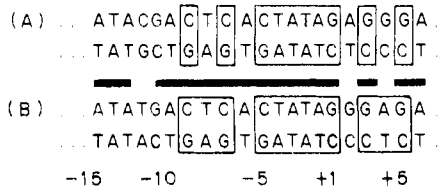
T7 RNA Polymerase: Promoter Structure and Polymerase Binding†

John L. Oakley, Ronald E. Strothkamp, Andreas H. Sarris, and Joseph E. Coleman*

ABSTRACT: The sequences of two promoters recognized by the phage-specified T7 RNA polymerase are presented. The two are identical in sequence but for one base pair from the initiation point (as determined by the 5′ sequence of the transcripts), denoted +1, to position −15. The common



sequence also includes a region of hyphenated twofold symmetry indicated by the boxes, with the twofold axis as the center of the six base-pair box. The heavy line indicates the extent of homology. The first promoter (A) is demonstrated to lie within gene 1, the gene for the polymerase itself, and 40 bases into the message transcribed from this promoter is found the RNase III site separating genes 1 and 1.1. Binding of T7 RNA polymerase to these promoters is associated with a hyperchromic blue shift of the base chromophores consistent with partial melting of the base pairs at the promoter. Binding of T7 RNA polymerase to these promoters disappears at low pH and low temperature and is accompanied by a consequent loss of polymerase activity. The pH dependence of the binding step is adequately described by a single pK of 7.0. Polymerase catalytic activity, but not promoter binding, requires a single free sulfhydryl group of the enzyme with a pK_a of ~7.8.

T7 RNA polymerase is a single polypeptide chain of 107 000 which initiates specific transcription both in vivo and in vitro from eight or nine major promoters located in the rightmost 85% of the phage T7 genome (Chamberlin & Ring, 1973; Golomb & Chamberlin, 1974; Niles et al., 1974; Oakley et al., 1975; Oakley & Coleman, 1977). With the exception of zinc, additional cofactors or polypeptide chains are not required (Coleman, 1974). The enzyme is the product of phage gene 1 which consists of ~2500 base pairs extending from ~8.5 to 15% of the genome and is transcribed by the *Escherichia coli* RNA polymerase which terminates at ~20% of the genome.

We have now isolated two of these promoters contained on small Hpa II restriction fragments from T7 DNA which are 138 and 140 base pairs in length. Both form active transcription complexes with the T7 polymerase and direct the synthesis of truncated mRNAs of 56 and 85 bases, respectively. The primary structures of these promoters, the sequences of the transcripts, and additional properties of the enzyme and its DNA complexes are described in this paper. The structure of one of these promoters has been described in a preliminary report (Oakley & Coleman, 1977).

Materials and Methods

Chemicals. [γ-^{32}P]GTP was from Amersham Corp., Arlington Heights, IL. [γ-^{32}P]ATP, [α-^{32}P]GTP, and [^{14}C]iodoacetamide were purchased from New England Nuclear Corp., Boston, MA. [^{3}H]ATP was purchased from Schwarz/Mann, New York, NY.

T7 Bacteriophage. Strain ΔH3 kindly supplied by William Summers was used to prepare T7 RNA polymerase and in the investigation of transcription from Hinf fragments of T7. Strain d14, used for all other experiments, was also from Dr. Summers. Pure T7 phage and DNA were prepared as previously described (Oakley et al., 1975).

Protein concentration was determined by absorbance at 280 nm using $E_{280}^{0.1\%}$ 0.74 for T7 RNA polymerase (Niles et al., 1974).

Spectroscopy. Ultraviolet spectra were taken on a Cary Model 118 spectrophotometer with thermostated cells at a full-scale reading of 0.02 ± 0.001 OD unit.

Enzymes. T7 RNA polymerase and Hpa II restriction endonuclease were prepared and assayed as previously described (Oakley et al., 1975). Hae III and Hinf restriction endonucleases were purchased from New England Biolabs. Bacterial alkaline phosphatase was the gift of J. F. Chlebowski. T4 polynucleotide kinase was a gift of J. A. Steitz or purchased from Bethesda Research Laboratories, Bethesda, MD.

^{14}C Labeling of T7 RNA Polymerase. T7 RNA polymerase was labeled with [^{14}C]iodoacetamide by a method analogous to that described by Applebury et al. (1970) for alkaline phosphatase in which $H^{32}PO_4^{2-}$ was the labeling agent.

Tris buffer (10 mM) was used for the pH range 7–11.6 and phosphate (10 mM) for pH 4.5–7. The reaction was carried out at 10^{-3} M [^{14}C]iodoacetamide at 4 °C for 3 h.

Nitrocellulose Filter Binding Assays. Filter binding assays to detect complexes between T7 DNA and T7 RNA polymerase were performed according to a modification of the method of Jones & Berg (1966), as previously described (Oakley et al., 1975).

Gel Electrophoresis of Duplex DNA Fragments. Polyacrylamide gel electrophoresis was carried out in slab gels in either TPE[1] buffer (0.04 M Tris, 0.03 M NaH_2PO_4, 0.001 M EDTA, pH 8.5) or TBM buffer (0.045 M Tris-borate, 2.5 mM $MgCl_2$, pH 8.3). DNA bands were located either by soaking the gel in buffer plus 0.5 μg/mL of ethidium bromide and visualizing with ultraviolet light, or by autoradiography for labeled DNA.

Purification of DNA Fragments from Acrylamide Gels. Bands were excised from acrylamide gels and the DNA was eluted and concentrated by one of two methods.

Electrophoretic extraction was performed for large fragments (>200 base pairs) by placing gel slices in a 10-mL disposable pipet plugged with siliconized glass wool. A dialysis bag containing about 1–2 mL of buffer was placed over the end, the pipet filled with buffer, and the DNA electrophoresed into the bag at 150 V for 12 h. The buffer used was 0.05 M

† From the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Conneticut 06510. *Received July 27, 1978.* This work was supported by National Institutes of Health Grants GM 21919-04 and AM 09070-13 and National Science Foundation Grant GB 43481. R.E.S. was supported by NIH Fellowship CA06135. A.H.S. was supported by the Life Insurance Medical Research Fund.

[1] Abbreviations used: TPE buffer, 0.04 M Tris, 0.03 M NaH_2PO_4, 0.001 M EDTA, pH 8.5; TBM buffer, 0.045 M Tris-borate, 2.5 M $MgCl_2$, pH 8.3.

0006-2960/79/0418-0528$01.00/0

Tris-borate, 0.001 M EDTA, pH 8.3. Following extraction the contents of the bag were brought to 0.3 M sodium acetate and the DNA was precipitated by addition of 2.5 vol of cold ethanol and chilling to −70 °C. The DNA was pelleted by centrifugation (12000*g*, 30 min), washed with ethanol, dried, and resuspended.

"Crush and soak" elutions in high salt buffer were performed for small DNA and RNA species according to a modification of the procedure described by Maxam & Gilbert (1977). Gel slices were crushed and elution buffer (0.5 M ammonium acetate, 0.01 M magnesium acetate, 0.1% sodium dodecyl sulfate, and 0.1 mM EDTA) was added. Approximately twice the volume of the crushed gel was used. Elution was carried out at 37 °C for 12 h. Gel pieces were filtered out with siliconized glass wool in a disposable syringe, and the DNA in the filtrate precipitated with 2.5 vol of ethanol and chilling to −70 °C. After centrifugation (12000*g*, 30 min) the pellet was dried, resuspended in 0.3 M sodium acetate, and reprecipitated. The final pellet was washed with ethanol, dried, and resuspended. Fragments were used without further purification.

