

---

**Transcription initiation site of rat ribosomal DNA**

---

Lawrence I. Rothblum, Ramachandra Reddy and Brandt Cassidy

---

Department of Pharmacology, Baylor College of Medicine, Houston, TX 77030, USA

---

Received 13 April 1982; Revised and Accepted 18 October 1982

---

**ABSTRACT**

The sequence of 1,100 nucleotides surrounding the transcription initiation site of a cloned rat ribosomal RNA gene (rDNA) has been determined. The location of the 5' terminus of 45S pre-rRNA was determined by S1 nuclease mapping, reverse transcriptase elongation and confirmed by *in vitro* capping of 45S rRNA and *in vitro* transcription. Two different plasmid subclones, from two separate genomic clones of rat rDNA, contained the identical sequence surrounding the transcription initiation site:  $-^{10}\text{GGAGATATAT } ^1\text{GCTGACACGC TGTCCTTTTG}^{+20}$ .

Relatively long, greater than 15 base pairs, regions of sequence homology were found when the sequences of the initiation regions of rat and mouse rDNA (Urano, Y., Kominami, R., Mishima, Y., and Muramatsu, M., *Nucleic Acids Res.* 8, 6043-6058, 1980) were compared. When both the rat and mouse sequences were compared to that of human rDNA (G. Wilson, personal communication) a sequence of 15 nucleotides immediately following the initiation sites were found to be preserved.

**INTRODUCTION**

The transcription initiation sites of several ribosomal RNA (rDNA) genes have been sequenced (2,6,18-21). Although Miller and Sollner-Webb (1) noted that 5 nucleotides of the initiation sites of the mouse and *Xenopus* genes were homologous, there was little homology between the remainder of the initiation regions of those two species. The negligible homology among the transcription initiation regions of the ribosomal RNA genes of yeast, *Drosophila*, *Xenopus*, and mouse (2), suggests that the nucleotide signal that directs RNA polymerase I to initiate the transcription of ribosomal DNA varies in a species specific manner (3,4). However, since the nucleotide sequence of the initiation site of only one mammalian ribosomal RNA gene has been reported (2,5) homologies may exist that have not been reported.

In this study, we have determined the nucleotide sequence surrounding the transcription initiation site of the rat ribosomal RNA gene. One difficulty in determining the initiation sites of the rat and mouse ribosomal RNA genes has been the apparent rapidity of processing of the 5' terminus of the initial

transcript (1,3,6) and the presence of long A-T regions in the DNA. These result both in a very low content of the 5' terminus of the 45S preribosomal RNA and reduce the availability of hybrids that could be used in either S1 or primer elongation experiments. By using high concentrations of either total nuclear RNA or nucleolar 45S pre-ribosomal RNA, and by varying the temperatures at which the hybrids were digested with S1 nuclease, we have been able to map the transcription initiation site of the rat ribosomal RNA gene. The putative in vivo transcription initiation site was determined by S1 nuclease mapping and reverse transcriptase elongation experiments. These experimental findings were supported by in vitro capping of 45S rRNA and in vitro transcription using truncated templates.

### MATERIALS AND METHODS

Preparation of Recombinant DNA, Southern Hybridization and DNA Sequencing. Separate plasmids, pB4-5.1 and pB7-2.0, were constructed from two genomic clones of rat ribosomal DNA,  $\lambda$ ChR-B4 and  $\lambda$ ChR-B7E12 (7), respectively, by standard methods (8). Plasmid isolation, Southern hybridization, and DNA sequencing were carried out as described previously (9). Restriction endonucleases were used as recommended by the suppliers, either Bethesda Research Laboratories or New England Biolabs.

Total Nuclear RNA 45S and 18S rRNA Isolation. Total nuclear RNA was isolated from nuclei prepared by the citric acid method (10) and 45S pre-rRNA was isolated from the nucleoli of Novikoff hepatoma ascites cells and fractionated by sucrose density gradient centrifugation as described earlier (11). 18S rRNA was isolated from 40S subunits as described by Choi and Busch (12). RNA was labeled in vitro with  $\gamma$ - $^{32}$ P-ATP and polynucleotide kinase after partial alkaline hydrolysis (13).

S1 Nuclease Protection and Reverse Transcriptase Elongation Mapping. 5' end-labeled, double-stranded DNA was mixed with either 45S RNA or total nuclear RNA, precipitated with ethanol and lyophilized. Lyophilized mixtures of DNA and RNA were resuspended in 80% formamide, 0.4 M NaCl, 0.001 M EDTA, 0.04 M Pipes (pH 6.4), heated to 65°C for five minutes and hybridized at 57°C for 16 h (14). For S1 mapping, solutions containing hybrids were diluted into 9 volumes of S1 digestion buffer, (0.25 M NaCl, 0.03 M Na Acetate, 0.002 M Zn Acetate, pH 4.6) digested with 400-1,000 units of S1 nuclease (Boehringer-Mannheim) at the temperatures indicated for 1 h, and precipitated with ethanol. For reverse transcriptase elongation, hybrids, prepared as described above, were precipitated with ethanol and treated essentially as described by Bina-

Stein et al (15). In the primer elongation experiments, the amount of RNA was adjusted so that the concentration during hybridization was between 150 and 200 mg/ml. After precipitation, the hybrids were resuspended in buffer containing 0.1 M KCl, 0.01 M MgCl<sub>2</sub>, 0.01 M  $\beta$ -mercaptoethanol, 0.05 M Tris HCl (pH 8.3), and the four deoxyribonucleotides at 0.1 mM. The hybrids were then incubated with 100 units of AMV reverse transcriptase (Life Sciences) at 30°C for 30 min, followed by alkaline hydrolysis and ethanol precipitation. The ethanol precipitates were dissolved in 80% formamide and electrophoresed on 8% polyacrylamide-7 M urea sequencing gels.

