Pausing and Termination of Human RNA Polymerase II Transcription at a Procaryotic Terminator

G. WESLEY HATFIELD,^{1*} JANICE A. SHARP,¹ AND MARTIN ROSENBERG²

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, California 92717,¹ and Laboratory of Biochemistry, Smith Kline & French Laboratories, King of Prussia, Pennsylvania 19406²

Received 9 May 1983/Accepted 18 July 1983

Kinetic analyses of runoff transcription in a cell-free eucaryotic transcription system revealed that the bacteriophage lambda 4S RNA terminator caused human RNA polymerase II to pause on the template and partially terminate transcription of transcripts initiated by the adenovirus 2 major late promoter. Analogous to the procaryotic RNA polymerase, the eucaryotic enzyme terminated just beyond the guanine-plus-cytosine-rich region of dyad symmetry in the terminator sequence. These results suggest that the eucaryotic RNA polymerase II may respond to transcription termination sequences similar to those used by the procaryotic enzyme. However, similar templates containing lambda t_{int} or lambda t_{R_1} terminators did not elicit pausing or termination, suggesting that other features, such as sequence specificity, may also be involved.

In eucaryotes the 3' termini of most RNA polymerase II transcripts are generated by RNA processing and subsequent polyadenylation (5). In yeasts, the polyadenylation of the iso-1-cytochrome *c* transcript is coupled to the termination event (17). In higher eucaryotes, however, it appears that polyadenylation and termination are not coupled, and in fact, with few exceptions, it is not known where termination of these transcripts occurs. The *in vivo* 3' termini of nuclear RNA polymerase II transcripts have been accurately determined for a number of histone genes (7), human U1 RNA (12), and the simian virus 40 major late transcript (6, 15). In these cases, transcription appears to terminate after a region of dyad symmetry in the DNA template. This same type of DNA structure is found in procaryotic termination sequences and has been well defined (13).

Termination sites at the end of procaryotic genes are specified by a DNA sequence, including a guanine-plus-cytosine (G+C)-rich region of dyad symmetry, followed by a string of thymidine residues. Termination of transcription occurs within the string of thymidine residues, resulting in an RNA transcript with a G+C-rich stem loop structure, followed by a short string of uridine residues at the 3' terminus of the RNA transcript. A wide variety of experimental results support the hypothesis that the G+C sequences in the DNA impede the progress of the procaryotic RNA polymerase; the stem loop formation in the transcript triggers the termination event, and the instability of the deoxyaden-

osine-uridine base pairing between the DNA template and the RNA transcript facilitates the release of the transcript. Thus, it appears that the mechanism of transcription termination in procaryotes is based upon the physical chemistry of nucleic acid interactions specified by the DNA sequence and the secondary structural properties of the RNA transcript (1, 3, 4, 10, 16). It is not unreasonable to assume that these same physicochemical parameters might influence the termination of RNA polymerase II transcription in eucaryotes. Therefore, we investigated whether procaryotic termination sites also function as eucaryotic transcription terminators. We found that the procaryotic phage lambda 4S terminator sequence does, in fact, cause human RNA polymerase II to terminate transcription in a cell-free system.

MATERIALS AND METHODS

The transcription reactions were performed essentially as described by Manley et al. (9), except that the ribonucleoside triphosphate concentrations were reduced and a preincubation in the presence of 50 μ g of heparin per ml was included to reduce nonspecific transcription initiations. The standard 25- μ l reaction mix contained the following in microliters: HeLa cell lysate (Bethesda Research Laboratories), 15; 7 mM EDTA, 0.5; 0.25 M creatine phosphate, 1.0; plasmid DNA template, 2.5 (1.0 mg/ml); water, 3.5. This mixture was preincubated at 30°C for 20 min, 2.5 μ l of heparin (0.5 mg/ml) was added, and preincubation was continued for another 10 min. The reaction was initiated by transferring the preincubated reaction mix to a tube containing dried (in vacuo) ATP, CTP, GTP (2.5

FIG. 1. Plasmid vectors used for *in vitro* assay of RNA polymerase II transcription and termination. The expected nucleotide length of the runoff transcripts from *Sma*I-cut pWES01 and *Hinc*II-cut pWES1816 are indicated.

nmol each; final concentration, 100 μ M), UTP (0.25 nmol, final concentration, 10 μ M), and 25 μ Ci of [α - 32 P]UTP (Amersham Corp.).

