Volume 8 Number 19 1980 Nucleic Acids Research

Nucleotide sequence analysis of precursor 5S RNA from Bacillus licheniformis

Willem J.Stiekema, Rita de Leede-Twente, Hendrik A.Raué and Rudi J.Planta

Biochemisch Laboratorium, Vrije Universiteit, de Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Received 28 July 1980

ABSTRACT

The complete nucleotide sequences of the various precursor 5S RNA species occurring in Bacillus licheniformis have been elucidated. The B. licheniformis precursors contain a 5'-precursor-specific segment of 95 nucleotides which is four times as long as the corresponding segment of the p5S RNAs from the closely related strains B. subtilis (Sogin, M.L., Pace, N.R., Rosenberg, M., Weissman, S.M. (1976) J. Biol. Chem. 251, 3480-3488) and Bacillus Q (Stiekema, W.J., Raué, H.A., Planta, R.J. (1980) Nucl. Ac. Res. 8, 2193-2211). However, fourteen of the sixteen nucleotides at the 5'-end are identical in the precursors from all three strains. These conserved nucleotides can form a stem and loop structure which is likely to play an important role in the biosynthesis of 5S RNA.

Extensive secondary and tertiary structure is present in the 5'-precursor-specific segment as concluded from the results of digestion with RNAase T_1 both of the isolated segment and the intact precursors.

No sequence homology exists between the 3'-precursor-specific segments of the B. licheniformis precursors and those of the other two strains except for a stretch of U residues at the 3'-terminus. This stretch of U residues is not immediately preceded by a hairpin loop, however, as expected for a transcription termination signal (20). The question whether the precursors have already undergone processing at the 3'-end, therefore, remains open.

The total number of genetically distinct precursor species in B. licheniformis is at least five and at most ten. Most likely each ribosomal RNA cistron produces a separate p5S RNA as is also the case in Bacillus Q.

INTRODUCTION

The final step in the biosynthesis of 5S ribosomal RNA in Bacilli is the removal of fairly long extra sequences from both ends of precursor 5S RNA molecules by the endonuclease RNAase M5 (1,2,3). The precursors can be isolated, manipulated and subjected to structural analysis relatively easily. Moreover a simple in vitro maturation system is available (1,2,4). All this has made maturation of Bacillus 5S RNA an attractive model system for studying the highly specific interaction between a maturation enzyme and its substrate.

The studies so far carried out (3,5-8), have led to the conclusion that, apart from one or two nucleotides immediately adjacent to the 5'-terminus of the mature domain, the precursor-specific sequences play no part in the recognition of the precursor by RNAase M5. Nevertheless a strong conservation of the 5'-precursor-specific sequence can be observed in the various precursor species of the closely related strains Bacillus Q and B. subtilis (3).

We have previously shown (2) that B. licheniformis, which is also closely related to Bacillus Q and B. subtilis (9), contains precursor 5S RNAs that are appreciably longer than those of the other two strains. The mature 5S RNAs of all three strains have virtually identical sequences (9), so the difference in size is due to the length of the precursor-specific segments. It was therefore interesting to determine the primary structure of the precursor-specific segments of B. licheniformis p5S RNA to see whether the sequence conservation observed in the Bacillus Q and B. subtilis precursors extends to this third strain.

MATERIALS AND METHODS

Bacillus licheniformis (laboratory strain S244) was used throughout this study. Isolation of ^{32}P-labeled precursor 5S RNA, sequencing techniques, isolation of crude RNAase M5 and *in vitro* maturation of p5S RNA have been described before (3).

RESULTS

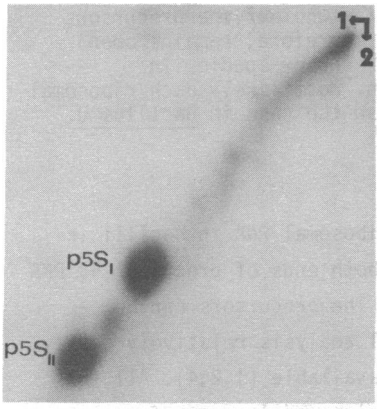

Fig. 1. Two-dimensional slab gel electrophoresis of low-molecular weight RNA from B. licheniformis cells.

