Transcription Unit Mapping in Bacteriophage T7

II. Proportionality of Number of Gene Copies, mRNA, and Gene Product

ALAN R. BRÄUTIGAM AND WALTER SAUERBIER

Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado 80220

Received for publication 24 January 1974

The effect of UV irradiation of bacteriophage T7 on in vivo early RNA synthesis has been studied by direct quantitation of the gene-specific RNA transcripts. The results show that the early region of phage T7 is transcribed from left to right as a single unit. Furthermore, gene inactivation, the UV sensitivity of synthesis of gene-specific RNA, and the UV sensitivity of synthesis of the corresponding proteins all follow pseudo first-order kinetics in multiply infected cells, demonstrating a random statistical correlation between both transcriptional sampling of gene copies and translational sampling of the resultant RNA transcripts. In addition, these simple kinetics imply an absence of positive feedback mechanisms compensating for the differential decline of individual early gene products in cells multiply infected with phage T7.

Recently we presented a radiological method for locating the promoter sites for the in vivo transcription of bacteriophage T7 DNA by *Escherichia coli* RNA polymerase (2). This method would appear to be of general validity, requiring only that the RNA polymerase be released at the site of UV lesions and that UV irradiation not create artificial promoter sites. These requirements are fulfilled in in vitro transcriptions (17) and in the four in vivo systems which we have recently studied: T7 (2), T4 (6), *E. coli* (see also ref. 13), and mammalian cells (P. Hackett and W. Sauerbier, manuscript in preparation). Application of the method requires a known genetic map and a quantitative assay of a transcription product of each gene. This assay could in principle be either a measure of the transcribed RNA itself or protein translated from this RNA. Quantitation of either can be achieved after resolution by polyacrylamide gel electrophoresis.

In the first paper of this series (2), we applied this method to analysis of in vivo transcription of the early and late regions of phage T7 by *E. coli* RNA polymerase. Analysis of the relative rates of gene-specific protein synthesis versus UV dose to the phage was utilized to determine the distance of a particular gene from its promoter. We showed that in vivo the early region of T7 is transcribed from a single promoter located at the left end of the genome (2). We assumed proportionality between the rate of gene-specific protein synthesis and the total

number of copies of any given gene available for transcription in the infected culture. It was argued that the results verified all assumptions which were implicit in the assay methods. Because we used bacteria which were multiply infected by T7, the assumption of proportionality of number of active gene copies and rate of gene-specific protein synthesis was nontrivial.

Identification by gene number and map position of the mature, early T7 mRNAs by Summers et al. (22) allowed us to reinvestigate the mode of in vivo transcription of the early region of T7 by assaying for the transcription products themselves. The mature T7 mRNA's are produced by cleavage of the 2.2×10^6 molecular weight transcript by RNase III (4, 5). We reconfirm that there is only one promoter utilized in the in vivo transcription of the T7 early genes and that, even in a culture of multiply infected cells, there is strict proportionality between the number of DNA copies of a particular gene lost by UV irradiation, the reduction in the amount of RNA synthesized from these DNA copies in a given time, and the amount of protein synthesized from this RNA within the same time period.

MATERIALS AND METHODS

Bacteriophage and bacterial strains. The bacterial host used in all experiments was *E. coli* B₊; T7 am342 was generously provided by F. W. Studier (21). Am342 is an amber mutant in gene 1, the early gene coding for T7 RNA polymerase. It produces a slightly shortened, nonfunctional T7 RNA polymerase.

Bacteriophage lysates were prepared in M9 medium (1), using *E. coli* 011' (su⁺) as the host. Phage were concentrated by centrifugation and resuspended 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, whereupon 15 ml was irradiated with UV light (254 nm) at an incident dose of 5,800 erg/mm², after which the cells were incubated for an additional 5-min period at 37 C. UV irradiation and all subsequent operations were carried out in dim yellow light to avoid photoreactivation. To 2-ml samples was added 2×10^{10} T7 am342 phage which had been UV-irradiated at various doses. After a 1-min adsorption period, 5,6-³H]uridine (Schwarz/Mann; specific activity, 42 Ci/mmol) was added to each sample of phage-infected cells to a final concentration of 20 μ Ci/ml. After a 5-min labeling period, infected complexes were chilled, pelleted, washed once with cold M9 buffer, and resuspended in 0.4 ml of 3 mM EDTA containing 1% sodium dodecyl sulfate and 10% sucrose. Infected complexes were then lysed by heating to 85 C for 2 min.

