J Mol Evol (1985) 22:237–242

Journal of
Molecular Evolution
© Springer-Verlag 1985

Nuclease S_1 Analysis of Eubacterial 5S rRNA Secondary Structure

M. T. MacDonell* and R. R. Colwell

Department of Microbiology, The University of Maryland, College Park, Maryland 20742, USA

Summary. Single-strand-specific nuclease S_1 was employed as a structural probe to confirm locations of unpaired nucleotide bases in 5S rRNAs purified from prokaryotic species of rRNA superfamily I. Limited nuclease S_1 digests of 3′- and 5′-end-labeled [^{32}P]5S rRNAs were electrophoresed in parallel with reference endoribonuclease digests on thin sequencing gels. Nuclease S_1 primary hydrolysis patterns were comparable for 5S rRNAs prepared from all 11 species examined in this study. The locations of base-paired regions determined by enzymatic analysis corroborate the general features of the proposed universal five-helix model for prokaryotic 5S rRNA, although the results of this study suggest a significant difference between prokaryotic and eukaryotic 5S rRNAs in the evolution of helix IV. Furthermore, the extent of base-pairing predicted by helix IV needs to be reevaluated for eubacterial species. Clipping patterns in helices II and IV appear to be consistent with a secondary structural model that undergoes a conformational rearrangement between two (or more) structures. Primary clipping patterns in the helix II region, obtained by S_1 analysis, may provide useful information concerning the tertiary structure of the 5S rRNA molecule.

Key words: 5S rRNA — Secondary structure — Nuclease S_1 — RNA — Molecular evolution

Introduction

Elucidation of the secondary and tertiary structures of 5S ribosomal RNAs has proven surprisingly difficult, despite the accumulation of a substantial literature on the subject. More than 275 5S rRNA sequences have been reported since the initial publication of the sequence of the 5S rRNA of *Escherichia coli* by Brownlee et al. in 1967. Several factors have contributed to the slow progress in solving secondary and tertiary structures of 5S rRNAs. One is that it has been necessary to rethink concepts of permissible base pairing in RNA, and to include apparent noncanonical pairings, viz. G/U and, to a lesser extent, G/G, C/C, and A/C, for which there is now experimental evidence, in transfer and ribosomal RNAs. Ironically, the classical Watson–Crick base-pairing rules, derived for DNA, appear to represent a special subset of a more general set of base-pairing rules (Noller 1984).

Comparative analysis of the 5S rRNA sequences that have become available in recent years provides substantial information on secondary structure, notably through indicating highly conserved sequence regions and corroborating the existence of helices, based on evidence for conservation of base pairing across base-paired regions. Single- and double-strand-specific structural probes have been successfully applied in the characterization of secondary structures of ribosomal RNAs. Two types of structural probes are commonly employed: (1) chemical, such as the single-strand-specific probes kethoxal (Noller 1974) and diethyl pyrocarbonate (Peattie and Gilbert 1980) and the double-strand-specific psoralen derivatives (Rabin and Crothers 1979; Garrett-Wheeler et al. 1984); and (2) enzymatic, i.e., the single-strand-specific probe nuclease S_1 (Ross and Brimacombe 1979) and the double-strand-specific cobra venom RNase (Douthwaite and Garrett 1981).

Using cobra venom RNase, Douthwaite and Garrett (1981) confirmed the basic features of the five-helix model for prokaryotic 5S rRNAs. This analysis also raised interesting questions, however, since the

Offprint requests to: R. R. Colwell
* *Present address:* Center of Marine Biotechnology, University of Maryland, College Park, MD 20742

experimentally observed patterns of hydrolysis obtained with single-strand-specific and double-strand-specific probes are not entirely consistent with any of the currently accepted secondary structure models.

Limited digestion of populations of uniquely end-labeled rRNAs by nuclease S_1, a single-strand-specific endoribonuclease that produces 5′ phosphoryl products, has been shown to be a particularly powerful tool for determining secondary structure in rRNAs (Ross and Brimacombe 1979). Moreover, when employed in conjunction with standard sequencing methods, nuclease S_1 allows for the generation of sequence ladders that not only identify each nucleotide base, but also indicate whether each base is paired or unpaired.