In Vitro Capping. 45S rRNA, purified through two rounds of sucrose density gradient centrifugation, was capped in vitro using vaccinia virus guanylyl transferase (Bethesda Research Laboratories) in 50  $\mu$ l at 37°C under the conditions recommended by the supplier. After 30 minutes, 25  $\mu$ g of carrier yeast tRNA was added and the reaction mixture was extracted as described (17) and chromatographed on a Sephadex G-50 column that had been equilibrated with 20 mM ammonium acetate. The void peak was lyophilized, resuspended in 20 mM sodium acetate, 0.05% SDS and precipitated with ethanol. The resulting pellet was rinsed with 70% ethanol, dried, suspended in 20 mM sodium acetate and digested with P1 ribonuclease or T1 ribonuclease followed by bacterial alkaline phosphatase. The products of the capping reaction and subsequent nuclease digestions were identified by electrophoresis on DE-81 paper in 0.9 N ammonium acetate, pH 3.5 beside standard cap structure (P-L. Biochemicals).

In Vitro Transcription. Cell free extracts were prepared from Novikoff hepatoma ascites cells five days after transplantation, as described by Weil et al (16), with the modification that cell and nuclear lysis was performed in a Dounce homogenizer equipped with a type B pestle.

In vitro transcription was carried out in 100 $\lambda$  reaction volumes containing 5-10  $\mu$ g plasmid DNA essentially as described by Miller and Sollner-Webb (1), in the presence of 300  $\mu$ g/ml  $\alpha$ -amanitin. After incubation for 30 minutes at 30°C, the transcription products were extracted twice with phenol:chloroform:isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1), ethanol precipitated, resuspended in 99% formamide, heated at 65°C for ten minutes and analyzed by electrophoresis on 4% or 8% polyacrylamide, 7 M urea gels (17). The molecular size markers were terminally labeled Hae III digests of  $\phi$ X 174.

Fingerprinting of In Vitro Transcription Products. After polyacrylamide gel electrophoresis, in vitro transcription products were visualized by autoradiography, extracted from the gel and precipitated (31). The RNA was digested

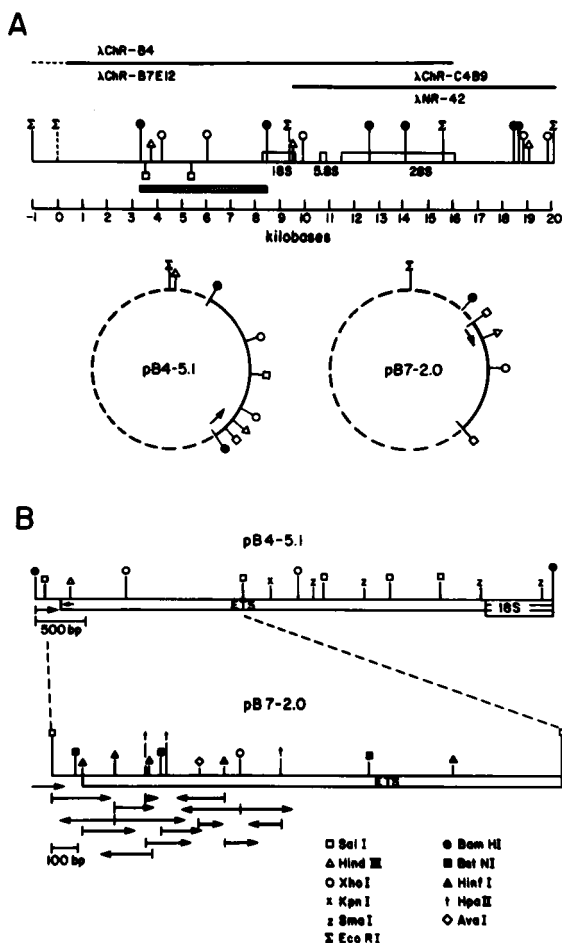
with T1 RNase and fingerprinted by the method of Brownlee et al. (32), as previously described (31).

### RESULTS

Figure 1A depicts a portion of the repeat unit of the rat rDNA and the two subclones that contain the transcription initiation site. pB4-5.1, a subclone of  $\lambda$ ChR-B4, contains a 5.1 kb Bam HI fragment which extends from 3.3 kb to 8.4 kb on the kilobase scale. pB7-2.0, a subclone of  $\lambda$ ChR-B7E12, contains a 2.0 kb Sal I fragment that extends from 3.5 to 5.5 on the kilobase scale. Figure 1B shows the strategy used to determine the nucleotide sequence of the transcription initiation region. For the sake of orientation, the Hind III site in this region is 4 kb upstream of the 5' end of the 18S, which is approximately 650 bp from the Bam HI site at the 3' end of pB4-5.1. Southern hybridization experiments (Figure 2) placed the 5' terminus of 18S rRNA near the Sma I site, and the 5' end of 45S rRNA near the Xho I site (Figure 1B).

Southern Hybridization Experiments. The 650 bp Sma I fragment of pB4-5.1 hybridized to 18S rRNA (Figure 2B, lanes 3 and 4) placing the 5' terminus of 18S rRNA within that fragment, this result has since been confirmed by DNA sequencing (data not shown). When the same sheet of nitrocellulose was incubated with total nuclear RNA, several fragments hybridized which had not hybridized to 18S rRNA. In particular, a 2.0 kb Sal I fragment (Figure 2C, lane 6) and a 1.7 kb Xho I fragment (Figure 2C, lane 5) hybridized. Both of these fragments are found within the region above the shaded bar in Figure 1A. The additional fragments that hybridized to total nuclear RNA mapped between the Xho I (+637) site within the 2.0 kb Sal I fragment and the 5' terminus of 18S rRNA. Neither the 512 bp Xho I-Hind III (Figure 2C, lane 5) fragment, nor the 400 base pairs of DNA 5' to the Hind III site (+124) (Figure 2C, lanes 2 and 5) hybridized to total nuclear RNA. These results indicated that the 5' terminus of 45S RNA was probably within the 2.0 kb Sal I fragment, cloned in pB7-2.0, and that other procedures would be required to further define it.

