Structural Analysis of Ternary Complexes of *Escherichia coli* RNA Polymerase: Ribonuclease Footprinting of the Nascent RNA in Complexes

Sandra Milan, Linda D'Ari, and Michael J. Chamberlin*

Division of Biochemistry and Molecular Biology, University of California, Berkeley, Berkeley, California 94720-3202†

Received July 31, 1998; Revised Manuscript Received October 12, 1998

ABSTRACT: Ternary complexes of RNA polymerase containing the DNA template and nascent RNA are the intermediates in transcript elongation in all cells. We have footprinted the RNA transcript with single-strand-specific ribonucleases in ternary complexes of *Escherichia coli* RNA polymerase. When complexes are treated with elevated levels of ribonucleases A and T1, the nascent transcript can be cleaved to within 3−4 nucleotides of the 3′-terminus. Ternary complexes containing ribonuclease-cleaved transcripts as short as 3 nucleotides remain stable and active, ensuring that the cleavage occurred within an active ternary complex. However, cleavage by ribonuclease I is restricted, and gives a limited digest product of about 16 nt. At lower concentrations of ribonuclease T1, two regions of partial protection are seen. The first region extends through the first 15−16 nucleotides from the 3′-OH terminus; the second region extends from position 30 out to position 45. We interpret these regions of partial protection as defining two RNA product binding sites on the RNA polymerase that bind the product to the enzyme during elongation. Our results rule out the existence of a stable RNA-DNA hybrid in these ternary complexes of greater than 3 base pairs in length.

Transcription by prokaryotic RNA polymerases involves a cycle of four steps: (i) promoter binding and activation, (ii) RNA chain initiation and promoter escape, (iii) RNA chain elongation, and (iv) RNA chain termination and release. RNA synthesis is regulated in vivo at each of these steps. Regulation of transcription during the chain elongation and termination phases involves interaction of regulatory factors and sequences with ternary complexes. These interactions can alter the ability of the RNA polymerase to continue normal elongation, leading it to pause, become transcriptionally arrested, or terminate (*1*, *2*, *3*), or they may suppress these same processes.

Our views of transcript elongation have been strongly influenced by minimalist models for the process that postulate a structure in which the configuration of the elongating enzyme remains unchanged during the process, moving monotonically along the DNA (*4*, *5*). In these models, the RNA is bound to the ternary complex through the formation of about 12 base pairs between the DNA template strand and the 3′-end of the nascent RNA.

Over the past 10 years evidence has accumulated that raises significant questions as to the validity of this model. These findings have led to the formulation of new models that incorporate data from many recent biochemical studies (for review, see ref *3*). In these models, the RNA polymerase is postulated to have two DNA binding sites that can alternately lock and slide on the DNA. In one model, it is postulated that the polymerase has at least two RNA binding sites, each interacting with 8−10 nucleotides of the RNA, that are responsible for stable binding of the RNA to the complex. This model includes a very short RNA−DNA hybrid with only 2−3 base pairs at the 3′-terminus of the transcript (*6*, *7*). More recently a modified form of this model has been proposed in which there is an 8−9 bp DNA−RNA hybrid present at the 3′-terminus of the nascent RNA in place of the leading RNA binding site (*8*, *9*). In principle, these models can be distinguished through mapping of the 3′-terminal ends of the nascent RNA using single-strand-specific ribonucleases (RNases).

EXPERIMENTAL PROCEDURES

Enzymes and Reagents. RNA polymerase holoenzyme was purified from *Escherichia coli* DG 156 cells as described previously (*10*). The preparation was >90% pure and contained 40−70% active molecules as measured by the method described by Chamberlin et al., 1979 (*11*). Reagents were purchased from the following sources: [α-^{32} P] UTP (NEN); HPLC grade nucleoside triphosphates and Sephadex G-50 (Pharmacia); ApU (Sigma); Ribonuclease A and T1 (Boehringer Mannheim); Ribonuclease I (Promega); Biospin 30 columns (BioRad); and Dynabeads (Dynal International, Great Neck, NY). Because of the variability in the formation of ternary complexes we have not used other sources (*12*).

Buffers. SDS−urea gel loading buffer contained the following: 7 M urea, 10 mM EDTA, 0.5% SDS, and 0.05% each xylene cyanol and bromophenol blue. Transcription buffer contained the following: 44 mM Tris-HCl (pH 8.0), 14 mM $MgCl_2$, 14 mM 2-mercaptoethanol, 20 mM NaCl, 2% w/v glycerol, 0.04 mM EDTA, and 40 μg/mL acetylated bovine serum albumin. TBE buffer contained the follow-

† This research was supported by a grant (GM12010) from the National Institute for General Medical Sciences.

* To whom correspondence should be addressed. Fax: (510) 642-7846.

10.1021/bi9818422 CCC: $18.00 © 1999 American Chemical Society
Published on Web 12/08/1998

ing: 89 mM Tris-borate, pH 8.3, and 2.5 mM EDTA.

Preparation of U21 and U40 Ternary Complexes. Stable ternary complexes were formed by the method of Krummel and Chamberlin (*12*, *13*). This procedure involves the formation of defined complexes halted at a specific site along the DNA template by using a dinucleotide primer and a limited subset of NTPs. The U21 complexes were formed by initially preparing A20 complexes using bacteriophage T7 D123 DNA as template (*12*). The A20 complexes were formed by a 2 min incubation at 30 °C with 50 μM ApU, and 5 μM ATP, CTP, GTP. For the removal of the unincorporated nucleotides, the reaction mixture was passed through two gel filtration columns (1 × 4 cm packed with Sephadex G-50 at room temperature and equilibrated with transcription buffer without BSA and magnesium). The fractions containing the excluded volume were pooled and stored at 4 °C. The 3′-end-labeled U21 complexes were formed by walking (*12*, *13*) the A20 complexes to position 21 with [α-^{32}P] UTP. The 3′-end-labeled U40 complexes were formed using intact, supercoiled pM19 plasmid DNA as template (*14*). The synthesis and walking procedures were the same as those for the U21 complexes.

