Comparison of pausing during transcription and replication

Susan E. LaFlamme, Fred Russell Kramer and Donald R. Mills

Institute of Cancer Research and Department of Human Genetics and Development, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

Received 9 September 1985; Revised and Accepted 1 November 1985

ABSTRACT

Pausing during the transcription of MDV-1 cDNA by Escherichia coli RNA polymerase was compared with pausing during the replication of MDV-1 RNA by Q6 replicase. MDV-1 RNA is able to form many strong hairpin structures, and Q6 replicase pauses after the synthesis of each [Mills et al. (1978) Cell 15, 541-550]. Although the transcripts were virtually identical to MDV-1 RNA, the locations at which RNA polymerase paused were different and apparently were not related to sequences that can form hairpins. These results indicate that hairpin stability, per se, cannot be used to predict the occurrence of pausing during transcription.

Four pauses that occur within a 5-nucleotide region were studied in detail. Insertions and deletions were made in the template DNA to determine the contribution made by the surrounding sequences to these pauses. The results indicate that some of the pauses require the presence of particular upstream sequences, while others are unaffected by the template modifications. Thus, there are at least two different transcriptional pausing mechanisms: one depends on the nature of upstream sequences, while the other is independent of upstream sequences.

INTRODUCTION

The formation of secondary structures strongly influences the rate of chain elongation during RNA-directed RNA synthesis by Q6 replicase (1). Electrophoretic examination of the distribution of the partially synthesized strands that occur during the synthesis of MDV-1 RNA, a template for Q6 replicase (2), indicated that the replicase pauses at a relatively small number of specific sites. Nucleotide sequence analysis of the partially synthesized strands indicated that the pauses occur immediately after the synthesis of sequences that can form strong hairpins (1). This suggested that pausing is due to the formation of hairpins.

A similar correlation between pauses and secondary structures has been observed during DNA-directed RNA synthesis by Escherichia coli RNA polymerase (3-6). Furthermore, during transcription from the E. coli trp operon, the formation of a particular hairpin was shown to be required for a particular pause to occur (4,7-9), and the length of that pause was shown to
be dependent on the stability of the hairpin (10). Several models have been proposed to explain how the formation of hairpins might result in pausing during transcription: hairpins could interfere sterically with the polymerase (4); the formation of hairpins might disrupt the template-product hybrid (10); or hairpins could serve as signals that are interpreted by the polymerase as instructions to pause (11).

In this paper, we compare the pausing that occurs during the synthesis of MDV-1 RNA by Q8 replicase to the pausing that occurs during the synthesis of the same RNA by _E. coli_ RNA polymerase. Both polymerases synthesize a single-stranded RNA product (12,13). The availability of a complete cDNA copy of MDV-1 RNA (14) provided the opportunity to see if the mechanism of pausing was indeed similar during transcription and replication. We therefore constructed a plasmid in which a promoter for _E. coli_ RNA polymerase was inserted immediately adjacent to MDV-1 cDNA, where it would promote the synthesis of MDV-1 (-) RNA. Since the same RNA was synthesized in these two different polymerization systems, the role of the product strand in the mechanism of pausing could be explored. Furthermore, recombinant DNA techniques afforded us the opportunity to alter the DNA template, to determine the contribution made by specific nucleotide sequences and secondary structures to the pausing pattern.

**MATERIALS AND METHODS**

**Enzymes, Nucleotides, and Oligonucleotides** — The following enzymes were purchased: _E. coli_ RNA polymerase (Sigma), Klenow fragment of _E. coli_ DNA polymerase I and bacteriophage T₄ polynucleotide kinase (Boehringer Mannheim), bacteriophage T₄ DNA ligase and nuclease S1 (BRL), restriction endonucleases (BRL and New England Biolabs), ribonuclease T₁ (Calbiochem), and ribonuclease A (Worthington). Unlabeled ribonucleoside triphosphates were purchased from P-L Biochemicals. \([\alpha^{-32P}]GTP\) and \([\gamma^{-32P}]ATP\) were purchased from Ameraham. Cytidylyl (3'-5') adenosine (CpA) was obtained from Sigma. The following linkers were purchased: Eco RI (GGAATTCC) and Xho I (CCTCGAGG) from Collaborative Research and Stu I (CAGGCCTG) from Worthington.

