The secondary and tertiary structure of the 5 S rRNA from the horsetail *Equisetum arvense*

Martin Digweed, Norbert Ulbrich and Volker A. Erdmann*

Institut für Biochemie, Freie Universität Berlin, Thielallee 63, Otto-Hahn-Bau, D-1000 Berlin 33 (Dahlem), Germany

Received 20 August 1984

The secondary and tertiary structures of the 5 S rRNA from *Equisetum arvense* were examined using double-strand and single-strand specific nucleases, the bisulphite mediated deamination of unpaired cytidines and the ethylnitrosourea modification of accessible phosphodiesters. The results are evaluated in terms of the consensus eukaryotic 5 S rRNA structure, the occurrence of 'looped-out' nucleotides, structure of loops and a putative tertiary structure.

Eukaryotic 5 S rRNA *Single-strand specific nuclease* *Double-strand specific nuclease* *Chemical modification* *RNA structure* *Tertiary structure*

1. INTRODUCTION

In recent years, the powerful technique of sequence comparison has been applied to the question of 5 S rRNA structure leading to a consensus structure for this molecule [1–3]. This approach has been accompanied by a refinement in the scoring values for base-pair stacking interactions made in [4,5] based on the thermodynamic measurements in [6,7].

What is now required is experimental data on an appropriately large selection of 5 S rRNAs in order to evaluate the consensus models and define those structures occurring in natural RNA molecules. We have recently reported S1-nuclease [8], cytidine modification (submitted) and adenine modification [9] data for a range of 5 S rRNAs in addition to studies on spinach chloroplasts 5 S rRNA [10] and *E. coli* 5 S rRNA ([11], reviewed in [14,15]).

Here, we present the structural characterisation of the 5 S rRNA from *Equisetum arvense*, the sequence of which we have recently determined [14], using the two complementary nucleases S1 and CSV which are single- and double-strand specific, respectively. Furthermore, we have used the sensitive bisulphite modification of cytidines to confirm base-pairing and monitor dynamic effects such as 'helix-fraying'. Finally, we have examined the tertiary structure of the molecule by analysis of the accessibility of phosphodiester bonds to alkylation by ethylnitrosourea.

2. EXPERIMENTAL

5 S rRNA was isolated as in [14] and 3′- or 5′-end labelled using standard techniques ([15,16] respectively). S1-nuclease digestion was carried out as in [10] and RNase CSV digestion with our own enzyme preparation (unpublished) as follows: 4 μg carrier RNA with (3′-^{32}P)-labelled *E. arvense* 5 S rRNA were incubated with 0.08 units RNase CSV in 100 μl buffer, 50 mM Tris–HCl (pH 6.8), 50 mM NaCl, 5 mM $MgCl_2$ for 30 min at 25°C. One unit of enzyme activity produces an increase of 10^{-3} A_{260} units·min^{-1} in the hydrolysis of 0.737 A_{260} units *E. coli* 16 S and 23 S rRNA at 20°C in TR-buffer (30 mM Tris–HCl, pH 7.4, 23°C, 20 mM $MgCl_2$, 350 mM KCl, 6 mM 2-mercaptoethanol). The fragmented 5 S rRNA was precipitated with ethanol (−20°C, 2 h) and loaded onto a sequenc-

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

ing gel next to the S1-nuclease digest and guanosine and base-unspecific ladders generated as in [17].

The cytidine modification and identification of modifiable sites were carried out as in [10]. Modification of the 3′-terminal cytidine was detected by RNase T_2 digestion of modified (3′-^{32}P)-labelled 5 S rRNA and PEI thin-layer chromatography of the 5′-nucleoside phosphates as in [18]. The radioactive spots were excised and counted in order to quantify the extent of modification. The ethylnitrosourea modification was carried out as in [19].

