

Processivity in Early Stages of Transcription by T7 RNA Polymerase[†]Craig T. Martin,[‡] Daniel K. Muller,[§] and Joseph E. Coleman*

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

Received November 24, 1987; Revised Manuscript Received February 10, 1988

ABSTRACT: Immediately following initiation of transcription, T7 RNA polymerase enters a phase in which dissociation of the enzyme-DNA-RNA ternary complex significantly competes with elongation, a process referred to in the *Escherichia coli* enzyme as abortive cycling [Carpousis, A. J., & Gralla, J. D. (1980) *Biochemistry* 19, 3245-3253]. Characterization of this process in the T7 RNA polymerase system under various reaction conditions and on templates with differing message sequences reveals that conversion to a highly processive ternary complex occurs after incorporation of eight bases and that the relative competition between dissociation and elongation up to this point is influenced by several different forces. In particular, the sequence dependence of abortive falloff suggests that dissociation is favored immediately following incorporation of UMP and is less likely following incorporation of GMP into the RNA message. Abortive cycling is unchanged in transcription from a synthetic oligonucleotide template which is double-stranded in the promoter region but single-stranded throughout the entire message region. This result proves that melting and reannealing of the DNA duplex in the coding region do not contribute to abortive cycling. Furthermore, weakening of promoter binding by an order of magnitude affects abortive cycling only slightly, suggesting that strong interactions with the promoter are not the major cause of abortive cycling. Kinetic analyses show that conversion to a highly processive ternary complex after the incorporation of eight bases may reflect a large decrease in the unimolecular rate of dissociation of the complex due to increased contacts between the nascent RNA and the DNA template and between RNA and enzyme. Finally, it is shown that when transcribing a message of initial sequence GGG..., in the presence of GTP as the sole nucleotide, T7 RNA polymerase synthesizes a ladder of poly[r(G)] transcripts due to a slipping of the RNA message along the DNA template. This ladder is largely truncated at a length of 14 bases, providing information on steric constraints in the ternary complex positioned at +3 on the DNA template.

T7 RNA polymerase is the most well-characterized of the single-subunit bacteriophage RNA polymerases (Chamberlin & Ryan, 1982). Its small size (*M*_r 98 856), stringent promoter specificity, and lack of external regulation make T7 RNA polymerase an ideal system in which to study the fundamental aspects of transcription. We have recently developed a kinetic assay for the study of transcription initiation by this enzyme (Martin & Coleman, 1987). The assay follows the steady-state synthesis of a five-base RNA transcript and has allowed the determination of individual promoter base contributions to specific promoter recognition and initiation by T7 RNA polymerase.

Most recent *in vitro* transcription studies of T7 RNA polymerase have employed transcription from a long linear DNA template, with termination of the transcript occurring at the end of the double-stranded DNA template, so-called runoff transcription. It is well-known that T7 RNA polymerase is highly processive. In the absence of specific terminator sequences in the DNA, observed transcripts are full-length runoff transcripts. However, under conditions limiting one or more of the four ribonucleotides, it has been shown that the RNA polymerases from T7 (Morris et al., 1987) and from the related bacteriophage T3 (McAllister et al., 1973) can produce a significant amount of very short transcripts. Recent studies have also reported synthesis of short oligomers in the presence of all four nucleotides (Lowary et al., 1986; Milligan et al.,

1987). The abortive synthesis of short RNA transcripts immediately following initiation under *normal* nucleotide triphosphate levels has been observed in RNA polymerase from *Escherichia coli* (Carpousis & Gralla, 1980), as well as in eukaryotic RNA polymerases (Furuichi, 1981; Yamakawa et al., 1981; Ackerman et al., 1983). In this work, we report that T7 RNA polymerase synthesizes abortive transcription products even at saturating levels of all four ribonucleotides. Thus, it appears that abortive cycling is a fundamental aspect of early transcription in all RNA polymerases.

To characterize the early events following initiation of transcription, we now extend the principles of our previous initiation assay to the study of transcription of longer, yet relatively short RNA messages. The steady-state synthesis of transcripts in the range of 5-50 bases allows a detailed investigation of the progression from a newly initiated to a fully processive transcription complex. In particular, the use of synthetic oligonucleotide DNA templates provides opportunities not available in more traditional assays. In the present study, we use both plasmid-derived and synthetic DNA templates to examine some of the details of abortive cycling and show that it is an intrinsic part of the progression to a fully processive ternary transcription complex.

MATERIALS AND METHODS

T7 RNA polymerase was prepared from *E. coli* strain BL21 containing plasmid pAR1219 (kindly supplied by William Studier and John Dunn), with T7 gene 1 (RNA polymerase) cloned under inducible control of the *lac* UV5 promoter (Davanloo et al., 1984). Enzyme was purified as previously described (King et al., 1986) by fractionation with Polymix P (less than 1.25%) and ammonium sulfate, followed by

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

[‡] C.T.M. is the recipient of National Institutes of Health Postdoctoral Fellowship GM10902.

[§] D.K.M. was supported by National Institutes of Health Training Grant GM07223.

Table I: Assignments of Major Abortive and Full-Length Transcription Products Observed in Figure 1