Gel Electrophoresis of RNA Transcripts. Transcripts were labeled with either [γ-^{32}P]GTP, [α-^{32}P]GTP, or ATP in the normal assay mix (Oakley et al., 1975). Whole DNA or isolated restriction fragments were added as template. After incubation, reaction mixtures were deproteinized, and the RNA was precipitated with ethanol. The pellet was resuspended in electrophoresis buffer and loaded onto slab gels. Transcripts were run in one of three gel systems. (a) Agarose–acrylamide gels: 2.5% acrylamide–0.5% agarose gels were prepared in 0.04 M Tris, 0.02 M sodium acetate, 0.002 M EDTA, pH 7.5, and run in this buffer plus 0.2% sodium dodecyl sulfate, according to Summers (1969). (b) Acrylamide gels: 10% acrylamide gels were prepared and run in the same buffer. (c) Acrylamide–urea gels: denaturing, 12% acrylamide gels were prepared and run according to Heyden et al. (1972). Pelleted RNA transcripts were heated at 60 °C for 5 min in the gel buffer before loading.

In Vitro Transcription of Restriction Fragments. Restriction fragments were eluted from polyacrylamide gels as described by Maxam & Gilbert (1977). Transcription of these fragments by T7 RNA polymerase was carried out in a standard assay mix (Oakley et al., 1975) containing [α-^{32}P]GTP or [γ-^{32}P]GTP.

DNA Sequencing. Labeling of the 5′ end, strand separation, and sequence determination were according to Maxam & Gilbert (1977). Certain regions of these sequences were confirmed by two-dimensional separation of the products of partial digestion with snake venom phosphodiesterase and pancreatic DNase (Maniatis et al., 1975).

RNA Sequencing. The nucleotide sequences of transcripts made with isolated DNA restriction fragments were determined according to the method of Donis-Keller et al. (1977). RNA species were end-labeled with T4 polynucleotide kinase (Donis-Keller et al., 1977) after removal from a gel or labeled during transcription with [γ-^{32}P]GTP.

Results

Activity of T7 RNA Polymerase, DNA Binding, and Alkylation of the Single Sulfhydryl Group of the Enzyme as Functions of pH. Binding of T7 polymerase to T7 DNA as determined by the filter-binding technique is pH sensitive (Figure 1A). Very little binding is observed at pH 6 and binding increases rapidly from pH 6 to 9 along a sigmoid curve which is adequately described by a single ionization process with an apparent pK_a of ~7. No binding of DNA occurs in

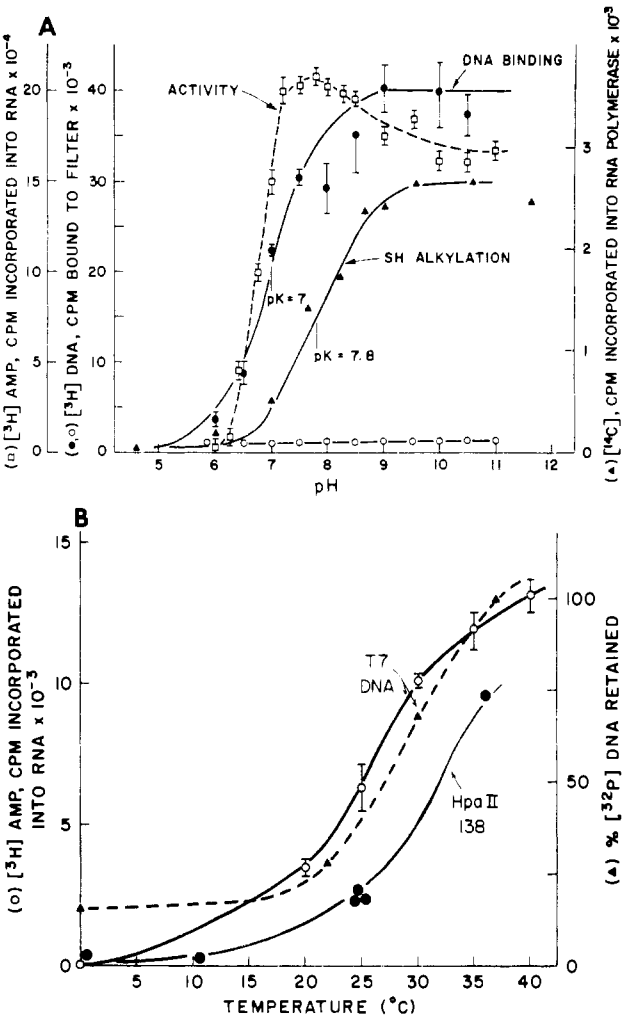


FIGURE 1: (A) pH dependencies of polymerization activity (□), DNA binding (●), and alkylation of the sulfhydryl group (by iodoacetamide) (▲) of T7 RNA polymerase. (○) Filter binding of ^{32}P-labeled T7 DNA in the absence of T7 RNA polymerase. Conditions as in Oakley et al. (1975). (B) Effects of temperature on the binding to T7 DNA and activity of T7 RNA polymerase. (▲) Percent of [^{32}P]T7 whole DNA retained on nitrocellulose filters by native T7 RNA polymerase. (○) Polymerization activity with whole T7 DNA as template. (●) Polymerization activity with Hpa II 138 as template. Binding assays were performed with 5 μg of polymerase/0.25 mL, sufficient to specifically bind all the DNA at 37 °C but not to induce appreciable nonspecific binding. The specific activity of the [^{3}H]ATP was increased for assay of transciption from Hpa II 138 over that used with T7 DNA in order to enhance sensitivity.

the absence of enzyme throughout the pH range. Activity of the enzyme using whole T7 DNA as the template shows a much steeper dependence on pH, sigmoid to high pH, again with a midpoint near pH 7, but the curve cannot be described by a single ionization (Figure 1A).

Previous work from this laboratory (Oakley et al., 1975) has shown that alkylation of the single titratable SH group of the enzyme abolishes activity. In order to estimate the pK_a of this sulfhydryl group, the pH dependency of the alkylation reaction was determined by measuring ^{14}C incorporation into the enzyme from [^{14}C]iodoacetamide as a function of pH. Incorporation follows a pH function sigmoid to high pH with a midpoint of 7.8. The reaction is describable by a single ionization process. Because of the limited amount of enzyme available, this experiment was done by allowing the reaction to proceed for an extended time, 3 h, rather than the more accurate method of measuring initial reaction rates. Hence 7.8 represents an upper limit for the pK_a of the sulfhydryl. The

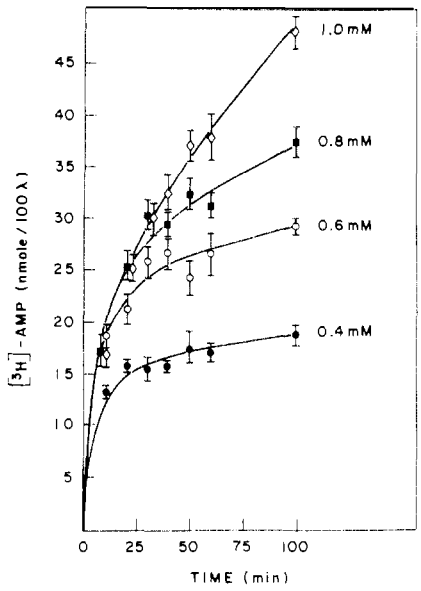


FIGURE 2: Polymerization activity of T7 RNA polymerase, [^{3}H]AMP incorporation, as a function of NTP concentrations: (●) 0.4 mM, (○) 0.6 mM, (■) 0.8 mM; (◇) 1.0 mM. Standard assays were carried out at pH 8, 37 °C, and concentrations refer to that for each nucleoside triphosphate and were the same for all four.

true pK_a may be significantly lower.

