Volume 8 Number 21 1980 Nucleic Acids Research

Nucleotide sequence of Crithidia fasciculata cytosol 5S ribosomal ribonucleic acid

Ron M.MacKay, Michael W.Gray and W.Ford Doolittle

Department of Biochemistry, Faculty of Medicine, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Received 25 August 1980

ABSTRACT

The complete nucleotide sequence of the cytosol 5S ribosomal ribonucleic acid of the trypanosomatid protozoan *Crithidia fasciculata* has been determined by a combination of T1-oligonucleotide cataloging and gel sequencing techniques. The sequence is: GAGUACGACCAUACUUGAGUGAAAACACCAUAUCCCGUCCGAUUUGUGAAGUUAAGCACC CACAGGCUUAGUUAGUACUGAGGUCAGUGAUGACUCGGGAACCCUGAGUGCCGUACUCCCOH. This 5S ribosomal RNA is unique in having GAUU in place of the GAAC or GAUC found in all other prokaryotic and eukaryotic 5S RNAs, and thought to be involved in interactions with tRNAs. Comparisons to other eukaryotic cytosol 5S ribosomal RNA sequences indicate that the four major eukaryotic kingdoms (animals, plants, fungi, and protists) are about equally remote from each other, and that the latter kingdom may be the most internally diverse.

INTRODUCTION

5S ribosomal ribonucleic acid (rRNA) primary sequences provide valuable information on phylogenetic relationships [1,2] and on 5S rRNA function [3,4] and secondary structure [5-8]. Of the more than twenty eukaryotic cytosol 5S sequences so far reported [9,10,11], only two (those for *Chlorella pyrenoidosa* [6] and *Tetrahymena thermophila* [11] are from the Kingdom Protista (unicellular algae and protozoa), although this kingdom is likely the most ancient and diverse of all eukaryotic kingdoms [12]. Here we report the sequence of the 5S rRNA of the trypanosomatid protozoan *Crithidia fasciculata* and compare it to other eukaryotic 5S rRNA sequences.

MATERIALS AND METHODS

Preparation and T1-oligonucleotide cataloging of uniformly-labeled [^{32}P] 5S rRNA. *Crithidia fasciculata* was grown at 26°C for four days in medium [13] containing 0.1 mCi/ml [^{32}P] orthophosphate (New England Nuclear). RNA was extracted from ca 1 g of cells suspended in 30 ml 50 mM Tris-Cl, pH 8.0 as described previously [14]. That RNA fraction soluble in 3 M NaCl at 4°C (5S rRNA and tRNA) was, after ethanol precipitation, dried, dissolved (to 5 µg/µl) in loading buffer [15] and resolved on 20 x 20 x 0.3 cm 10% polyacrylamide slab gels [16]. The 5S rRNA band (identified by autoradiography

or UV-absorbance) was "electroeluted" [17] from gel slices using an Isco electroelution device. The buffer chamber contained TBE (50 mM Tris, 50 mM boric acid, 1 mM EDTA, pH 8.3). The elution cups contained 1/10 x TBE, and electrophoresis was for 90 min at 120 v. Eluted RNA was extracted once with TBE-saturated phenol, and precipitated and washed with ethanol before drying. T1-oligonucleotide catalogs were prepared with ribonuclease T1, U_2 (Calbiochem) and A (Worthington) as described previously [18-21].

Preparation of terminally-labeled 5S rRNAs. Unlabeled 5S rRNA prepared as described above was 5'-labeled (after digestion with calf intestine alkaline phosphatase [Sigma, 0.1 units/nmol terminal phosphate]) using [γ-^{32}P] ATP (prepared as described in [22]) and polynucleotide kinase (P-L Biochemicals) as described by Donis-Keller *et al.* [16]. 3'-labeled 5S rRNA was prepared as described by Peattie [15] using [5'-^{32}P] pCp synthesized to a specific activity of 5 x 10^3 Ci/mmol [22] with a donor : acceptor : ATP molar ratio of 1 : 4 : 15 and RNA ligase (P-L Biochemicals) at 500 units/ml. Terminally-labeled RNAs were purified electrophoretically as described above.

