JOURNAL OF VIROLOGY, Oct. 1973, p. 882-886
Vol. 12, No. 4

Transcription Unit Mapping in Bacteriophage T7

I. In Vivo Transcription by *Escherichia Coli* RNA Polymerase

ALAN R. BRÄUTIGAM AND WALTER SAUERBIER
Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado

Received for publication 21 May 1973

Premature termination of transcription at UV lesions in DNA permits a study of the sequential order of transcription by *E. coli* RNA polymerase of the genes in the early region of T7. This analysis is extended to the transcription by *E. coli* polymerase of the late region of T7 by deleting the early terminator. The results demonstrate the existence of a single large transcription unit spanning the early region, with a promotor located at the left end of the T7 genome. Furthermore, there are no initiation sites for *E. coli* RNA polymerase in the late region. When *E. coli* polymerase transcribes the late region, it does so exclusively by initiation at the early promotor.

The early region of the bacteriophage T7 genome is transcribed by *E. coli* RNA polymerase, whereas the late region is exclusively transcribed by T7 RNA polymerase, which is the product of gene 1 (2). The early region comprises the leftmost 20% of the map (7) and codes for at least six proteins. When it is transcribed in vitro by using *E. coli* RNA polymerase a single RNA species of molecular weight 2.2×10^6 is obtained which can be cleaved by RNase III to yield a set of RNA species which is apparently identical to that produced when the early region is transcribed in vivo (5).

This is consistent with the finding by Millette (9) and by Davis and Hyman (4) of a single in vitro promotor at the left end of the early region, and with the observations of Schweiger et al. (11) on the sequential synthesis of early T7 proteins.

This investigation reports an analysis of the in vivo transcription of T7 DNA by the RNA polymerase of *E. coli*. Genes are localized relative to the promotor through the polar effect of ultraviolet light-induced photoproducts in T7 DNA on the transcription of promotor-distal genes. This method was suggested by the observation of Sauerbier et al. (10), that UV irradiation of bacteriophage T4 DNA led to release of the *E. coli* RNA polymerase at the site of the UV photoproduct, without re-initiation beyond the photoproduct. Shortened mRNA molecules resulted, the length being inversely proportional to dose of irradiation. Similar polar effects of UV irradiation of phage DNA on transcription were observed with T7 DNA (W. Sauerbier, unpublished data). Extremely high UV doses have been shown to competitively inhibit the initiation of transcription by T7 RNA polymerase in vitro (3), however, these doses are approximately 1,000 times larger than those used here.

MATERIALS AND METHODS

Bacteriophage and bacterial strains. The bacterial host used in all experiments was *E. coli* B_{s-1}, and infections were carried out in dim yellow light to avoid photoreactivation. T7 *am*342 and T7 *am*342, LG37 were generously provided by F. W. Studier (12). *Am*342 is an amber mutant in gene 1, the early gene coding for T7 RNA polymerase. It produces a slightly shortened, nonfunctional T7 RNA polymerase. The double mutant *am*342, LG37 contains the above amber mutation together with LG37, a deletion mutation which deletes gene 1.3 (ligase) and the early terminator located at the far right end of the early region. Bacteriophage lysates were prepared in M9 medium (1) by using *E. coli* 011′ (su⁻) as the host, concentrated by centrifugation, and suspended in M9 buffer.

Conditions of infection and labeling. *E. coli* B_{s-1} was grown at 37 C to a density of approximately 5×10^8 cells per ml in M9 medium and was irradiated with UV light to 3,500 ergs/mm^2. To 1-ml samples were added 10^{10} T7 phage which had been UV irradiated at various doses.

For experiments with T7 *am*342, a 1.5-min adsorption period at 37 C was allowed, followed by a 3-min labeling period. For experiments with the double mutant *am*342, LG37, a 4-min adsorption period was allowed, followed by a 7-min labeling period. Proteins were labeled with ^{14}C amino acids from an algal protein hydrolysate (Schwarz/Mann) present at 5 μCi/ml final concentration. To ensure completion of labeled protein chains, Casamino Acids (Difco) were

added at the end of the labeling period to a concentration of 0.5% and incubation was continued for 2 min.

UV irradiation of bacteriophage. T7 phage stocks were diluted to a concentration of 7.5×10^{10} ml in M9 buffer and a 4-ml volume was irradiated with UV in doses ranging from 0 to 80 ergs/mm^2 (in increments of 20 ergs/mm^2) for *am*342, LG37, and ranging from 0 to 320 ergs/mm^2 (in increments of 20 ergs/mm^2 and larger) for *am*342.

