Processivity of Proteolytically Modified Forms of T7 RNA Polymerase†

Daniel K. Muller,‡ Craig T. Martin,§ and Joseph E. Coleman*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510

Received December 15, 1987; Revised Manuscript Received March 17, 1988

ABSTRACT: Two proteolytically modified forms of T7 RNA polymerase have been characterized with respect to transcription initiation and processivity. One species, denoted 80K–20K, is singly cleaved within the region of the polypeptide chain between amino acids 172 and 180. The second species, denoted 80K, is generated by extensive proteolysis of the N-terminal 20K domain by trypsin. The 80K–20K form is fully active in initiation and escape from abortive cycling. It is deficient only in processivity on long DNA templates. Likewise, the 80K species shows initiation kinetics and abortive product synthesis similar to those of the native enzyme. This latter species, however, is unable to escape abortive cycling and shows no synthesis of transcripts longer than about eight bases. Studies of RNA and DNA binding to the three different forms of the enzyme by gel retention assays reveal that the native (98K), the 80K–20K, and the 80K species all form specific complexes with promoter-containing DNA. In addition, the native enzyme binds nonspecifically to double-stranded DNA, while the 80K–20K and 80K enzymes do not. The native enzyme also binds RNA. This RNA binding is reduced in the 80K–20K enzyme and is absent in the 80K species. We suggest a model for T7 RNA polymerase wherein the 20K N-terminal domain of the protein or a shared region between the N- and the C-terminal domains of the protein forms a nonspecific polynucleotide binding site. Binding of the nascent mRNA to such a site may be involved in the observed transition of the native enzyme from the newly initiated complex, susceptible to abortive cycling, to a more stable highly processive ternary complex.

RNA polymerase from the bacteriophage T7 is a small monomeric enzyme which is highly specific for transcription of a set of small, well-conserved promoters within the T7 genome (Chamberlin et al., 1970; Chamberlin & Ring, 1973; Niles et al., 1974; Oakley & Coleman, 1977; Dunn & Studier, 1981). In contrast to the large multisubunit bacterial or eukaryotic RNA polymerases, the T7-encoded enzyme is much less complex both structurally and functionally and requires no zinc or external cofactors for transcription (King et al., 1986). With the cloning and overexpression of T7 gene 1 (the polymerase gene), large amounts of the enzyme have become available for protein chemical studies. These facts combine to make the T7 system an excellent choice in which to study the mechanisms of transcription.

Proteolysis of cloned T7 RNA polymerase has been observed during purification and results in the nicking of the enzyme to produce an 80K- and a 20K-size fragment (Davanloo et al., 1984; Tabor & Richardson, 1985). In a similar fashion, the RNA polymerase from the related bacteriophage T3, which is over 80% homologous to the T7 enzyme, has also been observed to undergo proteolysis (Küpper, 1974). For both enzymes it has been observed that this nicking results in a loss of approximately 80% of the polymerase activity as measured in the standard assay (Oakley et al., 1975). We have recently developed a kinetic assay for initiation by T7 RNA polymerase which allows the direct determination of Michaelis–Menten kinetic parameters from transcription of a short oligonucleotide promoter template (Martin & Coleman, 1987). Initiation kinetics for the native enzyme fit well to the simple steady-state Michaelis–Menten equation:

$$\text{enzyme} + \text{DNA} \overset{K_m}{\leftrightarrow} \text{enzyme–DNA} \xrightarrow{k_{cat}} \text{enzyme} + \text{DNA} + \text{message} \quad (1)$$

The rate constant, k_{cat}, reflects the slowest step or steps in the initiation process. In this assay the production of a well-defined five-base message avoids contributions to the observed incorporation rate from the elongation phase of transcription. Since the five-base message is the predominant product, complications in the measurement of transcription initiation due to premature termination are avoided.

In the current study, we have exploited the single-site proteolysis of T7 RNA polymerase by whole *Escherichia coli* cells (Dunn and Studier, personal communication) to produce a homogeneous preparation of uniquely cleaved "80K–20K" species analogous to that recently studied by Ikeda and Richardson (1987a,b). The latter study concluded that the reduced activity of the singly cleaved enzyme was due to both a lower initiation rate and premature chain termination. We now show that single-site cleavage between the 80K and 20K polypeptides does not affect specific promoter recognition or initiation of transcription; however, the "processive complex" shows decreased processivity relative to the native form. In addition, we have exploited proteolysis by trypsin to generate a new form of the enzyme, denoted "80K", in which the 20K N-terminus of the protein is extensively digested. This proteolyzed form of the enzyme has an only slightly reduced initiation rate, suggesting that the catalytic determinants of transcription in T7 RNA polymerase reside solely in the larger C-terminal polypeptide.

We have recently characterized abortive cycling by native T7 RNA polymerase (Martin et al., 1988). Under normal transcription conditions the enzyme produces many small transcripts (less than eight bases long) for each full-length transcript made. We observe here that all three species of the enzyme synthesize the abortive products but only the native and 80K–20K enzymes produce full-length transcripts. In the

† This work was supported by NIH Grant GM21919-13.

‡ D.K.M. was supported by National Institutes of Health Training Grant GM07223.

§ C.T.M. is the recipient of National Institutes of Health Postdoctoral Fellowship GM10902.

0006-2960/88/0427-5763$01.50/0

80K species, the transition from an initiated to a fully processive ternary complex is inhibited. Comparison of these two proteolytically modified forms of T7 RNA polymerase with the native enzyme provides valuable insight into the domain structure and mechanism of T7 RNA polymerase.

Materials and Methods

Purification. T7 RNA polymerase was purified from *E. coli* strain BL21 containing plasmid pAR1219 (kindly supplied by William Studier and John Dunn, Brookhaven National Laboratories), according to W. Studier as described by King et al. (1986). A molar extinction coefficient of $\epsilon_{280} = 1.4 \times 10^5$ M^{-1} cm^{-1} was used to determine enzyme concentrations (King et al., 1986). Both the preparation of oligonucleotide and the kinetic assays of transcription on the oligonucleotide template were performed as previously described (Martin & Coleman, 1987). Concentration of oligonucleotide and plasmid DNA was determined by the absorbance at 260 nm by assuming an optical density of 1.00 represents 50 μg/mL double-stranded DNA. Double-stranded synthetic oligonucleotides were prepared by heating the single strands to 90 °C and allowing the solution to slowly cool to room temperature.

Proteolysis by Whole Cells. T7 RNA polymerase was digested according to Dunn and Studier (personal communication) to produce the singly-cleaved 80K–20K enzyme form. Briefly, *E. coli* strain HMS174 was grown overnight on agar plates with Luria broth. The cells were then suspended in 50% glycerol to form a stock solution. The cleavage reaction was carried out in 20 mM potassium phosphate, 1 mM EDTA,[1] and 150 mM NaCl with HMS174 cells added to a final OD_{600nm} of 1.5. The mixture was then incubated at 37 °C for approximately 6 hr. The reaction was stopped by centrifugation and filter sterilization of the supernatant. The extent of the cleavage was monitored by SDS–PAGE. We have found this procedure to produce uniform samples of the singly cleaved enzyme.

Proteolysis by Trypsin. Digestion of the enzyme with TPCK-trypsin (Worthington) was carried out in 20 mM potassium phosphate, 1 mM EDTA, and 100 mM NaCl. To produce the 80K enzyme form, a 1:2000 (w/w) ratio of trypsin to protein was incubated at room temperature. Aliquots were removed at the indicated times and spotted into tubes containing a twofold excess of soybean trypsin inhibitor (Sigma) over trypsin.