Temperature Dependence of DNA Binding and Activity of T7 RNA Polymerase. We have previously shown the steep temperature dependence of the activity of the T7 RNA polymerase using whole T7 DNA as template. Activity is lost completely at ice temperature. Loss of activity as the temperature is lowered is paralleled by the loss of binding of the enzyme to the DNA as determined by the filter-binding technique (Figure 1B). The isolation of small restriction fragments containing single functional promoters for T7 polymerase will be described below. Data on the temperature dependence of polymerase activity on one of these, Hpa II 138, producing the 56 base long mRNA is given along with the data on whole T7 DNA in Figure 1B. The temperature dependence of this activity parallels that observed when whole T7 DNA is the template.

Transcription by T7 RNA Polymerase as a Function of NTP Concentration. The standard assay conditions for T7 RNA polymerase employ 0.4 mM ribonucleoside triphosphates in volumes of 0.1 to 1.0 mL (Chamberlin & Ring, 1973). Incorporation of nucleotides by this enzyme is so rapid, however, that approximately half of the total amount of NTPs may be depleted in ~10 min in a 1-mL reaction under the conditions shown in Figure 2. The reaction rate shows a rapid fall off by 10 min (Figure 2). This is not reaction inactivation, since if the NTP concentration is increased to 1 mM, transcription proceeds at a relatively undiminished rate for ~100 min. These data strongly suggest that the K_m values for elongation are in the range of 0.1 mM, somewhat larger than the K_d values estimated for the *E. coli* enzyme which are as low as 0.015 mM (Chamberlin, 1976). This shows that, if large amounts of T7 transcripts are required, either a higher concentration of NTPs should be used or the reaction mix supplemented with substrates during the progress of the reaction.

Isolation and Characterization of Restriction Fragments of T7 DNA Which Serve as Templates for Transcription by T7 RNA Polymerase. In the search for short restriction fragments of T7 DNA carrying intact promoters, a number of restriction enzymes have been tried. The restriction enzyme

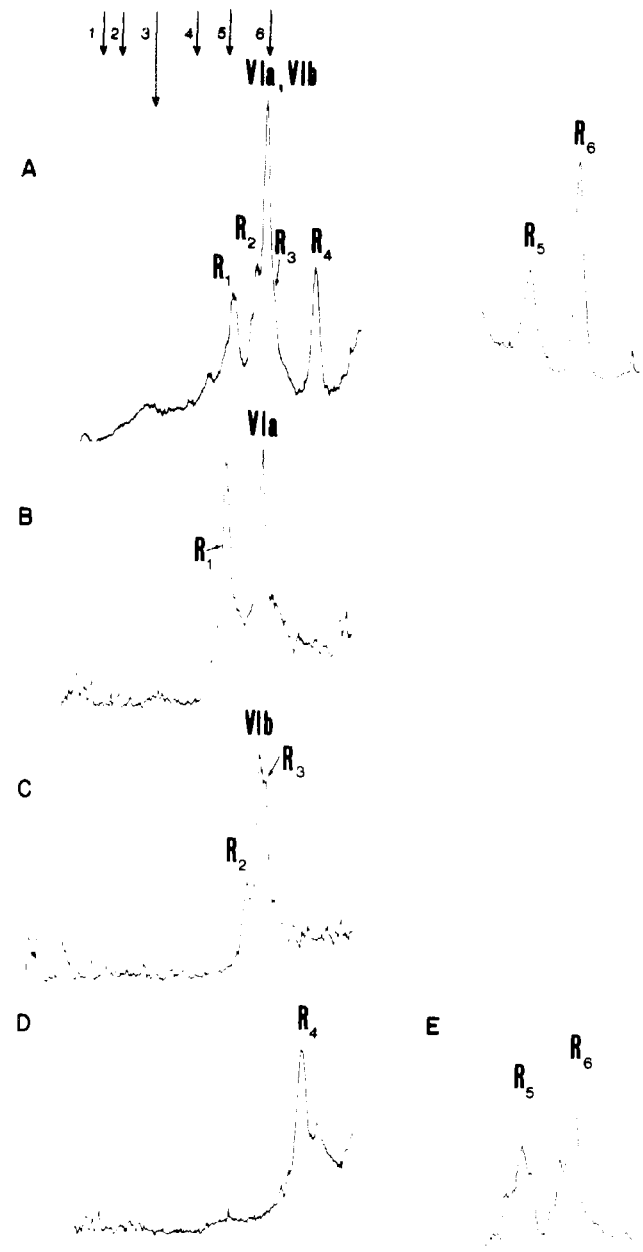


FIGURE 3: Densitometer traces of RNA gels of transcripts from (A) total mix of Hpa II restriction fragments; (B) Hpa II fragments >1300 base pairs; (C) Hpa II fragments from 1150 to 550 base pairs; (D) Hpa II fragments from 500 to 250 base pairs; (E) Hpa II fragments from 200 to 56 base pairs. Arrows (1–6) at the top indicate gel positions of the seven major in vitro transcripts from whole T7 DNA (I–VI, including IIIa and IIIb which comigrate). VIa and VIb indicate original band VI which is unaffected by Hpa II restriction and a new band of the same length. [Which of these is the native band VI is not certain. The Hpa II fragment containing the right-hand end of the T7 genome has an estimated size of ~1500 bases (Ludwig, 1975); therefore, VIa probably corresponds to native band VI.]

Hpa II from *Hemophilus parainfluenzae* makes from 50–60 cuts in double-stranded T7 DNA. Previous work from this laboratory has shown that the promoters for T7 RNA polymerase are preserved in the Hpa II fragment mix. A densitometer trace of the radioautograph of an agarose–acrylamide gel of the transcription products using the whole fragment mix as template is shown in Figure 3A and compared to the positions in the same gel of the seven major transcripts produced in vitro from whole T7 DNA (Golomb & Chamberlin, 1974). It now appears likely that one or more additional large transcripts initiated at about 14% of the genome do not enter the gel (Oakley & Coleman, 1977). Transcripts ap-

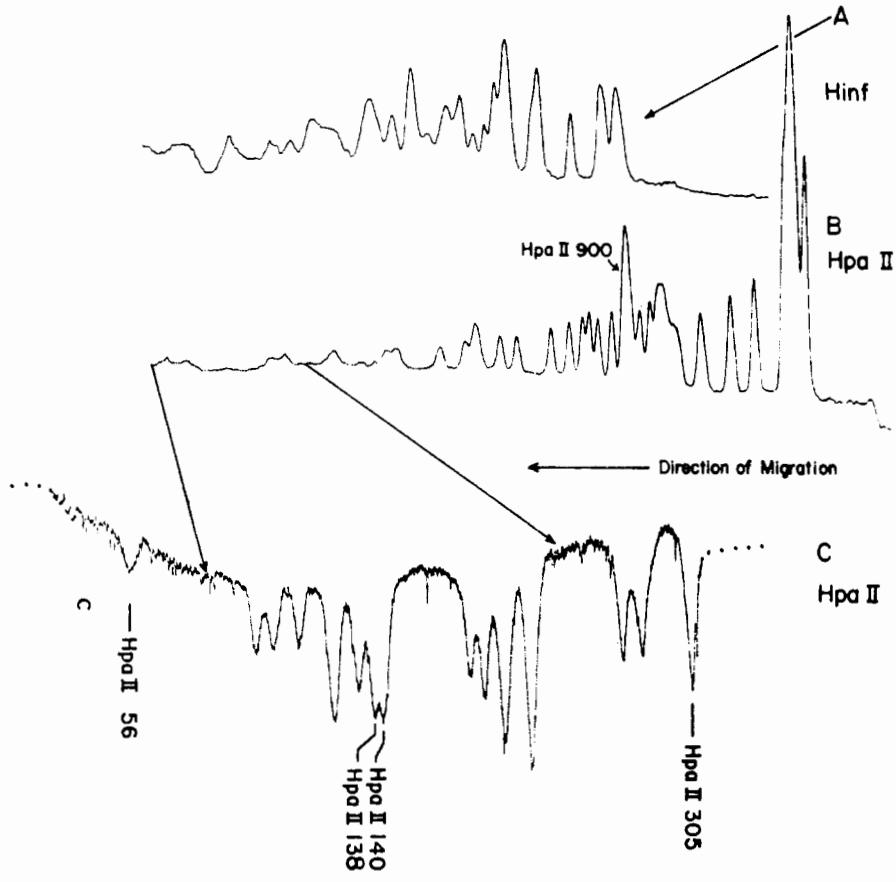


FIGURE 4: Densitometer traces of DNA gels of restriction fragments of T7 DNA: (A) after digestion by Hinf; (B) after digestion by Hpa II; (C) high resolution of Hpa II fragments below 300 base pairs. All gels are 5% TBM–polyacrylamide gel.

pearing only on transcription of Hpa II fragments are labeled R_n, while two transcripts appearing at the position of the "normal" band VI are labeled VIa and VIb. The small transcripts R_5 and R_6 required a separate, higher percentage gel.