Gel electrophoresis. The ^{14}C-labeled protein samples from T7-infected cells were prepared for and banded by sodium dodecyl sulfate polyacrylamide gel electrophoresis according to the slab-gel method of Studier (13). Samples of 20 μliters were loaded in each slot. All buffers and conditions were those of Studier (13), except that a lower voltage was used (40 V for 10% gels, 50 V for 15% gels). Gels were dried and subsequently exposed to 2DC X-ray film (Du Pont Cronex) for periods of 2 or more days, taking care not to exceed the linear region of film sensitivity. The films were scanned on a Beckman analytrol film densitometer, and areas under the curves were determined by using the integrator supplied with the instrument.

Graphical presentation of data. With UV-irradiated template DNA of T7, the probability of transcription proceeding from an RNA polymerase binding and initiation site (promotor) through a length L along the DNA is simply the probability that no RNA chain terminating UV hit has occurred on the DNA within this length. This probability is given by the 0th term of the Poisson distribution which may be written e^{-Kd} where d is the UV dose, K is the exponential slope of the UV inactivation curve, and K is proportional to L. Consequently, for a given gene which terminates at length L from its promotor, the fraction of mRNA molecules whose transcription has been completed through the end of that gene should decrease exponentially with dose. Synthesis of a polypeptide which will band on polyacrylamide gel electrophoresis at the position of the complete gene product requires the existence of the completed mRNA of that gene as a template for translation. Therefore, protein production for the gene in question, as measured by band density of polyacrylamide gel electrophoresis, should also decrease exponentially with UV dose. Thus, a semilog plot of amount of a given gene product versus UV dose should yield a straight line with negative slope K.

Since K is proportional to L, a graph of the K value for each of a set of genes within the same transcription unit versus physical distance along the T7 map should be linear. The slope of this line would be characteristic of the sensitivity of the RNA polymerase to termination of RNA chain elongation by UV photoproducts in the DNA. The slope would be identical for all transcription units utilizing that polymerase provided that the UV sensitivities of the DNA segments are equal. Consequently, two types of graphs will be presented in this paper: (i) semilog plots of protein production for individual genes as a function of UV dose, (ii) a graph in linear coordinates of the slopes resulting from these UV inactivation curves for each of a number of T7 genes versus physical position of the genes along the DNA. The graph should be linear within any one transcription unit and should have the same slope for all such units transcribed by a single polymerase. From this graph the location of each promotor is inferred.

RESULTS

Transcription map of the early region of T7. When cells were infected with T7 *am*342 phage which had been ultraviolet irradiated at a series of increasing doses, a differential reduction of synthesis of early proteins was observed (Fig. 1). (An average of 1 lethal hit on the entire T7 genome is achieved at 18 ergs/mm^2. This corresponds to 90 ergs/mm^2 for 1 lethal hit in the early region alone.) At a dose of 320 ergs/mm^2 less than 2% of the initial amount of the gene 1.3 protein is still being completed. This is in contrast to the gene 0.3 protein which is essentially unaffected by the same dose. Comparison of the UV-inactivation curves of the proteins coded by genes 0.7 and 1.3, which are nearly identical in molecular weight but quite different

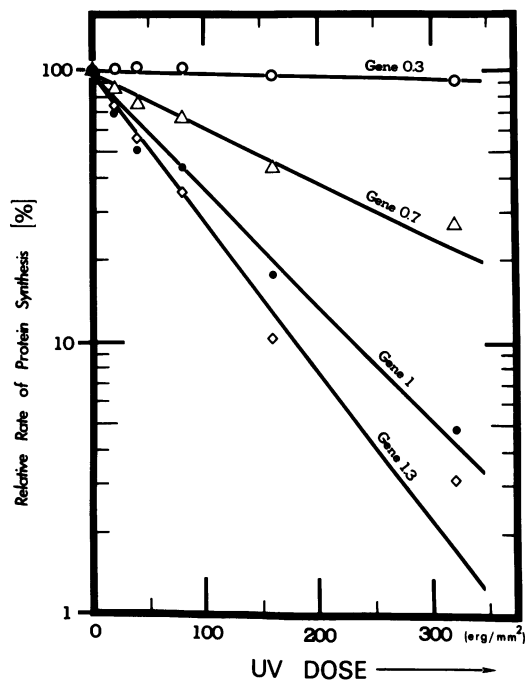

FIG. 1. *Relative rates of synthesis of early T7 proteins in am342 infected E. coli B_{s-1} as a function of UV dose received by the phage prior to infection. The points of the graph are obtained from evaluating autoradiograms of proteins synthesized after infection with gene 1 deficient T7. The abscissa gives the UV dose to the extracellular phage; the ordinate the normalized rate of synthesis of several early proteins. Symbols: open circles, gene 0.3; open triangles, gene 0.7; solid circles, gene 1; open diamonds, gene 1.3.*

in the rate at which their synthesis falls off with UV dose, demonstrates that the variation in sensitivity to UV is not merely a function of gene size.