The reactions were terminated at the indicated times by the addition of 175 μ l of stop mix containing 66 mM sodium acetate (pH 4.8), 0.3% sodium dodecyl sulfate, and 100 μ g of *Escherichia coli* K-12 tRNA per ml.

The inclusion of 50 μ g of heparin per ml in the *in vitro* transcription reaction significantly reduces non-specific background incorporation of radioactivity without affecting RNA polymerase II-specific incorporation.

The RNA transcripts were isolated by three phenol and one chloroform extractions and ethanol precipitation. The pellet was dried *in vacuo* and dissolved in 15 μ l of TBE buffer (0.1 M Tris-borate [pH 8.3], 2.3 mM EDTA). Before electrophoresis, 10 μ l of a solution containing 10 mM Tris (pH 7.4), 0.5% sodium dodecyl sulfate, and 7.5 M urea was added, and each sample was denatured by heating at 95°C for 3 min. Electrophoresis was performed in 6% polyacrylamide gels containing 7.5 M urea (acrylamide to bisacrylamide, 29:1). The electrophoresis buffer was 0.1 M TBE buffer.

RNAse T₁ fingerprinting of the gel-purified transcripts was performed by the methods of Sanger et al. (14).

RESULTS

Construction of plasmid vectors for *in vitro* assay of RNA polymerase II transcription termination. A plasmid vector (pWES01, Fig. 1) was constructed for assessing *in vitro* the ability of a DNA sequence to terminate RNA polymerase II-directed transcription. The vector contains a strong RNA polymerase II promoter site which initiates transcription 211 base pairs (bp) upstream of a unique blunt-end endonuclease restriction site. This restriction site (*Sma*I) can be used for inserting restriction fragments suspected of possessing a transcription termination sequence. To construct this vector, a 2,000-bp *Sma*I-*Hind*III restriction fragment from adenovirus 2 (map position 11.7 to 17.0), which contains the major late promoter (map position 16.45), was first inserted into a pBR322 derivative which contained unique *Sma*I and *Hind*III sites (11). This plasmid was then opened with *Sma*I, and a *Hind*III linker was inserted, thereby allowing us to obtain the adenovirus 2 fragment containing the major late promoter on a *Hind*III restriction fragment. This *Hind*III fragment was

FIG. 2. Nucleotide sequence of the DNA template and RNA transcript from 53 bases before the adenovirus II major late promoter transcription start site (+1) through the *Sma*I site of pWES01 and the *Hinc*II site of pWES1816. Nucleotides 248 through 278 encode the lambda 4S terminator sequence. The brackets above the RNA transcript indicate the major T₁ RNase oligonucleotides shown in Fig. 6.

inserted into the unique *Hind*III site of pK01 (11) to yield plasmid pWES01 (Fig. 1), which contains a unique *Sma*I site 16 bp downstream of the insert (Fig. 1). This is a convenient site to linearize the plasmid for runoff transcription experiments and to insert DNA sequences suspected of possessing an RNA polymerase II termination sequence. A 190-bp *Dpn*I restriction fragment derived from bacteriophage lambda and known to contain the 4S RNA terminator was inserted into the *Sma*I site of pWES01. This plasmid (pWES1816) contains a *Hinc*II site positioned ca. 50 bp downstream of the termination sequence suitable for generating a linear DNA template for runoff transcription experiments. The in vitro RNA polymerase II-directed transcripts expected from *Sma*I-cut pWES01 and *Hinc*II-cut pWES1816 are shown schematically at the bottom of Fig. 1.

Characterization of in vitro runoff transcription. pWES01 was linearized with restriction endonuclease *Sma*I (Fig. 1 and 2). This DNA was used as a template to determine the optimal conditions (described above) for the in vitro synthesis of RNA polymerase II-directed runoff transcripts. The RNA products of this DNA template, resolved by electrophoresis on a polyacrylamide-urea gel, are shown in Fig. 3. Figure 3A shows that a major transcription product of 211 nucleotides was synthesized in the in vitro transcription reaction. Figure 3B shows that this

211-nucleotide-long transcript is RNA polymerase II dependent since its synthesis was inhibited by low concentrations of α -amanitin. Incorporation of label into this transcript is linear for at least 60 min (data not shown). In a separate experiment it was shown that ribonucleoside triphosphates do not become limiting during 1 h of transcription since the same amount of radioactivity was incorporated into the transcript during the first 10 min of the 1-h transcription reaction as during the last 10 min.