Ribonuclease Treatment of Complexes and Free RNA. U21 and U40 ternary complexes and the corresponding free RNAs were digested by ribonucleases (digests range from 500 to 0.5 units/mL for RNase T1, from 100 to 0.1 μg/mL for RNase A, and from 100 to 0.0005 units/mL for RNase I) for one minute at 30 °C. After one minute, the digests were divided in half. One-half was stopped immediately by extraction with a mixture (25:24:1) of phenol, chloroform, and isoamyl alcohol. The other half was "chased" by incubating the truncated transcripts in the presence of all four NTP substrates (100 μM) and excess cold carrier RNA (1 mg/mL) for an additional minute. The chase reaction was stopped by either organic extraction as described above or by the addition of urea−SDS gel loading buffer.

To determine the length of the fragments produced by RNases A and T1 on transcripts longer than U21 and U40, one must know the RNA sequence beyond positions U21 and U40. The sequence of T7 D123 RNA reading from the T7 A1 promoter start site is the following: pppAUCGAGAGGG-ACACGGCGAAU*AGUGAGA here the asterisk is placed just following the U21 that forms the 3′-terminus of U21 RNA. The sequence of plasmid M19 RNA reading from the T7 A1 promoter start site is: pppAUCGAGAGGGAA-GAGAAGAAGAGAGAGAGGCACGGCGAAU*AGCCAU here the asterisk is placed just following the U40 that forms the 3′-terminus of U40 RNA.

Piperidine Treatment of Free U21 and U40 RNAs. Piperidine ladders of 3′-end-labeled RNAs were generated by incubating the purified U21 or U40 transcripts plus carrier RNA (0.3 mg/mL) in the presence of 1% piperidine at 92 °C for 2 min. The piperidine was removed by evaporation, and the samples were resuspended in urea−SDS gel loading buffer.

Hybrid Controls. The free U40 and U21 transcripts were hybridized to a 50 nt DNA oligomer whose sequence is identical to that of the T7 D123 template strand. The hybridization procedure involved heating the RNA and DNA to 65 °C and allowing it to cool slowly to room temperature. Once the template strand and transcript were annealed, the hybrids were treated with ribonuclease under the same conditions as the intact complexes.

Analysis of Transcripts by Denaturing Polyacrylamide Gel Electrophoresis. Samples were resuspended in urea−SDS gel loading buffer and directly loaded onto a 1 X TBE gel (0.3 mm thick, 40 cm long) containing 7 M urea, 20% (w/v) acrylamide, and 3% (w/v) bis-acrylamide. Electrophoresis was carried out at 1200 V until the xylene cyanol had migrated 20 cm from the wells (approximately 16 h). The gels were left on the plate, covered with plastic wrap, and autoradiographed in the presence of an intensifying screen at −70 °C for approximately 24 h.

RESULTS

Rationale. The goal of these studies was to map the structure of the nascent transcript in defined elongation complexes formed with *E. coli* RNA polymerase. These defined complexes were synthesized by walking the RNA polymerase to specific points along the DNA by the use of dinucleotide primers and limited subsets of NTPs (*12*, *13*). It was desirable to have the nascent RNA terminally labeled at the 3′-OH end, to detect cutting along the transcript in this region. Hence, the major part of the transcript was synthesized in a first reaction using unlabeled NTPs; this complex was purified, and then incubated with a single, labeled NTP, to form a ternary complex bearing a 3′-terminally labeled transcript. Two complexes, designated U21 and U40, were ultimately footprinted. The initial transcript sequences for each of these RNAs are shown below in summary figures; each of the RNAs are transcribed from the bacteriophage T7 A1 promoter, and both transcripts share the same sequence for the first 11 nucleotides. Transcript U21 is read from the wild-type T7 DNA sequence; transcript U40 is read from a modified sequence in which a 19 bp stretch of A and G residues has been inserted (*14*).

RNase footprinting was carried out using three single-strand-specific RNases. Digestion with these RNases was limited to one minute reaction periods to avoid interference by the transcript cleavage reaction carried out by RNA polymerase itself (*15*). RNase A is base-specific and cuts on the 3′ side of uridine or cytidine residues to give 3′-phosphoryl termini. RNase T1 cuts on the 3′ side of guanosine residues to give 3′-phosphoryl termini. In contrast, RNase I cuts on the 3′ side of all four nucleosides (*16*).

Note that, in the nomenclature we use for the synthesis of ternary complexes, we designate complexes on the basis of the length of the RNA, beginning at the 5′-terminus and reading toward the 3′-terminus. However, for the discussion of the RNase mapping experiments, we will designate the digestion products relative to the 3′-end which is labeled. Hence production of a labeled 4 nt fragment must result from cleavage of the RNA between the fourth and fifth nucleotides from the 3′-terminal end.

Structural Analysis of U21 Complexes. Polyacrylamide gel analysis of transcripts extracted from a typical 3′-end-labeled U21 complex shows that the U21 RNA obtained in our walking procedure is nearly homogeneous (Figure 1A, lane 2) with a small trace of a 24 nt contaminant produced by readthrough. The sites of digestion by RNases that generate 3′-phosphoryl terminal products are easily mapped by

