TRANSCRIPTIONAL TERMINATION IN VITRO: THE 3'-TERMINAL SEQUENCE OF COLIPHAGE T7 "EARLY" RNA

Gordon G. Peters and Richard S. Hayward
MRC Molecular Genetics Unit and Department of Molecular Biology,
University of Edinburgh, EH9 3JR, Scotland, U.K.

Received October 3,1974

RNA of molecular weight 2.5×10^6 has been transcribed from the "early" region of phage T7 DNA, using highly purified RNA polymerase from Escherichia coli. Its 3'-terminal trinucleotide sequence has been identified as -CpCpC, by a method involving oxidation with periodate and terminal labelling with [^{3}H] $NaBH_4$, coupled with a sequential chemical degradation technique. The sequence found is in agreement with results which we have previously obtained by an independent method, although it conflicts with some earlier reports. These results are discussed in relation to possible features of signals for termination of transcription.

The principal mode of transcription of phage T7 DNA by the RNA polymerase of its host, Escherichia coli, is now well established (1-3). In vivo and in vitro, initiation occurs mainly at a promoter located near the left end of the genome. Termination occurs with high efficiency at a site, t1, located at the distal end of the early region (20.2% from the left end in terms of total DNA length). In vivo, due to the presence of rho factor, some polymerase molecules probably terminate at sites between the promoter and t1 (3,4). With highly purified polymerase in vitro, however, there is little or no termination proximal to t1. The main product is, therefore, RNA corresponding to the whole early region, of molecular weight $2.4\text{-}2.5 \times 10^6$ (2,5).

We are interested in the mechanism of simple (factor-independent) transcriptional termination, and are seeking to determine the DNA nucleotide sequences which signal for it. To this end we prepared T7 RNA terminated at site t1, and hybridised it to its purified template DNA strand. We then extended the 3'-terminus of the RNA, using E. coli DNA polymerase I and [^{32}P] deoxynucleotide substrates, so as to produce a ^{32}P-labelled complementary copy of the DNA template strand sequence, distal to the transcriptional stop site. Our initial results with this approach were highly specific: the RNA

molecule ends with a C residue, and the first base coded by the template beyond this terminus is guanine (5).

Previous workers had published different results for the 3'-terminal nucleoside of RNA transcribed from the identical strain of T7 DNA, and ostensibly terminated at the same (t1) site. Millette and his colleagues observed about 75% U, 10-11% each of A and G, and only 3.5% C (6). Maitra et al presented data indicating about 60% U and 27% C (7), or (elsewhere) 88% U and 9% C (8). Although our controls showed that our result was unlikely to have been biassed by the method used (5), it was desirable to check the 3'-terminal sequence of the RNA by an independent, well-established technique. We have therefore, like Millette (6), applied periodate oxidation and reduction with tritiated borohydride to label the terminal nucleoside. Our results again show that the final nucleoside is C (at least 90%). The two preceding nucleosides are also C.

EXPERIMENTAL

Bacteriophage T7 was obtained from Dr. W. Studier. The isolation of T7 DNA, and the preparation of highly purified RNA polymerase core and sigma from E. coli MRE600 (ribonuclease I^-) have been described previously (5). For use, core was recombined with sufficient sigma to give maximum activity on T7 DNA. The endoribonucleolytic activity of this preparation was negligible (5). T7 RNA labelled with ^{14}C, and terminated at site t1, was prepared in a reaction containing (per ml) 50 μmoles Tris. HCl (pH 7.8), 150 μmoles KCl, 1 μmole EDTA, 2 μmoles dithiothreitol, 0.4 μmoles each CTP, GTP, UTP and [^{14}C] ATP (Amersham; 3.1 mCi/mmole), 125 μg T7 DNA (5 pmole), and 24 μg reconstituted RNA polymerase holoenzyme (approx. 15 pmoles active enzyme (5)). After 3 min at 37°C, $MgCl_2$ was added (to 10 mM), to initiate transcription. 10 min later rifampicin (Lepetit) was added (50 μg/ml) to prevent further initiation, while allowing completion of nascent RNA molecules during a further 5 min incubation. After addition of EDTA to 50 mM, and Sarkosyl NL35 (Geigy) to 1% (v/v), and a further 2 min incubation at 37°C,

