Volume 9 Number 22 1981 Nucleic Acids Research

Mapping of a mouse ribosomal DNA promoter by in vitro transcription

Ingrid Grummt

Institut für Biochemie der Universität Würzburg, Röntgenring 11, D-8700 Würzburg, GFR

Received 7 September 1981

ABSTRACT

An in vitro transcription system that provides proper initiation of RNA polymerase I on cloned rDNA has been used to identify the start site for rDNA transcription. Different subclones that span defined regions of the 5' terminal region of the ribosomal gene have been constructed and assayed in the cell-free system for their ability to promote specific initiation of pre-rRNA synthesis. It is shown that rapid processing at the 5' end of the primary transcript occurs both in vivo and in vitro which in former studies has led to a wrong interpretation of the S1 nuclease mapping data (1 - 3). RNA polymerase I starts in vitro at a unique point on the rDNA yielding run-off transcripts that have a triphosphorylated 5' end pppApC. If multiple copies of the promoter-containing rDNA fragment were placed in head-to-tail orientation in front of the transcribed region distinct RNA products were synthesized that have been started at the tandem initiation sites. Removal of sequences upstream the initiation site indicates that 5' flanking regions are essential for specific transcription.

INTRODUCTION

Ribosomal genes provide an attractive system to study gene expression in eukaryotes since their rate of transcription is regulated in response to the proliferation rate of the cells (4 - 8). In order to investigate the molecular mechanisms involved in the control of rRNA synthesis we have established a cell-free system for the specific transcription of ribosomal genes from mouse (2). It has been shown that the in vitro system which contains crude cytoplasmic (S-100) extracts and cloned rDNA mimics the transcriptional activity of the cells from which the S-100 extracts have been prepared. Thus in vitro studies on the transcription of ribosomal DNA may reveal both regulatory proteins involved in the modulation of the initiation frequency of RNA polymerase I and signal sequences required for

the specific read-out of rDNA. For the identification of DNA sequences which may be necessary for promotion of rRNA synthesis the precise site of transcription initiation must be known.

In a previous paper we have reported the sequence of 1100 nucleotides surrounding the initiation site of pre-rRNA transcription. The 5' end of 45S pre-rRNA has been located on the mouse rDNA fragment by primer extension and the S1 nuclease protection technique (1). However, subsequent experiments revealed some discrepancies between the length of *in vitro* transcripts and the S1 mapping data. It turned out that there is a rapid processing at the 5' end of 45S pre-rRNA which takes place both *in vivo* and *in vitro*. Thus the mapping of the 5' end has revealed a processing site instead of the initiation site of transcription. Probably the same holds true for the initiation site which has been mapped by Urano et al. (3) 200 bp further downstream. The processing site we have identified is characterized by a long stretch of thymidines in the non-coding strand of the DNA (1).

In this communication the use of the cell-free transcription system for the identification of the initiation site of RNA polymerase I on mouse rDNA is described. If the rDNA template was truncated upstream the processing site transcripts bearing a triphosphate end were generated which start at a unique point 448 bp upstream the processing site.

MATERIALS AND METHODS

Construction of recombinant plasmids

The plasmid pMrSal B used for transcription studies represents a Sal I fragment of mouse rDNA cloned in pBR322 (2, 9). The pMrSal B subclones studied were constructed as follows: pMrSP has been constructed by cleaving pMrSal B with Pvu II and circularization of the 3.4 kb fragment containing the replication origin, the ampicillin resistance gene and 460 bp of mouse rDNA. For transcription the DNA was truncated with Pvu II. pMrTaq I contains a 940 bp Taq I fragment inserted into the Cla I site of pBR322. The DNA was linearized with Hind III or Eco RI, respectively, depending on the orientation of the insert. The clone pMrPX was constructed by inserting the 970 bp Pvu II-

Xho I fragment into the BamH I site of pBR322 via synthetic linkers. For transcription the recombinant DNA has been cleaved with BamH I. For the construction of the plasmids containing multiple rDNA-promoter sites a 231 bp Hae III-Sma I fragment was isolated and inserted into the Sma I site of pMrSal B. Clones containing one or two inserted fragments in head-to-tail orientation were selected and used as templates for transcription studies.

