Unusual structural features of the 5S ribosomal RNA from *Streptococcus cremoris*

Harold Neimark, Janet Andersen⁺ and Nicholas Delihis⁺

Department of Microbiology and Immunology, College of Medicine, State University of New York, Brooklyn, NY 11203, and ⁺Department of Microbiology, State University of New York, Stony Brook, NY 11794, USA

Received 9 August 1983; Revised and Accepted 10 October 1983

ABSTRACT

The nucleotide sequence of the 5S ribosomal RNA of *Streptococcus cremoris* has been determined. The sequence is 5' UUUUGGUCAUCAUUGCGAUGGAGAUACACCUGUCCCAUGUCGAACACAGAAGUUAAGUCC AUCUACGCGGGAAGUACUUGGGGUUGCCCCUGGGAGAUAGCGAGUGGCCAAGU^{OH} 3'. Comparison of the *S. cremoris* 5S RNA sequence to an updated prokaryotic generalized 5S RNA structural model shows that this 5S RNA contains some unusual structural features. These features result largely from uncommon base substitutions in helices I, II and IV. Some of these unusual structural features are shared by several of the known 5S RNA sequences from mycoplasmas. However, the characteristic bloc of deletions found in helix V of these mycoplasma 5S RNAs is not present in the 5S RNA of *S. cremoris*.

INTRODUCTION

The 5S ribosomal RNA has been studied extensively from the viewpoint of molecular function (1) and this structurally conserved molecule also has been utilized for phylogenetic studies (2-7). Here we report the nucleotide sequence of the 5S RNA from the Gram positive lactic acid bacterium *Streptococcus cremoris* and characterize several unusual features in its structure that have not been described before. Our analysis is based on a generalized 5S RNA structural model for prokaryotes which was derived from generalized models developed previously (8-10). We also utilized our model to analyze the published sequences of the 5S RNAs found in each of three fermentative mycoplasmas (11,12) that are believed to be descendants of Gram positive bacteria.

MATERIALS AND METHODS

S. cremoris (ATCC 11603) was grown in MRS medium (Difco), the cells collected, ground with twice their wet weight of

alumina (Alcoa, bacteriological grade A-305) and extracted with 10 mM $MgCl_2$, 60 mM NH_4Cl , 6 mM 2-mercaptoethanol in 10 mM Tris-HCl, pH 7.8. Crude ribosomes were obtained from the supernatant (13), taken up in 1 mM $lgCl_2$, 0.5% SDS, 10 mM Tris-HCl, pH 7.6, and treated with 10 μg proteinase K/ml for 5 min at room temperature. An equal volume of 1.0 M KOAc was added, the suspension kept on ice for 30 min and centrifuged. The RNA fraction was precipitated from the supernatant with ethanol and purified by chromatography on Sephadex G-100 and by electrophoresis in a non-denaturing 10% polyacrylamide gel. The purified 5S RNA was labeled at the 5'-end or the 3'-end and sequenced as described previously (14-18). All positions were confirmed using at least two independent sequencing methods.

Areas of band compression on sequencing gels caused difficulties in resolving the sequence of certain regions of the S. cremoris 5S RNA molecule (e.g. helix V). Other regions of the molecule contained unusual sequences for a 5S RNA (e.g. helix IV). The sequences of these regions were resolved by two dimensional mobility shift analysis (17,18). Unlabeled 5S RNA (1 μg) was partially hydrolyzed in d. formamide- $MgCl_2$ (4 mM) at 100°C for 5 minutes and then dried overnight. The hydrolyzed RNA was subsequently labeled using [γ - ^{32}P] ATP and T4 polynucleotide kinase (P.-L. Biochemicals). The labeled 3'-fragments were denatured in 20 mM NaCitrates, pH 5, 1.3 mM EDTA, 7 M urea at 80°C for 1 min and were separated on 8% and 12% polyacrylamide 7 M urea sequencing gels using 0.5 M Tris-Borate, pH 8.3, 10 mM EDTA buffer. Using an autoradiogram as a template, bands that corresponded to the region of the molecule where the sequence needed to be confirmed were cut from the gel. The RNA fragments in these bands were eluted from the gel in 0.1% SDS, 0.3 M NaCl with 0.2 O.D. carrier tRNA/ml by shaking overnight at 37°C. The eluant was pipetted away from the gel slice and the RNA was precipitated with ethanol (16). These 5'-labeled RNA fragments were hydrolyzed in d. formamide- $MgCl_2$ (4 mM) at 100°C for 30 min. After being dried overnight, the RNA was resuspended in 2 μl distilled water and assayed using mobility shift analysis (17,18). Parts of the sequence needing special attention were confirmed using this method.