3. RESULTS AND DISCUSSION

The sequencing gel for location of the S1-nuclease and CSV RNase cleavage sites is shown in fig.1. The cytidines converted to uridines after reaction with bisulphite are identified as additional bands in the uridine specific sequence-ladder of fig.2. The terminal cytidine is represented by the bisphosphate 3′,5′-[5′-^{32}P]pCp which, due to its high charge to mass ratio, is easily lost from sequencing gels. Thus, modification of this terminal cytidine was established by RNase T_2 digestion followed by PEI thin-layer chromatography as shown in fig.3.

3.1. *Helix I*

In the report of the *E. arvense* sequence [14] we showed two alternative conformations for helix I:

OH P		OH
C–G_1		C_{120}
C U		C–G^P
C–G		C–G^U
C–G		C–G
A–U	⇄	A–U
C–G		C–G
G–C		G–C
C–G		C–G
C–G		C–G
G–U		G–U
↑ ↓		↑ ↓
A		B

In conformation A there is a mismatch at the second position whilst in B there is a 'looped-out' uridine, as implicated in other helices in this and

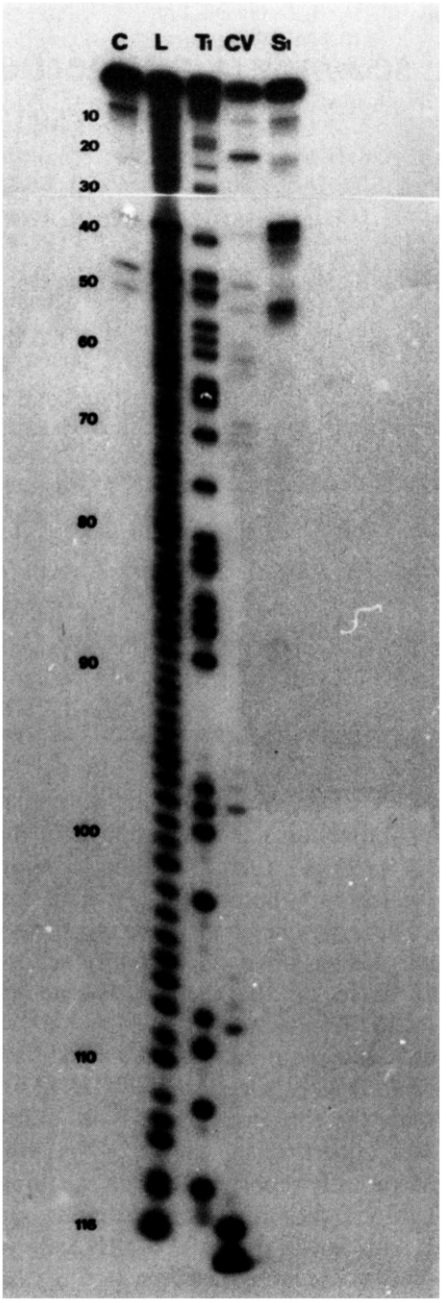

Fig.1. Sequencing gel for the location of nuclease S1 and RNase CSV cleavage sites in (3′-^{32}P)-labelled 5 S rRNA. C, control incubation in the absence of enzyme; L, base unspecific ladder; T1, guanine specific ladder; S1, nuclease S1 digest; CSV, RNase CSV digest. The nucletodide residues are numbered on the left. Electrophoresis was at 150 W on a 60 cm × 0.2 mm, 10% polyacrylamide gel in 8 M urea.

