Proc. Nat. Acad. Sci. USA
Vol. 70, No. 10, pp. 2911–2915, October 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

J. Y. LE TALAER, M. KERMICI, AND PH. JEANTEUR

Unité de Biochimie-Enzymologie, Institut Gustave Roussy, 94800 Villejuif, France; and Laboratoire de Biochimie, Centre Régional de Lutte contre le Cancer, Cliniques Saint Eloi, 34000 Montpellier, France

Communicated by André Lwoff, June 15, 1973

ABSTRACT **Previous isolation and analysis of *E. coli* RNA polymerase (EC 2.7.7.6) binding sites on λ DNA had demonstrated the existence of a sigma-dependent process of recognition of A-T-rich DNA sequences. We have now extended this finding to T5 and T7 DNA and have provided evidence for the double-strandedness of the isolated binding sites. The possible equation of these sites to the genetically defined promoters is discussed.**

We have reported (1) the isolation of *Escherichia coli* RNA polymerase binding sites on λ DNA by taking advantage of their resistance to nuclease digestion. Their nucleotide composition was enriched in A-T (57%) when a low polymerase-to-DNA ratio was used, although at high ratio it was not very different from that of total λ DNA (50%). This observation reflected the existence of at least two types of sites (2), the ones with the highest affinity being selected for at low polymerase-to-DNA ratios. Indeed, when using tRNA as a competitor for RNA polymerase during the digestion step, we observed (3) that, when the binding was done with holoenzyme at 37°, a fraction of protected DNA fragments was resistant to displacement by tRNA, as assayed by the filter retention technique (4). This resistant fraction was enriched in A-T. In contrast, no resistance was observed with core enzyme or with holoenzyme incubated at 0°. By further purification on acrylamide gels, we were able to resolve two populations of protected DNA fragments (5). The first type of fragments, referred to as peak I, contained 45–52 nucleotide residues, was considerably enriched in A-T (up to 67%), and, most importantly, was observed only when sigma was present during the binding step. The second type of fragments (peak II), containing 7–10 nucleotide residues, showed no dependence on sigma nor was it enriched in A-T. On the basis of its sigma dependence the first population of protected fragments was operationally defined as promoters. The sigma-dependent selection of A-T-rich DNA sequences by RNA polymerase along with an additional temperature requirement in order to form a stable complex, as demonstrated with T7 (6) and λ DNA (3), fits well with the idea of a local melting of DNA. Direct evidence on that point has been obtained recently by Saucier and Wang (7), who have detected a limited unwinding (about four base pairs) upon binding of RNA polymerase to circular λ DNA template.

It would be of great interest to know whether the sigma-directed recognition of A-T-rich DNA sequences is a general phenomenon. Recent papers (8, 9) report results very similar to ours with replicative form of phage fd DNA. The present paper extends to T5 and T7 our previous results on λ (5) and reports further studies on two questions that had been raised by this work and are now answered unequivocally.

(*1*) The enrichment in A-T observed in the "promoters" (peak-I fragments) decreased with increasing polymerase-to-DNA ratios. This observation had been interpreted in the following way: at high ratios, peak I is contaminated with nonspecific binding sites of average base composition. Although of lower affinity than specific ones, these sites still bind tightly enough not to be displaced by DNase at high concentration of polymerase. Therefore, nonspecific sites of average composition should be detectable as peak-I molecules even in the absence of sigma.

(*2*) Although the nucleotide composition of peak-I fragments was consistent with their being double stranded, this point had to be unequivocally established. Because of the high number of T5 promoters (10), making them more easily available in quantities possible to work with, the experiments designed to answer this latter question were done on T5 DNA.