We report here the results of a study in which limited nuclease S_1 digests of 3′- or 5′-end-labeled 5S rRNAs were resolved in parallel with reference riboendonuclease digests on thin sequencing gels. 5S rRNAs were prepared from a total of 11 eubacterial species of rRNA superfamily I (De Vos and De Ley 1983) for which the evolutionary histories had been determined by comparative sequence analysis (MacDonell and Colwell 1985). The method permitted identification of single-stranded regions of the 5S rRNAs as follows. Analysis of limited digests of RNAs labeled, in aliquots, on each terminus, provided a means by which primary nuclease cuts, attributable to secondary structure, could be distinguished from secondary cuts, which derive from rearrangements that are a consequence of the primary cuts (Douthwaite and Garrett 1981).

Materials and Methods

5S rRNAs were isolated from *Aeromonas media* ATCC 33907™, *Photobacterium leiognathi* ATCC 25521, *P. logei* ATCC 15382, *Plesiomonas shigelloides* ATCC 14029, *Vibrio alginolyticus* ATCC 17749, *V. carchariae* ATCC 35084, *V. diazotrophicus* ATCC 33466, *V. fischeri* ATCC 7744, *V. gazogenes* ATCC 19988, *V. natriegens* ATCC 14048, and *V. psychroerythrus* ATCC 27364 by the phenol method and purified by electrophoresis on 5% thiol-soluble acrylamide [bis-acrylylcystamine (BAC); Bio Rad, Richmond, California] gels using methods described by Hansen (1981). Separate aliquots of purified 5S rRNAs were end labeled on dephosphorylated 5′ termini using polynucleotide kinase and 0.5 mCi [γ-^{32}P]ATP and on 3′ termini with 0.5 mCi [^{32}P]cytidine-bis-phosphate and RNA ligase. End-labeled terminal bases were identified by polyethylenimine-cellulose thin-layer chromatography of exhaustive nuclease P_1 or RNase T_2 digests, following the methods of Randerath and Randerath (1967). Sequence determination was accomplished using the enzymatic method described by MacDonell and Colwell (1984a,b).

Nucleases T_1, Phy M, U_2, B.c., M_1, P_1, T_2, and S_1; polynucleotide kinase; and RNA ligase were purchased from P.L. Biochemicals (Milwaukee, Wisconsin) and [^{32}P]ATP and [^{32}P]pCp from ICN Radiochemicals (Irvine, California). Limited nuclease S_1 digests were carried out for 2 min at 25°C under the conditions described by Vogt (1980). Limited enzyme digests, as well as limited alkaline hydrolysates, of [^{32}P]5S rRNAs were resolved on thin 8% (w/v) acrylamide, 8.5 M urea sequencing gels (40 × 40 × 0.04 cm) and electrophoresed at 2000 V as described by Sanger and Coulsen (1978).

Results and Discussion

The results of this study indicate the locations of nuclease S_1 hydrolysis sites in 5S rRNAs from eubacterial species of rRNA superfamily I. Locations of S_1-labile sites in the 5S rRNA sequences examined in this study indicate that hydrolysis patterns generated from the 5′-end-labeled RNAs are similar, but not identical, to those of 3′-end-labeled RNAs. Primary S_1 cuts, i.e., those attributable to RNA secondary structure, were distinguishable from secondary S_1 cuts, because whereas the former occur at fixed positions in a sequence regardless of which terminal carries the label, the latter give the appearance of being "terminal label specific," since some clips appear only in 3′-end-labeled digests, and others only in 5′-end-labeled digests. A primary cut may allow rearrangement of the secondary structure; therefore some secondary S_1 cuts appeared in the 5′-end-labeled fragments, while others appeared in 3′-end-labeled fragments, but none appeared in both, as observed by Douthwaite and Garrett (1981). Similar S_1 primary cutting patterns were observed for all the 5S rRNAs from the 11 species examined. A composite of these in which the sites of strong primary cuts and all primary cuts are superimposed is shown in Fig. 1.

Relationships between the universal five-helix 5S rRNA secondary structure model (De Wachter et al. 1981) and nuclease S_1 cuts are shown superimposed on the *V. carchariae* sequence (as an example) in Fig. 2. Sites of secondary cuts are represented by dashed lines and those of primary cuts by solid lines.

