Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6290–6294, October 1984
Biochemistry

Proximal and distal domains that control *in vitro* transcription of the adenovirus *IVa2* gene

(non-"TATA"-box promoter/G- and C-rich regions/polymerase II entry site/supercoiled DNA templates)

V. NATARAJAN, M. J. MADDEN, AND N. P. SALZMAN*

Laboratory of Biology of Viruses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20205

Communicated by James E. Darnell, Jr., June 25, 1984

ABSTRACT **The adenovirus *IVa2* gene, which is expressed at an intermediate time in the viral infectious cycle, is separated from the adenovirus major late promoter (MLP) 5′ start site by 210 base pairs and is transcribed from the opposite strand. In contrast to the MLP, the *IVa2* gene does not contain a "TATA" box upstream from its 5′ start sites. By using a series of deletion mutants, two upstream control regions that are rich in cytidine residues, one proximal to the cap site at nucleotide positions −39 to −48 and a distal domain between nucleotide positions −152 and −242 have been identified as essential for *IVa2* transcription (IVa2 cap site is nucleotide position +1). Transcription efficiency is decreased by 70–90% after the deletion of a proximal C-rich domain when either linear or supercoiled DNAs were used as template. However, distal sequences functioned as transcriptional control domains only with covalently closed DNA templates. The deletion of both the proximal and distal regions from covalently closed DNA templates reduces the levels of *IVa2* transcription by a factor of 100–150. When the plasmid pAd242 that contains the 5′ start sites of adenovirus MLP and *IVa2* is transcribed, there is essentially a complete suppression of transcription of the adenovirus *IVa2* gene. The transcription efficiency of *IVa2* is increased 10-fold after deletion of the MLP cap site. A model based on a shared entry site for RNA polymerase II and competition between major late and *IVa2* promoters is proposed to explain the *in vitro* transcriptional results.**

Most genes transcribed by RNA polymerase II have promoter regions that contain a "TATA" sequence located 30 base pairs upstream of the transcription initiation site (1). When transcription is assayed *in vitro*, deletion of the TATA box abolishes transcription, and the introduction of point mutations within it also reduces the efficiency of transcription (2–6). *In vivo*, deletion of the TATA box leads to heterogeneity in 5′ start sites used for initiation, and the level of transcription is also depressed (7–9). Mutants containing base substitutions within the TATA box have been found to occur in nature, and they also reduce the efficiency of transcription (10). Overall, the evidence establishes the TATA box as one critical domain in control of eukaryotic transcription.

Genes that do not have TATA-like sequences in their promoters include polyoma and simian virus 40 (SV40) late genes, the hepatitis surface antigen gene, and the adenovirus *IVa2* gene (11–14). Studies from our laboratory have demonstrated that in the SV40 late promoter, the G-G-T-A-C-C-T-A-A-C-C sequence present at the −30 region may have similar functions both *in vivo* and *in vitro* as that of the TATA box (13). A sequence with extensive structural homology to this SV40 sequence has also been located in the same position relative to the 5′ start site of the hepatitis B surface antigen (14). Studies of other non-TATA-box-containing promoters will lead to a fuller understanding of recognition signals that are important in the mechanisms of regulation of transcription. We recently reported conditions for the efficient transcription of adenovirus *IVa2* promoter *in vitro* (15). The analysis of the control sequences of this promoter was undertaken by construction of promoter deletion mutants that were transcribed *in vitro*. In this paper, we report the identification of two upstream sequences that are necessary for efficient *in vitro* transcription. The distal control domain only functions when the template is in a covalently closed form.

MATERIALS AND METHODS

Enzymes and Reagents. All restriction endonucleases used in this study were purchased either from New England Biolabs or from Bethesda Research Laboratories. T4 polynucleotide kinase and T4 DNA ligase were obtained from Boehringer Mannheim. BAL-31 nuclease and T4 DNA polymerase were obtained from Bethesda Research Laboratories. [α-^{32}P]UTP and [γ-^{32}P]ATP were obtained from Amersham. All enzymatic reactions were carried out as specified by the manufacturer.