The data of Fig. 1 permit evaluation of K, the slope of the UV-inactivation curve, for each of the early genes represented. When these K values are plotted as a function of gene position along the physical map (Fig. 2) a straight line results, which begins at or near the left end of the early region and continues through the end of gene 1.3, on the far right end of the early region. Thus, the data permit the conclusion that there is only one transcription unit spanning the early region from the end of gene 0.3 through the end of gene 1.3. This entire region is not interrupted by any small transcription units since if any of these had existed they would have interrupted the continuity of the large unit, and this would have been readily seen.

Extension of transcription by E. coli polymerase into the late region of T7. To confirm the method and to check for segments of the late DNA which could serve as promotors for host polymerase, we determined the differential UV sensitivity for the synthesis of proteins in the late region, when this region is transcribed by *E. coli* polymerase. For this purpose a double mutant deficient in gene 1 and containing LG37, a deletion which omits the early termination signal, was used. The time course of late protein synthesis in the *am*342, LG37 double mutant (in which all transcription is by the host polymerase) is greatly different from that observed when the late region is transcribed by the T7-coded late polymerase (13). The temporal appearance of late proteins in infections involving this double mutant appears to be consistent with a straightforward readthrough from left to right.

If this were the case then genes in the late region, when transcribed by *E. coli* RNA polymerase, should be markedly UV sensitive, and the gradient of K (the exponential slope, which is an index of the sensitivity of the synthesis of each individual protein to UV irradiation of the DNA) as a function of length should be a continuation of the slope presented in Fig. 2. Any sites in the late region which could be recognized as a promotor by the *E. coli* polymerase should be readily apparent in such a graph, unless extremely weak.

Use of the above double mutant for measurements of the sensitivity of late region protein synthesis to UV irradiation of phage DNA led to the data shown in Fig. 3. The solid line is a replotting of Fig. 2 for comparison. The slope of the line for the late region (broken line) is seen to be identical with that found independently for the early region. The data do not demonstrate late region promotor sites for *E. coli* polymerase; rather, they imply that transcription of the late region by the host polymerase is achieved only by read-through from the early region. The data of Fig. 3 also confirm that the slope obtained in Fig. 2 is not an artifact.

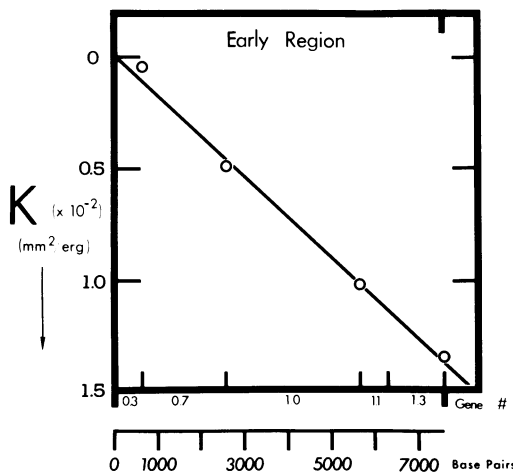

FIG. 2. *UV sensitivity of production of T7 early proteins versus physical position of genes in the early region. Abscissa: early region of the T7 genome in base pairs and position of early genes. Ordinate: UV sensitivity of early genes of T7 (slopes K). The K value for each gene was calculated from the data of Fig. 1 by using the equation* $K = (\ln P_0/P)/d$ *where* P_0 *= relative rate of protein synthesis at zero UV dose and* P *= relative rate of protein synthesis at UV dose* d. *Physical map is that of Dunn and Studier (5), in which gene sizes were determined by acrylamide gel electrophoresis.*

DISCUSSION

When E. coli is infected by increasingly UV irradiated T7 the amount of each individual gene product synthesized decreases exponentially as a function of UV dose (Fig. 1). A plot of the exponential slope K of each gene product versus map position of the gene yields a straight line (Fig. 2). This linear relationship of K versus map position for genes within the same transcription unit is predicted from the known effect of UV-induced photoproducts on the transcription of DNA by *E. coli* RNA polymerase (8, 10).

