Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 3092–3096, May 1982
Biochemistry

Human ribosomal RNA gene: Nucleotide sequence of the transcription initiation region and comparison of three mammalian genes

(rDNA/sequence analysis/initiation site/S1 nuclease mapping/*in vitro* capping)

ISTVAN FINANCSEK*†, KIYOHISA MIZUMOTO‡, YUKIO MISHIMA*, AND MASAMI MURAMATSU*

*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan; and ‡Department of Chemistry, Institute of Medical Science, University of Tokyo, Shiroganedai, Minato-ku, Tokyo 108, Japan

Communicated by Susumu Ohno, January 29, 1982

ABSTRACT **The transcription initiation site of the human ribosomal RNA gene (rDNA) was located by using the single-strand specific nuclease protection method and by determining the first nucleotide of the *in vitro* capped 45S preribosomal RNA. The sequence of 1,211 nucleotides surrounding the initiation site was determined. The sequenced region was found to consist of 75% G and C and to contain a number of short direct and inverted repeats and palindromes. By comparison of the corresponding initiation regions of three mammalian species, several conserved sequences were found upstream and downstream from the transcription starting point. Two short A+T-rich sequences are present on human, mouse, and rat ribosomal RNA genes between the initiation site and 40 nucleotides upstream, and a C+T cluster is located at a position around −60. At and downstream from the initiation site, a common sequence, T-$^{G}_{A}$-C-T-G-A-C-A-C-G-C-T-G-T-C-C-T-$^{T}_{C}$-T, was found in the three genes from position −1 through +18. The strong conservation of these sequences suggests their functional significance in rDNA. The S1 nuclease protection experiments with cloned rDNA fragments indicated the presence in human 45S RNA of molecules several hundred nucleotides shorter than the supposed primary transcript. The first 19 nucleotides of these molecules appear identical—except for one mismatch—to the nucleotide sequence of the 5′ end of a supposed early processing product of the mouse 45S RNA.**

The ribosomal RNA genes (rDNA) in eukaryotes are transcribed by RNA polymerase I (RNA nucleotidyltransferase, EC 2.7.7.6). This enzyme is in many respects different from the other two RNA polymerases present in eukaryotic cells (1, 2). The regulation of transcription initiation by RNA polymerase II and III is beginning to be understood. Conserved short sequences at fixed distances from the initiation site were found in bacterial genes (3) and in eukaryotic genes transcribed by RNA polymerase II (4, 5), although their function may not be necessarily the same (6, 7). Very limited conserved sequences were found in the prelude region of genes transcribed by RNA polymerase III (8–10). Rather, sequences downstream from the initiation point were found to be needed for correct initiation (11, 12).

Contrary to polymerases II and III, virtually nothing is known about the regulatory sequences of transcription by RNA polymerase I. Nucleotide sequences of ribosomal RNA transcription initiation regions were determined in several eukaryotes (13–19). Only very limited similarity was found in those regions of the compared genes where some regulatory sequences might exist (19).

In the present work, we have determined the transcription initiation site of the human rRNA gene and then compared the surrounding nucleotide sequences with those of the mouse and rat ribosomal genes that were determined previously (16, 20, 21). By comparison of the corresponding regions from closely and distantly related mammals, we were able to identify some conserved regions that may have functional significance.

MATERIALS AND METHODS

Enzymes. Restriction endonucleases were purchased from Bethesda Research Laboratories or from Takara Shuzo, Kyoto; bacterial alkaline phosphatase (EC 3.1.3.1) from Worthington and polynucleotide kinase (polynucleotide 5′-hydroxyl-kinase, EC 2.7.1.78) from Boehringer Mannheim were used according to the supplier's instructions. P1 nuclease was the gift of A. Kuninaka and M. Fujimoto (Yamasa Shoyu, Choshi). Rat liver guanylyltransferase was purified as described (22).

Preparation of 45S RNA. 45S RNA was prepared from a human mammary tumor cell line, MX-1, inoculated into nude mice. RNA was extracted from isolated nuclei or nucleoli, and the 45S RNA was purified by two or three cycles of sucrose gradient centrifugation as described (23).

