NUCLEOTIDE SEQUENCE OF *BACILLUS MEGATERIUM* 5 S RNA

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

and

Mitchell SOGIN and Norman PACE
National Jewish Hospital, Denver, Colorado, USA

Received 11 May 1974

1. Introduction

5 S ribosomal RNA has exhibited a remarkable stability during evolution. It is the smallest molecule which spans the gap, although not in a completely straightforward manner [1], between the procaryotes and eucaryotes. The architecture of this RNA, perhaps more so than tRNA, may be typical of the larger ribosomal RNAs and even RNAs in general. 5 S RNA contains base pairing, single stranded stretches and regions of both RNA-protein [2, 3], and RNA–RNA interactions [4, 5]. Sequencing investigations should ultimately illuminate the general principles behind 5 S RNA structure and its relationship to function. Already comparative sequence studies have convincingly shown (1) the presence of a base paired region the existence of which had previously been speculative at best, and (2) the important relation between environmental conditions and primary sequence [6]. Further inquiry of this type should be fruitful if the choice of organisms is judiciously made.

As a preliminary to our work pancreatic and T1 RNase oligonucleotide catalogs were developed for several Bacilli 5 S RNAs. Of these organisms *Bacillus megaterium* was selected for sequencing since its potential for yielding information seemed greater than that of the other organisms considered. Now that this sequence is available, it is possible to construct by analogy, complete or nearly complete sequences of these other Bacilli. Such a procedure will allow a family analysis similar to that which has been made for the Enteric group [7]. These results will be reported elsewhere [8].

2. Materials and methods

^{32}P-labelled 5 S RNA was prepared from *Bacillus megaterium* KM (obtained from J. T. Wachsman, University of Illinois) as described previously [9]. The sequence was derived by established methods [6,10] based on two-dimensional paper electrophoresis. Purified 5 S RNA was digested to completion with pancreatic RNase or with RNase T1. The digestion products were fingerprinted and the resulting oligonucleotides sequenced by secondary and tertiary procedures utilizing pancreatic, T1 or U_2 RNase [11]. In order to obtain larger oligonucleotides and thus create sufficient overlaps to obtain the sequence, partial digestions with T1 or U_2 RNase were performed [6,10]. The resulting fragments were again separated by two dimensional electrophoresis using an 'extended resolution' second dimension [6,11] and sequenced by conventional means [6,10]. The details of these procedures appear elsewhere [8].

* To whom requents for reprints should be sent.

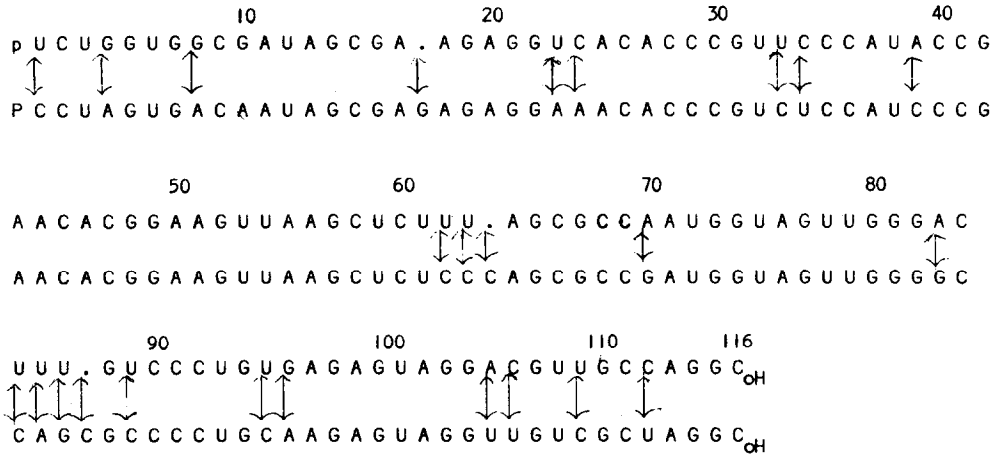

Fig. 1. Nucleotide sequence of *B. megaterium* (upper) with points of comparison to the published *B. stearothermophilus* sequence (below). The unsequenced portions of latter have been arranged to maximum homology.

3. Results and discussion

The complete sequence of *B. megaterium* 5 S RNA is shown in fig. 1. This sequence is clearly related to that of *B. stearothermophilis* 5 S RNA, but differs from the latter by at least twenty-three base replacements and is three nucleotides shorter in overall length [12]. How much of this represents 'fortuitous' evolutionary variation, how much represents adaptation to a thermophilic vs mesophilic niche, and how much represents mere experimental error remains to be determined.

Three interesting insights provided by the *B. megaterium* 5 S RNA sequences are: (1) The sequence – U G G G G U C U C C C C A – of the Enteric and Pseudomonad 5 S RNAs [7,10,13] which is believed to form a tight (i.e. three base) hairpin loop is here replaced by – G G G A C U U U G U C C C – which is substantially different in primary sequence but is still able to form such a loop. Thus it is virtually certain that this loop is present in at least one functionally meaningful conformation of the procaryotic 5 S RNA molecule. (2) This is the first procaryotic sequence to show any variation whatever, in what are here positions 66–78. The oligonucleotide – C G C C *A* A U G G U A G U – found here has one base change from the normal – C G C C *G* A U G G U A G U. (3) The ten base pairs of the 'stalk' region of the molecule include three contiguous G–U base pairs.

References

[1] Fox, G. E. and Woese, C. R. (1974) manuscript to be submitted to the J. of Molec. Evolution.
[2] Gray, P. N. and Monier, R. (1972) Biochimie 54, 41–45.
[3] Gray, P. N., G. Bellemare, and R. Monier (1973) J. Mol. Biol. 77, 133–152.
[4] Erdmann, V. A., M. Sprinzl and Pongs, O. (1973) Biochem. Biophys. Res. Commun. 54, 942–948.
[5] Richter, D., Erdmann, V. A. and Sprinzl, M. (1973) Nature New Biology 246, 132–135.
[6] Woese, C. R., Pribula, C., Fox, G. E. and T. Uchida (1974) Manuscript submitted to the J. of Molec. Evolution.
[7] Sogin, S. J., Sogin, M. L. and Woese, C. R. (1972) J. Molec. Evolution 1, 173–184.
[8] Pribula, C., Fox, G. E., Woese, C. R., Sogin, M. L. and Pace, N. R. Manuscript in preparation.
[9] Pace, N. R., Pato, M. L., McKibbin, J. and Radcliffe, C. W. (1973) J. Mol. Biol. 75, 619–631.
[10] Brownlee, G. G., Sanger, F. and Barrell, B. G. (1968) J. Mol. Biol. 34, 379–412.
[11] Uchida, T., Bonen, L., Schaup, H. W., Lewis, B. J., Zablen, L. and Woese, C. R. (1974) J. Molec. Evolution 3, 63–77.
[12] Marotta, C. A., Levy, C. C., Weissman, S. M. and Varicchio, F. (1973) Biochemistry 12, 2901–2904.
[13] Duboy, B. and Weissman, S. M. (1971) J. Biol. Chem. 246, 747–761.