UV irradiation of bacteriophage. A T7 am342 phage stock was diluted to a concentration of 10^{11} phage/ml in M9 buffer, and a 4-ml volume was UV-irradiated with incident doses of 0, 40, 80, 160, and 320 erg/mm².

Gel electrophoresis of RNA. The lysed ³H-labeled samples from infected complexes were loaded directly onto 2.6% polyacrylamide disc gels (final concentrations were 2.6% acrylamide and 0.12% *N,N'*-methylene-bis-acrylamide prior to polymerization) containing no agarose. The procedure is that of Loening (11), except that plastic gel tubes, 8 mm in diameter and 60 mm long, were substituted for the smaller glass ones. Samples of 50 μ liters were loaded onto each gel and electrophoresed for 7 h at 70 V (10 to 11 mA/gel). Gels were sliced (1 mm) and digested at 50 C in 0.5 ml of 30% H₂O₂ containing 1% concentrated NH₄OH. Thereafter, 10 ml of Kinard's scintillator (9) were added, and the slices were counted by liquid scintillation spectroscopy. RNA bands were identified from molecular weights given by Simon and Studier (19), by using mouse L cell 28S and 18S ribosomal RNA and *E. coli* 23S and 16S ribosomal RNA as internal markers. The electrophoretic mobility as measured on our gels, when plotted versus logarithm of molecular weight, yields a single straight line for the four T7 early RNA species and the four markers. Relative amounts of RNA were calculated from gel peaks by summing the ³H counts per minute contained in the peak (after appropriate background correction).

RESULTS

Effect of UV irradiation of T7 on early region RNA synthesis in vivo. In order to test in vivo for the sensitivity of gene-specific RNA production as a function of distance between the gene and its promoter, we performed experiments of the following type. Samples of *E. coli*

B_{s-1}, in which bacterial RNA synthesis was greatly reduced by UV irradiation prior to infection, were infected with T7 am342 which had been UV-irradiated at a series of increasing doses. RNA was pulse-labeled with [³H]uridine, and then isolated and electrophoresed on 2.6% polyacrylamide gels. Typical gel electrophoresis patterns of RNA for several UV doses are shown in Fig. 1. In looking at these patterns, it is obvious that overall RNA synthesis decreases with increasing UV irradiation of the phage; however, it is also clear that the rate of decline of RNA synthesis with dose is vastly different from one RNA species to another. RNA transcribed from gene 1.3, for example, falls rapidly with increasing UV dose, this peak having completely disappeared by 320 erg/mm², whereas that from gene 0.7 shows much less sensitivity, in spite of the similar size of these two RNA species.

FIG. 1. Polyacrylamide gel electrophoresis patterns of T7 early RNA from am342-infected *E. coli* B_{s-1}. Infecting phage had been irradiated at incident UV doses of 0, 160, and 320 erg/mm² (see righthand side of Fig.). RNAs were labeled with [³H]uridine present at 20 μ Ci/ml final concentration in the medium (M-9) from 1 min to 6 min after infection. Direction of electrophoresis is from left to right. RNAs are identified by gene number (bottom of figure); peak G is supposedly the larger of the two guanine triphosphate-initiated cleavage products transcribed from the beginning of the genome. The molecular weights of the T7 early messenger RNAs are: gene 0.3, 208,000; gene 0.7, 598,000; gene 1, 975,000; gene 1.1, 208,000; and gene 1.3, 403,000 (Table 4 of ref. 19). The ordinate gives the ³H counts per minute in the gel fractions. Host RNA synthesis was reduced by UV irradiation prior to infection, and RNA from uninfected bacteria shows a flat distribution of radioactivity not exceeding 100 counts per minute (not shown).