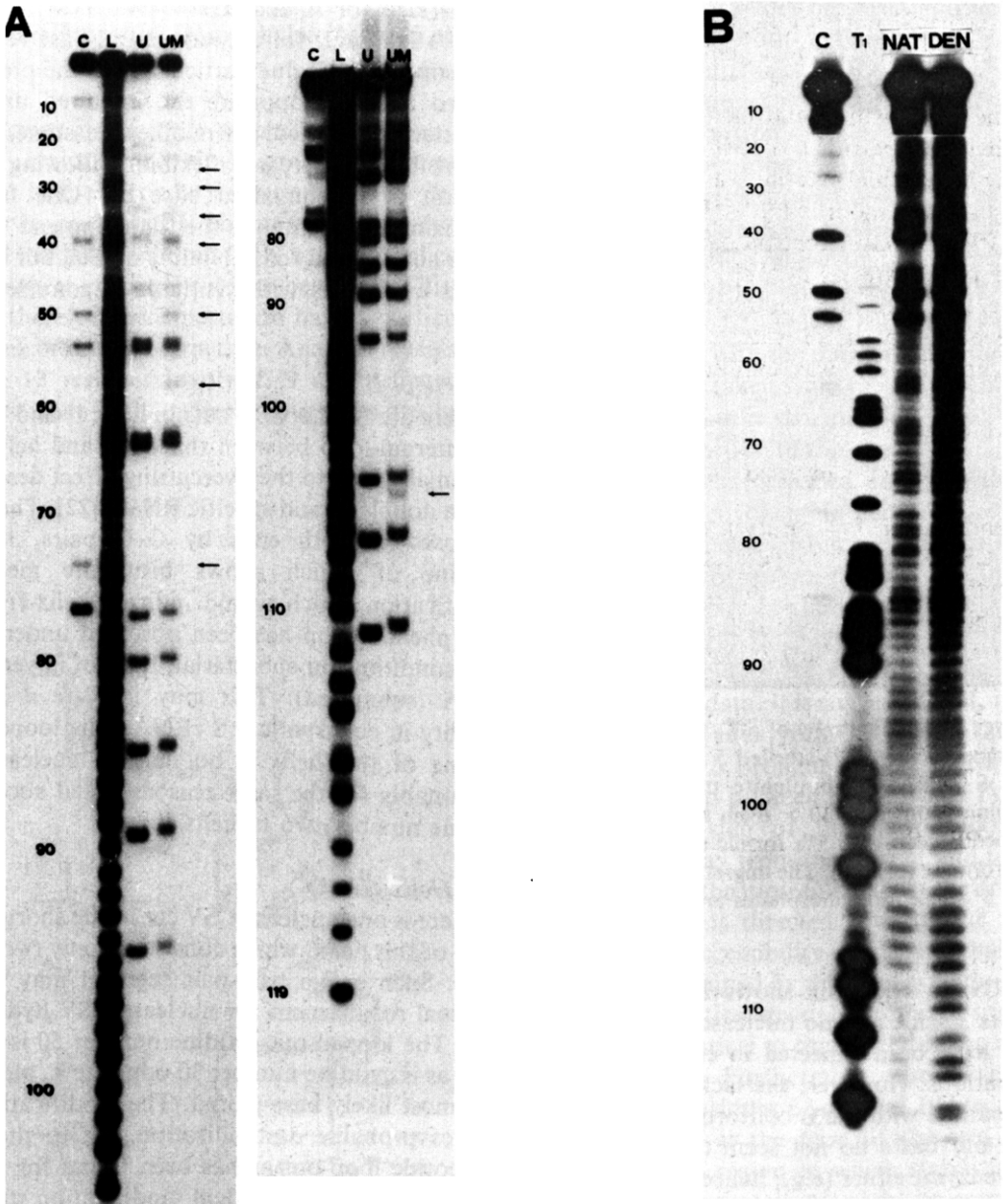

Fig.2. Sequencing gels for the location of: (A) modifiable cytidines and (B) alkylation resistant phosphodiester bonds in (3′-^{32}P)-labelled 5 S rRNA. C, control incubation in the absence of reagent; L, base unspecific ladder; U, uridine-specific sequencing reaction; UM, uridine-specific sequencing reaction on bisulphite-modified 5 S rRNA; Tl, guanine-specific ladder; DEN, 5 S rRNA alkylated with ethylnitrosourea under denaturing conditions followed by alkaline hydrolysis at modified phosphodiester bonds; NAT, 5 S rRNA alkylated under native conditions followed by alkaline hydrolysis at modified phosphodiester bonds.

