

Location, function, and nucleotide sequence of a promoter for bacteriophage T3 RNA polymerase

(restriction map of T3 DNA/Southern hybridization/restriction fragment transcription by T3 RNA polymerase/DNA sequencing/
RNA sequencing)

SAMIT ADHYA, SHANTANU BASU, PROBIR SARKAR, AND UMADAS MAITRA

Department of Developmental Biology and Cancer, Albert Einstein College of Medicine, Bronx, New York 10461

Communicated by Jerard Hurwitz, September 18, 1980

ABSTRACT The major promoters for bacteriophage T3 RNA polymerase on the T3 genome have been mapped by DNA-RNA filter hybridization. One promoter is located in a 300-base-pair *Hpa* I restriction fragment near the genetic "left" end of T3 DNA. The sequence in the vicinity of the major initiation site of transcription in this region has been determined. A part of the (-)strand sequence is 5' T-A-T-T-T-A-C-C-C-T-C-A-C-T-A-A-A-G-⁺¹G-G-A-A-U 3'. Comparison of this sequence with the prototype 23-base-pair promoter sequence for bacteriophage T7 RNA polymerase shows a striking pattern of homology and divergence. Between positions -9 and +4, the sequences are virtually identical, whereas between positions -17 and -10, the sequences are quite different. It is postulated that these sequence subsets may perform different functions in transcription initiation by the phage RNA polymerases.

The development of coliphage T3 in infected cells is controlled at the transcriptional level by two distinct RNA polymerases. First, a set of "early" genes is transcribed by the host *Escherichia coli* RNA polymerase. Then, one of the early gene products (that of gene 1) that is itself an RNA polymerase copies the "late" T3 genes (1-4).

T3 RNA polymerase is interesting for several reasons. It is a single-subunit enzyme that, by itself, can catalyze all the partial reactions of transcription. Also, it has remarkable template specificity and efficiently transcribes only T3 DNA (2, 5). Even DNA from the related bacteriophage T7 is a poor template for it. Finally, recent studies (6) have shown that the set of 5'-terminal sequences for RNA chains synthesized by T3 RNA polymerase is restricted, implying a high degree of initiation specificity.

Thus, it was of interest to determine the DNA sequences of one or more promoters for T3 RNA polymerase. In this paper, we present data for a promoter located in the early region of the T3 genome. We show that this promoter, although not recognized by T7 RNA polymerase, has striking similarities (as well as differences) to T7 prototype promoter (7-9). The implications of these results for promoter function are discussed.

MATERIALS AND METHODS

Restriction Map of T3 DNA. The cleavage sites for a number of restriction enzymes on T3 DNA have been mapped by standard methods (ref. 10; Fig. 1). A similar map has recently been published by Bailey *et al.* (15).

Isolation of Restriction Fragments. The 300-base-pair (bp) *Hpa* I-N fragment and fragments derived from it were isolated from *Hpa* I digests of T3 DNA by electrophoresis on 5% polyacrylamide gels, followed by elution of the ethidium bromide-

stained (or ³²P-labeled) band by the "crush-and-soak" method of Maxam and Gilbert (12).

Isolation of [γ -³²P]GTP-Labeled T3 RNA Polymerase Transcripts. T3 RNA polymerase transcripts were prepared in reaction mixtures (0.1 ml) containing 50 mM Tris·HCl (pH 8.0); 20 mM MgCl₂; 4 mM dithiothreitol; 0.4 mM each of ATP, CTP, UTP, and [γ -³²P]GTP (11,000 cpm/pmol); 0.3 pmol of T3 DNA; and 1-3 pmol of T3 RNA polymerase. Incubation was for 30 min at 37°C. The reaction mixtures were digested with RNase-free DNase I (10 μ g/ml) for 30 min at 0°C and then extracted with phenol. The extracts were treated with 2.5 vol of ethanol at -20°C, and the products were chromatographed on Sephadex G-50 columns to remove unincorporated nucleoside triphosphates.