Cell-free transcription system

The S-100 extract was derived from cultured Ehrlich ascites cells (2) and was prepared as described by Weil et al. (10). The final volume of the standard RNA synthesis reaction mixture was 50 µl, half of the volume being contributed by the S-100 extract. The concentration of components in the assay were: 10 mM HEPES (pH 7.9), 75 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM dithiothreitol, 10% glycerol, 10 mM creatine phosphate, 600 µM each of ATP, CTP, UTP, 25 µM GTP, 3 µCi of α-(^{32}P) GTP (spec. act. 10 - 25 Ci/mmol) and 10 - 20 µg/ml DNA. After incubation for 1 hr at 26^{o}C the samples were processed as described before (2) and analyzed on 4% acrylamide gels. For size determination the transcripts and ^{32}P-labelled DNA size markers were denatured by incubation for 2 hr at 37^{o}C in 1 M glyoxal, 50% (vol/vol) dimethyl sulfoxide, and 250 mM HEPES (pH 7.9) according to Mc Master and Carmichael (11).

S1 nuclease protection experiments

^{32}P-labelled nucleolar RNA was prepared from Ehrlich ascites cells. For this 3 x 10^{8} cells were incubated in 100 ml phosphate-free medium with 2 mCi ^{32}P-orthophosphate for 60 min. The nucleoli were isolated and the RNA prepared as described previously (8). For hybridization 50 000 cpm nucleolar RNA or 1000 cpm of the in vitro transcript, respectively, were hybridized to 1 µg denatured 225 bp Taq I fragment (-169 to +56) for 3 hr at 55^{o}C (2, 12). After treatment with nuclease S1 (100 U/200 µl assay) the hybrids were run on a 8 % non-denaturing acrylamide gel together with size markers.

Analysis of the 5' terminal nucleotides of the in vitro transcripts

For 5' end analysis of the transcripts the standard trans-

cription assay was scaled up 3-fold and 150 µCi $\alpha(^{32}P)$ATP or CTP, respectively, were added. After incubation for 75 min at 26° the samples were processed as usual and run on a 4 % acrylamide gel. The run-off transcripts were excised from the gel, eluted and digested in 10 µl 10 mM NH_4-acetate (pH 4.5) with 2.5 units of ribonuclease T_2 for 90 min at 37°C. Digests were spotted on thin layer plates together with pppAp and mononucleotide markers. Polyethyleneimine coated cellulose plates were chromatographed in 0.8 M LiCl, 0.8 M acetic acid. The solvent for normal cellulose plates was saturated ammonium sulfate: 1 M Na-acetate: iso-propanol (80:18:2).

RESULTS AND DISCUSSION

Previously we have reported the nucleotide sequence of 1100 nucleotides of a Sal I-fragment of mouse rDNA (pMrSal B) that contains the boundary between the non-transcribed spacer and the transcribed region (1). Nuclease S1 mapping located the 5' end of 45S pre-rRNA on this DNA 810 bp upstream the Xho I site. *In vitro* transcription of this Xho I-truncated Sal I fragment (Fig. 1A) yielded a 810 nucleotides long run-off transcript (2). This coincident position of the 5' end of cellular 45S pre-rRNA and the *in vitro* transcript suggested that RNA polymerase I initiates transcription 618 bp downstream the Sal I site. However, we frequently observed in addition to the 810 bases RNA a larger transcript approximately 1300 bases long in the *in vitro* assay. If the template was truncated with Pvu II a short transcript 290-300 nucleotides long was produced (Fig. 1A). This result suggests that there are either two start sites for RNA polymerase I on the rDNA template or that the 810 bases transcript represents a processing product which is found both *in vivo* and *in vitro*.