Fig. 1 shows an autoradiogram of the second dimension slab gel used for purification of the p5S RNA. Fingerprint analysis (2) demonstrates that spots I and II are p5S RNA species. The RNAase T_1 fingerprints of the precursors shown in Fig. 2 are identical to those published previously by Stoof *et al.* (2). At that time products tV and tT (Fig. 2A) as well as tW (Fig. 2B) were not recognized as precursor-specific oligonucleotides, however. As concluded previously (2), extra nucleotides are present at both ends of the precursors since the mature terminal oligonucleotides pUUUG (t18) and $AAGC_{OH}$ (p7) are lacking from the RNAase T_1

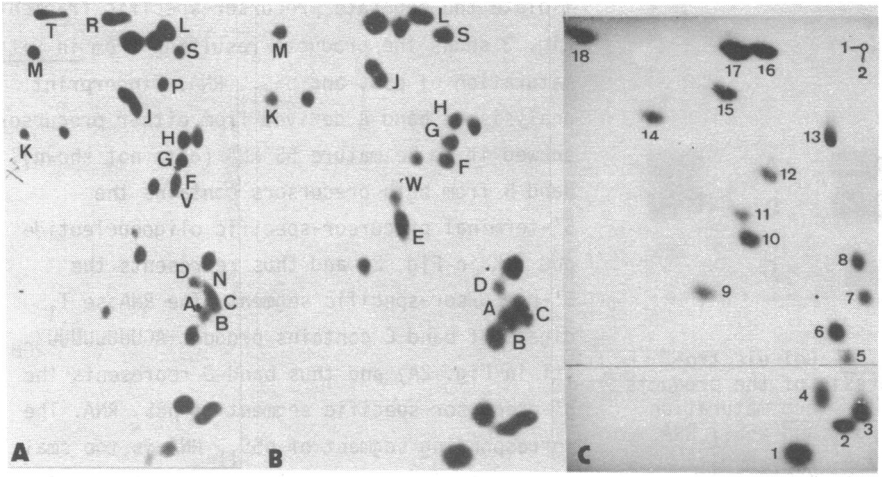

Fig. 2. Ribonuclease T_1 fingerprints of $p5S_I$ RNA (panel A), $p5S_{II}$ RNA (panel B) and m5S RNA (panel C) of B. licheniformis.

and RNAase A digest of both precursor species respectively (Fig. 2).

The lengths of the two precursors is about 240 and 220 nucleotides respectively as estimated from their electrophoretic mobility (Fig. 1) and sequence complexity*. This makes the B. licheniformis p5S RNAs 40-50% longer than their counterparts from either B. subtilis (5) or Bacillus Q (3).

The molar yields of the RNAase T_1 and RNAase A oligonucleotides of $p5S_I$ RNA provide clear evidence for the fact that this RNA consists of multiple species having slight differences in sequence. E.g. products tN and tV (cf. Fig. 4) are each present in a molar yield of 0.5 only. $p5S_{II}$ RNA, however, represents a unique RNA species, since all digestion products are present in unimolar yield. Moreover, $p5S_{II}$ is genetically distinct from $p5S_I$ RNA. Among other things it contains a 3'-terminal oligonucleotide (tW) different from that of $p5S_I$ RNA (tT).

It has been shown, both by Pace and coworkers (1,4) and by us (3) that crude Bacillus RNAase M5 is able to maturate Bacillus p5S RNA correctly in vitro. This maturation produces, in addition to mature 5S RNA, the intact precursor-specific segments. Sequence analysis of the relatively long B. licheniformis p5S RNAs was enormously facilitated by the possibility to

* Because of space limitations the detailed data on the sequence analysis are not included. They are available upon request, however.

