FIGURE 4: Temperature dependence of self-cleavage reaction. Temperature dependence of the self-cleavage reaction. The reactions were performed for 60 min at the following temperatures: lanes 1–5 at 4 °C, lanes 6–10 at 20 °C, lanes 11–15 at 37 °C, and lanes 16–20 at 50 °C. Except for the differences in temperature, the reaction conditions are the same as those for lanes 6–10 in Figure 2.

optimization of conditions. The experiment offers a rapid and precise measurement of the extent of RNA unavailable for folding in the transcription bubble, using readily available reagents and standard techniques. Several potential applications to other systems can be proposed. For example, polymerase accessory factors have been proposed to interact with the RNA transcript, and if any of these molecules change the transcript or complex structure, length- or sequence-dependent changes may be observed in the cleavage of the transcript. We have done preliminary experiments on the NusA-modified *E. coli* RNA polymerase without observing strong effects. It is also possible that the unmodified polymerase·DNA·RNA ternary complex can exist in altered, stable conformational states, as suggested by the recent work of Telesnitsky and Chamberlin (1989) and Goliger et al. (1989) on factor-independent antitermination. Changes in the transcript structure at the target terminator might be detectable using a modified terminator including a self-cleaving structure. The system could easily be modified to include factor recognition sequences within the stems or loops of the "hammerhead" or upstream of the 5′ end of the "hammerhead" structure. Coupled transcription and translation of a "hammerhead" could be used to study the interaction of the ribosome with the nascent transcript. The method could be used for less well-defined transcription complexes such as eukaryotic RNA polymerase II. A more practical application is in the production of transcripts with defined 5′ or 3′ ends (Dzianott & Bujarski, 1988).

Our results are consistent with the generally accepted model of the transcription bubble, which holds that an RNA–DNA hybrid of 12 ± 2 nucleotides forms upstream of the active site for polymerization. The RNA in the hybrid would be unavailable for RNA–RNA structure formation, and therefore the transcript cannot self-cleave until the essential nucleotides of its final stem emerge from the hybrid. Our results actually require only that the RNA be restrained by interaction with another component of the ternary complex. Though these results do not directly demonstrate the hybrid's presence, in the remainder of the discussion we assume its existence. It is clear that the RNA is free to form structure immediately upon exiting the hybrid, suggesting that it does not form strong, specific interactions with the RNA polymerase outside this region. Photoaffinity results showing efficient labeling of RNA polymerase subunits by RNA's up to 94 nt in length indicate only that the RNA is spatially close to the protein; they do not indicate specific interactions (Bernhard & Meares, 1986).

Models for transcription termination propose that an RNA–RNA hairpin is required for termination. Competition between RNA–RNA and RNA–DNA structure has been proposed to cause pausing and destabilization of the ternary complex. Our results suggest that the 6 base final stem of the self-cleaving "hammerhead" structure cannot disrupt the RNA–DNA hybrid and that RNA–RNA and RNA–DNA structures of comparable stability do not interconvert rapidly. If such interconversion were facile, we would expect to see cleavage starting when transcripts were elongated a minimum

