Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6908–6911, November 1982
Biochemistry

Nucleotide sequence requirements for specific initiation of transcription by RNA polymerase I

(rDNA promoter/cell-free transcription/deletion mutants/upstream sequences)

INGRID GRUMMT

Institut für Biochemie der Universität Würzburg, Röntgenring 11, D-8700 Würzburg, Federal Republic of Germany

Communicated by W. Beermann, August 23, 1982

ABSTRACT **The nucleotide sequence(s) specifying RNA polymerase I initiation has been investigated by studying the transcription of deleted and nondeleted mouse ribosomal RNA gene (rDNA) templates *in vitro*. The deletion of 5′-flanking sequences upstream from position −39 did not affect transcriptional activity, but removal of sequences between positions −39 and −34 resulted in a 90% decrease of rDNA transcription. The template activity was completely eliminated by the further deletion of nucleotides −33 to −13. It is concluded that sequences between −34 and −12, upstream from the transcribed region, represent an essential control region for the initiation of transcription *in vitro*. Therefore, this region may be functionally analogous to the T-A-T-A box of RNA polymerase II promoters. In addition to this control region, sequences located further upstream (between positions −45 and −169) may also exert some function in efficient transcription initiation as revealed by competition experiments between wild-type and mutant rDNA templates.**

Ribosomal RNA genes (rDNA) are an attractive experimental system in which to study gene expression in eukaryotic cells because the rate of transcription is efficiently regulated according to the growth rate of the cells (1–4). Thus, the existence of regulatory DNA sequences can be anticipated that define where, along the rDNA repeat, transcription is to begin and how frequently this event is to occur. The elucidation of the molecular mechanisms involved in the regulation of gene expression requires the identification and functional analysis of both transcription factors and DNA sequences that are essential for the readout of the gene.

The role played by specific DNA sequences in promoting transcription can be examined by systematically deleting sequences in the proximity of the initiation site and testing the resulting templates for their transcriptional competence in an *in vitro* assay. Such studies have revealed the importance of the T-A-T-A box (located about 30 nucleotides upstream from the start site of mRNA coding genes) both for the efficiency and specificity of the initiation reaction by RNA polymerase II. On the other hand, the promoter elements of genes transcribed by RNA polymerase III have been shown to be located at two positions within the structural gene itself (for review, see ref. 5).

In contrast, almost nothing is known about the sequence requirements for the transcription of rDNA by RNA polymerase I. In the present study, the 5′ flanking sequences necessary for accurate and quantitative transcription of mouse rDNA have been analyzed. By constructing deletion mutants of cloned rDNA *in vitro* and measuring the ability of the altered templates to direct the synthesis of specific run-off transcripts in cell-free extracts, an essential region located between 12 and 34 nucleotides upstream from the initiation site has been identified. However, competition experiments with wild-type and mutant genes indicate that, in addition to this control region, sequences located further upstream are important for the relative efficiency of the transcription process.

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.

MATERIALS AND METHODS

Construction of Deleted Templates. The nondeleted rDNA clone pMrSP used in this study has been described (6). It contains a 461-base pair (bp) insert of mouse rDNA with 169 bp of the nontranscribed spacer (NTS) region and 292 bp of the external transcribed spacer (ETS) region. Derivatives of this plasmid lacking defined segments of the NTS were prepared as follows. pMrSP (20 μg) linearized with *Sal* I was digested at 20°C with 5 units of BAL-31 for different periods of time. After phenolization, the digested DNA was cleaved with *Pvu* II, and the DNA fragments were fractionated on a 5% polyacrylamide gel. Appropriate gel regions were cut out, and the recovered fragments were inserted between the *Sal* I–*Pvu* II sites of plasmid pBR322 by blunt-end ligation. The extent of deletion in the resulting clones pMrΔ-86, pMrΔ-45, and pMrΔ-34 was determined by DNA sequence analysis. Clone pMrΔ-39 was constructed by inserting a 687-bp *Sau*3A fragment (−39 to +648) from clone pMrSalB (7) into the *Bam*HI site of pBR322. The clone pMrΔ-12 was obtained by inserting a 743-bp *Eco*RII fragment from pMrSP, which extends from position −12 of the rDNA insert to the pBR322 *Eco*RII site at position +2500 (8) into the *Sma* I site of pUC9 by blunt-end ligation.