MATERIALS AND METHODS

Enzymes. RNA polymerase holoenzyme (EC 2.7.7.6.) was purified from *E. coli* A 19 according to the phase-partition method of Babinet (11) followed by the glycerol gradient centrifugation steps of Burgess (12) to remove μ and τ bands. Although we could detect no difference when using preparations still containing these impurities, pure ones (over 97% β, β', α_2, and σ subunits) were used throughout this work. Core enzyme was also obtained as previously (5). Both holoenzyme and core enzyme were kept at −20° in Burgess' storage buffer (12). Pancreatic DNase (EC 3.1.4.5.) was obtained from Worthington. Venom phosphodiesterase (EC 3.1.4.1.) from Worthington or Calbiochem was exposed to pH 3.6 for 3 hr (13) in order to remove any residual 5′-nucleotidase activity. Single-strand specific endonuclease from *Neurospora crassa* was a kind gift of Dr. P. Sheldrick.

Phages and DNAs. Labeled T5 phages were grown by infection of *E. coli* in MGM medium (14) containing 0.4 phosphate and 10 μCi/ml of $H_3{}^{32}PO_4$. T7 phages were grown and labeled in the same medium supplemented with 0.05% casamino acids and containing 10 μCi/ml of $H_3{}^{32}PO_4$. Both phages were concentrated by phase partition and purified by banding in cesium chloride (15). DNA was phenol-extracted from phages as described (5), with especially gentle stirring and wide-bore pipettes for T5 DNA. Storage of DNA samples was

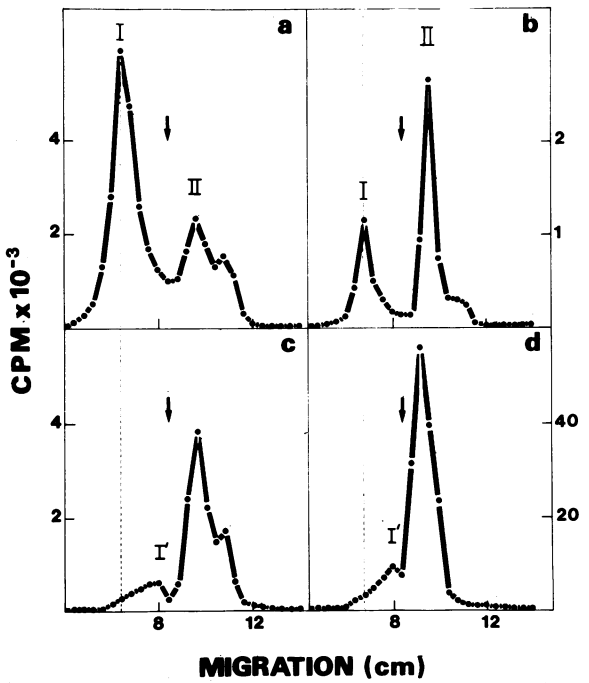

FIG. 1. Radioactivity profiles on acrylamide gels of protected DNA fragments from phage T5 DNA. Fragments obtained in the presence (*a*, *b*) or absence (*c*, *d*) of sigma factor were prepared and analyzed. The binding mixtures contained per ml: (*a*) 40 μg of DNA and 20 μg of holoenzyme; (*b*) 40 μg of DNA and 80 μg of holoenzyme; (*c*) 50 μg of DNA and 25 μg of core enzyme; (*d*) 50 μg of DNA and 250 μg of core enzyme. *Arrows* indicate the position of the marker dye (bromophenol blue).

in 40 mM Tris·HCl (pH 8.0)–1 mM EDTA. Carrier calf-thymus DNA was obtained from Worthington. Tritiated *E. coli* DNA was kindly supplied by Dr. Toutain.

Isolation and Analysis on Acrylamide Gels of Protected DNA Fragments. Polymerase-binding, nuclease-digestion, phenol-extraction, and ethanol-precipitation steps were done as described (1, 5). Acrylamide-gel slices were eluted by simple diffusion instead of electrophoresis. The gel slices were cut into small pieces and left overnight in 0.5–1 ml per slice of 10 mM Tris·HCl (pH 7.5)–1 mM EDTA at room temperature (23°).

