Volume 8 Number 19 1980 Nucleic Acids Research

The 3'-terminal primary structure of five eukaryotic 18S rRNAs determined by the direct chemical method of sequencing. The highly conserved sequences include an invariant region complementary to eukaryotic 5S rRNA

Ahmed A.Azad and Nicholas J.Deacon

Molecular Biology Unit, Research School of Biological Sciences, Australian National University, Canberra, A.C.T. 2601, Australia

Received 4 August 1980

ABSTRACT

The 3'-terminal sequences of 18S rRNA from chicken reticulocyte, mouse sarcoma, rat liver, rabbit reticulocyte and barley embryo were determined by the direct chemical sequencing method. The regions sequenced show complete homology for the first 73 nucleotides. A sequence 5'-proximal to the $m_2^6Am_2^6A$ residues that is complementary to eukaryotic 5S RNAs is totally conserved. This supports the hypothesis that base-paired interaction between 5S and 18S rRNA, which are present in the large and small ribosomal subunits respectively, may be involved in the reversible association of ribosomal subunits during protein synthesis.

INTRODUCTION

It has previously been shown that there is efficient and selective hybridization, *in vitro*, between eukaryotic 5S and 18S rRNAs (1-4). An examination of primary structures has shown that identical or very similar sequences present at an identical position near the 3'-ends of eukaryotic 5S rRNAs are complementary to sequences that constitute the 5'-strands of the $m_2^6Am_2^6A$ hairpin structures near the 3'-ends of yeast and rat liver 18S rRNAs (5). These 5S rRNA sequences are identical or very similar to the sequence of the wheat embryo 5S rRNA fragment (6) that is obtained by RNAse digestion of an *in vitro* hybrid formed between wheat embryo 5S and 18S rRNA. Sequences present at an identical position in prokaryotic 5S rRNAs are also complementary to the 5'-strand of the $m_2^6Am_2^6A$ hairpin structure near the 3'-end of *E. coli* 16S rRNA (5). The extent of base-pairing and the calculated stabilities of the hybrids that can be constructed between conserved sequences in 5S rRNAs and small ribosomal subunit RNAs are quite extensive and suggest that such base-paired interaction between 5S and 18S(16S) rRNA may have physiological significance.

Because the 5S and 18S(16S) rRNA are present in the large and small ribosomal subunits, respectively, it has been suggested (1, 5) that base-pairing between 5S and 18S(16S) rRNA might provide a mechanism whereby the ribosomal subunits reversibly associate during protein synthesis.

One important piece of evidence in support of the above hypothesis would be to show that the "5S RNA-binding site" in the small ribosomal subunit RNA is conserved. In the case of prokaryotes, comparative sequence analysis has shown (7) that the sequence of the 5'-strand of the $m_2^6Am_2^6A$ hairpin structure of 16S RNA that is complementary to prokaryotic 5S rRNAs (5), is highly conserved. In the case of eukaryotes, it has been shown that wheat embryo 5S rRNA (2) and 5S rRNA of mouse sarcoma and barley embryo (Azad, unpublished results) can hybridize with heterologous 18S rRNAs, suggesting that the "5S RNA-binding site" in 18S rRNA is conserved. Moreover, it has been demonstrated that the sequences that would constitute the 3'-strands of the $m_2^6Am_2^6A$ hairpin structures of eukaryotic 18S rRNAs are homologous (8), and if the hairpin structures are conserved, as seems likely, the 5'-strand, which is complementary to eukaryotic 5S rRNA, should also be conserved. This can, however, only be confirmed by sequence analysis.

In this communication we present the primary structures of the 3'-ends of 18S rRNA of various eukaryotes obtained by the direct chemical method of sequencing (9), and show that the "5S RNA-binding site" in eukaryotic 18S rRNA is indeed conserved.

MATERIALS AND METHODS

18S rRNAs of mouse sarcoma 180 Ascites cells, hen reticulocytes, rabbit reticulocytes, barley embryo, and rat liver, were isolated and purified as described earlier (10).

Labelling of RNA, purification of labelled RNA, chemical reactions, and gel electrophoresis were carried out essentially as described by Peattie (9), with some modifications.

