HOMOLOGY OF THE 5′-TERMINAL SEQUENCE OF 28 S rRNA OF MOUSE WITH YEAST AND *XENOPUS*

Implication for the secondary structure of the 5.8 S–28 S RNA complex

Bernard MICHOT, Jean-Pierre BACHELLERIE*, Françoise RAYNAL and Marie-Hélène RENALIER
Centre de Recherches de Biochimie et de Génétique Cellulaires du CNRS, 118, route de Narbonne, 31062 Toulouse Cédex, France

Received 25 February 1982

1. Introduction

Ribosomal RNA sequences have undergone little divergence during evolution. Indirect evidence, involving heterologous hybridizations [1,2] showed that strongly conserved regions do exist within rRNA molecules, which could be of particular functional importance. Determination of the complete nucleotide sequence of small ribosomal subunit RNA of yeast *Saccharomyces cerevisiae* [3] and amphibian *Xenopus laevis* [4] have revealed extensive regions of high homology interspersed with divergent tracts and a very strong phylogenetic stability has been observed for the 3′-terminal region of a variety of small rRNAs [5–7]. The sequence of 5.8 S rRNA, that is stably associated with eukaryotic large rRNA, has also been highly conserved, with complete identity between mammals and >70% homology between yeast and mammals [8]. Sequence data available so far on eukaryotic large rRNAs are scarce. Yeast 26 S RNA [9,10] is the only major species with a known complete sequence.

Our interest in the rRNA maturation process and identification of recognition signals for processing endonucleases in higher cells led us to determine the primary sequence of ribosomal transcribed spacers and adjacent regions of mature rRNAs in mouse (in preparation).

Here, we have mapped at sequence resolution the 5′-terminus of 28 S rRNA within a 144 nucleotide segment of cloned mouse rDNA. The sequence of this region has been compared with the other eukaryotic homologs available so far, yeast [11,12] and *X. laevis* [13]. A high divergence is observed for the transcribed spacer immediately upstream 28 S 5′-terminus, although a common feature could be recognized in mouse and *X. laevis*. On the contrary, the first 100 nucleotides of 28 S RNA are highly conserved. This region has been proposed to participate directly in the 5.8 S–large-rRNA complex in yeast [9], due to significant base complementarity with the 3′-terminal sequence of 5.8 S RNA. We show here that an homologous junction complex, even more energetically favourable than yeast's, can be proposed for mouse rRNAs.

2. Methods

2.1. *Isolation of DNA*

The 3.7 kilobase *Eco*RI–*Bam*HI fragment of mouse ribosomal DNA encompassing the 5′-terminus of 28 S RNA was inserted into the *Eco*RI + *Bam*HI cleaved plasmid pBR 322, giving rise to a pMEB3 recombinant plasmid. Plasmid DNA was isolated from *E. coli* HB101 by the clear lysate method [14], followed by CsCl equilibrium ultracentrifugation. Supercoiled closed circular plasmid DNA was further purified by ultracentrifugation on sucrose gradients. As a prerequisite for the complete sequence determination of mouse ribosomal transcribed spacers (in preparation), we have established a detailed restriction map of the cloned *Eco*RI–*Bam*HI region through multiple digestions and partial digestions of 5′-end labelled fragments.

2.2. *Sequencing procedures*

The 1060 basepair *Sma*I '1', encompassing 28 S

* To whom correspondence should be addressed

Published by Elsevier Biomedical Press
00145793/82/0000–0000/$02.75

RNA 5′-terminus, was purified from a total plasmid DNA *Sma*I digest by electrophoresis on a 5% acrylamide gel [15]. After redigestion by *Hae*II (see fig.1c), dephosphorylation by alkaline phosphatase and 5′-^{32}P-end labelling, the 300 basepair fragment containing 28 S 5′-terminus was purified by electrophoresis on an 8% acrylamide gel and strand separation was achieved by electrophoresis on a 6% acrylamide gel (acrylamide/bis = 50/1). Identification of coding strand and mapping of 5′-terminus at sequence resolution were carried out by S1 nuclease experiments. DNA extraction, purification and sequencing procedures were as in [15], except for additional DE52-cellulose chromatography immediately before chemical DNA sequencing.