When the amount of each RNA species as a function of UV dose was quantitated from such gels, the pattern, as shown in Fig. 2, of differential reduction of early region RNA synthesis was obtained (Fig. 2). The curves are seen to follow simple exponential kinetics (see Discussion). If each gene were transcribed as a separate transcription unit, target theory would predict UV sensitivities for each species proportional to gene size. This would imply that RNA production from gene 1.0 should be approximately three times more sensitive than that from gene 1.3; however, in actuality, it is gene 13 transcription which is more sensitive (Fig. 2).

Transcription map of the early region of T7 as assayed by direct quantitation of RNA.

FIG. 2. Relative rates of synthesis of early T7 RNA species in am342-infected *E. coli* B_{s-1} as a function of incident UV dose received by the phage prior to infection. Points of the graph were obtained by calculating the relative amounts of ³H RNA, synthesized during a 4-min pulse, in polyacrylamide disc gel patterns such as the ones shown in Fig. 1. The abscissa gives the UV dose to the extracellular phage; the ordinate gives the normalized rate of synthesis of several early RNAs. Uppermost curve represents a composite of two coelectrophoresing RNA species. Dashed line is the curve through the experimental points for the RNA of gene 1.3; solid line shown for this RNA species represents the actual exponential decay of this RNA after correcting the experimental points (see Appendix and Fig. 6). Symbols: ○, genes 0.3 + 1.1; △, gene 0.7; ●, gene 1; ◇, gene 1.3.

The exponential slope of the UV-inactivation curve, K , may be evaluated from Fig. 2 for each RNA species and these K values plotted versus the physical map. The result (Fig. 3) is seen to again yield a straight line of slope similar to that found previously (2) when UV sensitivities were determined by measuring gene-specific protein synthesis. This implies the existence of a single large transcription unit for the early region of T7, with the possible exception that gene 0.3 could be transcribed separately (see ref. 2).

DISCUSSION

Shapes of dose-effect curves. When UV-irradiated DNA is transcribed by an RNA polymerase, the probability of transcription of a gene which terminates at length x' from the promotor is given by e^{-Kd} , where K is the exponential slope of the UV-inactivation curve and is proportional to x' , and d is the UV dose. Due to the accumulation of shorter-length RNA fragments with increasing dose (see Appendix), this simple exponential dependence is not experimentally observable; the curves expected are shown in Fig. 6. The only RNA band which would be appreciably affected by this accumulation of RNA fragments is that corresponding to gene 1.3. The experimental data shown in Fig. 2 confirm these predictions. Curves are simple exponentials, except that the points corresponding to the rate of synthesis of gene 1.3 mRNA describe a gradually upward curving line.

Cleavage by RNase III of partial length transcripts. The full-length transcript of the T7 early region is cleaved by RNase III to yield the mature early mRNAs (4, 5). It is assumed in the consideration given in the Appendix that RNA molecules need not be full-length transcripts of the entire early region in order to be cleaved by RNase III. This assumption is certainly correct, because the UV sensitivity of production of all T7 early RNA species shown in Fig. 2 would have to be the same, and the curves would all have the exponential slope found for gene 1.3, if only full-length transcripts could be cleaved. There also is no indication in our experiments (Fig. 2) that the rate of cleavage is reduced. Such an effect would have led to "shouldered" dose-effect curves.

Comparative evaluation of gene survival using protein assays (2) and RNA assays. Comparison of the present results with the corresponding results obtained from analysis of the differential reduction of individual early gene protein synthesis caused by UV irradiation of T7 (Fig. 1 of ref. 2) shows that the results obtained by using the two different assays are

very similar. The relative order of UV sensitivity of the early genes is the same by both assays, and curves for genes 0.7, 1.0, and 1.3 are nearly identical, whereas that for RNA of genes 0.3 + 1.1 is somewhat steeper than the curve obtained by analysis of protein of gene 0.3, as would be expected because this curve represents a composite of the slopes of two RNA species which coelectrophorese.

