The nucleotide sequence of the 5S rRNA from the archaebacterium Thermoplasma acidophilum

Kenneth R.Luehrsen*, George E.Fox*, Michael W.Kilpatrick**, Richard T.Walker**, Horst Domdey[+†], Guido Krupp[+†] and Hans J.Gross[+†]

*Department of Biophysical Sciences, University of Houston, Houston, TX 77004, USA, **Chemistry Department, Birmingham University, Birmingham B15 2TT, UK, and [+]Max Planck Institut für Biochemie, 8033 Martinsried, GFR

Received 31 December 1980

ABSTRACT

The complete nucleotide sequence of the 5S ribosomal RNA isolated from the archaebacterium Thermoplasma acidophilum has been determined. The sequence is: pG GCAACGGUCAUAGCAGCAGGGAAACACCAGAUCCCAUUCCGAACUCGACGGUUA AGCCUGCUGCGUAUUGCGUUGUACUGUAUGCCGCGAGGGUACGGGAAGCGCAA UAUGCUGUUACCAC(U)OH. The homology with the 5S rRNA from another archaebacterial species, Halobacterium cutirubrum, is only 60.6% and other 5S rRNAs are even less homologous. Examination of the potential for forming secondary structure is revealing. T. acidophilum does not conform to the usual models employed for either procaryotic or eucaryotic 5S rRNAs. Instead this 5S rRNA has a mixture of the characteristic features of each. On the whole this 5S rRNA does however appear more eucaryotic than eubacterial. These results give further support to the notion that the archaebacteria represent an extremely early divergence among entities with procaryotic organization.

INTRODUCTION

Thermoplasma acidophilum is considered a genus of uncertain affiliation and is listed with the mycoplasmas (1). This is inappropriate since the mycoplasmas seem related to certain Clostridia (2-4) while T. acidophilum is an archaebacterium (5). In order to explore further the phylogenetic position of this thermoacidophile, we have determined the complete nucleotide sequence of its 5S rRNA. This work was conducted independently in the USA and Europe. Only at a late stage have we compared our data and the identical results have been consolidated into a single communication.

MATERIALS AND METHODS

Low molecular weight RNA was isolated from Thermoplasma acidophilum (strain 122-1B2 or strain 122-1B3) cells (6). Fractionation on a BD-cellulose column with a linear gradient of NaCl (0.425M - 0.6M) containing 10mM $MgCl_2$ gave a crude preparation of 5S rRNA (7). Final purification was by electrophoresis at 1.0 kV for 7 hours

on a 40 cm x 20 cm x 0.02 cm 8% polyacrylamide slab gel.

Both 5'-(5) and 3'-(8) ^{32}P *in vitro* end-labeled 5S rRNA were utilized in sequence determination. Sequence analysis in the 5'-region was by partial digestion of 5'-labeled material with nuclease P_1 or alkali, followed by two dimensional homochromatography (9). The terminal nucleotide was identified by complete digestion with nuclease P_1 followed by TLC in the presence of UV markers. Alternatively, aliquots of 5'-labeled RNA were digested with ribonucleases T_1, U_2, or A and separated on DEAE-paper as previously described (10).

In order to sequence the interior regions, both 5'- and 3'-labeled 5S rRNA and a 5'-labeled fragment obtained by partial digestion with RNAse T_2 were analyzed by rapid gel sequencing methods. The pyrimidine distinction in the enzymatic method was enhanced with nucleases from *Staphylococcus aureus* and *Neurospora crassa* (11) or *Physarum polycephalum* (12). The chemical method (13) was also employed on 3'-labeled 5S rRNA. Sequencing ladders were run on thin (0.02 cm) 20% polyacrylamide gels at 2.7 kV.

