The secondary structures of the *Xenopus laevis* and human mitochondrial small ribosomal subunit RNA are similar

D. Dunon-Bluteau and G. Brun*

Laboratoire de Biochimie, Section de Biologie, Institut Curie, 26 rue d'Ulm, 75231 Paris Cedex 05 and **Laboratoire de Biochimie des acides nucléiques, Ecole Normale Supérieure de Saint-Cloud, 92211 Saint-Cloud Cedex, France*

Received 22 January 1986; revised version received 31 January 1986

Extensive corrections of the nucleotide sequence of the *Xenopus laevis* mitochondrial small ribosomal subunit RNA gene [Roe et al. (1985) J. Biol. Chem. 260, 9759–9774] are reported. We found an additional fragment of 142 nucleotides and describe 25 nucleotide differences scattered in the gene. The nucleotide sequence of the *X. laevis* mitochondrial 12 S rRNA gene presents 80% homology with that of the same gene of bovine mitochondrion. We propose a new secondary structure for the product of the *X. laevis* gene. Contrary to the finding of Roe et al., we observed the same general organization of stems and loops as for the human mitochondrial 12 S rRNA gene product. On the other hand, the structural homology observed between the mitochondrial and cytoplasmic small subunit rRNAs of *X. laevis* appears much lower. These results strongly suggest that animal vertebrate mitochondrial DNAs have followed the same evolutionary pathway.

Mitochondria *rRNA* *Secondary structure*

1. INTRODUCTION

The mt DNA of all eucaryotes is a closed double-stranded molecule, whose size ranges from 14.5 to 19.5 kb. In mammals, the mt genome codes for 22 tRNAs and 2 ribosomal RNAs which are required for the mt translational apparatus [1,2]. These genomes also encode 3 subunits of the cytochrome *c* oxidase, ATPase subunit 6, apocytochrome *b*, 7 subunits of the NADH dehydrogenase [3] and possibly one polypeptide coded in another potential unidentified reading frame (URF6).

The organization of higher eucaryote mt genomes is extremely economical. With the exception of the D-loop region, there are very few non-coding nucleotides. The unique properties of the mt genetic code [4], as well as the mechanisms of its replication [1] and gene expression [2] are conserved from amphibians to mammals.

The complete nt sequence for the *Xenopus laevis* mt genome has recently been made available [5,6] and allowed an extensive comparison with those of mammals [7–9]. All genes show remarkable conservation both of their nt sequence and of their respective location on the mt genome. The main differences between *X. laevis* and mammals are observed in the size of the D-loop and in the 5′-region of the mt 12 S rRNA gene. The secondary structure of the *X. laevis* mt 12 S rRNA established by Roe et al. [5] reveals that it lacks a large loop-stem structure which is present in human mt 12 S rRNA.

We have identified and sequenced an additional fragment of 142 nt in the 5′-region of the *X. laevis* mt 12 S rRNA gene lacking in the sequence already reported [5]. We have analysed the product of that gene through secondary structure modeling

Abbreviations: D-loop, displacement loop; kb, 1000 base pairs; mt, mitochondrial; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; mt 12 S rRNA, small ribosomal subunit RNA of vertebrate mitochondria; 16 S rRNA, small ribosomal subunit RNA of *Escherichia coli*; 18 S rRNA, cytoplasmic small ribosomal subunit RNA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

[10–13]. This study leads to the conclusion that mt 12 S rRNA is highly conserved from amphibians to mammals.

2. MATERIALS AND METHODS

The entire *X. laevis* mt DNA cloned at the *Bam*HI site of pBR322 (pXlm31, a gift from I. Dawid, NIH, Bethesda) was used for sequencing. We purified the *Bgl*II-B fragment located at position 1.7–2.36 and a *Bgl*II-*Hpa*I fragment located at position 2.36–4.75 of our restriction map [14].

The different fragments were obtained by preparative 1% agarose gel electrophoresis, electroelution, chromatography on NACS columns (BRL) and ethanol precipitation. They were further digested with restriction endonucleases yielding 5′-protruding ends: *Ava*II, *Eco*RI and *Hin*fI. The digests were 3′-end labelled with [^{32}P]deoxyribonucleotide triphosphates using the Klenow polymerase, each labelled fragment was secondarily digested and each piece separated by PAGE and then sequenced following the Maxam and Gilbert procedure [15].