Figure 1. The sequence of the *Streptococcus cremoris* 5S RNA drawn in accord with the prokaryotic secondary structure model (Fig. 2A).

RESULTS AND DISCUSSION

The *S. cremoris* 5S RNA is 117 nucleotides long, has a GC content of 53% (G:C:A:U = 34:26:27:30) and like most other 5S RNAs sequenced to date, appears to contain no modified nucleotides. The nucleotide sequence of the 5S ribosomal RNA from *S. cremoris* is shown in Fig. 1. The sequence is drawn in accordance with the revised prokaryotic generalized secondary structural model shown in Fig. 2A. A number of unusual features occur in the *S. cremoris* 5S RNA and these are described with respect to the generalized model.

The prokaryotic generalized structure (Fig. 2A) was developed from alignment of the known bacterial 5S RNA sequences (19,20) and from the generalized 5S RNA structure published earlier (8,9). As pointed out previously (9), this model contains five double-helical regions as proposed by Studnicka and co-workers (21) and by Nishikawa and Takamura (22) for prokaryotic 5S RNAs. This model is a modification of the Fox and Woese base-pairing scheme originally derived by comparative sequence analysis of 5S RNAs (23) but includes extensions of the base paired regions and an additional helix (helix IV) (22). The base paired regions have been extended still further by the concept that guanine and adenine can form base pairs (24,25,20). The prokaryotic generalized structure contains a number of common

positions distributed throughout the molecule that occur in all prokaryotic 5S RNAs in addition to the universal positions (those positions found in all 5S RNA classes) (9).

More specific generalized structures were derived for Gram positive bacteria (Fig. 2B) and Gram negative bacteria (Fig. 2C) and these contain five definitive Gram signature positions at which either of two alternate bases specify the Gram type. The five Gram positive signatures are: A₂₇, U₃₄, A₄₈, U₅₆ and G₅₉; the five Gram negative signatures are: C₂₇, A₃₄, U₄₈, G₅₆ and A₅₉. In addition to these signatures, "partial" Gram signature positions occur, i.e. those where the position is always occupied by a single specific base in one of the Gram types but may be occupied by any base in the other Gram type. The Gram positive partial signatures are A₁₃ and G₇₂; Gram negative partial signatures are C₁₁ and G₈₁. Hori and Osawa (4) have already emphasized that two nucleotide length classes occur among eubacterial 5S RNAs, 116-N or 117-N for Gram positive bacteria and 120-N for Gram negative bacteria. Exceptions where sequences contain features intermediate between typical Gram positive and Gram negative bacterial 5S RNAs occur (26,27).

Comparison of the *S. cremoris* 5S RNA structure (Fig. 1) to the prokaryotic generalized structural model (Fig. 2A) and the Gram positive generalized structural model (Fig. 2B) shows that this molecule conforms to the prokaryotic generalized structural model in that it contains most of the common prokaryotic positions and all of the Gram positive signatures are present. However, this comparison also reveals a number of features that

Figure 2. Generalized models for (A) Prokaryotic, (B) Gram positive and (C) Gram negative bacterial 5S RNA secondary structures. Nucleotide position numbers correspond to the *Escherichia coli* 5S RNA. The double-helices are labeled with Roman numerals and the single stranded loop regions are labeled with the letters "a" through "g" starting at the 5'-end. The positions that are common among the known prokaryotic 5S RNA sequences are indicated as explicitly as possible and are representative of more than 90% of the known bacterial sequences (19,20). Bold letters (A) denote universal positions and brackets indicate generalized chain lengths between universal or prokaryotic common positions. The Gram positive (B) and Gram negative (C) signature positions are indicated by black arrow heads. The partial Gram signature positions are indicated by open arrowheads. Abbreviations: Pu, purine; Py, pyrimidine.

are either unusual or unique to the S. cremoris 5S RNA. These include:

