Structural Determinant of the Species-Specific Transcription of the Mouse rRNA Gene Promoter

GEZA SAFRANY,[†] NOBUYUKI TANAKA, TOSHIMITSU KISHIMOTO, YOSHIAKI ISHIKAWA, HIROYUKI KATO, RYO KOMINAMI,[‡] and MASAMI MURAMATSU^{*}

Department of Biochemistry, The University of Tokyo Faculty of Medicine, Hongo, Bunkyo-ku, Tokyo 113, Japan

Received 23 June 1988/Accepted 2 September 1988

Mammalian ribosomal DNA (rDNA) transcription has a certain species specificity such that, both in vivo and in vitro, human rDNA cannot be transcribed by mouse machinery and vice versa. This is due to a species-dependent transcription factor, TFID (Y. Mishima, I. Financsek, R. Kominami, and M. Muramatsu, Nucleic Acids Res. 10:6659–6670, 1982). On the basis of the information obtained from 5' and 3' substitution mutants, we prepared a chimeric gene in which the mouse sequence from positions -32 to -14 was inserted into the corresponding location of the human rDNA promoter. The chimeric gene could be transcribed by mouse extracts nearly as efficiently as the wild-type mouse promoter. The chimeric gene could also sequester transcription factor TFID at an efficiency similar to that for the mouse promoter. Partially purified mouse TFID that could not protect the human rDNA promoter against DNase I produced a clear footprint on this chimeric gene that was similar to that on mouse rDNA promoter. The basic structure of the mouse rDNA core promoter is discussed in relation to the interaction with TFID.

RNA synthesis by RNA polymerase I enzyme has a special characteristic; i.e., a certain species specificity such that the human extract cannot substitute for the mouse factor in transcription of mouse ribosomal DNA (rDNA) and vice versa (10, 24, 28; for a recent review, see reference 34). This principle appears to hold true in vivo, since mousehuman hybrid cell lines often produce only one rRNA species from the predominant chromosome group, even if the cell line retains nucleolar organizers of the minor chromosome group (5, 27, 32; our unpublished observation), and the capability is reflected in the transcription ability of their cell extracts in vitro (31). Although the mouse extract could transcribe rat rDNA (15, 28) and the monkey factor could substitute for the human factor in human rDNA transcription (22), it appeared that any polymerase I machinery generally could transcribe only a closely related rDNA. More recently, however, Culotta et al. (4) have shown that mouse extract can start correct transcription on Xenopus rDNA, indicating that under certain conditions this specificity may not be so rigorous as was first postulated. Nevertheless, the coevolution of the rDNA and the polymerase I machinery offers a unique opportunity in which the mode of interaction between a gene and the transcription apparatus could be studied (6, 36).

The determination of the sequences that define the species specificity is of special biological interest, because it is to this region of the rDNA that a crucial transcription factor binds first to form a stable preinitiation complex (1, 14, 37, 39). We have already shown that this factor is fractionated into the high-salt (1 M KCl) eluate on a phosphocellulose column and designated as TFID (14, 28). A similar factor was reported by others (2, 17, 24, 37). Since the in vitro transcription system requires only a part of the promoter

region designated as a core promoter for its rather high efficiency under certain standard conditions and since it requires the homologous TFID fraction for its activity, at least the major sequences that define the species-specific promotion of transcription initiation will be located within the core promoter.

Studies with various deletion and substitution mutants have led to a general conclusion that the promoter may be divided at least into two portions: one required for minimal accurate and efficient transcription initiation and the other augmenting efficiency further (9, 12, 13, 16, 18, 20, 21, 24, 35, 40, 41). The former promoter element lies proximal to the transcription start site and is called a core promoter. For the mouse rDNA, this region was shown to be around nucleotides -40 to +23 by deletion experiments (3, 25, 41). Further upstream sequences are required for higher transcription efficiency, which is apparently revealed under certain stringent conditions in vitro (26). Comparison of human (7) and mouse (38) promoter regions reveals three well-conserved sequences in the core region (Fig. 1A); i.e., from positions -38 to -33 (ATCTTT, designated here distal conserved sequences), from positions -20 to -12 (T_T^ATTGG_G^ACC, proximal conserved sequence), and from positions -1 to +18 ($T_G^ACTGACACGCTGTCCT_C^TT$, start site conserved sequence) (by mouse coordinates and with human substitutions underneath). Homologous nucleotides in the same positions are presumably not responsible for the species specificity. Rather they are likely to function as general non-species-specific signals for promotion. The sequences that determine the species-specific recognition by the species-dependent initiation factor should be sought in the nonhomologous core promoter regions (from positions -32to -21 and -11 to -2 in the mouse) and at the nonhomologous nucleotides in the conserved sequences (positions -19, -14, +1 and +17).

