Volume 8 Number 22 1980 Nucleic Acids Research

A mammalian mitochondrial serine transfer RNA lacking the "dihydrouridine" loop and stem

M.H.L.de Bruijn, P.H.Schreier*, I.C.Eperon and B.G.Barrell
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK, and

E.Y.Chen, P.W.Armstrong, J.F.H.Wong and B.A.Roe
Chemistry Department, Kent State University, Kent, OH 44242, USA

Received 22 September 1980

ABSTRACT

A unique transfer RNA has been identified in human and bovine mitochondria that lacks the "dihydrouridine" loop and stem structure. This tRNA is mitochondrially coded as shown by DNA sequence analysis of the human and bovine mitochondrial DNA. Sequence analysis of the RNA shows that it is post-transcriptionally modified by the addition of CCA at the 3' terminus and that at least one base is modified. As predicted by its anticodon (GCU, corresponding to the serine codons AGU/C) this tRNA can be aminoacylated with serine when purified mitochondria are incubated in a medium containing ^{3}H-serine.

INTRODUCTION

Analysis of the complete sequence of human mitochondrial DNA (mtDNA) and also of the near complete bovine mtDNA sequence [1,2] has enabled us to identify the genes for 22 tRNAs while the gene for a further tRNA ($tRNA_m^{Met}$) has not yet been unambiguously identified. These 22-23 mt tRNAs are thought to be sufficient to translate the different mammalian mitochondrial genetic code [1,3; I.G. Young and S. Anderson, manuscript in preparation). The sequences of some of these tRNAs have already been published [2-6]. They differ from all other known tRNAs [7] in that sequences and sizes of the "dihydrouridine" and "TψC" loops vary considerably, even when homologous human and bovine species are compared.

We report here the sequence of an extreme example of these different tRNAs, a serine tRNA from human and bovine mitochondria that lacks the entire "dihydrouridine" arm. This tRNA is identical to $3S_E$ RNA, the small mtRNA component which has previously been suggested to be an equivalent of cytoplasmic 5S RNA [8,9].

MATERIALS AND METHODS

DNA sequence analysis

Restriction enzyme fragments of human mtDNA were cloned in pBR322 [10] and sequenced using the chain termination method of Sanger et al. [11,12]. Single stranded template DNA was prepared using exonuclease III [13].

Restriction enzyme fragments of beef heart mtDNA were similarly cloned (I.G. Young and S. Anderson, manuscript in preparation). The cloned fragments were excised and isolated by electrophoresis in low melting point agarose gels (Bethesda Research Laboratories, Rockville, Md. 20850) using a buffer containing 40 mM Tris, 20 mM sodium acetate, 2 mM EDTA, adjusted to pH 8.3. Bands were excised and melted at 70° C for 10 min in an equal volume of electrophoresis buffer and then phenol-extracted three times prior to ethanol precipitation. Sequence analysis of these DNA fragments was by randomly "shotgunning" restriction enzyme digests into the single stranded vector M13mp2 [14,15] as described by Sanger et al. [5]. These clones were sequenced using a "universal" primer [15,16]. The sequences obtained by these methods were collated and analysed using the computer programs of Staden [17].

Isolation of mt tRNA for sequence analysis

Mitochondria were isolated from fresh bovine livers as described [18,19]. Total mitochondrial nucleic acid was isolated by extraction with phenol in 0.14 M NaCl, 0.14 M sodium acetate, pH 5.0 and the mt tRNA fraction purified by chromatography on DEAE-cellulose [20]. The mt tRNAs were labelled at their 3' termini with ^{32}P-pCp (Amersham/Searle) and T4 RNA ligase [21]; the molar ratio of tRNA to ^{32}P-pCp was 3:1. The mt tRNAs showed no contamination with cytoplasmic tRNAs when analysed by two-dimensional polyacrylamide gel electrophoresis (PAGE) before and after hybridisation to bovine mtDNA.

For sequence analysis the 3' ^{32}P-pCp-labelled total mt tRNA was fractionated by two-dimensional PAGE (0.35 mm x 20 cm x 40 cm) in TBE buffer (90 mM Tris, 90 mM boric acid, 2.5 mM EDTA, pH 8.3) using 20% polyacrylamide-7 M urea as the first and 20% polyacrylamide-3 M urea as the second dimension. The tRNA was dissolved in 10 µl of formamide containing 1.5 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol and loaded onto the first gel. After PAGE at 1.8 kV for 24-36 hr at 4° C the tRNAs were located by autoradiography, the bands excised and embedded in a 5 cm deep stacking gel which was poured above a polymerised 20% polyacrylamide-3 M urea gel. The stacking gel consisted of 5% polyacrylamide, 50 mM Tris, pH 3.5 and 0.001% riboflavin-5'-phosphate; it was polymerised by illumination after embedding the first dimension gel slice. PAGE was at 1.8 kV for 20-24 hr at 4° C and was followed

by autoradiography. For ^{3}H-base analysis unlabelled mt tRNAs were located by staining with methylene blue.

The three fastest migrating tRNA species, which only differed in the number of residues at the 3' terminus, were excised from the gel and extracted by soaking in 10 vol of 0.5 M NaCl, 10 mM Tris, pH 7.6 and 1 mM EDTA for 3 hr at 37°C. The gel slices were removed and the RNA, containing typically 200,000 dpm, ethanol-precipitated in the presence of about 0.1 A_{260} of carrier E. coli tRNA.

RNA sequence analysis

PAGE for ladder sequence analysis was performed on 20% polyacrylamide-7 M urea gels [22-24]; the conditions for partial alkaline hydrolysis were as described by Lockard [25]. In addition to RNase T_1 and RNase U_2, an avian liver ribonuclease, RNase Cl_3, was used for locating cytosine residues [26].