Terminal nucleotide analyses. Not more than 2 µg 5'-labeled 5S rRNA was digested with snake venom phosphodiesterase (0.25 mg/ml, prepared according to [23]) in 20 µl 0.125 M ammonium formate, pH 9.2, for 16-24 h at 37°C. A small amount (< 2µg) of 3'-labeled RNA was hydrolyzed (90 h, 22°C) in 10 µl 1 m NaOH and neutralized by the addition of 1 µl glacial acetic acid. 5'- and 3'-labeled nucleotides were identified by thin layer chromatography on Eastman Chromagram cellulose plates with fluorescent indicator (first washed with 10% saturated ammonium sulfate and then thoroughly dried). Appropriate unlabeled 5' or 2'(3') nucleotide markers (pA, pC, pG, pU, pAm, pCm, pGm, pUm, and pΨ, or Ap, Cp, Gp, Up and Ψp) were spotted with the labeled material. 95% ethanol: water, 4 : 1 [24] was used in the first dimension, and saturated ammonium sulfate : propan-2-ol, 40 : 1 [25] was used in the second. Labeled material was identified autoradiographically, while markers were visualized under UV light.

Sequencing of terminally-labeled RNAs. The enzymatic method of Donis-Keller *et al.* [16] was used with 5'- or 3'-labeled RNA, while the chemical method of Peattie [15] was used only on 3'-labeled material. Electrophoresis was on thin (0.05 x 33 x 40 cm or 0.05 x 33 x 100 cm) 6.0 or 10.5% polyacrylamide-urea gels (19 : 1, acrylamide : bis-acrylamide).

RESULTS

Fig. 1 shows the sequence of *Crithidia fasciculata* 5S rRNA derived from gel-sequencing techniques and T1-oligonucleotide catalog analyses. All but five of the 120 residues could be unambiguously identified by the partial

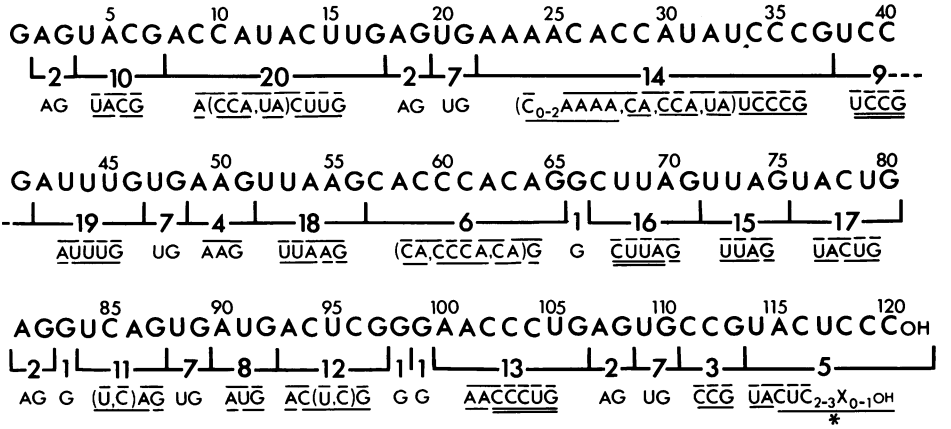

FIGURE 1. Nucleotide Sequence of *Crithidia fasciculata* 5S rRNA. Numbering is from the 5'- to 3'-terminus. The degree of *in vivo* phosphorylation at the 5'-terminus is unknown. The ribonuclease T1-generated oligonucleotides are aligned below the sequence. Lines above and below these oligonucleotides indicate respectively the observed products of digestion with ribonucleases A and U_2. Double lines indicate confirmation of the sequence by digestion of oligonucleotide with ribonuclease U_2 under such conditions that U_2 acts as a 5'- to 3'-exonuclease [18]. The sequence of the oligonucleotide above the asterisk (*) was inferred from (i) the primary position of T5 after two-dimensional electrophoresis (mobility of T5 was consistent with the presence of two U residues and a 3'-hydroxyl), (ii) the ribonuclease A digestion products and (iii) the observation that ribonuclease T1 used under "overcutting" conditions [19] generated the dinucleotide CU. Quantitation of the T1-oligonucleotides yielded approximately one molar copy of each, except for five copies each of T1, T2 and T7.