Nucleotide Composition Analysis. The gel eluates were brought to 0.1 NaCl and precipitated by two volumes of ethanol in the presence of 10 μg of calf-thymus DNA as a carrier. After 4 hr at −20°, the protected fragments were spun down for 30 min at 11,000 rpm (14,000 × *g*) and dissolved in 0.1 ml of water. After addition of 20 μl of 0.05 M $MgCl_2$ and 50 μl of a 1.27-mg/ml solution of calf-thymus DNA in water, the pH was adjusted to 7.5 by 4 μl of 0.1 N NaOH. Hydrolysis was started by addition of 20 μl of RNase-free DNase (2.5 mg/ml in 1 mM $CaCl_2$). The incubation was done at 37° for 3 hr after which the pH was raised to 8.5 by addition of 0.1 N NaOH. 20 μl of phosphodiesterase (75 units/ml in water) was then added and the incubation was continued for three additional hours. The hydrolysate was evaporated to dryness under a stream of nitrogen and dissolved in 20 μl of water and 10 μl of a 3-mg/ml solution of both dGMP and dTMP. Nucleotides were separated by thin-layer electrophoresis on cellulose sheets (Eastman Kodak 6064) under 1500 V in 0.2 M sodium acetate buffer (pH 3.8). The spots were then located under UV light, cut out, and counted in a toluene-based scintillator.

Treatment of Protected Fragments with N. crassa Endonuclease. Gel eluates were precipitated with ethanol as above except for the absence of added DNA carrier, and dissolved in 10 mM Tris·HCl (pH 7.5)–1 mM EDTA. Denaturation was performed by heating for 10 min at 100° in the above buffer followed by quenching in ice. The incubation was done according to Linn and Lehman (16) in the presence of 4.5 μg/ml of either native or heat-denatured tritiated *E. coli* DNA. 40-μl aliquots were withdrawn after different times and precipitated by 100 μl of 0.35 N $HClO_4$ and 50 μl of 2.5 mg/ml of bovine-serum albumin for 5 min in ice. After a 10-min centrifugation at 17,000 × *g*, 150-μl aliquots of the supernatant were counted in 10 ml of Bray's solution.

RESULTS

Gel analysis and nucleotide composition of protected fragments from T5 DNA

In Fig. 1*a* and *b* are presented the migration profiles of two preparations of protected fragments obtained at different polymerase/DNA ratios. Holoenzyme (containing sigma) was used in both binding mixtures. These two profiles with distinct peaks I and II, are exactly reminiscent of the ones previously obtained with λ DNA (5). Moreover, these peaks have the

TABLE 1. *Nucleotide composition of protected fragments from T5 DNA*

Exp.	Binding conditions	Peak I or I'					Peak II				
		C	A	G	T	A-T%	C	A	G	T	A-T%
1	Total DNA (no polymerase)	(20.2)	(29.8)	(20.1)	(29.9)	(59.7)	(20.2)	(29.8)	(20.1)	(29.9)	(59.7)
2	+ σ, R = 0.5	15.5	32.6	18.1	33.8	66.4	20.4	29.6	20.0	30.0	59.6
3	+ σ, R = 1	18.1	33.0	16.6	32.3	65.3	20.3	31.0	19.8	28.9	59.9
4	+ σ, R = 2	20.9	30.0	19.0	30.1	60.1	22.1	29.0	18.8	30.1	59.1
5	− σ, R = 0.5	19.7	27.8	20.5	32.0	59.8	20.0	29 8	20.0	30.2	60.0
6	− σ, R = 5	21.6	29.2	20.0	29.2	58.4	22.7	30.2	20.0	27.1	57.3

Exps. *2*, *4*, *5*, and *6* correspond, respectively, to patterns *a*, *b*, *c*, and *d* of Fig. 1. The binding mixture of Exp. *3* contained 50 μg of DNA and 50 μg of holoenzyme in 1.5 ml. *R* refers to the weight ratio of polymerase to DNA. For easier comparison the nucleotide composition of unprotected total T5 is shown in parentheses on top of both peak I and peak II columns. All figures presented are averages of several determinations.

same relative mobilities and, consequently. the same size as their λ counterparts (see *Discussion*). The occasional presence of a shoulder on the light side of peak II is considered insignificant and results from some carry-over during ethanol precipitation. In any case, its migration corresponds to that of mononucleotides, obtained by omitting polymerase from the binding mixture.