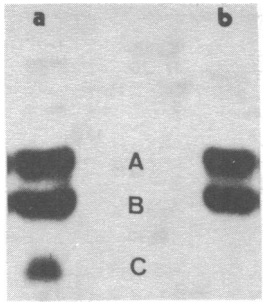

Fig. 3. Gel electrophoresis of the products of in vitro maturation of $p5S_I$ and $p5S_{II}$ RNA. Lane a: $p5S_I$ RNA; lane b: $p5S_{II}$ RNA.

isolate the separate precursor-specific fragments. Fig. 3 shows the products resulting from in vitro maturation of $p5S_I$ and $p5S_{II}$ RNA. Fingerprint analysis of band A derived from either precursor showed it to be mature 5S RNA (data not shown). Band B from both precursors contains the 5'-terminal precursor-specific oligonucleotide pUG (tK in Fig. 2) and thus represents the 5'-precursor-specific segment. The RNAase T_1 digest of band C contains product $ACUUUUUUUUX_{OH}$ (tT in Fig. 2A) and thus band C represents the 3'-precursor-specific segment of $p5S_I$ RNA. The corresponding segment of $p5S_{II}$ RNA is too small to be retained on the gel (see below). The RNAase T_1 digests of bands B and C of $p5S_I$ RNA together account for all precursor-specific products present in the corresponding digests of $p5S_I$ RNA, taking into account the fact that RNAase M5 produces 5'-phosphate and 3'-OH termini. Thus bands B and C represent the intact precursor-specific segments.

Although analysis of the complete RNAase T_1 and RNAase A digests of the 5'-precursor-specific segments allowed us to derive a partial sequence it was not possible to deduce the complete sequence from these data. Therefore the isolated segments were subjected to limited digestion with RNAase T_1 and the products were separated by urea gel electrophoresis. Digestion at an enzyme to substrate ratio of 1:2000 for 15 min at 0^oC was found to produced two fragments that together account for the complete 5'-precursor-specific segment. The RNAase T_1 cut was located between residues 47 and 48 (cf. Figs. 4 and 5) as deduced from analysis of the resulting fragments.

The final ordering of the oligonucleotides of the 5'-non-conserved segment of $p5S_I$ and $p5S_{II}$ RNA was done by analysis of smaller partial RNAase T_1 fragments obtained by digestion at an enzyme to substrate ratio of 1:500 at 0^oC for 15 min. Fig. 4 shows the fragments obtained and the sequence ultimately derived. There is heterogeneity at positions 9 and 10 where in 70% of the molecules the sequence was found to occur whereas the remaining 30% contain the sequence CA at these positions in $p5S_I$ RNA.

Derivation of the 3'-precursor-specific sequence of $p5S_I$ RNA from the analysis of the RNAase T_1 and RNAase A digests was fairly uncomplicated despite the occurrence of some sequence heterogeneity. The derivation of the