S1 and P1 Nuclease Protection Mapping. The protocol of Berk and Sharp (24) was followed. Briefly, 5′-end-labeled DNA fragments (100–200 ng) and 45S RNA (6–8 μg) were hybridized at 61°C for 4 hr in 80% formamide/0.4 M NaCl/0.04 M 1,4-piperazinediethanesulfonic acid, pH 6.8/1 mM EDTA (total volume, 25 μl). S1 and P1 nucleases were used at 500 units/ml and 15–60 μg/ml, respectively. S1 digestion was at 45°C and P1 digestion was at 50°C for 1 hr. P1 buffer (suggested by M. Fujimoto) was 0.3 M NaCl/0.05 M sodium acetate, pH 5.3/0.1 mM $ZnSO_4$. In these experiments we found that the P1 nuclease (25)—which was more stable than the mung bean nuclease (26)—was apparently superior to the S1 nuclease, generating a single-length protected DNA instead of the short "ladder" found in most S1 protection experiments. After phenol extraction, DNA that had been precipitated with ethanol was electrophoresed either in alkaline agarose (27) or in polyacrylamide urea gels (28).

***In Vitro* Capping of 45S RNA and Cap Analysis.** The capping reactions were carried out by using purified rat liver guanylyl-

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: rDNA, ribosomal RNA gene; kb, kilobase(s); bp, base pair(s).

† On leave from the Department of Biochemistry (Head, E. J. Hidvegi), Frederic Joliot-Curie National Research Institute for Radiobiology and Radiohygiene, Budapest, under a cooperation program between Japan Society for the Promotion of Science and Hungarian Academy of Sciences.

transferase and [α-^{32}P]GTP (22) as detailed elsewhere (21).

After treatment with P1 nuclease and alkaline phosphatase, the resulting products were subjected to paper electrophoresis on Whatman DE81 paper (22).

DNA Sequence Determination. This was done as described by Maxam and Gilbert (28).

rDNA Clone and Plasmid DNA Preparation. The human rDNA fragment containing the initiation site, cloned in pBR322 (29) and designated pBE2, was the gift of R. D. Schmickel. Plasmid DNA was prepared as described (30). Guidelines set by the Japanese government were followed.

RESULTS

Localization of the Initiation Site by S1 Protection Mapping. Electron microscopic measurements of human rDNA·45S RNA hybrids showed that the 5′ end of the 45S RNA was about 3.8 kilobase(s) (kb) upstream from the beginning of the 18S RNA coding region (31, 32). This site falls on the *Eco*RI/*Bam*HI 2.1-kb fragment located at the extreme left of the cloned rDNA (Fig. 1*A*). We first used this fragment in an S1 protection experiment with 45S RNA, obtaining two protected fragments (Fig. 2*A*). The longer of these (1.6 kb) ends at about 3.9 kb from the 18S coding region, in good agreement with the electron microscopic observation. Another protected DNA of 1.1 kb was also found. We next examined the same region by using the *Eco*RI/*Sal* I 1.2-kb fragment, the left end of which was the same as that of the 2.1-kb DNA except that 0.88 kb was removed by the *Sal* I cut from the opposite side. The two protected bands that appeared after S1 treatment (Fig. 2*B*) were found to be 0.88 kb shorter than those in the previous experiment. These experiments locate the 5′ end of the 45S RNA at about 5.6 kb from the right *Eco*RI site. The second site that was found several hundred nucleotides downstream could correspond either to the second initiation site or to an early processing site.

To localize more precisely the 5′ end of the 45S RNA, we used shorter 5′-labeled DNA fragments generated by cutting the *Eco*RI/*Sal* I 1.2-kb DNA with *Hinf*I enzyme (Fig. 1*B*). When the rightmost *Sal* I/*Hinf*I 536-bp fragment was used for protection, a band of 280 nucleotides was detected as expected by the previous experiment (Fig. 2*C*). On the other hand, the