For the construction of the hybrid plasmid pMrSPΔ39, the 687-bp *Sau*3A rDNA insert from pMrSalB was isolated and inserted into the *Bam*HI site of pMrSP. Plasmids containing the *Sau*3A fragment in either orientation were cleaved with *Hin*fI before being assayed in the cell-free transcription system.

***In Vitro* Transcription Assays.** The preparation of S-100 extracts, the cell-free transcription system, and the analysis of the transcribed RNA have been described (6, 7). Because the S-100 extracts contained enough endogenous RNA polymerase I activity, no exogenous enzyme was added in these experiments.

RESULTS

Template Activity of Wild-Type rDNA and Deletion Mutants. The nondeleted rDNA clone used in the present study is the recombinant plasmid pMrSP (Fig. 1*A*). It contains a 461-bp mouse rDNA fragment that extends from a *Sal* I site within the NTS at position −169 to a *Pvu* II site located 292 bp downstream from the initiation site. The nucleotide sequence of this region and the precise localization of the start site of RNA polymerase I on the cloned rDNA fragments have been determined (6, 9, 10). In the presence of S-100 extracts from exponentially growing mouse cells, faithful initiation of transcription is ob-

Abbreviations: bp, base pair(s); rDNA, ribosomal RNA gene; NTS, nontranscribed spacer; ETS, external transcribed spacer.

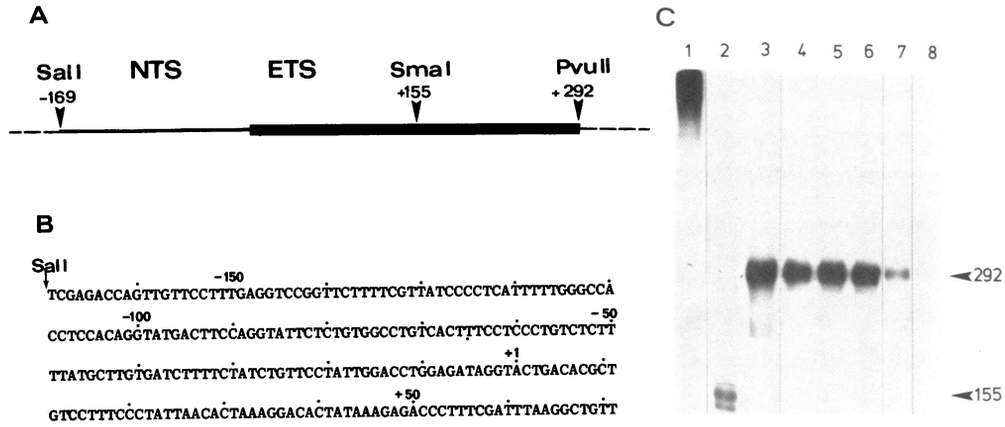

FIG. 1. Structure of the 5′ region of mouse rDNA. (*A*) rDNA sequences in pMrSP. ■, transcribed region (external transcribed spacer, ETS; +1 to +292); ——, NTS (−1 to −169); – – –, vector pBR322 DNA. (*B*) DNA sequence of the 5′ region of mouse rDNA. The position +1 corresponds to the initiation point of RNA polymerase I. (*C*) Transcription of wild-type and 5′ deletion-mutant DNAs. The autoradiogram shows the gel analysis of α-amanitin-resistant RNAs synthesized *in vitro* in the presence of S-100 extracts from Ehrlich ascites cells. DNA templates (20 μg/ml) were supercoiled pMrSP (lane 1), pMrSP truncated with *Sma* I (pMrSP/*Sma* I) (lane 2), pMrSP/*Pvu* II (lane 3), pMrΔ-86/*Pvu* II (lane 4), pMrΔ-45/*Pvu* II (lane 5), pMrΔ-39/*Pvu* II (lane 6), pMrΔ-34/*Pvu* II (lane 7), and pMrΔ-12/*Pvu* II (lane 8). The arrowheads indicate the position of the +292 and +155 run-off transcript synthesized after truncation of the template with *Pvu* II or *Sma* I, respectively.