The pattern of primary clipping obtained using nuclease S_1 in this study provides confirmation of and augments the results of earlier structural probe analyses (Douthwaite and Garrett 1981; Digweed et al. 1984; Pieler et al. 1984). The use of nuclease S_1 rather than base-specific chemical probes, as were employed by Douthwaite and Garrett (1981) and Pieler et al. (1984), allows the secondary structure to be elucidated in greater detail, since except in those cases in which the enzyme approach is sterically hindered, all bases recognized as single stranded by the structural probe are hydrolyzed. Enzymatic probes, unlike chemical probes, are subject to hindrance by tertiary structure. Consequently, enzymatic probes also provide useful information for the inference of tertiary structure based on an effect similar to that exploited in "footprinting" experiments, i.e., mapping of locations where the approach of an enzyme to a substrate is obstructed.

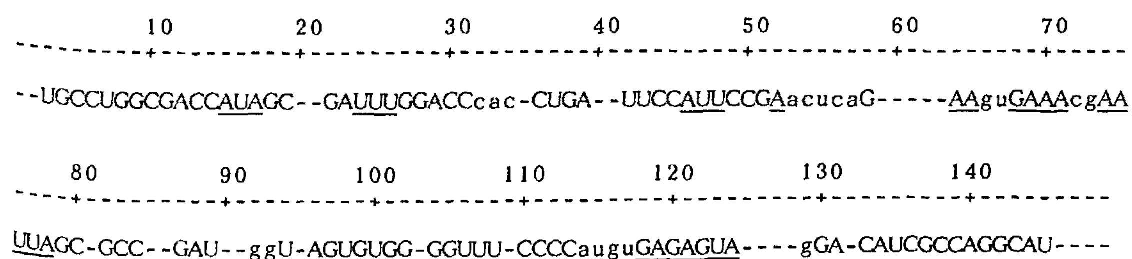

Fig. 1 Nucleotide base sequence of 5S rRNA from *V. carchariae* (MacDonell and Colwell 1985), indicating locations of nuclease S_1 hydrolysis. Underscored bases correspond to strong S_1 cuts. Bases listed in lower-case letters correspond to weak S_1 cuts

Fig. 2. 5S rRNA from *V. carchariae* folded to conform to the universal five-helix secondary structure model (De Wachter et al. 1982). Boxed-in regions indicate sites of S_1 hydrolysis: ——, primary cuts; – – –, secondary cuts. Helices are indicated with roman numerals I–V. Loops are indicated by letters A–G

Sequence	Species
-GCGNNNUR- -CGYNNNGY- A	Consensus[a]
-GCGUUUUG- -CGCAAAGC- A	V. gazogenes[b]
-GCGAUUUG- -CGUUAAGC- A	V. diazotrophicus[b] V. carchariae[b]
-GUGCUGUA- -CGUGACGU- A	V. psychroerythrus[b]
-GCGUUAUG- -CGUAAUGC- A	V. logei[b] P. leiognathi[b]
-GUGCGGUG- -CGCGCCGC- A	P. shigelloides[b]
-GCGCCGUG- -CGUGGCGC- A	A. media[b]
-GCGUUUUG- -CGUAAAGC- A	V. natriegens[b] V. alginolyticus[b]

Fig. 3. The helix II region of prokaryotic 5S rRNA. Nucleotide base sequences are shown for helix II of ten eubacterial species. The high frequency of adjacent weak base pairs, i.e., A/U and G/U, may help to explain the tendency for the distal portion of helix II (right half, as shown here) to be hydolyzed by nuclease S_1. References: [a] Erdmann et al. (1984); [b] MacDonell and Colwell (1985)

Tertiary structure interaction appears to be involved in inhibition of hydrolysis of two apparently single-stranded regions flanking the ascending side of helix III. These regions correspond to positions G_{28}–C_{31} and U_{40}–C_{45} in Fig. 2. Attack by chemical probes was observed by Douthwaite and Garrett (1981) in both of these regions. The exact nature of the steric hindrance in the enzymatic approach of these sites is unknown, but conceivably could be determined using space-filling molecular models. The nuclease S_1 attack on the ascending and descending sides of helix II, corresponding to bases U_{24}–U_{26} and A_{74}–U_{77} (Fig. 2) contradicts the results of Douthwaite and Garrett (1981), in which the only attack observed in this region was achieved using a double-strand-specific structural probe, cobra ven-