Construction of Adenovirus *IVa2* Promoter Deletion Mutants (pAd Series). The plasmid pϕ4 (Fig. 1*a*) was used to construct deletion mutants upstream of the IVa2 RNA initiation site. Twenty micrograms of pϕ4 DNA was linearized with *Hin*dIII and treated with 3.5 units of BAL-31 exonuclease. Aliquots of this reaction were taken at 45 sec to 5 min, and they were pooled. After the DNA was digested with *Eco*RI, fragments ≤ 433 base pairs that were generated from the *Eco*RI–*Hin*dIII adenovirus insert were purified by agarose gel electrophoresis. pBR322 digested with *Eco*RI and *Bam*HI was ligated with the BAL-31-treated fragments at the *Eco*RI site. This was then treated with T4 DNA polymerase, followed by a blunt-end ligation. This DNA was used to transform the HB101 strain of *Escherichia coli*, and DNA isolated from ampicillin-resistant colonies was screened for deletions in the *IVa2* promoter by using restriction enzymes. The sequence of plasmids used in the transcription studies was determined to identify the exact extent of the deletion. For sequence analyses, the fragment between the *Eco*RI and *Nae* I cleavage sites for the pAd plasmids (Fig. 1*b*) was cloned into the *Eco*RI and *Hin*dII sites of plasmid M13mp8, and the sequence was determined by the dideoxy method (16). All plasmid DNAs used in transcription analysis were amplified in minimal medium and purified by two cesium chloride/ethidium bromide gradient centrifugations (17). The pAd plasmid series designates the pBR322 plasmid with the adenovirus *IVa2* promoter region. Each of the pAd plasmids contains 200 base pairs downstream and various lengths of DNA upstream from the *IVa2* cap site. The number designation following pAd indicates the number of upstream

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.

Abbreviations: MLP, major late promoter; np, nucleotide position(s); SV40, simian virus 40.

*To whom reprint requests should be addressed.

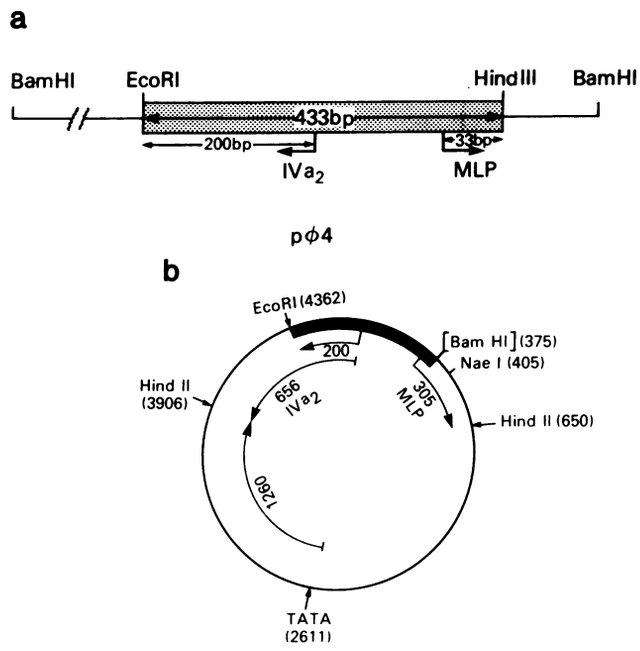

FIG. 1. Construction of *IVa2* promoter deletion mutants (*a*) Structure of the pϕ4 plasmid. The DNA fragment of adenovirus, serotype 2 DNA, from map unit 15.3 to 16.5 was cloned in the *Eco*RI and *Hin*dIII sites of pBR322. This plasmid, pϕ4, was provided by J. Manley. The adenovirus DNA is represented by the stippled bar, and pBR322 DNA is presented as a thin line. The relevant restriction enzyme sites used in the study are indicated. The arrows show the directions of transcription of the *IVa2* and the major late promoters (MLP). (*b*) Structure of pAd plasmid. The construction of the pAd plasmid series is described in *Materials and Methods*. The thick line represents adenovirus DNA, while pBR322 is represented by a thin line. The position of the RNA start site and the direction of transcription from the MLP and *IVa2* promoters are indicated. The position of the TATA-box-like sequence present in pBR DNA and the length of expected run-off transcripts are shown when *Hin*dII-digested DNA is used as the template for transcription. A 200-base-pair length of S1-nuclease-resistant DNA is obtained when an *Eco*RI–*Hin*dII probe is used in the S1-nuclease mapping of *IVa2* RNA. Each of the pAd plasmids contain 200 base pairs downstream and various lengths of DNA upstream from the *IVa2* cap site. The number designation following pAd indicates the number of upstream base pairs.

