Volume 10 Number 2 1982 Nucleic Acids Research

Secondary structure of Bombyx mori and Dictyostelium discoideum 5S rRNA from S1 nuclease and cobra venom ribonuclease susceptibility, and computer assisted analysis

Anthony Troutt, Thomas J.Savin, William C.Curtiss, James Celentano and John N.Vournakis

Department of Biology, Syracuse University, 130 College Place, Syracuse, NY 13210, USA

Received 11 September 1981; Revised and Accepted 27 November 1981

ABSTRACT

The 5S rRNAs from Bombyx mori and Dictyostelium discoideum were end-labeled with [^{32}-P] at either the 5' or 3' end and sequenced using enzymatic digestion. The secondary structure of these molecules was studied using the single-strand specific S1 nuclease and the base-pair specific cobra venom ribonuclease. Computer analysis of these results was performed and was used to generate a consensus secondary structure for each molecule. A comparison of these results with those of other workers is presented.

INTRODUCTION

The secondary and tertiary structure of RNA is vital in determining its biological function. There is extensive experimental documentation of this for such molecules as tRNA (1,2), and the rRNAs (3). Predictions of secondary structure for these molecules have generally employed one of the following criteria: computer modeling based upon thermodynamic data (4), structure-specific enzymatic cleavage (5), and phylogenetic comparisons (6). However, current thermodynamic data is insufficient to yield a correct structure in most cases (7), enzymatic structure - mapping alone often yields ambiguous results (8), and phylogenetic comparisons cannot aid in the elucidation of any secondary or tertiary interactions which are particular to a given organism's RNA.

In the work presented here, we have utilized a combination of all three of these methods to predict the secondary structure of the 5S rRNA molecules from the silk moth Bombyx mori and the slime mold Dictyostelium discoideum. Consensus secondary structures for these molecules are presented, and are compared to 5S rRNA structural models obtained by others (9-21).

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.
0305-1048/82/1002-0653 $2.00/0

MATERIALS AND METHODS

Enzymes: Calf intestine alkaline phosphatase (Boehringer-Mannheim) was purchased and further purified by Sephadex-G75 chromatography prior to use in order to eliminate contaminating ribonucleases (22). T4 polynucleotide kinase (Bethesda Research Laboratories), T1 (Sankyo), P1 (P-L Biochemicals), T2 (Sigma Chemical Co.), U2 (Research-Plus) and CL3 (Bethesda Research Laboratories) ribonucleases were used without additional purification. Physarum nuclease, was a gift from H. Donis-Keller. T4 RNA ligase was kindly provided by N. Pace. Bacillus. cereus pyrimidine specific ribonuclease (23), S1 nuclease and cobra venom ribonuclease were purified as described (8).

Purification of 5S rRNA: Bombyx mori wild type fifth instar larvae were a gift of Marian R. Goldsmith. Total RNA was prepared from these insects by phenol/meta-cresol/8-hydroxyquinolone extraction as follows: approximately 5 grams of wild type larvae were ground to a fine powder in liquid nitrogen. The powder was suspended in 200 ml of homogenization buffer (25 mM Tris-HCl pH 7.4, 50 mM $MgCl_2$, 250 mM NH_4Cl, 250 mM NaCl, 0.25 M sucrose (nuclease-free, Schwartz/Mann), 0.5 mM EGTA, 0.5% v/v ß-mercaptoethanol, 50 microgram/ml spermine, 100 microgram/ml sodium heparin) and homogenized with a Dounce homogenizer. This suspension was then centrifuged at 3500 rpm, for 60 min., at 4°C, in a Beckman J-21B centrifuge using a JA-14 rotor. The supernatant was saved and centrifuged at 5000 rpm, for 30 min., at 4°C. This supernatant was extracted 5 times with equal volumes of a phenol mix (1 liter phenol, 140 ml meta-cresol, 1.0 gm 8-hydroxyquinolone, H_2O saturated). The aqueous phase was then extracted 3 times with equal volumes of H_2O saturated diethyl ether. Residual ether was removed by evaporation under vacuum, the solution was made 0.3 M in NaCl, and 3 volumes of 95% ethanol were added. The resultant RNA precipitate was stored at -20°C. RNA was recovered by centrifugation, and was redissolved in 9 M urea, 10 mM EDTA, 0.5% xylene cyanol (XC), 0.5% bromophenol blue (BPB). 5S rRNA was isolated from this mixture by electrophoresis of the total RNA on a 6 mm thick 6% polyacrylamide-8.3 M urea preparative gel. Bands were visualized by UV shadowing, excised, and eluted in 300 mM NH_4OAc pH 5.5, 10 mM EDTA, at 37°C for 8-16 hrs.