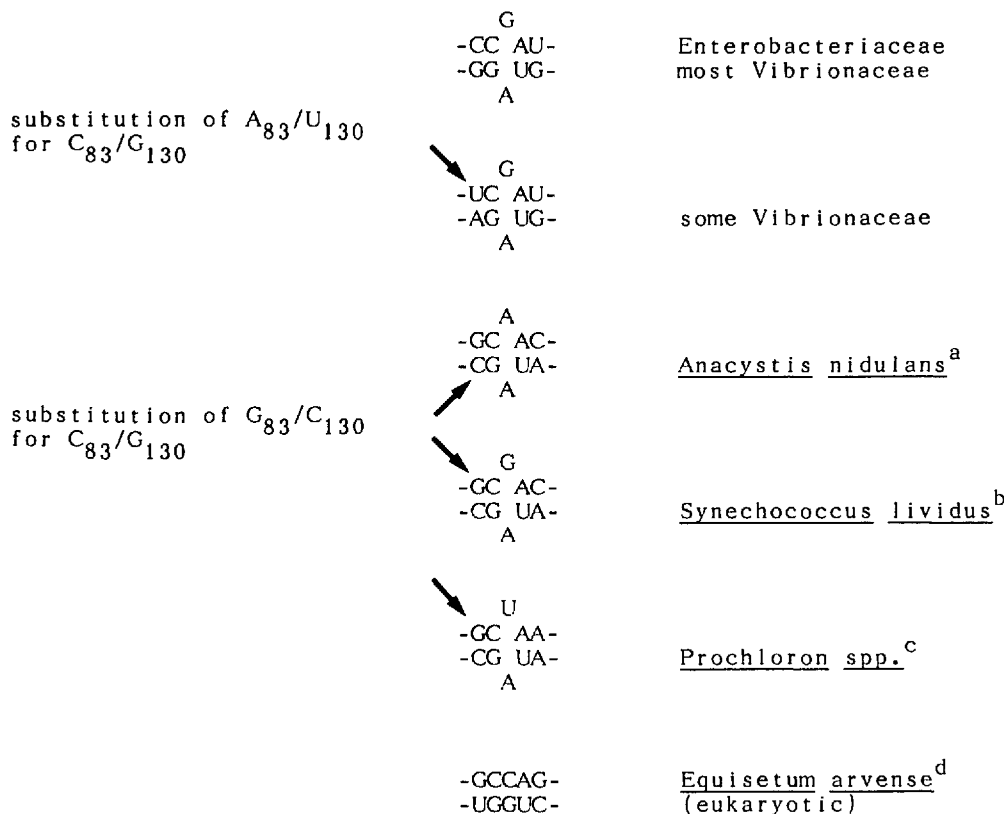

Fig. 4. Helix IV region of the universal five-helix model is shown for prokaryotic species of rRNA superfamily I (Vibrionaceae and Enterobacteriaceae) and rRNA superfamily III (*A. nidulans, S. lividus,* and *Prochloron* spp.). The helix IV region of a eukaryotic 5S rRNA (*E. arvense*) is shown for comparison. References: [a]Corry et al. (1984); [b] Delihas et al. (1982); [c] MacKay et al. (1982); [d] Digweed et al. (1984)

om RNase. Nevertheless, since primary cutting by nuclease S_1 was symmetric with respect to helix II, it seems unlikely that the observation is an artifact. Helix II is a relatively weak structure in many eubacterial species, due to the high frequency of G/U and A/U base pairs (Fig. 3).