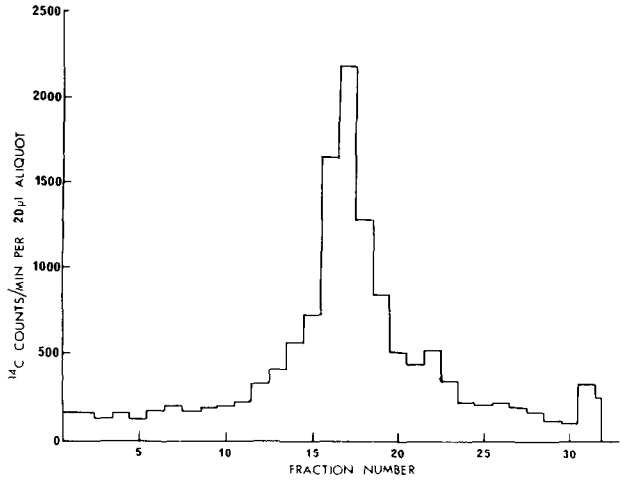

Fig. 1. Purification of T7 RNA. After treatment with EDTA and Sarkosyl (see text) the reaction mixture was applied to two 54 ml linear gradients of 15-40% glycerol, 0.1 M NaCl, 50 mM Tris. HCl (pH 7.8), 1% Sarkosyl. After centrifugation (60000g, 17h, 4°C) fractions were collected, and their acid-insoluble [^{14}C]RNA content was determined in aliquots. The bottom of the gradient is on the left.

---

the solution was chilled and the RNA purified as shown in Fig. 1. Sterile, siliconised glassware and sterilised reagents were used wherever possible, throughout all the experiments to be described.

The peak fractions from the glycerol gradient (16 to 18) were pooled, and the RNA was precipitated with 0.3 M potassium acetate (pH 5.4) and 66% ethanol (-20°C, 2h), collected by centrifugation, redissolved in 10 mM Tris. HCl (pH 7.5), 10 mM NaCl, 1 mM EDTA, and reprecipitated to remove traces of glycerol. Analysis of a sample of the purified RNA by agarose-acrylamide gel electrophoresis indicated the expected molecular weight of $2.4\text{-}2.5\times10^6$, as observed previously (5). We have also shown that RNA prepared as above hybridises to the expected, r-strand of T7 DNA, and close to one end of that strand (5).

Terminal labelling was carried out essentially as described by Mansbridge et al. (9). The RNA was redissolved in 150 µl 10 mM sodium phosphate (pH 6.0), and 15 µl of fresh 50 mM $NaIO_4$ was added to oxidise the 3'-terminal ribose. After 1h at 20°C (in the dark) 15 µl of ethylene glycol was added to destroy