Labelling and recovery of 3'-end labelled RNA The reactions were carried out in 20 µl volumes containing 50 mM Hepes pH 7.5, 3.3 mM DTT, 15 mM $MgCl_2$, 10% DMSO, 10 µg/ml BSA, 75 µg 18S rRNA (108 pmole), 35 pmole [5' ^{32}P] pCp (Radiochemical Centre, 1000-

3000 Ci/m mole), 1000 pmole of ATP, and 7.4 units of T_4 RNA ligase (P.L. Biochemicals). The reaction mixtures were held at 4° for 20 h, and then 40 µl of gel loading buffer containing 10 M urea, 5 mM Tris-borate buffer, pH 8.3, 0.1 mM EDTA, 0.05% Xylene Cyanol, 0.05% Bromophenol blue, was added. The reaction mixture was incubated at 37° for 3 min and then subjected to electrophoresis in 1.2% agarose slab gel in 50 mM Tris-borate buffer, pH 8.3, 1 mM EDTA, for 2-3 h at 150 volts. The RNA bands were located by brief exposure to Kodak X-Omat film, and excised from the gel. The RNA was electro-eluted from the gel into dialysis bags in ¼-strength electrophoresis buffer containing 1% SDS, and recovered by ETOH-precipitation in the presence of carrier E. coli tRNA. The RNA precipitates were washed (3x) with 67% ETOH to remove salt and SDS, dried under vacuum, and dissolved in deionized water.

Chemical reactions were carried out exactly as described by Peattie (9).

Sequencing gels After the final lyophilisation, the chemical reaction products were dissolved in 8 M Urea, 20 mM Tris HCl, pH 7.4, 1 mM EDTA, 0.05% Xylene Cyanol, 0.05% Bromophenol blue, denatured at 90° for 45 sec and analysed on thin 10% or 20% polyacrylamide gels (11) (400 x 300 x 0.3 mm) in 8 M Urea, 100 mM Tris-borate pH 8.3, 1 mM EDTA. Gels were autoradiographed using Kodak RP Royal X-Omat film and Dupont Lighting Plus enhancement screen at -70° overnight.

RESULTS AND DISCUSSION

In order to see if a sequence complementary to eukaryotic 5S rRNA is conserved in eukaryotic 18S rRNAs we were interested in sequencing the 3'-ends of 18S rRNA of a number of eukaryotic organisms.

The 3'-terminal sequences of the small ribosomal subunit RNA of E. coli (12), Bombyx mori (13), Zea mays chloroplast (14,15), have been obtained by DNA sequencing methods using specific restriction fragments from cloned genomic rDNA. Because of the relative ease of preparing sufficient quantities of pure rRNAs, an alternative sequencing strategy would be to prepare cDNA after first polyadenylating the rRNA (8). However, in the case

of 18S RNA, reverse transcriptase cannot proceed past the $m_2^6Am_2^6A$ residues so that only short cDNA copies result (8,16). In order to determine the nucleotide sequence beyond the site of the modified bases we decided to adopt the direct chemical RNA sequencing method (9) after first 3'-terminally labelling the RNA with [5'-^{32}P] pCp using T_4 RNA ligase (17).

Typical cleavage patterns of the 3'-ends of 18S rRNAs obtained by the direct chemical sequencing method are shown in Fig. 1. Sequences were read essentially as described by Maxam and Gilbert (18). The chemical sequencing method is largely independent of secondary structure effects if the gels are run in buffer containing 8 M Urea and at about 60°. However, we were not able to completely eliminate sequence compression (19) which is an inherent problem with sequencing nucleic acids on polyacrylamide gels. By using the direct chemical sequencing method it is possible to obtain sequences on the 5'-side of the $m_2^6Am_2^6A$ residues in 18S rRNA. At positions 20 and 21 (Fig. 1) we found gaps or extremely faint A bands suggesting that the 18S rRNAs are not efficiently cleaved at these positions. DNA sequences show AA at these positions (12-15), but because of the gaps we think that in the 18S rRNA molecules these adenosines are modified and are likely to be the $m_2^6Am_2^6A$ residues that have been reported to be present near the 3'-ends of 16S and 18S rRNAs (20-24). Peattie found that there was no cleavage of the yeast 5S rRNA at position 50 where pseudouridine is present (9).