The overall pattern of nuclease S_1 hydrolysis observed in this study generally supports the five-helix model of De Wachter et al. (1982). However, modification of the definition of helices II and IV is required. As depicted by the universal five-helix model, the eubacterial helix IV consists of five base pairs arranged as two two-base-pair helices flanking a conserved G/A mismatch (Fig. 4). There is strong evidence of evolutionary conservation of at least a portion of this helix, since in approximately one-third of the 5S rRNAs from species of the Vibrionaceae (MacDonell and Colwell 1985) the base pair C_{83}/G_{130} is replaced with U_{83}/A_{130}, whereas in *Anacystis nidulans* (Corry et al. 1974), *Synechococcus lividus* (Delihas et al. 1982), and *Prochloron* spp. (MacKay et al. 1982) G_{83}/C_{130} is substituted (Fig. 4). We did not detect any primary S_1 clipping in the N_{83}/N_{130}–C_{84}/G_{129} region of helix IV, although the putative base pairs A_{88}/U_{133}–U_{89}/G_{132} were consistently clipped.

This observation is inconsistent with the proposed universal five-helix model (De Wachter et al. 1982) and suggests that unlike in the eukaryotic case, only a portion of the universal helix IV has been evolutionarily conserved in prokaryotes. There has been a tendency, on the part of some investigators, to consider G/A a legitimate, although noncanonical, base pair, since it occurs in nearly all eubacterial 5S rRNAs at the center of helix IV. Indeed, it can be argued that the high "penalty" assessed for G/A base pairing according to Ninio's rules (Ninio 1979) is a consequence of the rarity of G/A in the primary structure of tRNAs, from which these empirical rules were derived. The results presented herein, however, suggest that (i) G_{87}/A_{124} are unpaired in solution and (ii) helix IV is abbreviated in eubacterial species (Fig. 5).

The calculated negative free energy (Ninio 1979) indicates that a two-base-pair helix IV would be stable, even if only weakly ($\Delta G < 2$ kcal/mol). Helix IV may be stabilized, however, by tertiary interactions. The "chair" model for 5S rRNA proposed by

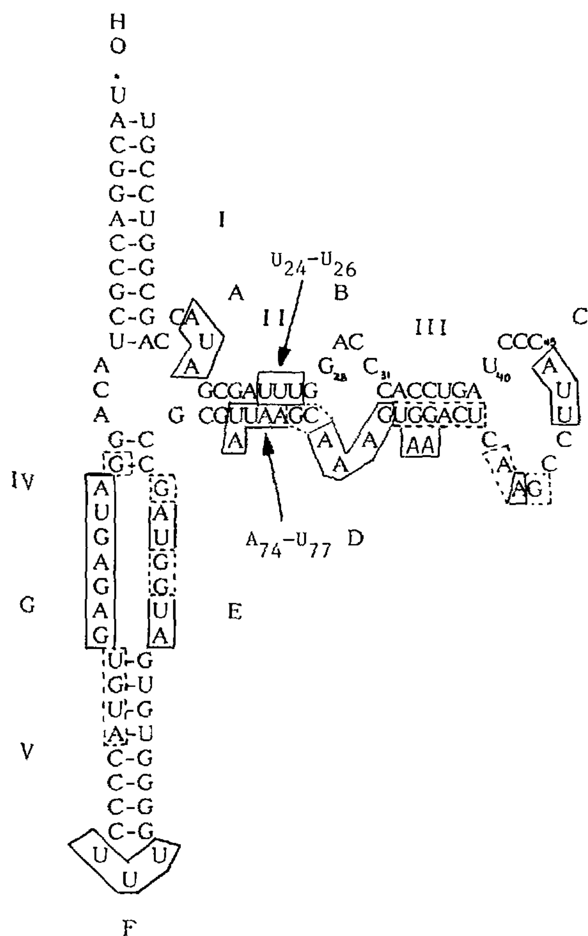

Fig. 5. The 5S rRNA from *V. carchariae* folded to conform to a variation of the five-helix model. Note the minimal base pairing in helix IV and the extension of loops E and G compared with the universal five-helix model (Fig. 2). S_1 hydrolysis sites, loops, and helices are indicated as in Fig. 2.