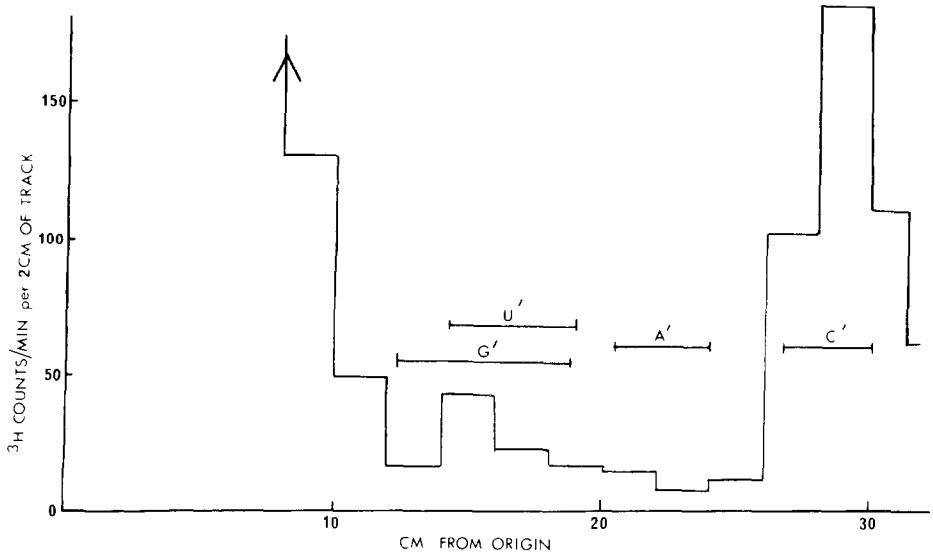

Fig. 2. Identification of the 3'-terminal nucleoside of T7 RNA. The ^{3}H-labelled terminal nucleoside trialcohol present in the neutralised hydrolysate (see text) was identified by ascending chromatography on DE81 paper, using water as solvent. The four marker trialcohols, synthesised as described by Khym and Cohn (10), were co-chromatographed in the same track, and subsequently located under UV light (shown by lettered bars on drawing). Segments of track were counted in the scintillant of Petri (11) using ^{3}H / ^{14}C discrimination settings. Background counts were subtracted before plotting. All ^{14}C and some unknown ^{3}H -labelled material remained near the origin. No ^{3}H was found in the C-trialcohol region unless the RNA had first been hydrolysed with alkali.

---

excess periodate. 30 min later, a portion containing 20 μg RNA was dried *in vacuo*. The RNA was redissolved in 100 μl of 0.5 M sodium phosphate (pH 7.0) containing about 10 mCi of freshly dissolved [^{3}H]$NaBH_4$ (Amersham, 5.9 Ci/mmole). The mixture was left overnight at 20°C in a dark fume cupboard, then passed through a 5 ml column of Sephadex G75, equilibrated with 20 mM NH_4HCO_3 (pH 7.5). The peak fractions of excluded RNA were pooled and dried *in vacuo*. After hydrolysis with 0.2 M KOH, 18h x 37°C, to liberate the ^{3}H-labelled 3'-terminal nucleoside trialcohol, the chilled mixture was neutralised with $HClO_4$, centrifuged to remove $KClO_4$, and analysed (Fig. 2). The results indicate that the 3'-terminal nucleoside is C (at least 92%), with at most 8.4% of U and/or G. From the ratio of ^{3}H in C-trialcohol to ^{14}C in 2'(3')AMP (near the origin), and assuming an RNA molecular weight of 2.5×10^6 with 25%

A residues, the efficiency of terminal labelling was about 42%. (This figure is approximate; the specific activity of $[^3H]NaBH_4$ particles is variable, according to the manufacturers). In another experiment, the terminal nucleoside was found to be at least 88% C, with no definite indication of any other residue. In this case the efficiency of labelling was about 90%.

To determine the nature of the penultimate nucleotide we used a method based on that of Steinschneider and Fraenkel-Conrat (12). 44 µg of the above periodate-oxidised RNA in 60 µl of 0.01 M sodium phosphate (pH 6.0) was mixed with 180 µl of 0.44 M aniline hydrochloride (pH 5.0, prepared with freshly distilled aniline), and 25 µl of 0.1 M sodium acetate (pH 5.0), and left for 5h at 20°C, to release the 3'-terminal oxidised nucleoside residue. The RNA was recovered by precipitation with ethanol/acetate, and drained thoroughly. It was redissolved in 200 µl of 50 mM Tris.HCl (pH 7.5), 10 mM $MgCl_2$, and incubated for 30 min at 37°C with 15 µg of ribonuclease-free *E. coli* alkaline phosphatase (Worthington BAPF) to remove terminal phosphates. After two extractions with redistilled phenol, and two with ether, the RNA was again ethanol-precipitated. It was then subjected to periodate oxidation as before, and half of the product was labelled with $[^3H]NaBH_4$, and analysed as described previously to identify the 3'-subterminal nucleoside of the original T7 RNA (Fig. 3A). The other half was retreated with aniline and phosphatase as above, and subjected to a third periodate oxidation, followed by ^3H-labelling and identification of the third nucleoside away from the original RNA end (Fig. 3B). Although we cannot exclude the possibility of some heterogeneity in the latter case, the efficiency of terminal labelling (assuming intact chains) was 90-100% in both experiments. Thus it seems clear that the second and third last nucleosides of this T7 RNA are C.