The Hpa II restriction fragment mix was electrophoresed on an acrylamide gel (Figure 4B), stained with ethidium bromide, and cut into sections containing arbitrary size classes of DNA as follows: >1300 base pairs, from 1150 to 550 base pairs, from 500 to 250 base pairs, and from 200 to 56 base pairs in length. The smallest Hpa II fragment, 46 base pairs in length, is not seen on the gels used. The fragments were then eluted from the gel and used as templates for separate transcription reactions. The transcripts whose synthesis is directed by templates in each size class are displayed in Figure 3B–3E as densitometer traces of the radioautographs of the RNA gels. The two smallest transcripts, R_5 (~85 bases) and R_6 (~50 bases), are of greatest interest for further structural work because of their small size and the fact that they also come from the set of smallest DNA fragments, 200–256 base pairs. A high-resolution gel of the smallest fragments of the Hpa II digest from which the two smallest transcripts, R_5 and R_6, come is shown in Figure 4C. The latter fragments were cut out separately and tested for template activity. Transcript R_5 comes from a template 140 base pairs long, Hpa II 140b, while R_6 comes from a template 138 base pairs long, Hpa II 138. The band labeled Hpa II 140 in Figure 4C can be shown by additional electrophoresis (and most clearly by sequencing gels, see below) to consist of two DNA fragments ~140 base pairs in length. Only one of these, designated Hpa II 140b, contains a promoter.

A gel of the isolated template Hpa II 138 is shown in Figure 5A and the gel of the products of its transcription by T7 RNA polymerase in Figure 5D. The major product is R_6. The small band 1 is probably due to slight contamination by Hpa II 140b which produces R_5 (Figure 5C) and small band 2 is probably a small amount of end-to-end transcription which we observe with some restriction fragments, especially when high concentrations of enzyme are used. Small amounts of premature termination may also occur since light bands corresponding to shorter fragments are present in some cases. The DNA template, Hpa II 138, can be further cut by the restriction enzyme Hae III. This shortens Hpa II 138 by 23 bases to give Hpa II, Hae III 115 and a small fragment Hpa II, Hae III 23 (Figure 5B). The transcript R_6', shorter than R_6 by 23 bases, is transcribed from Hpa II, Hae III 115 (Figure 5E). This Hae III site provided a marker to be confirmed in the final sequence (see below).

In contrast to digestion of the T7 DNA template by Hpa II, which does not reduce the initiation rate of RNA chains as measured by [γ-^{32}P]GTP incorporation (Oakley et al., 1975), digestion by Hinf drastically reduces [γ-^{32}P]GTP incorporation. The gel of the Hinf digest of T7 DNA is shown in Figure 4A. No DNA fragments longer than ~900 bases remain after digestion by Hinf. Transcription of the whole mixture of Hinf fragments produces two short mRNAs, F_1 and F_2, one in considerably larger amounts than the other (Figure 5G). Rough estimates of size based on migration of a few RNAs of known size suggest that the two transcripts from the Hinf digest are between 150 and 250 bases long. The loss of most of the initiation sites on treatment with Hinf anticipated the Hinf sites discovered eight bases from the initiation point in at least two of the T7 promoters (see below).

Changes in Ultraviolet Absorption of the DNA on Binding of T7 RNA Polymerase to the Promoters on Hpa II 140 and Hpa II 138. Both Hpa II 140b and Hpa II 138 are short

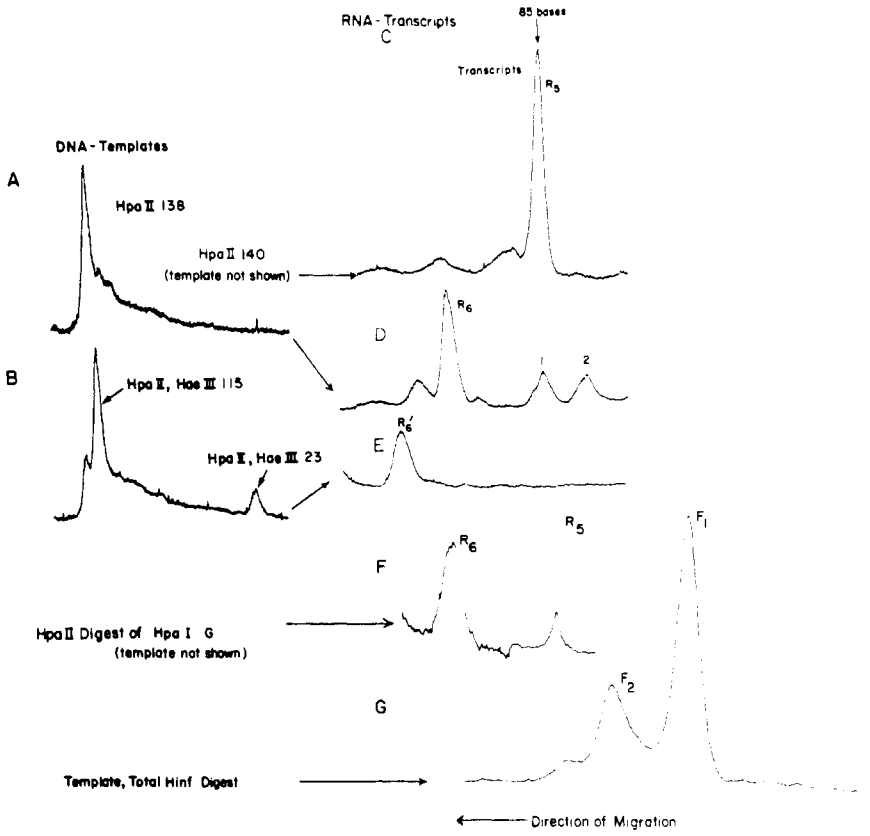


FIGURE 5: Densitometer traces of various DNA templates and RNA transcripts made from them: (A) Hpa II 138; (B) Hpa II, Hae III 115 plus Hpa (II), Hae III 23; (C) RNA transcript of Hpa II 140; (D) RNA transcript of Hpa II 138; (E) RNA transcript of Hpa II, Hae III 115; (F) RNA transcript of a Hpa II digest of Hpa IG; (G) RNA transcripts from the total Hinf digest of T7 DNA.

enough that the promoter, an interaction site for a protein of M_r 107 000, must constitute a significant percentage of the total length of the fragment. Thus it should be possible to detect any changes in the absorption of the base chromophores in the promoter region which are induced by binding of the polymerase. Spectral changes induced in the DNA on binding of T7 RNA polymerase to a mixture of Hpa II 140b and 138[2] were followed in the Cary 118 spectrophotometer. Binding reactions were carried out in thermostated (37 °C), 1-mL cuvettes in 0.04 M Tris-HCl, 0.01 M $MgCl_2$ (pH 8.0). Measurements were made at very high sensitivity, and it was found necessary to clarify the solutions used by centrifugation of the enzyme stock and nitrocellulose filtration of the DNA solutions before use to eliminate light-scattering problems. The results are shown in Figure 6.