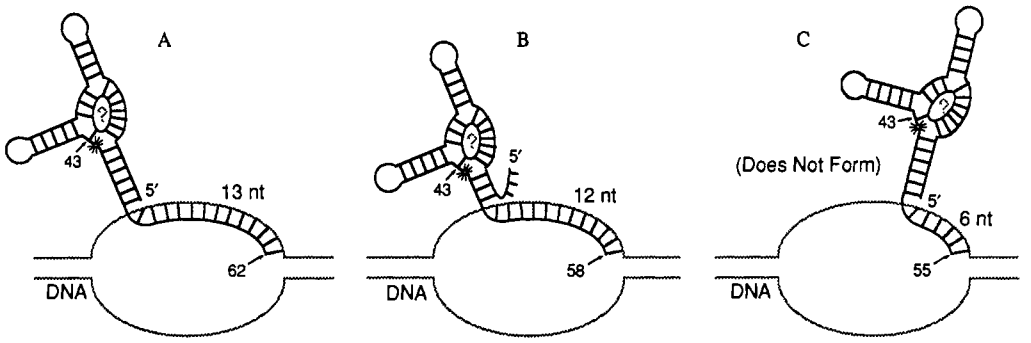

FIGURE 5: Proposed self-cleaving structures. The nascent RNA transcript, the DNA of the transcription bubble, and the RNA–DNA hybrid are shown. The cleavage point is indicated between nucleotides 43 and 44. The question mark represents the tertiary structure of the "hammerhead". The melted region of the DNA helix is shown as approximately 18 base pairs (Gamper & Hearst, 1982), and the active site for polymerization is shown at the front of the transcription bubble (Shi et al., 1988); neither of these assumptions affect the interpretation of our results. The lengths of the RNA–DNA hybrid and the 3'-nucleotide are indicated for each structure. (A) The proposed structure of transcript C_{62}, the shortest transcript which cleaves completely at 22 °C. (B) The structure of C_{58}, the shortest transcript which can cleave at temperatures up to 50 °C. (C) A hypothetical structure, not believed to form, in which the RNA–RNA stem invades the RNA–DNA hybrid, yielding cleavage of the U_{55} transcript. The total number of base pairs is the same as would be found in a complex with a 12-base RNA–DNA hybrid.

of 12 nucleotides beyond the cleavage site, when the nucleotides composing the final stem of the "hammerhead" reach the upstream end of the RNA–DNA hybrid. When less than 12 nucleotides have been synthesized beyond the cleavage point, disruption of the hybrid would require melting base pairs at the upstream end of the bubble for which no apparent Watson–Crick complement for the RNA is available, and it is reasonable that no cleavage would be observed even if there were equilibration. However, once 12 nucleotides have been made, formation of the final stem only requires that 6 base pairs of Watson–Crick RNA–RNA hybrid replace 6 base pairs of RNA–DNA hybrid, and since these helices would have comparable stability the cleaving structure should form readily. These ideas are illustrated in Figure 5. The shortest transcript which cleaves completely, C_{62}, is shown in Figure 5A with a 13-nt RNA–DNA hybrid, and the shortest transcript which cleaves under any conditions within the complex, C_{58}, is shown in Figure 5B with a 12-nt hybrid. If the RNA–DNA hybrid were being disrupted during cleavage of transcript C_{58}, we would expect that the structure shown in Figure 5C could also form, and such a structure would give cleavage at position U_{55}, which is not observed. Our model is consistent with recent studies on pausing during transcription of *Salmonella typhimurium his* operon mutants which suggest that the RNA–DNA hybrid is not completely disrupted at a strong pause site (Chan & Landick, 1989).

Since cleavage is only observed at lengths for which the hybrid can be essentially undisturbed, we propose that disruption of the hybrid may be prevented by the polymerase. These considerations and the stability of transcripts shorter than 57 nucleotides within the complex, even at 50 °C, suggest that the polymerase may specifically bind to and stabilize the RNA–DNA hybrid in the region before the proposed "separator" which disrupts the hybrid (Yager & von Hippel, 1987). We believe that the data on the temperature dependence of cleavage, showing approximately a four-base range in which cleavage becomes more efficient at higher temperatures, reflect more rapid cleavage of molecules with truncated final stems at higher temperatures rather than more rapid equilibration between complexes with complete final stems, because experiments on the RNA freed from the complex indicate that three nucleotides of the final stem are sufficient to allow cleavage (Figure 3).

Of course, terminator must eventually cause the disruption of the RNA–DNA hybrid. Our results imply that the hairpins in pause and termination sequences may need to have some minimum stability before they can overcome the binding energy of the polymerase for hybrid and thereby disrupt the hybrid. Perhaps hairpins of some minimum length are required in order to allow disruption, and there may be kinetic as well as thermodynamic barriers to termination depending on the specific hybrid base pairs which must be disrupted during hairpin formation.