Dictyostelium discoideum total RNA was obtained by Alan Kimmel. D. discoideum 5S rRNA was also prepared by electrophoresis of the total RNA on a 6mm thick 8% polyacrylamide-8.3 M urea preparative gel, followed by UV shadowing, excision and elution of the proper band.

End-labeling: Both 5' and 3' end labeling of the 5S rRNAs was performed as described by Vournakis et al. (8), and Pavlakis et al. (24).

End analysis: 5'[^{32}P]-labeled and 3'[^{32}P]-labeled RNAs were digested exhaustively with P1 and T2 nucleases, respectively. The end products of these digests were identified using PEI-cellulose thin layer chromatography as reported (5).

Sequencing and structure analysis reactions: Sequencing reactions and S1 structure mapping reactions were performed according to procedures established by Vournakis et al. (8). In addition, standard S1 reactions were performed at 75°C in order to generate cleavages at all nucleotide positions, thus assisting in the proper identification of bands in the structure-mapping reactions. Cobra venom ribonuclease (CVR) digests were performed essentially as described, using a 1:200 enzyme dilution (8). The reactions were loaded onto thin (40 X 33 X 0.40 cm. or 90 X 30.5 X 0.028 cm.) 10% to 20% polyacrylamide-8.3 M urea gels and electrophoresed as described (8,25).

Computer predictions of secondary structure: Computer models of secondary structure were obtained using a version of the program developed by Zuker and Stiegler (26), which was a gift from Richard J. Feldman. S1 accessibility data was used to generate models as discussed below (25). In addition, a second approach was used to obtain predictions of 5S secondary structure. The PROPHET (27) interactive graphics system, which is a national resource for biomedical research, has the capability of generatting RNA secondary structures that are unbiased by thermodynamic considerations. Both S1 and CVR data were input via a light pen/data tablet system to construct structures that are in agreement with experimental evidence. Phylogenetic considerations were used as well. Details of this approach are described in a forthcoming publication (28).

RESULTS AND DISCUSSION

End-Analysis

Bombyx mori 5S rRNA is equally homogenous at both ends as determined by end-analysis. The 5' end consists 100% of G and the 3' end is 100% U. Dictyostelium discoideum 5S rRNA labeled at the 5' terminus was found to yield no minor species upon end analysis, and consists 100% of G.

Sequence of the 5S rRNAs

The Bombyx mori 5S molecule was completely sequenced from both ends using 5' and 3' end-labeled RNA and the rapid RNA sequencing method (29,30)

with thin polyacrylamide gels (see Materials and Methods). An example of a set of sequencing reactions is illustrated on the left side of Figure 1. The major species is 119 nucleotides long and identical to the first 119 nucleotides of the B. mori 5S molecule sequence determined by Komiya et al. using Peattie's chemical sequencing method (31). Light bands one nucleotide longer than the major species are seen on sequencing gels of 3' end-labeled RNA, suggesting the presence of a small percentage of molecules with one additional uracil residue at the 3' end. This minor species is the same as the 120 nucleotide molecule sequenced by Komiya et al. (31). Several other 5S RNAs appear to have terminal uracil moieties present in less than one mole per mole of 5S RNA (32-37).