In panel A are shown the spectral changes induced in a mixture of fragments Hpa II 138 and Hpa II 140b on addition of increasing amounts of T7 RNA polymerase stock solution (3 mg/mL). Enzyme was added to buffer alone in the reference cuvette to cancel the contribution of enzyme absorbance. The decrease in amplitude of the peak is due to dilution of the sample. There is a small but distinct blue shift of 1–2 nm induced in the DNA at low enzyme–DNA ratios.

In panel B is shown a parallel experiment using nonpromoter-containing fragments, Hpa II 195–175. In this case there is no change at low enzyme concentrations and a red shift at high concentrations. Completion of this latter change at high enzyme concentrations required several minutes.

The remainder of the figure shows the results of two difference experiments. In the first, enzyme was added to Hpa II 140b and 138 in one cuvette and to an equal amount of nonpromoter-containing fragment (at the same optical density) in the reference cuvette. As would be predicted from the spectra in the upper part of the figure, there is an increase in absorption at 257 nm of 0.005 OD, complete at active enzyme to promoter ratios of about 1, an hyperchromicity probably associated with specific promoter binding (trace A, Figure 6). With the addition of more enzyme (additions in microliters shown), a time-dependent decrease in OD is seen, corresponding to the red shift associated with nonspecific DNA binding.

To confirm the absence of any change in nonpromoter fragments at low enzyme concentrations, a parallel experiment was performed with nonpromoter fragments in the experimental cuvette and buffer alone in the reference. The lower curve, B, shows the result. At low enzyme concentrations dilution of the sample is the only result. The decrease in OD at 257 nm is not detectable until the addition of >10 μL of RNA polymerase stock solution and again requires several minutes for completion.

Strand Separation and Sequencing of Hpa II 138 and 140. Strand Separation of Hpa II 138. Examples of strand separation gels of Hpa II 138 are shown in Figure 7. Strand separation of Hpa II 138, labeled at both 5′ ends, consistently yielded a pattern of four bands (Figure 7A). Band A corresponds to one strand and bands B, C, and D are due to the other strand running at three different positions. The A and G patterns produced when material from bands B, C, and D is subjected to methylation and chain cleavage according to Maxam & Gilbert (1977) confirm that all three have the same

[2] Prolonged electrophoresis begins to separate the two Hpa II 140 fragments, suggesting that they are not identical in length. Hpa II 140b is the more slowly migrating band. Hence the nonpromoter-containing Hpa II 140a is a contaminant of the promoter-containing fragments used for these experiments.

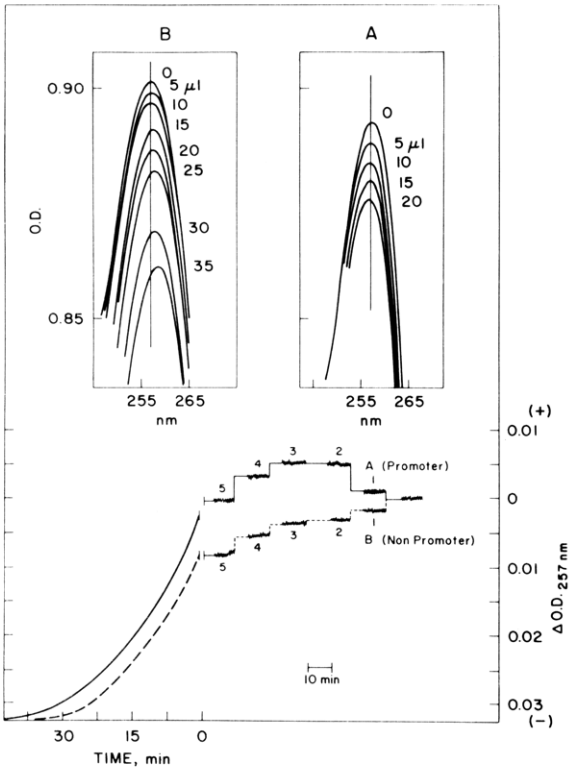


FIGURE 6: UV absorption spectral changes induced in promoter-containing and nonpromoter-containing DNA restriction fragments by T7 RNA polymerase. Panels A (promoter-containing) and B (nonpromoter-containing) show spectra from 250 to 265 nm as T7 RNA polymerase was added as described in the text. The lower curves show ΔOD at 257 nm during addition of T7 RNA polymerase to (A) promoter-containing fragments vs. nonpromoter-containing fragments in the reference cell; (B) nonpromoter-containing fragments vs. buffer. No further aliquots of enzyme were added after zero time.

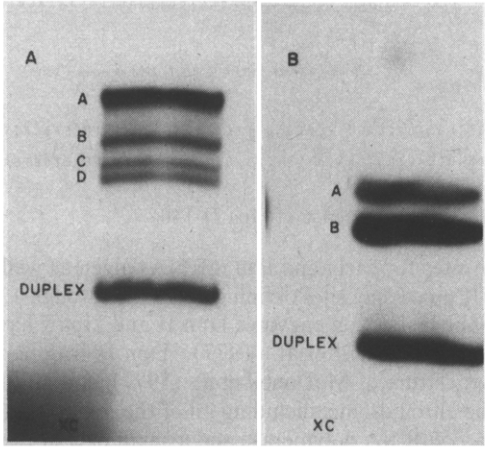


FIGURE 7: Examples of strand separation on 8% polyacrylamide gels (radioautographs): (A) strand separation of Hpa II 138; (B) strand separation of Hpa II, Hae III 115. Gel conditions as in Maxam & Gilbert (1977).

sequence, that of noncoding strand. This phenomenon is presumably due to the assumption of three different discrete, fairly stable conformations. Subtle differences in conformation of two complementary strands in the Tris-borate buffer are probably the reason this gel method is effective in separating strands in the first place. That the material in bands B, C, and D was of identical length and thus had not been subjected to attack by a contaminant nuclease was confirmed by running the eluted DNA on denaturing gels (not shown). In contrast to the behavior of whole Hpa II 138 on strand separation, Hpa II, Hae III 115 separates normally and only two conformers are observed (Figure 7B). The different conformations observed with Hpa II 138 might relate to the base-pairing structures possible in the region of the DNA corresponding to the RNase III site in the message (see below). Why there would be a greater tendency to this in the strand complementary to the template is not clear. This region is removed by Hae III. Such a conformational phenomenon shows that multiple bands on strand separation gels need to be interpreted with caution.

Strand Separation of Hpa II 140. Strand separation gels of the band corresponding to Hpa II 140 in Figure 4C resulted in four bands which clearly carried two different sequences, confirming the impression from a number of gel densitometer tracings that there exist two Hpa II fragments ~140 base pairs long. The Hpa II 140 fragment (Hpa II 140b) giving rise to the transcript R_5 was found to contain a Hinf site analogous to the one within the promoter on Hpa II 138 (see below), while the other Hpa II 140 fragment (Hpa II 140a) does not contain this particular Hinf site, although two others are present. Hence the mixture of the two Hpa II 140 fragments was treated with Hinf, and the multiple fragments were separated by electrophoresis and sequenced.

Base Sequences of Hpa II 138, 140a, and 140b. The sequences of both strands of Hpa II 138 were determined by the chemical methods of Maxam & Gilbert (1977). The sequence of Hpa II 138 is shown and compared to the promoter region of Hpa II 140b to be described below (Figure 8). The orientation corresponding to transcription is from left to right since transcription by T7 RNA polymerase has been shown to be exclusively rightward (Chamberlin et al., 1970). The Hae III site expected in Hpa II 138 was found in the sequence and sequencing of Hpa II, Hae III 115 provided a convenient confirmation of the original sequence. The sequence also revealed two Hinf sites and as predicted Hpa II 138 was cleaved at two points by Hinf. Sequencing of these fragments provided further verification. A sequencing gel of the initiation region of Hpa II 138 is shown in Figure 9A.