When the early and the late regions of T7 are transcribed in vivo by *E. coli* RNA polymerase a plot of the exponential slope K of each gene product versus map position of the gene yields two contiguous straight lines of identical slope for the entire genome. This slope is characteris-

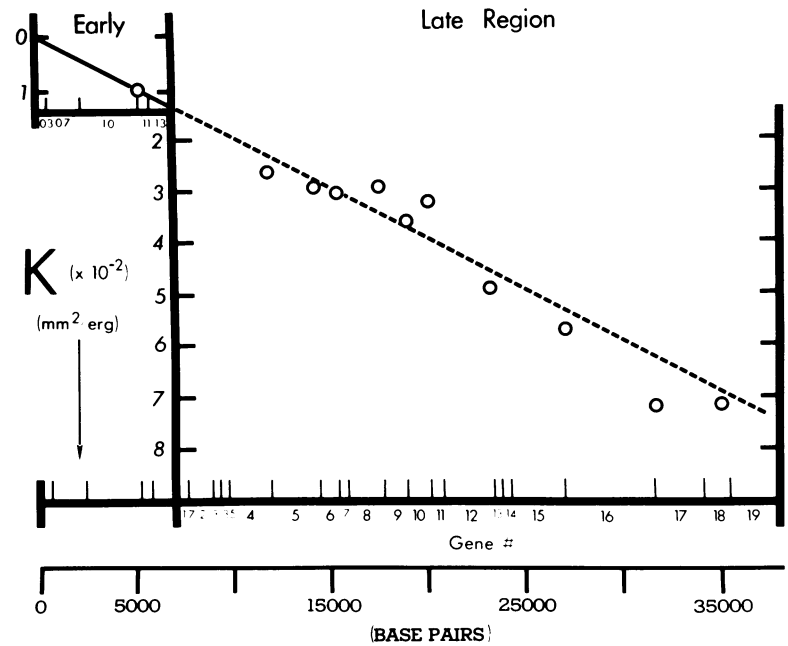

FIG. 3. *UV sensitivities of late genes (slopes K) in T7 am342, LG37 versus gene position on the physical map of the T7 genome. Abscissa: genome of T7 in base pairs and position of known T7 genes. Ordinate: UV sensitivities of T7 genes (slopes K). Figure 2 has been superimposed on Fig. 3 in the upper left hand corner. The broken line is a continuation of the slope from Fig. 2 into the late region. The early gene, gene 1, was evaluated together with the late genes to verify that the slope in the early region determined from am342, LG37 is identical to that obtained from am342 (Fig. 2). K values were calculated from the relative rates of protein synthesis as a function of UV dose by using the method of Fig. 2. Gene positions and gene sizes are those given by Studier (13), except for genes 2 and 18 which are of unknown size. We have arbitrarily assigned them molecular weights of 40,000. Since the proteins corresponding to genes 17 and 19 co-electrophoresed on our gels, the data point obtained for this band has been plotted at a position corresponding to the relative amounts of gene 17 protein versus gene 19 protein.*

tic of the host polymerase. A similar slope has been found when the late region is transcribed by T7 RNA polymerase (unpublished data). Linear slopes, identical to those of Fig. 2 and 3, have been obtained for the early region of T4 DNA in vivo, where transcription is by *E. coli* polymerase (6).

It should be pointed out that the results shown in Fig. 2 would not reveal a very small transcription unit of several hundred base pairs at the beginning of the genome because of the inaccuracy of the method. Since gene 0.3 is of this size, we cannot rule out the possibility that gene 0.3 is actually transcribed separately from the rest of the early region, with promotors both before and after this very small gene. However, if we take into account the possibility of a very short, separate transcription unit at the beginning of the genome, the data demonstrate that apart from this possibility, there is only one transcription unit spanning the early region from approximately gene 0.3 through the end of gene 1.3. This is so because any small transcription units located other than at the very beginning of the large unit would of necessity interrupt the continuity of the large unit, and this would be readily apparent. Consequently, with the above reservation in mind, we conclude that there exists one single transcription unit covering the entire early region, at the left of which is a single promotor.

This finding, of a single promotor region for *E. coli* RNA polymerase near the left end of T7 DNA, is in agreement with the fact that an RNA molecule of 2.2×10^6 molecular weight was transcribed from T7 DNA in vitro in the system used by Millette (9), and with the electron microscope analysis of the in vitro transcription complex by Davis and Hyman (4), in which a specific initiation site for *E. coli* polymerase near one end of the DNA was found.