Region that determines the mouse species specificity. To determine the exact location of species-specific promoter element of the mouse rRNA gene, a series of mouse-human chimeric genes was constructed in which parts of the mouse

^{*} Corresponding author.

[†] Present address: Department of Biochemistry, Frederic Joliot-Curie National Research Institute for Radiobiology and Radiohygiene, Budapest, Hungary.

[‡] Present address: 1st Department of Biochemistry, Niigata University School of Medicine, Niigata 951, Japan.



ribosomal promoter were replaced with corresponding human promoter sequences.

To prepare a vector for mouse-human chimeric genes, first a hybrid plasmid was constructed from a human 5'-deletion mutant (a generous gift of R. Tjian [21]) containing sequences from position -155 to the SalI site (position +697) in the coding region and from pMrSA plasmid harboring mouse ribosomal promoter sequences from positions +133 (AvaI site) to -167 (SalI site) between the EcoRI and SalI sites of pBR322 (Fig. 1). In both cases, the direction of transcription was counterclockwise. In pMrSA, an A-to-G point mutation was introduced at nucleotide +23 by oligonucleotide-directed mutagenesis (Kishimoto et al., manuscript in preparation) to obtain a new HpaI restriction site (+23 GTTAAC +28). In the human deletion mutant, the SalI site was changed to an HpaI site by the ligation of an HpaI linker. The shorter PstI-HpaI fragment of the human plasmid containing pBR sequences was replaced with the shorter PstI-HpaI fragment of mouse plasmid carrying pBR322 sequences from the PstI site to the EcoRI site and mouse sequences from position +133 to the newly prepared HpaI site (position +26). Then this hybrid plasmid was cleaved with HpaI and XmaI (cut at position -53 in the human sequence), and the human sequences between these two sites were replaced with a short oligonucleotide linker consisting of BamHI and HindIII recognition sites, reconstituting the HpaI and XmaI sites. To construct mouse-human chimeric genes, the hybrid vector was digested with HpaI and XmaI, and seven synthetic oligonucleotides with the desired mouse-human sequences were inserted into the vector in place of the linker. Because synthetic oligonucleotides were used for the construction of chimeric genes, we could prepare any kind of genes with any mouse/human sequence ratio and order between human position -53 and mouse position +26 nucleotides without any linker insertion or point mutation in the chimeric gene. Even the formerly prepared A-to-G mutation at nucleotide +23 was eliminated by restriction enzyme digestion during the hybrid gene preparation. Synthetic oligonucleotides were hybridized and ligated (11) to the hybrid vector; then the ligation mixtures were digested with BamHI to destroy uncut parent plasmid, and a sample was used to transform Escherichia coli HB101.

The in vitro promoter activity of each chimeric gene was determined by a runoff assay (14, 29) after the plasmid was truncated. Gradual expansion of the human sequence into the mouse promoter, from upstream toward the initiation



FIG. 2. In vitro transcription assay and sequestration analysis of the chimera pM-38/-12 (mouse-human) by FM3A S-100 extract. (A) Runoff transcripts from 0.2 µg of pM-38/-12 linearized with HindIII (lane 2) and from 0.2 µg of wild-type pMrSP harboring mouse rDNA from positions -167 (SalI) to +291 (PvuII), digested with AccI (lane 3). Every sample contained 0.1 μ g of control wild-type template (pMrSP digested with PvuII). Lane 1 shows an EcoRI-HinfI-digested pBR322 marker. Arrows indicate the faithful transcripts; the lowest is the control. (B) Sequestration assay of pM-38/-12 and wild-type pMrSP templates. Autoradiographs show RNA products from reporter template (pMrSP linearized with PvuII; 291 nucleotides) and wild-type pMrSP (linearized with AccI; 470 nucleotides; lanes 4 to 6) competitors and from pM-38/-12 (linearized with HindIII; 592 nucleotides; lanes 1 to 3). In vitro transcription (14, 29) and the sequestration assay (30) were carried out as described previously. Competitors were used at 0 µg (lanes 1 and 4), 0.2 µg (lanes 2 and 5), and 0.4 μ g (lanes 3 and 6).