Pausing and termination at the lambda 4S RNA terminator. In vitro transcription experiments were performed with plasmid pWES1816 under the same reaction conditions as those used for pWES01. In this case, the plasmid was linearized with restriction endonuclease *Hinc*II, which cuts 325 bp downstream of the adenovirus 2 major late transcription start site and about 50 bp beyond the lambda 4S termination site (Fig. 1 and 2), or *Eco*RI, which cuts approximately 3,920 bp beyond the lambda 4S termination site (Fig. 4). Transcription with templates cut with either *Hinc*II or *Eco*RI resulted in termination at the lambda 4S terminator. The same terminated transcript is observed when [α -³²P]GTP or [α -³²P]CTP is used for radiolabeling (data not shown). This indicates that termination was due to the lambda 4S terminator and was not an artifact resulting from the position of the *Hinc*II end of the template or the low concentration of

FIG. 3. In vitro runoff transcripts of *SmaI*-cut pWES01. (A) Left lane, 250- and 180-nucleotide-long attenuated *ilv* transcripts (8) used as markers; right lane, runoff transcription product of *SmaI*-cut pWES01. (B) Runoff transcription products of *SmaI*-cut pWES01 with (right lane) and without (left lane) 0.5 µg of α -amanitin per ml.

UTP in the transcription reaction mixture. Thus, subsequent experiments were performed with the shorter *HincII*-cut template to monitor both the terminated and read-through transcripts.

The products formed between 1 and 32 min of reaction time were monitored by electrophoresis on polyacrylamide-urea gels (Fig. 5). A 273-nucleotide RNA transcript appeared as a major product within 2 min of reaction time. In contrast, the 325-nucleotide runoff transcript was not apparent until 3 to 4 min after transcription was initiated. Relative intensities of these bands, as measured by scanning with a soft-laser densitometer, are plotted as a function of the time of the in vitro transcription reaction in Fig. 6.

The 273-nucleotide-long RNA transcript was completed within 1 to 2 min of the reaction (1 min as determined from the intercept on Fig. 6). This rate of appearance suggests an average in vitro transcription elongation rate of four to five nucleotides per s. The 325-nucleotide runoff transcript, however, did not appear until ca. 1 min later. Since the unimpeded transcription time from the 273-bp site to the end of the fragment should be about 10 s, the DNA polymerase II molecules must be pausing for ca. 50 s somewhere between the 273-bp site and the end. Since the 273-nucleotide transcript accumulated

in a linear fashion for another 30 min, some of the RNA polymerase molecules must have been terminating at this pause site. A comparison of the rate of formation of the terminated transcript with the rate of formation of the runoff transcript indicates that the efficiency of this termination is approximately 35%. That is, one RNA polymerase II molecule terminates transcription at the lambda 4S termination site for every two molecules that pause at this site and then read through to the end of the fragment.

Characterization of in vitro transcription products. To confirm the identity and further characterize both the 273-nucleotide terminated transcript and the 325-nucleotide runoff transcript, gel-purified [α - 32 P]UTP-labeled in vitro transcripts were subjected to T_1 RNase fingerprint analysis (Fig. 7). Each T_1 oligonucleotide was identified by its characteristic position in the two-dimensional fingerprint and by secondary analysis with pancreatic RNase (14). A comparison of the results with the known DNA sequence (Fig. 2) of the template allowed unambiguous identification and positioning of each oligonucleotide on the template. The results indicate that both the 273- and the 325-nucleotide transcripts initiate at the adenovirus 2 major late promoter with the expected cap-modified oligonucleotide 5A (Fig. 2 and 7). Moreover, the appearance of oligonucleotides 9B, 10, 11B, and

FIG. 4. In vitro runoff and terminated transcripts of *HincII*-cut (A) and *EcoRI*-cut (B) pWES1816. Transcriptions were performed as described in the text, except heparin was not included.