served (6, 11, 12). After truncation of pMrSP DNA at the *Sma* I (+155) or *Pvu* II (+292) site, α-amanitin-resistant run-off transcripts 155 or 292 nucleotides long were synthesized (Fig. 1*C*, lanes 2 and 3). In order to test whether sequences flanking the 5′ end of the rDNA are important for the initiation of specific RNA *in vitro*, a series of deletion mutants was constructed. Each mutant was truncated within the NTS in a 5′ to 3′ direction. All of the mutant genes except the plasmid pMrΔ-39 were identical to the undeleted rDNA pMrSP in the transcribed region (+1 to +292) but contained variable lengths of the NTS depending on the extent of deletion. The precise end points in the 5′-flanking region of the mutants were determined by DNA sequence analysis.

Fig. 1*C* also shows the template activities of the mutant clones pMrΔ-86, pMrΔ-45, pMrΔ-39, and pMrΔ-34 (numbers indicate the sites of deletion ends) in the cell-free transcription system. All templates produced the 292-nucleotide run-off transcript after truncation with *Pvu* II (Fig. 1*C*, lanes 4–8). However, major differences were observed in the amount of transcript synthesized. Templates that retained nucleotides up to position −39 synthesized the specific transcript in the same amounts as did the undeleted DNA (Fig. 1*C*, lanes 3–6), whereas further deletion of NTS sequences up to position −34 resulted in a marked reduction of the template activity by a factor of about 10 (Fig. 1*C*, lane 7). Analysis of the 5′ end of the transcripts by S1 nuclease mapping (13) did not show any differences in the initiation specificity of transcripts synthesized from deleted and nondeleted genes (data not shown). These results suggest that nucleotides between positions −39 and −34 of the NTS are essential to the efficiency of rDNA transcription but do not affect the specificity of initiation. Clone pMrΔ-12, in which all but 12 bp of the NTS immediately adjacent to the initiation site have been deleted and replaced by pBR322 DNA sequences, never showed detectable transcription (Fig. 1*C*, lane 8). This indicates that sequences between 12 and 34 nucleotides upstream from the initiation site are essential for transcription initiation *in vitro*.

Competitive Effect of Upstream Sequences on the Transcription of Deleted Genes. Experiments were carried out in which equimolar amounts of pMrSP and deleted DNAs were assayed simultaneously in the reaction mixture. Because the two templates could be differently truncated (with *Pvu* II or *Sma* I, respectively), it was possible to detect the transcription products from each rDNA template and to compare the efficiency of the initiation of transcription on each DNA by quantitating the radioactivities of the 292- and 155-nucleotide RNA on the gel.

Separate incubations of *Pvu* II-truncated pMrSP, pMrΔ-45, and pMrΔ-39 showed almost identical template activity (Fig. 2, lanes 1–3). A mixture of equimolar amounts of pMrSP truncated with either *Pvu* II or *Sma* I resulted in approximately equal transcription from both templates (Fig. 2, lane 5). If, however, equimolar amounts of pMrSP wild-type DNA cut with *Sma* I were added to reactions containing mutant DNAs truncated with *Pvu* II, the transcription from the mutant plasmids was severely impaired (Fig. 2, lanes 6–8). The decrease in initiation efficiency was more pronounced as more 5′ sequences were deleted. This result suggests that upstream regions might have a competitive effect on the efficiency of initiation.