	length	sequence	[γ -GTP] (μ M)	predicted and obsd ratios of nucleotide incorporation ^a	
				[α -UMP]/[γ -GTP]	[α -GMP]/[γ -GTP]
S1	3	GGA	3.68	0, 0.1 \pm 0.0	2, 1.7 \pm 0.4
	4	GGAC	9.13	0, 0.1 \pm 0.0	2, 1.9 \pm 0.4
	5	GGACU	38.2	1, 1.0 \pm 0.2	2, 2.0 \pm 0.4
S2	5	GGGAA	2.68	0, 0.1 \pm 0.0	3, 2.8 \pm 0.6
	6	GGGAAG	0.65	0, 0.3 \pm 0.2	4, 5.7 \pm 1.6
	7	GGGAAGU	1.80	1, 1.0 \pm 0.2	4, 3.5 \pm 0.8
	8	GGGAAGUC	0.36	1, 1.1 \pm 0.5	na ^b
	20a	GGGAAGUCUGUACCAGACGU	0.38	4, 5.6 \pm 1.9	7, 9.4 \pm 3.0
	20b	GGGAAGUCUGUACCAGACGU	0.43	4, 4.3 \pm 1.4	7, 8.3 \pm 2.6
P1	5	GGGAA	11.9	0, 0.0 \pm 0.0	3, 2.7 \pm 0.6
	6	GGGAU	5.10	1, 0.7 \pm 0.1	3, 3.6 \pm 0.8
	7	GGGAUU	5.05	2, 1.6 \pm 0.3	3, 2.1 \pm 0.5
	29	GGGAUUUCGAGC...GAUCC	0.99	5, 4.8 \pm 1.2	11, 11.0 \pm 2.7
P2	3	GGA	2.45	0, 0.1 \pm 0.0	2, 1.4 \pm 0.3
	4	GGAC	3.10	0, 0.1 \pm 0.0	2, 1.6 \pm 0.4
	5	GGACU	5.90	1, 0.9 \pm 0.2	2, 1.8 \pm 0.4
	38	GGACUGGGAUCC...CAUAG	1.49	9, 8.1 \pm 1.9	12, 10.3 \pm 2.4
P3	3	GGA	2.02	0, 0.1 \pm 0.1	2, 1.5 \pm 0.4
	4	GGAC	2.35	0, 0.1 \pm 0.0	2, 1.6 \pm 0.4
	5	GGACU	4.44	1, 0.9 \pm 0.2	2, 1.9 \pm 0.4
	29	GGACUGGGAUCC...CCAAG	0.89	4, 3.1 \pm 0.8	10, 8.7 \pm 2.2
P4	3	GAG	0.64	0, 0.3 \pm 0.2	2, 1.7 \pm 0.6
	5	GAGGG	0.34	0, 0.5 \pm 0.3	4, 3.6 \pm 1.4
	6	GAGGGA	0.20	0, 0.6 \pm 0.5	na
	64	GAGGGACAAACU...GAUCC	0.63	16, 17.4 \pm 4.7	18, 18.1 \pm 4.9

^a The terms α and γ refer to the position of ³²P in the incorporated nucleotide. ^b Not available.

RESULTS

In the presence of limiting concentrations of pyrimidine triphosphates, T7 RNA polymerase produces small (abortive) transcripts in addition to the predicted full-length transcripts (Morris et al., 1986). Recent studies describing the use of T7 RNA polymerase to synthesize large amounts of RNA have shown that short oligomers are also observed under normal levels of all four nucleotides (Lowary et al., 1986; Milligan et al., 1987). In studies of plasmid- and oligonucleotide-based transcripts from 20 to over 200 bases long, we have observed abortive products from T7 RNA polymerase and have consistently found that abortive transcripts are produced even in the presence of saturating concentrations of all four nucleotides (see below) and at very low turnover levels.

Characterization of Abortive Transcription Products. In order to determine whether abortive cycling is a general property of transcription by T7 RNA polymerase, we have compared abortive transcription products from a variety of DNA templates, including both synthetic oligonucleotides and DNA fragments derived from T7 promoters cloned into plasmids. All promoters used in this study (except as noted below) share the same T7 class III promoter consensus sequence from position -17 to position -1 (see legend to Figure 1); only the message sequence is changed. The predicted full-length message sequences are compared in Table I.

In order to quantitatively assay transcription in these studies, we generally follow transcript synthesis via incorporation of the initiating nucleotide labeled with ³²P at the γ -position (in this study, [γ -³²P]GTP). This provides an advantage over labeling via incorporation of α -³²P-labeled nucleotides in that every correctly initiated transcript is labeled and each transcript incorporates only one radioactive label, regardless of the length or composition of the transcript. However, in order to assign some of the shorter transcripts produced in this study, we have also run duplicate experiments incorporating either [α -³²P]UMP or [α -³²P]GMP. This provides quantitative

determination of the number of UMP or GMP ribonucleotide bases incorporated into each message.

The abortive transcription products from several templates containing different message sequences are compared in Figure 1. Comparison of lanes incorporating [γ -³²P]GTP with those incorporating [α -³²P]UMP and [α -³²P]GMP confirms that all significant transcription products begin with the correct nucleotide (G) at position +1. The control lanes show no significant RNA synthesis either in the absence of DNA or in the presence of random sequence (calf thymus) DNA. Although all samples were 0.20 μ M in template, the total amount (by weight) of DNA present in the synthetic template samples (0.005 mg/mL) was much less than that in the samples derived from plasmid (0.4 mg/mL). Nevertheless, for templates with similar initial sequences, the amounts of both abortive and full-length transcripts observed are similar.

The amount and length of abortive transcription products is dependent on the initial (ca. 8–10 bases) sequence of the message. In order to quantitatively assess the amount of abortive cycling in these experiments, regions of the gel corresponding to specific transcripts were excised and counted in a scintillation counter. From the predicted and observed ratios of incorporation of [α -³²P]UMP, [α -³²P]GMP, and [γ -³²P]GTP, most of the small transcripts can be unambiguously assigned. The results, summarized in Table I, confirm that abortive transcripts are the predominant products on a molar basis and that the extent of abortive transcription is dependent on the sequence of the message.