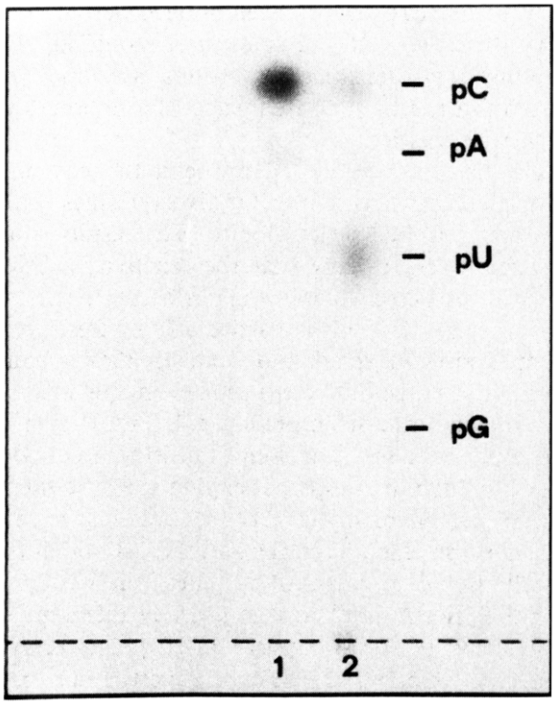

Fig.3. Thin-layer chromatography of a T_2 RNase digest of bisulphite modified (3′-^{32}P)-labelled 5 S rRNA. (1) unmodified 5 S rRNA; (2) bisulphite modified 5 S rRNA. Plate dimensions were 10 × 6 cm and the chromatography media were: (i) 0.5% formic acid; and (ii) 0.3 M lithium formate (pH 3.0). The migration positions of the four 5′-nucleotide phosphates are shown.

other 5 S rRNAs. The helix shows RNase CSV cleavage on its 3′-side and no nuclease S1 activity which might have been expected in either of the two conformations. However, the lack of S1 cuts is more compatible with the B conformation since other looped-out bases do not seem to be recognized by this enzyme either (e.g., helices II and V). This may be explained by intercalation of the unpaired base, rather than looping-out, as found in synthetic duplexes [20] and in our own results of nuclease S1-analyses of eubacterial 5 S rRNAs (submitted). Examination of the accessibility of cytidines shows that C_{120} is 70% modified whilst C_{119} is not modified at all. This is further evidence for the B conformation with the looped-out uridine at position 2. Calculation of the stacking energy of the two conformations according to Ninio's latest model (which is derived from an examination of 5 S rRNA sequences, [5]) gives ΔG = −21.4 kcal for A and ΔG = 24.3 kcal for B, a slight 'advantage' for the looped-out-base version. This superiority is due particularly to the presence of two cytidines opposite the unpaired uridine, poor stacking of these pyrimidines is assumed to be responsible for increased flexibility allowing intercalation of the unpaired base [5]. One further stabilizing effect may be the 'dangling end' of B; this is not allowed for in Ninio's model, but is suggested by studies with synthetic oligonucleotides [21].

3.2. *Helix II*

There are RNase CSV cuts in helix II and one in the internal-loop between this helix and helix III, presumably due to the overcutting effect described with a double-strand specific RNase [22]. The helix is closed at both ends by G–C pairs, neither cytidine of which shows bisulphite mediated modification which would indicate helix-fraying. This phenomenon has been observed under identical conditions in eubacterial, but not in yeast 5 S rRNA (submitted). This may indicate a higher stability in eukaryotic 5 S rRNA. The looped-out uridine of this helix is not cut by nuclease S1, presumably for the same reasons noted above for uridine number two in helix I.

3.3. *Helix III*

There is one nuclease CSV cut in the shorter segment of this helix, which consists of only two base-pairs. Such a two base-pair segment may be the minimal requirement for nuclease CSV hydrolysis [22]. The looped-out cytidine number 50 is modified, as is cytidine number 30 opposite it, although it is most likely base-paired. The modification indicates probable destabilization due to the two-nucleotide loop-out as has been found for helix I in a temperature-dependent modification study of *E. coli* 5 S rRNA (submitted).