1. In helix I, the base-paired positions C₈ and G₁₁₀ are reversed from the equivalent universal positions G₉-Py₁₁₁.
2. In loop a, a pyrimidine, U₁₄, is substituted at the universal purine 15 position. (The only other purine to pyrimidine substitution known at this position in eubacteria is in Rhodospirillum rubrum 5S RNA, also a U) (28).
3. In helix II, a substitution of an A at position 66 in place of the common prokaryotic G results in mispairing with the common prokaryotic C at position 16. This substitution is also found in some mycoplasma 5S RNAs (see below).
4. Helix IV and loops e and g adjacent to helix IV in particular contain a concentration of several alterations: In helix IV of the S. cremoris 5S RNA structure, the base-pair G₆₉-C₁₀₅ (Fig. 1) is reversed (see Fig. 2B, equivalent base-pair C₇₀-G₁₀₆). This same reversal has been observed in cyanobacteria and chloroplast 5S RNAs (29). In loop e, between positions A₇₃ and U₇₆ (Fig. 2B, equivalent positions A₇₃ and U₇₇), either the A/U₇₄ or one purine, 75 or 76, has been deleted. Similarly in loop g (Fig. 2B), one purine deletion has occurred between positions G₁₀₀ and U₁₀₃. In addition, in helix IV (see Fig. 1), it appears that the nucleotide-pair G₇₁-U₁₀₃ is probably an insertion. This insertion not only maintains the chain length of the entire limb between position G₆₈ and position G₁₀₈ but also maintains an extended helix IV (five pairs can be formed). Thus, major alterations have occurred in this region but the basic integrity of this limb of the 5S RNA molecule has been maintained. The conservation of structure and overall chain length in this region may not be fortuitous but rather may be crucial for the function of the 5S RNA molecule.

The S. cremoris 5S RNA was selected for sequencing because previously it had been shown that one large division of the mycoplasmas, called acholeplasmas, descended from streptococci and it was concluded also that the other mycoplasmas similarly evolved from various other walled bacterial progenitors (30,31). It was proposed that mycoplasmas, which are characterized by small genome sizes, arose from their walled progenitors through

losses of large amounts of genetic material including genes coding for portions of the cell envelope and for RNAs. Woese et al. (32) also concluded that mycoplasmas arose by degenerative evolution from bacteria and described two group III Clostridium species as the bacteria most closely related to mycoplasmas, spiroplasmas, and acholeplasmas.

5S RNA sequences from three mycoplasmas, Mycoplasma capricolum (11), M. mycoides subsp. capri, and Spiroplasma sp. BC3 (12) are available and each is 107 nucleotides in length. All three mycoplasma 5S RNAs were recognized as being more closely related to Gram positive bacteria than to Gram negative bacteria and a few of the deletions responsible for the short chain lengths were identified (11,12). The generalized models described here are particularly valuable for analyzing sequences where nucleotides are deleted as in the case of the 5S RNAs from mycoplasmas, chloroplasts and mitochondria. Comparison of the mycoplasma short structures to the Gram positive model shows that the most extensive alterations occur in the 5S RNAs from M. mycoides subsp. capri and M. capricolum and these are quite similar to one another in pattern. Each of these 5S RNAs as well as the Spiroplasma BC3 5S RNA has sustained a bloc deletion at the base of Helix V. These deletions are not present in the S. cremoris 5S RNA. On the other hand, the S. cremoris 5S RNA shares several unusual structures with the 5S RNAs from these mycoplasmas. One unusual base pair found in the S. cremoris and the two Mycoplasma 5S RNAs is a GC reversal in helix IV. All three of these mycoplasma 5S RNAs share with the S. cremoris 5S RNA at position 66 an unusual A which results in mispaired bases at C₁₆-A₆₆. They also share a U at looped-out position 65 which in most Gram positive bacterial 5S RNAs is an A. The 5S RNA comparisons are consistent with a phylogenetic relationship between S. cremoris and the mycoplasmas, but the data in hand suggest that S. cremoris may be related only distantly to these fermentative mycoplasmas (30). The divergence of the 5S RNA of S. cremoris from the generalized 5S RNA structure results in an inability to further evaluate the phylogenetic relationship between S. cremoris and the mycoplasmas.

ACKNOWLEDGMENTS

Supported by grants National Institutes of Health GM-20052, National Science Foundation PCM-83-02127 and a State University of New York faculty award.