point, established the 5' border of the mouse species-specific element at nucleotide -32, which is just downstream of the distal conserved sequence of the mouse core promoter. When the human sequence was expanded from the initiation point toward upstream, the 3' border of the mouse element was found at nucleotide -14, which lies inside the proximal conserved sequence (detailed data will be published elsewhere). The results suggest that only a rather short region (19 nucleotides) between the above-mentioned positions defines the mouse species-specific initiation.

To test this hypothesis a chimeric gene, pM-38/-12, was constructed, in which mouse nucleotides -32 to -14 were inserted in the complete human rDNA surrounding (Fig. 1B). To prepare pM-38/-12, the original human mutant with a 5' deletion mutation at position -155 was digested with *XmaI* and *SaII* (cut at the junction of human rDNA and pBR322 sequences), human sequences from nucleotides -53 to +697 were eliminated, and eight synthetic oligonucleotides were ligated to this vector, resulting in a chimeric gene with a human sequence from positions -155 to +40 but in which human nucleotides from positions -35 to -14 (human coordinates) were replaced by the corresponding mouse sequence from positions -32 to -14 (Fig. 1C). Nucleotides -12 and -13 and nucleotides -33 to -38 (distal conserved sequence) were conserved between the two species.

In good agreement with our hypothesis, pM-38/-12 was transcribed in vitro at about 90% of the wild-type level (Fig. 2A) by the mouse cell extract. Competition assays with this chimeric gene showed that it also sequestered transcription initiation factor(s) at least to three-fourths of the wild-type level (Fig. 2B).

Binding of transcription factor TFID to the chimeric gene. Recently, we found that partially purified mouse and human TFID fractions could protect both the core and an upstream region of the homologous rDNA promoter but not those of the heterologous promoter from DNase I digestion (Tanaka et al., unpublished observation). We therefore examined the footprints of a TFID fraction on the chimeric gene pM-38/-12.

TFID was partially purified by phosphocellulose chroma-

tography (1 M KCl fraction after a 0.6 M KCl wash) followed by DEAE-cellulose column chromatography (0.35 M KCl fraction after a 0.1 M KCl wash), both in a buffer system described previously (14). DNase I footprinting was carried out as described by Galas and Schmitz (8). Plasmid pM-38/ -12 was cleaved with *Sal*I at nucleotide +40, and mouse wild-type pMrSP was cleaved with *Ava*I at nucleotide +130. Binding reactions were carried out under the same conditions as in vitro transcription (detailed procedures will be published elsewhere).

The human rDNA in which only the core region (-14 to)-38) had been replaced with the mouse sequence could be protected by partially purified mouse TFID fraction in a manner remarkably similar to that of the wild-type mouse rDNA (Fig. 3A, lanes 1 through 4; Fig. 3B). Protection at the core region (positions -12 to -40) and at the upstream region (positions -40 to -140) was also seen. Enhanced cleavage at position -11 was as in the wild type. Cleavage upstream between positions -90 and -100, although shifted slightly in position from that of the wild type, was also characteristic of the factor binding. However, enhanced cleavage at around position -50, which is found in the wild-type mouse promoter, was not apparent in this chimera. By contrast, the human TFID fraction which could protect the human rDNA in a similar but not identical manner (data not shown) apparently failed to protect any part of the whole promoter region of this chimera (Fig. 3A, lanes 5 through 8).

The results are consistent with the observation that this chimera can be transcribed efficiently by mouse FM3A cell extract (Fig. 2) but not by human HeLa cell extract (data not shown). Although the TFID fraction may contain some other copurified protein(s) that might confer species specificity separately and also interact with upstream sequences, these data together with previous observations (14, 30) strongly suggest that the stretch from nucleotides -14 to -38 in mouse rDNA is required for specific binding of mouse TFID and thereby determines the mouse-specific rDNA transcription.