3.4. *Helix IV and helix V*

RNase CSV cuts are found on both sides of helix IV, which is unique to eukaryotic 5 S rRNA; there is no evidence of helix-fraying. Only one cut is found in helix V; the small hairpin loop which it closes is not cut by nuclease S1, nor is the looped-out uridine number 85; however, again the presence of two pyrimidines opposite the unpaired base may favour intercalation. Phosphates on the

3′-end of helix IV from position 109 upwards into helix I are protected from ethylnitrosourea alkylation.

3.5. *Loops and tertiary structure*

The majority of single-stranded regions indicated in the secondary structure shown in fig.4 are cut by nuclease S1; an interesting exception to this is the large internal-loop between helices IV and V where no single cut is to be found. This loop might therefore be expected to have a particularly complex or ordered structure. Analysis reveals that of the 13 residues involved, 9 are well-stacking purines; that such ordered structures are impervious to nuclease S1 has been noticed before [11]. Nevertheless, both cytidines are modified, indicating that their amino groups are open to destabilization as a consequence of sulphonation at the C6 position.

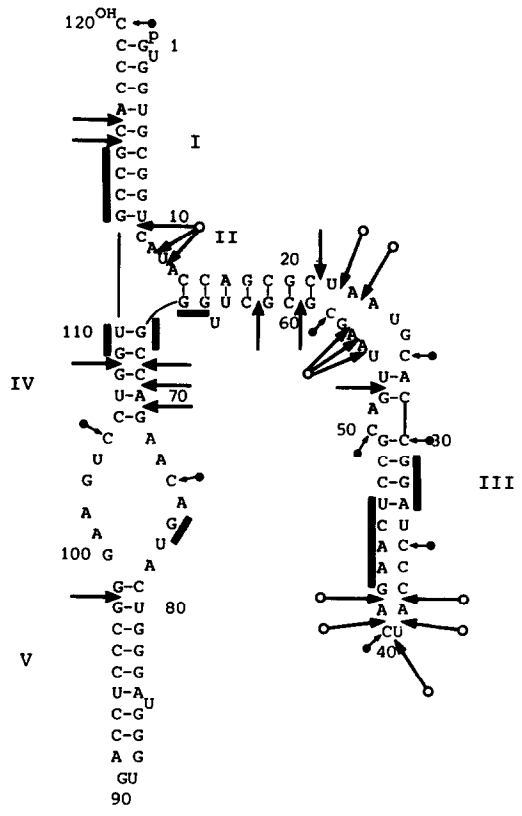

Fig.4. The secondary structure of *Equisetum arvense* 5 S rRNA with collected structural data; (→) CSV RNase; (○→) S1 nuclease; (●→) cytidine accessible to bisulphite mediated deamination; (——) phosphodiester bonds resistant to alkylation under native conditions.

The hairpin-loop closed by helix III may also have a complex structure. Of five cytidines, only two are modified by bisulphite. The results summarized in fig.4 show that the reactivity of this loop is restricted to the end distal to the loop-closing helix III. This is particularly interesting in comparison with the data on phosphodiester bond accessibility taken from the sequencing gel in fig.2 and shown in the structure in fig.4. The 3′-side of the upper portion of this loop is resistant to alkylation by ethylnitrosourea, implying that the phosphates are involved in tertiary bonding or are otherwise shielded from the solvent. Although the distribution of resistant phosphates in this eukaryotic 5 S rRNA is almost completely different to that found for rat 5 S rRNA in [19], the tertiary structure which these workers suggested may also explain our data. It is suggested that, in the tertiary structure of eukaryotic 5 S rRNA, the loop closed by helix III is brought in close proximity with the internal loop between helices V an IV. Particularly, the 3′-side of the hairpin-loop would be directed away from the solvent and towards other residues of the molecule, whilst the area we find to be reactive is directed towards the solvent. The relative inaccessibility of the internal-loop would also be explained by this conformation. The single, from alkylation strongly protected site in this loop (G_{76}) is identical to that found in the rat 5 S rRNA study [19]. Furthermore, the unreactive phosphates of residues 63–67 and 109 to (at least) 114 are located at the base of the two arms of the molecule, a region in which stabilizing non Watson–Crick interactions might be suspected.