chemical digestion method of Peattie [15], using 3'-labeled material (Fig. 2). The identity of the 5' residue as pGp was confirmed by snake venom phosphodiesterase digestion and thin layer chromatography of 5'-labeled material, which yielded 90% pG as terminal nucleotide. Residue 10, which could be read as either G or C (Fig. 2A), was concluded to be C, on the basis of analyses of oligonucleotide T20 (Fig. 1). Residue 29 (Fig. 2A) was similarly confirmed to be C on the basis of the sequence of oligonucleotide T14. Residue 87 gave no strong band in any track (Fig. 2B). T1-oligonucleotide catalog data (Fig. 1, T11) showed residue 87 to be G. Although absence of any strong band at this position might indicate residue modification, a Gp of ordinary mobility was observed when ribonuclease U_2 digests of oligonucleotide T11 were resolved either by electrophoresis on DEAE paper [18] or by thin layer chromatography (see Materials and Methods). (Furthermore, snake venom phosphodiesterase digests of uniformly-labeled 5S rRNA yielded only the four unmodified 5'-nucleotides.) The number of C residues between G_{110} and G_{113} and between

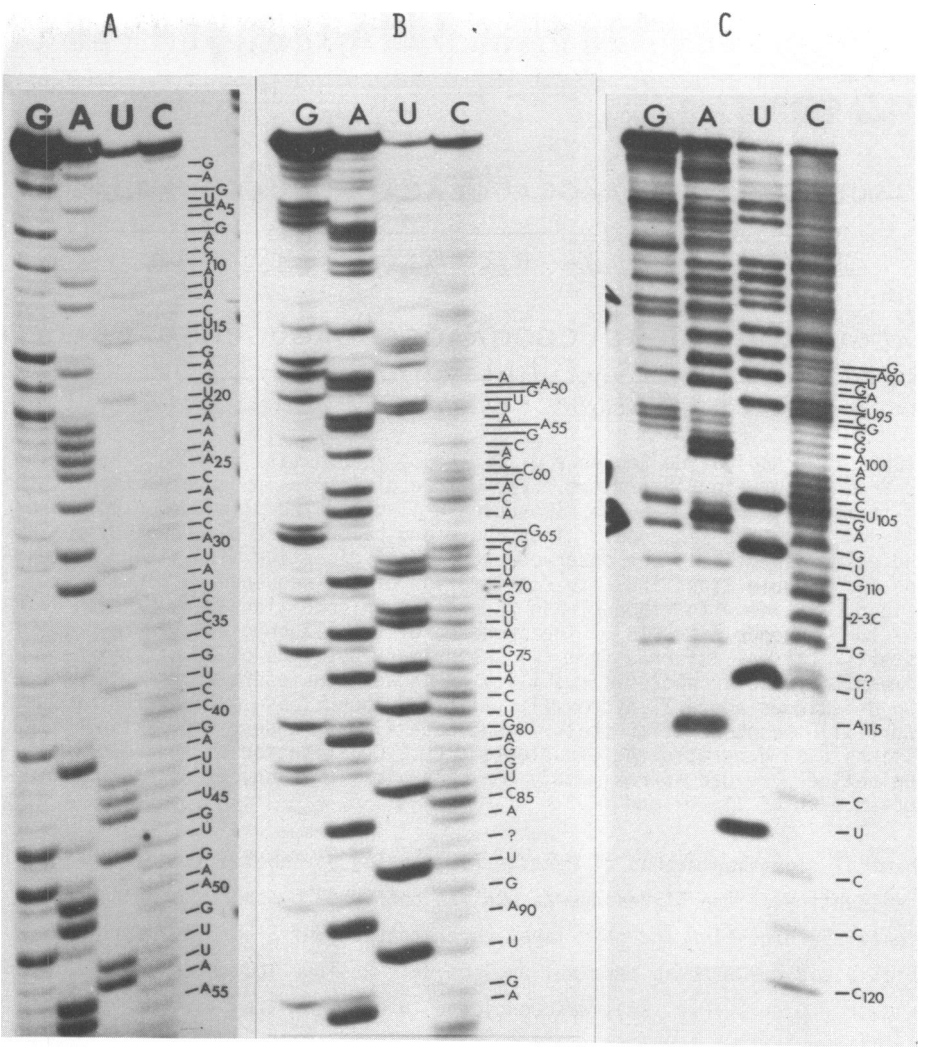

FIGURE 2. Autoradiograms of three sequencing gels showing the resolution of 3'-labeled RNA subjected to limited chemical cleavage [15]. Gels A and B (6% polyacrylamide) and gel C (10.5% polyacrylamide) span the entire molecule from 5'- to 3'-terminus.