**Preparation of MDV-1 cDNA** — MDV-1 DNA was prepared by using avian myeloblastosis virus reverse transcriptase to synthesize a cDNA from each complementary strand of wild-type MDV-1 RNA, hybridizing the cDNA products to each other to form an MDV-1 DNA duplex, then inserting this DNA into the Eco RI restriction site of pBR322 DNA with the aid of Eco RI linkers (14).
E. coli HB101 was transformed with this plasmid and a clone was selected as a source for the isolation of recombinant DNA.

Construction of pSL1 — MDV-1 double-stranded DNA was isolated by digestion with Eco RI, and purification by gel filtration chromatography on Sepharose 4B (Pharmacia). The single-stranded ends were removed by incubation with 2.4 units of nuclease S1 for every pmole of ends in 200 mM NaCl, 2 mM ZnCl₂, and 50 mM sodium acetate (pH 4.5) at 25 °C for 30 min. The DNA was isolated by extraction with phenol and precipitation with ethanol. Eco RI linkers were then ligated to the ends (15). Following digestion with Eco RI, this DNA was inserted at the Eco RI site of a pBR322-derived plasmid that contained an Xho I linker in place of base-pairs 3608 to 4360 of pBR322 DNA (16). Addition of the linkers restored the Eco RI ends and created a unique Sma I recognition site at one end of the MDV-1 DNA. Clones containing recombinant plasmids were identified by restriction analysis (17). Another DNA fragment, containing an E. coli tryptophan (trp) promoter (18), was obtained by digestion of pDR33 DNA (kindly provided by Dr. George Bennett of Rice University) with Pvu II and isolated from a low-melting-point agarose gel (19). This fragment was then inserted at the Sma I site. Recombinant plasmids containing the promoter directed towards the MDV-1 DNA were identified by restriction analysis. One recombinant was chosen and named pSL1. A template for the synthesis of MDV-1 (-) RNA was obtained by digesting this plasmid with Eco RI and isolating the 260 base-pair template fragment by gel filtration chromatography. The construction of pSL1 was confirmed by DNA sequence analysis (20) and RNA sequence analysis of its transcripts (21). All templates used in these studies were purified in this manner.

Construction of pSL2 — A 297 base-pair fragment containing the trp promoter and MDV-1 DNA was excised from pSL1 DNA by digestion with Xho I and Hind III. This fragment was partially digested with Bbv I, generating a number of smaller fragments. The single-stranded ends were then filled in by incubating 1 unit of the Klenow fragment of E. coli DNA polymerase with 4 pmoles of ends in 50 μl of 6.7 mM Tris-HCl (pH 8.0), 6.7 mM MgCl₂, 10 mM 2-mercaptoethanol, 200 μg per ml gelatin, and 200 μM of each of the four deoxyribonucleoside 5'-triphosphates, at 17 °C for 3 hrs. Xho I linkers were ligated to the ends, the fragments were digested with Xho I and Eco RI, and the fragments were then separated on a low-melting-point agarose gel. Two fragments were isolated: an 86 base-pair fragment containing the trp promoter and one end of MDV-1 DNA, and a 145 base-pair fragment containing...
the other end of MDV-1 DNA. The 145 base-pair fragment was inserted into the pBR322-derived plasmid, described above, in place of its 8 base-pair Xho I-Eco RI fragment, creating intermediate plasmid pSLa. Similarly, the 86 base-pair fragment was inserted into plasmid pMK2004 (22) in place of its 2,121 base-pair Xho I-Eco RI fragment, creating intermediate plasmid pSLb. Plasmid pSLa was digested with Xho I and Hind III to obtain a 174 base-pair fragment encompassing the original 145 base-pair Eco RI-Xho I segment. This fragment was then inserted into pSLb in place of its 400 base-pair Xho I-Hind III fragment, creating plasmid pSL2. This construction resulting in the replacement of a 29 base-pair internal Bbv I segment of MDV-1 DNA with an Xho I linker. The 231 base-pair Eco RI fragment from pSL2 was purified for use as a template. The sequence surrounding the deletion was confirmed by nucleotide sequence analysis of the transcripts.