The sequence of the promoter-containing Hpa II 140b was obtained applying the Maxam & Gilbert method to the Hpa II, Hinf fragments. In our hands it has been difficult to efficiently label and strand-separate the Hpa II, Hinf fragments. Hence we present the sequence of only 40 bases flanking the initiation site (Figure 8). The sequence beyond this is not certain enough to be reported at present. This sequence contains the essential region of the promoter, at least that homologous to the T7 promoter on Hpa II 138.

The nonpromoter-containing Hpa II 140a contains a Hha I site, and the sequence of 140a from the Hha I site to one of the Hpa II ends is shown in Figure 10. While this fragment has an AT-rich region and some homology, including twofold symmetry, with the proven promoter, this fragment gives no in vitro transcript with T7 RNA polymerase. A sequencing gel for Hpa II 140a is shown in Figure 9B. Important regions of the sequences were confirmed by two-dimensional electrophoresis and chromatography after digestion with snake venom phosphodiesterase and DNase I.

Initiation Points and Sequences of R_5 and R_6 mRNAs. The transcripts from Hpa II 140b and Hpa II 138 were isolated from the gels shown in Figure 5 and the sequences determined by the enzymatic method of Donis-Keller et al. (1977). Representative sequence gels are shown in Figure 11 and the complete sequences are shown in Figure 8 placed in register with the corresponding DNA sequence. Note that the initiation point for R_6 on Hpa II 138 is two bases to the left of

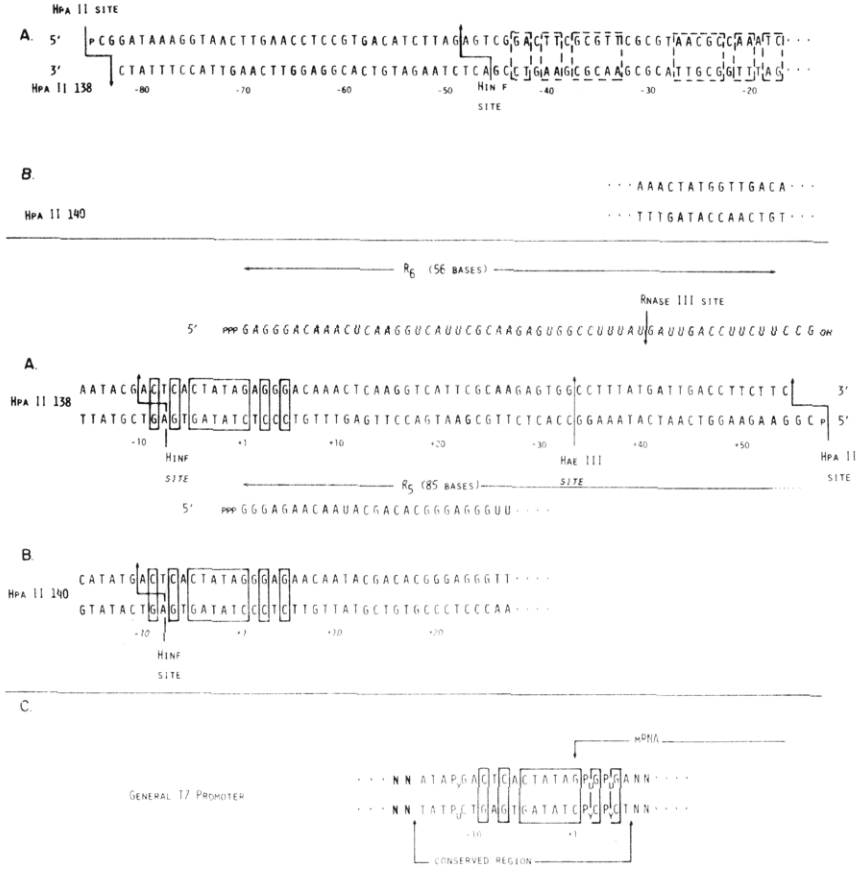


FIGURE 8: DNA sequences of Hpa II 138 (A) and Hpa II 140b (B) and their respective transcripts R_6 and R_5. (C) Conserved DNA sequence of the T7 promoter in the region of initiation.

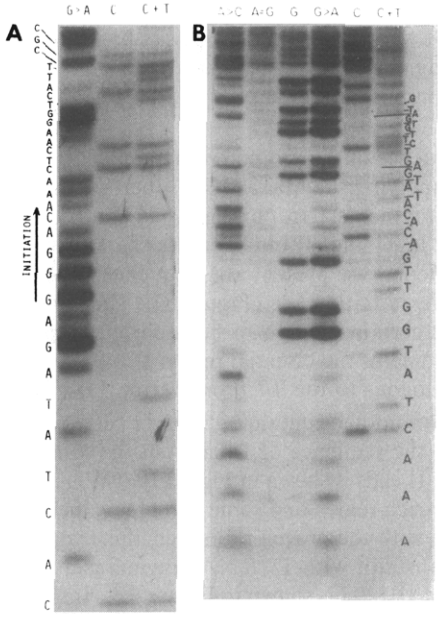


FIGURE 9: (A) Sequence gel (method of Maxam & Gilbert, 1977) in the region of RNA initiation for Hpa II 138. (B) Sequence gel for Hpa II 140a, the fragment not containing a major promoter (see Discussion).

the initiation point reported in our original publication (Oakley & Coleman, 1977; see Discussion). The numbering system in Figure 8 is based on +1 as the initiation point for transcription, omitting zero in the base numbering.

Mapping of Hpa II 138, 140a, and 140b. A map of the left-hand end of the T7 genome is given in Figure 12 with the

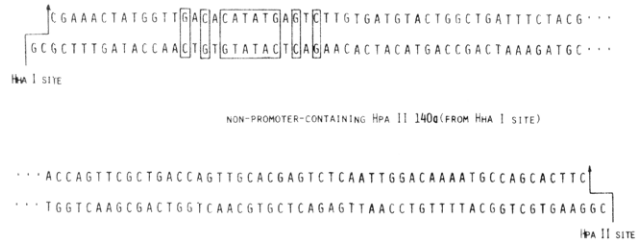


FIGURE 10: DNA sequence of Hpa II 140a.

initiation sites for early and late mRNAs given as well as the RNase III processing sites defining genes 0.3 through 1.3. The sites for the restriction enzymes Dpn II and Hpa I have been mapped by McDonell et al. (1977). Dpn II fragment B, in the nomenclature of McDonell et al. (1977), contains the left end of the chromosome, including all of the region transcribed by the *E. coli* RNA polymerase and a small part of the region transcribed by the T7 enzyme including the first promoter for the T7 polymerase near the end of gene 1. Gene 1.3, the T7 DNA ligase gene, is transcribed by both the *E. coli* and T7 RNA polymerases (Studier, 1972), establishing the fact that the first T7 polymerase promoter lies to the left of the terminator for the *E. coli* enzyme. The T7 enzyme reads through this terminator since transcription by the T7 enzyme continues to the end of Dpn II B (Dunn & Studier, personal communication). A terminal part of gene 1 and gene 1.1 is apparently also transcribed by both polymerases.