The length heterogeneity at the 3'-end was verified by electrophoresis of 3'-labeled 5S rRNA on a long 8% polyacrylamide gel [80 cm x 30 cm x 0.02 cm] for 20 hours at 1.2 kV which gave two bands. A 3'-end analysis of each band was conducted by digestion with RNAse T_2 and separation by TLC in the presence of UV markers. Complete digestion of 3'-labeled material with RNase T_1 followed by two dimensional separation (9) also confirmed the length heterogeneity.

EXPERIMENTAL RESULTS

Sufficient gel ladders were independently obtained by both groups to determine unequivocably the *Thermoplasma acidophilum* 5S rRNA sequence given in Figure 1B. The only heterogeneity identified was at the 3'-terminus, where at least 2 variants were found which differ by the presence or absence of the terminal uracil. The longer version is the more abundant ($\sim$10:1). As is usual in 5S rRNA, no post-transcriptionally-modified nucleotides were encountered in either the gel ladders or in preliminary oligonucleotide cataloging studies (Luehrsen and Woese, unpublished results).

COMPARISON WITH OTHER 5S rRNAs

Limited homology with other published 5S rRNAs is seen. *Halobacterium cutirubrum*, another archaebacterium, is the highest at 60.6%. Various eucaryotes show similar

levels of homology eg. Tetrahymena thermophilia - 57.8%; whereas typical eubacteria such as Escherichia coli - 44.4% and Bacillus subtilis - 44.8% are somewhat lower.

Two distinct but related models of secondary structure are available for eubacterial (14), Figure 1A, or eucaryotic cytoplasmic (15), Figure 1C, 5S rRNAs. We have examined possible pairing schemes in T. acidophilum and the arrangement shown in Figure 1B is the one which is most consistent with these structures. The T. acidophilum sequence, like that of H. cutirubrum (16), does not conform to either model exactly. The most striking exception is that one of the three helices that can be formed in all other 5S rRNAs cannot be formed in T. acidophilum. Sequence homology implies that position 29 - 32 would be expected to pair with 46 - 49. In fact only two pairs can be made which would not permit a stable helix unless other factors such as protein interaction were present in vivo. The loop that would be defined by the missing helix III is however quite typically eubacterial. It contains 13 nucleotides which is usual (eucaryotic 5S rRNAs contain 12) and includes the common $\overset{41}{\text{pyGAAC}}$ sequence. Indeed, this entire segment, -UCCCAUUC$\overset{41}{\text{C}}$GAAC- is found exactly in Thermus aquaticus and Clostridium pasteurianum 5S rRNAs.

In the structural domain containing helical Region IV, the similarity is to the eucaryotic cytoplasmic 5S rRNAs. In eubacteria this helix is a strongly base-paired hairpin which defines a small loop (three or four nucleotides). In the eucaryotes this structure is replaced by two helices separated by an interior loop. Helix IV is then less strongly paired and typically defines a larger loop of 9 - 12 nucleotides. The additional helix, region V, can be envisaged as being coaxial with the molecular stalk; helix I. The T. acidophilum sequence fits this eucaryotic prototype reasonably well. The loop defined by helix IV is 11 bases long (assuming that the terminal U_{84}, G_{94}, pair does not form) and the helical region includes the conserved U_{80}, G_{98} base pair. The supporting helix, V, can be coaxial with helix 1 (which contains a C_3, A_{118} mispair) but deviates from the eucaryotes in being ten pairs in length.

Since T. acidophilum 5S rRNA does not conform to either structural class, comparative arguments regarding the correctness of Figure 1B are not compelling and this structure should be regarded as speculative and in any case several alternative models for 5S rRNA have been proposed (17 - 19). Indeed, as shown, the status of positions 64 and 65 is uncertain. In most 5S rRNA sequences helix II can be extended to eight pairs if a singly looped-out nucleotide is proposed. In T. acidophilum this extension can be made