Whereas Komiya et al. (31) reported that the chemical sequencing technique they employed was unable to detect G87, our enzymatic technique yielded unambiguous results in this region when reactions were performed under proper conditions and gels were run maintaining total denaturation of the RNA as discussed elsewhere (8). Other workers (13) have reported difficulty in making clear distinctions between cytosine and uracil in enzymatic digests. However, with judicious use of such enzymes as Physarum nuclease, Bacillus cereus nuclease, and the cytosine specific chicken liver ribonuclease (CL3) (38,39), our laboratory has avoided such problems.

Dictyostelium 5S rRNA was completely sequenced using 5' end-labeled RNA as discussed above. The sequence thus obtained is identical to that determined by Hori et al. using chemical degradation (11).

Structure-Analysis of 5S rRNA with Enzyme Probes

S1 nuclease isolated from Aspergillus oryzae has been shown to be specific for single-stranded RNA (40) and has been used to locate non-base-paired regions in numerous RNAs of biological interest (5,8,24,25, 41-43), including 5S rRNA (8,10,20,21). Cobra venom ribonuclease is known to cleave RNA only in regions involved in secondary or tertiary interactions in several RNA molecules (8,44,45). Thus, through concomitant use of both of these enzymes one can obtain exhaustive direct information concerning the secondary structure of a given RNA. Examples of digests of B. mori 5S rRNA with these enzymes are shown on the right side of Figure 1. Care must be taken in the assignment of bands from such digests because both S1 and CVR generate 5'-oligonucleotides with 3' hydroxyl groups whereas alkaline digests as well as all of the sequencing enzymes generate oligonucleotides with 3'-phosphates. This causes S1 and CVR bands to migrate at a slower rate than the corresponding sequencing bands. In order to