The observation reported here, of a single early promotor for *E. coli* RNA polymerase on T7 DNA in vivo, combined with the recent report by Dunn (5) of specific in vitro cleavage of the single transcription product of the early

region by RNase III, suggests a mechanism whereby the spectrum of early mRNA species found in vivo could be achieved. Their production could be explained by transcription of the complete early region into RNA and subsequent specific cleavage. This mechanism would lead to equimolar amounts of message for each gene, unless there were selective degradation of certain mRNA species. Equimolar amounts of the early mRNA species corresponding to genes 0.7, 1.0, and 1.3 and a 10-fold excess of the RNA corresponding to gene 0.3 have recently been reported (R. L. Millette, personal communication; 14). Since gene 0.3 is very small and is the first protein of the early region, it is possible that it is contained within its own very short transcription unit. An extremely short unit would be undetectable by the present method if it were at the very beginning of the early region, although we can conclude that there are no such units at intermediate positions.

The early mRNA species from in vivo infections with T7 have recently been identified and mapped (14). We are now using the method described here to study the differential effect of UV irradiation of DNA on the in vivo synthesis of the different mRNA species by quantitating RNA production versus UV dose. This will then be correlated with the effect of UV irradiation of T7 DNA on protein production reported here. The transcription units for the T7-induced RNA polymerase are currently being investigated by the same methods.

ACKNOWLEDGMENTS

This investigation was supported by National Science Foundation grant GB-21413 and by a Public Health Service training grant no. GM00781 of the Department of Biophysics and Genetics (no. 545), University of Colorado Medical Center.

We are particularly indebted to F. W. Studier for supplying us with mutants of T7.

ADDENDUM IN PROOF

E. Minkley and D. Pribnow have recently suggested that in vivo the early region of T7 consists of five separate scriptons (J. Mol. Biol., vol. 77, p. 255-277, 1973). Also, R. I. Young and G. Smith (Biochem. Biophys. Res. Commun., vol. 53, p. 952-959, 1973) present results which they interpret to be in favor of a polyscriptonic transcription of the early region T7.

LITERATURE CITED

1. Anderson, E. H. 1946. Growth requirements of virus-resistant mutants of *Escherichia coli* strain "B." Proc. Nat. Acad. Sci. U.S.A. **32**:120-128.
2. Chamberlin, M., J. McGrath, and L. Waskell. 1970. New RNA polymerase from *E. coli* infected with T7. Nature (London) **228**:227-231.
3. Chamberlin, M., and J. Ring. 1973. Characterization of T7-specific ribonucleic acid polymerase. II. Inhibitors of the enzyme and their applications to the study of the enzymatic reaction. J. Biol. Chem. **248**:2245-2250.
4. Davis, R. W., and R. W. Hyman. 1970. Physical locations of the *in vitro* RNA initiation site and termination sites of T7M DNA. Cold Spring Harbor Symp. Quant. Biol. **35**:269-282.
5. Dunn, J. J. and F. W. Studier. 1973. T7 early RNAs are generated by site specific cleavages. Proc. Nat. Acad. Sci. U.S.A. **70**:1559-1563.
6. Hercules, K. and W. Sauerbier. 1973. Transcription units in bacteriophage T4. J. Virol. **12**: 872-881.
7. Hyman, R. W. 1971. Physical mapping of T7 messenger RNA. J. Mol. Biol. **61**:369-376.
8. Michalke, H., and H. Bremer. 1969. RNA synthesis in *Escherichia coli* after irradiation with ultraviolet light. J. Mol. Biol. **41**:1-23.
9. Millette, R. L., C. D. Trotter, P. Herrlich, and M. Schweiger. 1970. *In vitro* synthesis, termination, and release of active messenger RNA. Cold Spring Harbor Symp. Quant. Biol. **35**:135-142.
10. Sauerbier, W., R. Millette, and P. B. Hackett, Jr. 1970. The effects of UV irradiation on the transcription of T4 DNA. Biochim. Biophys. Acta **209**:368-386.
11. Schweiger, M., P. Herrlich, and R. L. Millette. 1971. Gene expression *in vitro* from DNA of bacteriophage T7. J. Biol. Chem. **246**:6707-6712.
12. Studier, F. W. 1969. The genetics and physiology of bacteriophage T7. Virology **39**:652-574.
13. Studier, F. W. 1972. Bacteriophage T7. Science **176**:367-375.
14. Summers, W. C., I. Brunovskis, and R. W. Hyman. 1973. The process of infection with coliphage T7. VII. Characterization and mapping of the major *in vivo* transcription products of the early region. J. Mol. Biol. **74**:291-300.