Ratios of Abortive Products from Different Templates. At each base in the sequence, two events are possible during transcription: either the next base is incorporated (elongation), or the complex dissociates to release an abortive product. The fraction of transcription complexes which dissociate at each position in the message can be obtained as the ratio of the amount of observed abortive product to the total amount of products greater than or equal to that length.² The analysis

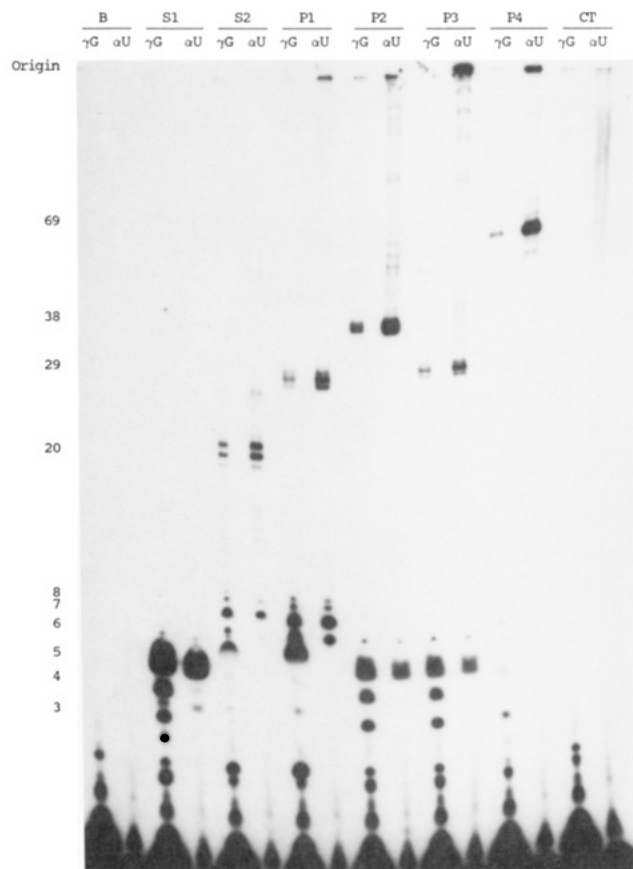


FIGURE 1: Comparison of transcription products from various double-stranded promoter templates. All templates share the same T7 class III promoter consensus sequence from position -17 to position -1:TAATACGACTCACTATA (except P4, which contains C at position -17). Templates S1 and S2 are double-stranded synthetic oligonucleotides containing the promoter consensus sequence upstream to position -17. Templates P1 through P4 are plasmids cut with a restriction enzyme to yield a linearized template coding for a short runoff transcript. P1 is plasmid pIBI24 cut with *Bam*HI. Templates P2 and P3 are derived from plasmids pUCM22 and pUCM22R, respectively, which contain template S1 cloned into the *Bam*HI site of pUC8 in opposite directions. Both were linearized with *Alu*I. P4 is plasmid pUDG8, which contains the T7 ϕ 1.1a promoter cloned into the *Bam*HI site of pUC8, and was cut with *Bam*HI. Lanes marked γ G and α G contained [γ - 32 P]GTP and [α - 32 P]UTP, respectively, as the labeling nucleotide. The assignment of abortive transcripts as discussed in the text and presented in Table I is indicated in the figure. Reaction conditions were as described under Materials and Methods and contained 0.10 μ M enzyme and 0.20 μ M template DNA incubated 10 min at 37 $^{\circ}$ C.

presented in Figure 2 suggests that early abortive termination occurs most often immediately after the incorporation of UMP and least often immediately following incorporation of GMP into the message. Very little abortive products larger than eight bases in length are observed for any of the templates, despite the observations that the product of template S2 has U's at message positions 9 and 11, those of P1 and P2 have U's at position 10, and the full-length message from template P4 has one run of UUU and four occurrences of UU (all beyond position 20).

Studies over a range of enzyme and DNA concentrations (0.01–0.50 μ M) show no significant change in the ratios of abortive to full-length products, provided that nucleotide pools

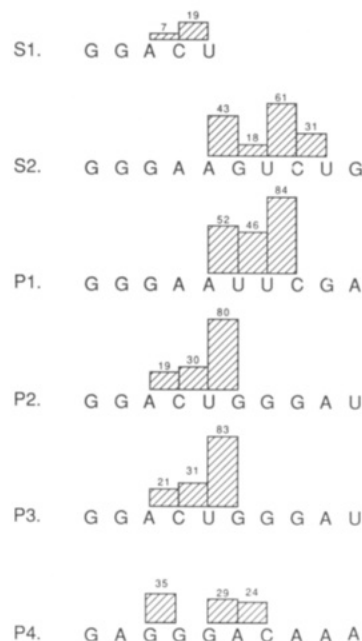


FIGURE 2: Percent termination for each position in the message. The percent of abortive termination at a given position in the message is determined as the amount of abortive product of a given length divided by the total amount of products of that length or longer. A histogram bar placed between two bases in the sequence denotes the percent of transcripts that abortively terminated just prior to formation of the bond between the two bases. The values indicated above for various templates are derived from the data in Figure 1 and Table I. Note that the error in these numbers increases for longer abortive message lengths since fewer total products are being compared.

are not significantly depleted (data not shown). Similarly, comparison of samples taken at times from 1 to 10 min show no measurable time dependence in the steady-state ratios of abortive to full-length products (data not shown). These results suggest that abortive cycling is a general phenomenon that accompanies transcription by T7 RNA polymerase.

Characterization of Runoff Transcripts. We note that for template S2, in all of the studies presented here, the expected unique runoff transcript is in fact heterogeneous, consisting of at least two messages near 20 bases in length. Although the exact nature of these transcripts is not yet known, it is readily seen from the data in Figure 1 and Table I that they are not the result of imprecise starts. The two major runoff transcripts are produced in approximately equal amounts. If this heterogeneity were due to imprecise initiation, a similar doubling of abortive products would be expected; this is not observed.

Although we have not been able to assign these bands precisely to individual runoff products, it is highly probable that these species represent the full-length runoff transcript and prematurely terminated transcripts one or two bases shorter. This result seems to be a general one, since multiple transcripts are detected for some of the other templates studied here [see also Lowary et al. (1986)]. Further studies with different short templates may provide a better understanding of this effect.

Dependence on Reaction Conditions. In order to understand the factors which contribute to abortive cycling, we have compared the effects of temperature, ionic strength, and nucleotide levels on the extent of abortive cycling. It has been previously observed that overall transcription rates for T7 RNA polymerase decrease with temperature (Chamberlin & Ring, 1973; Oakley et al., 1975). We have found that initiation similarly has a strong temperature dependence (Martin and Coleman, unpublished results). It has also been recently

² The total production of transcripts (under identical reaction conditions) varies with the sequence of the template. The sequence dependence of these variations is currently under investigation. However, for the purposes of this study, the relevant variable is the ratio of dissociation to the sum of dissociation and elongation at a given point in the sequence.