In order to demonstrate the preferential transcription of the nondeleted DNA in the presence of the deleted clones, the

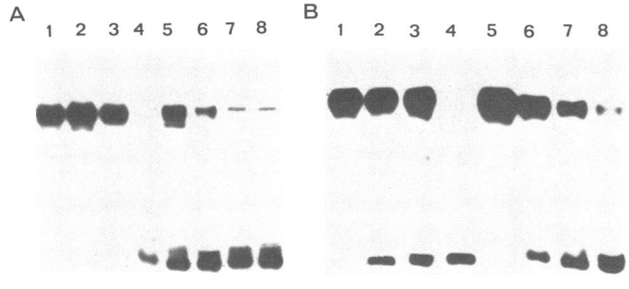

FIG. 2. (*A*) Mixed incubations of deleted and nondeleted genes. The autoradiogram shows the run-off transcripts synthesized in the S-100 system from pMrSP and deleted clones in the absence and presence of approximately equimolar ratios of nondeleted DNA. Each template was present at 0.75 μg/50 μl. Lanes: 1, pMrSP truncated with *Pvu* II (pMrSP/*Pvu* II); 2, pMrΔ-45/*Pvu* II; 3, pMrΔ-39/*Pvu* II; 4, pMrSP/*Sma* I; 5, pMrSP/*Pvu* II plus pMrSP/*Sma* I; 6, pMrΔ-86/*Pvu* II plus pMrSP/*Sma* I; 7, pMrΔ-45/*Pvu* II plus pMrSP/*Sma* I; and 8, pMrΔ-39/*Pvu* II plus pMrSP/*Sma* I. (*B*) Transcription of wild-type or pMrΔ-45 DNA in the presence of increasing amounts of pMrSP DNA. Each assay contained 0.4 μg of pMrSP/*Pvu* II DNA (lanes 1–3) or pMrΔ-45/*Pvu* II DNA (lanes 5–8) per 50 μl, as well as increasing amounts of competitor DNA pMrSP/*Sma* I. Lanes: 1 and 5, no competitor DNA; 6, 0.1 μg of pMrSP/*Sma* I; 2, 4, and 7, 0.25 μg of pMrSP/*Sma* I; and 3 and 8, 0.4 μg of pMrSP/*Sma* I.

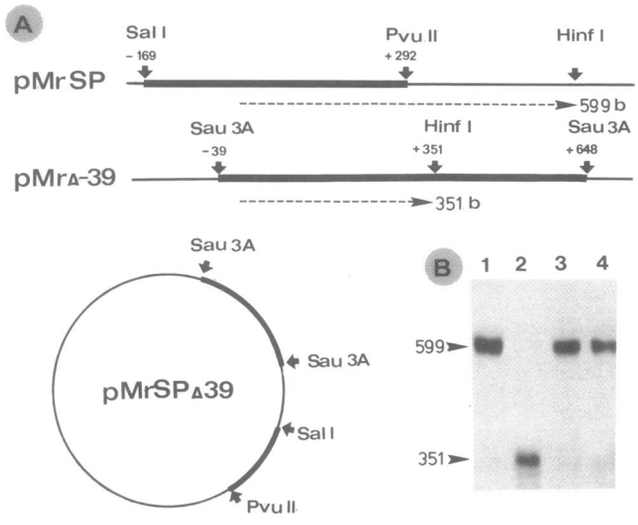

FIG. 3. Transcription of a recombinant plasmid that contains both the rDNA insert of the nondeleted gene (pMrSP) and the deleted clone pMrΔ-39. (*A*) *Hin*fI restriction sites in pMrSP and pMrΔ-39 and structure of the hybrid clone pMrΔ-39. ■, rDNA sequences; ——, pBR322 sequences; – – –, run-off transcripts at the *Hin*fI site. (*B*) Run-off transcripts synthesized in the S-100 system from *Hin*fI-truncated templates (template/*Hin*fI). Lanes: 1, pMrSP/*Hin*fI; 2, pMrΔ-39/*Hin*fI; 3, pMrSPΔ-39/*Hin*fI; and 4, same as lane 3, but the *Sau*3A fragment was inserted in opposite orientation. The arrowheads indicate the 599- and 351-nucleotide run-off RNA from pMrSP and pMrΔ-39, respectively.