12 in the 325-nucleotide RNA but not in the 273-base transcript indicates that this larger RNA extends through the 4S termination signal to the *HincII* end of the template (i.e., the larger RNA is the "runoff" RNA). Most importantly, the data indicate that the 273-nucleotide-long RNA stops within oligonucleotide 12. T₁ oligonucleotides UpUpGp (position 259 to 261, Fig. 2) and CpGp (position 269 to 270) immediately preceding oligonucleotide 12 are clearly present in the fingerprint of the 273-nucleotide-long RNA (see legend to Fig. 7). Thus, at least the first uridylylate residue of oligonucleotide 12 (i.e., position 271) must be incorporated into the transcript. However, we were unable to detect the very-3'-terminal oligonucleotide of this RNA. Thus, we presume that the 273-nucleotide-long transcript stops before position 275 since a 3'-terminal oligonucleotide longer than this [e.g., (U)₄ < $\frac{\text{OH}}{\text{OH}}$,

FIG. 6. Rate of formation of the terminated and runoff transcripts from *HincII*-cut pWES1816.

(U)₅ < $\frac{\text{OH}}{\text{OH}}$, (U)₆ < $\frac{\text{OH}}{\text{OH}}$] would have been readily identified. We conclude that the terminal transcript stops within the stretch of uridylylate residues in the 4S terminator between template positions 271 and 275 (Fig. 2).

DISCUSSION

We have demonstrated the selective pausing and termination of human RNA polymerase II transcription in a cell-free transcription system. Further, a procaryotic transcription termination sequence causes efficient pausing and partial termination of eucaryotic RNA polymerase II in the G+C-rich region of the bacteriophage lambda 4S transcription termination sequence. Biochemical and physical characterization of the terminated transcript shows that termination takes place near the beginning of the string of thymidine residues in the DNA sequence at approximately bp 273 (Fig. 8). Thus, the eucaryotic polymerase termination site precedes the normal procaryotic termination site (predominantly the adenine residue at bp 278) by only a few nucleotides (Fig. 2 and 7). These results suggest that interactions similar to those involved in procaryotic transcription termination (1, 3, 4, 10, 13, 16) may be operative in eucaryotes. However, it is possible that unidentified DNA sequences surrounding the lambda 4S termination site are required for RNA polymerase II termination since other procaryotic termination sequences, such as the lambda t_{int} and lambda t_R terminators, do not affect RNA polymerase II pausing and termination (data not shown).

Although we have demonstrated the ability of a procaryotic transcription terminator to terminate eucaryotic RNA polymerase II transcription in vitro, we do not know whether this terminator functions in vivo. However, recent

FIG. 5. In vitro runoff and terminated transcription products formed between 1 and 32 min of transcription from *HincII*-cut pWES1816. The 273-nucleotide transcript is the terminated transcript. The 325-nucleotide transcript is the runoff transcript. The origin of other bands is unknown. Linearizing pWES1816 at the unique *EcoRI* site (Fig. 1) eliminates the appearance of the runoff transcript but does not affect the rate of formation or size of the terminated transcript (Fig. 4).

FIG. 7. Two-dimensional T_1 fingerprints of the runoff (A) and terminated (B) transcripts from *HincII*-cut pWES1816. T_1 oligonucleotides U^*UpGp (right uppermost) and CpG^* (left uppermost) labeled by $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ (*) are found in the same relative molar amount in each fingerprint. Thus, template positions 259 to 261 and 269 to 270 must be represented in the 273- and 325-nucleotide-long RNAs (see text).

reports (6, 15) suggest that a DNA sequence in simian virus 40 resembling a procaryotic transcription terminator is an *in vivo* transcription termination site for eucaryotic RNA polymerase II (Fig. 8). Also, the U1 RNA is thought to

terminate *in vivo* at a DNA sequence possessing the properties of a procaryotic terminator (12) (Fig. 8). Finally, Birchmeier et al. (2) have shown that a G+C stem loop structure at the 3' terminus of the histone H2A gene of sea urchins

FIG. 8. Secondary structures of the 3' termini of eucaryotic RNA polymerase II transcripts.