In the context of the tertiary structure discussed here for eukaryotic 5 S rRNA, it is of interest to point out that we have previously proposed a tertiary, Watson–Crick-base-paired interaction in eubacterial 5 S rRNA involving similar regions of the molecule [11]. Whilst colinearity of helices I and IV in the eubacterial molecule leads to a conformation different to the eukaryotic model, interaction of the large hairpin-loop and the internal-loop between helices V and IV may be one characteristic common to both.

ACKNOWLEDGEMENTS

We would like to thank Dr T. Pieler for constructive discussion of the manuscript, A. Schreiber for preparing the figures and I. Brauer for typing the manuscript. The financial support of the Deutsche Forschungsgemeinschaft (SFB 9/B5) and the Fonds der Chemischen Industrie are gratefully acknowledged.

REFERENCES

[1] DeWachter, R., Chen, M.-W. and Vandenberghe, A. (1982) Biochimie 64, 311–329.
[2] Delihas, N. and Andersen, J. (1982) Nucleic Acids Res. 10, 7323–7344.
[3] Studnicka, G.M., Eiserling, F.A. and Lake, J.A. (1981) Nucleic Acids Res. 9, 1885–1904.
[4] Ninio, J. (1979) Biochimie 61, 1133–1150.
[5] Papanicolaou, C., Gouy, M. and Ninio, J. (1984) Nucleic Acids Res. 12, 31–44.
[6] Tinocco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40–41.
[7] Borer, P.N., Dengler, B., Tinocco, I. and Uhlenbeck, O.C. (1974) J. Mol. Biol. 86, 843–853.
[8] Pieler, T., Schreiber, A. and Erdmann, V.A. (1984) Nucleic Acids Res. 12, 3115–3126.
[9] Pieler, T., Kumagai, I. and Erdmann, V.A. (1982) Zbl. Bakt. Hyg. I Abt. Orig. C3, 69–78.
[10] Pieler, T., Digweed, M., Bartsch, M. and Erdmann, V.A. (1983) Nucleic Acids Res. 11, 591–604.
[11] Pieler, T. and Erdmann, V.A. (1982) Proc. Natl. Acad. Sci. USA 79, 4599–4603.
[12] Erdmann, V.A., Digweed, M. and Pieler, T. (1983) J. Biomol. Struct. Dyn. 1, 657–667.
[13] Pieler, T., Digweed, M. and Erdmann, V.A. (1983) Alfred Benzon Symp. 19, 353–376.
[14] Ulbrich, N., Digweed, M. and Erdmann, V.A. (1984) Nucleic Acids Res. 12, 1577–1580.
[15] England, T.E. and Uhlenbeck, O.C. (1978) Nature 275, 560–561.
[16] Silberklang, M., Gillum, A.M. and RajBhandary, U.L. (1979) Methods Enzymol. LIX, 58–109.
[17] Donis-Keller, J., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527–2538.
[18] Volkaert, G. and Fiers, W. (1977) Anal. Biochem. 83, 222–227.
[19] McDougall, J. and Nazar, R.N. (1983) J. Biol. Chem. 258, 5256–5259.
[20] Patel, D.J., Pardi, A. and Itakura, K. (1982) Science 216, 581–590.
[21] Neilson, T., Romaniuk, P.J., Alkema, D., Hughes, D.W., Everett, J.R. and Bell, R.A. (1980) Nucleic Acids Res. 7, 293–311.
[22] Auron, P.E., Weber, L.D. and Rich, A. (1982) Biochemistry 21, 4700–4706.