base pairs. The plasmid pAd242 contains the same number of nucleotides upstream from the *IVa2* 5′ start site [nucleotide position (np) +1] as are contained in pϕ4, which originally was used as a transcription template, and it is considered the wild type in this study.

***In Vitro* Transcription Assays.** The whole cell extract (extract A) used in all of the transcription assays was prepared as described (15) with one modification. During the preparation of extract A, a modified buffer A (50 mM Tris·HCl, pH 7.9/12.5 mM $MgCl_2$/40 mM $(NH_4)_2SO_4$/0.1 mM EDTA/2 mM dithiothreitol/17% glycerol) was used for final dialysis. The dialysis was carried out for 16 hr with a change of buffer after 4 hr. The extract contained 10 mg of protein per ml as estimated by a dye-binding assay (18). The DNAs of different deletion mutants were digested with *Hin*dII prior to their use as templates for run-off assays. Standard reaction mixtures in a final total volume of 60 μl contained 25 mM Tris·HCl (pH 7.9); 6.25 mM $MgCl_2$; 20 mM $(NH_4)_2SO_4$; 1 mM dithiothreitol; 0.05 mM EDTA; 8.5% glycerol; 400 μM each ATP, CTP, and GTP; 40 μM UTP; 10 μCi (1 Ci = 37 GBq) of [α-^{32}P]UTP (410 Ci/mmol); and 30 μl of whole-cell extract. The concentrations of DNA used are shown in the figure legends. The reaction was incubated at 30°C for 60 min. The purification

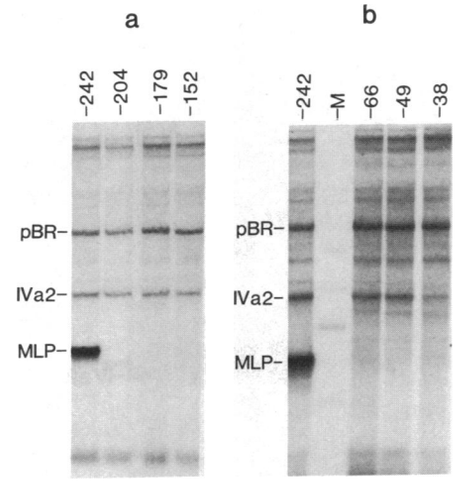

FIG. 2. *In vitro* transcription of pAd deletion mutants. (*a* and *b*) Two separate experiments. The *Hin*dII-digested templates (10 μg/ml) were transcribed, and purified RNA was denatured with glyoxal and run on a 1.4% agarose gel as described (15). The position of transcripts initiating from a TATA sequence in pBR322 (designated pBR), RNA transcribed from adenovirus *IVa2* promoter (IVa2), or RNA transcribed from MLP are indicated. Lane M had DNA markers, and the other lanes are designated by the pAd construct used in the transcription. Transcription of pAd242 (242) is shown in both panels for comparison with other mutants. With pAd204, the run-off transcript from the MLP is expected to be 38 nucleotides shorter than from pAd242, and this is observed (*a*).

and characterization of RNAs were carried out as described earlier (15). When RNAs were transcribed for S1-nuclease mapping, the [α-^{32}P]UTP was omitted.

For S1-nuclease mapping of the 5′ end of IVa2 RNA, pAd242 was digested with *Eco*RI and 5′-end-labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase, and the fragment from the *Eco*RI to the *Hin*dII site [np 650 (see Fig. 1*b*)] was isolated. The conditions used for hybridization of this probe to RNA and S1-nuclease digestion have been described (15). The S1-nuclease-protected DNA was fractionated by electrophoresis on a denaturing polyacrylamide gel (19) (30 × 15 × 0.1 cm) at 250 V for 16 hr and was autoradiographed.