SV40 RNA 5'...cGUAGUG-3b-GGACUGa-28b-3' terminus
 Human U1 RNA 5'...gGUAGUG-2b-GGACUGc-23b-3' terminus
 Lambda 4S RNA 5'...aGUAaUG-6b-GaACUcca-42b-3' terminus

FIG. 9. Comparison of sequence homologies between the 3' terminus of the simian virus 40 major late, human U₁, and lambda 4S transcripts. Nucleotide spacing between sequences is given in bases (b).

is required for transcription termination. Therefore, on the basis of these *in vivo* observations and the *in vitro* data presented here, it appears that RNA polymerase II may, in part, respond to similar structural features as the procaryote RNA polymerase to terminate transcription. However, since the terminators lambda t_{int} and lambda t_R do not terminate RNA polymerase II transcription, the possibility that unidentified DNA sequences surrounding the lambda 4S terminator may be responsible for RNA polymerase II termination is raised. In fact, in the case of the histone genes, a highly conserved sequence of 23 nucleotides has been found at the 3' terminus of all histone gene transcripts. It has been demonstrated that this sequence is required for termination of histone gene transcription (2). This sequence does not exist in the simian virus 40 major late, human U₁, or lambda 4S transcripts. A comparison of these transcripts for sequence homologies, however, shows a sequence common to each (Fig. 9). Although there are, as yet, no data to support that these sequences are important in the termination of these transcripts, it is interesting that such homologies exist.

ACKNOWLEDGMENTS

We gratefully acknowledge the expert technical assistance of Twyla Miner. G.W.H. is also grateful to U. Schmeissner for assistance with the T₁ fingerprint analyses. This work was supported in part by Public Health Service grant GM 24330 from the National Institutes of Health to G.W.H.

LITERATURE CITED

- Adhya, S., and M. Gottesman. 1978. Control of transcription termination. *Annu. Rev. Biochem.* 47:967-996.
- Birchmeier, C., R. Grosschedl, and M. L. Birnstiel. 1982. Generation of authentic 3' termini of an H2A mRNA *in vivo* is dependent on a short inverted DNA repeat and spacer sequences. *Cell* 28:739-745.
- Biro, P. A., and S. M. Weissman. 1979. Recognition and control sequences in nucleic acids, p. 177-245. *In* J. H. Taylor (ed.), *Molecular genetics*, part 3. Academic Press, Inc., New York.
- Crawford, I. P., and G. V. Stauffer. 1980. Regulation of tryptophan biosynthesis. *Annu. Rev. Biochem.* 49:163-195.
- Fitzgerald, M., and T. Shenk. 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. *Cell* 24:251-260.
- Hay, N., H. Skolnik-David, and Y. Aloni. 1982. Attenuation in the control of SV40 gene expression. *Cell* 29:183-193.
- Hentschel, C. C., and M. L. Birnstiel. 1981. The organization and expression of histone gene families. *Cell* 25:301-313.
- Lawther, R. P., and G. W. Hatfield. 1980. Multivalent translational control of transcription termination at attenuator of *ilvGEDA* operon of *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. U.S.A.* 77:1862-1866.
- Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gelfer. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. U.S.A.* 77:3855-3859.
- Martin, F. H., and I. Tinoco. 1980. DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Res.* 8:2295-2299.
- McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg. 1981. A system to study promoter and terminator signals recognized by *E. coli* RNA polymerase, p. 383-417. *In* J. G. Chirikjian and T. Papas (ed.), *amplification and analysis. 2. Analysis of nucleic acids by enzymatic methods*. Elsevier/North-Holland Publishing Co., Amsterdam.
- Murphy, J. T., R. R. Burgess, J. E. Dahlberg, and E. Lund. 1982. Transcription of a gene for human U1 small nuclear RNA. *Cell* 29:265-274.
- Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* 13:319-353.
- Sanger, F., G. G. Brownlee, and B. G. Barrell. 1965. A two-dimensional fractionation procedure for radioactive nucleotides. *J. Mol. Biol.* 13:373-398.
- Skolnik-David, H., N. Hay, and Y. Aloni. 1982. Site of premature termination of late transcription of simian virus 40 DNA: enhancement by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole. *Proc. Natl. Acad. Sci. U.S.A.* 79:2743-2747.
- Yanofsky, C. 1981. Attenuation in the control of expression of bacterial operons. *Nature (London)* 289:751-758.
- Zaret, K. S., and F. Herman. 1982. DNA sequences required for efficient transcription termination in yeast. *Cell* 28:563-573.