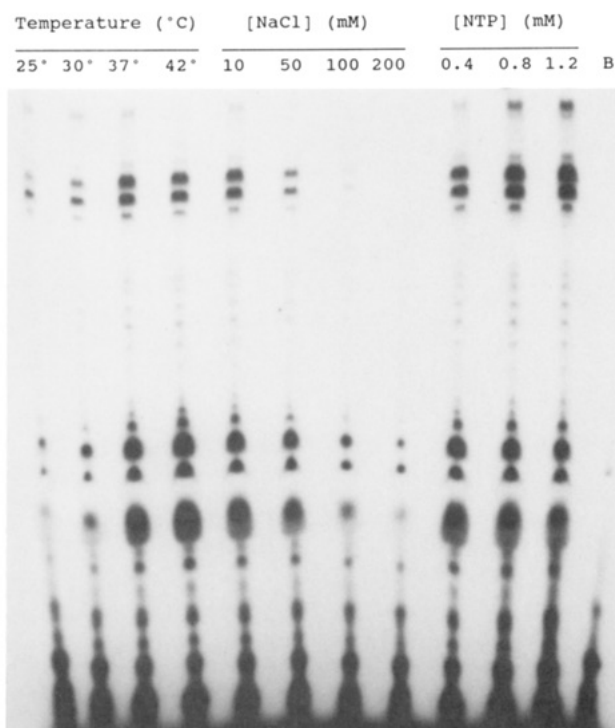


FIGURE 3: Temperature dependence of transcription. All reactions transcribed template S2 under conditions as for Figure 1, except that the incubation temperatures were as indicated in the figure (25, 30, 37, and 42 °C). DNA was preincubated in the assay mixture at the indicated temperature for 3 min (total volume 9 μ L). Enzyme (3 μ L at 4 °C) was then added, and the reaction was incubated for 10 min at the indicated temperature. Ionic strength dependence of transcription. All reactions transcribed template S2 under conditions as for Figure 1, except that the final concentrations of NaCl in the reaction mixture were as indicated in the figure (10, 50, 100, and 200 mM). Effect of nucleotide concentrations on transcription. All reactions transcribed template S2 under conditions as for Figure 1, except that the final concentrations of each of the four nucleotides (the same within a given sample) were as indicated in the figure (0.4, 0.8, and 1.2 mM each ribonucleotide).

Table II: Effects of Reaction Conditions on Abortive and Full-Length Transcription

	total [γ -GTP] (μ M)	% falloff ^a				
		5	6	7	8	20
25 °C	0.7	41	8	38	15	64
30 °C	1.0	39	12	44	17	57
37 °C	3.5	41	16	56	19	49
42 °C	4.6	41	21	64	27	52
10 mM NaCl	3.1	40	17	55	18	51
50 mM NaCl	2.0	45	21	62	25	50
100 mM NaCl	0.8	48	32	71	37	51
200 mM NaCl	0.3	47	37	79	57	62
0.4 mM each NTP	3.5	40	19	56	19	49
0.8 mM each NTP	4.2	34	18	42	13	49
1.2 mM each NTP	4.0	31	18	38	12	47

^a Percent falloff for a transcript of length n is determined as the ratio of the amount of transcripts which dissociate at length n vs the total amount of transcripts with lengths greater than or equal to n . Consequently, the error increases for longer transcripts, since fewer transcripts successfully reach these lengths. In general, the average error for these figures is about 5%.

noted that the relative amount of abortive cycling by the *E. coli* enzyme decreases slightly from 22 to 28 °C and increases significantly from 28 to 37 °C (Straney & Crothers, 1987a). For some of the abortive transcripts produced by T7 RNA polymerase, the relative amount of abortive cycling increases uniformly with temperature, as shown in Figure 3 and summarized in Table II. However, other abortive products (the

Table III: Effects of Template Structure on Abortive and Full-Length Transcription

	total [γ -GTP] (μ M)	% falloff ^a			
		5	6	7	8
S2	4.3	38	23	50	22
S2a	7.5	45	33	48	22
S3	6.2	32	23	46	23
S3a	3.4	23	25	42	17
S3b	1.2	22	16	26	10

^a As defined in Table II. The average estimated error is about 5%, except for template S3b, which shows lower overall incorporation and therefore will have somewhat higher error levels. K_m and k_{cat} for templates S3, S3a, and S3b are predicted to be 0.02, 0.35, and 0.32 μ M and 50, 24, and 18 min^{-1} , respectively (Martin & Coleman, 1987).

five-base transcript in particular) show little temperature dependence. This position and/or sequence dependence suggests that multiple interactions are involved in determining the stability of the ternary complex and that these factors may have different temperature dependencies.

In our normal transcription buffer, the concentrations of NaCl and MgCl_2 are 10 mM and 20 mM, respectively. The dependence of abortive cycling on the concentration of NaCl is shown in Figure 3. Increasing the concentration of NaCl not only decreases overall transcription but significantly increases the probability of abortive termination for only some of the abortive products. In particular, it seems that early falloff products are less sensitive to increasing ionic strength.

Finally, in these studies the individual concentration of nucleotides (0.4 mM each) were well above the measured K_m 's for traditional transcription assays (Oakley et al., 1979). To determine whether the K_m 's for nucleotide incorporation are similar during the early phase of transcription, we uniformly increased all nucleotide concentrations to 0.8 or 1.2 mM each. The results presented in Figure 3 show that a slight decrease in abortive cycling is achieved at 0.8 mM ribonucleotides but that no significant change occurs at 1.2 mM. In all cases, there are still substantial short RNA products, verifying that abortive cycling occurs even at saturating nucleotide levels.