Transcription of *Hpa* I-N Fragment. Typical transcription mixtures had the same composition as above, except that 0.5 pmol of *Hpa* I-N fragment was used in place of the T3 DNA and 1 pmol of T3 RNA polymerase was used per 50 μ l of mixture. After 30 min of incubation at 37°C, reactions were stopped by addition of NaDodSO₄ (final concentration, 0.1%) and EDTA (final concentration, 20 mM), followed by heating at 65°C for 1 min. The mixtures (or aliquots thereof) were treated with ethanol, and the pellets were rinsed once with 95% ethanol, dried, and suspended in 50 μ l of 87% formamide containing xylene cyanol and bromophenol blue before loading on 5% acrylamide/8 M urea gels for electrophoretic analysis.

Preparation of 5'-³²P-Labeled *Hpa* I-N Transcripts. The following procedure was used to end label the major *Hpa*-N transcript. Transcription reactions containing only unlabeled ribonucleoside triphosphates were prepared. The products were subjected to electrophoresis on a 5% acrylamide/8 M urea gel alongside *Hpa* I-N transcripts labeled with [α -³²P]UTP as markers. After autoradiography, the regions in the unlabeled lanes corresponding to the ³²P-labeled major transcript were excised and eluted by the "crush-and-soak" method (12). The 5'-triphosphate end of the RNA was dephosphorylated with bacterial alkaline phosphatase and then labeled with [γ -³²P]ATP and polynucleotide kinase as described (16). The labeled transcript was separated from unreacted ATP and any degradation products by electrophoresis on 5% acrylamide/8 M urea gels and eluted as above.

RESULTS

Mapping T3 RNA Polymerase Start Sites on T3 DNA. The first step in the determination of the structure(s) of the promoters for T3 RNA polymerase was to locate the major transcriptional start sites in the restriction fragments of known map position. To do this, use was made of the fact that T3 RNA poly-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair(s).