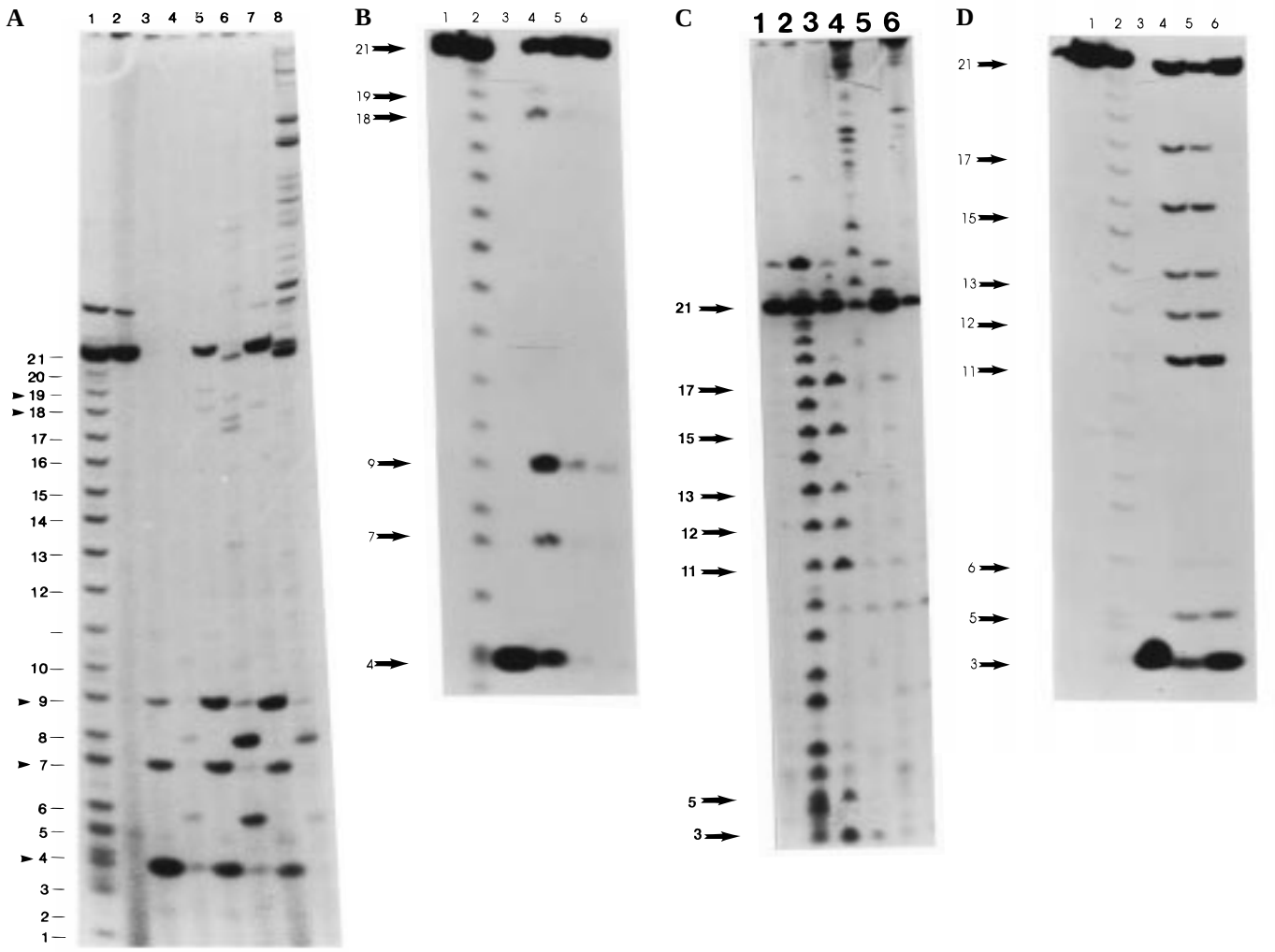

FIGURE 1: RNase footprinting analysis of active U21 complexes and free U21 RNA. Complexes were prepared and subjected to RNase treatment as described in Materials and Methods. All digests were carried out at 30 °C, for 1 min. The resulting fragments were analyzed by denaturing PAGE and autoradiography. Sites of digestion are shown on the left of each figure. (A) Active U21 complexes digested by RNase A: lane 1, piperidine ladder; lane 2, U21 starting material; lanes 3, 5, and 7 correspond to digests done at 100, 10, and 1 μg/mL of RNase A, respectively; lanes 4, 6, and 8 correspond to the chases of the complexes after digestion. (B) RNase A digestion of free U21 RNA. After phenol extraction of the U21 complexes, the free RNA was treated as described in (A): lane 1, untreated RNA; lane 2, piperidine ladder; lanes 3−6, digests at 10, 1.0, 0.1, and 0.01 μg/mL of RNase A, respectively. (C) Active U21 complexes digested by RNase T1. Complexes were formed and treated as described in Materials and Methods: lane 1 corresponds to RNA transcripts from untreated U21 complexes; lane 2, piperidine ladder; lanes 3 and 5 correspond to RNase treatment with 500 and 50 units/mL of RNase T1, respectively; lanes 4 and 6 correspond to the chases of the complexes after digestion. (D) RNase T1 treatment of free U21 RNA. After phenol extraction of the U21 complexes, the free RNA was treated as described in (C): lane 1, U21 starting material; lane 2, Piperidine ladder; lanes 3−6, RNase T1 digests done at concentrations of 500, 50, 5, and 0.5 units/mL, respectively.

comparison to a ladder of randomly cleaved 3′ phosphoryl terminated products (Figure 1A, lane 1) generated by alkaline (piperidine) treatment of the 3′-end-labeled RNA, and by comparison with the known RNA sequence.

The products resulting from treatment of 3′-end-labeled U21 ternary complexes with decreasing amounts of RNase A (100, 10, and 1 μg/mL), are shown in Figure 1A (lanes 3, 5, and 7, respectively). From the specificity of RNase A and the sequence of the U21 RNA, products are expected of lengths 4, 7, 9, 18, and 19 nt (see map of cleavage sites on the U21 sequence in Figure 2). Digestion products of lengths 4, 7, and 9 nt are major products at all three RNase A concentrations. Digestion products of lengths 18 and 19 nt are minor products seen in Figure 1A, lanes 5 and 7.

Digestion at 1 μg/mL leaves about 50% of the initial U21 RNA in the complex unaffected (Figure 1A, lane 7); hence the products of digestion at this RNase concentration are probably generated primarily from single hits. Digestion at 100 μg/mL leads to the complete disappearance of the original U21 RNA and the accumulation of the 4 nt RNA as the major labeled product (Figure 1A, lane 3).