To decide between these two possibilities I have constructed subclones of pMrSal B that contain different regions of the Sal B fragment between the Sal I and the Xho I restriction site and used them as a template in the cell-free transcription system. Only the clone pMrSP that contains 460 bp from the left Sal I to the Pvu II site promoted α-amanitin resistant transcription yielding a 290 - 300 nucleotides long run-off RNA

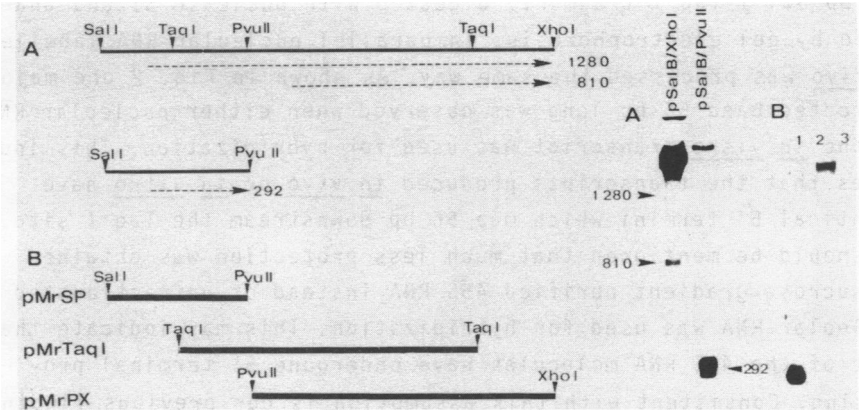

Fig. 1. rDNA transcripts synthesized in the S-100 system.
A) Length of run-off transcripts produced after truncation of pMrSal B with Xho I and Pvu II, respectively. 0.75 µg of DNA was transcribed in the presence of S-100 extracts from cultured Ehrlich ascites cells as described previously (2).
B) Transcription of cloned fragments derived from the 5' region of the rDNA. 1) pMrSP truncated with Pvu II; 2) pMrTaq I linearized with Hind III; 3) pMrPX cleaved with BamH I. The construction of the subclones is described in Materials and Methods.

after cleavage with Pvu II. The clones pMrTaq I and pMrPX which contain sequences downstream the Taq I site did not support specific transcription (Fig. 1B). The same result was obtained whether the DNA fragments alone or fused to pBR322 sequences were assayed in the cell-free system. These in vitro experiments indicate that a strong initiation site for RNA polymerase I is located 290 - 300 bp upstream the Pvu II site. If there is a second promoter about 450 bp further downstream its functioning in the cell-free system requires sequences upstream the Taq I site in the DNA. It therefore appears that previous experiments which have been performed to map the 5' end of pre-rRNA (1 - 3) have revealed 5' terminal processing sites instead of true initiation points.

A more precise localization of the 5' end of the in vitro transcripts was obtained by the S1 nuclease mapping procedure. The labelled pMrSalB/Pvu II run-off RNA was hybridized to the

225 bp Sal I-Taq I fragment, digested with nuclease S1 and analyzed by gel electrophoresis. In parallel nucleolar RNA labelled *in vivo* was processed the same way. As shown in Fig. 2 one major protected band 56 bp long was observed when either nucleolar RNA or the *in vitro* transcript was used for hybridization. This indicates that the transcripts produced *in vivo* or *in vitro* have identical 5' termini which map 56 bp downstream the Taq I site. It should be mentioned that much less protection was obtained if sucrose-gradient purified 45S RNA instead of unfractionated nucleolar RNA was used for hybridization. This may indicate that most of the 45S RNA molecules have undergone 5' terminal processing. Consistent with this assumption is our previous finding that only 10-15 % of the 45S RNA molecules could be capped *in vitro* (1) indicating a very low amount of di- or triphosphate residues on the 5' end of 45S RNA.