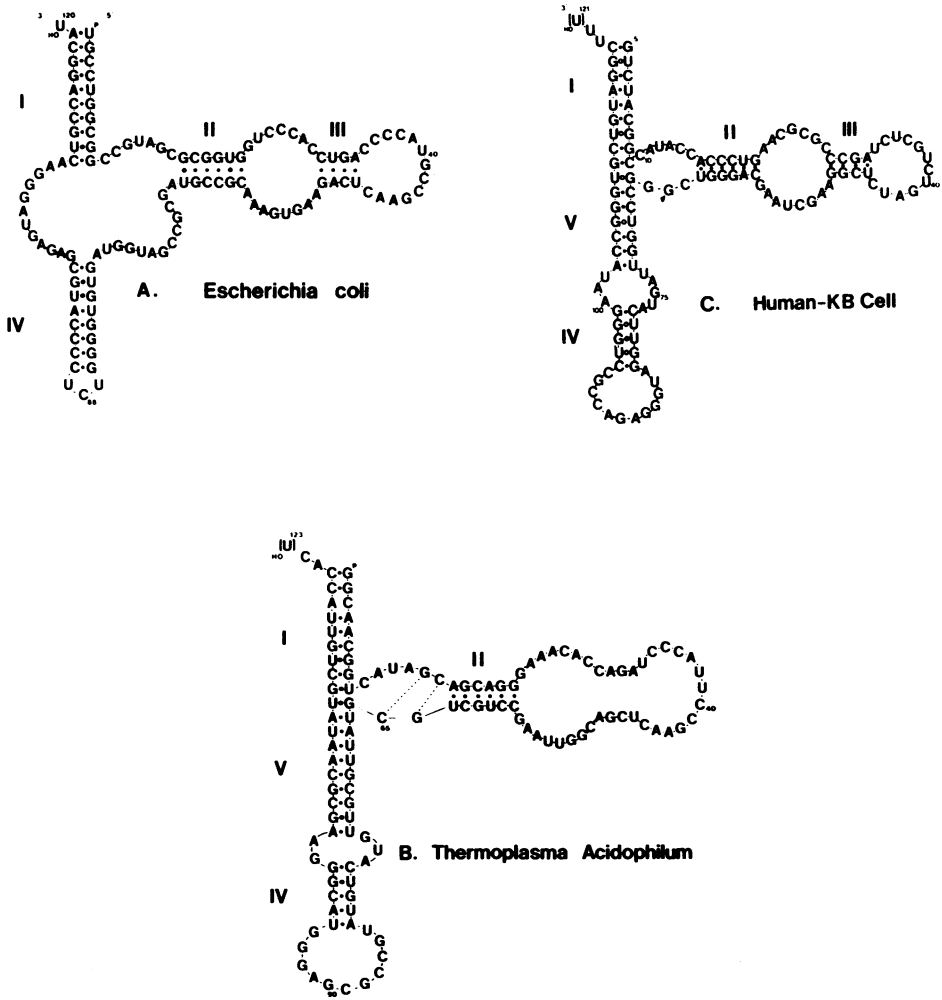

Figure 1.

The nucleotide sequence of Thermoplasma acidophilum 5S rRNA is shown in an hypothetical secondary structural arrangement that facilitates comparison with the two major classes of 5S rRNA. The eubacterial types are represented here by the Escherichia coli sequence and the eucaryotic cytoplasmic types by the human KB cell sequence. The secondary structure shown for the E. coli sequence is the usual one (12) but it has been redrawn to make the comparison easier.

directly since positions 14 and 15 are complementary to 64 and 65 (17). Nevertheless this eight base helix might not exist since 64 and 65 are also complementary to 112 and 113. This latter arrangement would allow the extension of helix V to twelve pairs but would shrink helix I.