The data obtained from these kinds of digestion protection experiments can help describe the structure of the RNA in the ternary complex only if there is evidence that the RNA transcript remains bound to the complex during digestion, and that the complex remains catalytically active. To address these issues, we carried out chase experiments in which all four NTP substrates were added to the complexes after RNase treatment to permit elongation of any of the truncated transcripts that remained bound to the complexes (Figure 1A, lanes 4, 6, and 8). This nucleotide addition led to the disappearance of nearly all (>90%) of the primary RNase A digestion products of 4, 7, and 9 nt and the appearance of longer transcripts in all of the RNase A-treated lanes (see Figure 1A, lanes 4, 6, and 8; note that in lane 4 much of the sample appears to have been lost, because the total amount

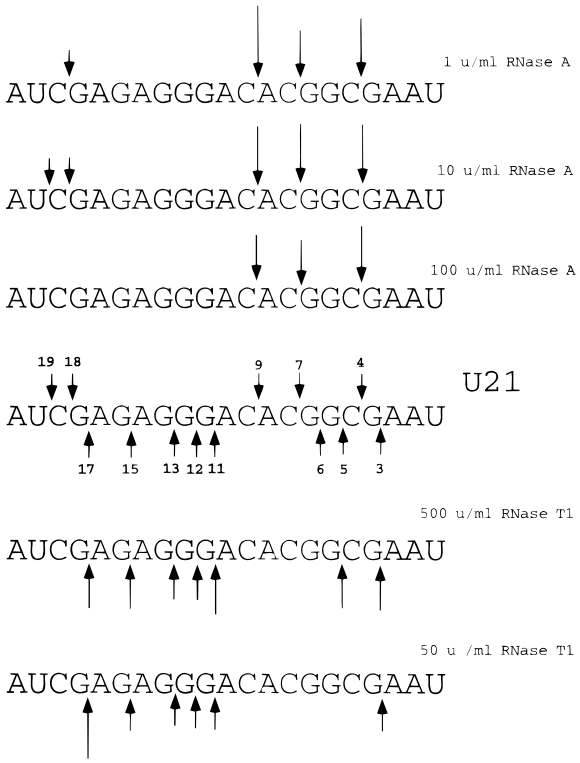

FIGURE 2: RNA sequence and summary of digestion pattern of U21 complexes. The numbered arrows above represent the expected sites of digestion by RNase A. Numbered arrows below represent the expected sites of digestion by RNase T1. Nonnumbered arrows represent a summary of the results obtained after treatment with RNase A and T1. Larger arrows indicate sites of greater frequency cutting, and smaller arrows indicate sites of reduced cutting.

of radioactivity should be identical to that in lane 3). Hence virtually all of these 3′-terminal fragments remained bound to the RNA polymerase in an active complex during the period of the experiment and could be elongated to longer transcripts. (In separate experiments it was found that RNase-cleaved ternary complexes bearing short 3′-end-labeled fragments (3−4 nt) are only stable for 5−10 min. Quite similar results have been reported (*7*) for mammalian RNA polymerase II ternary complexes.) Our ability to chase RNase-cleaved 3′-labeled RNA fragments into larger RNAs by adding NTP rules out the possibility that the short transcripts seen in these experiments were the result of the digestion of RNAs released from the ternary complex due to preferential binding of the RNA to the RNase A. Since RNase A is highly specific for single-stranded RNA at concentrations of 1 μg/mL, it appears that the U21 transcript is single-stranded 4 nt from the 3′-terminus.

We have compared the specificity and frequency of RNase A cutting of U21 in the ternary complex with that of free U21 RNA by extracting U21 RNA from ternary complexes and treating it with variable amounts of RNase A (Figure 1B). The extent of cutting was significantly greater for the free RNA as compared to the RNA in the complex under identical conditions. At 10 μg/mL RNase A (Figure 1B, lane 3), the free U21 RNA is quantitatively cut to the smallest 4 nt product, while significant amounts of U21 RNA remain intact in the ternary complex (Figure 1A, lane 5). These differences suggest that binding of U21 RNA in the ternary complex reduces the rate of RNase A digestion at all sites along the transcript. There are also minor differences in the frequency of cutting between free U21 RNA and U21 RNA in the complex. In the complex at lower RNase A concentrations there is somewhat more of the 9 nt product, as compared with the 4 and 7 nt products. With the free RNA the 9 and 4 nt products are about equal in amounts, and there is less of the 7 nt product. This may be due to the formation of weak secondary structure in the free RNA. Similarly, cutting at position 18 appears somewhat enhanced in the reactions containing free RNA as compared to complexes (compare Figure 1A, lane 7, and Figure 1B, lane 4).

The 3′-terminally labeled U21 complex was also treated with RNase T1. At a concentration of 500 units/mL, RNase T1 cleavage of U21 RNA in complexes is also efficient throughout the length of the transcript (Figure 1C, lane 3). Under these digestion conditions over half of the full-length 21 nt transcript remains; hence most of the products are due to single-cleavage events. The amounts of the cleavage products are present in approximately equal concentrations, with expected products of lengths 3, 11, 12, 13, 15, and 17 nt easily identified. The expected products at lengths 5 and 6 nt are seen, but are very faint. Once again, there is nearly quantitative chasing of all the cleavage products when NTPs are added after RNase T1 digestion. Hence all of the 3′-end-labeled RNA products are present in ternary complexes, and these complexes are active. By the criterion of RNase T1 cleavage, then, the maximum length of any stable DNA−RNA hybrid at the 3′-terminus of the nascent RNA transcript is 3 bp or less.

The free U21 RNA was treated with RNase T1 under the same conditions as those for the ternary complex. (Figure 1D; lanes 3−6 correspond to digests performed with 500, 50, 5, and 0.5 units/mL of RNase T1, respectively.) There are dramatic differences in the rate of cutting of the free U21 RNA and the U21 RNA in the complex. At 500 units/mL (Figure 1D, lane 3), the free RNA is completely degraded to its limit of 3 nt product. By contrast, more than 50% of the RNA in the ternary complex remains intact at the same concentration (Figure 1C, lane 3). At 50 units/mL, less than 10% of the RNA in the ternary complex is cut while more than 50% of the free RNA is cut (Figure 1C, lane 5, and Figure 1D, lane 4). These differences suggest that the RNA transcript is being protected by the polymerase while in the complex.