Construction of pSL3 -- The 231 base-pair Eco RI fragment from pSL2 was inserted into the Eco RI site of a pBR322-derived plasmid, in which base-pairs 3608 to 4360 of pBR322 DNA were replaced by an Eco RI linker. This new plasmid was linearized by digestion with Xho I. The single-stranded ends were filled in by incubation with the Klenow fragment of DNA polymerase, as described above. Stu I linkers (CAGGCCTG) were then ligated to the ends. Since the addition of these linkers created a Pst I recognition site whenever two linkers were added to the same end, the DNA was digested with Pst I to ensure that only one copy of Stu I linker was added at each end. The recircularization of this DNA at the Pst I site created plasmid pSL3. The 260 base-pair Eco RI fragment from this plasmid was purified for use as template. The sequence surrounding the alteration was confirmed by nucleotide sequence analysis of the transcripts.

Construction of pSL4 -- The 260 base-pair Eco RI fragment from pSL1 was purified and digested with Mbo II, generating a 143 base-pair fragment (containing the trp promoter and a portion of the MDV-1 sequence) and a 117 base-pair fragment (containing the remainder of the MDV-1 sequence). The Eco RI ends were filled in and the Mbo II ends were trimmed by incubating the DNA with the Klenow fragment of E. coli DNA polymerase, as described above. Xho I linkers were then ligated to the ends. Following digestion with Xho I, the fragments were inserted into the Xho I site of the pBR322-derived plasmid containing an Xho I linker, described above. A plasmid containing the 143 base-pair fragment was isolated and named pSL4. A 137 base-pair Xho I fragment from this plasmid was purified for use as template.
Construction of pSL5 — A plasmid (kindly donated by L. J. Brunet of Columbia University), containing a Hha I-Pvu II fragment of adenovirus DNA from the region of the major late promoter within map units 16.4 and 16.6 (23), was modified so that it contained an Xho I site and a Sal I site at the ends of the adenovirus sequence. The 54 base-pair Xho I-Sal I fragment was then inserted at the Xho I site of pSL2, creating plasmid pSL5. The orientation of the inserted fragment was determined by restriction analysis. A 294 base-pair Eco RI fragment from this plasmid was purified for use as a template. The sequence surrounding the insertion was confirmed by nucleotide sequence analysis of the transcripts.

In vitro Transcription — 10 to 20 nM template DNA was preincubated with 175 nM RNA polymerase in 150 mM KCl, 15 mM MgCl$_2$, 0.1 mM EDTA, 0.5 mg/ml bovine serum albumin, 1 mM dithiothreitol, 0.25 mM CTP (24), and 60 mM Tris-Cl (pH 8.0) in 20 µl at 37 °C for 5 min to allow the formation of open promoter complexes (25). RNA synthesis was initiated by the addition of 20 µl (prewarmed to 37 °C) containing 125 µM of each of the four ribonucleoside 5'-triphosphates, 150 mM KCl, 15 mM MgCl$_2$, 0.1 mM EDTA, 0.5 mg/ml bovine serum albumin, 1 mM dithiothreitol, and 60 mM Tris-Cl (pH 8.0). The reaction was then incubated at 37 °C. When more than one timepoint was analyzed, a master reaction was prepared and divided into aliquots. Each aliquot was then initiated and terminated separately. RNA synthesis was terminated by the addition of 5 µl containing 150 mM EDTA and 2 mg/ml yeast RNA as carrier. The reaction products were extracted with phenol, precipitated with 2.5 vol of 4 M acetic acid at 0 °C for 20 min, washed twice with 2 M acetic acid, once with 70 % ethanol, and dried in a vacuum. Although these reactions were carried out under conditions of polymerase excess, preliminary experiments indicated that the pausing patterns are identical to those obtained when there is an excess of template.