Analysis of RNA Transcripts. Transcripts produced from whole T7 DNA or from restriction fragments of T7 DNA were analyzed by electrophoresis on a composite gel system according to Chamberlin et al. (1979). The transcripts were made radioactive by including [α-^{32}P]UTP (ICN) in the reaction mixture. The reaction was stopped by the addition of 5 × TAE buffer (1× TAE buffer is 40 mM Tris-HCl, 20 mM sodium acetate, and 2 mM EDTA, pH 7.5) containing 0.1% SDS, and an aliquot was loaded directly onto a 0.4 mm thick 0.5% agarose–2.75% polyacrylamide–0.08% bis(acrylamide) gel buffered in 1× TAE. Small transcripts produced from synthetic oligonucleotides or restriction digests of pIBI24, a T7 promoter containing plasmid (International Biotechnologies Inc., New Haven, CT), were analyzed by electrophoresis on a 0.2 mm thick, 20% acrylamide–7 M urea gel system in 0.5× TBE buffer (where 1× TBE buffer is 89 mM Tris, 89 mM boric acid, and 20 mM EDTA, pH 8.0). After a 20-min incubation in the standard assay buffer (Oakley et al., 1975) with 70 nM enzyme and 110 nM DNA, the samples were made 1 M in ammonium acetate, and 2 μg of tRNA was added. The transcripts were precipitated with 4 volumes of ethanol. After centrifugation the samples were dried and redissolved in 7 M urea (Bethesda Research Laboratories). Alternatively, the reaction was quenched by the direct addition of an equal volume of 90% formamide, 50 mM EDTA, and 0.01% bromphenol blue. Both methods produced the same results. In either case, the samples were heated to 90 °C, quick cooled, and then loaded directly onto a preelectrophoresed gel. All denaturing gels were electrophoresed at approximately 1200 V. These and subsequent gels were vacuum dried and autoradiographed.

Gel Retention Assay To Monitor DNA Binding. Binding between enzyme and DNA was analyzed with a nondenaturing polyacrylamide gel system [for a review, see Hendrickson (1985)]. Electrophoresis of a mixture of protein and DNA results in the separation of free DNA from that bound to protein and hence provides a qualitative assay for DNA binding. The protein–DNA mixtures were incubated at 37 °C at concentrations of protein and DNA as indicated in Figure 3. All binding experiments were carried out in 10 mM potassium phosphate, pH 7.8, 1 mM EDTA, 20 mM NaCl, and 4% glycerol. The reaction mixtures were incubated for 10 min, loaded onto a preelectrophoresed 8% gel, and run at room temperature at 150 V. Both the gel and running buffer contained 1× TBE buffer.

Purification of RNA. A double-stranded synthetic 37 base pair oligonucleotide (see Results) was used as a template to produce large amounts of 20-base RNA transcript. The transcription reaction was carried out in 400 μL of the standard reaction buffer except that all four NTPs were made 2 mM instead of 0.4 mM. For the production of radioactive RNA, 40 μCi of [α-^{32}P]UTP was added. The reaction was carried out at 37 °C for 30 min. After the addition of glycerol and bromphenol blue dye, the reaction mixture was directly loaded onto a 20% nondenaturing polyacrylamide gel buffered in 0.5× TBE. The gel was autoradiographed to locate the 20-base RNA band, which was cut out and electroeluted according to Maniatis et al. (1982). The sample was made 1 M in ammonium acetate, and 4 volumes of 95% ethanol was added to precipitate the RNA. After centrifugation, the pellet was redissolved in 50 μL of 10 mM Tris–1 mM EDTA, pH 7.8, buffer. The concentration was measured by specific activity, if ^{32}P-labeled nucleotides were used in the transcription reaction, or by the absorbance at 260 nm. The two measurements generally agreed within 10%.

Results

Proteolysis of T7 RNA Polymerase. The specific cleavage of T7 RNA polymerase into 80K and 20K fragments during purification of the overproduced enzyme from *E. coli* strain HMS174/pAR1219 has been previously reported (Davanloo et al., 1984). Native (98K) T7 RNA polymerase can be digested in a controlled manner into two polypeptide fragments with molecular weights of approximately 80K and 20K by incubation with *E. coli* HMS174 cells at 37 °C as shown in Figure 1A. To determine the exact position of the cut site, we analyzed the fully cleaved material by gas-phase peptide sequencing. Although sequencing is complicated by the presence of the original N-terminus of the protein, two sequences can be clearly identified and show that the bacterial protease cleaves after Lys-179, in agreement with the cleavage site determined by Dunn and Studier (unpublished observations). We have not observed any significant cleavage at other

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane.

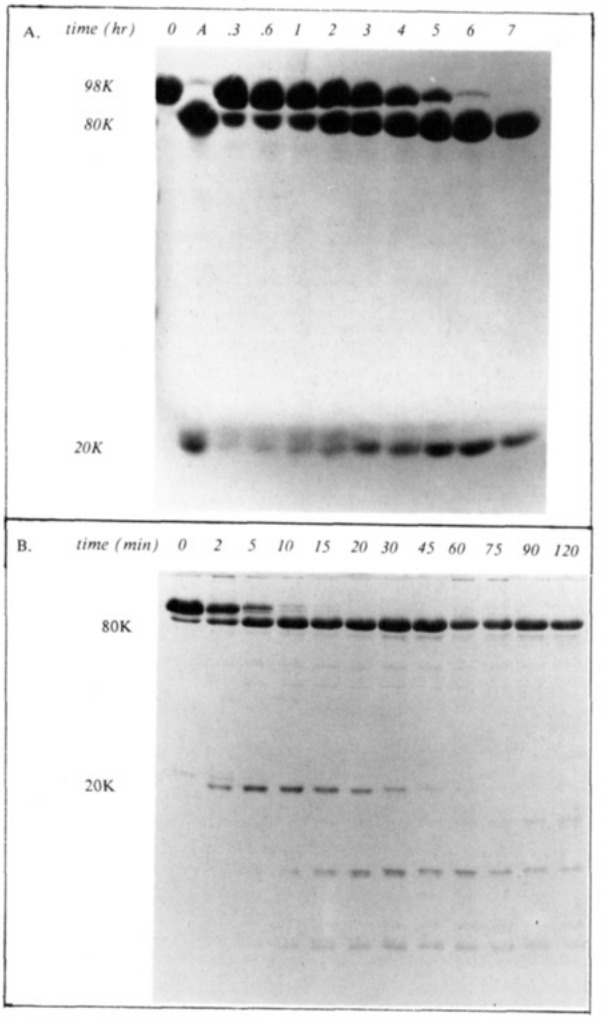

FIGURE 1: Time course of proteolysis of T7 RNA polymerase. (A) Whole *E. coli* (strain HMS174) cells were added to purified T7 RNA polymerase. After 7-h incubation, the enzyme is separated into 80K and 20K polypeptides by SDS–PAGE (13%) due to a single peptide bond cleavage after Lys-179. (B) Digestion of T7 RNA polymerase with trypsin. After 15 min of incubation with trypsin, the enzyme is separated into 80K and 20K polypeptides by SDS–PAGE (15%) due to peptide bond cleavage after Arg-173 and Lys-180 (see text). After 1.5 h there is further digestion with the disappearance of the intact 20K polypeptide. The 80K species shows a N-terminal Ala (181) and has undergone some minor additional cleavage. This may be cleavage near the C-terminus (see text). Lane A contains a standard for 80K and 20K polypeptides, produced by long-term incubation with whole cells.

sites within the protein when the purified enzyme is digested with whole *E. coli* strain HMS174. Tabor and Richardson (1985) have also reported cleavage of the enzyme during purification from *E. coli* strain HMS273, but after Lys-172. We have found that some cleaved enzyme is produced upon cell lysis and purification of the enzyme from *E. coli* BL21. In this case the cleavage occurs in this same region of the polypeptide chain but after Tyr-178 rather than Lys-179 as shown by gas-phase peptide sequencing.