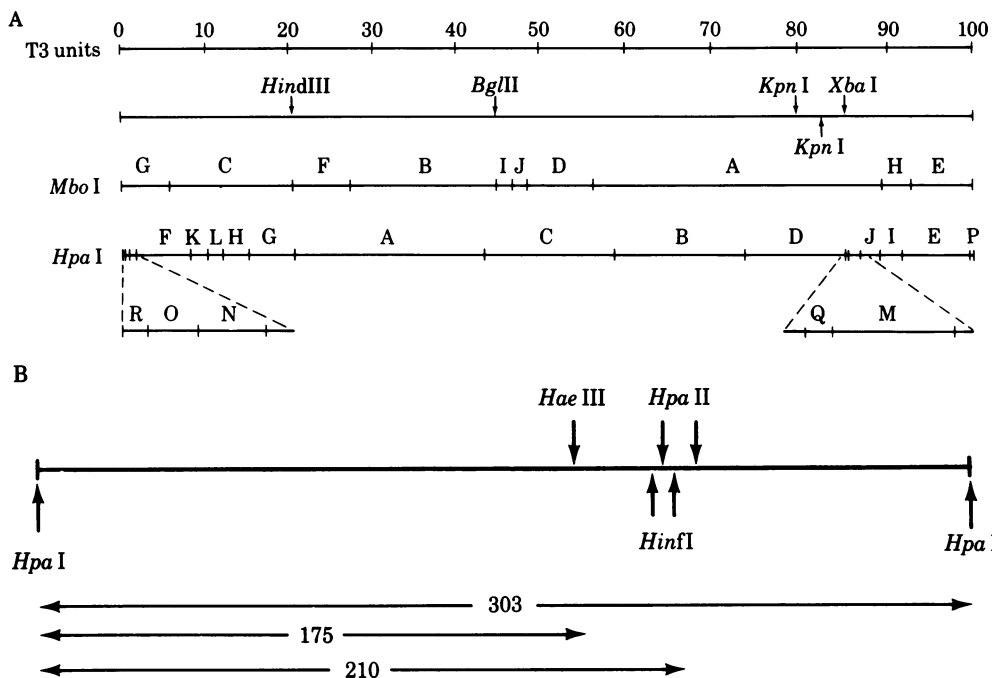


FIG. 1. Restriction endonuclease cleavage sites on T3 DNA. (A) Sites of T3 DNA cleavage by *Hind*III, *Xba*I, *Bgl*II, *Kpn*I, *Mbo*I, and *Hpa*I nucleases. Mapping was by a combination of standard techniques, including (i) use of deletions R4 and R14 (11) in the early region for orientation with the genetic map, (ii) 5'-end labeling to identify terminal fragments (12), (iii) double and partial digestions (13), (iv) combined exonuclease III-exonuclease VII digestion for deletion of restriction sites from both ends of the DNA (13), and (v) the partial-digestion mapping method of Smith and Birnstiel (14). The horizontal scale calibrates percentage units from the genetically defined left end of T3 DNA. One T3 unit equals 380 bp. (B) Sites of cleavage by *Hinf*I, *Hae*III, and *Hpa*II nucleases within the *Hpa*I-N fragment. The map was constructed from end-labeling experiments and DNA sequencing. Distances of various sites (in bp) are from the leftward *Hpa*I site.

merase initiates solely with GTP (17). 5'-³²P-labeled RNA was prepared in an *in vitro* reaction mixture containing T3 DNA, T3 RNA polymerase, and nucleoside triphosphates having [γ -³²P]GTP as the only labeled nucleotide. This RNA was hybridized to Southern blots of denatured *Hpa*I and *Mbo*I restriction fragments of T3 DNA (18). After hybridization, excess filter-bound RNA, as well as nonhybridized portions of DNA-bound RNA, were digested with RNase A. Autoradiography of the treated filters was used to locate the bands corresponding to those DNA fragments containing internal RNA start sites (and thus protecting the 5'-³²P-label of the probe). The following *Hpa*I bands were found to contain transcription start sites: *Hpa*I-A, *Hpa*I-B, C (not resolved by this procedure), *Hpa*I-E, *Hpa*I-G, and *Hpa*I-N. *Mbo*I fragments hybridizing to 5'-³²P-labeled RNA included *Mbo*I-A, *Mbo*I-B, *Mbo*I-C, *Mbo*I-D, *Mbo*I-E, *Mbo*I-G, and *Mbo*I-I (Fig. 2). Under different conditions of transcription (i.e., 10 mM Mg²⁺/50 mM KCl), an additional *Hpa*I fragment—*Hpa*I-H—was observed to hybridize to [γ -³²P]GTP-labeled T3 RNA polymerase transcripts (data not shown). Therefore, these fragments were suitable for promoter-structure analysis.

Template Activity of *Hpa*I-N Fragment. One of the restriction fragments found to contain a T3 RNA polymerase start site was *Hpa*I-N, a 300-bp fragment located about 330 bp from the genetic left end of T3 DNA [in front of the 0.3 gene for SAMase (19); see Fig. 1]. However, the hybridization assay used does not detect complete promoter elements, including sequences upstream from the start site known to be important for *E. coli* RNA polymerase (20–23). Furthermore, it is not possible to conclude solely on the basis of this assay whether one or more than one promoter is present in the fragment. Therefore, the promoter activity of the *Hpa*I-N fragment was examined by carrying out *in vitro* transcription reactions with T3 RNA polymerase using the fragment as template. Gel analysis of the transcription products, labeled with [α -³²P]UTP, showed the presence of one major band (transcript A) and a few minor bands (Fig. 3). Transcript A is 230 ± 10 nucleotides long and represents the major promoter within the *Hpa*I-N DNA fragment. The other bands probably represent minor start sites for T3 RNA polymerase and are not considered further in this report.