Nucleotide compositions of both peaks from Fig. 1*a* and *b* are presented in Table 1 along with data pertaining to a similar gel pattern (not shown here) obtained at a still higher polymerase/DNA weight ratio (R = 2). The maximum A-T content (66.4%) of fragments obtained at the lowest ratio used was identical for T5 and λ. The decrease of this A-T content towards that of total DNA with increasing concentrations of polymerase is also a common feature of both systems. These observations are interpreted in the following way: at least two types of binding sites exist for which the polymerase has different affinities (2, 3, 17). Only those sites of highest affinity, selected for at low concentrations of polymerase, are A-T rich and assumed to be the specific ones. On the other hand, the sites with lower affinity are nonspecific and have a normal base composition. They become increasingly occupied as the ratio of polymerase to DNA is raised. Their lower affinity for polymerase is still quite high (the complex is able to withstand a sedimentation run of several hours in sucrose) compared to their affinity for DNase (K_m = 1 mg of DNA/ml; De Traversay, M. & Le Talaer, J. Y., unpublished results). Under conditions of excess polymerase, these nonspecific sites would not be displaced and split by DNase and would therefore behave in our system as peak-I molecules. On the assumption that these sites are nonspecific, the prediction can be made that core enzyme (lacking the sigma specificity factor) should also bind to them, although they have a nucleotide composition like that of the total DNA. The distributions on gels of protected fragments obtained in the absence of sigma at two polymerase/DNA ratios are shown in Fig. 1*c* and *d*. In addition to peak II, a heavier peak overlapping the peak-I region (referred to as peak I′) is also observed. Peak I′ is much wider than peak I and centered at a significantly lighter position. It therefore contains an heterogeneous population of fragments smaller than those in peak I. However, in contrast with those protected fragments obtained in the presence of sigma, the base composition of fragments in peak I′ is in no case different from that of total T5 DNA, regardless of the polymerase-to-DNA ratio. All peaks II in Fig. 1*a*, *b*, *c*, and *d* also have normal composition. The selection of

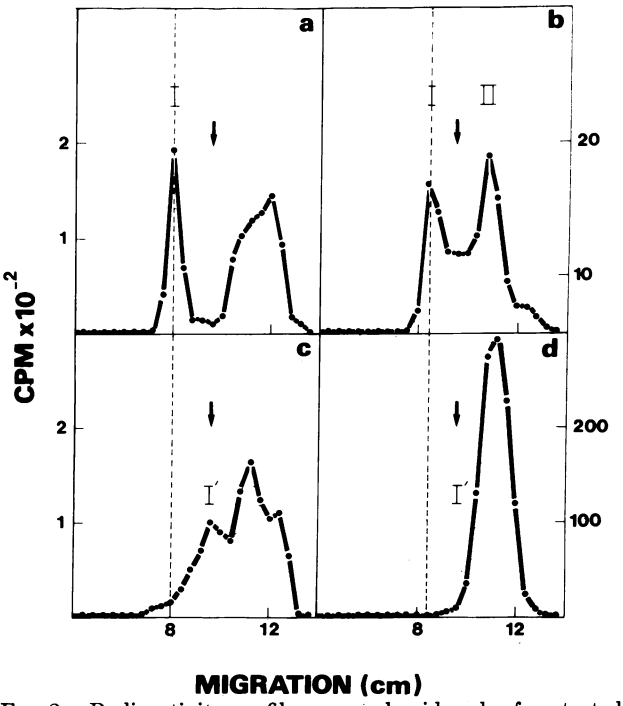

FIG. 2. Radioactivity profiles on acrylamide gels of protected DNA fragments from phage T7 DNA. Fragments obtained in the presence (*a*, *b*) or absence (*c*, *d*) of sigma factor were prepared and analyzed. The binding mixtures contained per ml: (*a*) 89 μg of DNA and 13 μg of holoenzyme; (*b*) 40 μg of DNA and 200 μg of holoenzyme; (*c*) 43 μg of DNA and 10 μg of core enzyme; (*d*) 43 μg of DNA and 250 μg of core enzyme. *Arrows* indicate the position of the marker dye.