2.3. *Purification of rRNA*

RNA from mouse liver ribosomes was extracted by phenol–SDS treatment at room temperature and 28 S rRNA was purified by two successive sedimentation runs in (17–34%) linear sucrose gradients in 5 mM Tris (pH 7.4), 20 mM NaCl, 10 mM EDTA, 0.1% SDS. Before the last sedimentation, RNA was further deproteinized by proteinase K followed by phenol–chloroform extraction; it was heat-denatured (65°C, 5 min) immediately before loading onto sucrose gradient.

2.4. *S1 nuclease mapping*

This was done essentially as in [16,17]. Double-stranded or single-stranded [5′-^{32}P]DNA fragment (0.1–0.5 pmol) in 90% formamide was heated for 3 min at 90°C, quickly chilled on ice and incubated with 0.4–2 μg purified 28 S mouse rRNA (or an equal amount of *E. coli* tRNA for control) in 80% formamide, 2 SSC for 10 min at 65°C followed by 2 h at 60°C. After ethanol precipitation and redissolution in 0.2 M NaCl, 50 mM Na-acetate (pH 4.5), 1 mM $ZnCl_2$, samples were incubated for 30 min at 37°C with 250 units nuclease S1 (Sigma), in the presence of 2 μg *E. coli* tRNA. Digestion was stopped by adding 5 mM EDTA. After ethanol precipitation with 3 μg *E. coli* tRNA carrier, S1 nuclease-resistant [^{32}P]DNA fragments were analyzed onto 8% acrylamide/7 M urea sequencing gels [15] after redissolution in 80% formamide, 10 mM NaOH, 1 mM EDTA and heat-denaturation for 1 min at 90°C.

3. Results and discussion

3.1. *Mapping 28 S RNA 5′-terminus*

The strategy (fig.1), exploited the single-strand specificity of nuclease S1 to identify the gene junction by analyzing potential hybrids between 28 S rRNA and 5′-end, labelled DNA fragments. Each *Sma*I fragment (fig.1b) of recombinant plasmid pMEB3 was assayed for 28 S RNA protection. For DNA fragments located on the 'right' side of *Sma*I '1', fully protected DNA bands were detected while no protection from S1 nuclease was observed for DNA fragments located on the 'left' side of *Sma*I '1'. Conversely, a partial protection (~900 nucleotide long) was found for *Sma*I '1' fragment. The more precise location of 28 S 5′-end within this fragment was determined by S1 nuclease experiments done on single-stranded shorter DNA segments (fig.1e) obtained after *Hae*II digestion and strand-separation [15].

After hybridization with 28 S rRNA and digestion with nuclease S1, the coding strand of the *Sma*I–*Hae*II subfragment, labelled at the *Hae*II site (fig.1c), was sized in parallel with the chemical DNA sequencing ladder obtained from the same full length subfragment. Two adjacent bands of similar intensity were observed with mobility corresponding to 111 and 112 nucleotides, after a single nucleotide correction [18].

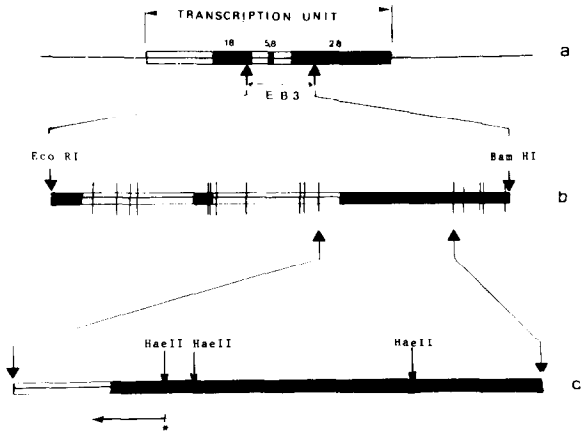

Fig.1. Restriction map of cloned mouse rDNA fragments encompassing 28 S RNA 5′-terminus: (a) location of the 3.7 kilobase *Eco*RI–*Bam*HI region within mouse ribosomal transcription unit; (b) *Sma*I restriction map of *Eco*RI–*Bam*HI fragment cloned into pMEB3; (c) expanded map of the 1060 basepair *Sma*I '1' fragment containing 28 S RNA 5′-terminus. The position of the 5′-end labelled fragment used for sequencing the 5′-terminal region of 28 S RNA is shown by the arrow.