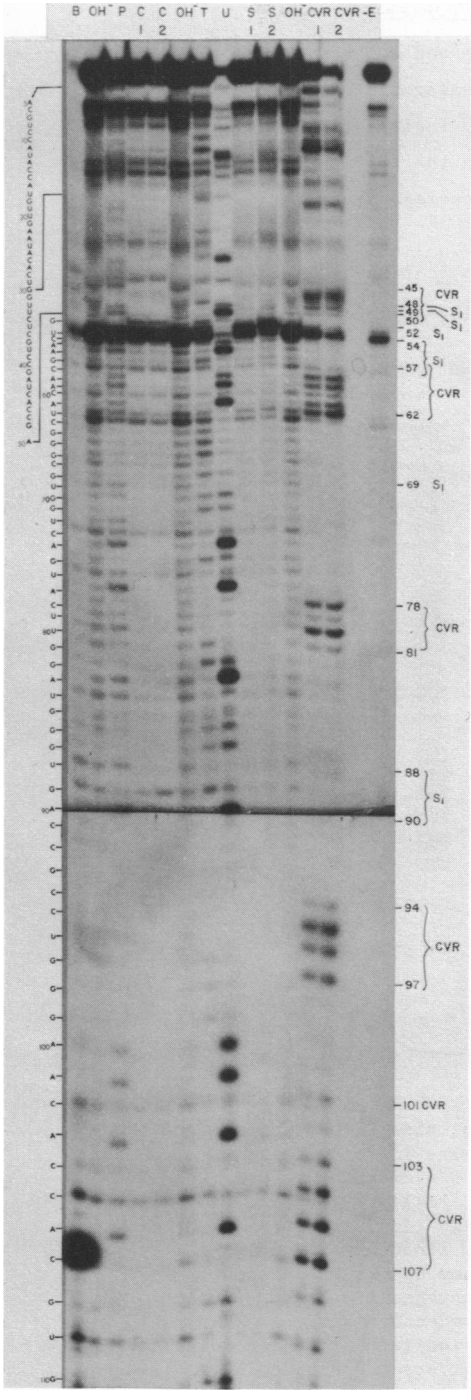

Figure 1. Electrophoretic patterns of partial enzymatic digestions for sequencing and structural analysis of 3' end-labeled Bombyx mori 5S rRNA on a 10% polyacrylamide-8.3M urea gel. Reactions were loaded onto the gel as follows: BC, Bacillus cereus nuclease, 15 min; OH^-, alkaline digestion, 12 min at 90 C; P, Physarum nuclease, 15 min; C(1), CL3 ribonuclease, 0.8 U/microgram RNA, 30 min; C(2), CL3 ribonuclease, 0.8 U/microgram RNA, 90 min; T, T1 nuclease, 0.08 U/microgram RNA, 15 min; U, U2 ribonuclease, 0.018 U/microgram RNA, 15 min; S(1), S1 nuclease, 0.006 U/microgram RNA, 10 min; S(2), S1 nuclease, 0.006 U/microgram, 30 min; CV(1), cobra venom ribonuclease, 10 min; CV(2), cobra venom ribonuclease, 25 min; -E, minus enzyme control. 9000 Cerenkov cpm of the 3'-labeled 5S was loaded per slot. The positions of cleavage by S1 and cobra venom ribonuclease are indicated along the right side and the sequence is shown along the left side of the figure.

aid in the assignment of structure-mapping bands, S1 reactions were performed at high temperature (see Materials and Methods). Reactions under these conditions are known to produce cleavages at all nucleotide positions (R. Riley, personal communication), facilitating localization of enzymatic cuts. Positions of cleavage in the 5S rRNAs from B. mori and D. Discioideum by these structure-specific nucleases are indicated in Figure 1 and are summarized in Figure 3.

The usefulness of the coordinate application of CVR and S1 in determination of RNA secondary structure is illustrated in Figure 2. Here a

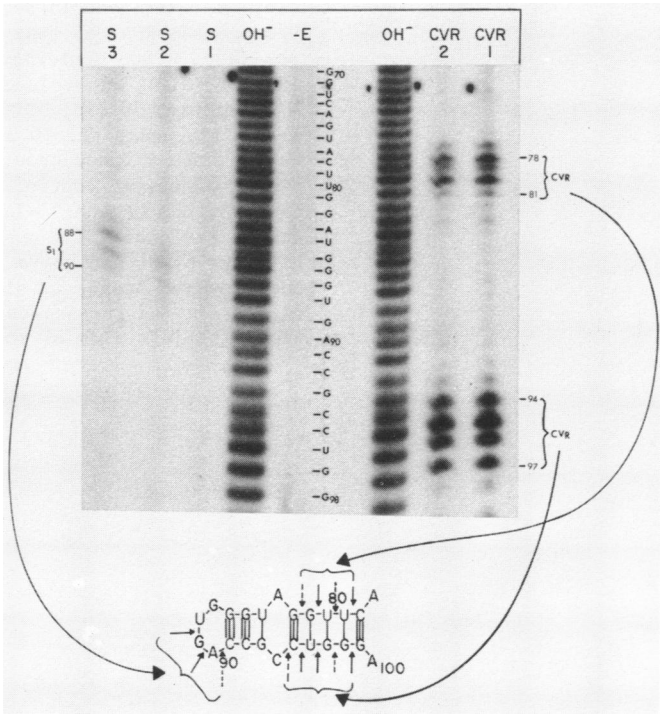

Figure 2. Localization of a hairpin loop using S1 and Cobra venom ribonuclease accessibility data. Reactions were loaded onto a 10%-8.3M urea gel as follows: S(3), S1 nuclease, 0.05 U/microgram RNA, 15 min; S(2), S1 nuclease, 0.05 M/microgram RNA, 5 min; S(1), S1 nuclease, 0.05 U/microgram RNA, 1 min; OH^-, alkaline digestion, 12 min at 90 C; -E, minus enzyme control; CV(2), Cobra venom ribonuclease, 15 min; CV(1), Cobra venom nuclease, 5 min. 13000 Cerenkov cpm of 3'-labeled 5S was loaded per slot. The positions of S1 cleavages are indicated along the left side, cobra venom ribonuclease cleavages are shown at the right, and the nucleotide sequence of the region is indicated along the alkaline digest lane. The proposed secondary structure of this region is illustrated at the bottom of the figure.

hairpin loop is localized by the presence of a series of S1 cleavages flanked both 5' and 3' by CVR cleavages. When data such as this is combined with sequence comparisons between various 5S rRNA molecules (data not shown) the structures illustrated in Figure 3 are generated.