Melting of the DNA Duplex in the Message Region. In order to determine whether melting or reannealing of the DNA double strand in the message region of the DNA template contributes to abortive cycling, we constructed synthetic templates (see Table III and Chart I) containing double-stranded DNA in the promoter region (−17 to −1), but which are single-stranded (coding strand only) throughout the message region. One such construction, template S1a, is analogous to the 22 base pair template, template S1, used in our previous study of initiation (Martin & Coleman, 1987), but contains a single-stranded coding region from +1 to +5. Comparison of the transcription products reveals that, as for the duplex template, the predominant product is the full-length five-base message (data not shown). Steady-state kinetics of this template further reveal that both k_{cat} and K_m are similar to those observed for transcription from the analogous fully duplex template (Martin & Coleman, 1987).

To investigate the effect of a single-stranded template on abortive cycling, we prepared a synthetic template similar to template S2 but which contains single-stranded DNA (coding strand) throughout its 20-base message (template S2a, Chart I). We also prepared a template which is base paired at positions +1 and +2, but which has mismatches at positions +3, +4, and +5 and is single-stranded from +6 to +20 (template S3, Chart I). Comparison of the transcription products derived from these templates with products derived from the fully duplex template reveals very little change in

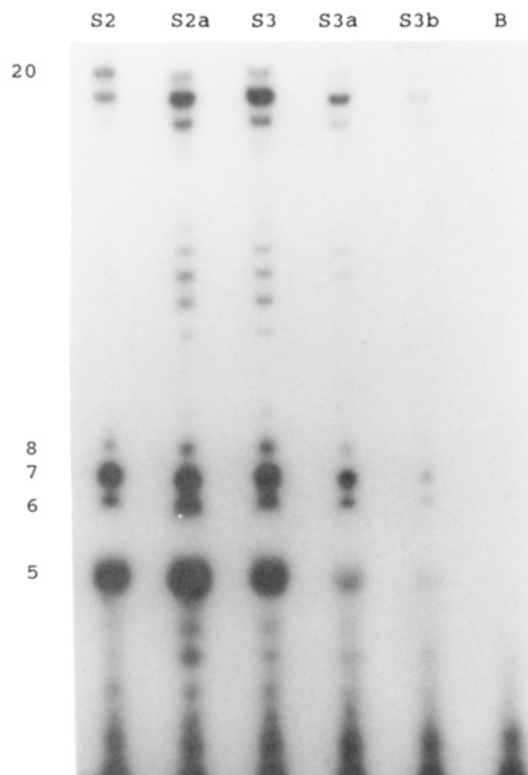


FIGURE 4: Comparison of transcription from templates with altered structure. Templates S2, S2a, and S3 share the consensus promoter sequence, but have altered DNA structure downstream from position +1, as shown in Chart I. Templates S3a and S3b are identical with S3, except that S3a contains a G replacing T in the noncoding strand (only) at position -10 and S3b contains the noncoding strand upstream only to position -14. Both S3a and S3b are predicted to have weakened binding relative to S3 (see text).

the ratio of abortive to full-length transcripts, as presented in Figure 4. This result demonstrates that the absence of the noncoding DNA strand in the message region neither increases nor decreases abortive cycling. Some increase in prematurely terminated products longer than 10 bases is observed, however, possibly due to secondary structure in the single-stranded DNA template.

Effects of Weakened Promoter Binding. It has been proposed for the RNA polymerase from *E. coli* that abortive cycling is caused by strong binding interactions between the enzyme and promoter bases, which serves to inhibit forward procession of the enzyme along the DNA template during transcription of the first bases of the message (Carpousis & Gralla, 1980; Straney & Crothers, 1987a,b). We have previously characterized two deviant T7 promoter constructs which both display weakened promoter binding (Martin & Coleman, 1987). In the first such species the consensus A at position -10 of the noncoding strand is replaced by C. The introduction of this single mismatched base pair results in an order of magnitude increase in the K_m for steady-state kinetics, with only a small decrease in k_{cat} . In the second deviant promoter, the coding strand is extended full length upstream to position -17, while the noncoding strand is truncated upstream at position -14. This change again produces an order of magnitude increase in K_m , while decreasing k_{cat} by about a factor of 2.

We have prepared longer templates with analogously weakened promoter regions (S3a and S3b, respectively, see Chart I), which have message regions identical with template S3 of the current study. Template S3 was shown in the previous section to produce the same abortive products as native template, and so serves as an appropriate control for

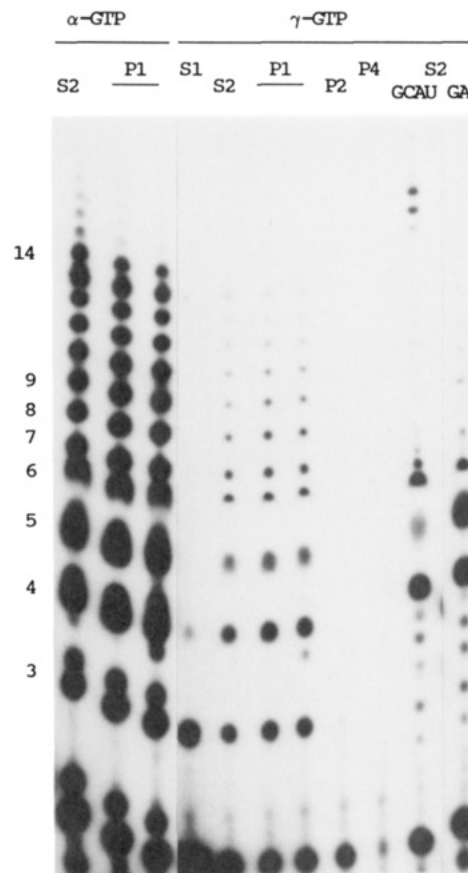


FIGURE 5: Comparison of transcription from various templates in the presence of GTP as the sole nucleotide. The initial message sequences of the templates shown here are GA... (P4), GGA... (S1, P2), and GGGA... (S2, P1), as described in Table I. All reactions were under conditions as for Figure 1, except that for the first nine lanes only the nucleotide GTP was present (with the radioactive label indicated). For comparison, the last two lanes show transcription of template S2 in the presence of all four ribonucleotides and in the presence of only GTP and ATP, respectively.

the following experiments. The results for the weakened promoters S3a and S3b (compared with those of the control in Figure 4 and Table III) suggest that although promoter strength may affect abortive cycling slightly, it cannot be the major determinant in causing premature termination of transcription in T7 RNA polymerase.