In order to make sure that the 5' end of the transcripts

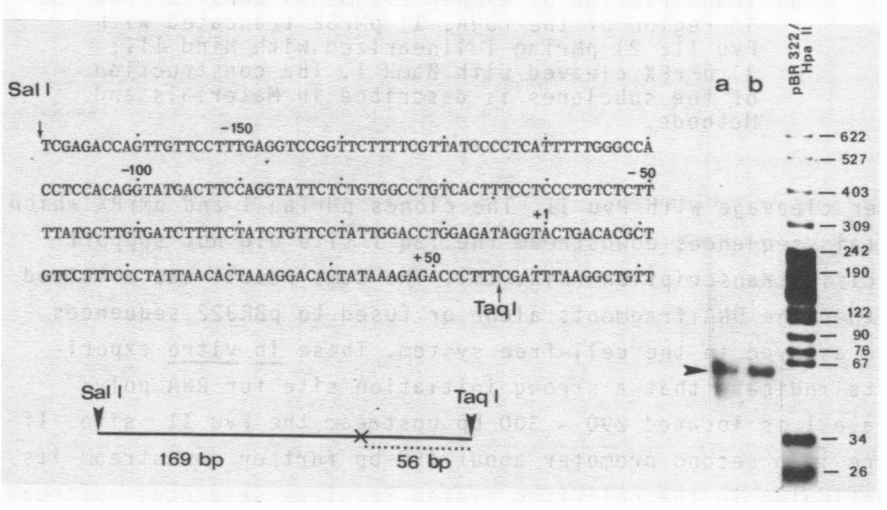

Fig. 2. Nucleotide sequence of the initiation region of mouse rDNA and S1 nuclease mapping of 5' terminal sequences of pre-rRNA labelled either *in vivo* or *in vitro*. ^{32}P-labelled nucleolar RNA prepared from Ehrlich ascites cells (a) and the 292 bases run-off transcript from pMrSal B-Pvu II; (b), respectively, were hybridized to 1 µg denatured 225 bp Taq I fragment (-169 to +56) as described in Materials and Methods. The S1 nuclease resistant hybrids were analyzed on a 8% non-denaturing acrylamide gel.

produced in vitro corresponds to the site of transcription initiation I have constructed recombinant plasmids that contain one or two additional Hae III-Sma I fragments in head-to-tail orientation and used these DNAs as templates in the cell-free transcription system (Fig. 3). After truncation of the plasmids with Pvu II one, two or three distinct RNA bands, respectively, were produced the lengths of which correspond to the distance between the Pvu II site and the first, second or third initiation site. Obviously RNA polymerase I has started at the tandem initiation sites and used the different promoters with different efficiency. The fact that tandem promoter-containing fragments yielded distinct transcripts of defined length clearly rules out the possibility that the 5' termini of the RNA were generated by a rapid processing event.

A direct proof that the in vitro product represents a primary transcript would be the demonstration of a conserved triphosphate residue on the 5' terminus. For this the 292 bases long pMrSalB/Pvu II run-off transcript was digested to comple-

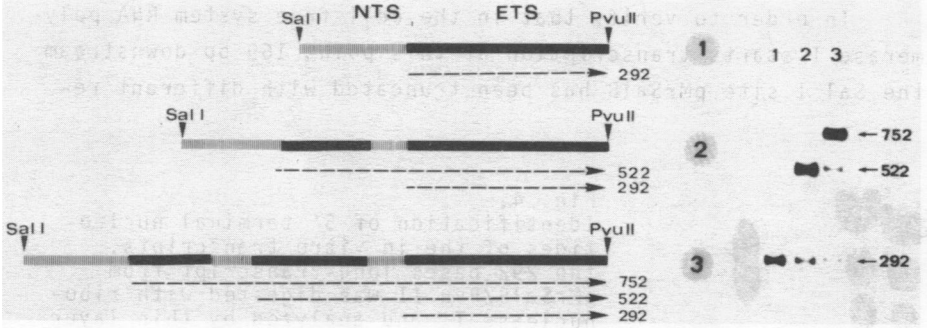

Fig. 3. Construction and transcription of rDNA containing tandemly repeated promoter fragments in head-to-tail orientation.
The recombinant plasmids with one or two additional promoters were constructed by inserting the 231 bp Hae III/Sma I fragment (-74 to +155) into the Sma I site of pMrSal B (I). The plasmids were truncated with Pvu II and used as templates in the S-100 system. pMrSalB without (1), with one (2) or two (3) additional Hae III/Sma I fragments. The solid bars indicate the transcribed region of the rDNA clone the hatched bar represents non-transcribed spacer sequences. The arrows indicate the run-off transcripts originating from the individual promoters.