A comparison between the cutting pattern of the RNA in the complex digested at 500 units/mL and the free RNA digested at 50 units/mL shows that cutting to give products of 3 and 11 nt is somewhat enhanced in the free RNA, and that cutting to give the 17 nt product is somewhat reduced. Note that formation of the 6 nt product is not seen in either the RNA in complexes or the free RNA. This finding eliminates the possibility that there is protection by the polymerase or DNA base pairing at this site. This lack of sensitivity at this position could be due to secondary structure; however, this transcript has no predicted secondary structure as determined by computer analysis, and hence it seems likely that this site is intrinsically resistant to digestion.

Structural Analysis of U40 Complexes. Transcripts extracted from a typical U40 synthesis are shown in Figure 3, part A and part B, lane 1. Close examination of the starting material shows that the synthesis of U40 is homogeneous with the exception of a minor band at position 15 and several

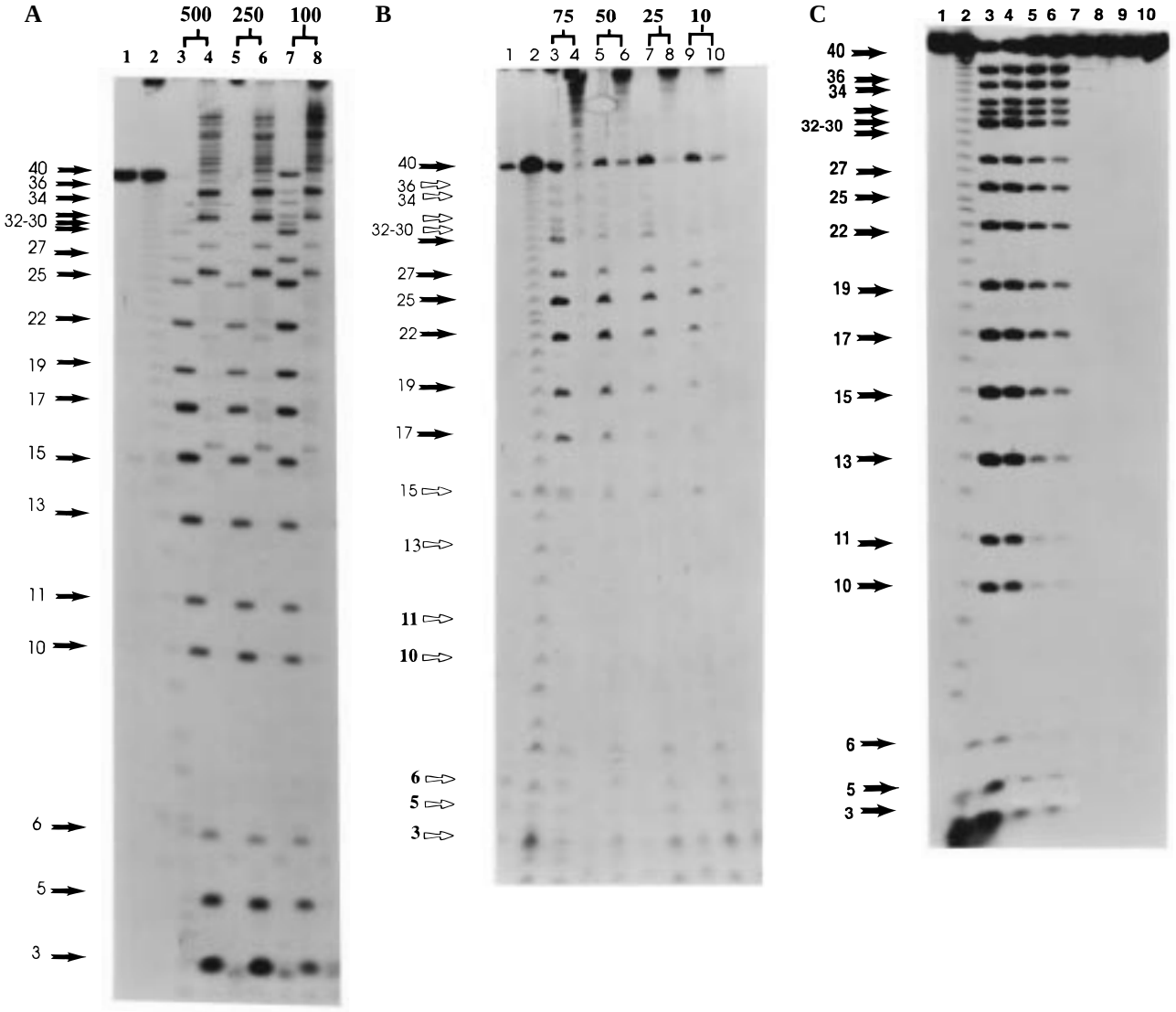

FIGURE 3: Analysis of U40 ternary complexes treated with RNase T1. The complexes were formed and treated as described in materials and methods. All digests were carried out at 30 °C for 1 min. (A) Sites of digestion by RNase T1 are indicated on the left: lane 1 corresponds to the U40 starting material; lane 2 corresponds to the U40 RNA after piperidine treatment; U40 ternary complexes shown in lanes 3, 5, and 7 were treated with RNase T1 concentrations of 500, 250, and 100 units/mL, respectively; lanes 4, 6, and 8 correspond to the chases of the complexes after treatment. (B) U40 complexes treated with lower concentrations of RNase T1: lane 1, starting material; lane 2, piperidine ladder; lanes 3, 5, 7, and 9 correspond to U40 RNA treated with 75, 50, 25, and 10 units/mL of RNase T1, respectively; lanes 4, 6, 8, and 10 correspond to the chases of the complexes after treatment. (C) Free U40 RNA and RNA−DNA hybrid treated with RNase T1. The free RNA was treated under conditions and RNase concentrations identical to those of the complexes in panel A: lane 1, starting material; lane 2, piperidine ladder; lanes 3−6 correspond to digests done at 75, 50, 25, and 10 units/mL, respectively; lanes 7−10 are RNA−DNA hybrids treated at 500, 250, 100, and 50 units/mL of RNase T1.

smaller products (less than 6 nt) visible at the bottom of the gel.