DISCUSSION

The above results indicate that T7 RNA, terminated at site t1, has

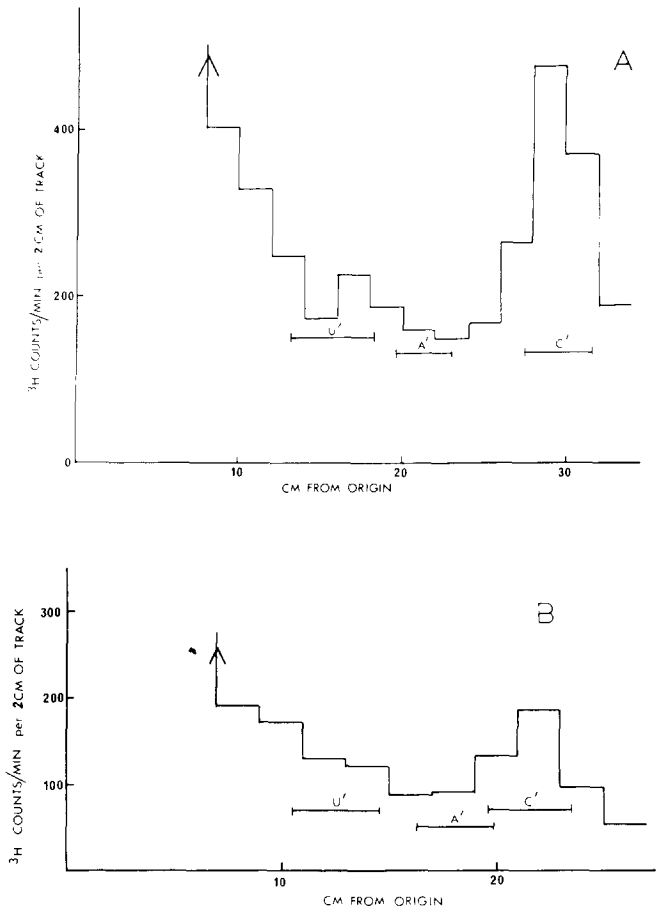

Fig. 3. Identification of the second and third nucleosides away from the 3'-terminus of T7 RNA. After one (A) or two (B) cycles of terminal nucleotide elimination, the new terminal nucleosides were labelled by periodate-oxidation and $[^3H]NaBH_4$-reduction. The labelled nucleoside trialcohols, released by alkali, were then analysed as in Fig. 2.

-CpCpC as its 3'-end sequence. The finding of C as the final nucleoside agrees with our previous conclusion, obtained by the independent "DNA extension" technique (5). Recently, further exploitation of that technique has shown that the penultimate nucleoside is also C (13), again agreeing with our present results. We do not understand why earlier workers found more U than C at ostensibly the same RNA terminus (6-8). We have previously discussed some possible explanations (5); they include genuine strain differences, or inaccuracies arising from insufficient characterisation of the product, and from possible nucleases or other

contaminating activities in the RNA polymerase used.

A general conclusion of our work is that RNA polymerase, by itself, terminates accurately at a unique stop site in the t1 region of T7 DNA. The RNA terminus produced in this case is -CpCpC (followed by an untranscribed G). In two other known cases of simple (factor-independent) transcriptional termination, on the DNA of coliphage λ, the RNA terminus produced is $(Up)_6A$ (14,15). This suggests that the base pairs coding for the last three nucleotides of the RNA are not involved in the DNA sequence which actually causes RNA polymerase to terminate. Alternatively, if involved in defining the stop site, these base pairs must be ambiguously recognised. A priori, the true termination signal sequence(s) might lie within the transcribed portion of the DNA, and/or a short distance beyond the end of the RNA-coding unit.