Electrophoretic Analysis — The partially synthesized transcripts were dissolved in 5 to 10 µl of deionized formamide containing 1 mg/ml xylene cyanol, heated at 100 °C for 3 min to free the transcripts from their templates, chilled to 0 °C, and then analyzed by electrophoresis at 3,000 V on 8 % polyacrylamide slab gels, 0.4 mm thick and 90 cm long, containing 7 M urea, and run in 89 mM Tris-borate (pH 8.3) and 2 mM EDTA.

Nucleotide Sequence Analysis — Each partially synthesized transcript was excised from the gel and eluted by incubation in 400 mM NaCl, 3 mM EDTA, and 20 mM Tris-Cl (pH 7.5) at 37 °C on a shaker overnight, and then precipitated with ethanol. The nucleotide sequence of each [³²P]RNA
intermediate was determined by the two-dimensional electrophoretic separation of its ribonuclease T₄ digest (21). Some of the resulting oligonucleotides were analyzed further by digestion with ribonuclease A or by alkaline hydrolysis (21).

RESULTS

Synthesis of MDV-1 (-) RNA by Transcription — Since secondary structure formation in the product has been implicated in pausing, both during RNA-directed RNA synthesis (1) and during DNA-directed RNA synthesis (3-6), we were interested in determining whether the synthesis of the same RNA in two different polymerization systems would result in similar pausing patterns. The pattern of pausing during the synthesis of MDV-1 (-) RNA by Q₈ replicase was already known (1). To observe pausing during the synthesis of MDV-1 (-) RNA from a DNA template by E. coli RNA polymerase, a plasmid, pSL1, was constructed containing a trp promoter immediately adjacent to and directed towards MDV-1 cDNA. A template for transcription was obtained by digesting pSL1 DNA with Eco RI and isolating the 260 base-pair fragment that contained the trp promoter and MDV-1 cDNA (Figure 1). By using the isolated fragment as a template, MDV-1 (-) RNA synthesis could be observed in the absence of the synthesis of other transcripts. Furthermore, transcription was initiated uniquely, two nucleotides upstream from the MDV-1 sequence. Computer analysis (26) indicated that these two additional nucleotides would not alter the secondary structures that could form in the MDV-1 sequence.

Pausing was observed by sampling a reaction at various times and analyzing the distribution of the partially synthesized transcripts by electrophoresis. Discrete-sized RNAs were observed. The RNAs increased in size over time. Moreover, they were true elongation intermediates, since they could be chased into full-length RNAs (Figure 2). Nucleotide sequence analysis of the transcripts indicated where in the sequence the polymerase had paused. Figure 3 shows the locations of the first seven major pauses, and compares them to the locations at which Q₈ replicase is known to pause. In general, E. coli RNA polymerase paused at different sites than Q₈ replicase, and there was no obvious correlation between the sites at which E. coli RNA polymerase paused and the formation of hairpin structures. Similar observations have been made by other investigators working with different templates (27,28). In our experiments, only one pause occurred at the 3' side of a hairpin stem; and this was the only pause that occurred in the same position during both transcription and replication. Apparently,
Fig. 1. Nucleotide sequence of the template obtained from pSL1 DNA. Bold letters identify the MDV-1 DNA sequence. The top strand is the sense strand. The large arrow shows the site at which transcription begins. The availability of a variety of restriction sites facilitated the construction of modified templates and simplified their verification by restriction analysis.

E. coli RNA polymerase responds to the presence of hairpin structures in a different manner than Q<sub>8</sub> replicase.

Effects of Site-directed Alterations on Pausing -- Pause 2 was the only pause that occurred at the base of a hairpin stem, and in the same location as a Q<sub>8</sub> replicase pause. It therefore seemed likely that, at the other sites, E. coli RNA polymerase responded to different aspects of sequence and structure. Since the template DNA was obtained from a plasmid, recombinant DNA techniques were used to alter the template in a defined manner in order to see the effect that the presence or absence of specific sequences and structures had on the pausing pattern. In particular, the region surrounding pauses 3, 4, 5, and 6 was amenable to the introduction of alterations. The availability of particular restriction sites, both upstream and downstream of this region, permitted the introduction of deletions and insertions.