A-T-rich sequences of T5 DNA is therefore sigma-dependent. Comparison of patterns *a* and *c* (Fig. 1) obtained in the presence or absence of sigma at the same polymerase/DNA ratio shows that, even at this low ratio, some contamination of peak I by nonspecific fragments does exist. However, this contamination seems very slight in comparison to the size of peak II.

Gel analysis and nucleotide composition of protected fragments from T7 DNA

The gel patterns obtained with T7 protected fragments (Fig. 2) are exactly analogous to those obtained with T5 (Fig. 1) and λ (5). The same sigma dependence of peak I is obvious when patterns in the presence of sigma (Fig. 2*a* and *b*) are

TABLE 2. *Nucleotide compositions of protected fragments from T7 DNA*

Exp.	Binding conditions	Peak I or I′					Peak II				
		C	A	G	T	A-T%	C	A	G	T	A-T%
1	Total DNA (no polymerase)	(24.5)	(25.5)	(24.6)	(25.4)	(50.9)	(24.5)	(25.5)	(24.6)	(25.4)	(50.9)
2	+ σ, R = 0.15	16.5	33.5	16.0	34.0	67.5	25.3	25.4	23.7	25.6	51.0
3	+ σ, R = 0.5	22.4	27.5	22.4	27.7	55.2	24.9	26.6	24.9	23.6	50.2
4	+ σ, R = 5	22.1	27.8	22.5	27.6	55.4	27.8	24.6	25.0	22.6	47.2
5	− σ, R = 0.15	24.3	25.9	24.3	25.5	51.4	23.9	26.9	25.0	24.2	51.1

Exps. *2* and *4* correspond, respectively, to patterns *a* and *b* of Fig. 2. Exp. *3* contained 80 μg of T7 DNA and 40 μg of holoenzyme in 1.6 ml and Exp. *5* contained 400 μg of T7 DNA and 60 μg of core enzyme in 0.8 ml. All figures presented are averages of several determinations.

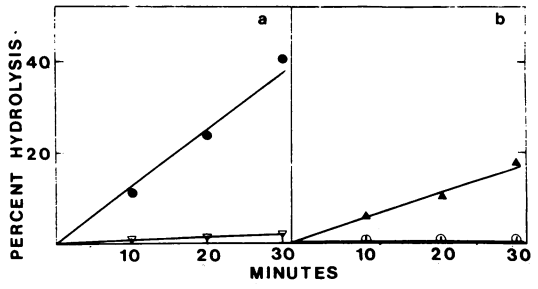

FIG. 3. Assay for sensitivity to *N. crassa* endonuclease of protected fragments from T5 DNA. Kinetics of release as a function of time of acid-soluble ^{32}P counts from protected T5 DNA fragments and ^{3}H counts from *E. coli* DNA. Conditions of incubation were described in *Methods*. The reaction mixture of (*a*) contained native [^{32}P]DNA (9500 cpm) Δ and denatured [^{3}H]DNA (32000 cpm) ●. The reaction mixture of (*b*) contained denatured [^{32}P]DNA (9500 cpm) ▲ and native [^{3}H]DNA (32000 cpm) ○.

compared to those obtained in its absence (Fig. 2*c* and *d*). The nucleotide compositions of the different peaks are shown in Table 2. The A-T enrichment (67.5%) observed here is also very close to the ones obtained with λ (66.8%) and T5 (66.4%).

Compared to the present results, the absence of peak I′ in λ might appear discrepant. However, a closer look at Fig. 1*d* of ref. 5 revealed the presence of a significant amount of radioactivity in the region of peak I–I′ which we had overlooked because of the absence of a defined peak. Had it been performed, we are quite confident that the base-composition analysis of these fragments would have also proved them to be like total DNA. These results provide complete fulfillment of the above prediction. They again point out strongly the stringent need for using low polymerase/DNA ratios.