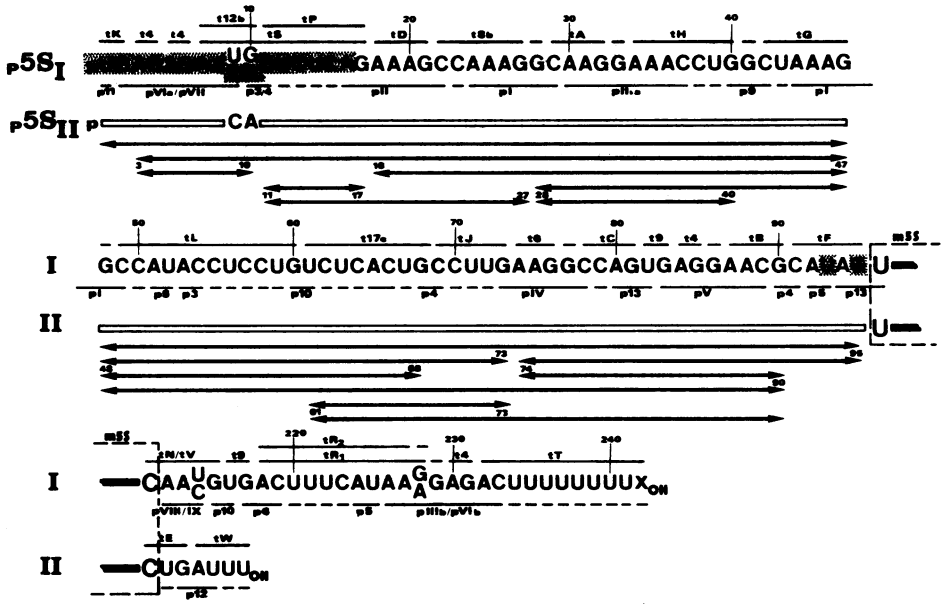

Fig. 4. Nucleotide sequence of the precursor-specific segments of p5S$_I$ and p5S$_{II}$ RNA. The positions of the RNAase T_1 and RNAase A digestion products are indicated above and below the sequence respectively. The partial RNAase T_1 fragments used for deriving the sequence are shown by arrows. The numbers correspond to the positions of the 5'- and 3'-terminal residues. The open bar indicates sequence homology between p5S$_I$ and p5S$_{II}$ RNA. The closed bar represents the sequence of the mature domain (cf. Fig. 5) of which only the 5'- and 3'-terminal nucleotides are shown. Sequence conservation between the p5S RNAs from B. licheniformis, B. subtilis and Bacillus Q is indicated by shading.

complete sequence of the 3'-precursor-specific segment of p5S$_I$ RNA is shown in Fig. 4. Heterogeneity occurs at positions 214 and 228. In both cases the alternative nucleotides are present in 50% of the molecules.

The 3'-precursor-specific segment of p5S$_{II}$ RNA could not be recovered after in vitro maturation (Fig. 3). Therefore the 3'-"half" of this precursor extending from residue C^{135} of the mature domain to the 3'-terminus was prepared by limited RNAase T_1 digestion of intact p5S$_{II}$ RNA, followed by urea gel electrophoresis of the digest (3). Sequence analysis showed this fragment to contain only two precursor-specific oligonucleotides: tE (CUG) and tW ($AUUU_{OH}$). The sequence of the 3'-precursor-specific segment of p5S$_{II}$ RNA is thus established (Fig. 4).

DISCUSSION

A. Number of precursor 5S RNA species in B. licheniformis

Like Bacillus Q (3) and B. subtilis (5), B. licheniformis cells contain two types of p5S RNA differing in length (Fig. 1). The $p5S_{II}$ type represents a unique p5S RNA species genetically distinct from the $p5S_{I}$ type as is evident from its sequence analysis (Fig. 4). The $p5S_{I}$ type however consists of multiple species. The 70/30 heterogeneity at positions 9/10 together with the 50/50 heterogeneity at both position 214 and 218 leads to a minimum of three and a maximum of eight genetically distinct precursor-species of the $p5S_{I}$ type. Neither $p5S_{I}$ nor $p5S_{II}$ RNA contains a precursor to minor B. licheniformis RNA (10). No oligonucleotides specific for minor m5S RNA occur in the fingerprints of the precursors (Fig. 2) nor do they show up in the fingerprints of m5S RNA obtained by in vitro maturation of $p5S_{I}$ and $p5S_{II}$ RNA (data not shown). The total number of distinct p5S RNA species in B. licheniformis cells thus is at least five and at most ten. The number of rRNA cistrons in B. licheniformis is estimated to be six or seven based on the ratio of minor to major 5S RNA (10). It seems likely, therefore, that as in Bacillus Q (3), each rRNA cistron in B. licheniformis produces a separate p5S RNA. The sequence differences between the various p5S RNA species in B. licheniformis, however, are much less extensive than in Bacillus Q. In this respect B. licheniformis more closely resembles B. subtilis (5).