Similar proteolysis of T7 RNA polymerase can be achieved by digestion with trypsin, but with cleavage at slightly different peptide bonds. As shown in Figure 1B, trypsin initially cleaves the polymerase into 80K and 20K polypeptide fragments; however, with time the 20K fragment is slowly degraded. In contrast, the 80K fragment remains mostly uncleaved over the entire time course (Figure 1B). Gas-phase peptide sequencing of the product mixture reveals that after 10 min (when the protein is mostly 80K–20K) there are two predominant proteolysis sites, with cleavage after Arg-173 and after Lys-180. After extensive digestion with trypsin (90 min at room temperature), sequence and gel analyses of the product reveal complete cleavage occurring after Lys-180, with extensive cleavage occurring within the 20K polypeptide.

Although the region of the polypeptide between amino acid residue 172 and residue 179 is readily cleaved by proteases, we have been unable to separate the 20K and 80K fragments by any means short of denaturation. The two domains appear to be held together by strong hydrophobic interactions. In the following, we will refer to uncleaved enzyme as the native form and to the two proteolytically modified forms of T7 RNA polymerase as the 80K–20K form, containing a unique cleavage site following Lys-179 (produced by cleavage with whole cells), and the 80K form, in which the N-terminal 20K fragment (the region of the sequence preceding Ala-181) is degraded by extensive proteolysis with trypsin.

Standard Activity Assay. Polymerase activity on whole T7 DNA was measured for the three different enzyme samples (native, 80K–20K, and 80K) as previously described (Chamberlin & Ring, 1973; Oakley et al., 1975). Native enzyme typically shows a standard activity of 300000 units/mg (1 unit = 1 nmol of ATP incorporated per hour at 37 °C). In this assay, the singly nicked species has been shown to possess approximately 30% the activity of native enzyme (Ikeda & Richardson, 1987a). Our results confirm this finding for the 80K–20K species, with a standard activity on whole T7 DNA of approximately 100000 units/mg. In contrast, the 80K species is completely inactive as judged by this assay (<1% the activity of the native polymerase).

Transcription Products from Long Templates. RNA transcripts produced by the three enzyme species on templates of varying lengths were compared by gel electrophoresis according to Chamberlin et al. (1979). All but 2 of the 17 expected runoff transcripts from whole T7 DNA are predicted to be longer than 10000 bases, and so remain near the top of a 0.5% agarose–1.75% acrylamide gel, as shown for the native enzyme in Figure 2, column 1. In contrast, the 80K–20K species transcribing from the same template produces a distribution of much shorter transcripts. This decrease in average transcript size must result from premature termination by the 80K–20K species and fully accounts for its observed 30% activity in the standard (whole T7 DNA) assay.

Digestion of T7 DNA with the restriction endonuclease *Kpn*I produces a set of DNA fragments in which the predicted runoff transcripts range in length between 100 and 15000 bases. Comparison of the transcription products from this template for the native and 80K–20K species reveals that the distribution of transcripts for the modified form of the enzyme is skewed toward shorter transcripts (Figure 2, columns 2 and 3). Digestion of T7 DNA with *Hpa*II produces templates with runoff transcripts ranging from 50 to 1350 bases (Oakley et al., 1979). With the small *Hpa*II fragments as templates, both the native enzyme and the 80K–20K species produce comparable amounts of all transcripts.

Comparison of these results shows that the 80K–20K species exhibits slightly reduced processivity which is only apparent in synthesis of transcripts longer than about 2000 bases. In contrast, the 80K species produces no detectable transcripts on any of these templates including whole T7 DNA (Figure 2, column 4). Very short transcripts (less than 50 bases), however, are not expected to be observed on this gel system (see below).

Characterization of Promoter Recognition by the Gel Retention Assay. In order to assess the ability of the native, 80K–20K, and the 80K enzymes to form complexes with DNA

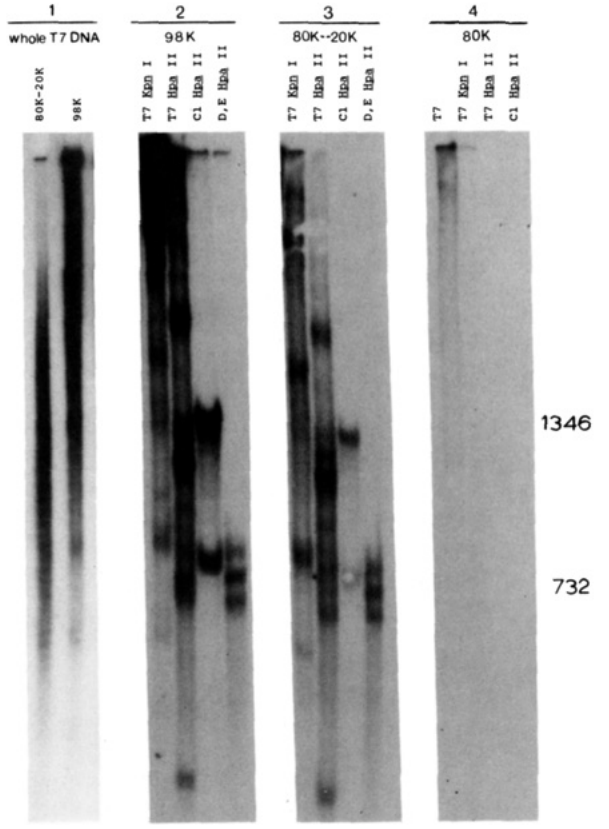

FIGURE 2: Comparison of transcripts produced by native (98K), 80K–20K, and 80K species of T7 RNA polymerase. (Column 1) Whole T7 DNA was the template for the 80K–20K enzyme (left) and native (98K) enzyme (right). Columns 2–4 compare transcripts from native (column 2), 80K–20K (column 3), and 80K (column 4) on various templates derived from T7 DNA. Lanes marked T7 *Kpn*I and T7 *Hpa*II contain the transcripts from synthesis on whole T7 DNA digested with the indicated restriction enzyme. Lanes marked C1 *Hpa*II and D,E *Hpa*II contain the corresponding purified fragment(s) (Dunn & Studier, 1981). The predicted runoff lengths of some of the transcripts are shown on the right. Column 1 is a 0.5% agarose–1.75% polyacrylamide gel; columns 2–4 are 0.5% agarose–2.75% polyacrylamide gels.

containing the T7 promoter sequence, we have utilized a nondenaturing gel electrophoresis system in which the migration of DNA is retarded by complex formation with the enzyme (Hendrickson, 1985). Two 5′-^{32}P-end-labeled synthetic double-stranded DNA constructions were used, the 24 base pair consensus promoter and a second 24 base pair DNA of "random" sequence:

promoter	5′	- ATTAATACGACTCACTATAGGACT - 3′
sequence	3′	- TAATTATGCTGAGTGATATCCTGA - 5′
random	5′	- CTATGTATTCTGTAACTAGATTGC - 3′
sequence	3′	- GATACATAAGACATTGATCTAACG - 5′