The specificity of initiation from the major *Hpa*I-N promoter was examined in transcription reactions containing [γ -³²P]GTP or [γ -³²P]ATP as the labeled substrate. Transcript A was labeled

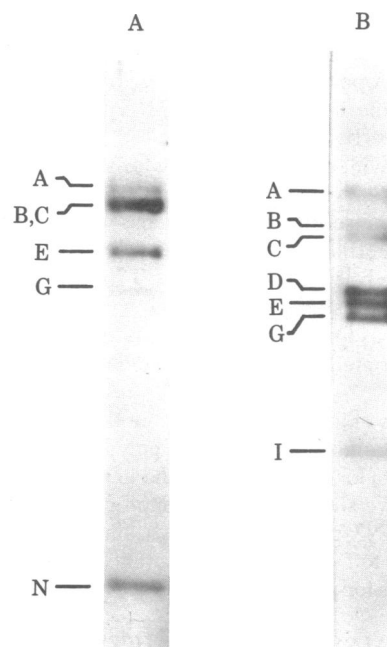


FIG. 2. Hybridization of [γ -³²P]GTP-labeled T3 RNA polymerase transcripts to T3 restriction fragments. Restriction digests of T3 DNA and *Hpa*I (A) or *Mbo*I (B) were subjected to electrophoresis on 1% agarose gels alongside ³²P-labeled markers (not shown). The separated fragments were denatured with alkali, transferred to nitrocellulose filters by the technique of Southern (18), and hybridized for 15 hr at 60°C in 1 ml of 0.3 M NaCl/0.03 M sodium citrate/0.1% NaDodSO₄/100 μ g of *E. coli* tRNA/225,000 cpm of [γ -³²P]GTP-labeled T3 RNA polymerase products. The filters were then washed with 0.3 M NaCl/0.03 M sodium citrate, incubated in the same mixture with RNase A (19 μ g/ml) for 30 min at room temperature, washed again, dried, and autoradiographed. *Hpa*I and *Mbo*I fragments hybridizing to the probe are indicated.

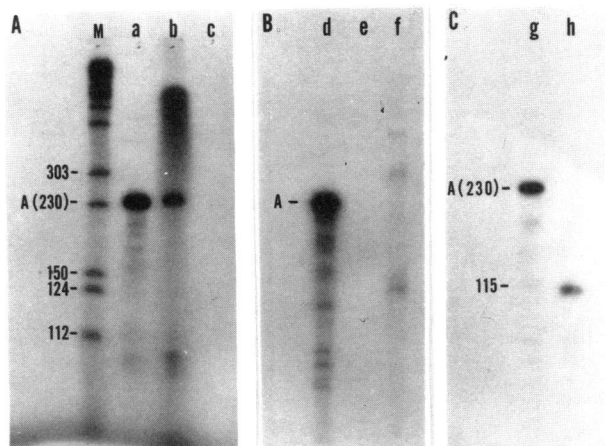


FIG. 3. Template activity of *Hpa* I-N DNA fragment. (A) *Hpa* I-N (0.5 pmol) was transcribed with T3 RNA polymerase (1 pmol) as described using the following labeled substrates: lane a, [α - 32 P]UTP at 5000 cpm/pmol; lane b, [γ - 32 P]GTP at 12,000 cpm/pmol; and lane c, [γ - 32 P]ATP at 10,000 cpm/pmol. (B) *Hpa* I-N DNA (0.5 pmol) was transcribed in the presence of [α - 32 P]UTP at 1900 cpm/pmol using the following enzymes: lane d, T3 RNA polymerase at 16 units (1 pmol); lane e, T7 RNA polymerase at 3 units; and lane f, *E. coli* RNA polymerase at 6.5 units. (C) T3 RNA polymerase (1 pmol) was used to transcribe 0.5 pmol of *Hpa* I-N (lane g) or 0.5 pmol *Hpa* I-N digested with *Hae* III [i.e., a mixture of *Hpa* I/*Hae* III (175 bp) and *Hae* III/*Hpa* I (115 bp); (see Fig. 1)] (lane h). [α - 32 P]UTP at 1700 cpm/pmol was used as the labeled substrate. All reactions were processed as described and run on 5% acrylamide/8 M urea gels for 3 hr at 120 V. Lane M, denatured 32 P-labeled DNA molecular-length markers. Sizes (in nucleotide residues) are indicated at the left of each panel. The high molecular weight smear in lane b probably represents an aggregate of [γ - 32 P]GTP with reaction components.