A computer model of the secondary structure of the B. mori and D. discoideum 5S rRNAs, generated using thermodynamic considerations alone, are shown in Figure 4. It is seen, by comparing Figure 4 with Figure 3 that the basic folding rules and thermodynamic information available at present are inadequate to predict a correct secondary structure with assurance. This is indicated by the disagreement of the computer-generated structure with data from structure-mapping digests as well as by the fact that lowest-free-energy models calculated for several 5S rRNA (data not shown) bear little or no resemblance to one another (compare the B. mori and the D. discoidium structures in Figure 4). Clearly, structure specific enzyme accessibility data and phylogenetic comparisons must be taken into consideration in any determination of an RNA molecule's secondary structure.

The models in Figure 3 are in excellent agreement with our enzymatic data on the B. mori 5S except for the S1 cleavage between U69 and G70 and following G116 and G117, and the CVR cuts located between C39 and C40, C40 and G41, and A54 and A55. It is possible that the two CVR cleavages at C39 and C40 may be a result of tertiary interactions between the two hairpin loops at positions 89-91 and 32-45. The S1 cleavages at the 3'-end of the molecule (stem I) may be an end-effect due to breathing of the hairpin helix. S1 and CVR digests were performed on D. discoideum 5S in essentially the same manner as those on B. mori 5S rRNA. Positions of these structure-mapping cleavages are also indicated in Figure 3. A computer model of the secondary structure of this molecule, constructed utilizing thermodynamic constraints, alone, is illustrated in Figure 4. It shows even more clearly than the model for B. mori 5S rRNA that thermodynamic computer modeling alone is insufficient for the prediction of secondary structure.

It is evident that the proposed secondary structures for B. mori and D. discoideum 5S rRNAs are nearly identical to one another. Additionally, it can be observed that the 5S rRNAs from such diverse sources as Human KB cells, Xenopus laevis, Drosophila melanogaster, Tetrahymena thermophilia, Torulopsis utilus and Triticum vulgare can form a secondary structure which correlates very well with our proposed models. Additional evidence