Previous studies have suggested that native T7 RNA polymerase cannot bind specifically to promoter in the absence of the initiating nucleotide GTP and Mg^{2+} (Basu & Maitra, 1986; Ikeda & Richardson, 1986). More recent studies, however, have shown that T7 RNA polymerase can bind specifically to promoter in the absence of all nucleotides (Gunderson et al., 1987). The gel in Figure 3A, column 1, shows formation of the complex between native enzyme and the consensus promoter sequence in the absence of nucleotides. There are two relatively rapidly migrating complexes that are not well resolved on an 8% polyacrylamide gel but are easily resolved on a 4% gel (insert to Figure 3A). We will define

Table I: Comparison of Kinetic Parameters of Oligonucleotide Template

	k_{cat} (min^{-1})	K_m (μM)	% active
native	64 (60–68)	0.022 (0.016–0.029)	93 (83–102)
80K–20K	59 (53–66)	0.019 (0.013–0.027)	98 (85–111)
80K	36 (33–40)	0.024 (0.019–0.030)	53 (50–56)

the lower and upper of these two complexes as the fast- and slow-migrating complexes, respectively. The slower moving complex is more readily displaced by a threefold excess of calf thymus DNA than the fast-migrating complex (insert to Figure 3A). However, for the 80K–20K (Figure 3A, column 2) and 80K enzymes (Figure 3B, column 2) only one band appears for the promoter DNA complex, even at high enzyme to DNA ratios. The complex observed for the two proteolyzed species migrates similar to the fast-migrating complex of the native enzyme. At high enzyme to DNA ratios, a third even slower migrating complex is observed with promoter DNA which may represent aggregate. The free DNA sometimes appears to migrate as two species on the gel, with the faster moving of the two probably representing single-stranded DNA. The quantity of this contaminant varied with each annealing reaction without any effect on the protein pattern observed; i.e., the presence of this contaminant is not the cause of the two slowly migrating protein complexes (data not shown).

In addition to binding promoter DNA, the native enzyme also forms a complex with DNA of random sequence, although to a lesser extent than with promoter containing DNA (Figure 3A, column 3). In this case, the complex migrates like the slow complex observed in the promoter–enzyme interaction. This slow-migrating species formed with nonpromoter DNA, like the slow-migrating complex formed with promoter DNA, is readily displaced by a threefold excess of calf thymus DNA. In contrast, the two cleaved forms of the enzyme interact with nonpromoter DNA to a much lesser extent, since no complex is observed between either form and the random sequence oligonucleotide (Figure 3A,B, columns 4).

Binding of RNA to T7 RNA Polymerase. In order to determine whether T7 RNA polymerase binds RNA, we employed a similar gel retention assay. As shown in Figure 4A, native enzyme readily forms a complex with a purified 20-base RNA transcript. Cleavage of a single peptide bond between the 20K and 80K domains severely weakens the RNA binding affinity. The trypsin-cleaved 80K enzyme has no deetectable RNA binding under the conditions used. RNA bound to native enzyme is competitively displaced by the binding of double-stranded promoter DNA (Figure 4B). In contrast, double-stranded nonpromoter DNA competes poorly with RNA binding.

Initiation of Transcription by Proteolytically Modified T7 RNA Polymerase. We have recently developed a steady-state kinetic assay to measure initiation by T7 RNA polymerase. In this system the enzyme transcribes a short synthetic oligonucleotide template to produce a single five-base message. This assay allows direct determination of the Michaelis–Menten parameters K_m and k_{cat} which characterize promoter recognition and initiation (Martin & Coleman, 1987). Kinetic data for the native, 80K–20K, and 80K species are summarized in Table I. Within the error of the measurement, the 80K–20K and native enzyme species possess identical initiation kinetics. The 80K species also initiates with fidelity on this synthetic promoter and produces the five-base transcript (see Figure 5); however, k_{cat} is somewhat lower than that of the native enzyme. Note that this assay can also determine the fraction of inactive enzyme in a sample and that the 80K species typically possessed approximately 50% inactive enzyme[2]

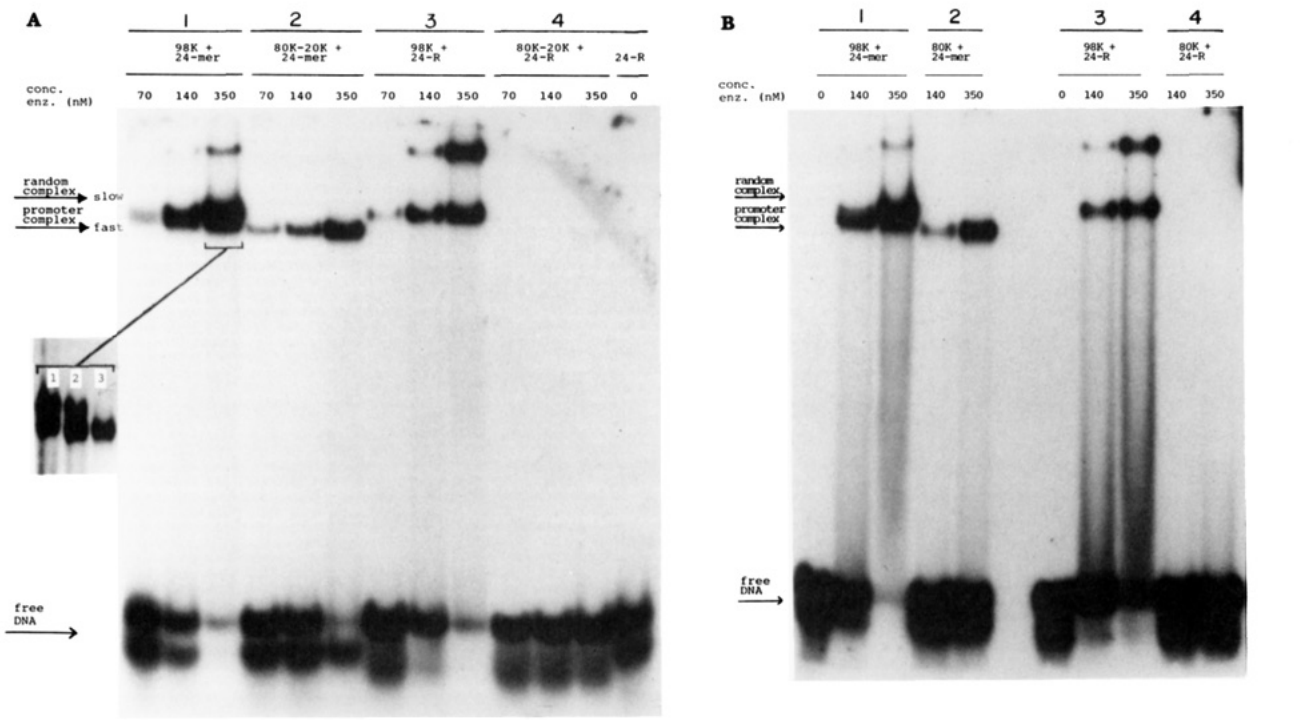

FIGURE 3: Retention of promoter- and nonpromoter-containing DNA by the native (98K), 80K–20K, and 80K species of T7 RNA polymerase. Enzyme–DNA complexes were analyzed on a nondenaturing polyacrylamide gel. (A) (Column 1) Retention of a 24 base pair promoter containing DNA (24-mer) by increasing concentrations (70–350 nM) of native (98K) enzyme on an 8% polyacrylamide gel. The large band at 350 nM enzyme represents two complexes that are easily resolved on a 4% polyacrylamide gel as shown by the insert on the left. The upper of these two complexes is readily displaced by calf thymus DNA as shown by lanes 2 and 3 in the insert representing 1:1 and 3:1 ratios of calf thymus to promoter DNA, respectively. (Column 2) Retention of the 24-mer promoter DNA by increasing concentrations (70–350 nM) of 80K–20K species of T7 RNA polymerase. (Column 3) Retention of a 24 base pair nonpromoter containing DNA (24-R) by increasing concentrations (70–350 nM) of native (98K) enzyme. (Column 4) Retention of the 24-R nonpromoter containing DNA by increasing concentrations (70–350 nM) of the 80K–20K enzyme. The last lane is a control electrophoresis of 24-R alone. (B) Retention of the 24-mer promoter by the native (98K) enzyme (column 1) compared to retention of the 24-mer promoter by the same concentrations of the 80K enzyme (column 2). Retention of the 24-R nonpromoter DNA by the native (98K) enzyme (column 3) compared to retention of 24-R by the same concentrations of the 80K enzyme (column 4). The gel in part B was 8% acrylamide. The concentration of DNA in all samples was 50 nM.