when [γ - 32 P]GTP was present but not in the presence of [γ - 32 P]ATP. These observations are consistent with the reported specificity of T3 RNA polymerase for GTP as the initiating nucleotide (17).

We also determined whether the major *Hpa* I-N promoter was recognized by RNA polymerases other than the T3 enzyme. In *in vitro* transcription reactions containing *E. coli* RNA polymerase, no band corresponding to transcript A was detected, although other bands were present. Furthermore, T7 RNA polymerase was completely inactive in the reaction. These results confirm the unique specificity of the T3 promoter for its cognate (T3) RNA polymerase.

The direction of transcription from the major promoter was deduced as follows. *Hpa* I-N was digested with *Hae* III to generate two fragments, *Hpa* I/*Hae* III (175 bp) and *Hae* III/*Hpa* I (115 bp), and the mixture was then used as template for T3 RNA polymerase. Under these conditions, the size of the major transcript A was reduced from 230 to 115 nucleotides, as expected if transcription proceeds from left to right as shown in Fig. 1—i.e., the same as the general direction of T3 RNA polymerase transcription on T3 DNA.

DNA Sequence of the Promoter Region. To sequence the promoter region in *Hpa* I-N, the restriction fragment was labeled at its 5'-ends by using [γ - 32 P]ATP and T4 polynucleotide kinase (12). The labeled fragment was then recut with *Hae* III, *Hinf*I, or *Hpa* II to generate fragments labeled at a single 5'-end (see Fig. 1B). Fragments *Hpa* I/*Hinf*I (210 bp), *Hinf*I/*Hpa* I (92 bp), *Hpa* I/*Hae* III (175 bp), and *Hae* III/*Hpa* I (115 bp) were subjected to DNA sequencing by the Maxam-Gilbert method (12) to obtain the entire sequence of *Hpa* I-N DNA fragment.

The DNA sequence of a portion of *Hpa* I/*Hae* III (175 bp) fragment is shown in Figs. 4 and 5. The start point for transcrip-