Slippage at a Repeating Initial Nucleotide. We have shown in Figure 5 that when transcribing from a template whose message begins GGGA, in the presence of only the nucleotide GTP, T7 RNA polymerase synthesizes a ladder of transcripts with lengths increasing from three bases. In order to better understand this reaction, we have examined transcripts produced from templates which start with one, two, or three repeating nucleotides. The results presented in Figure 5 compare transcription from templates S2 and P1 (both GGGA...), P2 (GGAC...), and P4 (GAGG...) in the presence of GTP as the sole nucleotide. Since significant ladder synthesis occurs only for templates S2 and P1, we conclude that for each addition of GMP beyond 3, the growing RNA chain slides back by one base relative to the DNA, retaining two rG-dC base pairs and reexposing the coding strand C at position +3 of the DNA template.

Stable slippage requires that a minimum of two rG-dC base pairs remain, since no significant ladder production is observed for templates which produce transcripts beginning GGA... or GA.... Addition of the next base (A) allows elongation to position +6 on template S2, and effectively halts ladder pro-

duction, as demonstrated in Figure 5. Slippage at this point would require breaking up to six base pairs in order to re-form three base pairs. This result suggests that slippage will not occur significantly in vivo but is an artifact of limiting the subsequent nucleotide under very specific conditions. We have also found that T7 RNA polymerase is unable to incorporate 3'-*O*-methyl-AMP into transcript (presumably due to steric constraints) and that a similar ladder is produced for transcription from template P1 in the presence of GTP, CTP, UTP, and 3'-*O*-methyl-ATP (data not shown). In contrast, T7 RNA polymerase will incorporate 3'-deoxy-ATP, and as expected, incorporation of this modified nucleotide halts the synthesis of a poly(G) ladder (data not shown).

In order to determine how far the synthesis of an (rG)_n message proceeds, we followed incorporation of [α -³²P]GMP. Although there is a decrease in the amount of message remaining at each step, incorporation of an α label compensates by making longer messages correspondingly more radioactive. The data presented in Figure 5 show that the ladder of (rG)_n products ends abruptly at $n = 14$. While the two templates S2 and P1 share the same message sequence only as far as position +5, they both show abrupt termination of the slippage ladder at $n = 14$. In addition, synthesis on template P1 truncated at position 11 by digestion with *AluI* similarly produces a ladder pattern up to $n = 14$ (data not shown). This further supports the interpretation that the enzyme remains situated at position +3 throughout ladder synthesis and that the product RNA slips in register by one at each cycle. This result also suggests that the growing RNA chain must encounter some blockage at a length of 14 bases.

Similar studies (not shown) were compared for the synthetic promoters S2, S2a, S3, S3a, and S3b (all contain a message starting GGA... and are single stranded in the coding strand region from +1 to +20). These templates all produce results as for P1, demonstrating that slippage is a general property of transcription under the appropriate limited nucleotide pools and that this effect is not influenced by the nature or presence of the noncoding DNA message strand.

DISCUSSION

Transcription by T7 RNA polymerase is both highly specific and highly processive (Chamberlin & Ryan, 1982). Until recently its characterization has relied on transcriptional assays involving the synthesis of very long RNA transcripts. In a previous study, we examined promoter recognition and initiation with a transcription assay involving synthesis of a short five-base RNA transcript (Martin & Coleman, 1987). We now extend these studies to systems with longer, but still relatively short, transcripts. This new system has provided insight into a recently discovered property of transcription by T7 RNA polymerase—the abortive synthesis of very short RNA transcripts. Similar abortive cycling has been observed for *E. coli* RNA polymerase (Carpousis & Gralla, 1980, 1985; Grachev & Zaychikov, 1980; Straney & Crothers, 1987a,b), for RNA polymerase II in HeLa cells (Ackerman et al., 1983), and also for the synthesis of RNA from the polyhedrosis virus and reovirus by the eukaryotic host RNA polymerase (Furuichi, 1981; Yamakawa et al., 1981). It may be that all RNA polymerases will share a common mechanistic origin for this property of transcription. In the current study we examine the significance of abortive cycling in the early stages of transcription by T7 RNA polymerase.

Abortive Cycling Is a Normal Part of Transcription. The results presented in Figure 1 and Table I show that abortive cycling is a major feature of transcription from a T7 RNA polymerase promoter. Contrary to a previous report (Morris

et al., 1986), abortive products are not observed in the absence of template, nor are they produced significantly from non-promoter-containing DNA. In addition, all abortive products appear to be derived from correct initiation at the initial G in the message, as evidenced by their incorporation of [γ -³²P]GTP.

Throughout the process of transcription there must exist a kinetic competition between dissociation of the enzyme–DNA–RNA ternary complex³ and incorporation of successive nucleotides. The results presented in Figures 1 and 2 show that immediately following initiation, significant dissociation is observed. After a minimum number of bases are incorporated (approximately eight for T7 RNA polymerase), the ternary complex proceeds to a much more stable one in which abortive dissociation is greatly reduced. The extent to which this dissociation reduces the number of transcription complexes available for transition to a fully processive ternary complex depends on the sequence of the transcript in this initially transcribed region. From the data summarized in Figure 2, it appears that abortive products are produced preferentially immediately following incorporation of UMP into the message. This may be a result of the uniquely weak interaction characteristic of a dA–rU base pair (Martin & Tinoco, 1980), which might provide a weakening of the forces which correctly orient the last base in the waiting RNA chain. This effect would lead to slower catalysis at that step, allowing dissociation of the complex to compete more effectively with incorporation of the succeeding base.

Dependence of Abortive Cycling on Ionic Strength and Temperature. We have characterized the dependence of abortive cycling on both ionic strength and temperature. The results presented in Figure 3 and Table II confirm that overall initiation decreases with both increasing ionic strength and decreasing temperature. For some positions in the message, the amount of abortive cycling increases coincidentally. However, the extent of this increase varies with position; e.g., abortive falloff following position 5 is almost independent of ionic strength and temperature over the range studied. The fact that these variables affect abortive falloff differently for distinct sequence positions shows that multiple forces must be involved in the competition between procession and dissociation.