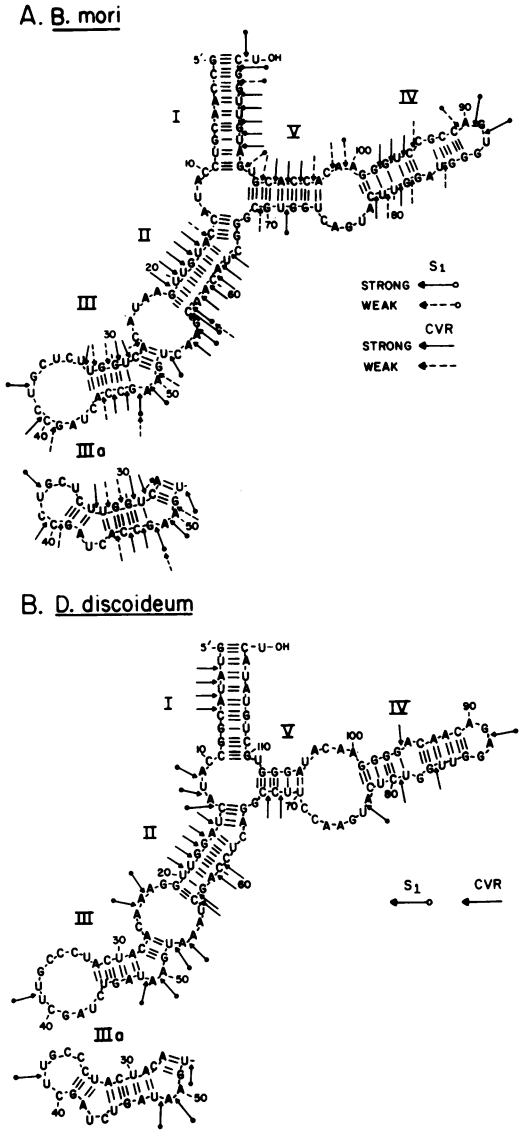

Figure 3. Secondary structure models of B. mori and D. discoideum 5S rRNA, using S1 nuclease and cobra venom ribonuclease accessibility data. Helical regions are labeled I-V using the nomenclature of Leuhrsen and Fox (13). An alternative for helical region III, labeled IIIa, is indicated for each molecule. The models were obtained using the PROPHET interactive computer graphics system developed by Bolt, Beranek and Newman, Inc., Cambridge, Mass (27). The nuclease susceptibility data (both S1 and CVR) shown were used as input information and the models were obtained without imposing thermodynamic criteria (28).

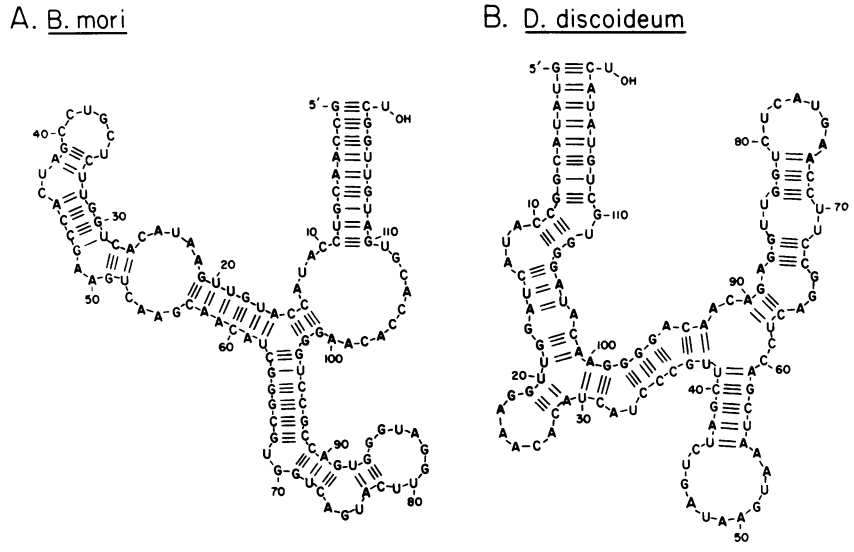

Figure 4. Optimal lowest free-energy structures of B. mori and D. discoideum 5S rRNAs generated using a version of the secondary structure prediction program developed by Luker and Stiegler (26), using only RNA sequence information as input data.

for the generality of our model comes from its agreement with structure-specific probing of other 5S molecules including: hydroxymethyltrimethylpsoralen crosslinking in Drosophila (17), RNAses T1, T2 and A, modification of exposed guanine residues in T. utilis (18), RNAses T1 and A in Chlorella (19), and S1 digestion in T. utilus (10), wheat embryo (20),and Yeast (21). Infrared and Raman spectroscopy on rat 5S RNA (12) indicates that it folds into a secondary structure similar to those proposed here.

Several attempts at describing a general 5S rRNA secondary structure model have been published in recent years (9-16). Luehrsen and Fox (13) have developed a useful nomenclature, designating base-paired and non-paired regions by numbers and letters, respectively, in a proposed general model. We describe our results on B. mori and D. discoideum 5S rRNAs, shown in Figure 3, using the Luehrsen and Fox nomenclature, in Table one. The most interesting differences between the Luehrsen and Fox (13) model and our results include the following: we do not have a loop A-2 only A-1 which we name A in Table 1; the bulge bases at positions 63 and 49-50 make stems II and III longer in our model; bulge bases 43-44 can generate an even longer stem III, that we call IIIa; and bulge bases at 83 and 94 in

Table 1. Architectural features of B. mori and D. discoideum 5S rRNA using the nomenclature of Luehrsen and Fox (13).