Structural features of the mouse rDNA core promoter. Our findings on the structural features of the mouse rDNA core promoter may be summarized as follows. The mouse core promoter, the basic unit of minimal transcription initiation, is located between nucleotides -38 and +18. Previous deletion experiments have shown that the start site conserved sequence is not required to fix the start point at nucleotide +1, but it can increase the promoter activity by 1 order of magnitude (41). At least for mouse cell extract, the replacement of this region with that of humans did not affect promoter activity, although the base substitutions between them are only two (position +1, A to G; position +17, T to C).

Mouse transcription factor TFID recognizes a sequence from positions -38 to -12, including the distal and proximal conserved sequences, and binds to this region. This is supported by the footprints as well as competition data described above. The conserved guanine residue at nucleotide -16 has an important role in this binding, as shown by previous studies (15, 30, 33; Tanaka et al., unpublished observation). On the other hand, conserved nucleotides -1and -7 are presumably crucial for the relatively speciesindependent recognition processes of rDNA such as the binding and functioning of polymerase I enzyme or other factor(s) (3, 14, 30). The role of upstream sequences was not clarified in this study, except that some protein in the TFID fraction bound with the region of nucleotides -40 to -140 in a core sequence-dependent manner.

The interplay of an upstream sequence called the up-



FIG. 3. DNase I footprint analysis. (A) DNase I footprinting of the chimeric gene pM-38/-12 with partially purified TFID fraction. Patterns of the coding strand are shown. Lanes: 1 and 5, with no protein; 2 through 4, with 2, 4, and 6 μ g, respectively, of mouse TFID fraction; 6 through 8, with 2, 4, and 6 μ g, respectively, of human TFID fraction. The protected regions are indicated by brackets, and enhanced cleavages are indicated by arrowheads. Numbers along the borders indicate the nucleotide positions relative to the transcription start site. M, Maxam-Gilbert (23) sequence ladders of G (left) and G+A (right) as size markers. (B) DNase I footprinting of mouse wild-type rDNA promoter with mouse TFID fraction. Lanes: 1, with no protein; 2 through 4, with 2, 4, and 6 μ g, respectively, of mouse TFID fraction. For details, see text.

stream control element with the core promoter element in the transcription of human rDNA has been reported by Learned et al. (20). They found that a species-specific factor, SL1, purified from HeLa cells could not bind with the core promoter by itself (19) but bound to a much further upstream region (positions -120 to -165) in the presence of another factor, UBF1, contained in the polymerase I fraction, providing species-specific augmentation of transcription to human rDNA (20). SL1, however, conferred species specificity even in the absence of the upstream control element, suggesting a dual function for this factor (19). TFID does bind to the core promoter and program the initiation even in the absence of upstream sequences (14, 24, 30). Whether TFID itself is responsible for the species specificity of this fraction, however, remains to be determined. Since the footprints indicate the binding of protein in TFID fraction also to the upstream sequence, there are at least two possibilities: one, that TFID binds with both core and upstream sequences, and the other, that this fraction contains upstream binding proteins in addition to TFID.

In any event, this study has clearly demonstrated that the relatively narrow region of the core promoter determines at least a significant portion of species specificity of the mouse rDNA promoter against human rDNA via binding of TFID. Further participation of the upstream sequences may also be present, although the relationship of TFID to UBF1 and SL1 is not clear at present. Reconstitution experiments with more purified factors are certainly required for further elucidation of interaction of transcription factors and the rDNA promoter.