We cannot decide at this moment if the presence of 2 bands is an inherent feature of the S1 nuclease experiment or is indicative of a terminal heterogeneity of 28 S rRNA, in line with a greater accessibility to ribonucleases of this unbasepaired 5′-terminus (see fig.4). These positions have been numbered, respectively, +1 and −1 in the sequence shown in fig.3.

3.2. *Conservation of 28 S rRNA terminal sequence*

The alignment of mouse 28 S rRNA sequence (determined as in fig.2) with its yeast and *X. laevis* counterparts shows a striking conservation despite the large evolutionary distance (fig.3). A 80% homology is observed between yeast and mouse for the 100 first nucleotides and a 91% homology in the same region is found between amphibian *X. laevis* and mouse. Even when considering such a short region, it appears that the few mutations that have occurred are not randomly distributed but are rather concentrated in clusters interspersed by perfectly conserved tracts. Most of the mutations between *X. laevis* and mouse are a subset of those that have occurred between yeast and mouse.

3.3. *Divergence of adjacent spacer*

The transcribed spacer immediately upstream 28 S rRNA could possess common sequence features possibly involved in the recognition process for maturation cleavages. The alignment of fig.4 shows an extensive divergence of these regions in contrast with the high conservation of 28 S rRNA 5′-terminal. However, common features can be found between *X. laevis* and mouse: the region −4 to −15 is an uninterrupted

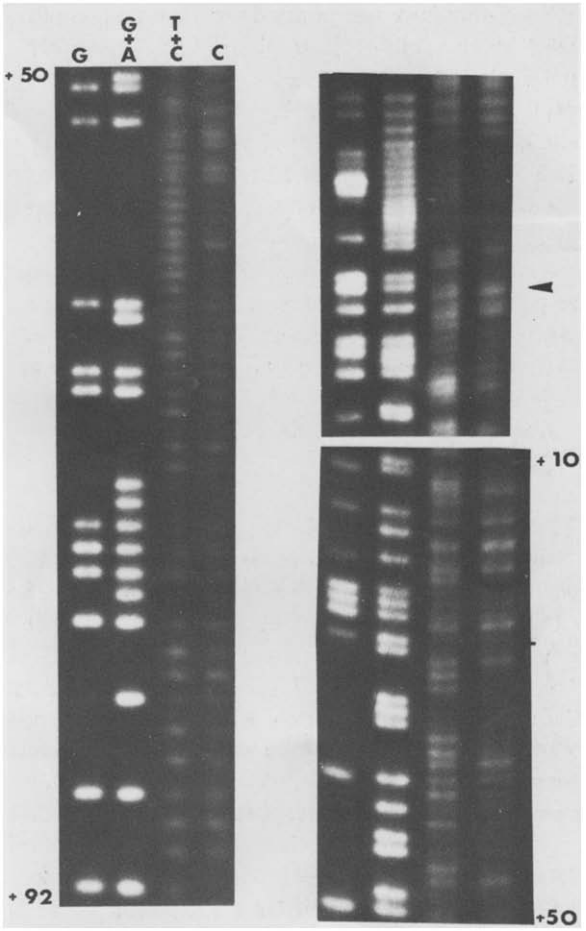

Fig.2. 8% sequencing gels of the mouse 28 S 5′-terminal region. The sequence is read on the coding strand from a *Hae*II–*Sma*I fragment 5′-^{32}P-labelled at *Hae*II site (see fig.1). The 28 S-transcribed spacer boundary is shown by an arrow.