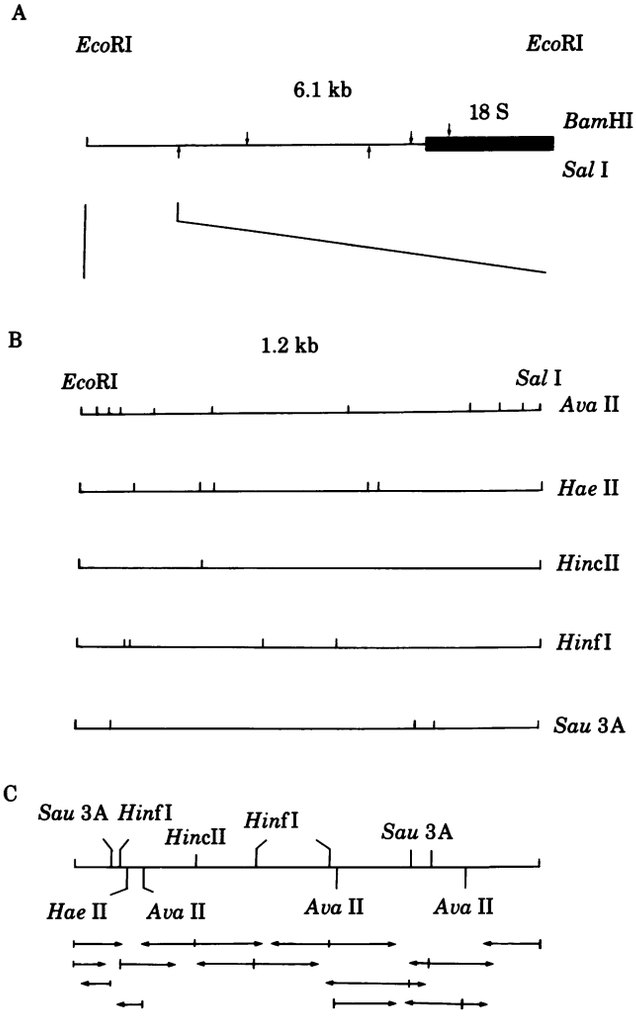

FIG. 1. (*A*) Restriction map of the 6.1-kb human rDNA fragment cloned in pBR322 (29). (*B*) Restriction map of the *Eco*RI/*Sal* I 1.2-kb fragment by the enzymes used for sequence analysis. (*C*) Sequence analysis strategy.

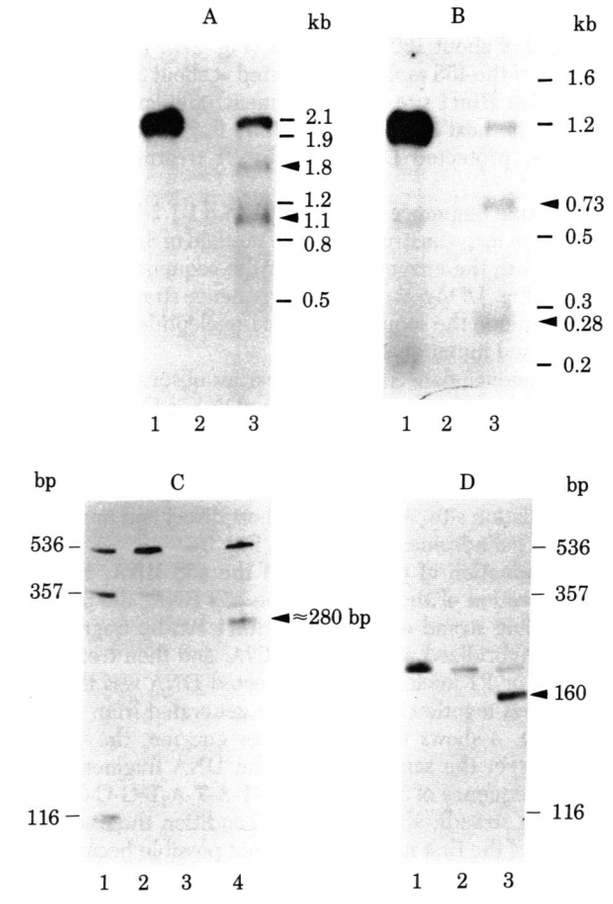

FIG. 2. Localization of the initiation site by S1 protection mapping. (*A*) The 5′-labeled *Eco*RI/*Bam*HI 2.1-kb fragment (lane 1) was hybridized with 45S RNA (lane 3) or with yeast RNA (lane 2), treated with S1 nuclease, and electrophoresed in a 1.2% alkaline agarose gel. The arrowheads point to the protected DNA bands. (*B*) The *Eco*RI/*Sal* I 1.2-kb fragment (lane 1) was hybridized with 45S RNA (lane 3) or with yeast RNA (lane 2), S1-treated, and electrophoresed in a 1.6% alkaline agarose gel. (*C*) Determination of the 5′ end of the shorter 45S RNA. The *Hinf*I/*Sal* I 536-base pair(s) (bp) fragment (lane 2) was hybridized with 45S RNA (lane 4) or with yeast RNA (lane 3), S1-treated, and electrophoresed in a 5% polyacrylamide/7 M urea gel. Lane 1 shows the fragments that are size markers. The arrowhead points to the protected DNA band. (*D*) Determination of the 5′ end of the longer 45S RNA. The 5′-labeled *Hinf*I/*Hinf*I 195-bp fragment (lane 1) was hybridized with 45S RNA (lane 3) or with yeast RNA (lane 2), S1-treated, and electrophoresed in a 5% polyacrylamide/7 M urea gel. The arrowhead indicates the protected DNA band. The positions of size markers are shown on the right.