Melting and Reannealing of Duplex DNA Do Not Contribute to Abortive Cycling. T7 RNA polymerase will not recognize promoter DNA which is completely single-stranded (Chamberlin & Ring, 1973; Osterman & Coleman, 1981; Martin & Coleman, 1987); initiation of transcription requires the presence of at least part of both strands of the DNA in the promoter region. In the normal process of transcriptional elongation, RNA polymerase must melt the DNA duplex downstream of its current position, and the DNA duplex must then reseat after the transcription “bubble” has passed. Conceivably, difficulties in either process could destabilize the ternary complex and lead to abortive termination. This cannot be the cause of the observed abortive cycling in T7 RNA polymerase, since similar abortive cycling occurs for a template which is lacking the noncoding DNA strand in the message region from +1 to +20 [Figure 4 and Table III, see also Lowary et al. (1986) and Milligan et al. (1987)]. Although possible secondary structure in the resulting single-stranded DNA template might impede efficient elongation of a pro-

³ Note that dissociation of any one component of the ternary complex is sufficient to halt processive transcription. Although the product RNA must dissociate at some point prior to the next round of transcription, it is unknown whether the enzyme and DNA remain associated.

cessive complex (note the slight increase in prematurely truncated products longer than eight bases), the lack of the complementary noncoding strand does not seem to favor either procession or dissociation in early transcription. This result may be useful to workers using T7 RNA polymerase as a tool to synthesize small RNA transcripts from synthetic oligonucleotide templates. One need only synthesize the promoter region (-17 to -1) of the noncoding DNA once, for use with any number of full-length coding strand templates.

Strong Enzyme-Promoter Interactions Are Not the Major Cause of Abortive Cycling. T7 RNA polymerase is specific for a highly conserved set of promoters (Oakley & Coleman, 1977; Dunn & Studier, 1983; Carter et al., 1981). To achieve this selectivity, the enzyme must have strong favorable interactions with at least some of the bases within the promoter. It has been proposed for the *E. coli* enzyme that such strong interactions might act to keep the enzyme from leaving the promoter during the initial stages of transcription, leading to abortive initiation (Carpousis & Gralla, 1980). If retained during the abortive cycling phase, such strong promoter contacts would have to be broken before a stable elongation complex could form (Straney & Crothers, 1987a,b). Comparison of the data for templates S3, S3a, and S3b in Figure 5 and Table III reveals that decreasing promoter binding affinity by an order of magnitude (as reflected in the steady-state K_m) has only a small effect on the extent of abortive cycling observed. While promoter strength may influence abortive cycling somewhat, it cannot be the major contributor to abortive dissociation in T7 RNA polymerase.

Slippage at a Repeating Initial Base. In the course of these studies, a previously unknown property of T7 RNA polymerase was discovered. During transcription of a template containing at least three repeats of the first base, in the absence of the ensuing (nonhomologous) nucleotide, the first nucleotide will continue to be added to form a growing homopolymer (Figure 5). Since homopolymer synthesis is independent of the sequence beyond the nonhomologous nucleotide, we conclude that after incorporation of the third ribonucleotide, the nascent RNA chain must slip in register -1 relative to the coding DNA strand. This allows multiple "correct" incorporations of the nucleotide at position +3 on the DNA template. In general, for a run of n identical nucleotides at the beginning of a template, the n base RNA-DNA duplex can slip by one and still retain $n - 1$ stable base pairs. Since slippage is not observed for templates with $n = 1$ or $n = 2$, we find that significant slippage occurs only for $n \geq 3$; i.e., a minimum of two rG-dC base pairs is required in the heteroduplex.

At each cycle in ladder synthesis there is a significant probability of dissociation, as evidenced by the decay of the ladder pattern. However, the synthesis of (rG) $_n$ decreases significantly at $n = 14$, independent of the sequence of the template (Figure 5). It may be that after progression of the enzyme to position +3, substantial contacts with the promoter are retained which sterically block RNA chain growth beyond 14 nucleotides. After the enzyme has moved further from the promoter these restrictions must disappear, since normal chain growth is unimpeded for even very long transcripts.

Runoff Transcripts Are Not Always Unique. For most of the templates studied, more than one transcript is observed near the expected position of the full-length runoff transcript. This effect has been previously assigned to sequence-dependent imprecise termination, in transcription from plasmid templates containing overhanging 5' ends (Lowary et al., 1986; Milligan et al., 1987). We confirm that slippage at the initial nucleotide or imprecise initiation cannot explain these multiple runoff

transcripts for two reasons. First, in the studies presented here slippage has been observed only when the initial nucleotide is repeated at least 3 times, and then only under conditions limiting addition of the ensuing nucleotide. In contrast, multiple runoff transcripts are observed under conditions in which all four nucleotides are saturating and for various initial sequences. Second, for template S2 approximately equal amounts of two different runoff transcripts are produced, yet no similar doubling is observed in the abortive transcripts. Multiple runoff transcripts are also not an artifact of transcription from DNA with staggered ends, since they are observed for the blunt-end templates S2, P2, and P3 (Figure 1). However, the distribution of runoff transcripts is different for template S2 in the presence and absence of the noncoding strand (Figure 4). In conclusion, multiple runoff transcripts must be a result of imprecise termination at the end of the DNA template, possibly due to a general weakening of the ternary complex from the loss of downstream DNA contacts. The sequence-dependent premature dissociation most likely reflects a kinetic competition similar to that occurring in abortive cycling.

Kinetic Considerations. During all stages of transcription there exists a competition between incorporation of successive nucleotides (elongation) and dissociation of the ternary complex. The data presented in this study show that immediately following initiation dissociation competes favorably with elongation. However, after transition to the fully processive complex, elongation must be strongly favored. It is not immediately clear whether this transition involves a decrease in the unimolecular rate of complex dissociation and/or an increase in the average (pseudo-) unimolecular rate of elongation. However, the observed rates of formation of the various transcripts allow us to place limits on the changes in the rates which accompany the transition to a fully processive complex.