The 3'-terminal sequences of the five different 18S rRNAs obtained by the direct chemical sequencing methods, and the published 3'-terminal sequences of Bombyx mori 18S rRNA (13), and E. coli 16S rRNA (12,21) are shown in Fig. 2. The sequences on the 3'-side of the $m_2^6Am_2^6A$ residues (positions 20 and 21) obtained by direct chemical sequencing agree with published sequences (8, 25), except for the presence of a G, and sometimes also a U, at position 7 (Fig. 1). Since there has been no other report of the presence of a G(U) at this position, we do not know whether this is an artefact of the direct chemical sequencing technique. The presence of a fainter band indicating a U residue might represent some heterogeneity of sequence at this position. In the case of rabbit reticulocyte and barley 18S rRNA there was insufficient

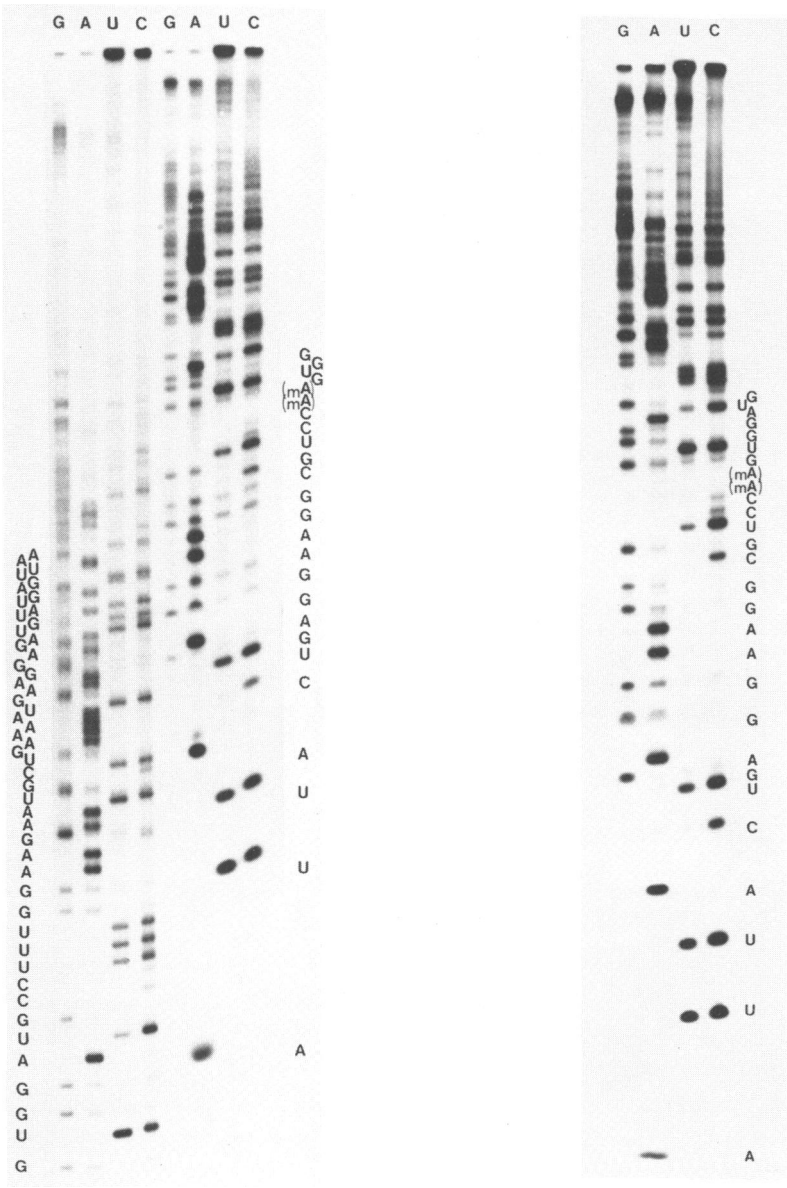

Fig. 1 Typical cleavage patterns of 18S rRNA of (a) chicken reticulocyte and (b) mouse sarcoma 180 Ascites cells, obtained by the direct chemical method of sequencing (9). The rRNAs were labelled at the 3'-termini with [α-^{32}P] pCp using T_4 RNA ligase, subjected to chemical reactions, and analysed on 20% acrylamide gels (8 M Urea) as described in Materials and Methods.