ACKNOWLEDGMENTS

RNA polymerase holoenzyme and NusA protein were generously provided by the Chamberlin laboratory, UC Berkeley. D. Koh synthesized all the oligonucleotides. We thank O. Uhlenbeck, I. Tinoco, and members of the Hearst laboratory for helpful advice and discussion.

Registry No. RNA polymerase, 9014-24-8.

REFERENCES

Berkhout, B., Silverman, R. H., & Jeang, K.-T. (1989) *Cell 59*, 273–282.

Bernhard, S. L., & Meares, C. F. (1986) *Biochemistry 25*, 6397–6404.

Brennan, C. A., Dombroski, A. J., & Platt, T. (1987) *Cell 48*, 945–952.

Buzayan, J. M., Feldstein, P. A., Segrelles, C., & Bruening, G. (1988) *Nucleic Acids Res. 16*, 4009–4023.

Chamberlin, M. J. (1974) in *The Enzymes* (Boyer, P. D., Ed.) pp 333–374, Academic Press, New York.

Chan, C. L., & Landick, R. (1989) *J. Biol. Chem. 264*, 20796–20804.

Dzianott, A. M., & Bujarski, J. J. (1988) *Nucleic Acids Res. 16*, 10940.

Eperon, L. P., Graham, I. R., Griffiths, A. D., & Eperon, I. C. (1988) *Cell 54*, 393–401.

Ericson, G., & Wollenzien, P. (1989) *J. Biol. Chem. 264*, 540–545.

Farnham, P. J., & Platt, T. (1980) *Cell 20*, 739–748.

Faus, I., Chen, C.-Y. A., & Richardson, J. P. (1988) *J. Biol. Chem. 263*, 10830–10835.

Forster, A. C., & Symons, R. H. (1987) *Cell 49*, 211–220.

Gamper, H. B., & Hearst, J. E. (1982) *Cell 29*, 81–90.

Hanna, M. M., & Meares, C. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4238–4242.

Hanna, M. M., Dissinger, S., Williams, B. D., & Colston, J. E. (1989) *Biochemistry 28*, 5814–5820.

Kahn, J. D., & Hearst, J. E. (1989) *J. Mol. Biol. 205*, 291–314.

Keese, P., & Symons, R. H. (1987) in *Viroids and Viroid-Like Pathogens* (Semancik, J. S., Ed.) pp 1-47, CRC Press, Boca Raton, FL.
Kramer, F. R., & Mills, D. R. (1981) *Nucleic Acids Res. 9*, 5109-5124.
Kumar, S. A., & Krakow, J. J. (1975) *J. Biol. Chem. 250*, 2878-2884.
Landick, R., & Yanofsky, C. (1987) *J. Mol. Biol. 196*, 363-377.
Levin, J. R., Krummel, B., & Chamberlin, M. J. (1987) *J. Mol. Biol. 196*, 85-100.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mulligan, M. E., Brosius, J., & McClure, W. R. (1985) *J. Biol. Chem. 260*, 3529-3538.
Polisky, B., Zhang, X.-Y., & Fitzwater, T. (1990) *EMBO J. 9*, 295-304.
Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R., & Bruening, G. (1986) *Science 231*, 1577-1580.
Puglisi, J. D., Wyatt, J. R., & Tinoco, I. (1990) *Biochemistry 29*, 4215-4226.
Rothman, J. E. (1989) *Cell 59*, 591-601.
Shi, Y., Gamper, H., Van Houten, B., & Hearst, J. E. (1988) *J. Mol. Biol. 199*, 277-293.
Stackhouse, T. M., & Meares, C. F. (1988) *Biochemistry 27*, 3038-3045.
Uhlenbeck, O. C. (1987) *Nature 328*, 596-600.
von Hippel, P. H., Bear, D. G., Morgan, W. D., & McSwiggen, J. A. (1984) *Annu. Rev. Biochem. 53*, 389-446.
Williams, A. L., Jr., & Tinoco, I., Jr. (1986) *Nucleic Acids Res. 14*, 299-315.
Yager, T. D., & von Hippel, P. H. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Ed.) pp 1241-1275, American Society of Microbiology, Washington, D.C.
Yager, T. D., & von Hippel, P. H. (1990) *Biochemistry* (in press).
Yang, X., & Roberts, J. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5301-5305.