The elongation rate constant for T7 RNA polymerase on long templates has been measured to be about 230 s^{-1} (Golomb & Chamberlin, 1974; Muller et al., 1988). If it is assumed that approximately 90% of transcripts which have escaped abortive cycling successfully produce full-length messages on a 10000-base template, the dissociation rate would have to be less than 0.003 s^{-1} . Were this also the average dissociation rate during the abortive cycling phase of transcription, the observation that less than 50% of the initiated transcripts proceed beyond eight bases to processive complexes places an *upper* limit on the elongation rate during this phase at 0.02 s^{-1} . However, the overall rates of transcript production observed in the current study lie near 0.2 s^{-1} , placing a *lower* limit on the average elongation rate during abortive cycling at about 1 s^{-1} . Thus, the average rate of dissociation of the ternary complex during the abortive phase must be significantly greater than 0.003 s^{-1} , and a large decrease in this rate must accompany the transition to a fully processive complex.

The kinetic considerations presented here argue that immediately following initiation the enzyme-DNA-RNA ternary complex possesses a much higher probability of dissociation than after its transition to a fully processive complex. It may also be that a lower rate of elongation during the abortive phase, especially following the incorporation of UMP, further allows dissociation of the complex to compete more effectively with elongation.

Nature of the Transition to a Highly Processive Ternary Complex. Finally, these data provide insight into what molecular interactions may be responsible for the transition from an unstable initiated complex to a highly processive ternary complex. A simple explanation is that formation of the

DNA-RNA duplex stabilizes the ternary complex. This would predict a steady increase in the stability of the ternary complex with increasing length of RNA, up to the point at which the origin of the DNA-RNA duplex begins to dissociate. The results presented here suggest a more abrupt change in stability of the ternary complex.

An alternate explanation for the sudden change to a more processive complex would include a conformational change in the protein which converts it to an altered form with higher processivity. Recent footprinting studies of transcription in the presence of limiting nucleotides have in fact led to the proposal that a conformational change in the protein occurs after synthesis of between 6 and 15 bases (Ikeda & Richardson, 1986). In these studies, and in similar studies of the RNA polymerase from *E. coli* (Carpousis & Gralla, 1985; Straney & Crothers, 1987a,b), it is assumed that transcription in the presence of limited nucleotides or in the presence of a 3'-blocked nucleotide produces a stably paused ternary complex. The current results show that, although this assumption may well be valid for complexes containing RNA longer than eight bases, for shorter RNA sequences footprinting may depict a time average of the abortive cycling process. Such an averaging would explain the *apparent* increase in size of the protein-DNA interface, which dramatically decreases after synthesis of an 8-10-base message. Consequently, although a conformational change in the protein is possible, it may not be necessary to explain the footprinting results.

Finally, it is possible that nonspecific interactions between the growing RNA transcript and specific regions of the protein stabilize the ternary complex. If the RNA binding region of the protein is distinct from the active site, stabilization would not occur until the RNA chain reaches sufficient length to allow the stabilizing interaction. Evidence for this latter possibility has been obtained recently in a study of proteolytically modified forms of T7 RNA polymerase (Muller et al., 1988).

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

REFERENCES

- Ackerman, S., Bunick, D., Zandomeni, R., & Weinmann, R. (1983) *Nucleic Acids Res.* **11**, 6041-6064.
- Carpousis, A. J., & Gralla, J. D. (1980) *Biochemistry* **19**, 3245-3253.
- Carpousis, A. J., & Gralla, J. D. (1985) *J. Mol. Biol.* **183**, 165-177.
- Carter, A. D., Morris, C. E., & McAllister, W. T. (1981) *J. Virol.* **37**, 636-642.
- Chamberlin, M., & Ring, J. (1973) *J. Biol. Chem.* **248**, 2235-2244.
- Chamberlin, M., & Ryan, T. (1982) *Enzymes (3rd Ed.)* **15**, 87-108.
- Chamberlin, M., McGrath, J., & Waskell, L. (1970) *Nature (London)* **228**, 227-231.
- 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. (1983) *J. Mol. Biol.* **166**, 477-535.
- Furuichi, Y. (1981) *J. Biol. Chem.* **256**, 483-493.
- Golomb, M., & Chamberlin, M. (1974) *J. Biol. Chem.* **249**, 2858-2863.
- Grachev, M. A., & Zaychikov, E. F. (1980) *FEBS Lett.* **115**, 23-26.
- Ikeda, R. A., & Richardson, C. C. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3614-3618.
- King, G. C., Martin, C. T., Pham, T. T., & Coleman, J. E. (1986) *Biochemistry* **25**, 36-40.
- Lowary, P., Sampson, J., Milligan, J., Groebe, D., & Uhlenbeck, O. C. (1986) in *Structure and Dynamics of RNA* (van Knippenberg, P. H., & Hilbers, C. W., Eds.) pp 69-76, Plenum, New York.
- 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, F. H., & Tinoco, I., Jr. (1980) *Nucleic Acids Res.* **8**, 2295-2299.
- McAllister, W. T., Kupper, H., & Bautz, E. K. F. (1973) *Eur. J. Biochem.* **34**, 489-501.
- Milligan, J. F., Groebe, D. R., 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, Amsterdam.
- Muller, D. K., Martin, C. T., & Coleman, J. E. (1988) *Biochemistry* (in press).
- Oakley, J. E., & 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.
- Stefano, J. E., & Gralla, J. D. (1979) *Biochemistry* **18**, 1063-1067.
- Straney, D. C., & Crothers, D. M. (1987a) *J. Mol. Biol.* **193**, 267-278.
- Straney, D. C., & Crothers, D. M. (1987b) *J. Mol. Biol.* **193**, 279-292.
- Yamakawa, M., Furuichi, Y., Nakashima, K., La Fiandra, A. J., & Shatkin, A. J. (1981) *J. Biol. Chem.* **256**, 6507-6514.