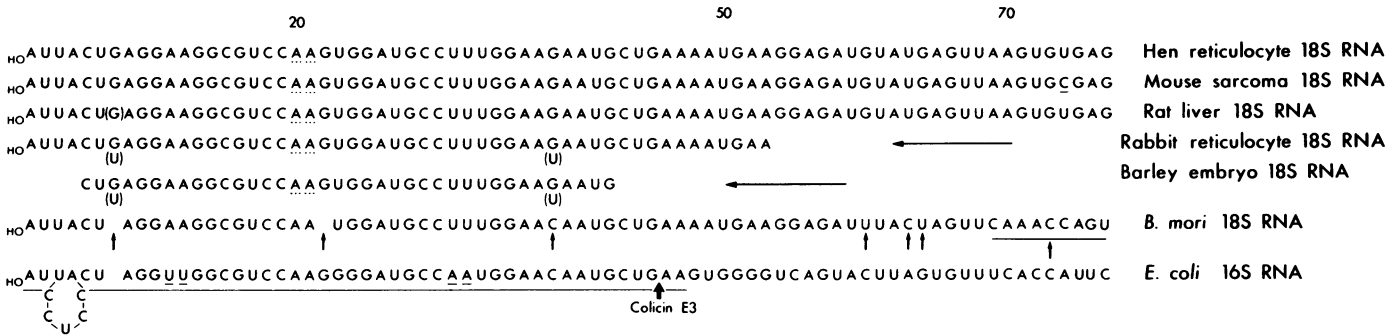

Fig. 2 The 3'-terminal sequences of 18S rRNAs of chicken reticulocyte, mouse sarcoma, rat liver, rabbit reticulocyte and barley embryo obtained by the direct chemical method of sequencing. The regions of B. mori 18S rRNA sequence (13) that differ from the above sequences are shown by vertical arrows. The region of E. coli 16S rRNA (12) that shows strong homology to eukaryotic 18S rRNA is underlined.

material to extend the sequences shown, and in addition the first few residues from the 3'-end of barley 18S RNA had run off the gel. As shown in Fig. 2, the sequences of the various eukaryotic 18S rRNAs obtained by the direct chemical sequencing technique are identical. Sequences of 18S rRNA from chicken reticulocyte, mouse sarcoma, and rat liver are shown up to position 77, and in this entire sequence there is a difference only at position 74. Chicken reticulocyte and rat liver 18S rRNA have a U at this position, while mouse sarcoma 18S rRNA and the published sequence of Bombyx mori 18S rRNA (13) have a C at this position. The sequence of a 32 nucleotide fragment of rat liver 18S rRNA has been published (25) and is identical to our sequence of rat liver 18S rRNA (positions 5-37) except that the published sequence did not have a G at position 7. The positions where the published Bombyx mori 18S rRNA (13) differs from the sequences of 18S rRNA of higher eukaryotes are shown by vertical arrows (Fig. 2). It should also be pointed out here that the 3'-terminal residue in 18S rRNA of higher plants is G and not A (26,27).

The primary structure of the 3'-end of E. coli 16S rRNA (12, 21) is strikingly similar to that of eukaryotic 18S rRNA. As shown in underlined region (Fig. 2), the only major differences are two double inversions (AA↔UU) and the presence of an extra sequence CCUCC (shown looped out) in E. coli 16S rRNA. The 3'-terminal sequence of E. coli 16S rRNA seems to be very important in protein synthesis (5). It is a surface component (28-30) and is probably present at the ribosome interface (28,30,31), and contains the site of colicin E3 cleavage (32,33). The nucleotide sequence around the colicin E3 cleavage site in E. coli 16S rRNA is conserved in eukaryotic 18S rRNA (Fig. 2). As in the case of E. coli (32,33), colicin E3 completely abolishes protein synthesis in eukaryotes (34,35). The sequence 5'-proximal to the $m_2^6Am_2^6A$ residues in eukaryotic 18S rRNA or E. coli 16S rRNA, is complementary to eukaryotic 5S RNA and prokaryotic 5S RNA, respectively (5). As can be seen in Fig. 2, this sequence (positions 23-41) is totally conserved in eukaryotic 18S rRNA.