The above λ species provide examples of the repeatedly observed occurrence of runs of uridylic acid residues near the 3'-ends of RNA molecules. Other examples have been reported in small molecules isolated from Ø80-infected (16) or uninfected (15) E. coli, although in these cases one cannot yet be sure whether the termini are produced at the time of transcription, or by some post-transcriptional cleavage in vivo. Lebowitz et al (14) suggested that formation of a run of A·U and T·A basepairs in the template·product complex might facilitate dissociation of product from template at termination. Our T7 result would not generate any such hypothesis. However, we have found that there are at least three pyrimidines at the end of the T7 RNA. There is evidence that RNA polymerase may transcribe certain regions of DNA more slowly than others (17), and that a series of U or especially C residues in the product would be likely to produce some deceleration (18,19). It is highly unlikely that such a mechanism could, by itself, produce the kind of accurate factor-independent termination observed on T7 and λ- DNA. However, it is conceivable that runs of transcribed pyrimidines might be an auxiliary feature of termination regions. The consequent deceleration of the polymerase might well help it to "recognise"

(bind) the hypothetical true termination signal, which would then define the stop site. If so, deceleration sequences might also be expected near rho or other factor-dependent termination sites. Moreover, their length might show some correlation with the efficiency of termination at particular sites. We hope to extend our studies to test such hypotheses.

We thank Drs. W. Brammar, K. Murray, and J. Scaife for discussion and help, and Lepetit Pharmaceuticals for the gift of rifampicin.

REFERENCES

1 Dunn, J. J., and Studier, F. W. (1973) Proc. Natl. Acad. Sci. U.S.A. 70, 1559-1563.
2 Simon, M. N., and Studier, F. W. (1973) J. Mol. Biol. 79, 249-265.
3 Summers, W. C., Brunovskis, I., and Hyman, R. W. (1973) J. Mol. Biol. 74, 291-300.
4 Davis, R. W., and Hyman, R. W. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 269-281.
5 Peters, G. G., and Hayward, R. S. (1974) Eur. J. Biochem. 48, 199-208.
6 Millette, R. L., Trotter, C. D., Herrlich, P., and Schweiger, M. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 135-142.
7 Maitra, U., Lockwood, A. H., Dubnoff, J. S., and Guha, A. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 143-156.
8 Maitra, U. (1970) Biochem. Biophys. Res. Commun. 41, 1255-1260.
9 Mansbridge, J. N., Crossley, J. A., Lanyon, W. G. and Williamson, R. (1974) Eur. J. Biochem. 44, 261-269.
10 Khym, J. X., and Cohn, W. E. (1960) J. Am. Chem. Soc. 82, 6380-6386.
11 Petri, W. H. (1972) Analytic. Biochem. 48, 442-447.
12 Steinschneider, A. and Fraenkel-Conrat, H. (1966) Biochemistry, 5, 2735-2743.
13 Peters, G. G., and Hayward, R. S. (1974) Symposium on Regulation of Transcription, Ninth FEBS Meeting, Budapest (in press)
14 Lebowitz, P., Weissman, S. M., and Radding, C. M. (1971) J. Biol. Chem. 246, 5120-5139.
15 Ikemura, T., and Dahlberg, J. E. (1973) J. Biol. Chem. 248, 5033-5041.
16 Pieczenik, G., Barrell, B. G., and Gefter, M. L. (1972) Arch. Biochem. Biophys. 152, 152-165.
17 Darlix, J.-L., Sentenac, A., and Fromageot, P. (1968) Biochim. Biophys. Acta 166, 438-458.
18 Manor, H., Goodman, D., and Stent, G. S. (1969) J. Mol. Biol., 39, 1-29.
19 Krakow, J. S., and Fronk, E. (1969) J. Biol. Chem., 244, 5988-5993.