Helical Region	No. of Base-pairs	Non-paired Region*	No. of Bases
A. B. mori			
I	9	A	4
II	8	B1	5
III(IIIa)	6(8)	B2	4
IV	8	C(Ca)	12(6)
V	6	D	4
		E1	5
		E2	3
B. D. discoideum			
I	9	A	4
II	8	B1	5
III(IIIa)	6(8)	B2	4
IV	8	C(Ca)	12(6)
V	4	D	4
		E1	7
		E2	5

*1 or 2 bulge bases can exist at positions 43-44, 49-50, 63, 83 and 94 in B. mori, and at positions 43-44, 49-50, 63, 82 and 95 in D. discoideum. Single, non-paired, hinge bases can exist at positions 69 and 109 in both molecules (see Fig. 3).

B. mori and at 82 and 95 in D. discoideum make stem IV longer in our models, being identical to the "extended" helix IV of Barnett et al. (14) and MacKay and Doolittle (16). Our models agree very well with that proposed in 1974 by Nisikawa and Takemura (9), and are consistent with the considerations of Studnicka et al., (15), among others (14, 16). We have displayed the 5S molecules with a new orientation, locating the 5'-terminus to the top left, in an attempt to be consistent with the usual display mode of tRNAs. Studies are underway to further elaborate the general features and important species differences of 5S rRNA molecules.

ACKNOWLEDGEMENTS

Many thanks to Calvin Vary, Michael J. Lane and George Pavlakis for enlightening discussions. The work was supported by N.I.H. grant GM22280. Special thanks are given to Wayne Rindone and Harold Perry of Bolt,Beranek and Newman, Inc. of Cambridge, MA, for assistance with the use of the PROPHET computer system for secondary structure analysis.

1. Gartland, W. J. and Sueoka, N. (1966) Proc. Natl. Acad. Sci. USA 55, 948-956.
2. Fresco, J., Adams, A., Ascione, R., Henley, D. and Lindahl, T. (1966) Cold Spring Harbor Symp. Quant. Biol. 36, 527-537.
3. Pace, N. R. (1973) Bacteriol. Rev. 37, 562-603.
4. Pipas, J. M. and McMahon, J. E. (1975) Proc. Natl. Acad. Sci. USA 72, 2017-2021.
5. Wurst, R., Vournakis, J. N., and Maxam, A. (1978) Biochem. 17, 4439-4499.
6. Hori, H. (1976) Molec. Gen. Genet. 145, 119-123.
7. Tinoco, D., Uhlenbeck, O. C., and Levine, M. D. (1971) Nature 230, 362-367.
8. Vournakis, J. N., Celentano, J., Finn, M., Lockard, R., Mitra, T., Pavlakis, G., Troutt, A., van den Berg, M. and Wurst, R., (1981), in "Gene Amplification and Analysis" (J. G. Chirikjian and T. S. Papas, editors), Vol. 2, Analysis of Nucleic Acids by Enzymatic Methods, pp 267-298.
9. Nishikawa, K. and Takemura, S. (1974) J. Biochem. 76, 935-947.
10. Nishikawa, K. and Takemura, S. (1977) J. Biochem. 81, 995-1003.
11. Hori, H., Osawa, S. and Iwabuchi, M. (1980) Nucl. Acids Res. 8, 5535-5539.
12. Bohm, S., Fabian, H., Welfle, H. and Bielka, H. (1980) Acta Biol. Med. Germ. 39, K1-K6.
13. Luehrsen, K. R. and Fox, G. E. (1981) Proc. Nat. Acad. Sci. USA 78, 2150-2154.
14. Garnett, R. A., Douthwaite, S. and Noller, H. F. (1981) TIBS 6, 137-139.
15. Studnicka, G. M., Eiserling, F. A. and Lake, J. A. (1981) Nucl. Acids Res. 9, 1885-1904.
16. MacKay, R. M. and Doolittle, W. F. (1981) Nucl. Acids Res. 9, 3321-3334.
17. Thompson, J., Wegney, M. R. and Hearst, J. E. (1981) J. Mol. Biol. 87, 205-225.
18. Nisikawa, K. and Takemura, S. (1978) J. Biochem. 84, 259-266.
19. Jordan, B. R., Galling, G. and Jourdan, R. (1974) J. Mol. Biol. 87, 205-225.
20. Barber, C. and Nichols, J. L. (1978) (Can. J. Biochem. 56, 357-364.
21. Nichols, J. L. and Welder, L. (1979) Biochem. Biophys. Acta 561, 445-451.
22. Efstratiadis, A., Vournakis, J. N., Donis-Keller, H., Chaconas, G., Dougall, D., and Kafatos, F. (1977) Nucl. Acids Res. 4, 4165-4174.
23. Lockard, R. E., Alzner-DeWeerd, B., Heckman, J. E. MacGee, J., Tabor, M. W., and RajBhandary, U. L. (1978) Nucl. Acids Res. 5, 37-56.
24. Pavlakis, G. N., Jordan, B. R., Wurst, R. M., Vournakis, J. N. (1979) Nucl. Acids Res. 7, 2213-2238.
25. Pavlakis, G. N., Lockard, R. E., Vamvakopoulos, N., Riser, L. RajBhandary, U. L., and Vournakis, J. N. (1980) Cell 19, 91-102.