tion with ribonuclease T_2 and analyzed by thin layer chromatography (Fig. 4). In addition to the four mononucleotides a nucleoside tetraphosphate could be identified which comigrated with a pppAp marker in three different solvent systems. This tetraphosphate spot was found after labelling the RNA with [α-^{32}P]ATP or [α-^{32}P]CTP, respectively, but not with [α-^{32}P]GTP or UTP. If the pppAp spot was labelled with [α-^{32}P]ATP and was treated with nuclease P1 the tetraphosphate was converted to [α-^{32}P]ATP. If the nucleoside tetraphosphate was labelled with [α-^{32}P]CTP nuclease P1 digestion released radioactive phosphate (not shown). This indicates that the ribosomal transcription unit starts with pppApC. If run-off transcripts of different lengths (see Fig. 5) were analyzed the same 5' terminal sequence pppApC was found (unpublished result). The initiation sequence ACXYZ is compatible with the position of the 5' end of the RNA as revealed by S1 mapping. This method has located the initiation site 169 bp downstream the Sal I site. At this place an adenosine followed by a cytosine is found in the non-coding DNA strand.

In order to verify that in the cell-free system RNA polymerase I starts transcription at this point 169 bp downstream the Sal I site pMrSalB has been truncated with different re-

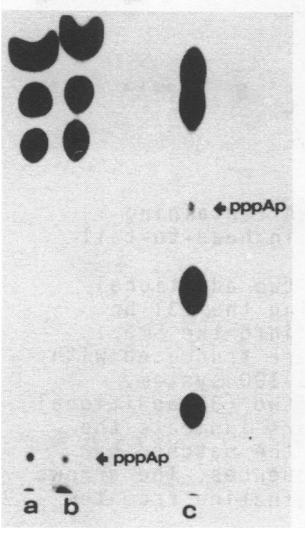

Fig. 4.
Identification of 5' terminal nucleotides of the in vitro transcripts. the 292 bases long transcript from pMrSalB/Pvu II was digested with ribonuclease T_2 and analyzed by thin layer chromatography as described in Materials and Methods. The RNA was labelled in vitro with either α-(^{32}P)ATP (a and c) or α-(^{32}P-)CTP (b), respectively. Chromatography was on either polyethyleneimine coated thin layer plates (a and b) or cellulose plates (c).