In comparing these two sets of data, it should be kept in mind that conditions of host irradiation and label period were not identical, and the pool kinetics for the two types of labeled precursors are probably different. Nonetheless, the two different experimental assays yield the same results (cf. Fig. 1 of ref. 2 with Fig. 2 of this paper).

The transcriptional map of the T7 early genes. Values for the exponential slope, K , defined as $(1/d) \ln(R_0/R)$ may be calculated for each RNA species from the data of Fig. 2. R_0 represents the amount of RNA synthesized from nonirradiated, and R from UV-irradiated templates at dose d . According to theory (see Appendix), a plot of these slopes versus the location of the respective genes on the physical map should be linear for any group of genes transcribed from the same promoter, with intercept at or near $K = 0$. Figure 3 shows a plot of the K values determined from Fig. 2 versus the physical map of the early region of T7.

A linear relationship is seen to occur and nonspecific depression of RNA synthesis by UV light appears to be minor. Comparison of Fig. 3 with the analogous graph obtained by analysis of the differential effect of UV irradiation of T7 on protein synthesis (Fig. 2 of ref. 2) reveals that the two are quite similar. The slope of this line corresponds to one chain-terminating hit per 1,000 base pairs per 630 erg/mm² for the RNA assay, whereas the value found for the protein assay is 1/1,000/560. Thus, there is good agreement between the two assay methods, the difference being within experimental error. The conclusion drawn from Fig. 3 is that the early region of T7 is transcribed *in vivo* as a single large unit, at least under the experimental conditions used here.

Proportionality of number of gene copies, amount of mRNA, and amount of gene product. Our assay for the reduction of synthesis of gene-specific RNA, or gene-specific proteins in a large population of multiply infected cells, reveals pseudo first-order loss of gene expression as a function of UV dose to the extracellular phage. We must, therefore, conclude that whenever transcription of a gene is blocked on any of the multiple DNA copies per cell, the result is a reduction in synthesis of the corre-

FIG. 3. UV sensitivity of production of T7 early RNA species versus physical map positions of early RNAs (22, 5). The exponential slope for each RNA was calculated from the data of Fig. 2 by using the equation $K = (\ln R_0/R)/d$ where R_0 = rate of RNA synthesis at zero UV dose and R = rate of RNA synthesis at UV dose d . Because the uppermost curve in Fig. 2 corresponds to the sum of the RNA produced from genes 0.3 and 1.1, it is characteristic of the relative molar amounts of each. For a transcription unit of length $L = 7,600$ base pairs, a transcription rate of 50 base pairs traversed per second, and a 5-min labeling period, a simple calculation shows that (assuming stable RNA) there will be about twice as many RNA copies of gene 0.3 as of gene 1.1. The K value shown above at the position of gene 0.3 has been obtained from the uppermost curve of Fig. 2 by using this correction factor.

sponding RNA and protein. Thus, templates for either transcription or translation are utilized at random within the cell, and measurements of either process represent the differential loss of functional genes on the DNA template. There is no compensatory mechanism either on the transcriptional or the translational level to restore normal expression of those genes, the transcription of which had been blocked on a fraction of the co-infecting genomes.

The above considerations are applicable to all early T7 genes. They are also applicable to most T4 genes (6), except for those which are self-regulatory (16) or regulated by other T4 products. Transcription mapping of self-regulatory genes is facilitated by single infection. Of most importance for application to systems where proteins have been mapped but RNA transcripts have not, is the agreement found between protein and RNA assays.

Comparison with previous data on early region transcription. Our finding of a single

early transcription unit agrees with results of others. Milette et al. (14) found that a 2.2×10^6 molecular weight RNA was transcribed from T7 DNA in vitro by using highly purified *E. coli* RNA polymerase. Analysis by electron microscopy of the in vitro transcription complex by Davis and Hyman (3) revealed a specific initiation site for *E. coli* polymerase near the left end of the DNA, and a similar conclusion is inferred from the observations of Schweiger et al. on the sequential synthesis of early T7 proteins in vitro (18). Marrs and Yanofsky (12) argued from the time course of early transcription that a single unit of about 6,000 base pairs must be present.