G_{113} and U_{114} could not be unambiguously determined by electrophoretic resolution of partial chemical degradation products. Comparison with T1-oligonucleotides T3 and T5 indicated that the sequence in the region beginning with G_{110} is 5'-GCCGU-3'. The observation of a 3' terminal C_{OH} was supported by terminal nucleotide analysis of 3'-labeled RNA, which gave > 90% Cp after alkaline hydrolysis.

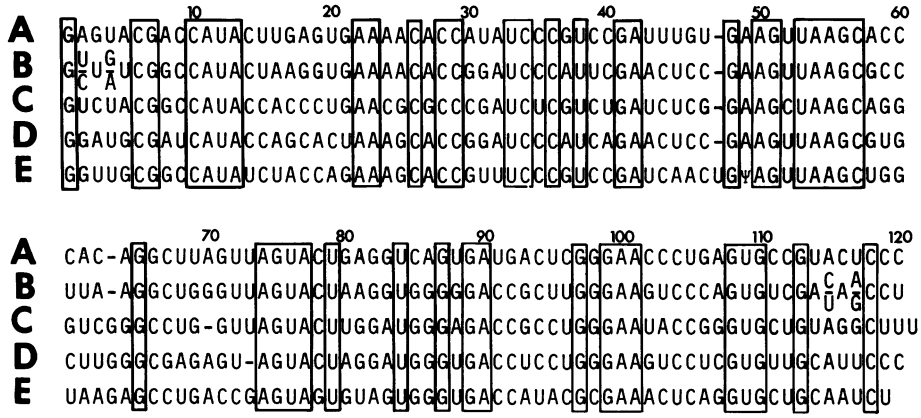

FIGURE 3. Alignment for maximum overall homology of the 5S rRNA sequences of A, *Crithidia fasciculata*; B, *Tetrahymena thermophila* [11]; C, human KB cells [26]; D, wheat embryo [27]; and E, *Saccharomyces carlsbergensis* [28]. Boxed regions outline nucleotide positions of universal homology in these five species. Numbering of positions is that of *C. fasciculata* (Fig. 1).

All of the above ambiguities were also independently resolved using the enzymatic sequencing method of Donis-Keller *et al.* [16] with either 5'- or 3'-labeled material (data not shown). Sequencing gels obtained with partial ribonuclease T1 or U_2 digests and partial alkaline hydrolysates showed that residues 10 and 29 are pyrimidines and *not* Gs, that residue 87 *is* G, that there are only two residues (neither G nor A) between G_{110} and G_{113}, and that there is no nucleotide between G_{113} and U_{114}. Thus the complete sequence of this, and we hope most other 5S rRNAs, can be determined by a combination of these two gel-sequencing techniques, and oligonucleotide catalogs may not in future be necessary for confirmation.

DISCUSSION

The sequence of the cytosol 5S rRNA of *Crithidia fasciculata* is aligned with those of the ciliated protozoan *Tetrahymena thermophila* [11], human KB cells [26], wheat embryo [27] and *Saccharomyces carlsbergensis* [28] in Fig. 3, the alignment being designed to maximize the total number of identities in all pairwise comparisons. Calculations of percent homology (identical residues ÷ average length) reveal the four eukaryotic kingdoms to be approximately equally divergent. *C. fasciculata* 5S rRNA shows 60.6, 59.2 and 59.7% homology with the 5S rRNAs of KB cells, wheat embryo and *S. carlsbergensis*, respectively. For *T. thermophila*, the comparable values are 63.9, 68.3 and 58.9%. Other interkingdom comparisons show similar values (67.2% for animals and plants, 61.2% for animals and fungi, 60.6% for plants and fungi).