Hpa II 138 has been mapped to the junction of genes 1 and 1.1 (Studier, personal communication), a position confirmed by the finding of the RNase III site for this gene junction on Hpa II 138 (Figure 8) (Oakley & Coleman, 1977; Rosenberg et al., 1974; Robertson et al., 1977). One of the Hpa II 140

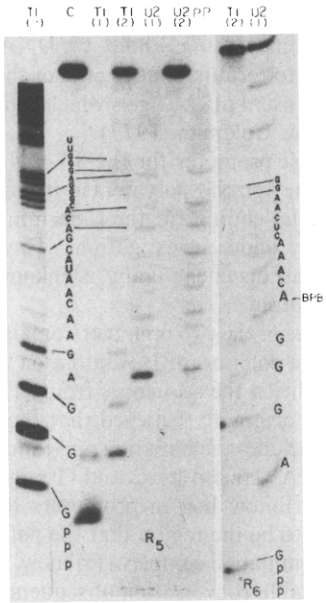


FIGURE 11: RNA sequence gel (method of Donis-Keller et al., 1977) of R_5 (left) and R_6 (right). The track PP displays the products of partial digestion with 1 unit of the RNase I from *Physarum polycephalum*, an enzyme cleaving at all bases but C (Simonscits et al., 1977). For the leftmost T1 track, R_5 RNA labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP was used. Others were labeled during transcription with [γ-^{32}P]GTP.

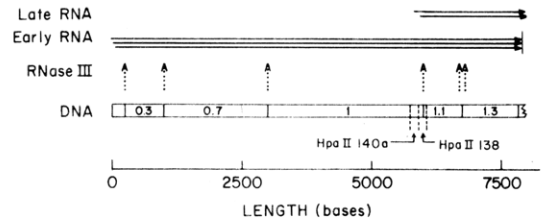


FIGURE 12: Map of the left-hand end of the T7 genome. The solid horizontal arrows indicate early and late mRNAs. The vertical dashed arrows indicate RNase II sites on the polycistronic mRNA. Positions of the genes as defined by the RNase III cleavages are indicated by the numbers on the DNA.

fragments maps immediately to the left of Hpa II 138.

In order to determine which of the Hpa II 140 fragments this is, we isolated the restriction fragment known as Hpa IG (McDonell et al., 1977) which contains this region of the chromosome and which on further digestion with Hpa II produces Hpa II 138 and one of the Hpa II 140 fragments. Transcription of the latter digest with T7 polymerase produces a standard amount of R_6 and a small amount of R_5 (Figure 5F). The latter proved to result from contamination of Hpa IG by Hpa IE, the fragment which carries the promoter for R_5. Using DNA from the mutant phage ΔH_3, which lacks one of Hpa I sites and hence produces fewer bands in the Hpa IE to G region, Hpa IE can be completely separated from Hpa IG and use of the Hpa II digest of Hpa IE as template produces only R_5, while the Hpa II digest of Hpa IG produces only R_6. Hence the nonpromoter-containing Hpa II 140a is the fragment immediately to the left of Hpa II 138. The promoter on Hpa II 140b is carried on Hpa IE which maps to ~70% of the genome and carries the promoter for band II mRNA from whole T7 DNA (M. Rosa, personal communication).

Discussion

T7 RNA polymerase is a single polypeptide chain which carries out all steps of transcription. The fidelity of promoter recognition and initiation of transcription is conserved even on small double-stranded DNA fragments of the T7 chromosome containing a single promoter as has been indicated by previous work (Oakley et al., 1975; Oakley & Coleman, 1977) and is further documented here. These properties should make T7 RNA polymerase and its promoter complexes a particularly useful system for the study of the molecular mechanisms of transcription.

pH Dependence of Transcription. Few studies of the pH dependency of transcription by any RNA polymerase have been published. Most of the data is on the *E. coli* enzyme. Rhodes & Chamberlin (1974) showed that the elongation reaction, i.e., elongation of the ternary complex between the enzyme, the template (poly[d(A-T)]), and poly[r(A-U)], shows a broad pH maximum centered near pH 8.2, activity rising from 50% at pH 7.0 and falling to 60% at pH 9.4. The complete reaction was earlier reported as 13% at pH 6.1, 62% at pH 7.0, and 84% at pH 8.9, with a maximum from 7.8 to 8.2 (Chamberlin & Berg, 1962). For T7 polymerase the pH dependence of the complete reaction is generally similar, although the very rapid fall below pH 7.0 is more clearly apparent (Figure 1A). While activity falls off somewhat at high pH, the general profile is more sigmoidal than bell-shaped. This pH function, with a midpoint near pH 6.8, cannot be described by a single ionization; two or more ionization processes must be involved.

One of the pH-dependent processes influencing the pH–rate profile of transcription is clearly binding of the enzyme to the DNA template (Figure 1A). Binding falls off rapidly below pH 8 and the sigmoid dependency of binding can be described by a single apparent pK_a of ~7. Since there is not a group on the DNA with a pK_a in this region, it may be concluded that there is a group on the enzyme with a pK_a near 7 that affects template binding. This group might be a histidyl nitrogen, or possibly the single sulfhydryl group required for activity. The pH dependency of alkylation of the SH group (Figure 1A) shows it to have an abnormally low pK_a (Figure 1A). Since the reaction conditions that were practical did not measure initial alkylation rates, the actual pK of this group is possibly lower than 7.8. Preliminary data at relatively high polymerase concentrations showed that the carboxymethyl enzyme bound to T7 DNA just as efficiently as the native enzyme (Oakley et al., 1975). We have since obtained more extensive data on the alkylated polymerase from low to high protein concentrations and confirmed that the alkylated inactive enzyme binds the DNA template apparently normally at pH 8.0 (Oakley & Coleman, unpublished data). Thus the free SH group must be involved in the catalytic process and not DNA binding.

The inactivation of the enzyme by carboxymethylation does not seem to be solely a steric hindrance as found in the case of creatine kinase (Smith et al., 1975). Modification of the sulfhydryl with methyl methanethiosulfonate, which by the method of Smith et al. (1975) introduces a SCH_3 group on the sulfur, totally inactivates the enzyme (Sarris & Coleman, unpublished data). A parallel sham reaction without reagent did not affect activity. Removal of the SCH_3 group leads to only partial reactivation, but this is probably due to the general fragility of the enzyme.

If, as seems likely, the pK_a of the SH group is near 7, its ionization may be reflected in the pH–rate profile. A catalytic group with pK_a ~ 7 and the pH function for template binding would be sufficient to generate the observed pH–rate profile for transcription. The presence of the nucleoside triphosphates may stabilize DNA binding above pH 7.0 since activity (complete system) rises faster in this region than DNA binding

as measured by the filter-binding technique.

Temperature Dependence of Transcription. The sharp temperature dependence of transcriptional activity by RNA polymerase has reasonably been postulated to reflect the requirement for melting of a section of the promoter by the polymerase in the formation of the "open" complex in preparation for initiation (Mangel & Chamberlin, 1974). The fact that this temperature dependence correlates directly with the temperature dependence of tight specific binding of the polymerase to the DNA template is particularly clearly shown with the T7 enzyme (Figure 1B). This temperature dependence of transcriptional activity extends to small, ~140 base pair, promoter-containing fragments (Figure 1B). The conclusion that the specific binding of the enzyme to a promoter-containing fragment of the DNA is associated with melting of a section of the template is further supported by the changes in the base chromophores of Hpa II 138 and 140 on the binding of polymerase as shown in Figure 5. The promoter-containing fragments undergo a hyperchromic blue shift consistent with melting of the double strand, while the nonpromoter-containing fragments under exactly the same conditions show a red shift and hypochromia consistent with nonspecific binding that does not melt the double strand (Figure 5). While these experiments are preliminary and limited by the quantity of fragments that can be isolated from whole T7 DNA, the system is a promising one for further investigation, especially since attempts to insert promoter-containing fragments into small plasmids have been successful and should increase availability of these fragments (Panayotatos, 1978; Campbell et al., 1978).