Deletion of an Upstream Hairpin -- Hairpin C occurs upstream from pauses 3 through 6. To see if this hairpin was responsible for these pauses, it was deleted from the template. Restriction sites for Bbv I flank
Fig. 2. Electrophoretic analysis of transcripts. The first four lanes show the partially synthesized [\(^{32}\text{P}\)]transcripts that were present at various times (in seconds) after the initiation of synthesis. Numbers identify bands that are due to major pauses. The last lane (C) shows transcripts that were labeled during the first 20 seconds of synthesis and then incubated for 5 minutes in the presence of a ten-fold excess of unlabeled ribonucleoside triphosphates.
Fig. 3. Location of the pauses that occur during the synthesis of MDV-1 (-) RNA by E. coli RNA polymerase and Qβ replicase. Lines identify regions in which the exact location of a pause was not determined. Letters identify the hairpins that can form.

the 29 base-pair sequence that forms this hairpin, facilitating its deletion. Plasmid pSL2 contained this modified template. Electrophoretic analysis of the partially synthesized transcripts from the 231 base-pair Eco RI fragment of pSL2 is shown in Figure 4. Nucleotide sequence analysis of the elongation intermediates confirmed that the deletion resulted in the absence of pauses 3 and 4 (as well as pause 2, which is located within the deleted sequence). Significantly, pauses 5 and 6 were still present. Apparently, pauses 3 and 4 require the presence of the sequences that form hairpin C, while pauses 5 and 6 are due to some other element of sequence or structure.

Replacement of Hairpin C with a New Hairpin — Since pausing was observed after the synthesis of hairpin C, but not after the synthesis of hairpins B or D, we were interested in determining whether the particular nucleotide sequence of hairpin C was required for pauses 3 and 4. Therefore, a new hairpin sequence was inserted in place of hairpin C. This was accomplished by adding two copies of a Stu I linker (CAGGCCCTG) at the deletion in pSL2, creating pSL3. The 260 base-pair Eco RI fragment from this plasmid served as a template for the synthesis of transcripts that contained a strong hairpin (hairpin S) in place of hairpin C. Electrophoretic analysis of the partially synthesized transcripts is shown
in Figure 5. The pausing pattern was similar to that obtained with the unmodified template. Furthermore, nucleotide sequence analysis showed that pauses 3 and 4, as well as 5 and 6, occurred. An additional pause (pause 2') occurred at the base of hairpin S, in a position analogous to pause 2 at
Fig. 5. Effect of replacing hairpin C with hairpin S. The autoradiograph shows the elongation intermediates that were present after 10 seconds of transcription (lane M) and compares them to those that were present in a control reaction (lane C). Because strong secondary structures persist in the presence of 7 M urea and there are more base-pairs in hairpin S than there are in hairpin C, the modified transcripts form a tighter structure and migrate slightly faster (29,30). The diagram shows the sequence and probable secondary structure of the modified transcript. The bracket indicates the region that was modified. For clarity, hairpin S and hairpin D are separated in the diagram. Arrows show the locations of the pauses.

The base of hairpin C. These results indicate that those features of hairpin C that are responsible for pauses 3 and 4 are shared by hairpin S.

Deletion of Downstream Sequences — Pauses 5 and 6 were unaffected by the deletion of hairpin C or the substitution of hairpin S for hairpin C.
Fig. 6. Effect of deleting downstream sequences. The autoradiograph shows the elongation intermediates that were present after 10 seconds of transcription (lane M) and compares them to those that were present in a control reaction (lane C). The diagram shows the sequence and probable secondary structure of the truncated transcript. The bracket indicates the region of the deletion. Arrows show the locations of the pauses.

Since modifications to the upstream sequence had no effect on the occurrence of these pauses, we altered the downstream sequence. The entire region downstream from the Mbo II site (beginning 20 base-pairs from the location of pause 6) was deleted from the template, resulting in plasmid pSL4. Electrophoretic analysis of the partially synthesized transcripts from the
Fig. 7. Effect of displacing hairpins. The autoradiograph shows the elongation intermediates that were present after 10 seconds of transcription (lane M) and compares them to those that were present in a control reaction (lane C). The modified transcripts were longer than those in the control (pauses 5 and 6 occurred, even though 33 additional nucleotides had to be synthesized). Furthermore, no significant pauses occurred within the inserted sequence. These results indicate that polymerization is extremely fast in the absence of pausing. The diagram shows the modified transcript (brackets enclose the insertion). Arrows show the locations of the pauses.