Nucleotide Sequence of the Transcription Initiation Site. The sequence of 1,100 nucleotides surrounding the initiation site is shown in Figure 3. The initiation region of the rat rDNA has two long stretches of thymidines in the noncoding strand, from +354 to +550, which accounted for approximately 50% of that 200 bp sequence. This is strikingly similar to the mouse transcription initiation region (2,5). In addition, long portions (>15bp) of the initiation region of rat rDNA, other than the initiation site itself, were more than 75% homologous with regions both upstream and downstream of the



**Figure 1** Restriction endonuclease map of rat rDNA. **A.** Partial restriction endonuclease maps of four genomic clones of rat rDNA, and of two plasmid subclones, pB4-5.1 and pB7-2.0, of λChr-B4 and λChr-B7E12 respectively. The shaded area indicates the region subcloned in pB7-2.0 and pB4-5.1. The arrows in the plasmid circles indicate the direction of transcription. **B.** Sequencing strategy used in determining the nucleotide sequence of the transcription initiation region. The arrows indicate the direction and distance each fragment was sequenced. The transcription initiation site, shown as the 5' end of the ETS (external transcribed spacer) in Figure 1B is at 3.7 on the kilobase scale in A. It should be noted that the relative sizes of the inserts and pBR322 are not depicted in proportion.

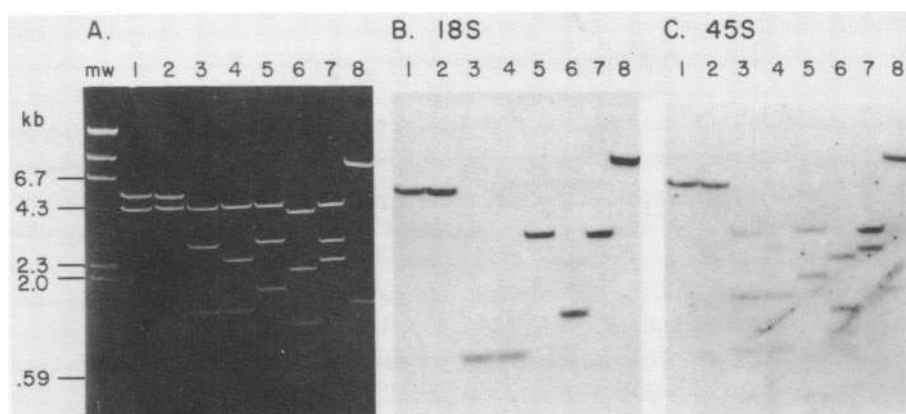


Figure 2 Southern hybridization patterns of pB4-5.1 after hybridization with (B) 18S rRNA and (C) 45S rRNA. Plate A is a photograph of the ethidium bromide stained agarose gel of pB4-5.1 after digestion with: 1. BamHI, 2. HindIII, 3. SmaI/BamHI, 4. SmaI/Hind III, 5. HindIII/XhoI, 6. SalI/BamHI, 7. KpnI/BamHI, and 8. KpnI/XhoI. The fragments displayed on the agarose gel in plate A were transferred to nitrocellulose and hybridized first to 18S rRNA (B) and then 45S rRNA (C).

initiation site of mouse rDNA (Figure 3). Little similarity was found either upstream or downstream of the initiation site when the sequence of the rat rDNA was compared to those of *Xenopus laevis* (18), *Drosophila melanogaster* (19), *Saccharomyces cerevisiae* (20), or *Tetrahymena pyriformis* (21).

**S1 Nuclease Protection Assays.** When a 810 bp Sal I-Xho I fragment labeled at the Xho I site (+637) was hybridized to either total nuclear RNA or 45S rRNA, the initial S1 nuclease assay results showed protection into the first of the stretches of adenine located on the coding strand (Figure 4A). These results suggested that 45S rRNA begins at this site. However, examination of the surrounding DNA sequence shows long stretches of A-T base pairs, thus the result may also be explained by the instability of A-U hybrids at 45°C (22,23) and resultant S1 nuclease digestion. Reduction of the temperature at which S1 digestions were carried out resulted in the detection of longer hybrids (Figure 4B), indicating that the initiation site was upstream of the A-T stretches.

To locate the transcription initiation site more precisely, the Bam HI-Hind III fragment (-292 to +124) was labeled at the Hind III site, hybridized to either 45S or total nuclear RNA and digested with S1 nuclease. The size of the protected fragment was determined by coelectrophoresis on a sequencing gel with sequencing reactions of the same fragment (Figure 4C). When the