Effect of Nucleotide Concentration on Transcription. The finding that the maximum rate of mRNA synthesis by the highly active T7 enzyme requires ~1 mM nucleoside triphosphates (Figure 2) suggests that the K_m's for substrate binding to the T7 enzyme are considerably higher than K_d's reported for the *E. coli* enzyme, especially for the initiating nucleotide which is generally reported as much less than 1 mM (Chamberlin, 1976). Since under the reaction conditions employed the mRNAs transcribed from T7 promoters by T7 polymerase are all produced in approximately equimolar amounts regardless of their length (e.g., those in Figure 3), the template binding or initiation step must be rate limiting. That initiation is much slower than elongation is supported by estimates of initiation rates made from measurements of [γ-^{32}P]GTP incorporation (all transcripts produced by T7 polymerase begin with G). Precise calculations of either elongation rate or initiation rate are hampered by variable amounts of inactive polymerase in all T7 polymerase preparations (Chamberlin & Ring, 1973; Oakley et al., 1975).

Location and Structure of the T7 Promoters. The DNA sequence and transcription data on Hpa II 138, 140a, and 140b clearly establish the promoter on Hpa 138 as the single major promoter beginning transcription by the T7 RNA polymerase. While there is some preliminary evidence for tandem transcriptional starts for T7 polymerase in the "early" region (Studier, personal communication), the possibility that such a tandem start is initiated from the promoter on Hpa II 138 and a second one on the adjacent Hpa II 140a (Oakley & Coleman, 1977) is clearly ruled out since Hpa II 140a does not contain a promoter. Since the promoter on Hpa II 140b is located at ~70% of the genome and therefore accounts for one of the seven major in vitro transcripts from whole T7 DNA, the current evidence suggests that eight major promoters for the T7 enzyme exist on the T7 genome including the one on Hpa II 138. A transcript from this promoter (at ~14% of the genome) using whole T7 DNA as template would be expected to terminate at ~60% of the genome and hence not enter the usual mRNA gels which show seven major mRNAs (Oakley & Coleman, 1977).

Thus the leftmost promoter for the T7 RNA polymerase lies within the gene for the polymerase itself. Its position relative to the triplet coding for the C-terminal amino acid residue is currently under investigation. The physiological significance of this arrangement, if any, is unknown and further information is required.

The base sequence of the two promoter-containing fragments is identical from the point of mRNA initiation to position −11 (Figure 8). In addition the sequences from −13 through −15 are identical. We originally believed that the boxed region (Figure 8) which includes the initiating nucleotide and contains the sequence TATA flanked by C and G might correspond to the so-called Pribnow box in promoters for the *E. coli* enzyme suggested to be the region that the polymerase melts in formation of the initiation complex (Pribnow, 1975). Except to note the remarkable conversation of sequence between the two promoters, specific features of the base structure and phosphodiester conformation required for initiation will require further exploration by chemical modification and protection experiments. It is true that to the left of the initiation point 14–15 out of the first 21 base pairs are AT, suggesting that this region would melt rather easily (Figure 8). Both promoters also show significant twofold symmetry as indicated by the boxed regions in Figure 8. The significance of this to promoter recognition is not known, but it may correspond to complementary topology on the enzyme surface.

Initial experiments incubating Hpa II 138 and T7 enzyme with [α-^{32}P]GTP alone showed the production of pppGpG and pppGpGpG (Oakley & Coleman, 1977). Since preliminary data suggested that T7 transcripts began with the sequence pppGGG, we assumed that the 5′ end of R_6 was coded by the CCC beginning at +3 of Hpa II 138 (Figure 8). The complete sequence of R_5 shows this assignment to be wrong and the transcript begins pppGAGGG (Figures 8 and 11). Since R_5 does begin with pppGGG, the original results might have reflected some contamination of the template Hpa II 138 with Hpa II 140b. It is also possible, however, that with only one nucleoside triphosphate present there is some "slippage" of the enzyme at the initiation point and the downstream CCC is picked up. This is currently under investigation.

Another structural feature common to both promoters is the presence of a Hinf restriction site covering positions −7 to −11. This must be within the region of the initiation complex, and indeed restriction of Hpa II 140b or 138 by Hinf abolishes transcription. Restriction of total T7 DNA by Hinf greatly reduces transcription, but transcription from at least one T7 promoter appears to continue (Figure 5G). The latter finding shows that the GACTC from −11 through −7, although preserved in the two T7 promoters shown here, may be altered in at least one of the others. To summarize, what is possibly a general T7 promoter sequence is shown at the bottom of Figure 8. A span of ~15 base pairs appears to be common to both promoters. This is a reasonable length for a DNA binding groove in a globular polymerase of ~M_r 107 000. Outside of this region major base substitutions occur between the promoters as indicated by N. Base methylation experiments in the presence of protecting polymerase show that major changes in the susceptibility of the bases to methylation are confined to the A and G residues from −6 to +6 (Oakley et al., in preparation).

Added in Proof

In an independent determination of the DNA sequences of the promoters for T7 RNA polymerase, Margaret Rosa has found the purine at position −12 in the Hpa II 140a promoter (located at ~70% of the genome) to be G rather than A as reported above. On reexamination of the sequence gels for the Hinf fragment, in which this purine is represented by a short nucleotide, there is abnormal migration of this nucleotide such that the dark G band appears in an anomolous position, now known to be a not infrequent occurrence for the short fragments. Hence position −12 is C-G in both promoters. Thus all 15 base pairs leftward of the initiation point are identical in the two T7 promoters. We thank Margaret Rosa for communicating this information prior to publication.

Acknowledgments

We gratefully acknowledge the technical assistance of Judy Pascale-Judd. We thank William Studier, Joan Steitz, and Margaret Rosa for many helpful discussions and suggestions. We additionally acknowledge Dr. Studier for communicating his data on mapping of the T7 genome.

References

Applebury, M. L., Johnson, B. P., & Coleman, J. E. (1970) *J. Biol. Chem. 245*, 4968–4976.

Campbell, J. L., Richardson, C. C., & Studier, F. W. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2276–2280.

Chamberlin, M. (1976) in *RNA Polymerase* (Losick, R., & Chamberlin, M., Eds.) pp 17–67, Cold Spring Harbor Laboratory, Cold Spring Harber, NY.

Chamberlin, M., & Berg, P. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 81–94.

Chamberlin, M., & Ring, J. (1973) *J. Biol. Chem. 248*, 2235–2244.

Chamberlin, M., McGrath, J., & Waskell, L. (1970) *Nature (London) 228*, 1160–1162.

Coleman, J. E. (1974) *Biochem. Biophys. Res. Commun. 60*, 641–648.

Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527–2538.

Golomb, M., & Chamberlin, M. (1974) *J. Biol. Chem. 249*, 2858–2863.

Heyden, B., Nusslein, C., & Schaller, H. (1972) *Nature (London), New Biol. 240*, 9–12.

Jones, O. W., & Berg, P. (1966) *J. Mol. Biol. 22*, 199–209.

Ludwig, R. A. (1975) Ph.D. Thesis, Yale University, New Haven, CT.

Mangel, W. F., & Chamberlin, M. (1974) *J. Biol. Chem. 249*, 3007–3013.

Maniatis, T., Jeffrey, A., & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1184–1188.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

McDonell, M. W., Simon, M. N., Studier, F. W. (1977) *J. Mol. Biol. 110*, 119–146.

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.

Panayotatos, N. (1978) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 37*, 1409.

Pribnow, D. (1975) *J. Mol. Biol. 99*, 419–443.

Rhodes, G., & Chamberlin, M. (1974) *J. Biol. Chem. 249*, 6675–6683.

Robertson, H. D., Dickson, E., & Dunn, J. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 822–826.

Rosenberg, M., Kramer, R. A., & Steitz, J. A. (1974) *Brookhaven Symp. Biol. 26*, 277–285.

Simonscits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R., & Guilley, H. (1977) *Nature (London) 269*, 833–836.

Smith, D. J., Maggio, E. T., & Kenyon, G. L. (1975) *Biochemistry 14*, 767–771.

Studier, F. W. (1972) *Science 176*, 367–376.

Summers, W. C. (1969) *Virology 39*, 175–182.