second *Hinf*I/*Hinf*I 195-bp fragment produced a protected DNA band of about 160 nucleotides (Fig. 2*D*), indicating that the 5′ end of the 45S molecule is located at about 30 nucleotides from the left *Hinf*I site of this fragment. Similar experiments that used the next 358-bp *Hinf*I/*Hinf*I fragment did not give rise to any protected DNA band by S1 treatment (data not shown).

Nucleotide Sequence of the *Eco*RI/*Sal* I 1.2-kb Fragment. A restriction map constructed by the method of Smith and Birnstiel (33) with the enzymes used for the sequence analysis are shown on Fig. 1*B* together with the sequence strategy (Fig. 1*C*). We determined the sequence of 1,211 nucleotides surrounding the supposed initiation site (Fig. 3).

The sequence data confirm the earlier observation (34) that the transcribed spacer of the human rDNA is G+C-rich. Seventy-five percent of the 1,211 nucleotides indeed consist of G·C pairs. The general G+C-rich character is interrupted by a few relatively short A+T-rich sequences, which precede the supposed initiation site. A number of short direct and inverted repeats and palindromes were found (Fig. 3).

Determination of the 5′ End of the 45S RNA. For more precise location of the 5′ end of the 45S RNA, the 5′-end-labeled coding strand of the *Hinf*I/*Hinf*I 195-bp fragment was isolated, hybridized with the 45S RNA, and then treated with either S1 or P1 nuclease. The protected DNA was then electrophoresed together with G ladder generated from the same DNA. Fig. 4 shows that, with either enzyme, the 45S RNA could protect the same length of the DNA fragment ending within a sequence of 5′-C-G-G-G-T-T-A-T-A-T-G-C-3′ (on the noncoding strand), although in our condition the exact determination of the first nucleotide was not possible because of the long length of the protected band. However, it can be read that the protected DNA ends much closer to the G-C at the 3′ end than to the 5′ end, especially when a theoretical correction is made by 1½ nucleotides downwards (35, 13). Because we found, by independent determinations, the first nucleotide of the 45S RNA to be G, there remains only one possible starting point—namely, the first G after the T-T-A-T-A-T sequence.

Determination of the First Nucleotide of the 45S RNA by *in Vitro* Capping. To demonstrate the first nucleotide of the 45S RNA, this molecule was capped by rat liver guanylyltransferase and [α-^{32}P]GTP. This enzyme can efficiently cap the 5′ di- or triphosphate end of RNA but not the monophosphate end of RNA or DNA (ref. 22; also unpublished). Approximately 15–20% of the human 45S RNA molecules could be capped (data not shown). Other primary transcripts of eukaryotic rDNA were shown to form the cap structure only up to 10–20%, probably due to rapid processing (18, 36, 37). By analysis of the cap structure, >95% of the product was GpppG (Fig. 5*A*) and the remaining 5% appeared to be GpppA. The presence of the latter