Quantitation of *in Vitro* Transcription. After gel electrophoresis and autoradiography, the RNA bands were scanned with a densitometer (BioMed Instruments, Fullerton, CA). Autoradiograms of different exposures from at least two different experiments were scanned to obtain an average value.

RESULTS

A series of deletion mutants upstream of the adenovirus *IVa2* transcription 5′ start sites were constructed by using the BAL-31 exonuclease method as described. These DNAs were digested with *Hin*dII and transcribed *in vitro* in a run-off assay. The results of such an experiment are shown in Fig. 2. If RNA initiates specifically from the *in vivo IVa2* start sites (20), transcripts about 650 nucleotides long would be expected in the run-off assay (Fig. 1*b*). Transcription from pAd242 showed an RNA of the expected length (Fig. 2*a*). This plasmid contains all of the adenovirus sequences that are present in plasmid pϕ4, which was used in previous *in vitro* transcription studies of the *IVa2* gene (15). The pAd242 also contains the adenovirus MLP initiation site, and the 305-nucleotide-long RNA that is expected from this promoter also was transcribed (Fig. 2 *a* and *b*). pBR322 also has a TATA-box-like sequence that is recognized by the *in vitro* transcription machinery (21). It was present in all of the pAd constructs; with *Hin*dII-digested templates, it generated a

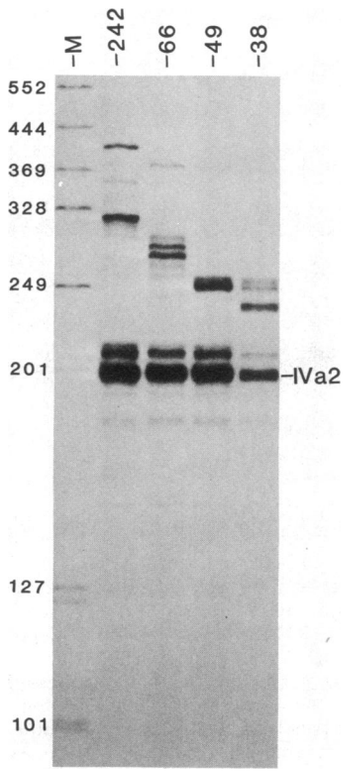

FIG. 3. S1-nuclease mapping of RNAs synthesized on linear DNAs of different mutants. *In vitro* transcriptions of different mutant DNAs (10 μg/ml) were carried out as in Fig. 2. The purified RNA was used for S1-nuclease mapping as described. Lane M had DNA markers. The other lanes are designated by the pAd construct used as the template for transcription. The position of the 5'-end-labeled DNA fragment protected by the IVa2 transcript is indicated.

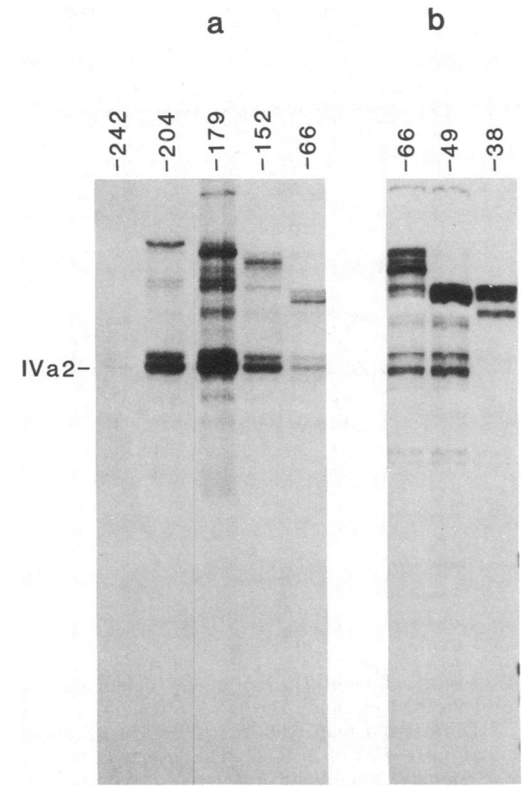

FIG. 4. S1-nuclease mapping of RNAs synthesized on supercoiled DNA of the pAd deletion mutants. (*a* and *b*) Two separate experiments. *In vitro* transcription of different mutants (12.5 μg/ml) were carried out with supercoiled DNAs as templates. The purified RNA was used for S1-nuclease mapping. The position of the 5'-end-labeled DNA fragment protected by the IVa2 transcript is shown. The designation noted above the lanes indicate the pAd construct used as the template for transcription. These results were reproduced in repeated experiments, and the same results were obtained with separate batches of whole-cell extract.