26. Zuker, M., and Stiegler, P., (1981), Nucl. Acids Res. 9, 133-148.
27. Bilofsky, H. (1974) Drug Information Journal May/September, 154-158.
28. Aaron, P., Rindone, W., Vary, C. P. H., Celentano, J. and Vournakis, J. N. (1982) Nucl. Acids Res., Special Computers Issue, Jan. 15, 1982, in press.
29. Donis-Keller, H. Maxam, A. M. and Gilbert, W. (1977) Nucl. Acids Res. 4, 2527-2531.
30. Simoncits, A., Brownlee, G. G., Brown, R. S., Nature 269, 833-836.
31. Komiya, H., Kawakami, M. and Takemura, S. (1981) J. Biol. Chem. 89, 717-722.
32. Averner, M. J. and Pace, N. (1972) J. Biol. Chem. 247, 4491-4493.
33. Wegney, M., Denis, H., Mayabrand, A. and Clerot, J. (1978) Devel. Biol. 62, 99-111.
34. Pace, N. R., Walker, I. A. and Pace, B. (1974) J. Mol. Evol. 3, 151-159.
35. Walker, T. A., Bety, J. L., Olah, J. and Pace, N. R. (1975) F.E.B.S. Lett. 54, 241-244.
36. Williamson, R. and Brownlee, G. G. (1969) FEBS Lett. 3, 306-308.
37. Forget, B. G. and Weissman, S. M. (1969) J. Biol. Chem. 244, 3148-3165.
38. Levy, C. C. and Karpetsky, T. P. (1980) J. Biol. Chem. 255, 2153-2159.
39. Boguski, M. S., Hierter, P. A. and Levy, C. C. (1980) J. Biol. Chem. 255, 2160-2165.
40. Vogt, V. M. (1973) Eur. J. Biochem. 33, 192-200.
41. Flashner, M. S. and Vournakis, J. N. (1977) Nucl. Acids Res. 4, 2307-2319.
42. Harada, F. and Dahlberg, J. E. (1975) Nucl. Acids Res. 2, 865-871.
43. Khan, M. S. N. and Maden, B. E. H. (1976) FEBS Lett. 72, 105-110.
44. Branlant, C., Krol, A. and Ebel, J. (1981) Nucl. Acids Res. 9, 841-858.
45. Stiegler, P., Carbon, P., Zuker, M., Ebel, J. and Ehresmann, C. (1981) Nucl. Acids Res. 9, 1253-1272.