(the native and 80K–20K enzyme forms were always approximately 100% active). This leads to the possibility that the "active" fraction of the 80K species may actually be heterogeneous with respect to initiation kinetics: thus, the data for trypsin-degraded enzyme must be viewed in this light.

As presented below, analysis of transcription products from the oligonucleotide template by gel electrophoresis reveals that a five-base message is the major transcription product for all three enzyme forms. Furthermore, none of the three enzyme species produces transcript when presented with similar duplex DNA of random sequence. These results further demonstrate that both modified forms of T7 RNA polymerase correctly recognize the T7 promoter for initiation of transcription.

Abortive Cycling by 98K, 80K–20K, and 80K Species of T7 RNA Polymerase. We have recently observed that native T7 RNA polymerase produces a significant amount of very short (≤8 bases) transcripts under normal turnover conditions (Martin et al., 1988). The RNA transcripts produced by the native, 80K–20K, and 80K species on the same set of promoter-templates are compared on a 20% acrylamide–7 M urea gel in Figure 5. Each enzyme transcribes the synthetic 24 base pair template to make the expected five-base transcript in similar amounts. The relative production of transcripts by the three forms of the enzyme is different, however, when a synthetic oligonucleotide with a longer message region is used. Shown is the sequence of such a synthetic 37 base pair template designed to produce a 20-base RNA transcription product:

```
-17                +1                +20
5' TAATACGACTCACTATAGGGAAGTCTGTACCAGACGT 3'
3' ATTATGCTGAGTGATATCCCTTCAGACATGGTCTGCA 5'
```

In transcription of this template the native enzyme synthesizes large amounts of five- to eight-base transcripts in addition to the correct runoff transcript (Figure 5, column 1). The 80K–20K species is similar to native enzyme in that it makes both abortive and runoff products. This result is in agreement with the transcription on longer templates presented in Figure 2 but in addition shows that the abortive cycling of the 80K–20K species is essentially identical with that of the native enzyme. In contrast, the 80K species produces very little full-length transcripts and only makes the abortive products. Apparently, the 80K enzyme can initiate RNA synthesis like the native enzyme and undergoes similar abortive cycling; however, it cannot escape abortive cycling.

Transcription from the plasmid pIBI24 digested with the restriction enzyme *Sma*I or *Pvu*II should produce runoff transcripts of 21 and 167 bases, respectively. While the native and 80K–20K enzymes produce the expected runoff transcripts, abortive cycling also occurs on these templates producing relatively large amounts of small transcripts five to eight bases long (Figure 5, column 2). As in the case of the 37 base pair template, the 80K species produces abortive products but synthesizes no observable transcripts longer than eight bases from either template (Figure 5, column 3). The abortive

[2] The inactivation of 50% of the enzyme molecules observed for the trypsin-cleaved enzyme may be due to cleavage of a small fragment of the protein from the C-terminus. Cleavage of the last two amino acids off the C-terminus is known to inactivate the polymerase (J. Dunn, personal communication).

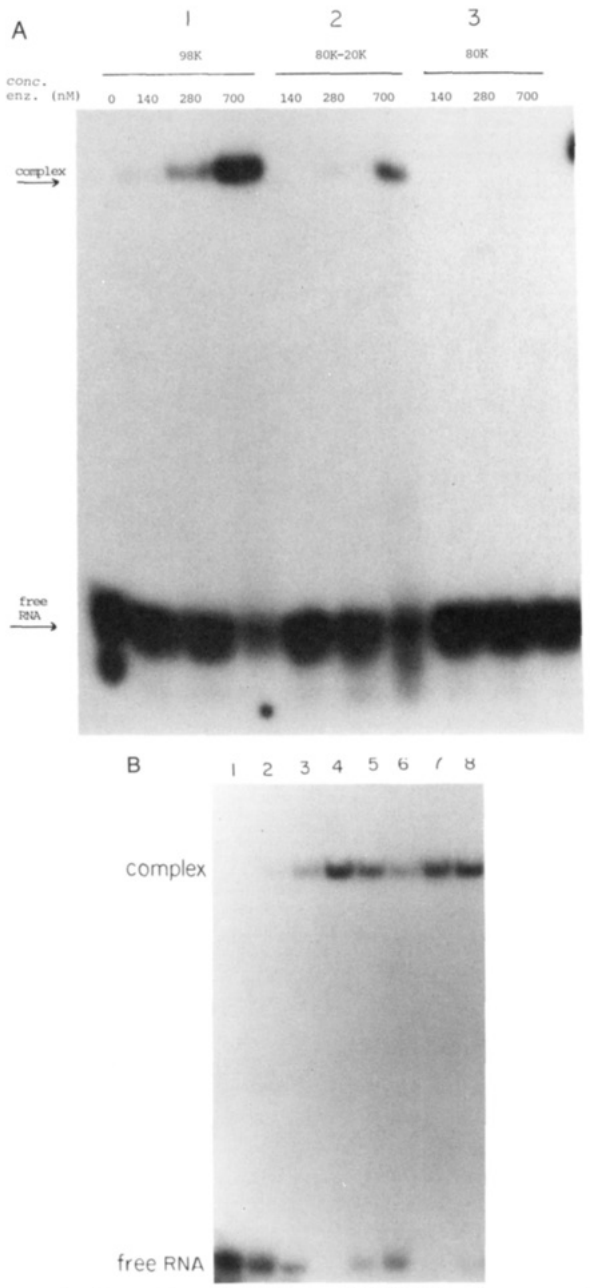

FIGURE 4: Formation of complexes of T7 RNA polymerase with [^{32}P]RNA. A 20-base RNA transcript was synthesized by T7 RNA polymerase with [α-^{32}P]UTP. The transcript was isolated and purified as described under Materials and Methods. The product was used for a gel retention assay with T7 RNA polymerase on a nondenaturing 8% polyacrylamide gel. (A) Retention of [^{32}P]RNA with 0–700 nM native (98K) enzyme (column 1), 140–700 nM 80K–20K enzyme (column 2), and with 140–700 nM 80K enzyme (column 3). (B) Displacement of [^{32}P]RNA from the native (98K) T7 RNA polymerase by double-stranded promoter containing DNA. [^{32}P]RNA (300 nM) was titrated with concentrations of native T7 RNA polymerase of 0, 110, 220, and 550 nM (columns 1–4). The double-stranded 24 base pair promoter containing DNA was then titrated in at 500 nM (column 5) and 1000 nM (column 6). The double-stranded nonpromoter DNA, 24-R, was also titrated in at 500 nM (column 7) and at 1000 nM (column 8).