Analyses of the sequences were carried out at CITI II computer facilities. Prediction of the secondary structure of the 12 S rRNA gene product was performed using the algorithm of Kanehisa and Goad [16] which gives thermodynamic constraints for each of the stems studied.

3. RESULTS AND DISCUSSION

3.1. *Sequence analysis*

Fig.1 presents the nt sequence of the *X. laevis* mt 12 S rRNA gene. This gene contains one *Eco*RI site located at position 2.95 of our previous restriction map [14]. Fig.2 presents the nt sequence around the *Eco*RI site. It shows that this site is

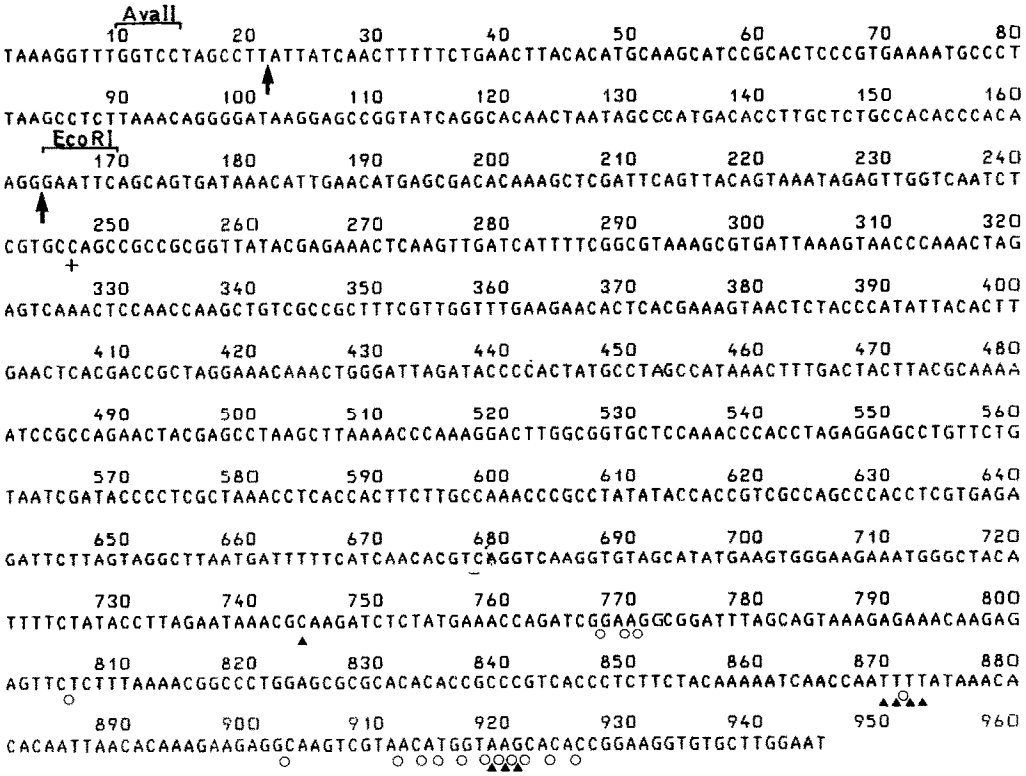

Fig.1. The nt sequence of the *X. laevis* mt 12 S rRNA gene. The sequence shown is that of the L strand of the mt DNA. The *Ava*II site is located at position 3.15 and the *Eco*RI site at position 2.95 in our restriction map [14]. The arrows bracket the new 142 nt-long fragment. Each symbol shows the position of a single nt either removed (○), added (+) or substituted (▲) from the sequence previously reported [5].

contiguous to a nt sequence absent from the sequence reported by Roe et al. [5]. This junction belongs to a 142 nt-long fragment omitted by these authors. This fragment is located at nt positions 22–163 of the mt 12 S rRNA gene. This finding together with other nt modifications extend the mt 12 S rRNA gene from 818 to 945 nt which is in better agreement with the length of the known mammalian mt 12 S rRNAs. All point modifications cannot be attributed to polymorphism since the same plasmid was used in both studies.

The *X. laevis* mt 12 S rRNA gene sequence presents 80% homology with that of mt 12 S rRNA of ox or mouse. This homology is slightly lower with the human mt 12 S rRNA gene but remains in the same range as that observed between the human gene and other mammalian 12 S rRNA genes.