1260-nucleotide-long RNA in the *in vitro* transcription system. This RNA transcript was used as an internal control in the run-off assays to determine the efficiency of *IVa2* gene transcription.

The level of transcription from the *IVa2* promoter containing upstream deletion mutations remained relatively constant until the 11 base pairs between np −49 and −38 were deleted (Fig. 2 *a* and *b*), and this led to a 70% drop in transcription efficiency (Fig. 2*b*). The sequence in this region that controls the level of transcription, C-C-C-C-T-C-C-C-A-C-T, has 8 cytidine residues and 10 pyrimidine residues out of 11 nucleotides (see Fig. 5).

Other deletions between np −242 and −49 did not affect the efficiency of *IVa2* transcription in the run-off assay with a linear DNA template (Fig. 2*a*). The mutant pAd204 had lost 38 nucleotides, including the MLP cap site, compared with pAD242, and the transcription efficiency from the MLP of pAd204 was decreased by 90%. However, this neither enhanced nor suppressed *IVa2* promoter transcription (Fig. 2*a*). Results obtained in the run-off assay were confirmed by the S1-nuclease-mapping method. If RNAs were to initiate at the specific initiation sites that are used *in vivo*, DNAs of 198- and 200-nucleotide length would be protected with the 5'-end-labeled DNA probe (Fig. 1*b*). The major band of 200 nucleotides that was observed (Fig. 3) could be further resolved into two bands of 198 and 200 nucleotides, corresponding to two *in vivo* initiation sites under more stringent gel conditions (15). As noted previously in the run-off assay, S1-nuclease mapping of RNA transcribed from linear templates also showed that the efficiency of transcription decreased by 70% when the C-rich region present between np −49 and −38 was removed.

It has been reported that when supercoiled DNAs are used in the *in vitro* transcription system, the role of sequences far upstream of the initiation site became apparent with promoters such as the adenovirus MLP (22) and the histone genes (3). To investigate this possibility in the *IVa2* promoter, transcription of various deletion mutants were carried out with supercoiled DNAs as templates. The quantities of RNA synthesized with different 5' deletion mutants were then estimated using the S1-nuclease-mapping method. The results of such an experiment are presented in Fig. 4.

The results with supercoiled templates are in sharp contrast to those obtained with linear templates with deletions between np −152 and −242. pAd204 and pAd179 were transcribed 10- to 15-fold more efficiently than was pAd242 (Fig. 4*a*), whereas pAd152 was 2- to 3-fold more efficient than pAd242. pAd66 and pAd49 were transcribed with an efficiency that is similar to that of pAd242. The results obtained with supercoiled deletion mutants pAd49 and pAd38 were similar to those obtained with linear templates, where the 11-base-pair deletion led to a 90% drop in the transcription level of pAd38 as compared with pAd49. Thus, this 11-base-pair region is important for efficient transcription with both linear and supercoiled DNA.

DISCUSSION

In the present study, we demonstrated that two regions upstream of the cap site are necessary for efficient *in vitro* transcription of the *IVa2* gene. One region is located between