Our result is also in complete agreement with the recent findings of Dunn and Studier (5), who have shown that RNase III cleaves the in vitro early region transcript of T7 DNA at five specific sites to yield RNA species which are essentially the same as those observed in vivo. In addition (4), they studied T7 transcription in vivo by an RNase III-deficient strain of *E. coli* (8). In this strain, 2.2×10^6 molecular weight RNA was produced and cleaved at the same five sites by RNase III in vitro.

Our results appear to conflict, however, with those of Minkley and Pribnow (15), who studied T7 transcription stimulated by dinucleotides or individual purine triphosphates in the presence of a 17:1 ratio of *E. coli* polymerase to DNA. In these reactions, the stimulating nucleotide was present at an excess of 10- to 200-fold over other ribonucleoside triphosphates, which were kept low (5 mM) to inhibit spontaneous initiation. They found that initiations occurred not only at the three sites contained in the early promoter, but also near the beginning of genes 0.7, 1.0, and 1.3, and even at one site outside the early region. Multiple binding sites for *E. coli* polymerase on T7 DNA have been reported (7, 10, 20), and their presence seems to depend upon the conditions used. The experimental conditions selected by Minkley and Pribnow (15) do not necessarily reflect intracellular conditions, and their conclusions are not necessarily applicable to in vivo transcriptions.

Recently, Young and Smith (23) have analyzed the end groups of early T7 mRNA transcribed in vivo and find, in addition to the known early promoter, only one other triphosphate-containing species; this is of a size which implies an initiation site before gene 1.0. They fail to show the identity of their RNA peaks on the gels and to find adenosine triphosphate-labeled termini. More data are needed to correctly interpret their results.

APPENDIX

Analysis of transcription of polycistronic RNA from UV-irradiated DNA templates. When DNA has been irradiated with UV light prior to transcription, premature termination of transcription at the site of UV lesions and release of the RNA polymerase occurs without reinitiation beyond the photoproduct (17). Such UV photoproducts can be formed at a large but finite number of discrete photosensitive sites (e.g., base pairs or intrastrand pyrimidine neighbors), and their distribution is assumed to be random. For irradiation of a large number of DNA molecules containing an identical transcription unit of length L , the probability of transcription of a gene whose promoter distal terminus is at length x' from the promoter is given by the term O^{th} , of the Poisson distribution. This may be written $e^{-\mu x'/L}$, where μ is the UV dose in RNA chain-terminating hits per transcription unit of length L (or equivalently, the mean of the Poisson distribution). This is the probability of the event that no RNA chain-terminating hit has occurred within the DNA segment beginning at the promoter and terminating at length x' . Thus, the number of transcripts of a specific gene decreases exponentially with UV dose. This implies that the exponential slopes, K , of UV-inactivation curves for a set of genes, transcribed by an RNA polymerase starting from the same promoter, will yield a straight line when plotted versus the physical map of the genome which will intercept at $K = 0$ at the origin of transcription (i.e., promoter or initiation site). The slope of these lines for different transcription units will be the same, provided that the composition of the DNA is invariant and that the units are all transcribed by the same polymerase or by polymerases which respond to UV lesions in an identical manner.

Furthermore, we have to consider the possibility that the presence of UV-irradiated phage DNA might lead to a dose-dependent lowering of the rate of overall RNA synthesis, apart from the premature termination of transcription considered above. We represent this effect by the term $e^{-C\mu}$, where C is a constant. The experimentally measured relative frequency of transcripts of a gene terminating at length x' from its promoter would be lower by this factor at any UV dose μ and would become

$$e^{-\mu x'/L} e^{-C\mu} = e^{-\mu[(x'/L) + C]}$$

The result of this is that the plot of exponential slope versus physical map will still be linear and will be of the same slope; however, the intercept of this line at the origin of transcription will be at C and not at $K = 0$.