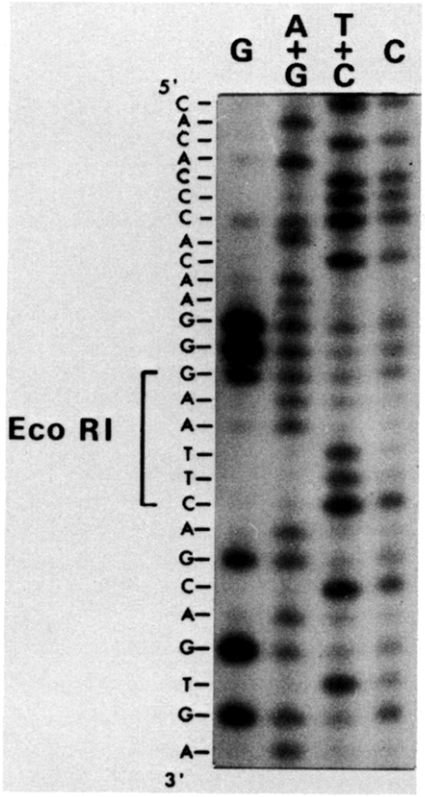

Fig.2. The nt sequence around the *Eco*RI site located in the *X. laevis* mt 12 S rRNA gene. The mt 12 S rRNA and the sequenced strand have the same polarity. The nt sequence has been determined following the procedure of Maxam and Gilbert [15]. The specificity of the reaction is indicated above each lane. The *Eco*RI site flanks the added fragment of 142 nt (nt 22–163 in the mt 12 S rRNA gene sequence). The extremity of this fragment is located upstream from the *Eco*RI site. A sequencing gel (8% acrylamide, 8 M urea) has been used.

3.2. *Secondary structure of the 5'-end of the mt 12 S rRNA*

Fig.3 shows a computerized model of the secondary structure obtained for the 5'-end of the *X.*

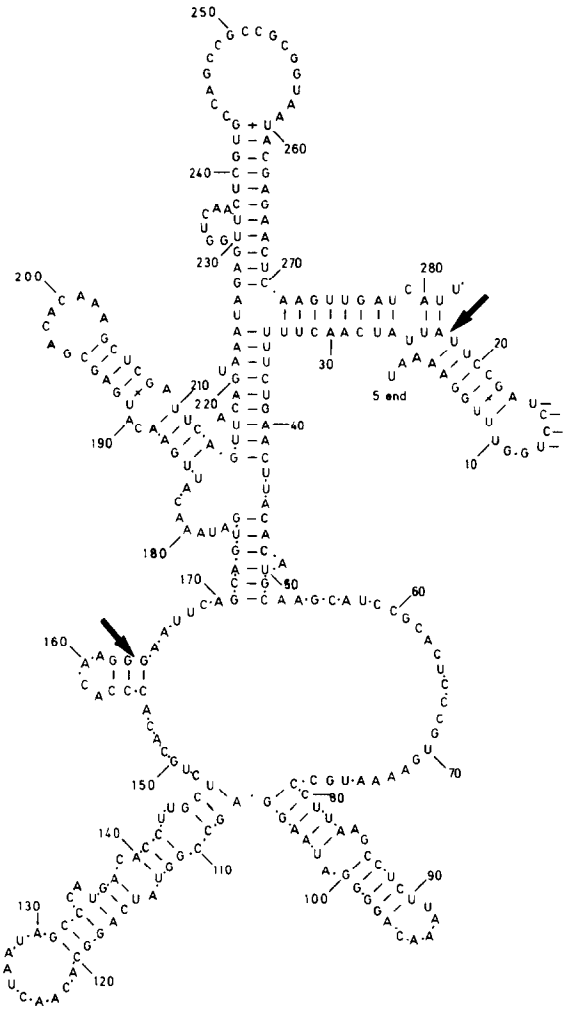

Fig.3. Putative secondary structure of the 5'-part of the *X. laevis* mt 12 S rRNA. Every tenth base is numbered from the 5'-end; (–) A-U and G-C base pairs, (+) a G+U base pair. Prediction of the secondary structure was performed using the algorithm of Kanehisa and Goad [16]. Arrows, extremities of the new 142 nt fragment.