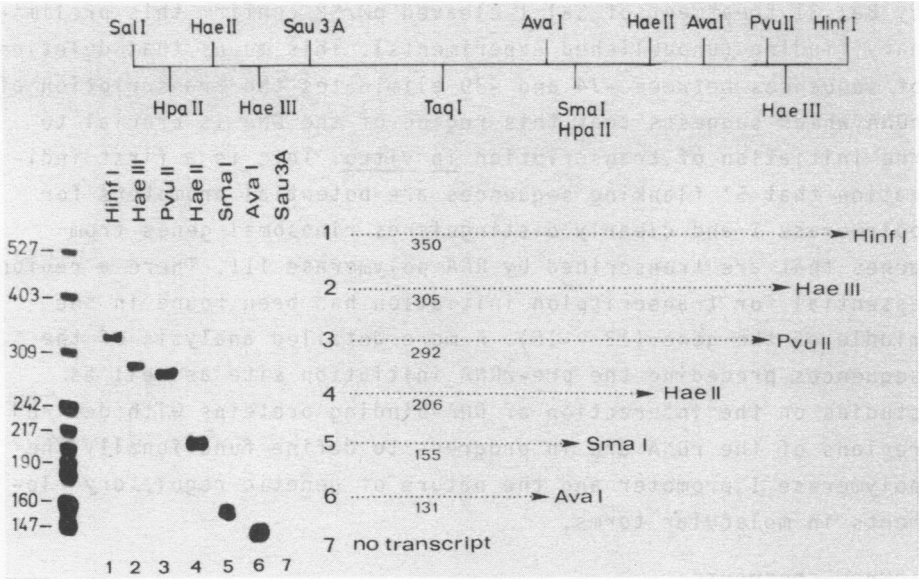

Fig. 5. Electrophoretic analysis of glyoxal-denatured RNAs synthesized in vitro from truncated mouse rDNA templates. The autoradiogram shows a polyacrylamide gel analysis of glyoxal-denatured RNAs synthesized in vitro. The length of the transcripts were determined experimentally by comparison with 3' labelled glyoxylated pBR322/Hpa II marker fragments. The numbers indicate the distance in bases from the putative initiation site to the restriction site as derived from sequence data.

striction enzymes and was assayed in the in vitro system. As shown in Fig. 5 the transcripts decrease in length as the 3' end of the template approaches the presumptive initiation site. The length of the different run-off transcripts corresponds to the sizes expected if RNA polymerase I had initiated transcription in vitro at the same place as in vivo.

Another important point comes up by this experiment. Removing 5' flanking sequences up to 74 nucleotides from the start point of RNA polymerase I (by cleavage with Hae III) does not affect the initiation of transcription. By contrast, specific transcription was not detectable if the template was cleaved with Sau 3A which removes upstream sequences beyond position -39. Analysis of the template activity of deletion mutants constructed

by Bal 31 treatment of Sal I cleaved pMrSP confirm this preliminary finding (unpublished experiments). This means that deletion of sequences between -74 and -39 eliminates the transcription of rDNA which suggests that this region of the DNA is crucial to the initiation of transcription in vitro. This is a first indication that 5' flanking sequences are potential promoters for polymerase I and clearly distinguishes ribosomal genes from genes that are transcribed by RNA polymerase III. There a region essential for transcritpion initiation has been found in the middle of the gene (13 - 15). A more detailed analysis of the sequences preceding the pre-rRNA initiation site as well as studies on the interaction of DNA-binding proteins with defined regions of the rDNA are in progress to define functionally the polymerase I promoter and the nature of genetic regulatory elements in molecular terms.

ACKNOWLEDGEMENTS

I am grateful to W. Schulz for the gift of the pppAp marker and to G. Krupp for advice in the 5' analysis of RNA. This work was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Bach, R., Grummt, I. and Allet, B. (1981) Nucl. Acids. Res. 9, 1559-1569.
2. Grummt, I. (1981) Proc. Natl. Acad. Sci. USA 78, 727-731.
3. Urano, Y., Kominami, R., Mishima, Y., and Muramatsu, M. (1980) Nucl. Acids Res. 8, 6043-6058.
4. Soeiro, R., Vaughan, M. H. and Darnell, J. E. (1968) J. Cell Biol. 36, 91-101.
5. Yu, F., L. and Feigelson, P. C. (1972) Proc. Natl. Acad. Sci. USA 69, 2833-2837.
6. Chesterton, C. J., Coupar, B. E. H., Butterworth, P. H. W., Bess, J. and Green, M. H. (1975) Eur. J. Biochem. 57, 79-83.
7. Gross, H.J. and Pogo, A.O. (1976) Biochemistry 15, 2082-2086.
8. Grummt, I., Smith, V. A. and Grummt, F. (1976) Cell 7, 439-445.
9. Grummt, I. and Gross, H. J. (1980) Mol. Gen. Genet. 177, 223-229.
10. Weil, P. A., Segall, J., Harris, B., Ng, S. Y. and Roeder, R. H. (1979) Cell 18, 469-484.
11. Mc Master, G. K. and Carmichael, G. C. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
12. Berk, A. and Sharp, P. (1977) Cell 12, 721-732.
13. Sakonju, S., Bogenhagen, D. F. and Brown, D. D. (1980) Cell 19, 13-25.
14. Bogenhagen, D. F., Sakonju, S. and Brown, D. D. (1980) Cell 19, 27-35.
15. Engelke, D. R., Ng, S. Y., Shastry, B. S. and Roeder, R. G. (1980) Cell 19, 712-728.