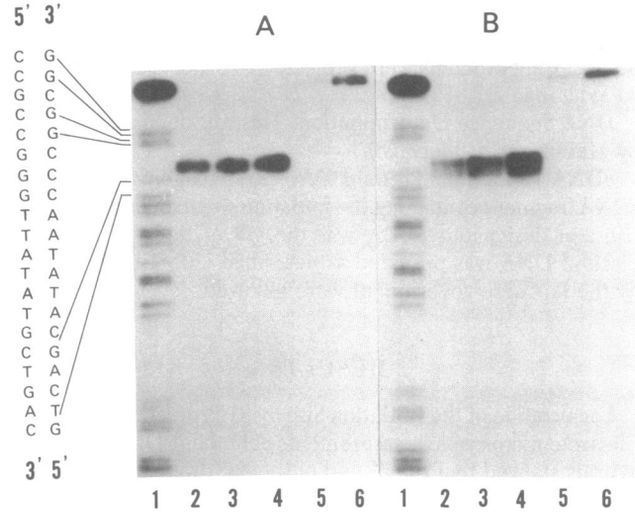

FIG. 4. Sizing of S1- and P1-nuclease generated fragments. The coding strand of the 195-bp *Hinf*I/*Hinf*I fragment was hybridized with 45S RNA and digested with P1 (*A*) or S1 (*B*) nuclease. The material was separated on a 0.4-mm thick 8% polyacrylamide/7 M urea gel. (*A*) Lane 1, G ladder generated from the coding strand; lanes 2, 3, and 4, P1-treated samples at concentrations of 60, 30, and 15 μg/ml, respectively; lane 5, mock-hybridized with yeast RNA and P1-treated at 30 μg/ml; lane 6, the 5′-labeled coding DNA strand. (*B*) Lane 1, G ladder of the coding strand; lanes 2, 3, and 4, S1 treatment at 600, 400, and 200 units/ml; lane 5, mock-hybridized with yeast RNA and S1 treatment at 400 units/ml; lane 6, the 5′-labeled coding DNA strand.

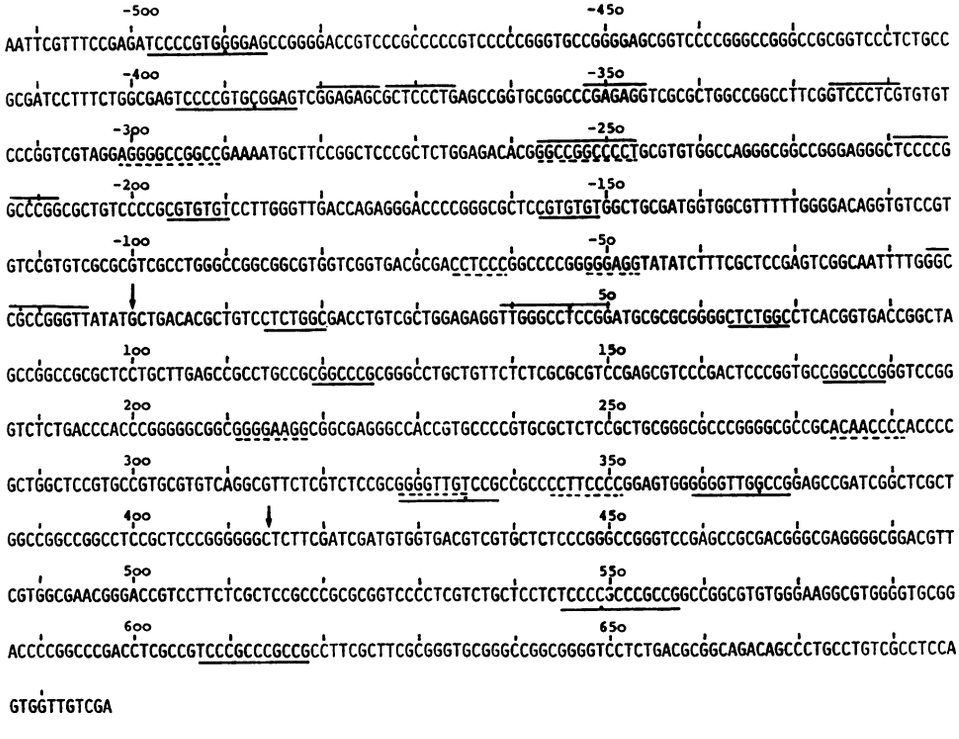

FIG. 3. The nucleotide sequence of the transcription initiation region of human rDNA. The sequence of the noncoding strand of the *Eco*RI/*Sal* I 1.2-kb fragment is shown in the 5′ → 3′ direction. Transcription proceeds from left to right. The arrow pointing to nucleotide 1 indicates the transcription starting point; the arrow at position 414 shows the probable 5′ end of the shorter 45S RNA molecules. The direct repeats are underlined, the inverted repeats are marked by the lines above the sequence, and the sequences with dyad symmetry (palindromes) are indicated by the dashed lines. Only those that are at least seven nucleotides long are marked.