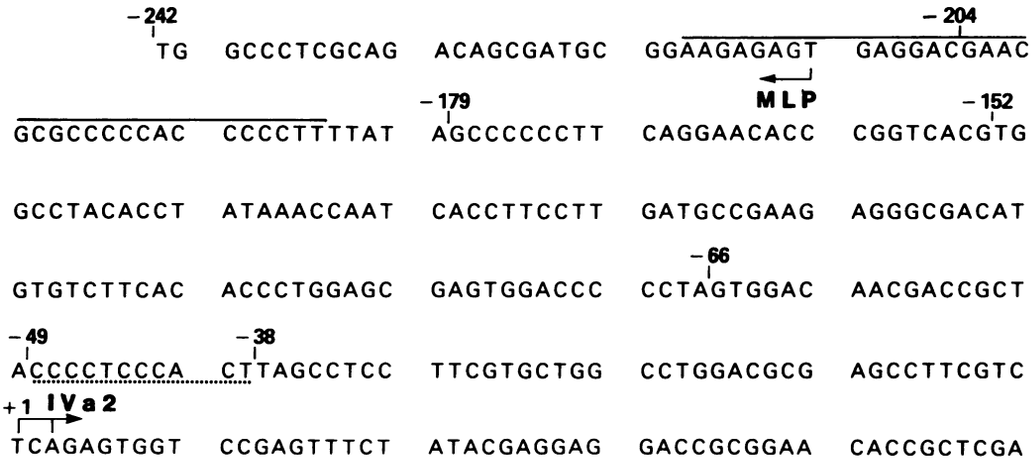

FIG. 5. The DNA sequence of the adenovirus *IVa2* promoter region. The nucleotide sequence is from Baker and Ziff (20). The direction and site of initiation of transcription of *IVa2* and MLP are shown by rightward and leftward arrows, respectively. The 5′ end point of each pAd construct is indicated with a corresponding number. The dotted underline shows the proximal control sequence of the *IVa2* promoter. The nucleotide sequence whose secondary structure has been studied (23, 24) is shown by an overline.

np −38 and −49, and the other is located between −152 and −242, and both are rich in cytidine residues (Fig. 5). The efficiency of transcription is reduced by 70–90% when the proximal domain is deleted with either linear or supercoiled DNA as template.

It is only when the supercoiled DNAs of various deletion mutants are added as templates in the transcription mixture that the effect of the distal C-rich domain 152–242 base pairs upstream from the 5′ start site is evident. Since this domain is dispensable for transcription of linear DNA, where the free ends of the DNA molecule can serve as the polymerase entry site, we propose that this distal domain is required as an entry site for the transcription machinery for covalently closed DNA templates. Similar observations have been made for the sea urchin histone gene promoter, which contains an upstream domain between −111 and −139 nucleotides from the capsite that is necessary for *in vitro* transcription of this gene when the template is in a supercoiled form but is dispensable when linear DNA templates are used for transcription (3). One interpretation of our data on the regulatory role of distal sequences is that a stem–loop structure is important for *IVa2* gene expression. In adenoviruses, a stem–loop structure has been proposed involving nucleotides located between np −26 and +8 from the MLP RNA initiation site (23). This site corresponds to np −185 to −214 for the *IVa2* promoter RNA initiation site (Fig. 5). Larsen and Weintraub (25) have observed an S1-nuclease-sensitive site in the MLP region when it is in a supercoiled configuration, which has been more precisely mapped by Goding and Russell (24) to the −30 (TATA) region of the MLP, corresponding to the −180 region of the *IVa2* promoter (Fig. 5). The presence or absence of nuclease-sensitive sites has been related to differences in the secondary structure of DNA (25). These findings provide experimental confirmation that DNA in this region exists in a different conformational form in supercoiled DNA. The region that exists in a different conformation in supercoiled DNA is part of the region (np −152 to −242) that is necessary for *IVa2* transcription.

Adenovirus MLP lies adjacent to the *IVa2* promoter in the adenovirus genome. The RNA initiation sites of these two promoters that are separated by 210 base pairs are transcribed in opposite directions (11, 15, 20). If DNA sequences at np −242 to −152 interact with transcriptional machinery and serve as an entry site for RNA polymerase II and/or transcription factors, we propose that, because of the proximity of this region to the MLP initiation site, this entry site also would be used by MLP. This is consistent with the suggestion advanced earlier by Ziff and Evans (23). The MLP has been shown to be the strongest promoter *in vitro* among the adenovirus promoters tested (11). The suppression of transcription from the *IVa2* gene that is observed with the pAd242 plasmid (see Fig. 4, lane 1) would result from ineffective competition of the *IVa2* promoter with the stronger MLP. When the MLP cap site and TATA box are deleted, there is a 10- to 15-fold enhancement of transcription from the *IVa2* promoter (pAd204 and pAd179), a finding that would be predicted based on competition of the two promoters. There are further decreases in transcription that occur when pAd152 and pAd66 are compared with pAd204 and pAd179. How the loss of these sequences could affect the entry site with a consequently reduced efficiency in initiation of transcription is presently not clear.