Frequency distribution of truncated RNA molecules as a function of length. We define the distribution function $f_x(x)$ as the relative frequency of RNA transcripts which are of length x . The template DNA from which such a molecule was transcribed, therefore, had a UV photoproduct at site $x + 1$, but none in the region from the promoter up to this site. This distribution function is readily derived from the frequency of transcripts which are of length x' (i.e.,

the probability of transcription through x'). Because all RNA molecules of length x' or longer contain the complete RNA message up to x' , the fraction of such RNA molecules plus the fraction of complete RNA transcripts (from unhit genomes) gives the total frequency of transcripts that contain the message up to x' . When x' is at the promotor distal end of a gene, the frequency of transcripts through x' equals the probability of transcription of this particular gene. The frequency of transcripts of length x' and larger, but not including those of full length L , is given by

$$\sum_{x=x'}^L f_{\mu}(x) = e^{-\mu x'/L} - e^{-\mu}$$

Because the number of photosensitive sites has been assumed to be reasonably large, i.e., at least several hundred, this sum may be replaced by an integral yielding

$$\int_{x'}^L f_{\mu}(x) dx = e^{-\mu x'/L} - e^{-\mu}$$

Because the integral is unique within a constant

$$f_{\mu}(x) = -\frac{d}{dx}(e^{-\mu x/L} - e^{-\mu}) = (\mu/L)e^{-\mu x/L} \quad (0 < x \leq L)$$

This is the UV-induced frequency distribution of truncated RNA molecules as a function of length.

UV-dependent distribution of labeled RNA molecules in polyacrylamide gel electrophoresis. The length distribution function $f_{\mu}(x)$ gives the relative frequency of RNA transcripts of length x resulting from unidirectional transcription of UV-irradiated DNA. However, analysis of polyacrylamide gels yields the number of disintegrations per minute occurring in radioactively labeled RNA bands. Assuming uniform composition of RNA, the number of disintegrations per minute should be proportional to RNA length, and the electrophoretic mobility should be given by the logarithm of the molecular weight.

Consequently, the gel distribution of truncated RNA molecules arising from transcription from a promotor at one end of UV-irradiated (dose μ) transcription units of length L is represented by a plot of $x \cdot f_{\mu}(x)$ versus $\ln(x)$.

This is shown in Fig. 4 for $\mu = 0.5, 1, 2$, and 4 hits; (these values for μ were chosen to correspond to the UV doses used experimentally). The peak of the distribution is given by $(d/dx) x \cdot f_{\mu}(x) = 0$ and occurs at position $x = L/\mu$.

The relative amount of RNA represented by this curve is given by

$$1/L \int_0^L x \cdot f_{\mu}(x) dx$$

and equals $1/\mu (1 - e^{-\mu} - \mu e^{-\mu})$. For μ larger than 7 the relative amount of RNA contained in this peak resolves to $1/\mu$ to 1% accuracy. This permits determination of L from the decline in RNA production with dose.

The consequence of the above is that UV irradiation of DNA at doses of $\mu = 1$ or larger causes, upon transcription, a broad peak of prematurely ter-

minated RNA molecules to appear on the gel at position $\ln(L/\mu)$. For μ larger than 1.25, the fraction of RNA represented by this peak becomes larger than that contained in the peak of full-length RNA molecules at position $\ln(L)$ on the gel. (The fraction of RNA molecules corresponding to unhit genomes is not plotted in Fig. 4.)

Frequency distribution of RNA fragments after cleavage by RNase III. Let us consider now the particular case of early region transcription of UV-irradiated T7 DNA. Transcription from a single promotor region at the left end of the genome is assumed to be followed by cleavage by RNase III of the 2.2×10^6 molecular weight transcripts. We further assume that such cleavage is complete and can occur on partial-length transcripts, and that all RNA species are stable. Because this cleavage occurs only at specific intergenic sites, it would yield equimolar amounts of each T7 early RNA species in the absence of irradiation. However, at UV dose μ not all of the early region transcripts will be full-length, and the prematurely terminated ones will be distributed according to length (Fig. 4) with L equal to the length of the T7 early region. Intergenic cleavage of each RNA molecule of this distribution will lead to one RNA corresponding to each of those genes whose transcrip-