B. Structure of the B. licheniformis p5S RNAs

The most striking feature of the B. licheniformis p5S RNAs, when compared to those from the closely related strains B. subtilis and Bacillus Q, is the great length of the 5'-non-conserved segment. In the B. licheniformis precursors this segment is four times as long as in the precursors from the two other strains (3,5). Therefore it is the more significant that at least fourteen of the sixteen 5'-terminal nucleotides are identical to those found at the 5'-end of the shorter B. subtilis and Bacillus Q precursors (cf. Fig. 4). This conserved sequence can form a stem and loop structure (Fig. 5). Evidence supporting the actual existence of such a structure has been presented (11). The strong conservation of this feature throughout all Bacillus p5S RNA species sequenced to date, makes it highly likely that it plays an important role in the biosynthesis of Bacillus 5S RNA probably as a recognition site for a processing event preceding maturation by RNAase M5.

As in B. subtilis and Bacillus Q p5S RNA (3,5) a G residue immediately precedes the 5'-terminus of the mature domain (Figs. 4 and 5). The base-

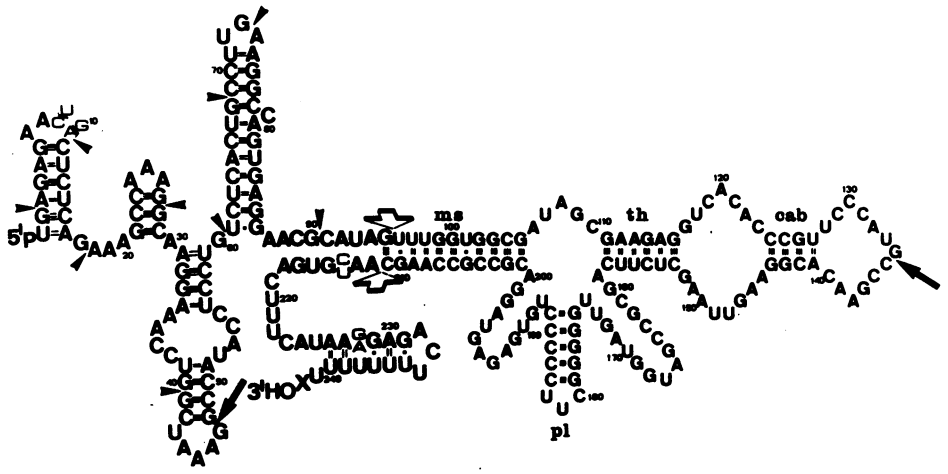

Fig. 5. Possible base-pairing scheme for p5S$_I$ RNA. The mature domain is folded according to the model proposed by Fox and Woese (15). The folding of the precursor-specific segments is maximized, taking into account the results of limited RNAase T_1 digestion. The large arrows indicate bonds cleaved by RNAase T_1 at an enzyme:substrate ratio of 1:2000, the small arrows denote the bonds cleaved at a ratio of 1:500, and the open arrows point to the RNAase M5 processing sites.

pairing of this G to the 3'-terminal C residue of the mature domain (Fig. 5) has been shown to be essential for efficient maturation of p5S RNA by RNAase M5 (7). The only other nucleotide in the immediate neighbourhood of the processing site conserved in all Bacillus p5S RNAs is residue U^{93} (Fig. 4). As pointed out before (3) this residue may play a role in determining the efficiency of maturation. Artificial p5S RNA containing CCCCCCG at the 5'-end is maturated with only about half the efficiency of p5S RNA carrying either UUUG or the complete 5'-non-conserved sequence (7).

The results of limited RNAase T_1 digestion of the 5'-precursor-specific segment indicate the presence of a relatively large amount of secondary and/or tertiary structure. Only the G residue at position 47, and to some small extent that at position 73, is cleaved by RNAase T_1 at an enzyme:substrate ratio of 1:2000 (Fig. 4). Additional cuts are introduced when the enzyme concentration is raised four fold but even then several G residues remain relatively resistant to cleavage (Figs. 4 and 5). Fig. 5 shows a secondary structure model in which base-pairing in the 5'-precursor-specific segment has been maximized. Residues G^{47} and G^{73} occur in loops in

this model in agreement with their relatively good accessibility to RNAase T_1. The model does not fully account for the observed sensitivities of these and other G residues relative to each other, however. Neither, for that matter, do other possible base-pairing schemes. Tertiary folding of the segment, therefore, must be responsible for the high relative resistance of several G residues to RNAase T_1 cleavage