The nucleotide sequence is:

```

      -250      -240      -230      -220      -210
      CCGGAGCT GGACTTCTGA GGCCGAGGGG AAGCTGTGTT CAGGCTATCT ACCATGGCCT

-200      -190      -180      -170      -160      -150
CCTCGGCTCG TGAGCACTCA GGGTTCTAAG GTCGACCACT TGTTCCTTTG CCGTCCGGTT

-140      -130      -120      -110      -100      -90
CTCTTTCTAC ATGGGGACCT CTTGGGGGAC ACGTCACCGA ACATGACTTC CAGACGTTCC

-80      -70      -60      -50      -40      -30
GTGTGGCCTG TCATGTTTAT CCTGTGTCT TTTACACTT TCATCTTTGC TATCTGTCTT

-20      -10      1      10      20      30      40
TATTGTACCT GGAGATATAT GCTGACACGC TGTCTTTTG ACTCTTTTTC TCATTAAGG

      50      60      70      80      90      100
ACGTTGGAAG AGGCTTGCAC CAAGGCTGTT TGCTTGTCCA GCCCTAGCTC TTTTCTTCTG

      110      120      130      140      150      160
CGCATGGGCC TCTTCGATGC TGAAAGCTTA GGCTCCCCC ATGAGTACGC GCTTCTGTCT

      170      180      190      200      210      220
TTCCCGTGCT TGCTTGCCTG TGCTCTCTGG GGCAGCTTTA TGACAACCGT CCCGCTGTCT

      230      240      250      260      270      280
AGGGCTTCCC GATTTCCTCC TGGTGGTTGT CGTCCGTTGC CGGTAGGAGT CGTTGGTGCC

      290      300      310      320      330      340
GAGTGGCAGT GGAAGGGTTT TCCCGTTTGC TGCTAGTGAC CCCCTGGCGT GCTCTTCTCT

      350      360      370      380      390      400
CGGGCGACCG GTTTTTTAT TTGTTTTTT TTTTGTGTT TTTTGTGTT TTTTGTGTT

      410      420      430      440      450      460
TTTTTGGAA GGAGTTCCCG AACCTCCGCT GCTTGGTGGT GTGTCCCTTT CTTTCTGTCT

      470      480      490      500      510      520
GTGTGCCTCC CGAGTTGCAC CTTTTCTCCT TCGAAGGGGA TTTTATTTTT TTTTATTTTT

      530      540      550      560      570      580
TTTTTTTTTT ATTTTTTATT TTTTGAAGG AGTTCCCGAA CCTCCGCTGC CCGTTGAGTC

      590      600      610      620      630      640
CCGTTCCTTC ACGCCACGTG CCTCCCGAGT GCAACGCCTC CTTTTTTTTT CTCGCCCTCGA

      650      660      670      680      690      700
GAAGGGCAAA TTTTTTTTGG TGTGTGTGTG GCAGTGTTAG CGACTTCTTC CCGTGCTCTC

      710      720      730      740      750      760
TCTCGCTCTT CTCGCTCGTA TTCCCGTCCA GTCGGTGTGA GAAAGCTCTC ACGCCCGTTG

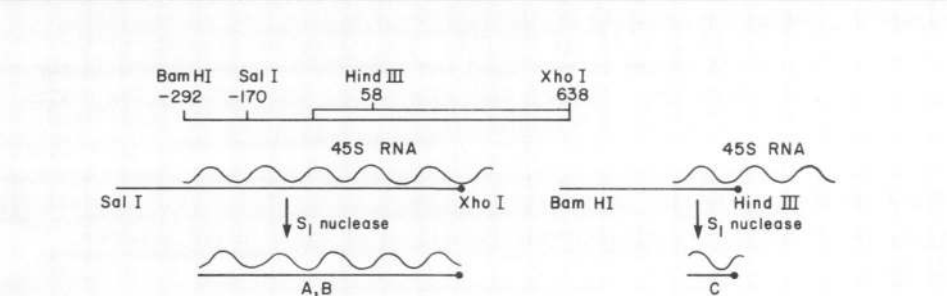
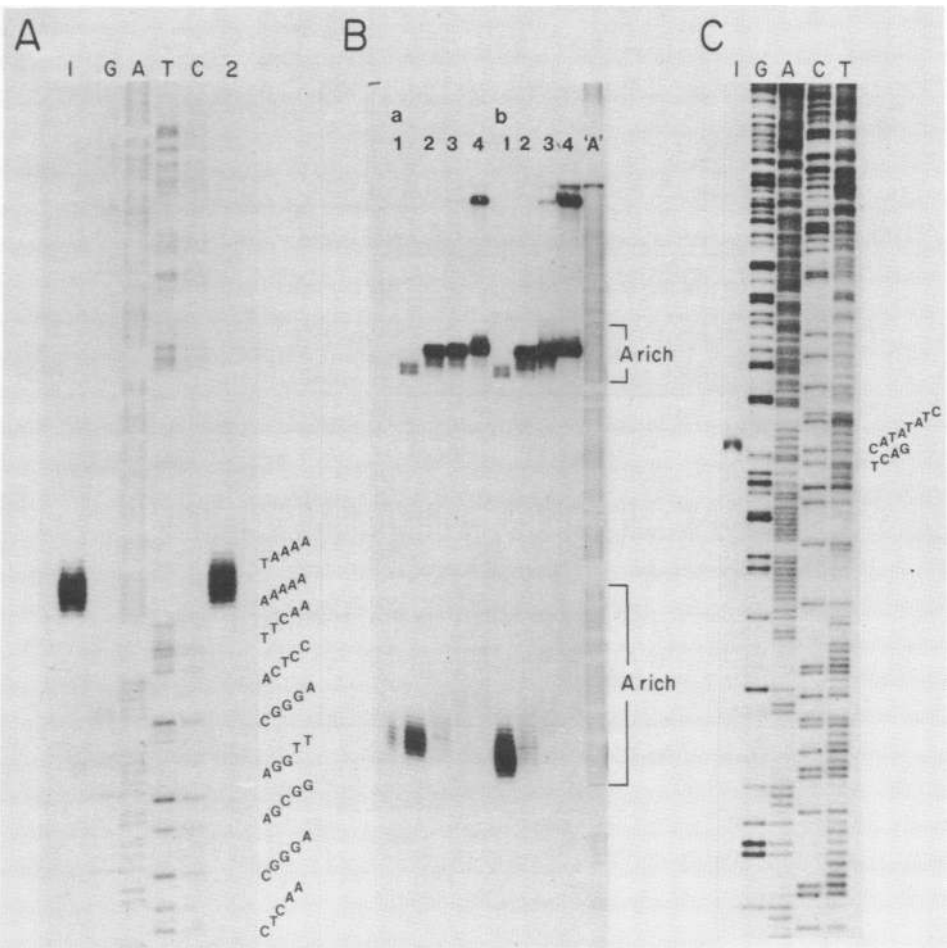
      770      780      790      800      810      820
TTCCCGATGC ATGGCGATGC TCGCTCCCGT TGGATCGATG TGGTGTCTGC CCGTTCTCTT

      830      840
CGGGCCGGGG CCAAGCCGCG CAGGCCGA
  
```

Figure 3 Nucleotide sequence of the region surrounding the transcription initiation site of rat 45S DNA. The nucleotide sequence of the noncoding strand is portrayed and numbered to correspond with the RNA transcript. The A-T rich regions are designated by the dotted lines. Regions of homology between rat and mouse are underlined.

migration distance of the S1 resistant fragment was corrected by  $1\frac{1}{2}$  nucleotides (18), the 5' terminal nucleotide was identified as the G in the sequence, TATGCT.

**Reverse Transcriptase Elongation Assays.** To examine this result, the 104 bp Hinf I-Hind III fragment (+23 to +124), labeled at the Hind III site, was hybridized to total nuclear RNA and the hybrids extended by reverse transcriptase. The sizes of the products of elongation were determined by coelectro-





phoresis on a sequencing gel with the sequence of the Bam HI-Hind III fragment (Figure 5). The Hinf I-Hind III fragment was extended 23 bp. The multiple bands of approximately the size of the initial Hinf I-Hind III fragment (Figure 5) may be due to premature termination by the reverse transcriptase. This could be due to the relatively low concentration of the template or to its secondary structure. These findings agreed with the S1 protection assay results. In control experiments when yeast tRNA was substituted for either 45S rRNA or nuclear RNA, or if RNA was omitted from the hybridization mixture, no S1 protection or reverse transcriptase elongation was detectable.