FIG. 4. Distribution of RNA as a function of transcript length (x) at four representative UV doses. The full length of the polycistronic transcription unit is represented by L (abscissa). This distribution function has been plotted in a manner characteristic of the way in which RNA is known to electrophorese on polyacrylamide gels. UV dose is represented by μ , the mean of the Poisson distribution, and curves are shown for doses of 0.5, 1, 2, and 4 chain-terminating hits per length L . These doses correspond to those used experimentally, with L = length of the T7 early region. Fraction of transcripts from unhit genomes is represented by the term $e^{-\mu}$ (not shown). Ordinate is representative of the absolute amount of RNA and is shown for the T7 early region, assuming $L = 7,000$ nucleotides. Thus, conversion of the ordinate to a scale representative of fractional amount of initial RNA at zero dose requires division by 7,000. Because the fraction of full-length transcripts decreases rapidly with dose ($e^{-\mu}$), the curves represent a rapidly increasing relative amount of RNA with dose.

tion had been completed within the uncleaved molecule. In addition, there will result one abnormal length segment of RNA corresponding to only the promotor proximal part of the next gene in the series, and this short segment is defined as an RNA "frag-

FIG. 5. Distribution of RNA fragments resulting from intergenic cleavage of the distribution of T7 early region transcripts (Fig. 4) by RNase III. The abscissa represent l , the RNA fragment length in bases; the ordinate is characteristic of the absolute amount of RNA and is the same as in Fig. 4. The four sets of curves correspond (from top to bottom) to $\mu = 0.5, 1, 2$, and 4 hits on the transcription unit of length L , where μ is the mean of the Poisson distribution. An RNA "fragment" is defined as that most promotor distal segment of RNA resulting from intergenic cleavage of the UV-truncated early region transcripts. Sites for cleavage were assumed to be between each of the known T7 early genes. Lengths corresponding to each of the early RNA species are shown at the bottom of the figure. Heavy black bars represent approximate bandwidths for gel electrophoresis of early RNA and were derived from Fig. 1. Those fragments whose length is such that they are within the band width would be expected to contribute extraneous RNA to experimental early RNA gel patterns, assuming them to be as stable as the early RNA species themselves. Curves have again been plotted in a manner characteristic of experimental gel patterns and are shown after subtraction of a background chosen to correspond approximately with the background corrections used for experimental data. The ordinate has a different meaning at each dose, because it represents with increasing dose a larger relative fraction of the transcripts of any given gene. This effect is most pronounced for RNA transcribed from the most promotor distal gene, gene 1.3, transcription of which has dropped to e^{-4} , or 1.83% of its initial value by $\mu = 4$ (see Fig. 6).

ment" (each molecule represented by the distribution $f_\mu(x)$ gives rise to only one fragment after cleavage).

The function $F_\mu(l)$ may now be defined as the frequency distribution of RNA fragments which results when cleavage by RNase III is applied to the frequency distribution of RNA transcripts, $f_\mu(x)$. In the special case in which L is the length of the T7 early region and RNA cleavage occurs between each of the five known early genes, $F_\mu(l)$ should be characteristic of the fragment distribution for transcription of UV-irradiated T7 by *E. coli* RNA polymerase. The distribution $F_\mu(l)$ for the early region of T7 in the absence of diffusion is shown after background correction in Fig. 5. It has been plotted versus $\ln(l)$ rather than l , in order to be characteristic of the expected polyacrylamide gel electrophoresis pattern. It is seen that spikes appear at the location of each gene. Experimentally determined bandwidths for the RNA of each gene, derived from Fig. 1, are also shown. The amount of RNA contained within the bandwidth in each spike can be determined and calculated as a ratio of the amount of normal length RNA expected

FIG. 6. Effect of the extraneous RNA fragments contained within the bandwidths of each RNA species (Fig. 5) on the (theoretical) UV-inactivation curves. The abscissa represents UV dose in chain-terminating hits, and the ordinate represents the probability of transcription of each individual gene. Solid lines are the simple exponential UV-inactivation curves expected from theory in the absence of extraneous RNA; dashed curves show the alteration induced by the consideration of the appearance of partial gene length cleavage products (fragments) at gel positions occupied by early RNA species.

at each position by theory at that UV dose. This will be extraneous RNA which will raise each point of the (formerly) exponential UV-inactivation curves slightly. This effect will be greater for points corresponding to higher doses and for transcription of those genes most promotor distal. The predicted effect is shown in Fig. 6, where the solid line for each RNA species is the simple exponential curve expected from the theory and the broken line is the alteration induced by the presence of RNA fragments. The inclusion of the effect of diffusion of the fragments in the gel would be complex and would not alter this figure significantly.