The 3'-terminal colicin E3 fragment of E. coli 16S rRNA (21), and similar regions of the small ribosomal subunit RNA of eukaryotes (13,25 and these studies) and chloroplast (14,15), can

be formed into a hairpin structure with the $m_2^6Am_2^6A$ residues in the loop. The existence of this hairpin structure within this region of E. coli 16S rRNA is supported by high resolution proton magnetic resonance study (36). The $m_2^6Am_2^6A$ hairpin structure near the 3'-end of eukaryotic 18S rRNA is shown in Fig. 3a. A part of the 18S rRNA sequence that is complementary to 5S RNA forms the 5'-strand of the hairpin structure. Some examples of hybrids that can be constructed between the conserved 18S rRNA sequence and sequences present at an identical position near the 3'-ends of eukaryotic 5S rRNAs are shown in Fig. 3b. As can be seen from Fig. 3b the calculated free energies (37) of the plant 5S-18S rRNA hybrids are greater than those of hybrids formed between 18S rRNA and 5S RNA from animal cells. This agrees with our previous observation that the wheat embryo (2,3) and barley embryo 5S-18S rRNA hybrids formed in vitro have a higher Tm than the in vitro hybrid formed between mouse sarcoma 5S and 18S rRNAs (4). Identical regions in E. coli 16S rRNA and prokaryotic 5S RNAs are also complementary (5). The conservation of complementary sequences in 5S and 18S(16S) rRNA, the extent of base-pairing and the calculated stabilities of the hybrids that can be constructed between 5S RNA and small ribosomal subunit RNA (5), and the fact that there is efficient, selective and stable hybridization, in vitro, between 5S and 18S rRNA (1-4), would suggest that base-pairing between 5S RNA in the large ribosomal subunit and 18S(16S) rRNA in the small ribosomal subunit may have physiological significance at the level of ribosomal subunit interaction (1,5). The mechanism by which 5S and 18S(16S) rRNA may interact within the ribosome has recently been described (5).

We have previously shown that there is very efficient and stable hybridization, in vitro, between chicken globin mRNA and eukaryotic 18S rRNA (38). The 3'-strand of the $m_2^6Am_2^6A$ hairpin structure has an invariant sequence that is complementary to the 5'-non-coding regions of most eukaryotic mRNAs (8,39), and a few of the hybrids that can be constructed between the 3'-strands of the 18S rRNA hairpin and the 5'-non-coding regions of eukaryotic mRNAs are shown in Fig. 3c. Thus, the 5'- and 3'-strands of the 18S rRNA hairpin can potentially base-pair with 5S RNA and mRNA, respectively. Although the base-paired interaction between