**In Vitro Capping.** The ability of an RNA molecule to act as the acceptor of  $\alpha^{32}\text{P}$  GTP in a reaction catalyzed by guanylyl transferase is an indication that that RNA contains a 5' polyphosphate terminus, and thus represents the initial transcription product (18). When Novikoff hepatoma cell 45S RNA was capped *in vitro*, approximately 5-10% of the molecules could be capped. The cap structure was identified following digestion of the RNA with ribonuclease P<sub>1</sub> or T<sub>1</sub> followed by alkaline phosphatase. Upon electrophoresis on DE-81 paper the products of P<sub>1</sub> digestion comigrated with authentic GpppG (Figure 6), as did the products of T<sub>1</sub> digestion (data not shown). Thus, the 5' terminal nucleotide of *in vivo* synthesized rat 45S RNA was found to be G, as predicted by the S1 and reverse transcriptase elongation experiments.

**In Vitro Transcription.** These results were consistent with data obtained following *in vitro* transcription of the cloned DNA. When either pB4-5.1 or pB7-2.0, two clones with the inserted fragments in reverse orientation, were cleaved with Hind III (+124), Xho I (+637) or Sau 3A (+793), the RNAs synthesized were approximately 124, 640 and 793 bases long, respectively (Figure 6). When pBR 322 DNA was transcribed under similar conditions, an RNA of those

Figure 4 Results of S1 nuclease mapping experiments to determine the 5' terminus of 45S rRNA.

- A. Results after hybridization (lanes 1 and 2) of the Sall-XhoI fragment with 45S rRNA and S1 digestion at 45C.
- B. Results after hybridization of the Sall-XhoI fragment with either 45S rRNA (a) or total nuclear RNA (b). S1 nuclease digestions were carried out at: (1) 45C, (2) 37C, (3) 30C and (4) 4C. Portrayed is the 'A' reaction of the coding strand of the Sall-XhoI fragment.
- C. Results after hybridization of BamHI-HindIII fragment to total nuclear RNA and S1 digestion at 37C. Lane 1 is the protected fragment displayed adjacent to the G, A, C and T reactions of the initial fragment. The nucleotide sequence of the coding strand immediately surrounding the initiation site is also displayed.

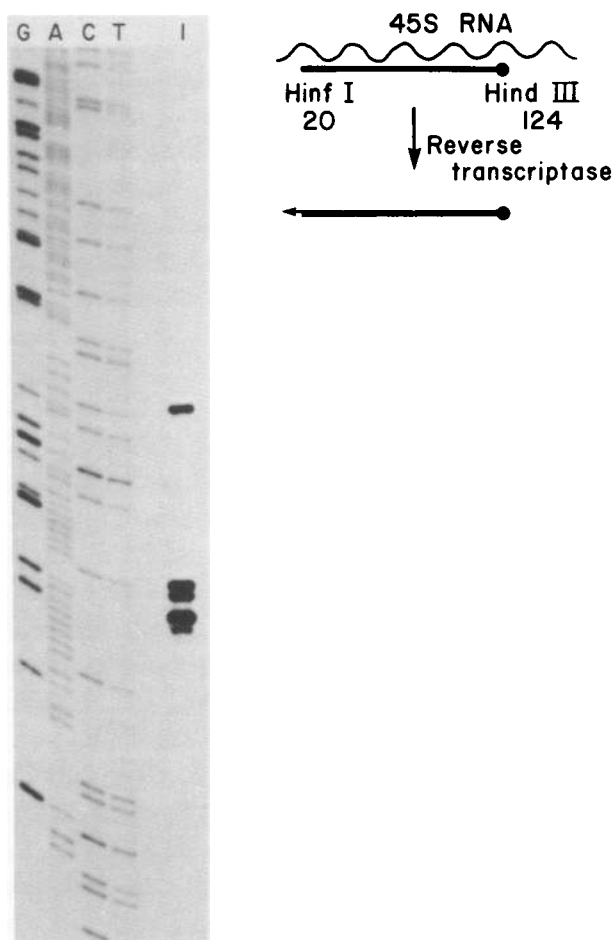


Figure 5 Results of reverse transcriptase elongation mapping experiments to determine the 5' terminus of 45S rDNA. Results after hybridizing total nuclear RNA with the 104 bp HinfI-HindIII fragment. The sequence displayed is that of the coding strand of the HindIII-BamHI fragment of pB4-5.1.

sizes was not synthesized (Figure 7). The second transcript of the Sau 3A template, of approximately 670 nucleotides, may be a processing product of the 793 base transcript. Although this result would be similar to the findings of Miller and Sollner-Webb (1), it requires further examination.

The transcripts from the Hind III truncated template was further characterized by fingerprinting the RNA synthesized *in vitro* in the presence of either  $\alpha$ - $^{32}\text{P}$  CTP or  $\alpha$ - $^{32}\text{P}$  UTP (Figure 8). When the  $T_1$  fingerprints of these trans-

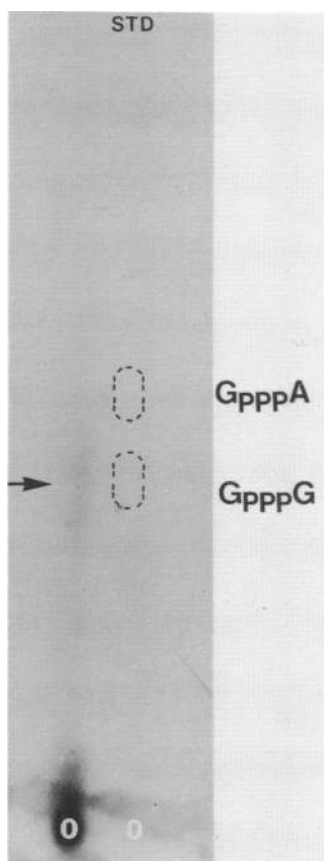


Figure 6 In vitro capping of 45S RNA. 45S RNA was capped and processed as described in Materials and Methods. The products of the in vitro capping reaction was analyzed by electrophoresis on DE-81. The migration of authentic GpppG is indicated.

cripts were compared to the DNA sequence of the initiation site (Figure 8), only the predicted oligonucleotides were found. A most striking characteristic of this experiment was the absence of oligonucleotides 7, 8 and 10 from the fingerprint of the transcript synthesized with  $\alpha$ - $^{32}\text{P}$  UTP as precursor, and the presence of spots A, B and C in the fingerprint of the same transcript (compare Figures 8A and B). Further, the presence of spot 16, (p)ppGp, which was found in the  $\alpha$ - $^{32}\text{P}$  CTP transcript (Figure 8A), would indicate that the transcript of the Hind III truncated template was not a product of the processing of a larger transcript, and that the 5' terminal G is adjacent to a C, as predicted by the results of the experiments using in vivo synthesized RNA.