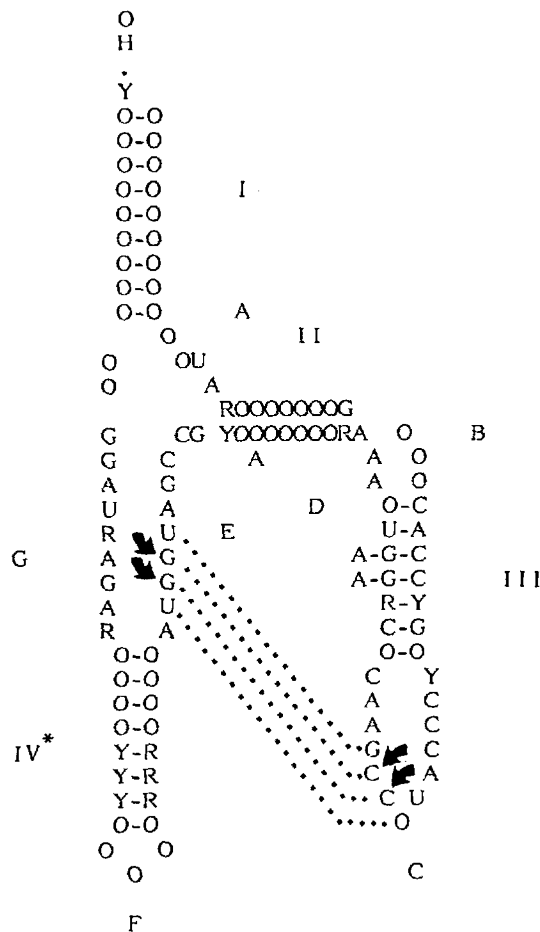

Fig. 6. Consensus sequence of prokaryotic 5S rRNA folded according to the chair model of Pieler and Erdmann (1982). Note that helix IV (*) in the chair model is equivalent to helix V in the universal five-helix model (Fig. 2). Arrows mark the locations of nucleotide bases speculated to be involved in tertiary interactions that are protected from hydrolysis by nuclease S_1. Designations of loops and helices are as in Fig. 2. A, adenine; C, cytosine; G, guanine; O, unspecified base (variable); R, purine; U, uracil; Y, pyrimidine

Pieler and Erdmann (1982) suggests a tertiary interaction between loop cLc' and loop dLe (Fig. 6). It is interesting to note that positions corresponding to the proposed site of tertiary interaction (indicated by arrows in Fig. 6) were not clipped by S_1.

An alternative explanation of the S_1 clipping pattern observed in the region of helix IV, specifically at bases A_{88}–U_{89} and G_{122}–U_{123} (Fig. 2), is that the secondary structure of prokaryotic 5S rRNA switches between two (or more) conformations. The notion that 5S rRNA may function as a modulator or "switch" in the ribosome is not new. In fact, direct experimental evidence of switchlike rearrangement in 5S rRNAs was reported by Kao and Crothers (1980), although there is, as yet, no information on the exact nature or specific location(s) of the molecular rearrangements. A computer-generated model fitting the general description of a conformational switch is the Y-form/P-form model described by Trifonov and Bolshoi (1983). From the calculated lowest global free energy of the secondary structure (Ninio 1979), it is concluded that the Y-form, modified slightly to accommodate formation of helix IV, and the P-form may both occur, since the secondary structures of these demonstrate approximately the same global free energy. Enzymatic probes, e.g., nuclease S_1, interacting with A_{88}–U_{89} and G_{122}–U_{123} region of helix IV are influenced by steric hindrance if the molecule is in the P-form. However, enzymatic attack would be possible during rearrangement.

The secondary and tertiary structures of 5S rRNAs may vary among and between prokaryotes and eukaryotes. The evolutionary significance of these structures is just beginning to be appreciated.

Acknowledgments. The authors wish to acknowledge V.A. Erdmann, G. Fox, T. Pieler, and E.N. Trifonov for helpful discussions. Support for this work was provided by National Science

Foundation Grant BSR-82-08418 and by Office of Naval Research Grant N000-14-81-K0638.