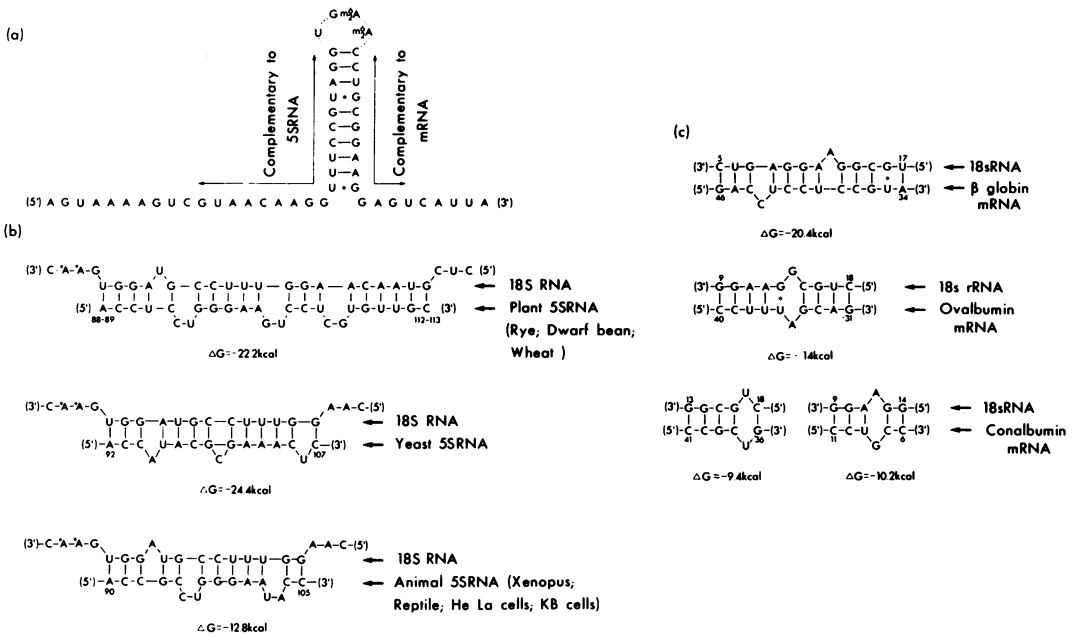

Fig. 3 a) The conserved $m_2^6Am_2^6A$ hairpin structure near the 3'-end of eukaryotic 18S rRNA.
b) A few of the hybrids that can be constructed between a conserved sequences present 5'-proximal to the $m_2^6Am_2^6A$ residues in 18S rRNA and 5S RNAs of plants, yeasts and animals. The numbers below the 5S RNA sequences indicate the position from the 5'-ends of 5S RNA molecules. The negative free energies were calculated as described by Tinoco et al. (37).
c) Some hybrids that can be constructed between the conserved region of 18S rRNA 3'-proximal to the $m_2^6Am_2^6A$ residues and sequences within the 5'-non-coding regions of chicken mRNAs for globin (40; NJD unpublished), ovalbumin (39), and conalbumin (45). Numbering is from the 3'-terminus of 18S rRNA and from the A of the ATG initiation sequence of the mRNAs.

eukaryotic mRNA and 18S rRNA may not be involved in the recognition of the mRNA initiation site (41,42) in a manner similar to that in prokaryotes (43,44), the binding of the 5'-non-coding region of eukaryotic mRNA to the 3'-strand of the $m_2^6Am_2^6A$ hairpin may facilitate the interaction between 5S and 18S rRNA by leaving the 5'-strand free to interact with 5S RNA, thus leading to the formation of the 80S initiation complex (5,38).

In conclusion, the 3'-terminal sequences of 18S rRNA of various eukaryotes have been determined by the direct chemical method of sequencing (9). A sequence 5'-proximal to the $m_2^6Am_2^6A$ residues that is complementary to eukaryotic 5S rRNAs is totally conserved in eukaryotic 18S rRNAs. This supports the hypothesis (1,5) that base-pairing between 5S and 18S rRNA may be involved in the reversible association of ribosomal subunits during protein synthesis. Other structural and biological evidence in support of this hypothesis has been described recently (5).

ACKNOWLEDGEMENTS

We would like to thank Ms. Jennifer Fernance for technical assitance and Ms. Gillian Philip for the preparation of this manuscript.

REFERENCES

1 Azad, A.A. and Lane, B.G. (1973) Can. J. Biochem. 51, 1669-1672
2 Azad, A.A. and Lane, B.G. (1975) Can. J. Biochem. 53, 320-327
3 Oakden, K.M., Azad, A.A. and Lane, B.G. (1977) Can. J. Biochem. 55, 99-109
4 Azad, A.A. (1978) Biochem. Biophys. Res. Commun. 83, 259-266
5 Azad, A.A. (1979) Nuc. Acids Res. 7, 1913-1929
6 Nichols, J.L. and Wijesinghe, W. (1978) Can. J. Biochem. 56, 760-764
7 Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B.J. and Stahl, D. (1975) Nature 254, 83-86
8 Hagenbuchle, O., Santer, M., Steitz, J.A. and Mans, R.J. (1978) Cell 13, 551-563
9 Peattie, D.A. (1979) Proc. Nat. Acad. Sci. U.S.A. 76, 1760-1764
10 Azad, A.A. (1978) Comp. Biochem. Physiol. 61B, 213-218
11 Sanger, F. and Coulson, A.R. (1978) FEBS Lett. 87, 107-110
12 Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. U.S.A. 75, 4801-4805