#### DISCUSSION

The transcription initiation site of rat ribosomal DNA was examined using four separate experimental protocols: S1 nuclease protection assays, reverse

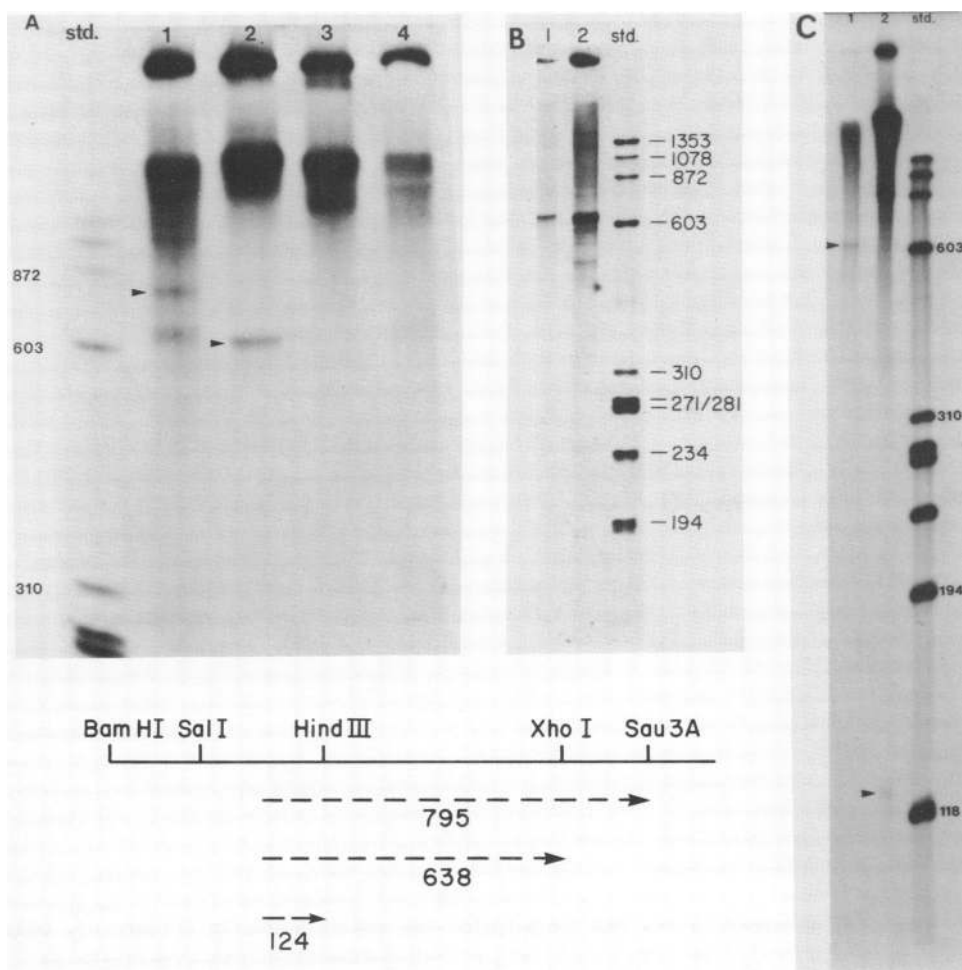


Figure 7 Autoradiographs of *in vitro* transcription products. A. *In vitro* transcripts of pB7-2.0 cleaved with Sau 3A (lane 1) or Xho I (lane 2) and pBR322 cleaved with Eco RI (lane 3) or Sau 3A (lane 4). B. *In vitro* transcripts of pB4-5.1 (lane 1) or pB7-2.0 (lane 2) cleaved with Xho I. C. *In vitro* transcripts of pB7-2.0 cleaved with Xho I (lane 1) or Hind III (lane 2). The arrowheads indicate transcripts that are of the predicted sizes.

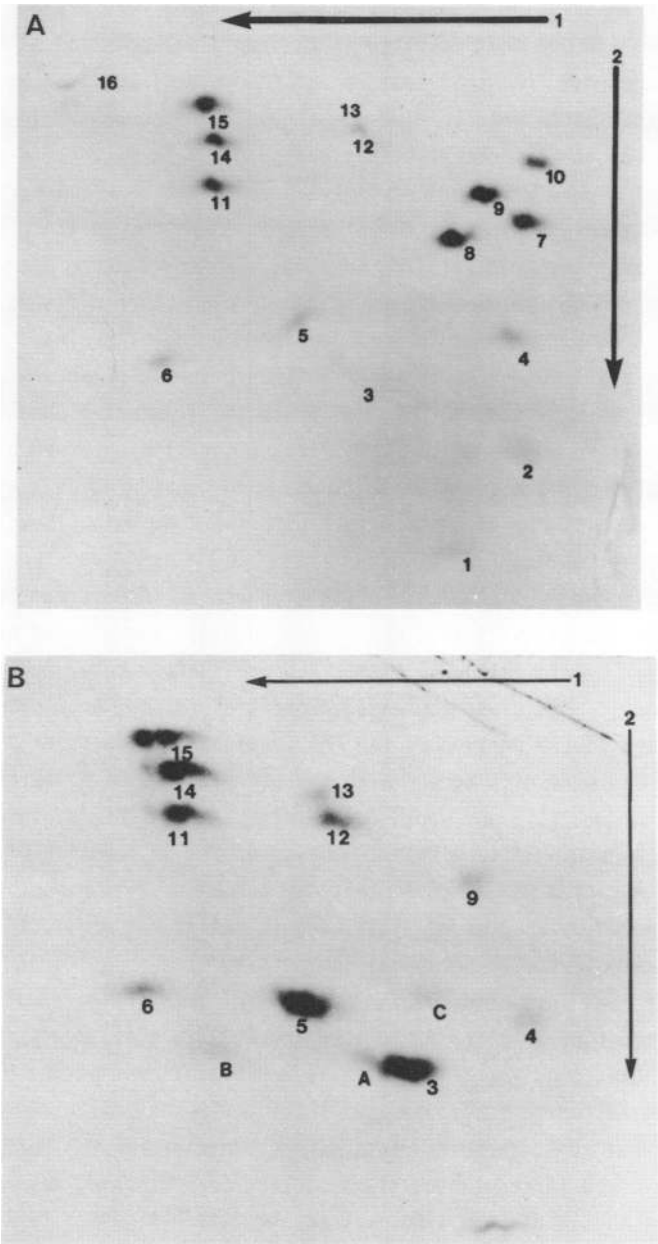
transcriptase elongation assays, *in vitro* capping, and *in vitro* transcription. The results of each protocol were consistent with the initiation site being TATAT<sup>\*</sup>GC.