There seems to be no extensive effect of the long 5'-precursor-specific segment on the conformation of the mature domain and vice-versa. The sensitivity of residue G^{134} to RNAase T_1 is about the same irrespective of whether precursor or mature 5S RNA is treated with the enzyme. The same holds true for residues G^{47} and G^{73} of the 5'-precursor-specific segment. The absence of a mutual effect on the conformation of the two segments is in agreement with earlier conclusions that RNAase M5 recognizes mainly the mature domain of the precursor (8) and the finding that the enzyme also recognizes mature Bacillus 5S RNA (12).

The relative insensitivity to RNAase T_1 of the G residues in the 3'-precursor-specific segment of $p5S_I$ RNA supports the base-pairing scheme shown for this segment in Fig. 5. Like the B. subtilis (5) and Bacillus Q (3) p5S RNAs those of B. licheniformis end in a stretch of U residues although in the case of $p5S_{II}$ RNA this stretch is only three nucleotides long (Fig. 4). Nevertheless this feature does not constitute a typical transcription termination signal due to the lack of a hairpin loop immediately preceding the stretch of U residues (13). In this respect the B. licheniformis precursors are different from B. subtilis $p5S_A$ RNA (5) and resemble those of Bacillus Q (3). It remains to be seen whether in B. licheniformis and Bacillus Q transcription stops immediately behind the p5S RNA gene, as it appears to do in B. subtilis, or whether the stop signal is located further downstream as in at least one ribosomal transcription unit of E. coli (14).

ACKNOWLEDGEMENTS

This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.). We gratefully acknowledge Miss M.M.C. Duin for skilful technical assistance. Mrs. P.G. Brink deserves special mention for maintaining a cheerful appearance throughout several retypings of the manuscript.

REFERENCES

1. Sogin, M.L. and Pace, N.R. (1974) Nature 252, 598-600
2. Stoof, T.J., de Regt, V.C.H.F., Raué, H.A. and Planta, R.J. (1974) FEBS Lett. 49, 237-241
3. Stiekema, W.J., Raué, H.A. and Planta, R.J. (1980) Nucl. Acids Res. 8, 2193-2211
4. Sogin, M.L., Pace, B. and Pace, N.R. (1977) J. Biol. Chem. 252, 1350-1357
5. Sogin, M.L., Pace, N.R., Rosenberg, M. and Weissman, S.M. (1976) J. Biol. Chem. 251, 3480-3488
6. Meyhack, B., Pace, B. and Pace, N.R. (1977) Biochemistry 16, 5009-5015
7. Meyhack, B., Pace, B., Uhlenbeck, O.C. and Pace, N.R. (1978) Proc. Natl. Acad. Sci. USA 75, 3045-3049
8. Meyhack, B. and Pace, N.R. (1978) Biochem. 17, 5804-5810
9. Raué, H.A., Rosner, A. and Planta, R.J. (1977) Mol. Gen. Genet. 150, 185-193
10. Raué, H.A., Heerschap, A. and Planta, R.J. (1976) Eur. J. Biochem. 68, 169-176
11. Stahl, D.A., Walker, T.A., Meyhack, B. and Pace, N.R. (1979) Cell 18, 1133-1143
12. Stiekema, W.J., Duin, M.M.C., Raué, H.A. and Planta, R.J. (1980) NAR, submitted
13. Adhya, S. and Gottesman, M. (1978) Ann. Rev. Biochem. 47, 967-996
14. Morgan, E.A., Ikemura, T., Lindahl, L., Fallon, A.M. and Nomura, M. (1978) Cell 13, 335-344
15. Fox, G.E. and Woese, C.R. (1975) Nature 256, 505-507