We have discussed regulation of transcription assuming that DNA with superhelical turns was serving as the template. However, when the whole cell extract is used to study *in vitro* transcription, it has been reported that superhelical DNA is converted within 5 min to a relaxed, covalently closed structure (22). It has not been determined if formation of a stable transcription complex occurs more or less rapidly than the loss of superhelical turns from the DNA template. If the transcription complex used a covalently closed template that lacks superhelical turns, then specific nucleotide sequences, rather than secondary structural features, would likely regulate the kinetics of formation and stability of a transcription complex. To decide between these models, it will be necessary to define the configuration of the template during *in vitro* transcription.

We thank Dr. J. Manley for his generous gift of the pϕ4 plasmid and acknowledge the valuable assistance of J. Carolan in the preparation of the manuscript.

1. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
2. Hu, S. L. & Manley, J. L. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 820–824.
3. Grosschedl, R. & Birnstiel, M. L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 297–301.
4. Schulz, T. Z., Tsai, S. Y., Itakura, K., Soberon, X., Wallace, R. B., Tsai, M. J., Woo, S. L. C. & O'Malley, B. W. (1982) *J. Biol. Chem.* **257**, 11070–11077.
5. Hirose, S., Takeuchi, K. & Suzuki, Y. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7258–7262.
6. Concino, M., Goldman, R. A., Caruthers, M. H. & Weinmann, R. (1983) *J. Biol. Chem.* **258**, 8493–8496.

7. Osborne, T. F., Gaynor, R. B. & Berk, A. J. (1982) *Cell* **29,** 139–148.
8. Grosveld, G. C., deBoer, E., Shewmaker, C. K. & Flavel, R. A. (1982) *Nature (London)* **295,** 120–126.
9. Dierks, P., Ooyen, A. V., Cochran, M. D., Dobkin, C., Reiser, J. & Weissman, C. (1983) *Cell* **32,** 695–706.
10. Orkin, S. H., Sexton, J. P., Cheng, T., Goff, S. C., Giardina, P. J. V., Lee, J. I. & Kazazian, H. H., Jr. (1983) *Nucleic Acids Res.* **11,** 4727–4734.
11. Fire, A., Baker, C. C., Manley, J. L., Ziff, E. B. & Sharp, P. A. (1981) *J. Virol.* **40,** 703–719.
12. Jat, P., Roberts, J. W., Cowie, A. & Kamen, R. (1982) *Nucleic Acids Res.* **10,** 871–887.
13. Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J. & Salzman, N. P. (1982) *Cell* **31,** 615–633.
14. Cattaneo, R., Will, H., Hernandez, N. & Schaller, H. (1983) *Nature (London)* **305,** 336–338.
15. Natarajan, V., Madden, M. J. & Salzman, N. P. (1983) *J. Biol. Chem.* **258,** 14652–14655.
16. Schreir, P. H. & Cortese, R. (1979) *J. Mol. Biol.* **129,** 169–172.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 75–96.
18. Bradford, M. (1976) *Anal. Biochem.* **72,** 248–253.
19. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87,** 107–110.
20. Baker, C. C. & Ziff, E. B. (1981) *J. Mol. Biol.* **149,** 189–221.
21. Fuhrman, S. A., Beveren, C. V. & Verma, I. M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5411–5415.
22. Hen, R., Sassone-Corsi, P., Corden, J., Gaub, M. P. & Chambon, P. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7132–7136.
23. Ziff, E. B. & Evans, R. M. (1978) *Cell* **15,** 1463–1475.
24. Goding, C. R. & Russell, W. C. (1983) *Nucleic Acids Res.* **11,** 21–36.
25. Larsen, A. & Weintraub, H. (1982) *Cell* **29,** 609–622.