LITERATURE CITED

- 1. Cizewski, V., and B. Sollner-Webb. 1983. A stable transcription complex directs mouse ribosomal RNA synthesis by RNA polymerase I. Nucleic Acids Res. 11:7043–7056.
- Clos, J., D. Buttgereit, and I. Grummt. 1986. A purified transcription factor (TIF-IB) binds to essential sequences of the mouse rDNA promoter. Proc. Natl. Acad. Sci. USA 83:604–608.
- Clos, J., A. Norman, A. Öhrlein, and I. Grummt. 1986. The core promoter of mouse rDNA consists of two functionally distinct domains. Nucleic Acids Res. 14:7581–7595.
- Culotta, V. C., J. K. Wilkinson, and B. Sollner-Webb. 1987. Mouse and frog violate the paradigm of species-specific transcription of ribosomal RNA genes. Proc. Natl. Acad. Sci. USA 84:7498-7502.
- Croce, C. M., A. Talavera, C. Bacilico, and O. J. Miller. 1977. Suppression of production of mouse 28S ribosomal RNA in mouse-human hybrids segregating mouse chromosomes. Proc. Natl. Acad. Sci. USA 74:694-697.
- 6. Dover, G. A., and R. B. Flavell. 1984. Molecular coevolution: DNA divergence and the maintenance of function. Cell 38:622– 623.
- Financsek, I., K. Mizumoto, Y. Mishima, and M. Muramatsu. 1982. Human ribosomal RNA gene: nucleotide sequence of the transcription initiation region and comparison of three mammalian genes. Proc. Natl. Acad. Sci. USA 79:3092–3096.
- Galas, D., and A. Schmitz. 1978. DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. 5:3157–3170.
- Grummt, I. 1982. Nucleotide sequence requirements for specific initiations of transcription by RNA polymerase I. Proc. Natl. Acad. Sci. USA 79:6908–6911.
- Grummt, I., E. Roth, and M. R. Paule. 1982. Ribosomal RNA transcription *in vitro* is species-specific. Nature (London) 296: 173–174.
- Grundstrom, T., M. W. Zenke, M. Wintzerith, H. W. D. Matthes, A. Staub, and P. Chambon. 1985. Oligonucleotidedirected mutagenesis by microscale shot-gun gene synthesis. Nucleic Acids Res. 13:3305-3316.
- Haltiner, M. M., S. T. Smale, and R. Tjian. 1986. Two distinct promoter elements in the human rRNA gene identified by linker scanning mutagenesis. Mol. Cell. Biol. 6:227–235.
- Iida, C. T., P. Kownin, and M. R. Paule. 1985. Ribosomal RNA transcription: proteins and DNA sequences involved in preinitiation complex formation. Proc. Natl. Acad. Sci. USA 82:1668– 1672.
- Kato, H., M. Nagamine, R. Kominami, and M. Muramatsu. 1986. Formation of the transcription initiation complex on mammalian rRNA. Mol. Cell. Biol. 6:3418–3427.
- Kishimoto, T., M. Nagamine, T. Sasaki, N. Takakusa, T. Miwa, R. Kominami, and M. Muramatsu. 1985. Presence of a limited number of essential nucleotides in the promoter region of mouse ribosomal RNA gene. Nucleic Acids Res. 13:3515–3532.
- Kohorn, B. D., and P. M. M. Rae. 1982. Accurate transcription of truncated ribosomal DNA templates in *Drosophila* cell-free system. Proc. Natl. Acad. Sci. USA 79:1501–1505.
- 17. Kownin, P., E. Bateman, and M. R. Paule. 1987. Eukaryotic RNA polymerase I promoter binding is directed by protein contacts with transcription initiation factor and is DNA sequence-independent. Cell 50:693-699.
- Kownin, P., C. T. Iida, S. Brown-Shimer, and M. R. Paule. 1985. The ribosomal RNA promoter of *Acanthamoeba castellani* determined by transcription in a cell-free system. Nucleic Acids Res. 13:6237-6247.
- Learned, R. M., S. Codes, and R. Tjian. 1985. Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. Mol. Cell. Biol. 5: 1358–1369.
- Learned, R. M., T. K. Learned, M. M. Haltiner, and R. T. Tjian. 1986. Human rRNA transcription is modulated by the coordinate binding of two factors to an upstream control element. Cell 45:847–857.