REFERENCES

1. Erdmann, V.A. (1976) Prog. Nucleic Acids Res. Mol. Biol. **18**, 45-90.
2. Woese, C.R., Luehrsen, K.R., Pribula, C.D., and Fox, G.E. (1976) J. Mol. Evol. **8**, 143-153.
3. Schwartz, R.M., and Dayhoff, M.O. (1978) Science **199**, 395-403.
4. Hori, H., and Osawa, S. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 381-385.
5. Kuntzel, H., Heidrich, M., and Piechulla, B. (1981) Nucleic Acids Res. **9**, 1451-1461.
6. Sankoff, D., Cedergren, R.J., and McKay, W. (1982) Nucleic Acids Res. **10**, 421-431.
7. Kuntzel, H., Piechulla, B., and Hahn, U. (1983) Nucleic Acids Res. **11**, 893-900.
8. Böhm, S., Fabian, H., and Welfle, H. (1981) Acta Biol. Med. Germ. **40**, k19-k24.
9. Delihas, N., and Andersen, J. (1982) Nucleic Acids Res. **10**, 7323-7344.
10. De Wachter, R., Chen, M., and Vendenberghe, A. (1982) Biochimie **64**, 311-329.
11. Hori, H., Sawada, M., Osawa, S., Murao, K., and Ishikura, H. (1981) Nucleic Acids Res. **9**, 5407-5410.
12. Walker, R.T., Chelton, E.T.J., Kilpatrick, M.W., Rogers, M.J., and Simmons, J. (1982) Nucleic Acids Res. **10**, 6363-6367.
13. Zimmermann, R.A. (1979) Meth. Enzymol. **59**, 551-583.
14. Andersen, J., Andresini, W., and Delihas, N. (1982) J. Biol. Chem. **257**, 9114-9118.
15. Donis-Keller, H., Maxam, A.M., and Gilbert, W. (1977) Nucleic Acids Res. **4**, 2527-2538.
16. Stanley, J., and Vassilenko, S. (1978) Nature (Lond.) **274**, 87-89.
17. Pirtle, R., Kashdan, M., Pirtle, I., and Dudock, B. (1980) Nucleic Acids Res. **8**, 805-815.
18. Silberklang, M., Prochiantz, A., Haenni, A.-L., and RajBhandary, U.L. (1977) Eur. J. Biochem **72**, 465-478.
19. Erdmann, V.A., Huysmans, E., Vendenberghe, A., and De Wachter, R. (1983) Nucleic Acids Res. **11**, r105-r133.
20. Dams, E., Vandenbergh, A., and De Wachter, R. (1983) Nucleic Acids Res. **11**, 1245-1252.
21. Studnicka, G.M., Eiserling, F.A., and Lake, J.A. (1981) Nucleic Acids Res. **9**, 1885-1905.
22. Nishikawa, K., and Takamura, S. (1974) J. Biochem. (Tokyo) **76**, 935-947.
23. Fox, G.E., and Woese, C.R. (1975) Nature **256**, 505-507.
24. Traub, W., and Sussman, J.L. (1982) Nucleic Acids Res. **10**, 2701-2708.

25. Stahl, D.A., Luehrsen, K.R., Woese, C., and Pace, N.R. (1981) Nucleic Acids Res. 9, 6129-6137.
26. Hori, H., Osawa, S., Murao, K., and Ishikura, H. (1980) Nucleic Acids Res. 8, 5423-5426.
27. Simoncsits, A. (1980) Nucleic Acids Res. 8, 4111-4124.
28. Newhouse, N., Nicoghossian, K., and Cedergren, R.J. (1981) Canad J. Biochem. 59, 921-932.
29. Delihias, N., Andresini, W., Andersen, J., and Berns, D. (1982) J. Mol. Biol. 162, 721-727.
30. Neimark, H. (1979) in The Mycoplasmas, M.F. Barile and S. Razin (Eds.), Vol. I, pp. 43-61. Academic Press, New York.
31. Neimark, H., and London, J. (1982) J. Bacteriol. 150, 1259-1265.
32. Woese, C.R., Maniloff, J., and Zablen, L.B. (1980) Proc. Natl. Acad. Sci., U.S.A. 77, 494-498.