The S1 nuclease and primer elongation results could, without the *in vitro*

capping and transcription data, be explained if there were a processing site at or near nucleotide one that resulted in the rapid dissociation of transcripts originating upstream from the remainder of the molecule. Although the primer elongation experiments were carried out at very high concentrations of nuclear RNA and no products were detected that extended beyond nucleotide one, the possibility of a very rapid processing step could not be eliminated. However, the identification of (p)ppG as the 5' end of 45S RNA by in vitro capping, the sizes of the in vitro transcription products, and the fingerprints of the products synthesized on the Hind III truncated template, were all consistent with initiation having occurred at the site defined by the S1 nuclease and primer elongation assays. Furthermore, in vitro transcription experiments using the 3.4 kb EcoR I-Bam HI fragment, which is 5' to the fragment cloned in pB4-5.1, did not demonstrate an RNA polymerase I initiation site upstream of the one defined in pB4-5.1 (data not shown). Thus, when considered as a whole, these experimental results are only compatible with the conclusion that TATATGC<sup>\*</sup> is the initiation site.

In complementary S1 nuclease protection assays, downstream fragments extending to the Xho I (+637) site were found to yield two classes of protected fragments after hybridization to either 45S pre-rRNA or total nuclear RNA (data not shown). One class indicated protection to the Xho I site and the other indicated that a portion of the RNA terminated <100 bases 3' to the Xho I site. These results agreed with the observations of Miller and Sollner-Webb (1) and Grummt (3), and would demonstrate that in rat, as in mouse, pre-rRNAs have a processing site within 700 bases of the 5' terminus of the initial transcript. A U rich RNA molecule that may represent this processing product has been reported by Choi et al. (24), and it yields the predicted S1 protection pattern (Choi, Y. C., personal communication).

The rat rDNA initiation region has several interesting structural features. The two stretches of dA-dT (+352 to +405 and +501 to +545) are strikingly similar to those found in the mouse (2,5) and reminiscent of the highly A-T rich nature of the Tetrahymena (21) and yeast (20) rDNA initiation regions. With respect to the rat rDNA, these findings suggest that either RNA polymerase I, or a factor(s) involved in regulating transcription or processing may interact with sites downstream of the initiation site. It may be possible that a factor that binds further than 300 base pairs from the initiation site affects gene regulation. Such a mechanism has been proposed (25,26) and investigations on testability suggest that highly dA-dT rich regions could affect the structure of adjacent regions (27).



RNA Sequence	G/CUG/ACACG/CUG/UCCUUUG/ACUCUUUUG/UCAUUAAG/G/ACG/UUG/G/AAG/AG/G/CUUG/									
Oligonucleotide	16	A	7	A	11	14	13	4	B	1 5
Radiolabelled Precursor:										
CTP	+	-	+	-	+	+	+	+	-	+
UTP	-	+	-	+	+	+	+	+	+	-

RNA Sequence	CACCAAG/G/CUG/UUG/CUG/UCCAG/CCCVAG/CUCUUUUCUUG/CGCAUG/G/G/CUCUUCG/											
Oligonucleotide	10	1	A	6	5	8	9	15	2	C	1	12
Radiolabelled Precursor:												
CTP	+	+	-	+	-	+	+	+	+	-	+	+
UTP	-	-	+	+	+	-	+	+	-	+	-	+

RNA Sequence	AUG/CUG'AA	
Oligonucleotide	3	A
Radiolabelled Precursor:		
CTP	+	-
UTP	+	+

Figure 8 Autoradiographs of RNase T<sub>1</sub> fingerprints of the 124 nucleotide transcripts of rDNA truncated at the Hind III site. Gel purified <sup>32</sup>P-RNA transcripts were digested with RNase T<sub>1</sub> and the oligonucleotides separated by electrophoresis on cellulose acetate (1) and homochromatography on PEI-cellulose plates (2). The transcripts were synthesized in the presence of either radiolabeled CTP (A) or UTP (B). The sequence of the transcript, derived from that of the DNA, is shown with slash marks indicating the sites of RNase T<sub>1</sub> digestion. The T<sub>1</sub> oligonucleotides that would be expected to be radiolabeled are indicated.

Immediately following the two dA-dT stretches, are identical 25 base pair sequences, +408 to +432 and +546 to +570. These sequences are of interest as they are recognized as dyad symmetries by the computer program of Korn *et al.* (28). In conjunction with the dA-dT stretches, these sequences may serve as regulatory elements either of transcription or processing.

The transcription initiation site of the rat rDNA is flanked by homologies to the analogous sites of the mouse and human rDNA (Figure 9), especially from +2 to +18. This finding raises the possibility that the sequences flanking the initiation site on both sides may constitute the recognition site for initiation of transcription. The presence of other regions of greater than 75% homology between the rat and mouse (Figure 3), for example from -167 to -131 and from +150 to +198, suggests that these sequences may also be involved in initiation or in processing. These homologies are large, greater than 20

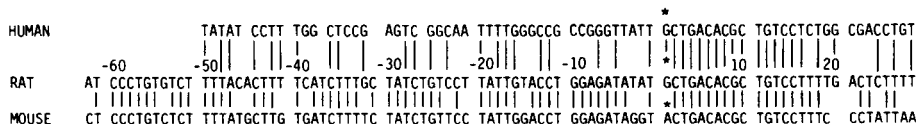


Figure 9 Sequence homologies of the transcription initiation sites of rat, mouse (2) and human (G. Wilson, personal communication) ribosomal RNA genes. The sequences have been aligned to demonstrate homology with respect to the rat sequence. Bars connecting the nucleotides designate sequences which are homologous; an asterisk indicates the first nucleotide of the respective transcripts.

bp, and are located at almost identical distances from the respective initiation sites. These areas of homology are interspersed with regions that are divergent. The possibility that they have been conserved due to functional roles should be considered.