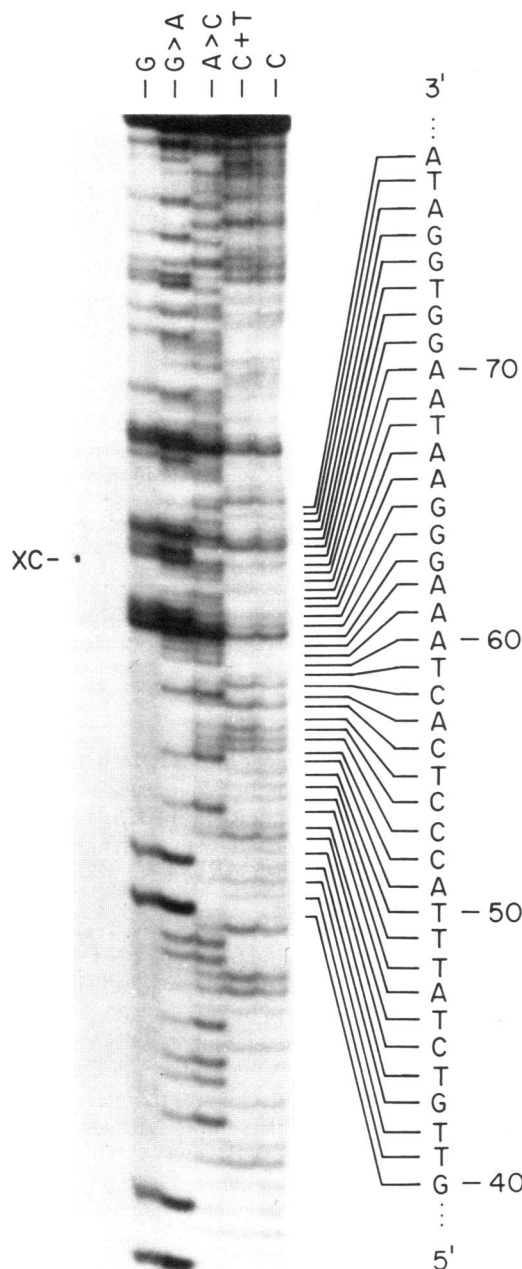


FIG. 4. DNA sequence of the major *Hpa* I-N promoter region. The *Hpa* I/*Hae* III (175-bp) fragment, labeled at the 5'-end of the (-)strand (i.e., the *Hpa* I site), was subjected to base-specific chemical cleavage procedures (12) and then subjected to electrophoresis on an 8% acrylamide/0.4% bisacrylamide/8 M urea sequencing gel with 0.1 M Tris-borate buffer at 25 W (constant power) until the xylene cyanol (XC) dye migrated two-thirds the length of the gel. The various base-specific reactions are indicated at the top, and a portion of the sequence is given at the right. Numbers indicate distance (in bp) from the *Hpa* I end.

tion from the major promoter was deduced as follows. Knowing the size of transcript A (230 ± 10 nucleotides) and assuming that transcription continues to the end of the *Hpa* I-N fragment in the left-to-right direction, we can tentatively assign the start site to 70 ± 10 bp from the left (*Hpa* I site) end of the *Hpa* I/*Hae* III (175 bp) fragment. The DNA sequence in this region showed the presence of two identical octameric sequences, 5' G-G-G-A-A-U-A-A 3', starting at positions 63 and 87, respectively. The hexamer 5' G-G-G-A-A-U 3' has been shown to be one of the

As the first step toward identification of the sequences necessary and sufficient for promoter recognition by T3 RNA polymerase, we have determined the DNA sequence around the major transcriptional start site in the *Hpa* I-N DNA fragment. Similar studies of a number of other promoter regions of T3 DNA will be necessary, however, before a common (consensus) sequence can be identified, although our results enable us to compare this system with that of the related phage T7. If the T3 and T7 sequences are aligned with the transcription start-points in register and, we assume that, in T3 as in T7, the 23-bp stretch from position -17 to position +5 (see Fig. 5) defines the promoter sequence, we find a striking pattern of sequence homology and divergence. Between positions -9 and +4, the T7 and T3 sequences are identical except for the single base pair at position -2. Upstream from this 13-bp stretch, the sequences diverge, corresponding at only 3 out of 8 positions (between -10 and -17). It is tempting to postulate that these two sets of sequences constitute the different elements required for transcription initiation in T3 and T7 systems.

The sequence between positions -9 and +4 may serve a basic function for initiation—e.g. helix unwinding by the RNA polymerase. It is interesting that, for the two *E. coli* promoters, *lac* UV5 and T7A3, similar stretches of DNA, located at similar positions (between -9 and +3 and between -9 and +2, respectively), are apparently unwound within the promoter complex (25). On the other hand, the sequence between -17 and -10 (or part thereof) could serve as a specificity element at which the polymerase initially binds. This would explain the lack of T3 promoter recognition by T7 RNA polymerase and vice versa (refs. 2 and 5; this report).