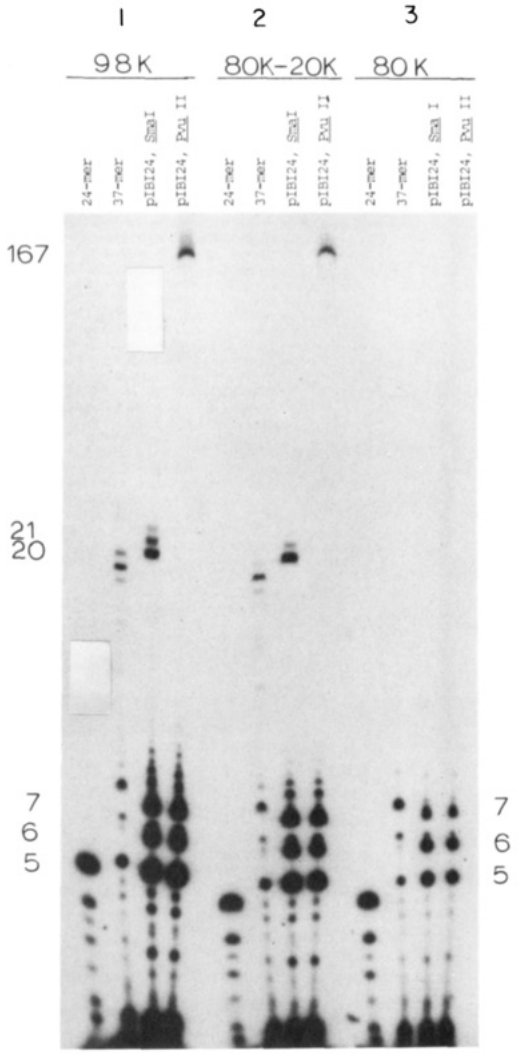

FIGURE 5: Transcripts produced by the native (98K) (column 1), 80K–20K (column 2), and 80K (column 3) forms of T7 RNA polymerase on a variety of templates. The templates are synthetic promoter containing DNAs and restriction digests of a T7 promoter containing plasmid. The 24-mer and 37-mer are dsDNA templates expected to produce a 5-base and 20-base RNA product, respectively (see text for sequences). The transcripts were separated on 20% polyacrylamide–7 M urea sequencing gels. The several bands at 20 ± 1 bases representing the 20-base RNA expected from the 37-mer template represent imprecise "falloff" [see Martin et al. (1988)]. pIBI24 *Sma*I and *Pvu*II are restriction digests of that plasmid and are expected to give runoff transcripts of 21 and 167 bases, respectively. All transcripts were labeled with [γ-^{32}P]GTP in the reaction mixture. The numbers in the margins represent the assigned transcript lengths.

cycling on both the 37 base pair and the plasmid templates shows that the presence or absence of DNA upstream from the promoter does not affect abortive cycling.

A T7 promoter need be double stranded only to position −1 in order to effect completely normal transcription of even very long messages; i.e., the complete coding region can be single stranded (Milligan et al., 1987; Martin et al., 1988). The absence of the noncoding strand also has no effect on the abortive cycling by the native enzyme. This result is illustrated by the data in Figure 6 which compare the transcripts synthesized by the native enzyme (column 1), the 80K–20K species (column 2), and the 80K species (column 3) from the 37 base pair synthetic promoter containing either double- or single-stranded message regions. The similar synthesis of a 5-base message by all three enzymes from the 24 base pair template is shown as a control. The qualitative aspects of transcription by the 80K–20K enzyme are not affected by the absence of the noncoding strand of the DNA. Quantitatively there is an effect in that the native enzyme appears to synthesize more total full-length transcript from the single-stranded than from the double-stranded template (Figure 6, column 1). On the other hand, the total amount of full-length message synthesized on the single-stranded template by the 80K–20K enzyme is reduced compared to that synthesized on the double-stranded template (Figure 6, column 2). The 80K

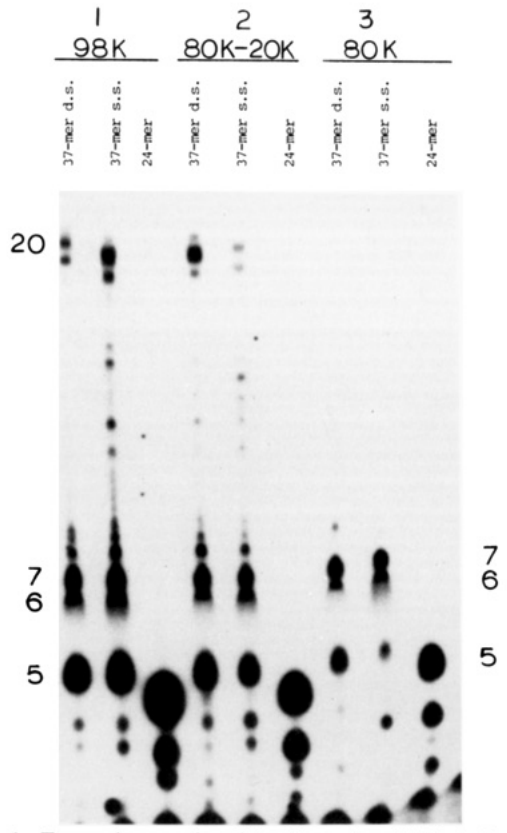

FIGURE 6: Transcripts produced by the native (98K) (column 1), 80K–20K (column 2), and 80K (column 3) forms of T7 RNA polymerase on double-stranded (ds) and single-stranded (ss) DNA templates. The synthetic templates are as described in Figure 5. In all cases the single-stranded templates refer to constructs that were double stranded in the promoter region to position −1, followed by a single-stranded message region beginning at position +1. The transcripts were separated on 20% polyacrylamide–7 M urea sequencing gels. Transcripts from the 24-mer double-stranded synthetic promoter expected to yield a five-base transcript are shown as a control for all three enzymes. The numbers in the margins represent the assigned transcript lengths.

enzyme cannot progress beyond the abortive cycling phase on either double- or single-stranded template (Figure 6, column 3). These results suggest that although the presence of the noncoding strand may very slightly stabilize the processive complex, melting and/or reannealing of the template DNA is not critical to the transformation of the initial complex into a fully processive ternary complex.

Discussion

Limited proteolysis of T7 RNA polymerase, first observed by Davanloo et al. (1984), provides a valuable approach to the study of structure and function in this prototypical RNA polymerase. In the current study, we show that controlled cleavage with either trypsin or an endogenous bacterial protease (located on the *exterior* of the bacterial cell wall) divides T7 RNA polymerase into two tightly associated polypeptide fragments with M_r ~20K and ~80K, respectively [see also Tabor and Richardson (1985) and Ikeda and Richardson (1987a,b)]. The facile cleavage of the polypeptide between residues 173 and 179 suggests that this small region of the polypeptide chain is exposed near the surface of the protein. More extensive proteolytic degradation of the enzyme by trypsin results in substantial digestion of the 20K N-terminal fragment of the protein with very little degradation of the 80K fragment (Figure 1B). This result suggests that the 20K polypeptide may represent a distinct domain of the enzyme. Comparison of the transcriptional properties of the singly cleaved enzyme (80K–20K) and the extensively digested enzyme (80K) with those of the native enzyme provides insight into the relationship between structure and function in T7 RNA polymerase.