13 Samols, D.R., Hagenbuchle, O. and Gage, L.P. (1979) Nuc. Acids Res. 7, 1109-1110
14 Schwarz, Zs. and Kössel, H. (1979) Nature 279, 520-522
15 Schwarz, Zs. and Kössel, H. (1980) Nature 283, 739-742
16 Engel, J.D. and von Hippel, R.H. (1974) Biochemistry 13, 4143-4158
17 Bruce, A.G. and Uhlenbeck, O.C. (1978) Nuc. Acids Res. 5, 3665-3677
18 Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 560-564
19 Brown, N.L. and Smith, M. (1977) J. Mol. Biol. 116, 1-30
20 Nichols, J.L. and Lane, B.G. (1966) Biochim. Biophys. Acta 119, 649-651
21 Ehresmann, C., Stiegler, P., Mackie, G.A., Zimmermann, A., Ebel, J.P. and Fellner, P. (1975) Nuc. Acids Res. 2, 265-278
22 Lau, R.Y., Kennedy, T.D. and Lane, B.G. (1974) Can. J. Biochem. 52, 1110-1123
23 Khan, M.S.N. and Maden, B.E.H. (1976) J. Mol. Biol. 101, 235-254
24 de Jonge, P., Klootwijk, D. and Planta, R.J. (1977) Nuc. Acids Res. 4, 3665-3664
25 Alberty, H., Raba, M. and Gross, H.J. (1978) Nuc. Acids Res. 5, 425-434
26 Azad, A.A. and Lane, B.G. (1973) Can. J. Biochem. 51, 1195-1202
27 Oakden, K.M. and Lane, B.G. (1976) Can. J. Biochem. 54, 261-271
28 Chapman, N.M. and Noller, H.F. (1977) J. Mol. Biol. 109, 131-149
29 Santer, M. and Shane, S. (1977) J. Bact. 130, 900-910
30 Schrieber, J.P., Hsiung, N. and Cantor, C.R. (1979) Nuc. Acids Res. 6, 181-193
31 Politz, S.M. and Glitz, D.D. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 1468-1472
32 Bowman, C.M., Sidikaro, J. and Monura, M. (1971) Nature New Biol. 234, 133-137
33 Senior, B.W. and Holland, I.B. (1971) Proc. Nat. Acad. Sci. U.S.A. 68, 959-963
34 Turnowsky, F., Drews, J., Eich, F. and Högenauer, G. (1973) Biochem. Biophys. Res. Commun. 52, 327-334
35 Suzuki, H. (1978) FEBS Lett. 89, 121-125
36 Baan, R.A., Hilbers, C.W., Van Charldorp, R., Van Leerdam, E., Van Knippenberg, P.H. and Bosch, L. (1977) Proc. Nat. Acad. Sci. U.S.A. 74, 1028-1031
37 Tinoco, I. jun., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40-41
38 Azad, A.A. and Deacon, N.J. (1979) Biochem. Biophys. Res. Commun. 86, 568-576
39 Kuebbing, D. and Liarakos, C.D. (1978) Nuc. Acids Res. 5, 2253-2266
40 Richards, R., Shine, J., Ullrich, A., Wells, J.R.E. and Goodman, H.M. (1979) Nuc. Acids Res. 7, 1137-1146
41 Baralle, F.E. and Brownlee, G.G. (1978) Nature, 274, 84-87
42 De Wachter, R. (1979) Nuc. Acids Res. 7, 2045-2054
43 Shine, J. and Dalgarno, L. (1974) Proc. Nat. Acad. Sci. U.S.A. 71, 1342-1346

44 Steitz, J.A. and Jakes, K. (1975) Proc. Nat. Acad. Sci. U.S.A. 72, 4734-4738
45 Cochet, M., Gannon, F., Hen, R., Maroteaux, L., Perrin, F. and Chambon, P. (1979) Nature, 282, 567-574