Secondary structure of protected fragments from T5 DNA

The nucleotide distribution of peak-I molecules observed with T5 (Table 1), T7 (Table 2), and λ (5) all show a fairly good A = T and G = C correlation. The same observation was made on protected fragments from the replicative form of phage fd DNA (8). However, this is no proof of double strandedness. The importance of this issue is such as to require an unambiguous answer. We therefore investigated the sensitivity of our fragments to the *N. crassa* endonuclease specific for single-stranded polynucleotides. Fig. 3 shows the rate of hydrolysis of peak-I fragments from T5 (polymerase/DNA weight ratio, R = 0.5) with or without previous heat denaturation. As an internal control of enzyme activity and specificity, denatured or native *E. coli* [^{3}H]DNA was mixed with native (Fig. 3*a*) and denatured (Fig. 3*b*) fragments of [^{32}P]DNA from T5, respectively. From comparison of these two figures, it emerges that the rate of release of acid-soluble ^{32}P counts from protected fragments is drastically increased when they have been previously heat denatured. However, there appears to be a very limited hydrolysis of native fragments (Fig. 3*a*), which might reflect the presence of a short single-stranded tail at one or both ends. Such a possibility arises from the proposal by Heyden *et al.* (9) that the two strands of protected DNA are of unequal length.

These experiments demonstrate that T5 peak-I fragments obtained at the lowest polymerase-to-DNA ratio (therefore containing mostly, if not exclusively, A-T-rich, sigma-dependent binding sites) are essentially double stranded. Should a hairpin structure have existed in the native sample, it would have been expected to reform on cooling, yielding again a resistant fragment.

We had observed (5) that peak-I fragments from λ had the same mobility in gels with or without previous heat denaturation. There are three different ways to interpret this behavior: (*1*) Fragments were already single stranded in the native state; (*2*) they are easily renaturable (hairpin structure); (*3*) single-stranded and double-stranded DNA molecules of the same length cannot be distinguished under these conditions. The present data have ruled out possibilities (*1*) and (*2*). From the third one, it stems that our former calibration curve made with single-stranded RNA markers should be valid for only single-stranded DNA. Our estimate of 45–52 nucleotides for peak-I molecules after denaturation therefore means that they contain 45–52 base pairs, in rather good agreement with the findings of Heyden and colleagues (9).

DISCUSSION

The present work provides clear evidence for the sigma-dependent recognition of A-T-rich sequences of T5 and T7 DNA by *E. coli* RNA polymerase. As demonstrated, in no case are any A-T-rich sequences preferentially selected unless sigma is present. This confirms and extends our earlier work on λ(5) and shows a striking constancy of A-T content (about 67%) of sites recognized by *E. coli* RNA polymerase on the DNAs from different phages. Other authors have also reported recently some A-T enrichment of binding sites from phage fd replicative form DNA both in the presence (8) and absence (9) of nucleoside triphosphates. It is tempting to correlate this recognition process of A-T-rich (67%) sequences with the structural transition observed by Bram (18) in DNAs containing more than 65% A-T.

Their resistance to the *N. crassa* endonuclease specific for single-stranded polynucleotides is conclusive evidence for the absence of any significant single-stranded regions or hairpin structure in sigma-dependent binding sites from T5 DNA. Using differents methods, Heyden *et al.* (9) have reached the same conclusion about phase fd replicative form DNA. With the possible exception of short single-stranded tails, we therefore conclude that double strandedness should be a general feature of all binding sites for *E. coli* RNA polymerase holoenzyme, when the binding sites have been isolated by the DNase excision method. It would be interesting to know whether nonspecific DNA fragments protected in the absence of sigma are also double stranded. However, the size of these fragments (7–10 nucleotide residues) is already too close to the limit of acid solubility for such an assay to be meaningful.