Early Stages of Transcription. We have recently developed a kinetic assay which allows the determination of K_m and k_{cat} for initiation of transcription (Martin & Coleman, 1987). The kinetic data summarized in Table I show that the steady-state K_m and k_{cat} for the 80K–20K species are both unchanged from those of the native enzyme and further, after correction for inactive enzyme, the 80K species possesses a similar K_m and an only slightly reduced k_{cat}. These data along with the gels of the transcripts produced from the 24 base pair template suggest that the 80K C-terminal domain contains all the essential determinants required for promoter recognition and initiation (Table I and Figures 5 and 6). Consistent with this conclusion are the results of genetic recombination experiments which produce hybrid T7 and T3 RNA polymerases (Ryan & McConnell, 1982; Morris et al., 1986). These studies show that base sequences from 0.25 to 0.6 in the polymerase gene and a short sequence at the very end of the gene code for the amino acid sequences responsible for the *differential* promoter recognition exhibited by these two closely related enzymes. Further confirmation comes from photochemical cross-linking of the complex between the 24 base pair promoter and 80K–20K complex, which results in cross-links to the 80K fragment only (Muller and Coleman, unpublished observations). It has also recently been shown that a reactive substrate analogue labels the 80K fragment (Schäffner et al., 1987).

We have shown that the 80K species can initiate transcription similar to the native enzyme. Therefore, the differences observed between the native enzyme and the 80K species must be due to aspects of transcription other than initiation. In a separate paper we have recently described a detailed study of abortive cycling by T7 RNA polymerase which shows that during normal transcription the native enzyme synthesizes a number of abortive transcripts (less than eight bases in length) for each full-length RNA that is produced (Martin et al., 1988). For example, in transcription of the 37 base pair template which codes for a 20-base message, the native enzyme makes 5-, 6-, and 7-base abortive products. The data presented in Figure 5 show that both the 80K–20K and 80K species produce abortive products similar to those produced by the native enzyme. However, in contrast to the other two forms of the enzyme, the 80K species is not able to escape abortive cycling and form the fully processive ternary complex. Hence, the 20K domain must participate in the transition of the enzyme complex to its fully processive form. Possible models for this interaction are discussed below.

DNA Binding Studies. Detection of the complexes formed between promoter DNA and the native, 80K–20K, and 80K forms of T7 RNA polymerase by the gel retention assay further supports the conclusion that the specific promoter-binding site and active center are located within the 80K domain. Although these gels represent a complicated combination of equilibrium and kinetic phenomena, their sensitivity often reveals several kinds of complexes which may form between a given protein and the DNA [for a review, see Hendrickson (1985)]. Several complexes with promoter-containing DNA have been identified for RNA polymerase from *E. coli* and have been assigned to open and closed forms of the promoter complex (Straney & Crothers, 1985). In those studies the closed complexes are operationally defined as being easily displaced by nonspecific DNA.

Native T7 RNA polymerase forms two enzyme–promoter

complexes which migrate close together on the gel (Figure 3A). Of these two complexes, the fast-migrating complex is not easily displaced by calf thymus DNA, whereas the slow-migrating one is readily competed away. The fast-migrating complex is formed by all three forms of the enzyme and may correspond to a specifically bound open promoter complex. Evidence for the existence of an open promoter complex in solution has been obtained by endonuclease digestion of enzyme–promoter complexes under similar conditions (Strothkamp et al., 1980; Osterman & Coleman, 1981). Although the slow-migrating complex may meet the operational definition of a closed promoter complex, this complex is not observed for either modified form of the enzyme, despite their abilities to correctly initiate transcription and therefore to form the open complex. A complex which migrates in the same position is observed between nonpromoter-containing DNA and the native enzyme but not for either modified enzyme species (Figure 3). As discussed below, the slow-migrating complex formed with promoter DNA may represent a nonspecific DNA binding mode of the native enzyme.

Transcription on Long Templates. In order for the enzyme to transcribe DNA, it must bind nonspecifically to DNA as it processes along the template making the transcript. Although the 80K–20K form of T7 RNA polymerase has only one peptide bond hydrolyzed, it is nevertheless less processive than the native enzyme, resulting in premature termination on long transcripts and a shorter average transcript length (Figure 2). Ikeda and Richardson (1987a) have suggested that both reduced initiation and premature termination contribute to the observed decrease in standard activity (on long templates). Our findings that the K_m and k_{cat} describing initiation for the 80K–20K and native enzymes are identical (Table I) and that the 80K–20K species is unchanged in its ability to escape abortive cycling suggest that the observed decrease in activity in the standard assay is solely due to premature chain termination. Presumably, the tertiary relationship between the 20K N-terminal and the 80K C-terminal regions is altered in the singly cleaved enzyme, resulting in a small increase in the probability of dissociation of the processive ternary complex.

This decrease in the ability of the processive 80K–20K complex to remain tightly bound to the template may also explain the observation that the 80K–20K enzyme does significantly less well in producing full-length transcripts from the single-stranded 37 base pair template than it does from the double-stranded 37 base pair template. Exactly the opposite is true of the native 98K enzyme (Figure 6). If weak interactions with the noncoding DNA strand add stability to the ternary complex, then the loss of these contacts could well affect the processivity of the already weakened 80K–20K complex more than the ternary complex of the native enzyme.

RNA Binding by T7 RNA Polymerase. It seems likely that one of several forces stabilizing the enzyme–DNA–RNA ternary complex is a direct interaction of the nascent RNA with groups within the enzyme. The results of the gel retention assay show that exogenous RNA (and nonspecific DNA) does bind to the native enzyme (Figures 3 and 4). The cleavage of a single polypeptide bond between the 80K and 20K domains significantly reduces this binding, while the 80K species appears to have no affinity for exogenous RNA or nonspecific DNA (Figures 3A and 4A). In both cases, specific promoter binding is not significantly reduced, suggesting that interactions distinct from the promoter binding site are affected by proteolysis. It is possible that determinants within the 20K polypeptide participate in forming a sequence-independent polynucleotide binding site, which serves to stabilize the growing RNA chain during processive transcription. The tight binding of exogenous RNA, however, would require interactions with both this site and regions closer to the active site, since promoter binding does displace exogenously bound RNA. During transcription, the growing RNA would span these two sites. It is possible that the transition to a more highly processive elongation complex which follows incorporation of about eight nucleotides reflects binding of the emerging RNA to a nonspecific polynucleotide binding site within the 20K polypeptide. Such a model would explain the decreased processivity observed for the two proteolytically modified forms of T7 RNA polymerase.

ACKNOWLEDGMENTS

We thank the Protein and Nucleic Acid Chemistry Facility at Yale University for performing the peptide sequencing.

Registry No. RNA polymerase, 9014-24-8; consensus promoter, 114956-78-4.