Active, 3′-end-labeled U40 complexes were examined with RNases A, T1, and I under different conditions. However, we will limit our discussion to the results obtained with RNase T1 and RNase I since RNase A gave no additional information. Products from digests done at high concentrations of RNase T1 (500, 250, or 100 units/mL) are shown in Figure 3A (lanes 3, 5, and 7m respectively). Under these conditions, every product expected from the transcript sequence is produced, although at different frequencies (for a summary, see Figure 4). Examination of the products from digests done at 500 and 250 units/mL shows that the full-length transcript has been digested almost completely, giving a limit product from the 3′-end that is 3 nt in length. At these concentrations the signals from the 36, 34, 32, 31, 30, and 27 nt products are very faint indicating that these transcripts have also been cut more than once. Hence, the intense signals from the 17, 15, 13, 5, and 3 nt products suggest that these shortened transcripts are the products of both single and double cutting events. Note that the production of the 6 nt product is very low. This is the same base and sequence context as position 6 of the U21 complex, which is relatively insensitive to RNase T1 digestion (see above).

Decreasing the concentration of RNase T1 to 100 units/mL (Figure 3A, lane 7) leaves about 15% of the full-length transcript intact, indicating that the products are still generated by double and single cutting events. At this concentration, the 25, 22, 19, and 17 nt products are most abundant, while the 30, 27, 13, 11, 10, 5, and 3 nt products are represented in intermediate quantities and the 36, 34, 32, and 31 nt products are produced at lower frequencies. Note that, under these lower concentrations of RNase T1, the

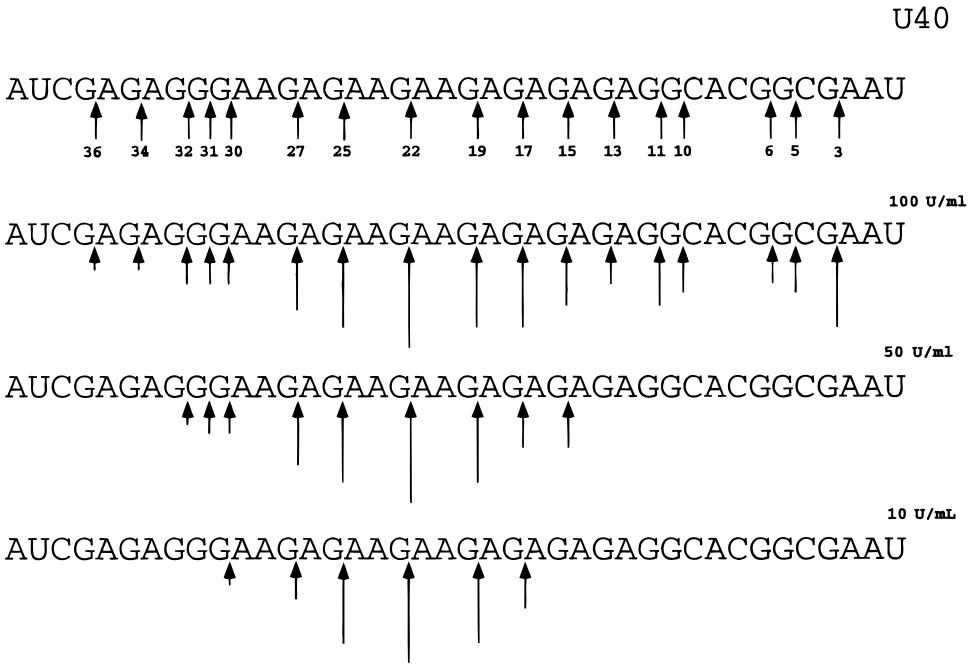

FIGURE 4: U40 RNA sequence and summary of RNase T1 digest. The top sequence with the numbered arrows represents the sites of expected cutting by RNase T1. The lower sequences represent the summary of results obtained from treatment of the complexes at 100, 50, and 10 units/mL final concentration.

transcript is still digested up to the last G in the sequence, which is 3 nucleotides upstream from the growing point.

Again, the chase experiments on the truncated transcripts show that the 3′-terminal fragments of lengths 3, 5, and 10 nt and larger sizes remained within the complexes in a configuration that permits them to be elongated almost completely to longer transcripts since these labeled bands are chased to larger products in the presence of the NTP substrates (Figure 3A, lanes 4, 6, and 8). These results complement the data obtained from the analysis of the U21 complex treated with RNase T1, which showed that the transcript is single-stranded, at least beyond 3 nt from the catalytic site.

We next analyzed the U40 complexes at lower RNase T1 concentrations in order to analyze products produced by single hits. The results obtained from these analyses showed dramatic changes in the digestion pattern which suggests that the transcript is protected in two regions (Figure 3B, lanes 3, 5, 7, and 9). The first region of protection extends from the 3′-end of the transcript up to position 15. The second region of protection extends from position 27 up to position 40. We are confident that the products obtained in this analysis are the products of single cutting events since a high percentage of full-length RNA remained intact.

The products resulting from RNase T1 digests done at 75, 50, 25, and 10 units/mL show that the 3′-end of the transcript is not digested (Figure 3B lanes 3, 5, 7, and 9, respectively). Gel analysis of all four digests at these lower RNase T1 concentrations shows loss of signal from the 15, 13, 11, 10, 6, 5, and 3 nt products. The loss of these products suggests that there is protection by the enzyme in the 3′ region of the transcript out to about 15 nt.

This conclusion is also supported by digestion of the 3′-end-labeled U40 complex with RNase I (Figure 5). Use of an excess of this nuclease gave a 15−17 nt protected RNA as a "limit product". It appears that the RNase I is sterically blocked from digesting sequences near the 3′-terminus of the RNA, while RNases A and T1 can enter the site, at least at higher enzyme concentrations.