Another finding of the present work is the existence of nonspecific binding sites of high affinity on T5 and T7 as well as λ DNA (5). Working with T7, Hinkle and Chamberlin (17) have actually found 8 sites of very high affinity ($K_a = 10^{12}$–10^{14} M^{-1}) for holoenzyme and a single population of sites for core enzyme of lower but still high affinity ($K_a = 2 \times 10^{11}$ M^{-1}). In order to select for the former, one should therefore use as low polymerase-to-DNA ratios as possible. The same conclusion has been reached by Dausse and colleagues (19), who have found that totally asymmetric transcription could only be obtained when less than four polymerase molecules were bound per T7 genome. This value corresponds to

R = 0.08 (weight per weight) and correlates very well with our own results.

Two recent papers by Blattner and colleagues (20, 21) on startpoints for RNA synthesis on λ DNA have made important points directly relevant to our previous (5) and present work. They found not only that the promoters and operators are not transcribed, but also that the actual initiation sites for RNA synthesis are located some 200 nucleotide pairs downstream with respect to the direction of transcription. These results raise the question of whether the sites we have isolated are the entry site (or any part of an "antenna" region), or the initiation site, or any sequence between these two (drift region). More work is obviously needed to critically select among these possibilities.

This work was supported by the Délégation Générale à la Recherche Scientifique et Technique (Contracts no. 72-7-0491 and 72-7-0389), the Institut National de la Santé et de la Recherche Médicale and the Ligue Nationale Française contre le Cancer. We are indebted for the interest and support of Dr. C. Paoletti in whose laboratory this work was done. We thank Miss S. Colin for her expert technical assistance.

1. Le Talaer, J. Y. & Jeanteur, Ph. (1971) *FEBS Lett.* **12**, 253–756.
2. Stead, N. W. & Jones, O. W. (1967) *J. Mol. Biol.* **26**, 131–135.
3. Le Talaer, J. Y. & Jeanteur, Ph. (1972) *FEBS Lett.* **28**, 305–308.
4. Jones, O. W. & Berg, P. (1966) *J. Mol. Biol.* **22**, 199–209.
5. Le Talaer, J. Y. & Jeanteur, Ph. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 3211–3215.
6. Hinkle, D. C. & Chamberlin, M. (1970) *Cold Spring Harbor Symp. Quant. Biol.* **35**, 65–72.
7. Saucier, J. M. & Wang, J. C. (1972) *Nature New Biol.* **239**, 167–170.
8. Okamoto, T., Sugiura, M. & Takanami, M. (1972) *Nature New Biol.* **237**, 108–109.
9. Heyden, B., Nusslein, C. & Schaller, H. (1972) *Nature New Biol.* **240**, 9–12.
10. Bautz, E. K. F. & Bautz, F. A. (1970) *Nature* **228**, 1219–1222.
11. Babinet, C. (1967) *Biochem. Biophys. Res. Commun.* **26**, 639–644.
12. Burgess, R. R. (1969) *J. Biol. Chem.* **244**, 6160–6169.
13. Sulkowski, E. & Laskowski, M. (1971) *Biochim. Biophys. Acta* **240**, 443–447
14. Lanni, Y. T. (1961) *Virology* **15**, 127–135.
15. Le Pecq, J. B. & Baldwin, R. L. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **33**, 609–620.
16. Linn, S. & Lehman, I. R. (1965) *J. Biol. Chem.* **240**, 1287–1293.
17. Hinkle, D. & Chamberlin, M. (1972) *J. Mol. Biol.* **70**, 157–185.
18. Bram, S. (1972) *Biochimie* **54**, 1005–1011.
19. Dausse, J. P., Sentenac, A. & Fromageot, P. (1972) *Eur. J. Biochem.* **26**, 43–49.
20. Blattner, F. R. & Dahlberg, J. E. (1972) *Nature New Biol.* **237**, 227–232.
21. Blattner, F. R., Dahlberg, J. E., Boettiger, J. K., Fiandt, M. & Szybalski, W. (1972) *Nature New Biol.* **237**, 232–236.