The presence of two functionally different subsets of DNA sequence in T3 and T7 promoters would imply a similar dichotomy at the level of the respective RNA polymerases. Distinct domains of the protein may interact, perhaps cooperatively, with the two subsets of the promoter. Homologous polymerase-promoter interaction would then be a result of binding of the specificity domain of the protein to the specificity element (between positions -17 and -10) of the promoter sequence. This model is supported by studies with T3-T7 hybrid RNA polymerases that show that a restricted region near the COOH-terminal end of the gene 1 protein determines its preference for T3 or T7 DNA as a template (26).

We are indebted to Dr. Edmund Benz of this Department for introducing us to DNA sequencing techniques. This work was supported by National Institutes of Health Grant GM15399-12.

- Maitra, U. (1971) *Biochem. Biophys. Res. Commun.* **43**, 443–450.
- Dunn, J. J., Bautz, F. A. & Bautz, E. K. F. (1971) *Nature (London New Biol.* **230**, 94–96.
- Chakraborty, P. R., Bandyopadhyay, P., Huang, H. H. & Maitra, U. (1974) *J. Biol. Chem.* **249**, 6901–6909.
- Dunn, J. J., McAllister, W. T. & Bautz, E. K. F. (1972) *Virology* **48**, 112–125.
- Chakraborty, P. R., Sarkar, P., Huang, H. H. & Maitra, U. (1973) *J. Biol. Chem.* **248**, 6637–6646.
- Maitra, U., Jelinek, W., Yudelevich, A., Majumder, H. & Guha, A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3908–3911.
- Rosa, M. (1979) *Cell* **16**, 815–825.
- Oakley, J. L., Strothkamp, R. E., Sarris, A. H. & Coleman, J. E. (1979) *Biochemistry* **18**, 528–537.
- Panayotatos, N. & Wells, R. D. (1979) *Nature (London)* **280**, 35–39.
- Sarkar, P., Adhya, S., Musich, P. & Maitra, U., *Gene*, in press.
- Studier, F. W. & Movva, N. R. (1976) *J. Virol.* **19**, 136–145.
- Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
- McDonnell, M. W., Simons, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119–146.
- Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
- Bailey, J. N., Dembinski, D. R. & McAllister, W. T. (1980) *J. Virol.* **35**, 176–183.
- Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977) *Nucleic Acids Res.* **4**, 2527–2538.
- Maitra, U. & Huang, H. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 55–59.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
- Hausmann, R. & Härle, E. (1971) *Proc. First Eur. Biophys. Congr.*, eds Broda, E., Locker, A. & Springer-Lederer, H. (Wien Med. Akad., Vienna), Vol. 1, pp. 467–488.
- Pribnow, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 784–789.
- Pribnow, D. (1975) *J. Mol. Biol.* **99**, 419–443.
- Takanami, M., Sugimoto, K., Sugisaki, H. & Okamoto, T. (1976) *Nature (London)* **260**, 297–302.
- Seeburg, P., Nusslein, C. & Schaller, H. (1977) *Eur. J. Biochem.* **74**, 107–113.
- Dunn, J. J., McAllister, W. T. & Bautz, E. K. F. (1972) *Eur. J. Biochem.* **29**, 500–508.
- Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980) *Cell* **20**, 269–281.
- Hausmann, R. & Tomkiewicz, C. (1976) in *RNA Polymerase*, eds Losick, R. & Chamberlin, M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 731–743.

Correction. In the article "Location, function, and nucleotide sequence of a promoter for bacteriophage T3 RNA polymerase" by Samit Adhya, Shantanu Basu, Probir Sarkar, and Umadas Maitra, which appeared in the January 1981 issue of *Proc. Natl. Acad. Sci. USA* (78, 147-151), the authors request that the following correction be noted. In the abstract, lines 7 and 8 should be "(⁻)strand sequence is 5' T-A-T-T-T-A-C-C-C-T-C-A-C-T-A-A-A-G-G-G-A-A-U 3'."