If the homologous sections of the initiation regions are regulatory or modulatory elements, the genes transcribed by RNA polymerase I may represent a transcription system comprised of both internal and external regulatory elements. This hypothesis, which predicts function based solely on sequence homology requires further study. The observation of Long *et al.* (19), that both active and inactive ribosomal RNA genes of *Drosophila* have identical transcription initiation regions, suggests that further studies will be necessary before the nature of the elements that regulate the transcription of rDNA are identified.

Our findings and the studies of mouse and human rDNA indicate that the mammalian ribosomal gene transcription initiation regions have common structural elements. As human ribosomal RNA genes can be expressed in human-mouse heterokaryons (29,30), it may be that there are both common regulatory regions as well as species specific sequences as suggested by the *in vitro* transcription experiments of Grummt *et al.* (4).

#### NOTE ADDED IN PROOF

While this manuscript was in preparation, Financsek *et al.* (33,34) reported the sequence of 300 nucleotides surrounding the transcription initiation sites of human and rat rDNA. Their observations are in agreement with those reported in this manuscript.

#### ACKNOWLEDGEMENTS

The authors wish to acknowledge the aid and encouragement of Drs. Harris



Busch and Subrahmanyam Chirala, and Dr. Golder Wilson for providing the sequence of the initiation site of human rDNA prior to publication. We would also like to thank Dr. Y. C. Choi for his generous gift of 45S rRNA and Drs. Michael Brattain, Lewis Rodriguez and Ming-jeer Tsai for their aid in the preparation of this manuscript. This research was supported by the Cancer Research Grant, CA-10893, P9, awarded by DHEW.

# REFERENCES

1. Miller, K.G. and Sollner-Webb, B. (1981) *Cell* 27, 165-174.
2. Urano, Y., Kominami, R., Mishima, Y., and Muramatsu, M. (1980) *Nucleic Acids Res.* 8, 6043-6058.
3. Grummt, I. (1981) *Nucleic Acids Res.* 9, 6093-6102.
4. Grummt, I., Roth, E. and Paule, M.E. (1982) *Nature* 296, 173-174.
5. Bach, B., Grummt, I. and Allet, B. (1981) *Nucleic Acids Res.* 9, 1559-1569.
6. Mishima, Y., Kominami, R., Honjo, T. and Muramatsu, M. (1980) *Biochemistry* 19, 3780-3786.
7. Rothblum, L.I., Parker, D.L., and Cassidy, B. (1982) *Gene* 10, 75-77.
8. Bolivar, F. and Backman, K. (1979) in *Methods in Enzymology*, R. Wu, Ed., Vol. 68, pp. 245-267, Academic Press, New York.
9. Subrahmanyam, C.S., Cassidy, B., Busch, H. and Rothblum, L.I. (1982) *Nucleic Acids Res.* 10, 3667-3680.
10. Busch, H. (1967) in *Methods in Enzymology*, L. Grossman and K. Moldave, Eds., Vol. 12A, pp. 434-439, Academic Press, New York.
11. Choi, Y.C. and Busch, H. (1970) *J. Biol. Chem.* 245, 1954-1968.
12. Choi, Y.C. and Busch, H. (1978) *Biochemistry* 17, 2551-2560.
13. Maizels, N. (1976) *Cell* 9, 431-438.
14. Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721-732.
15. Bina-Stein, M., Thoren, M., Salzman, N. and Thompson, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 731-735.
16. Weill, P.A., Segall, J., Harris, B., Ng, S.Y. and Roeder, R.G. (1979) *J. Biol. Chem.* 254, 6163-6173.
17. Tsai, S.Y., Roop, D.R., Stumph, W.E., Tsai, M.J., Stein, J.P., Means, A.R., and O'Malley, B.W. (1978) *Biochemistry* 17, 5773-5780.
18. Sollner-Webb, B. and Reeder, R.H. (1979) *Cell* 18, 485-499.
19. Long, E.O., Rebbert, M.L. and Dawid, I.B. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 1513-1517.
20. Bayev, A.A., Georgiev, O.I., Hadjilov, A.A., Kermekchiev, M.B., Nikolaev, N., Skryabin, K.G. and Zqkharyev, V.M. (1980) *Nucleic Acids Res.* 8, 4919-4926.
21. Niles, E.G., Sutiphong, J. and Haque, S. (1981) *J. Biol. Chem.* 256, 12849-12856.
22. Casey, J. and Davidson, N. (1977) *Nucleic Acids Res.* 4, 1539-1552.
23. Martin, F.H. and Tinoco, I. (1980) *Nucleic Acids Res.* 8, 2295-2299.
24. Choi, Y.C., Lee, K.M., and Busch, H. (1976) *Proc. Am. Assoc. Cancer Res.* 17, 38.
25. Burd, J.F., Wartell, R.M., Dodgson, J.B. and Wells, R.D. (1975) *J. Biol. Chem.* 250, 5109-5113.
26. Dickson, R.C., Abelson, J., Barnes, W.M. and Reznikoff, W.S. (1975) *Science* 187, 27-35.
27. Burd, J.F., Larson, J.E. and Wells, R.D. (1975) *J. Biol. Chem.* 250, 6002-6007.
28. Korn, L.J., Queen, C.L. and Wegman, N.M. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 4401-4405.

29. Soprano, K.J., Dev, V.G., Croce, C.M. and Baserga, R. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 3885-3889.
30. Soprano, K.J. and Baserga, R. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 1566-1569.
31. Reddy, R., Henning, D. and Busch, H. (1981) J. Biol. Chem. 256, 3532-3538.
32. Brownlee, G.G., Sanger, F. and Barrell, B.G. (1968) J. Mol. Biol. 34, 379-412.
33. Financsek, I., Mizumoto, K., Mishima, Y. and Muramatsu, M. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 3092-3096.
34. Financsek, I., Mizumoto, K. and Muramatsu, M. (1982) Gene 18, 115-122.