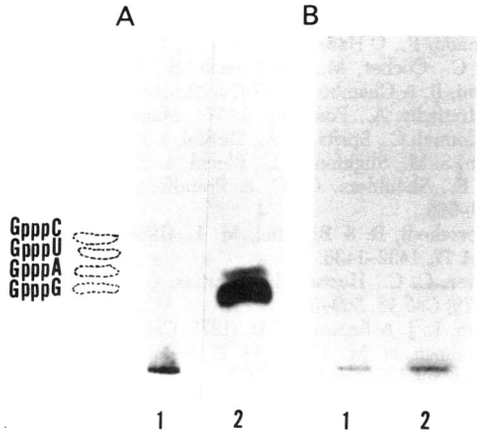

FIG. 5. Analysis of the cap structure formed *in vitro*. The positions of the markers are indicated on the far left. (*A*) Cap structure formed with 45S RNA isolated by sucrose gradient centrifugation. The capping reaction was carried out as described. After capping, RNA was treated with nuclease P1 and alkaline phosphatase and electrophoresed on Whatman DE81 paper at pH 3.4 (22). Lane 1, no RNA; lane 2, with 0.26 μg of 45S RNA. (*B*) Cap structure formed with the 45S RNA fragment isolated by hybridization. Conditions for the capping reaction and processing of the samples were the same as in *A*. Lane 1, no RNA; lane 2, with 45S RNA fragment isolated by hybridization. In this experiment, 45S RNA that was purified by sucrose gradient was hybridized to the *Eco*RI/*Sal* I 1.2-kb DNA fragment and, after P1 nuclease treatment, electrophoresed in 1.2% neutral agarose gel. To isolate the longer DNA·45S RNA hybrids, 5′-labeled DNA·RNA hybrids were run in the gel in parallel with the unlabeled hybrids. After autoradiography the appropriate band was cut out, recovered from the gel, and used in the capping experiment.

product may be due either to some contaminating RNA or to the real heterogeneity existing among the 45S RNA molecules.

To eliminate the possible contaminating RNAs in the 45S RNA obtained by sucrose gradient, this RNA was further purified by hybridizing to the *Eco*RI/*Sal* I 1.2-kb DNA and selecting the hybrids with the longer 45S RNA by the gel. Fig. 5*B* shows that the purified RNA can be capped, the first nucleotide again being G—confirming the result obtained by using 45S RNA from the sucrose gradient.

DISCUSSION

We have localized the transcription initiation site on human rDNA by mapping with single strand-specific nucleases and by determining the 5′ end of the 45S RNA with the *in vitro* capping method. The finding that 95% of the *in vitro* capped 45S RNA has G as the first nucleotide allows us to conclude that the great majority of 45S transcription starts at position 1 shown in Fig. 3. The absence of any longer RNAs than those ending at the site determined and the results of the *in vitro* capping of the 45S RNA fragment purified by hybridization confirm our assignment of the initiation site. The striking homology around and downstream from the initiation sites of the three mammalian rDNAs (see below) is also compatible with this conclusion.

We have cloned the rat rDNA, determined the transcription initiation site, and subjected the surrounding region to sequence analysis (21). Because the mouse initiation site was determined earlier (16, 20), we are now able to compare two phylogenetically closely related mammalian rDNAs (rat and mouse) and a more distant one (human). About 300 nucleotides surrounding the initiation site are compared in Fig. 6 *Upper*. Although the sequence near this region is very similar between the mouse and the rat, only several stretches are homologous or very similar among the three species.

Upstream from the initiation sites of three species, there are a few A+T-rich stretches that resemble the Pribnow or T-A-T-A box. These can be found at positions −21 to −18 and −43 to −38 on the human sequence. Although not underlined in Fig. 6 *Upper*, there is an A+T-rich short sequence at the starting point of the three genes (positions −5 to −2 on the human