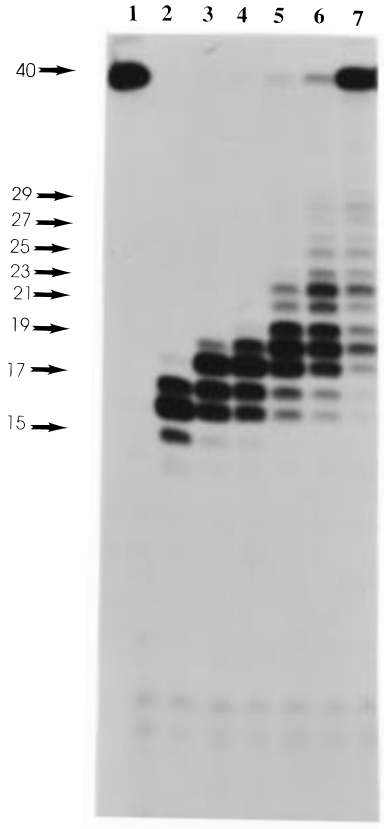

FIGURE 5: Analysis of U40 ternary complexes treated with RNase I. U40 ternary complexes were formed and treated as described in Materials and Methods. Sites of digestion are shown on the left side of the figure: lane 1, starting material; lanes 2−7 correspond to digests done at 100, 10, 1, 0.05, 0.001, and 0.0005 units/mL final concentrations of RNase I, respectively. All digests were carried out at 30 °C, for 1 min.

Close examination of the products obtained in the treatment of U40 ternary complexes with RNase T1 at lower concentrations (Figure 3B) suggests that there might be another region of protection within the 5′-end of the transcript. The low yields of the 36, 34, 32, 31, 30, and 27 nt products from digests done at 75, 50, and 25 units/mL and the complete disappearance of these products at 10 units/mL suggest either that this region of the transcript is protected by the polymerase or that it might contain secondary structure. The latter possibility can be ruled out by comparing the digestion pattern of the free RNA to that of the complex RNA (see below). These results are the first to show directly that there are two regions of protection for the nascent transcript offered by the RNA polymerase in the ternary complex.

Decreasing the concentration of RNase T1 revealed that particular areas of the transcript remain very accessible to digestion. The most abundant products formed are the 22 and 25 nt products. Decreasing the RNase concentration from 75 to 10 units/mL shows that this region continues to be sensitive to digestion, suggesting that this region of the transcript does not interact with the RNA polymerase. The stability and activity of the truncated complexes was tested by adding NTPs to the reaction after digestion (Figure 3B, lanes 4, 6, 8, and 10). These chase experiments demonstrate that digestion took place within active complexes, since all truncated transcripts elongate into longer RNAs.

To draw conclusions regarding the interactions of the RNA polymerase with the nascent transcript, we needed to analyze the free RNA in the same manner with which we analyzed the complexes. For this comparison, 3′-end-labeled U40 RNA was extracted from the ternary complexes, and was digested with 75, 50, 25, and 10 units/mL of RNase T1 as shown in Figure 3C, lanes 3, 4, 5, and 6, respectively. Digests were also done at 500, 250, and 100 units/mL which resulted in complete digestion of the transcript RNA into the limit product (data not shown).

Comparing the digests done at 75 and 50 units/mL RNase T1 (Figure 3, B and C) of both the free RNA and the RNA in complex, one can conclude that there is significant protection offered by the complex. In the free RNA the 3, 5, 10, 11, 13, and 15 nt products are formed with high frequency, indicating that these sites are strongly protected by the complex. Cutting at sites 22 and 25 is not enhanced in the free RNA (Figure 3C, lanes 3−7), suggesting that this region of the transcript is in a conformation which makes it more accessible to RNase digestion when it is in the complex. Cutting at positions 31, 32, 34, and 36 appears more intense in the free RNA than in the complex, indicating that these sites are also protected by the complex. In summary, the results from the free RNA digests show that the U40 transcript does not have significant secondary structure and confirms that the two regions of protection from RNase are indicative of enzyme−transcript interactions.

As a final and important control, we tested whether RNase T1 could cleave an authentic, 3′-end-labeled U40 RNA−DNA hybrid; Figure 3C shows the results of digestion with 500, 250, 100, and 50 units/mL, respectively (lanes 7−10). There is no detectable cleavage of the RNA at any of these RNase T1 concentrations.

DISCUSSION

We have synthesized defined populations of 3′-end-labeled elongation complexes of *E. coli* RNA polymerase and analyzed the structure of the nascent RNA by RNase footprinting experiments. By employing chase experiments, after RNase cleavage, the complexes used in these experiments were shown to remain stable and active throughout the manipulations. These results specifically exclude the possibility of rearrangement of the active site away from the 3′-end of the RNA (*12*, *13*, *17*). Consequently, the structural information obtained from these studies should reflect structures of elongation intermediates.

Although the complexes examined here vary in transcript sequence and length they show consistent features. Both ternary complexes showed cleavage of the 3′-proximal region with RNases A and T1 to give terminal fragments of 3−4 nt. Thus, no stable RNA−DNA hybrid region can be present in the ternary complex that is longer than 3 bp. Control experiments with a synthetic U40 RNA−DNA hybrid do not show cutting of the RNA, even at the highest RNase T1 concentrations used in the footprinting experiments. This result demonstrates that there is nothing unusual about the hybrid structure that would allow single-strand-specific RNases to cut. Similar results have been obtained with ternary complexes of RNA polymerase II (*7*) in which the RNA transcript was digested 3−4 nucleotides away from the growing point.

It might be argued that the RNases used can act as melting proteins, and therefore might disrupt a pre-existing DNA−RNA hybrid in the ternary complex in the process of cutting the nascent RNA. However, physical studies of the interaction of RNase A with nucleic acids have shown that it does not denature duplex DNA directly, but acts by trapping transiently melted regions. Since the binding site size is about 11 nt, this would require that the entire hybrid (for a hypothetical 12 bp hybrid) be melted in the ternary complex (*18*). Hence these considerations rule out the existence of a 12 bp stable, DNA−RNA hybrid.

In the case of RNase T1, structural studies of the protein bound to 2′-guanylic acid do not show a single-strand-specific RNA binding site. Rather, the catalytic site interacts directly with GMP, through stacking between two enzyme tyrosine residues, and by hydrogen bonding to the O6 and N1 of the GMP. Hence the single-strand specificity of RNase T1 is due to the catalytic site and not to the presence of a melting site on the protein (*19*).