competition experiments were repeated with different molar ratios of templates. Fig. 2*B* compares the effect of increasing concentrations of *Sma* I-truncated pMrSP on the template activity of either clone pMrΔ-45 DNA or pMrSP DNA, each cut with *Pvu* II. The experiment reveals an approximate 50% reduction of pMrΔ-45 transcription at an approximate 4:1 molar ratio of deleted to nondeleted DNA (Fig. 2*B*, lane 6). A further decrease of the template activity of the mutant DNA occurred at higher concentrations of *Sma* I-truncated pMrSP. At an equimolar ratio of mutant to pMrSP DNA, a 95% reduction of pMrΔ-45 DNA transcription was observed (lane 9). At the same ratio of the pMrSP/*Sma* I to pMrSP/*Pvu* II systems, transcription of the rDNA templates was not impaired (Fig. 2*B*, lanes 2 and 3).

In a series of experiments, it was found that the degree of template selection was variable from one extract to another depending on the activity of the S-100 extracts. If the extracts used were transcriptionally extremely active, higher concentrations of competitor DNA or lower amounts of S-100 proteins were required to obtain the same degree of competition. One explanation for this finding is that essential transcriptional components, the amount or activity of which varies from extract to extract, may interact with upstream regions in the nontranscribed spacer.

Because of the variability of the extracts and the difficulty of adequately controlling the effect of different DNA concentrations in the assay, a recombinant plasmid was constructed that contains the rDNA insert of both nondeleted pMrSP DNA and deleted pMrΔ-39 DNA (Fig. 3*A*). The use of this hybrid plasmid pMrSPΔ39 as template in the cell-free transcription system should allow a more accurate quantification of the transcription of wild-type versus mutant gene. A prerequisite of this approach is that the transcription products from each rDNA insert can be distinguished clearly. This can be achieved by restriction of pMrSPΔ39 DNA with *Hin*fI. Truncation of pMrSP with *Hin*fI will yield a 599-nucleotide RNA, whereas cleavage of pMrΔ-39 with *Hin*fI will produce a 351-nucleotide transcript (Fig. 3*A*). When the plasmids pMrSP and pMrΔ-39 were assayed separately in the cell-free system, both transcripts were synthesized in equal amounts (Fig. 3*B*, lanes 1 and 2). When the hybrid plasmid pMrSPΔ39 served as template, a strong preference for the transcription of the nondeleted rDNA insert of pMrSP was observed as compared to that of the deleted insert (Fig. 3*B*, lanes 3 and 4). The strong inhibition of transcription of the deleted gene in the presence of the wild-type gene is in accord with the competition experiments described above and suggests that sequences upstream from position −39 influence the relative transcription efficiency of rDNA in a competitive situation.

In order to define the position of the essential region within the 5′-flanking sequences more precisely, three overlapping regions of the NTS were cloned and used in competition experiments similar to those described in Fig. 2. The three clones contained sequences from −12 to −90, from −90 to −169, and from −39 to −169 of the NTS region. When assayed in the cell-free system, none of these NTS clones inhibited the transcription of the deleted rDNA template of pMrΔ-39 (Fig. 4, lanes 3–5), whereas the same amount of pMrSP DNA reduced transcription of the mutant DNA about 90% (lane 2). This suggests that the upstream regions have to be covalently joined to the gene to exert an effect on transcription, that is, they function in *cis*.