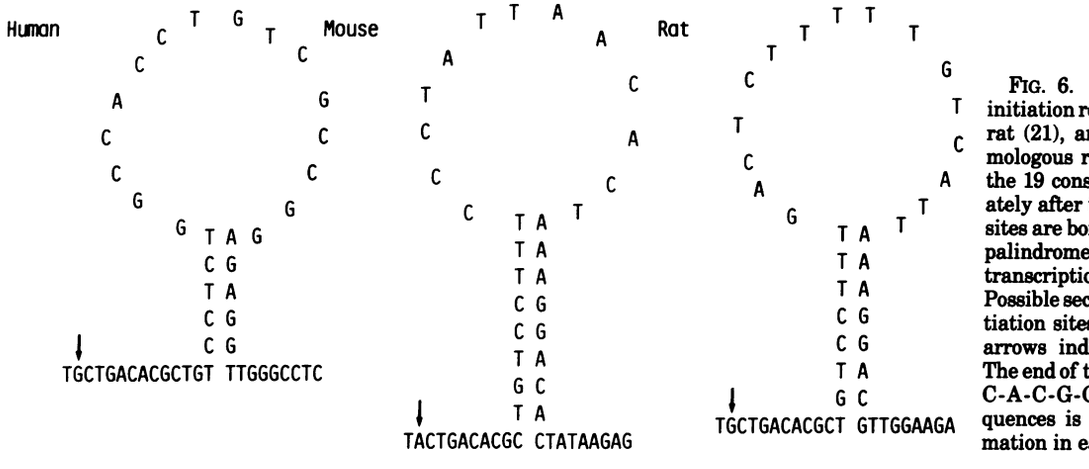

FIG. 6. (*Upper*) Comparison of the initiation regions of the mouse (16, 20), rat (21), and human rDNA. The homologous regions are underlined and the 19 conserved nucleotides immediately after the transcription initiation sites are boxed. Dotted sequences form palindromes. The arrow points to the transcription starting point. (*Lower*) Possible secondary structure at the initiation sites of the three rDNAs. The arrows indicate the starting points. The end of the conserved T-$\frac{G}{A}$-C-T-G-A-C-A-C-G-C-T-G-T-C-C-T-$\frac{T}{C}$-T sequences is involved in the stem formation in each case.

gene). The *Xenopus* rDNA also has a T cluster between −27 and −22 nucleotides. There is a short C-T cluster, located at about −60 on the three genes, that can also be found in the *Xenopus* gene at a similar position. A region found at about 120 nucleotides upstream from the initiation site is rather well conserved in the genes that were compared and may have some functional significance. The first 19 nucleotides at and downstream from the initiation point are almost identical among three species, suggesting the functional importance of this region—together with the purine-rich conserved sequence at around position 40—in transcription initiation. These conserved regions can form a base-paired structure in each case as shown by Fig. 6 *Lower*. Similar but more complex base-paired regions are found, among others, in the leader transcript of tryptophan operon in *Escherichia coli* (38), histidine operon of *Salmonella typhimurium* (39), and in the mRNA of the pE 194 plasmid encoding erythromycin resistance (40).

The presence of conserved sequences around the initiation point of the three species suggests the presence of common mechanisms in transcription regulation, although a part of the regulatory factors may be species specific (41).

Although the shorter class of 45S RNA molecules was not mapped precisely, we note the existence of homologous sequences of 19 nucleotides between the mouse and human genes at the position where the 5′ end of this shorter RNA seems to be localized. The position on the human gene is between 414 and 432 nucleotides after the starting point and between 651 and 669 in the case of the mouse. This sequence, if the shorter RNA is a processing product (16, 17, 20), might serve as a processing signal.

The sequences of the *Xenopus laevis* rDNA upstream from the initiation site—G-G-G-G-A-G-C-C-G-G (−281 to −272), T-C-C-C-C-C-G-G (−267 to −260), and T-C-C-C-C-G-G-G-C-C (−250 to −241)—are homologous to the nucleotide sequences of the human gene at −491 to −482, −464 to −457, and −442 to −433 in the same order although with different intervals. The sequence of G-G-G-G-A-A-G (2 to 8) at the initiation site of *Xenopus* and the C-G-G-G-C-G-C-C-C-G (26 to 35) sequence are homologous to the human sequences downstream from the initiation site (211 to 217 and 255 to 264), again in the same order but with slightly different interval. The presence of these homologous sequences on the human gene at different locations from those of *Xenopus* may be due to the rearrangement that occurred during the evolution. The short repeated sequences shown in Fig. 3 may provide a basis for recombinational events, similar to those found in bacterial (42) and eukaryotic (5, 43) genes. The conserved part resembling the *Xenopus* initiation region downstream from the human initiation site may either have some function or represent only a remnant of a former initiation site. On the other hand, the possibility that such similar sequences come about by chance alone cannot be ruled out completely. Further studies *in vivo* and *in vitro* with modified rDNA sequences will answer the question as to the role of the conserved sequences in mammalian rRNA genes.

The authors thank Dr. R. D. Schmickel for providing the human rDNA clone and Drs. R. Kominami and Y. Urano for their helpful advice during the work. They also thank Drs. M. Sakai and M. Fujimoto for S1 and P1 nucleases, respectively, Drs. K. Inoue and S. Fujimoto for human tumor cells, and Drs. Y. Fujii-Kuriyama and T. Taniguchi for critical reading of the manuscript. The capping experiments were carried out in the laboratory of Prof. Y. Kaziro to whom our thanks are due.