REFERENCES

Basu, S., & Maitra, V. (1986) *J. Mol. Biol. 190*, 425–457.
Chamberlin, M. J., & Ring, J. (1973) *J. Biol. Chem. 248*, 2235–2244.
Chamberlin, M. J., McGrath, J., & Waskell, L. (1970) *Nature (London) 228*, 227–231.
Chamberlin, M. J., Nierman, W. C., Wiggs, J., & Neff, N. (1979) *J. Biol. Chem. 254*, 10061–10069.
Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.
Dunn, J. J., & Studier, F. W. (1981) *J. Mol. Biol. 148*, 303–330.
Gunderson, S. I., Chapman, K. A., & Burgess, R. R. (1987) *Biochemistry 26*, 1539–1546.
Hendrickson, W. (1985) *BioTechniques May/June*, 198–207.
Ikeda, R. A., & Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3614–3618.
Ikeda, R. A., & Richardson, C. C. (1987a) *J. Biol. Chem. 262*, 3790–3799.
Ikeda, R. A., & Richardson, C. C. (1987b) *J. Biol. Chem. 262*, 3800–3808.
King, G. C., Martin, C. T., Pham, T. T., & Coleman, J. E. (1986) *Biochemistry 25*, 36–40.
Küpper, H. (1974) Ph.D. Thesis, University of Heidelberg, West Germany.
Maniatis, T., Fritsch, E. F., & Sambrook J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Martin, C. T., & Coleman, J. E. (1987) *Biochemistry 26*, 2690–2696.
Martin, C. T., Muller, D. K., & Coleman, J. E. (1988) *Biochemistry 27*, 3966–3974.
Milligan, J. F., Groebe, D. C., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.
Morris, C. E., McGraw, N. J., Joho, K., Brown, J. E., Klement, J. F., Ling, M. L., & McAllister, W. T. (1986) in *RNA Polymerases and the Regulation of Transcription* (Reznikoff, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, T., & Wickens, M. D., Eds.) pp 47–58, Elsevier, New York.
Niles, E. G., Conlon, S. W., & Summers, W. C. (1974) *Biochemistry 13*, 3904–3912.
Oakley, J. L., & Coleman, J. E. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4266–4270.
Oakley, J. L., Pascale, J. A., & Coleman, J. E. (1975) *Biochemistry 14*, 4684–4691.
Oakley, J. L., Strothkamp, R. E., Sarris, A. H., & Coleman,

J. E. (1979) *Biochemistry 18*, 528–537.
Osterman, H. L., & Coleman, J. E. (1981) *Biochemistry 20*, 4884–4892.
Ryan, T., & McConnell, D. J. (1982) *J. Virol. 43*, 844–858.
Schäffner, A. R., Jorgensen, E. D., McAllister, W. T., & Hartmann, G. R. (1987) *Nucleic Acids Res. 15*, 8773–8781.
Straney, D. C., & Crothers, D. M. (1985) *Cell (Cambridge, Mass.) 43*, 449–459.
Strothkamp, R. E., Oakley, J. L., & Coleman, J. E. (1980) *Biochemistry 19*, 1074–1080.
Tabor, S., & Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1074–1078.

Characterization of the Lead(II)-Induced Cleavages in tRNAs in Solution and Effect of the Y-Base Removal in Yeast tRNA^{Phe} †

W. J. Krzyzosiak,‡ T. Marciniec,‡ M. Wiewiorowski,‡ P. Romby,§ J. P. Ebel,§ and R. Giegé*,§

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland, and Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 Rue René Descartes, 67084 Strasbourg, France

Received November 10, 1987; Revised Manuscript Received February 16, 1988

ABSTRACT: The specificity of lead(II)-induced hydrolysis of yeast tRNA^{Phe} was studied as a function of concentration of Pb^{2+} ions. The major cut was localized in the D-loop and minor cleavages were detected in the anticodon and T-loops at high metal ion concentration. The effects of pH, temperature, and urea were also analyzed, revealing a basically unchanged specificity of hydrolysis. In the isolated 5′-half-molecule of yeast tRNA^{Phe} no cut was found in the D-loop, indicating its stringent dependence on T–D-loop interaction. Comparison of hydrolysis patterns and efficiencies observed in yeast tRNA^{Phe} with those found in other tRNAs suggests that the presence of a U59–C60 sequence in the T-loop is responsible for the highly efficient and specific hydrolysis in the spatially close region of the D-loop. The efficiencies of D-loop cleavage in intact yeast tRNA^{Phe} and in tRNA^{Phe} deprived of the Y base next to the anticodon were also compared at various Pb^{2+} ion concentrations. Kinetics of the D-loop hydrolysis analyzed at 0, 25, and 37 °C showed a 6 times higher susceptibility of tRNA^{Phe} minus Y base (tRNA^{Phe}-Y) to lead(II)-induced hydrolysis than in tRNA^{Phe}. The observed effect is discussed in terms of a long-distance conformational transition in the region of the interacting D- and T-loops triggered by the Y-base excision.

A number of studies have shown that tRNAs can undergo conformational changes, although they tend to modulate their spatial configuration rather than undergoing dramatic conformational rearrangements [e.g., Gassen (1980) and Moras et al. (1985, 1986) and references cited therein]. For instance, removal of the Y base in the anticodon region in yeast tRNA^{Phe} induces long-range conformational changes in this molecule [e.g., Cameron and Uhlenbeck (1973), Kearns et al. (1973), Wang et al. (1975), Davanloo et al. (1979), and Salemink et al. (1979)], a fact substantiated by an increased accessibility of its T-loop, leading to the modification of cytosine-60 by chloroacetaldehyde (Krzyzosiak & Ciesiolka, 1983). According to crystallographic studies of tRNA^{Phe}, C60 was found to be the closest ligand for lead ion binding (Brown et al., 1983, 1985; Rubin & Sundaralingam, 1983), which leads to specific cleavage of the tRNA in a spatially close region of the D-loop.

Considering these facts, and in agreement with the first observations by Werner et al. (1976), it seemed worthwhile to determine whether the excision of the Y base from yeast tRNA^{Phe} would have any influence on the specificity of the Pb^{2+}-promoted cleavage reaction or, in other words, would be sensitive enough to monitor conformational changes in the altered tRNA^{Phe}. However, the relative lack of information concerning kinetic and biochemical characteristics of the cleavage reaction in tRNA^{Phe} prompted us to examine the influence on the rate and specificity of this reaction of parameters such as the concentration of Pb^{2+} ions, pH, temperature, and tRNA structure. Once these variables had been optimized, the relative reactivities of yeast tRNA^{Phe} and tRNA^{Phe}-Y[1] were studied. The results will be discussed in the light of long-distance conformational transitions between the anticodon and the T–D regions in tRNAs, in relation to mechanistic aspects proposed for Pb^{2+}-induced strand scission in yeast tRNA^{Phe} as well as in the perspective of Pb^{2+} ions as a structural probe to test conformational features in tRNA and RNA molecules.

Materials and Methods

tRNAs. Yeast tRNA^{Phe}, tRNA^{Asp}, and tRNA^{Val} were prepared from crude brewer's yeast tRNA (Boehringer) by countercurrent distribution (Dirheimer & Ebel, 1967) followed by column chromatographies (Keith et al., 1971). In the case of tRNA^{Asp}, purification was achieved by chromatography on a Sepharose 4B column (Giegé et al., 1986a). Pure *Escherichia coli* tRNA^{Val} was from Boehringer. Yeast tRNA^{Phe} was a gift from P. Remy. The 5′-half-molecule of tRNA^{Phe} was prepared according to Phillipsen et al. (1968) by aniline/acetic acid (pH 4.5) treatment of 5′-end-labeled tRNA^{Phe}-Y.

† This work was partly supported by the Centre National de la Recherche Scientifique (CNRS, Paris), the Université Louis Pasteur (ULP) de Strasbourg, and the Polish Academy of Sciences (Project 04.12.1.3). W.J.K. was supported by two grants from FEBS and EMBO.

* To whom correspondence should be addressed.
‡ Institute of Bioorganic Chemistry.
§ Institut de Biologie Moléculaire et Cellulaire du CNRS.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; tRNA^{Phe}-Y, phenylalanine-specific tRNA minus Y base.

0006-2960/88/0427-5771$01.50/0