NUCLEOTIDE SEQUENCE OF *CLOSTRIDIUM PASTEURIANUM* 5S rRNA

Cheryl D. PRIBULA, George E. FOX and Carl R. WOESE*
Department of Genetics and Development, 515 Morrill Hall, University of Illinois, Urbana, Illinois 61801, USA

Received 3 March 1976

1. Introduction

Ribosomal RNAs (rRNA) are quite unique in that related and to some extent equivalent molecular species have been readily identified and isolated from every organism examined to date. Among the three species of rRNA, the 5S rRNA is the only one that can be readily sequenced with present techniques. It thus affords a unique opportunity, through sequence comparisons, to detail the constraints which have guided the evolution of this molecule. A comparative study of 5S rRNA sequences will, moreover, be useful in addressing the general problem of structure–function relationships in the translation apparatus and in identifying phylogenetic relationships.

All the procaryotic 5S rRNAs sequenced to date show a common general pattern of secondary structure and a remarkable constancy of sequence in several areas [1]. With the exception of *Anacystis nidulans* [2], a blue-green alga, the published sequences are all from aerobic bacteria. In order to expand the spectrum of 5S rRNA sequences and to explore further the universality of their previously identified architectural features [1,3], the anaerobic bacterium, *Clostridium pasteurianum,* has been selected for investigation. The choice of a *Clostridium* is especially enticing as its apparent relationship to the *Bacilli* suggests that the *Clostridia* have been pivotal organisms in the evolution of aerobic bacteria.

In this paper, the complete nucleotide sequence of *C. pasteurianum* 5S rRNA is reported. The sequence exhibits many of the architectural features found previously and is in general agreement with the universal procaryotic secondary structure proposed elsewhere [1,3]. It does, however, exhibit one unusual feature which raises the question that *C. pasteurianum* 5S rRNA may be representative of a second class of procaryotic 5S rRNAs.

2. Materials and methods

^{32}P-labelled 5S rRNA was prepared from *Clostridium pasteurianum* (ATCC 6013) by methods described previously [4,5]. The primary structure was derived by established methods [6,7] based on two-dimensional paper electrophoresis: the purified 5S rRNA was digested with either pancreatic RNase or RNase T_1. These digests were 'fingerprinted' and all the resulting oligonucleotides sequenced by 'secondary' and 'tertiary' procedures utilizing pancreatic, T_1, 'T_3', or U_2 RNases [8,9]. In order to obtain larger oligonucleotides and hence create sufficient overlaps to determine the sequence, partial digestions with U_2 and pancreatic RNase were performed [6,7]. The resulting fragments were again separated by two-dimensional paper electrophoresis using an 'extended resolution' second dimension [8,9] and sequenced by the usual secondary and tertiary procedures.

3. Results and discussion

The complete sequence of *C. pasteurianum* 5S rRNA is shown in fig.1 along with the major digestion products which lead to its determination. This sequence is 117 nucleotides in length, as are many *Bacillus* 5S rRNAs, and is consistent with the universal procaryotic secondary structure. When in this form,

* To whom all correspondence should be directed.

North-Holland Publishing Company – Amsterdam

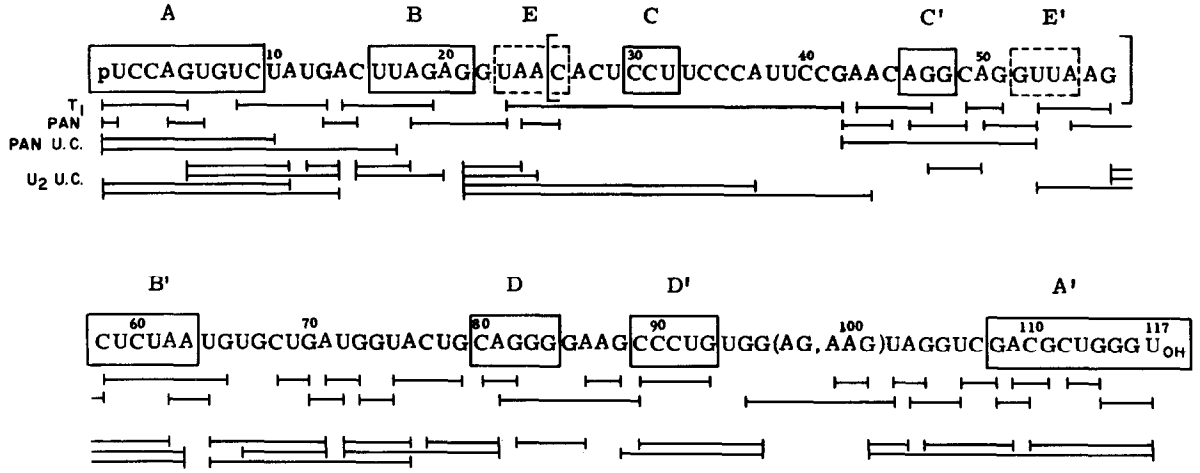

Fig.1. Sequence of *Clostridium pasteurianum* 5S ribosomal RNA. The boxes indicate regions of base pairing or suspected base pairing – A with A′, B with B′, etc. Brackets enclose the most phylogenetically conserved region in the molecule. The fragments used to deduce the sequence are designated by underlining: T_1 = T_1 ribonuclease, PAN = pancreatic ribonuclease, U_2 = ribonuclease U_2 from *Ustilago sphaergena*, and U.C. indicates partial digestion conditions.