1. Chambon, P. (1975) *Annu. Rev. Biochem.* **44**, 613–638.
2. Roeder, R. G. (1976) in *RNA Polymerases*, eds. Losick, R. & Chamberlain, M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 285–329.
3. Pribnow, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 784–788.
4. Gannon, F., O'Hare, K. O., Perrin, F., LePennec, J. P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B. & Chambon, P. (1979) *Nature (London)* **278**, 428–434.
5. Efstratiadis, A., Posakony, J. W., Maniatis, R., Lawn, R. M., O'Connel, C., Spritz, R. A., DeRiel, J. K., Forger, B. G., Weissman, S. M., Slighton, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
6. Grosschedl, R. & Birnstiel, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1432–1436.
7. Baker, C. C., Herisse, J., Courtois, G., Galbert, F. & Ziff, E. (1979) *Cell* **18**, 569–580.
8. Korn, L. J. & Brown, D. D. (1978) *Cell* **15**, 1145–1156.
9. Goodman, H. M., Olson, M. B. & Hall, B. D. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5453–5457.
10. Müller, F. & Clarkson, S. G. (1980) *Cell* **19**, 345–353.
11. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. (1980) *Cell* **19**, 13–25.
12. Bogenhagen, D. F., Sakonju, S. & Brown, D. D. (1980) *Cell* **19**, 27–35.
13. Sollner-Webb, B. & Reeder, R. H. (1979) *Cell* **18**, 485–499.
14. Valenzuela, P., Bell, G. I., Venegas, A., Sewell, E. T., Masiarz, F. R., DeGennaro, L. J., Weinberg, F. & Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 8126–8135.
15. Klemenz, R. & Geiduschek, E. P. (1980) *Nucleic Acids Res.* **8**, 2679–2689.
16. Urano, Y., Kominami, R., Mishima, Y. & Muramatsu, M. (1980) *Nucleic Acids Res.* **8**, 6043–6058.
17. Miller, K. G. & Sollner-Webb, B. (1981) *Cell* **27**, 165–174.
18. Bach, R., Grummt, I. & Allet, B. (1981) *Nucleic Acids Res.* **9**, 1559–1569.
19. Long, E. O., Rebbert, M. L. & Dawid, I. B. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1513–1517.
20. Mishima, Y., Yamamoto, O., Kominami, R. & Muramatsu, M. (1981) *Nucleic Acids Res.* **9**, 6773–6785.
21. Financsek, I., Mizumoto, K. & Muramatsu, M. (1982) *Gene*, in press.
22. Mizumoto, K. & Lipmann, F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4961–4965.
23. Muramatsu, M. (1973) *Methods Cell Biol.* **3**, 23–51.
24. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12**, 721–732.
25. Fujimoto, M., Fujiyama, K., Kuninaka, A. & Yoshino, H. (1974) *Agric. Biol. Chem.* **38**, 2141–2147.
26. Green, M. R. & Roeder, R. G. (1980) *Cell* **22**, 231–242.
27. McDonnel, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119–146.
28. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
29. Wilson, G. N., Hollar, B. A., Waterson, J. R. & Schmickel, R. D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5367–5371.
30. Colman, A., Byers, M. J., Primrose, S. B. & Lyons, A. (1978) *Eur. J. Biochem.* **91**, 303–310.
31. Wellauer, P. K. & Dawid, I. B. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2827–2831.
32. Wellauer, P. K. & Dawid, I. B. (1979) *J. Mol. Biol.* **128**, 289–303.
33. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
34. Robertson, J. A. & Maden, B. E. H. (1973) *Biochim. Biophys. Acta* **331**, 61–70.
35. Tomizawa, J., Ohmori, H. & Bird, R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1865–1869.
36. Reeder, R. H., Sollner-Webb, B. & Wahn, H. L. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5402–5406.
37. Nikolaev, N., Georgiev, O. I., Venkov, P. V. & Hadjiolov, A. A. (1979) *J. Mol. Biol.* **127**, 297–308.
38. Oxender, D. L., Surawski, G. & Yanofsky, C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5524–5528.
39. Johnston, H. M., Barnes, W. M., Chumley, F. G., Bossi, L. & Roth, F. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 508–512.
40. Horinouch, S. & Weisblum, B. (1981) *Mol. Gen. Genet.* **182**, 341–348.
41. Grummt, I. (1981) *Proc. Natl. Acad. Sci.* USA **78**, 727–731.
42. Farabaugh, P. J. & Miller, J. H. (1978) *J. Mol. Biol.* **126**, 847–863.
43. Kataoka, T., Miyata, T. & Honjo, T. (1981) *Cell* **23**, 357–368.