Despite the fact that RNases A and T1 can cleave the nascent RNAs within a few nucleotides of the 3′-terminus, there is compelling evidence that the enzyme partially protects the nascent RNA from the action of these RNases. The rate of cleavage of RNA in ternary complexes is much slower, and requires higher RNase concentrations, compared to the free U21 or U40 RNAs. In addition, at reduced RNase T1 concentrations cleavage of U40 within 15 nt of the 3′-end ceases to occur, and there is also diminished cleavage at a distance 30−40 nt from the 3′-end. These results are most plausibly explained by the presence of RNA binding sites on the RNA polymerase that partially block these regions. Finally, the results with RNase I on the U40 ternary complex, where there is complete blockage of cleavages shorter than 15 nt from the 3′-end, provide strong support

for the existence of a 3′-proximal RNA binding site on the RNA polymerase. The presence of two 8−10 nt RNA binding sites on RNA polymerase ternary complexes is predicted by the model of Chamberlin, 1995 (*6*).

There are many lines of direct evidence for RNA binding sites on the RNA polymerase (for review, see ref *20*). Recently, studies done on binary complexes (RNA and RNA polymerase) of yeast RNA polymerase II and *E. coli* RNA polymerase have clearly demonstrated that the enzyme binds RNA in an 1:1 stoichiometric ratio (*20*, *21*). In addition, these binary complexes carry out reactions expected of ternary complexes, such as nucleotide addition and transcript cleavage, that argue strongly that the RNA in binary complexes is positioned quite closely to the catalytic site of the RNA polymerase.

REFERENCES

1. Chan, C. L., and Landick, R. (1994) New perspectives on transcript elongation and termination by E. coli RNA polymerase, in *Transcription: Mechanisms and Regulation* (Conaway, J. W., and Conaway, R. C., Eds) pp 297−321, Raven Press, New York.
2. Kerppola, T. K., and Kane, C. M., (1991) RNA polymerase: regulation of transcript elongation and termination, *FASEB J. 5*, 2833−2842.
3. Uptain, S. M., Kane, C. M., and Chamberlin, M. J. (1997) Basic mechanisms of transcript elongation and termination, *Annu. Rev. Biochem. 66*, 117−172.
4. Gamper, H. B., and Hearst, J. E. (1982). A topological model for transcription based on unwinding angle analysis of *E. coli* polymerase binary, initiation and ternary complex, *Cell 29*, 81−90.
5. Yager, T. D., and von Hippel, H. P. (1987) *Escherichia coli and Salmonella typhimurim: Cellular and Molecular Biology* Niedhardt, F. C., Ingram, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarer, H. E., Eds., American Society for Microbiology, Washington, D.C., pp 1241−1275.
6. Chamberlin, M. J. (1995) New models for the mechanism of transcription elongation and its regulation, *Harvey Lect. 88*, 1−21.
7. Rice, G. A., Kane, C. M., and Chamberlin, M. J. (1991) Footprinting analysis of mammalian RNA polymerase II along its transcript: an alternative view of transcription elongation, *Proc. Natl. Acad. Sci. U.S.A. 88*, 4245−4249.
8. Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997) The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase, *Cell 89*, 33−41.
9. Nudler, E., Gusarov, I., Avetissova, E., Kozlov, M., and Goldfarb, A. (1998) Spatial organization of transcription elongation complex in *Escherichia coli*, *Science 281*, 424−427.
10. Gonzalez, N., Wiggs, J., and Chamberlin, M. J. (1977) A simple procedure for the resolution of *Escherichia coli* RNA polymerase holoenzyme from core polymerase, *Arch. Biochem. Biophys. 182*, 404−408.
11. Chamberlin, M. J., Nierman, W. C., Wiggs, J., and Neff, N. (1979) A quantitative assay for bacterial RNA polymerases, *J. Biol. Chem. 254*, 10061−10069.
12. Krummel, B., and Chamberlin, M. J. (1992) Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Individual complexes halted along different transcription units have distinct and unexpected biochemical properties, *J. Mol. Biol. 225*, 221−237.
13. Krummel, B., and Chamberlin, M. J. (1992) Structural analysis of ternary complexes of *Escherichia coli* RNA polymerase. Deoxyribonuclease I footprinting, *J. Mol. Biol. 225*, 239−250.
14. Metzger, W., Schickor, P., and Heumann, H. (1989) A cinematographic view of *Escherichia coli* RNA polymerase translocation, *EMBO J. 8*, 2745−2754.
15. Surratt, C. K., Milan, S. M., and Chamberlin, M. J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription, *Proc. Natl. Acad. Sci. U.S.A. 88*, 7983−7987.
16. Meador, J., Cannon, B., Cannistraro, V. J., and Kennel, D. (1990) Purification and characterization of *Escherichia coli* RNase I. Comparison with RNase M, *Eur. J. Biochem. 187*, 549.
17. Komissarova, N., and Kashlev, M. (1997) RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and RNA, *J. Biol. Chem. 272*, 15329−15338.
18. Jensen, D. E., and von Hippel, P. H. (1976) DNA "melting" proteins., I. Effects of bovine pancreatic ribonuclease binding on the conformation and stability of DNA, *J. Biol. Chem. 251*, 7198−7214.
19. Heinemann, U., and Saenger, W. (1982). Specific protein-nucleic acid recognition in ribonuclease T1-2′ guanylic acid complex: an X-ray study, *Nature 299*, 27−31.
20. Altmann, C., Solow-Cordero, D. E., and Chamberlin, M. J. (1994) RNA cleavage and chain elongation by *Escherichia coli* DNA-dependent RNA polymerase in a binary enzyme-RNA complex, *Proc. Natl. Acad. Sci. U.S.A. 91*, 3784−3788.
21. Johnson, T. L., and Chamberlin, M. J. (1994) Complexes of yeast RNA polymerase II and RNA are substrates for TFIIS-induced RNA cleavage. *Cell 77*, 217−224.

BI9818422