DISCUSSION

This paper describes the use of a cell-free transcription system derived from cultured Ehrlich ascites cells to determine the

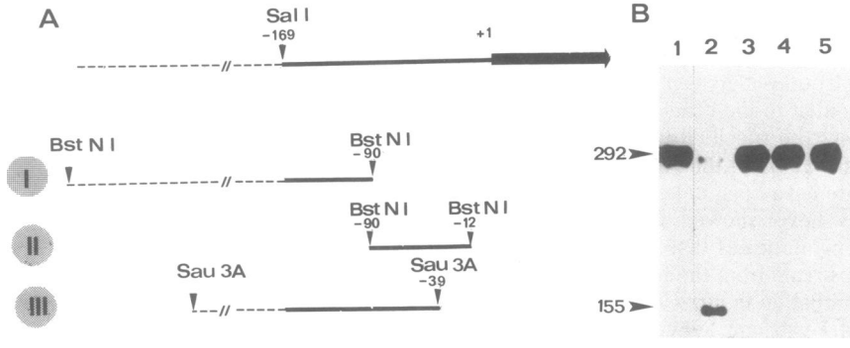

FIG. 4. Influence of cloned NTS DNA on the transcription of pMrΔ-39. (*A*) Diagram of the NTS clones pMrNTS90/169 (line I), pMrNTS12/90 (line II), and pMrNTS39/169 (line III). ■, part of the transcribed region; ——, NTS; – – –, pBR322 DNA. The clones were constructed by cleavage of pMrSP with either *Bst*NI or *Sau*3A, isolation of the desired fragments, and cloning into pUC9 (provided by J. Messing). (*B*) Transcription of pMrΔ-39 DNA in the presence of competitor DNA. Each transcription assay contained 0.4 μg of pMrΔ-39 DNA cut with *Pvu* II and 0.4 μg of competitor DNA: pUC9 (lane 1), *Sma* I-truncated pMrSP (lane 2), pMrNTS90/169 (lane 3), pMrNTS12/90 (lane 4), and pMrNTS39/169 (lane 5).

DNA sequences required for the specific initiation of RNA polymerase I on cloned mouse rDNA *in vitro*. For this, a series of deletion mutants was constructed that contain 292 bp of the transcribed region of the gene along with various lengths of contiguous 5′ flanking DNA. Mutant templates, in which vector DNA has replaced rDNA sequences to within 39 bp upstream from the initiation site, functioned like nondeleted DNA *in vitro*. Deletion of five more nucleotides to position −34 resulted in an approximate 90% reduction of initiation. The further removal of 5′ sequences to position −12 eliminated transcription completely. These results suggest that at least part of the sequence between nucleotides −12 and −34 is essential for transcription and implies that sequences required for faithful initiation are located 5′ proximal to the start site of the ribosomal transcription unit.

The localization of a transcriptional control region at this position of the rDNA resembles the position of the Goldberg–Hogness "T-A-T-A" consensus sequence required for the *in vitro* initiation of transcription by RNA polymerase II. It has been shown by several groups that the T-A-T-A box helps to determine the exact start site *in vivo* and is indispensable *in vitro*, whereas sequences further upstream, the C-C-A-A-T box at position −80 and more remote regions, determine the efficiency of transcription *in vivo* (5). Analogous experiments about the sequences required for the expression of RNA polymerase I genes *in vivo* and *in vitro* have not yet been published. The data presented in this paper indicate that the location of promoter elements relative to the initiation site may be similar for class I and II RNA polymerase genes. This is surprising because comparison of the sequences in the region of transcription initiation of ribosomal genes from yeast, *Tetrahymena pyriformis*, *Drosophila melanogaster*, three *Xenopus* species, and mouse did not reveal a T-A-T-A box-like sequence or any other canonical sequence common to classes I and II RNA polymerase genes (6, 9, 10, 14–18). This lack of homologous sequences in the rRNA genes is in accord with our finding that species-specific factors are probably involved in the transcription of rDNA by RNA polymerase I (19). The amount or the activity of the factor(s) that controls both the selectivity and the efficiency of the initiation reaction of RNA polymerase I has been shown to correlate with the growth rate of the cells (7). The mode of action of this RNA polymerase I transcription factor(s) is not yet understood, but possibly it may interact with the control region located directly in front of the transcribed region. However, it is equally possible that the factor or another protein binds to regions further upstream and functions by guiding the RNA polymerase I to the initiation site in a manner somewhat analogous to the σ factor of *Escherichia coli* RNA polymerase.