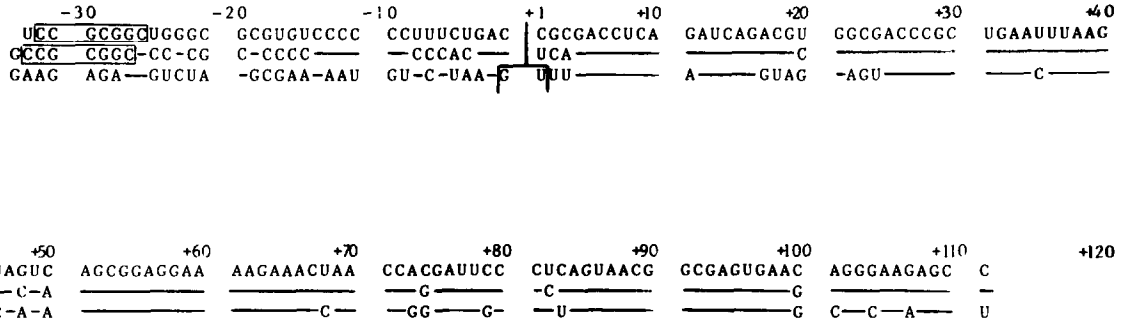

Fig.3. Sequence of the 5′-end of 28 S rRNA gene and beginning of the transcribed spacer of mouse. The sequence is numbered beginning at the 5′-terminus of 28 S RNA. Mouse sequence (top) is compared with yeast, *S. cerevisiae* [11] and *S. carlbergensis* [12], and *X. laevis* [13] homologous region. For yeast (bottom) and *X. laevis* (middle), the only written bases are those which are substituted in the mouse sequence.

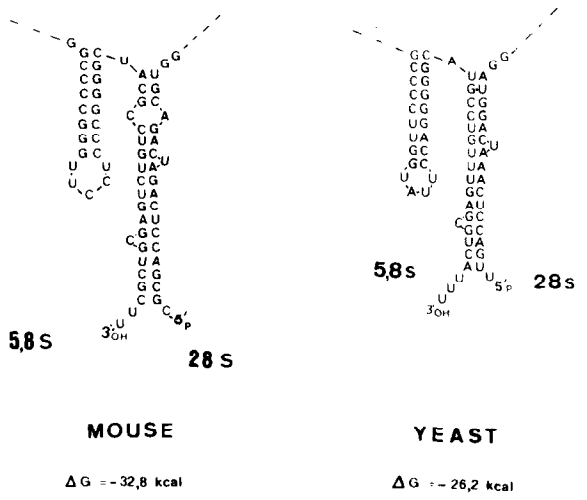

Fig.4. Model for the secondary structure of 5.8 S–28 S RNA complex. Free energy for the secondary structure involving the 3′-terminal region of 5.8 S RNA and the 5′-end of 28 S rRNA was determined as in [24]. The possible structure of the mouse complex is compared with its yeast (*S. carlbergensis*) counterpart [12].

pyrimide tract in both species and an identical 7 nucleotide sequence CCGCGGC is present in an almost identical position (boxed regions in fig.3) ~30 nucleotides upstream 28 S rRNA 5′-terminus. A complete sequence analysis of mouse internal-transcribed spacers will be described elsewhere (in preparation).

3.4. *5.8 S–28 S rRNAs junction complex*

Eukaryotic 5.8 S rRNA forms a specific complex with 28 S rRNA in the ribosome [19]. This RNA–RNA association is disrupted by treatment which suppress hydrogen bonding and can be restored in appropriate annealing conditions [20]. Direct experimental evidence have implicated the 3′-terminal 20–21 nucleotides of mouse 5.8 S RNA in the formation of an inter-molecular hairpin with 28 S rRNA [20]. Structural studies on isolated 5.8 S RNA [21,22] have led to propose basepairing interactions between this region and the 5′-end of 5.8 S RNA, in the absence of 28 S RNA. One may envision that a structural equilibrium between 2 kinds of interaction for the 3′-terminal region of 5.8 S RNA has a physiological significance, particularly during the conformational changes that elongating rRNA precursor must undergo during transcription and maturation. The complete sequence determination of yeast 26 S rRNA [9,10] has resulted in a secondary structure model [9] for the 5.8 S–large-rRNA junction complex, involving both 3′-terminal and 5′-terminal regions of 5.8 S RNA in basepairing interactions with the 5′-end and a more distal region (~400 nucleotides from 5′-terminus) of 26 S RNA, respectively. We have analyzed our sequence data in view of a potential conservation of the first of these two 5.8 S–28 S rRNA junctions, using the sequence for mouse 5.8 S in [23]. An homologous junction complex can be built for mouse rRNA (fig.4). It is important to observe that of the 13 mutations that have occurred between yeast and mouse in this interacting 5.8 S–28 S region, 12 correspond to compensatory base changes that result in an increased stability of the helix in mouse. It also seems noteworthy that the presence of 2 putative bulged nucleotides has been conserved at the same positions. As suggested with 5 S rRNA–L18 ribosomal protein in prokaryotes [25], such conserved bulged structure could have a key role in protein–RNA recognition.