ACKNOWLEDGMENTS

This investigation was supported by National Science Foundation grant no. GB-38085 and by Public Health Service training grant no. GM-00781 from the National Institute of General Medical Sciences to the Department of Biophysics and Genetics (no. 580), University of Colorado Medical Center, Denver.

LITERATURE CITED

- 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.
- Bräutigam, A. R., and W. Sauerbier. 1973. Transcription unit mapping in bacteriophage T7. I. In vivo transcription by *Escherichia coli* RNA polymerase. *J. Virol.* **12**:882-886.
- 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.
- Dunn, J. J., and F. W. Studier. 1973. T7 early RNAs and *E. coli* ribosomal RNAs are cut from large precursor RNAs *in vivo* by RNase III. *Proc. Nat. Acad. Sci. U.S.A.* **70**:3296-3300.
- 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.
- Hercules, K., and W. Sauerbier. 1973. Transcription units in bacteriophage T4. *J. Virol.* **12**:872-881.
- Hinkle, D. C., and M. J. Chamberlin. 1972. Studies of binding of *Escherichia coli* RNA polymerase to DNA. I. The role of sigma subunit in site selection. *J. Mol. Biol.* **70**:157-185.
- Keil, T. U., and P. H. Hofschneider. 1973. Secondary structure of RNA phage M12 replicative intermediates *in vivo*. *Biochim. Biophys. Acta* **312**:297-310.
- Kinard, F. E. 1957. Liquid scintillator for the analysis of tritium in water. *Rev. Sci. Instrum.* **28**:293-294.
- Le Talaer, J. Y., M. Kermici, and Ph. Jeanteur. 1973. Isolation of *Escherichia coli* RNA polymerase binding sites on T5 and T7 DNA: further evidence for sigma-dependent recognition of A-T rich DNA sequences. *Proc. Nat. Acad. U.S.A.* **70**:2911-2915.
- Loening, U. E. 1967. The fractionation of high molecular weight ribonucleic acid by polyacrylamide gel electrophoresis. *Biochem. J.* **102**:251-257.
- Marrs, B. L., and C. Yanofsky. 1971. Host and bacteriophage specific messenger RNA degradation in T7-infected *E. coli*. *Nature N. Biol.* **234**:168-170.
- Michalke, H., and H. Bremer. 1969. RNA synthesis in *Escherichia coli* after irradiation with ultraviolet light. *J. Mol. Biol.* **41**:1-23.
- 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.
- Minkley, E. G., and D. Pribnow. 1973. Transcription of the early region of bacteriophage T7: selective initiation with dinucleotides. *J. Mol. Biol.* **77**:255-277.
- Russel, M. 1973. Control of bacteriophage T4 DNA polymerase synthesis. *J. Mol. Biol.* **79**:83-94.
- 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.
- Schweiger, M., P. Herrlich, and R. L. Millette. 1971. Gene expression *in vitro* from DNA of bacteriophage T7. *J. Biol. Chem.* **246**:6707-6712.
- Simon, N. M., and F. W. Studier. 1973. Physical mapping of the early region of bacteriophage T7 DNA. *J. Mol. Biol.* **79**:249-265.
- Stevens, A. 1969. Studies of DNA-dependent RNA polymerase. *J. Cell. Physiol. Suppl.* **1** **74**:205-218.
- Studier, F. W. 1969. The genetics and physiology of bacteriophage T7. *Virology* **39**:652-674.
- 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.
- Young, R. J., and G. Smith. 1973. The end groups of T7 mRNA. *Biochem. Biophys. Res. Commun.* **53**:952-959.