Indeed, the importance of upstream DNA sequences for binding of proteins required for accurate transcription in the cell-free system can be revealed by competition experiments. The utilization of the deleted template was much less efficient in the presence of the nondeleted DNA if both templates were assayed simultaneously. The competitive effect was more pronounced when more 5′ flanking sequences had been deleted. This suggests that either the transcription factor(s), the RNA polymerase I, or both have a higher affinity for rDNA sequences than for plasmid DNA, which replaces the rDNA in the deleted clones. The mechanisms by which upstream sequences modulate transcription is not yet clear. However, an analogous effect of upstream sequences on gene expression *in vitro* has been described for both the transcription of the fibroin gene in silk gland extracts (20) and the silkworm alanine tRNA gene (21). In both cases, the effect of 5′ flanking sequences has been observed only in homologous cell-free extracts. This finding has been interpreted as a competitive capturing of specific transcription factors by defined 5′ flanking sequences.

The results in this paper are compatible with this interpretation. I suggest that more than one essential protein of the transcription machinery has to interact with defined regions of the rDNA. Such a binding of several proteins to the DNA might occur cooperatively. This working hypothesis could explain why DNA of defined subclones of the NTS were unable to compete with the deleted template pMrΔ-39.

I thank Anke Houpert for determining the sequence of the end points of the deletion mutants, Annette Barthel for technical assistance, and Stuart Siddell for help in the preparation of the manuscript. This work has been supported by the Deutsche Forschungsgemeinschaft.

1. Soeiro, R., Vaughan, M. H. & Darnell, J. E. (1968) *J. Cell Biol.* **36**, 91–101.
2. Yu, F. L. & Feigelson, P. C. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 2833–2837.
3. Gross, K. J. & Pogo, A. O. (1976) *Biochemistry* **15**, 2082–2086.
4. Grummt, I., Smith, V. A. & Grummt, F. (1976) *Cell* **7**, 439–445.
5. Shenk, T. (1981) *Curr. Top. Microbiol. Immunol.* **93**, 25–46.
6. Grummt, I. (1981) *Nucleic Acids Res.* **9**, 6093–6102.
7. Grummt, I. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 727–731.
8. Sutcliffe, J. G. (1978) *Nucleic Acids Res.* **58**, 2721–2728.
9. Bach, R., Grummt, I. & Allet, B. (1981) *Nucleic Acids Res.* **9**, 1559–1569.
10. Urano, Y., Kominami, R., Mishima, Y. & Muramatsu, M. (1980) *Nucleic Acids Res.* **8**, 6043–6058.
11. Miller, K. G. & Sollner-Webb, B. (1981) *Cell* **27**, 165–174.
12. Mishima, Y., Yamamoto, O., Kominami, R. & Muramatsu, M. (1981) *Nucleic Acids Res.* **9**, 6773–6785.
13. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12**, 721–732.
14. Bayev, A. A., Georgiev, O. I., Hadjiolov, A. A., Kermekchiev, M. B., Nikolaev, N., Skryabin, K. G. & Zakharyev, V. M. (1980) *Nucleic Acids Res.* **8**, 4919–4926.
15. Niles, E. G., Sutiphong, J. & Haque, S. (1981) *J. Biol. Chem.* **256**, 12849–12856.
16. Long, E., Rebbert, M. L. & Dawid, I. B. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1513–1517.
17. Sollner-Webb, B. & Reeder, R. H. (1979) *Cell* **18**, 485–499.
18. Bach, R., Allet, B. & Crippa, M. (1981) *Nucleic Acids Res.* **9**, 5311–5330.
19. Grummt, I., Roth, E. & Paule, M. R. (1982) *Nature (London)* **296**, 173–174.
20. Tsuda, M. & Suzuki, Y. (1981) *Cell* **27**, 175–182.
21. Sprague, K. U., Larson, D. & Morton, D. (1980) *Cell* **22**, 171–178.