C. pasteurianum 5S rRNA is seen to be typically procaryotic in most respects. The molecule contains a 'molecular stalk' [1,3] (helix A–A′) of nine base pairs including two of the G–U type. A 'tuned helix' (helix B–B′) (a stretch of six base pairs, the terminal one of which is G–C [1,3]) is found. Its estimated stability [10], –9.8 kcal, is that for a typical non-marine procaryote [6]. The present 5S rRNA also contains a 'procaryotic loop' (helix D–D′) [1,3]. This hairpin, found only in procaryotes, consists of a double helix of five base pairs capped by a loop of three or four (the present case) nucleotides. In *C. pasteurianum* 5S rRNA, as in *E. coli* 5S rRNA, the double-stranded stalk can be extended to eight pairs if G–U pairs are included [1]. The primary structural segment –CCGAAC–, which is universal in the procaryotes and appears to interact with the 'common arm' of the tRNA molecule [11,12] is present in *C. pasteurianum* 5S rRNA as well.

In spite of these similarities, two important differences have been detected. First, the highly conserved segment –GCGCCGAUGGUAGU– found in many aerobic bacteria is replaced by the rather different, but clearly related sequence –GUGCUGAUGGUACU–. The only other published procaryotic 5S rRNA which differs extensively in this region is that of *A. nidulans* [2], which has –GCGGCAACGAUAGC–. This region may be implicated in 5S rRNA interaction with the 23S rRNA [13].

The second difference found in *C. pasteurianum* 5S rRNA is that the 'common arm base' (helix C–C′) [1,3], contains only three contiguous base pairs instead of the usual four, though a fourth can be created if a looped out nucleotide is allowed. Helical stalks of three base pairs are not unique e.g., the dihydrouracil arm of some tRNAs. However, it is reasonable to consider in this instance that the single change noted signifies more extensive alteration in the molecule, perhaps an evolutionary redesign of the functional segment which includes the common arm base. This prospect, that the RNA in question may represent a second class of procaryotic 5S rRNA structures, is bolstered by two further observations: (1) A helix of four base pairs (helix E–E′) located between the tuned helix and the common arm base is consistent with the *C. pasteurianum* 5S rRNA primary structure. This potential helix is not found in any other 5S rRNA examined to date including *Clostridium thermosaccharolyticum* which has the customary four pairs in the common arm base [3,14]. (The E segment of the putative helix is in a rather variable portion of

the 5S rRNA sequence, while the E′ segment is highly conserved, which to us makes the helix suspect. The proof of helix E–E′ must be a direct experimental one.) (2) The sequence between positions 26–57 which includes the common arm base and the functionally significant segment –CCGAAC– is generally the most highly conserved in the molecule. Although clearly similar, the *C. pasteurianum* sequence in this segment is the least typical of all the procaryotes examined to date, including the *A. nidulans* and *C. thermosaccharolyticum* cases.

The detailed phylogenetic implications of the *C. pasteurianum* 5S rRNA sequence will be discussed elsewhere in a broader context [15]. However, it should be noted that the dissimilarity between the 5S rRNAs of *C. pasteurianum* and the *Bacilli* examined to date strongly suggests that the relationship between these two groups of organisms is not so close as has been customarily assumed.

Acknowledgements

This study was supported by NIH Grant AI-6457 and NASA Grant NSG-7044 to C.R.W. We thank K. Pechman for providing crude ^{32}P-labelled 5S rRNA from *Clostridium pasteurianum*.

References

[1] Fox, G. E. and Woese, C. R. (1975) Nature 256, 505–507.
[2] Corry, M. J., Payne, P. I. and Dyer, T. A. (1974) FEBS Lett. 46, 63–66.
[3] Fox, G. E. and Woese, C. R. (1975) J. Mol. Evol. 6, 61–76.
[4] Pechman, K. J. (1975) Ph.D. Thesis, University of Illinois.
[5] Pace, N. R., Pato, M. L., McKibbin, J. and Radcliffe, C. W. (1973) J. Mol. Biol. 75, 619–631.
[6] Woese, C. R., Pribula, C. D., Fox, G. E. and Zablen, L. B. (1975) J. Mol. Evol. 5, 35–46.
[7] Brownlee, G. G., Sanger, F. and Barrell, B. G. (1968) J. Mol. Biol. 34, 379–412.
[8] Woese, C. R., Sogin, M., Stahl, D., Lewis, B. J. and Bonen, L. J. Mol. Evol., in the press.
[9] Uchida, T., Bonen, L., Schaup, H. W., Lewis, B. J., Zablen, L. and Woese, C. R. (1974) J. Mol. Evol. 3, 63–77.
[10] Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. and Gralla, J. (1973) Nature, New Biology 246, 40–42.
[11] Erdmann, V. A., Sprinzl, M. and Pongs, O. (1973) Biochem. Biophys. Res. Commun. 54, 942–948.
[12] Richter, D., Erdmann, V. A. and Sprinzl, M. (1973) Nature, New Biology 246, 132–135.
[13] Herr, W. and Noller, H. F. (1975) FEBS Lett. 53, 248–252.
[14] Sutton, L. A. and Woese, C. R. (1975) Nature 256, 64–66.
[15] Fox, G. E. and Woese, C. R., in preparation.