References

Brownlee GG, Sanger F, Barrell BG (1967) Nucleotide sequence of 5S ribosomal RNA from *Escherichia coli*. Nature 215:735–736

Corry MJ, Payne PI, Dyer TA (1974) The nucleotide sequence of 5S rRNA from the blue-green alga *Anacystis nidulans*. FEBS Lett 46:63–66

Delihas N, Andresini W, Andersen J, Berns D (1982) Structural features unique to the 5S ribosomal RNAs of the thermophilic cyanobacterium *Synechococcus lividus* III and the green plant chloroplasts. J Mol Biol 162:721–727

De Vos P, De Ley J (1983) Intra- and intergeneric similarities of *Pseudomonas* and *Xanthomonas* ribosomal ribonucleic acid cistrons. Int J Syst Bacteriol 33:487–509

De Wachter R, Chen M-W, Vandenberghe A (1982) Conservation of secondary structure in 5S ribosomal RNA: a uniform model for eukaryotic, eubacterial, archaebacterial and organelle sequences is energetically favourable. Biochimie 64: 311–329

Digweed M, Ulbrich N, Erdmann VA (1984) The secondary structure of the 5S rRNA from the horsetail *Equisetum arvense*. FEBS Lett 176:255–260

Douthwaite S, Garrett RA (1981) Secondary structure of prokaryotic 5S ribosomal ribonucleic acids: a study with ribonucleases. Biochemistry 20:7301–7307

Erdmann VA, Wolters J, Huysmans E, Vandenberghe A, De Wachter R (1984) Collection of published 5S and 5.8S ribosomal RNA sequences. Nucleic Acids Res 12:r133–r166

Garrett-Wheeler E, Lockart RE, Kumar A (1984) Mapping of psoralen cross-linked nucleotides in RNA. Nucleic Acids Res 7:3405–3423

Hansen JN (1981) Use of solubilizable acrylamide disulfide gels for isolation of DNA fragments suitable for sequence analysis. Anal Biochem 116:146–151

Kao TH, Crothers DM (1980) A proton-coupled conformational switch of *Escherichia coli* 5S ribosomal RNA. Proc Natl Acad Sci USA 77:3360–3364

MacDonell MT, Colwell RR (1984a) The nucleotide base sequence of Vibrionaceae 5S rRNA. FEBS Lett 175:183–188

MacDonell MT, Colwell RR (1984b) Nucleotide sequence of 5S ribosomal RNA from *Vibrio marinus*. Microbiol Sci 1: 229–231

MacDonell MT, Colwell, RR (1985) Phylogeny of the Vibrionaceae and recommendation for two new genera *Listonella* and *Shewanella*. Syst Appl Microbiol 6:171–182

MacKay RM, Salgado D, Bonen L, Stackebrandt E, Doolittle WF (1982) The 5S ribosomal RNAs of *Paracoccus denitrificans* and *Prochloron*. Nucleic Acids Res 10:2963–2970

Ninio, J (1979) Prediction of pairing schemes in RNA molecules—loop contributions and energy of wobble and non-wobble pairs. Biochimie 61:1133–1150

Noller HF (1974) Topography of 16S RNA in 30S ribosomal subunits. Nucleotide sequences and location of sites of reaction with kethoxal. Biochemistry 13:4694–4703

Noller HF (1984) Structure of ribosomal RNA. Annu Rev Biochem 53:119–162

Peattie DA, Gilbert W (1980) Chemical probes for higher-order structure in RNA. Proc. Natl Acad Sci USA 77:4679–4682

Pieler T, Erdmann VA (1982) Three dimensional structural model of eubacterial 5S RNA that has functional implications. Proc Natl Acad Sci USA 79:4599–4603

Pieler T, Schreiber A, Erdmann VA (1984) Comparative structural analysis of eubacterial 5S rRNA by oxidation of adenines in the N-1 position. Nucleic Acids Res 12:3115–3126

Rabin D, Crothers DM (1979) Analysis of RNA secondary structure by photochemical reversal of psoralen crosslinks. Nucleic Acids Res 7:689–703

Randerath K, Randerath E (1967) Thin-layer separation methods for nucleic acid derivatives. Methods Enzymol 12A:323–347

Ross A, Brimacombe R (1979) Experimental determination of interacting sequences in ribosomal RNA. Nature 281:271–276

Sanger F, Coulsen AR (1978) The use of thin acrylamide gels for sequencing. FEBS Lett 87:107–110

Trifonov EN, Bolshoi G (1983) Open and closed 5S ribosomal RNA, the only two universal structures encoded in the nucleotide sequences. J Mol Biol 169:1–13

Vogt VM (1980) Purification and properties of S1 nuclease from *Aspergillus*. Methods Enzymol 65:248–264

Received February 21, 1985/Revised May 30, 1985