- Learned, R. M., S. T. Smale, M. M. Haltiner, and R. Tjian. 1983. Regulation of human ribosomal RNA transcription. Proc. Natl. Acad. Sci. USA 80:3558–3562.
- Learned, R. M., and R. Tjian. 1982. In vitro transcription of human ribosomal RNA genes by RNA polymerase I. J. Mol. Appl. Genet. 1:575-585.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499-560.
- Miesfeld, R., and N. Arnheim. 1984. Species-specific rDNA transcription is due to promoter-specific binding factors. Mol. Cell. Biol. 4:221-227.
- Miller, K. G., and B. Sollner-Webb. 1981. Transcription of mouse rRNA genes by RNA polymelase I: *in vitro* and *in vivo* initiating and processing sites. Cell 27:165–174.
- Miller, K. G., J. Tower, and B. Sollner-Webb. 1985. A complex control region of the mouse rRNA gene directs accurate initiation by RNA polymerase I. Mol. Cell. Biol. 5:554–562.
- Miller, O. J., D. A. Miller, V. G. Dev, R. Tantravahi, and C. M. Croce. 1976. Expression of human and suppression of mouse nucleolus organizer activity in mouse-human somatic cell hybrids. Proc. Natl. Acad. Sci. USA 73:4531-4535.
- Mishima, Y., I. Financsek, R. Kominami, and M. Muramatsu. 1982. Fractionation and reconstitution of factors required for accurate transcription of mammalian rRNA genes: identification of a species-dependent factor. Nucleic Acids Res. 10:6659–6670.
- Mishima, Y., O. Yamamoto, R. Kominami, and M. Muramatsu. 1981. In vitro transcription of cloned mouse ribosomal RNA gene. Nucleic Acids Res. 9:6773–6785.
- Nagamine, M., T. Kishimoto, J. Aono, H. Kato, R. Kominami, and M. Muramatsu. 1987. Sequestration analysis for RNA polymerase I transcription factors with various deletion and point mutations reveals different functional regions of the mouse rRNA gene promoter. Mol. Cell. Biol. 7:1486–1495.
- Onishi, T., C. Berglund, and R. H. Reeder. 1984. On the mechanism of nucleolar dominance in mouse-human somatic cell hybrids. Proc. Natl. Acad. Sci. USA 81:484–487.
- Perry, R. P., D. E. Kelley, U. Schibler, K. Huebner, and C. M. Croce. 1979. Selective suppression of the transcription of ribosomal genes in mouse-human hybrid cells. J. Cell. Physiol. 98: 553-559.
- 33. Skinner, J. A., A. Öhrlein, and I. Grummt. 1984. In vitro mutagenesis and transcriptional analysis of a mouse ribosomal promoter element. Proc. Natl. Acad. Sci. USA 81:2137-2141.
- Soliner-Webb, B., and J. Tower. 1986. Transcription of cloned ribosomal RNA genes. Annu. Rev. Biochem. 55:801-803.
- Sollner-Webb, B., J. A. K. Wilkinson, J. Roan, and R. H. Reeder. 1983. Nested control regions promote Xenopus ribosomal RNA synthesis by RNA polymerase I. Cell 35:199–206.
- 36. Tautz, D., C. Tautz, D. Webb, and G. A. Dover. 1987. Evolutionary divergence of promoters and spacers in the rDNA family of four Drosophila species. Implications for molecular coevolution in multigene families. J. Mol. Biol. 195:525–542.
- Tower, J., V. C. Culotta, and B. Sollner-Webb. 1986. Factors and nucleotide sequences that direct ribosomal DNA transcription and their relationship to the stable transcription complex. Mol. Cell. Biol. 6:3451-3462.
- Urano, Y., R. Kominami, Y. Mishima, and M. Muramatsu. 1980. The nucleotide sequence of the putative transcription initiation site of a cloned ribosomal RNA gene. Nucleic Acids Res. 8: 6043-6058.
- Wandelt, C., and I. Grummt. 1983. Formation of stable preinitiation complexes is a prerequisite for ribosomal DNA transcription *in vitro*. Nucleic Acids Res. 11:3795–3809.
- Windle, J. J., and B. Sollner-Webb. 1986. Two distant and precisely positioned domains promote transcription of *Xenopus laevis* rRNA genes: analysis with linker-scanning mutants. Mol. Cell. Biol. 6:4585–4593.
- Yamamoto, O., N. Takakusa, Y. Mishima, R. Kominami, and M. Muramatsu. 1984. Determination of the promoter region of mouse ribosomal RNA gene by an *in vitro* transcription system. Proc. Natl. Acad. Sci. USA 81:299–303.