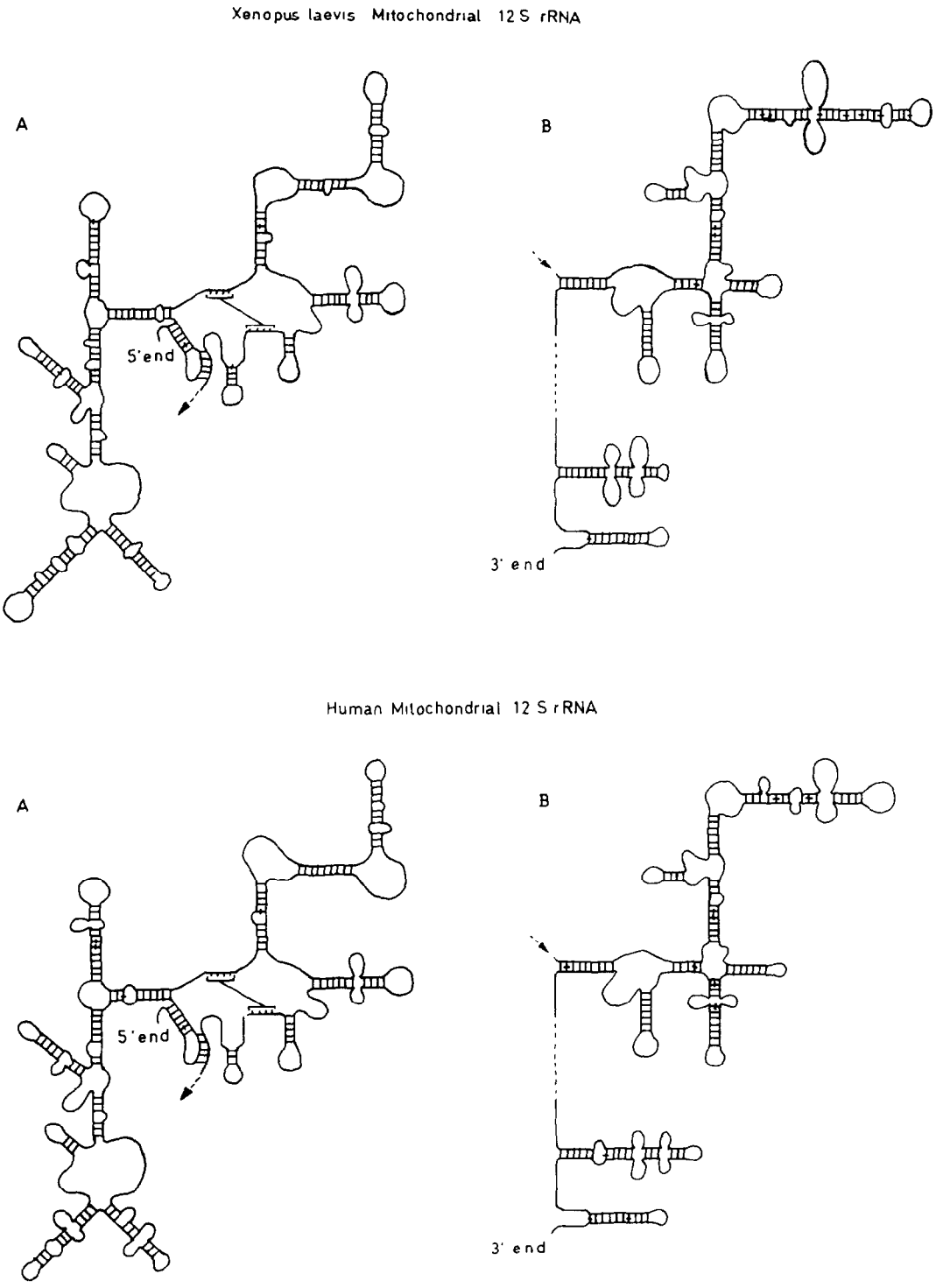

Fig.4. Comparison of the putative secondary structure of *X. laevis* and human mt 12 S rRNAs. These structures have been drawn in two domains (A and B) according to Maly and Brimacombe [11]. Arrows, connection between the two domains, (–) A-U and G-C base pairs, (+) G+U base pair.

laevis mt 12 S rRNA. This was drawn on the basis of the general model already established for numerous procaryotic or eucaryotic mt small ribosomal subunit RNAs [12]. It is worth noting that the single nt C, missing at position 246 in the previous report [5] is a conserved nt in this model.