137 base-pair Xho I fragment of pSL4 is shown in Figure 6. The pausing pattern was identical to the pausing pattern in the unmodified control. This result indicates that the deleted downstream sequences have no effect on the occurrence of pauses 5 and 6. Therefore, the results obtained so far
suggest that the sequences responsible for pauses 5 and 6 are contained within the local region in which they occur.

**Displacement of Hairpins Farther Upstream** — Pauses 5 and 6 occurred when hairpin C was present; they also occurred when, as a result of the deletion of hairpin C, hairpin B was moved into the position of hairpin C; and they occurred when hairpin S was substituted for hairpin C. To confirm that pauses 5 and 6 were due to local sequences, rather than upstream hairpins, we inserted a heterologous DNA that had little potential for secondary structure formation (26) in place of hairpin C. This modification (plasmid pSL5) placed 65 base-pairs between the region that forms hairpins B and the region in which pauses 5 and 6 occur (hairpin D). Electrophoretic analysis of the partially synthesized transcripts from the 294 base-pair Eco RI fragment of pSL5 is shown in Figure 7. The entire pausing pattern beyond the first pause was shifted upwards. Nucleotide sequence analysis of the transcripts in these slower migrating bands showed that pauses 5 and 6 occurred at the same sequence, now displaced 33 nucleotides downstream. This further suggests that the sequences responsible for the occurrence of pauses 5 and 6 are contained within the local region in which the pauses occur.

**DISCUSSION**

The results indicate that during transcription there is, in general, no correlation between where pauses occur and where structures can form. This was contrary to our expectations, since MDV-1 RNA contains many strong hairpin structures (31,32), and since Q8 replicase pauses after the synthesis of each hairpin (1). Furthermore, these structures are more stable than the extensively studied hairpin responsible for the pause during the transcription of the *trp* operon (10). The absence of a general correlation between pausing and structure formation during transcription of MDV-1 indicates that hairpin stability, per se, cannot be used to predict the occurrence of pausing.

We are aware that alternative secondary structures can form in RNA. Moreover, we have demonstrated that secondary structures can reorganize during replication (32). It is therefore possible that the formation of less-stable alternative structures could account for the pauses that occurred during transcription. However, an analysis of the alternate structures that could form indicates that this is unlikely.

The product strands were virtually identical during replication and
transcription. Therefore, differences in the nature of the template or the polymerase probably account for the different pausing patterns that we observed. Qb replicase copies single-stranded RNA (12), while E. coli RNA polymerase copies double-stranded DNA (13). Furthermore, Qb replicase is 215,000 Daltons (33), while E. coli RNA polymerase is 480,000 Daltons (13). Specific interactions among the product, template, and polymerase might therefore determine which potential hairpins result in a pause during transcription. In particular, the rate at which some structures form during transcription may be different from the rate at which they form during replication. Although pausing during transcription is clearly due to the formation of some hairpins (4,7-9), our results indicate that this is not generally true.

The template DNA was modified to determine the contribution made by specific sequences and potential secondary structures to pausing. The results indicate that some pauses (3 and 4) require the presence of an upstream sequence capable of forming a strong hairpin structure (either hairpin C or hairpin S). However, not every potential hairpin has this effect. For example, pauses 3 and 4 did not occur when hairpin B was moved into approximately the same position as hairpin C (see Figure 4). Thus, the nature of the upstream sequence, its location, or its rate of self-association, determines whether these pauses occur. On the other hand, some pauses were completely unaffected by upstream and downstream modifications (pauses 5 and 6). These results suggest that there are at least two different pausing mechanisms: one depends on the nature of upstream sequences, while the other is independent of upstream sequences.

ACKNOWLEDGEMENTS

We thank John Mack for the preparation of the illustrations. This work was supported by National Institutes of Health Grant GM-32044 and National Science Foundation Grant PCM-82-15902.

REFERENCES