DISCUSSION

It is apparent that T. acidophilum 5S rRNA is unusual. How much of this reflects its thermoacidophilic niche is not clear. Certainly comparison of both primary sequence and potential secondary structure suggests there is at least as much and probably more similarity to eucaryotic 5S rRNAs than to the usual eubacterial 5S rRNAs. This is consistent with the notion that there was an extremely early evolutionary divergence among entities with a procaryotic type of cellular organization resulting in at least two procaryotic lines of descent; the eubacteria and archaebacteria. The considerable difference between the 5S rRNAs of the two modern archaebacteria, H. cutirubrum and T. acidophilum likewise suggests a common ancestor that is remote in time. Finally the presence of eucaryotic features in these archaebacterial 5S rRNAs (herein and reference 16) joins other recent observations (20, 21) in suggesting that archaebacterial ribosomes are more similar to eucaryotic ribosomes than are eubacterial ribosomes.

ACKNOWLEDGEMENTS

We are grateful to Drs. T. A. Langworthy and D. G. Searcy for providing T. acidophilum cells. This work was supported by the Science Research Council and the National Aeronautics and Space Administration (NSG-7440). M. W. K. was a recipient of a CASE award with the P.H.L.S. Mycoplasma Reference Laboratory and an EMBO fellowship. We are grateful to R. H. Leach for advice and encouragement and to C. R. Woese for access to unpublished data.

†Present address: Institut für Biochemie, Universität Würzburg, Roentgenring 11, 8700 Würzburg, GFR

REFERENCES

1. Buchanan, R.E. and Gibbons, N.E. (ed.) (1974) Bergey's manual of determinative bacteriology, 7th edition, The William Wilkins Co. Baltimore.
2. Woese, C.R., Maniloff, J. and Zablen, L.B. (1980) Proc. Natl. Acad. Sci. USA 77, 494-498.
3. Walker, R.T. and RajBhandary, U.L. (1978) Nucleic Acids Res. 5, 57-70.

4. Kilpatrick, M.W. and Walker, R.T. (1980) Nucleic Acids Res. 8, 2783-2786.
5. Woese, C.R., Magrum, L.G. and Fox, G.E. (1978) J. Mol. Evol. 11, 245-252.
6. Jones, A.S., Tittensor, J.R. and Walker, R.T., (1965) J. Gen. Microbiol. 40, 405-411.
7. Domdey, H. and Gross, H.J. (1979) Anal. Biochem. 93, 321-328.
8. England, T.E. and Uhlenbeck, O.C. (1978) Nature 275, 560-561.
9. Silberklang, M., Gillum, A.M. and Rajβhandary, U.L. (1979) Methods Enzymol. 59, 58-109.
10. Uchida, T., Bonen, L., Schaup, H., Lewis, B.J., Zablen, L.B. and Woese, C.R. (1974) J. Mol. Evol. 3, 63-77.
11. Krupp, G. and Gross, H.J. (1979) Nucleic Acids Res. 6, 3481-3490.
12. Donis-Keller, H. (1980) Nucleic Acids Res. 8, 3133-3142.
13. Peattie, D. (1979) Proc. Natl. Acad. Sci. USA 74, 4537-4541.
14. Fox, G.E. and Woese, C.R. (1975) Nature 256, 505-507.
15. Luehrsen, K.R., and Fox, G.E. (1981) Proc. Natl. Acad. Sci. USA in press.
16. Matheson, A.T. (1978) in Energetics and Structure of Halophilic Micro-organisms ed. S.T. Caplan and M. Ginzburg, pp. 481-501.
17. Nishikawa, K. and Takemura, S. (1974) J. Biochem. (Japan) 76, 935-947.
18. Hori, H. (1976) Molec. Gen. Genet. 145, 119-123.
19. Vigne, R. and Jordan, B.R. (1977) J. Mol. Evol. 10, 77-86.
20. Kwok, Y. and Wong, J.T.F. (1979) Can. J. Biochem. 58, 213-218.
21. Matheson, A.T., Moller, W., Amons, R. and Yaguchi, M. (1979) in Ribosomes ed. G. Chambliss et al. pp. 297-332.