Use of Mono Q High-Resolution Ion-Exchange Chromatography To Obtain Highly Pure and Active *Escherichia coli* RNA Polymerase†

Dayle A. Hager,‡ Ding Jun Jin,§ and Richard R. Burgess*,‡
McArdle Laboratory for Cancer Research and Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706
Received November 1, 1989; Revised Manuscript Received May 16, 1990

ABSTRACT: A method for the purification of highly pure and active *Escherichia coli* RNA polymerase holoenzyme is described. This method is simple, reproducible, and can be performed at room temperature. The procedure involves the high-performance liquid chromatography of a partially purified RNA polymerase sample on a Mono Q ion-exchange column. Under the conditions used, RNA polymerase holoenzyme is well separated from the core RNA polymerase and other impurities. The purified RNA polymerase contains virtually no impurities as judged by SDS-polyacrylamide gel electrophoresis. The purified RNA polymerase holoenzyme contains the σ^{70} subunit in stoichiometric amounts and is at least 90% active.

The DNA-dependent RNA polymerase of *Escherichia coli* is an important enzyme that plays an essential role in gene expression and regulation. It is a multisubunit enzyme that exists in two forms. Initiation of transcription at defined promoter sites is carried out by RNA polymerase holoenzyme ($\alpha_2\beta\beta'\sigma^{70}\omega$) while elongation and termination are carried out by core RNA polymerase ($\alpha_2\beta\beta'\omega$) (Burgess et al., 1969; Burgess & Travers, 1970).

In order to study RNA polymerase in detail, it is often necessary to obtain highly purified RNA polymerase holoenzyme with the σ^{70} subunit present in stoichiometric amounts (that is, one σ^{70} subunit per core polymerase). The reason for this is that some impurities or even excess core RNA polymerase might interfere with the holoenzyme and therefore complicate the interpretation of results.

Only a few purification procedures can separate the core RNA polymerase from holoenzyme. Chromatography on a single-stranded DNA-agarose column (Nusslein & Heyden, 1972; Lowe et al., 1979) or on phosphocellulose in 50% glycerol (Gonzalez et al., 1977) separates the enzyme into core and holoenzyme-enriched fractions. However, the former method results in RNA polymerase which is only 60-70% saturated with σ^{70}. Although the latter method gives a higher molar ratio of σ^{70} subunit in the holoenzyme fraction, the increased viscosity caused by the high concentration of glycerol in the buffer makes it difficult to perform the chromatography. RNA polymerase holoenzyme can also be reconstituted by adding purified σ^{70} subunit (Lowe et al., 1979) to holoenzyme or core polymerase (Burgess & Jendrisak, 1977). However, it is time-consuming and requires access to purified σ^{70}.

In this paper, we describe a rapid purification method using Mono Q chromatography, which has proven very useful for obtaining highly pure and active *E. coli* RNA polymerase holoenzyme.

Materials and Methods

Materials. E. coli K12 cells (MG1655) were grown on 4 × LB (with 1 × NaCl) in a 10-L fermenter at the University of Wisconsin Biotechnology Center. Cells were harvested at A_{600} = 10 in late log growth phase. About 150 g wet cell paste was obtained and stored at −70 °C before use in the RNA polymerase preparation.

Most reagents were purchased and prepared as in Burgess and Jendrisak (1975). Nucleotides were purchased from

† This research was supported by NIH Grants GM28575, CA07175, and AI19635.
* Correspondence should be addressed to this author.
‡ McArdle Laboratory for Cancer Research.
§ Department of Bacteriology.

0006-2960/90/0429-7890$02.50/0