We have also compared the different putative helices with those obtained for the small ribosomal subunit RNA in different procaryotic and eucaryotic species [10,11]. The nt at positions 46–52 and 172–177 are conserved in all mammalian mt 12 S rRNAs as well as in *E. coli* 16 S rRNA. The nt 46–52 are located in the new 142 nt-long fragment.

In addition, we have determined the ΔG for each putative helix in this region using the algorithm of Kanehisa and Goad [16]. The results are consistent with a high degree of stability of helices 80–91/97–107, 108–121/132–148 and 184–195/204–214 and a poorer stability of helices 4–9/17–22 and 23–33/278–290. The stability of helix 34–42/215–227 is very weak but we have conserved that stem for reasons of homology with the mammalian mt 12 S rRNA.

3.3. *Comparative analysis of ribosomal subunit RNAs*

Comparison of the secondary structure of the mt 12 S rRNA of *X. laevis* and human depicted in fig.4, shows striking similarities. Analysis of the nt sequence of the stems shows that homology is higher between 12 S rRNA of the two species (human and *X. laevis*) than between cytoplasmic [17] and mt small ribosomal subunit of *X. laevis*. Interestingly, as with other eucaryotes [18], the *X. laevis* mt 12 S rRNA sequence is much closer to the *E. coli* small ribosomal subunit RNA than to the corresponding cytoplasmic 18 S rRNA from the same origin. This is an additional argument in favor of the possible bacterial endosymbiotic origin of mitochondria.

Taking into account the modification brought about by this report, it follows that all known coding sequences of the *X. laevis* mt genome are highly conserved throughout the animal vertebrate series. In contrast, the D-loop sequences are poorly conserved between all animal species studied so far. The D-loop contains the replication origin of the H strand of mt DNA and the transcription initiation sites of both H and L strands [2]. The newly replicated H strand of the D-loop is about 600 nt in mammals whereas it is between 1350 and 1500 nt in amphibians [1]. It would be of great interest to know the meaning of these unique distinctive features of the mt DNA of each vertebrate class. Since this region is known to be involved in the regulation of replication and transcription of the genome, this could reflect individual modes of control of the expression of genomes with the same basic organization.

ACKNOWLEDGEMENT

We thank Dr F. Michel for helpful discussion.

REFERENCES

[1] Clayton, D.A. (1982) Cell 28, 693–705.
[2] Clayton, D.A. (1984) Annu. Rev. Biochem. 53, 573–594.
[3] Chomyn, A., Mariottini, P., Cleeter, M.W.J., Ragan, C.I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R.F. and Attardi, G. (1985) Nature 314, 592–597.
[4] Barrell, B.G., Anderson, S., Bankier, A.T., De Bruijn, M.H.L., Chen, E., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1980) Proc. Natl. Acad. Sci. USA 77, 3164–3166.
[5] Roe, B.A., Ma, D.P., Wilson, R.K. and Wong, J.F.H. (1985) J. Biol. Chem. 260, 9759–9774.
[6] Dunon-Bluteau, D., Volovitch, M. and Brun, G. (1985) Gene 36, 65–78.
[7] Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167–180.
[8] Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature 290, 457–465.
[9] Anderson, S., De Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.C. (1982) J. Mol. Biol. 156, 683–717.
[10] Zwieb, E., Glotz, C. and Brimacombe, R. (1981) Nucleic Acids Res. 9, 3621–3640.
[11] Maly, P. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 7263–7286.
[12] Stiegler, P., Carbon, P., Ebel, J.P. and Ehresmann, C. (1981) Eur. J. Biochem. 120, 487–495.

[13] Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983) Microbiol. Rev. 47, 621–669.
[14] Cordonnier, A.M., Vannier, P.A. and Brun, G.M. (1982) Eur. J. Biochem. 126, 119–127.
[15] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[16] Kanehisa, M.I. and Goad, W.B. (1982) Nucleic Acids Res. 10, 265–278.
[17] Salim, M. and Maden, E.H. (1981) Nature 291, 205–208.
[18] Küntzel, H. and Köchel, H.G. (1981) Nature 293, 751–755.