Acknowledgements

The assistance of J. Feliu and M. Nicoloso in the isolation of plasmid DNA and ribosomal RNA was highly appreciated. The authors are grateful to Professor J. P. Zalta for his support. This investigation was financially supported by a grant from CNRS (ATP no. 4254), by a fellowship to B. M. from DGRST and by ADRC contract (no. 6107). Biohazards associated with the experiments were pre-examined by the French Control Committee.

References

[1] Gerbi, S. A. (1976) J. Mol. Biol. 106, 791–816.
[2] Gourse, R. L. and Gerbi, S. A. (1980) J. Mol. Biol. 140, 321–339.
[3] Rubstov, P. M., Musakhanov, M. M., Zakhariev, V. M., Krayev, A. S., Skryabin, K. G. and Bayev, A. A. (1980) Nucleic Acids Res. 8, 5779–5794.
[4] Salim, M. and Maden, B. E. H. (1981) Nature 291, 205–208.
[5] Hagenbüchle, O., Santer, M., Steitz, J. A. and Mans, R. J. (1978) Cell 13, 551–563.
[6] Jordan, B. R., Latil-Damotte, M. and Jourdan, R. (1980) FEBS Lett. 117, 227–231.
[7] Azad, A. A. and Deacon, N. J. (1980) Nucleic Acids Res. 8, 4365–4376.
[8] Erdmann, V. A. (1981) Nucleic Acids Res. 9/1, r25–r42.
[9] Veldman, G. M., Klootwijk, J., De Regt, V. C. H. F., Planta, R. J., Branlant, C., Krol, A. and Ebel, J. P. (1981) Nucleic Acids Res. 9, 6935–6952.

[10] Georgiev, O. I., Nikolaev, N., Hadjiolov, A. A., Skryabin, K. G., Zakharyev, V. M. and Bayev, A. A. (1981) Nucleic Acids Res. 9, 6953–6958.
[11] Bayev, A. A., Georgiev, O. I., Hadjiolov, A. A., Nikolaev, N., Skryabin, K. G. and Zakharyev, V. M. (1981) Nucleic Acids Res. 9, 789–799.
[12] Veldman, G. M., Klootwijk, J., Van Heerikhuisen, H. and Planta, R. J. (1981) Nucleic Acids Res. 9, 4847–4862.
[13] Hall, L. M. C. and Maden, B. E. H. (1980) Nucleic Acids Res. 8, 5993–6005.
[14] Goebel, W. and Bonewald, R. (1975) J. Bacteriol. 123, 658–665.
[15] Maxam, A. M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[16] Berk, A. J. and Sharp, P. A. (1977) Cell 12, 721–732.
[17] Bayev, A. A., Georgiev, O. I., Hadjiolov, A. A., Kermekchiev, M. B., Nikolaev, N., Skryabin, K. G. and Zakhariev, V. M. (1980) Nucleic Acids Res. 8, 4919–4926.
[18] Moss, T. and Birnstiel, M. L. (1979) Nucleic Acids Res. 6, 3733–3743.
[19] Pene, J. J., Knight, E. and Darnell, J. E. (1968) J. Mol. Biol. 33, 609–623.
[20] Pace, N. R., Walker, T. A. and Schroeder, E. (1977) Biochemistry 16, 5321–5328.
[21] Nazar, R. N., Sitz, T. O. and Busch, H. (1975) J. Biol. Chem. 250, 8591–8597.
[22] Kelly, J. M. and Maden, B. E. H. (1980) Nucleic Acids Res. 8, 4523–4534.
[23] Nazar, R. N., Sitz, T. O. and Busch, H. (1976) Biochemistry 15, 505–508.
[24] Tinoco, I. jr, Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. and Gralla, J. (1973) Nature New Biol. 246, 40–41.
[25] Peattie, D. A., Douthwaite, S., Garrett, R. A. and Noller, H. F. (1981) Proc. Natl. Acad. Sci. USA 78, 7331–7335.