The homology between the two protozoal 5S rRNAs (66.7%) is within this same low range (60-70%), and well below intrakingdom values (e.g. the 5S rRNAs of *Drosophila melanogaster* [8] and the echinoderm *Lytechinus variegatus* [10] show 76.3 and 82.2% homology, respectively, with KB 5S rRNA, and 85.0% homology with each other). Further evidence for protistan diversity is found in residues 41-44 of *C. fasciculata* 5S rRNA. This tetranucleotide is believed to be involved in complementary interactions with sequences in loop IV of tRNAs [34,29]. All previously sequenced 5S rRNAs show either GAAC (all bacteria, all plants, and *T. thermophila*) or GAUC (all fungi and animals) in this position. *C. fasciculata* is unique in showing GAUU. Further sequencing of protozoal and algal 5S (and 5.8S) rRNAs should yield invaluable information on constraints on rRNA structure and on evolutionary relationships within this phylogenetically diverse assemblage.

ACKNOWLEDGEMENTS

We thank G.E. Fox for communication of results prior to publication, the Natural Sciences and Engineering Research Council, and the Medical Research Council for support.

REFERENCES

1. Hori, H. and Osawa, S. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 381-385.
2. Schwartz, R.M. and Dayhoff, M.O. (1978) Science 199, 395-403.
3. Erdmann, V.A. (1976) in Progress in Nucleic Acids Research and Molecular Biology, Cohn, W.E. Ed., Vol. 18, pp. 45-90 Academic Press, New York.
4. Cox, R.A. (1977) in Progress in Biophysics and Molecular Biology, Butler, J.A.V. and Noble, D. Eds., Vol. 32, pp. 193-231 Pergamon Press, Oxford-New York.
5. Nishikawa, K. and Takemura, S. (1974) J. Biochem. 76, 935-947.
6. Jordan, B.R., Galling, G. and Jourdan, R. (1974) J. Mol. Biol. 87, 205-225.
7. Fox, G.E. and Woese, C.R. (1975) J. Mol. Evol. 6, 61-76.
8. Benhamov, J., Jourdan, R. and Jordan, B.R. (1977) J. Mol. Evol. 9, 279-298.
9. Erdmann, V.A. (1979) Nucl. Acids Res. 6, r29-r44.
10. Lu, A.-L., Steege, D.A. and Stafford, D.W. (1980) Nucl. Acids Res. 8, 1839-1853.
11. Luehrsen, K.R., Fox, G.E. and Woese, C.R. (1980) Current Microbiology, in press.
12. Whittaker, R.H. (1969) Science 163, 150-160.
13. Maclean, F.I. and Amiro, E.R. (1973) Can. J. Microbiol. 19, 878-880.
14. Gray, M.W. (1979) Can. J. Biochem. 57, 914-926.
15. Peattie, D.A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 1760-1764.
16. Donis-Keller, H., Maxam, A.M. and Gilbert W. (1977) Nucl. Acids Res. 5, 2527-2538.
17. Allington, W.B., Cordry, A.L., McCullough, G.A., Mitchell, D.E. and Nelson, J.W. (1978) Anal. Biochem. 85, 188-196.
18. Uchida, T., Bonen, L., Schaup, H.W., Lewis, B.J., Zablen, L. and Woese, C. (1974) J. Mol. Evol. 3, 63-77.

19. Woese, C., Sogin, M., Stahl, D., Lewis, B.J. and Bonen, L. (1976) J. Mol. Evol. 7, 197-213.
20. Sanger, F., Brownlee, G.G. and Barrell, B.G. (1965) J. Mol. Biol. 13, 373-398.
21. Brownlee, G.G. (1972) Determination of Sequences in RNA, North Holland/ American Elsevier, Amsterdam-London-New York.
22. Walseth, T.F. and Johnson, R.A. (1979) Biochim. Biophys. Acta, 562, 11-31.
23. Lane, B.G., Diemer, J. and Blashko, C.A. (1963) Can. J. Biochem. 41, 1927-1941.
24. Lane, B.G. (1963) Biochim. Biophys. Acta 72, 110-112.
25. Singh, H. and Lane, B.G. (1964) Can. J. Biochem. 42, 87-93.
26. Forget, B.G. and Weissman, S.M. (1967) Science 158, 1695-1699.
27. MacKay, R.M., Spencer, D.F., Doolittle, W.F. and Gray, M.W. (1980) submitted to Eur. J. Biochem.
28. Miyazaki, M. (1977) Nucl. Acids Res. Special Publication No. 3, 5153-5156.
29. Nishikawa, K. and Takemura, S. (1974) FEBS Lett. 40, 106-109.