Total base analysis

The RNA, prepared as above except that no carrier tRNA was added, was enzymatically digested to nucleosides, periodate oxidised, labelled with ^{3}H-borohydride (Amersham/Searle) and resolved on cellulose TLC (Brinkman) [27].

Aminoacylation of mt tRNAs

Mitochondria were isolated from fresh bovine hearts essentially using procedure number 3 of Smith [28]. Incubation and lysis of mitochondria was as described by Wallace and Freeman [29]. The medium contained 19 unlabelled amino acids and L-[3-^{3}H]serine (11 Ci/mmole, Amersham) at 25 μCi/ml or L-[4, 5-^{3}H]leucine (137 Ci/mmole, Amersham) at 10 μCi/ml. The charged RNA was isolated by phenol extraction and ethanol precipitation and then deaminated by treatment with nitrous acid as described by Clarkson et al. [30].

Fractionation of this material was by PAGE using 20% polyacrymalide-7 M urea gels (0.1 x 20 x 40 cm) in TBE buffer at 12 V/cm for 13 hr; the acrylamide to bisacrylamide ratio was 310:1 [31]. RNA samples were dissolved in 7.5 mM NaCl, 0.75 mM sodium citrate (pH 7.0), 50% formamide, 2.5 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol. Prior to loading they were incubated at 50°C for 3 min and cooled rapidly on ice. After PAGE the gels were stained with 0.5% methylene blue, photographed, prepared for fluorography [32] and exposed to Kodak X-Omat H film.

Amino acid analysis

Aminoacylated but non-deaminated RNA samples were ethanol-precipitated, dissolved in 6 N HCl, hydrolysed under vacuum for 24 hr at 105°C and evaporated to dryness [33]. The radioactive amino acids present in the samples were identified using a Durrum analyser modified for peak collection [34] and

quantitated by liquid scintillation counting.

RESULTS

Analysis of the complete sequence of human mtDNA has revealed only one candidate for a $tRNA^{Ser}$ capable of translating the 53 AGU/C serine codons found in the human mitochondrial genes for cytochrome c oxidase subunits I, II and III, cytochrome b, an ATPase and in eight so far unidentified putative genes [2]. The prediction of this unusual structure (Fig. 1) as the mt $tRNA^{Ser}_{AGU/C}$ is based on the fact that it has a classical anticodon loop sequence and that the aminoacyl, TψC and anticodon stems have the correct size and relationship to each other [7], although base pairing in the anticodon stem is somewhat less than normal. Furthermore, the two mt $tRNA^{Ser}$ genes are on different strands, which agrees with the results of Lynch and Attardi [18]. The position of this gene relative to others is also very suggestive: a $tRNA^{His}$ gene is adjacent to the 5' terminus and a $tRNA^{Leu}$ gene to the 3' terminus with no non-coding bases between them. As with all other mammalian mt tRNA genes the $C\text{-}C\text{-}A_{OH}$ 3' terminus is not encoded in the DNA sequence. The phenomenon of "butt-jointed" genes or of only a few bases

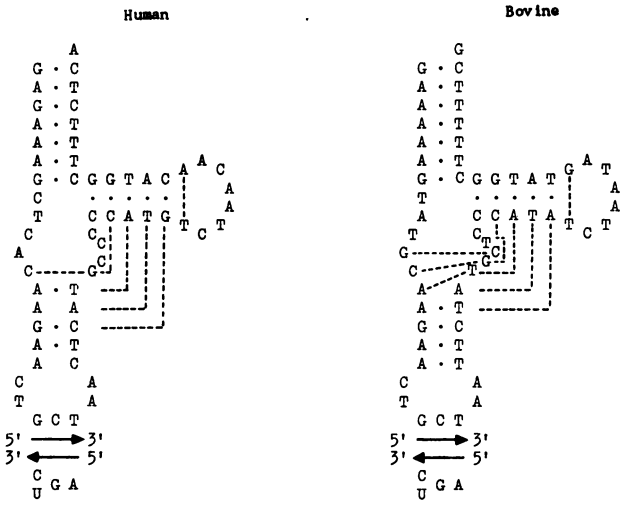

Figure 1. DNA sequences of human and bovine $tRNA^{Ser}$(AGU/C) arranged in the "cloverleaf" secondary structure showing additional anticodon base pairs and a possible alternative base-pairing. RNA sequencing data show that $C\text{-}C\text{-}A_{OH}$ is added post-transcriptionally at the 3' end and that the A next to the anticodon is modified to t^6A. The m^5C is tentatively located at position 35.

between them is typical of the mammalian mt genome and examples have already been published [3,4].

Sequence analysis of the bovine mtDNA provided important confirmation that this structure was conserved and was the only $tRNA^{Ser}_{AGU/C}$ candidate. The bovine gene, of which both strands have been analysed, is found in the same position with respect to strand and flanking tRNA genes except that it is separated by one base from the $tRNA^{Leu}$ gene.

If this potential gene was transcribed, then an obvious candidate was the small mtRNA component ($3S_E$ RNA [8]) which migrates faster than the bulk mt tRNA on PAGE. Sequence analysis of this bovine RNA, using the gel sequencing technique described, enabled us to show that its sequence corresponds to that predicted from the bovine mtDNA, and that it is post-transcriptionally modified by the addition of CCA at the 3' terminus. In addition, base analysis (average of six experiments) showed that N^6(N-threonyl-carbamoyl)adenosine (t^6A) was present in 0.9 molar yield and that 5-methylcytidine (m^5C) was present in 0.43 molar yield.

The RNA sequencing technique applied is useful for obtaining rapid sequence information though, by itself, is not sufficient for complete sequence determination as the U residues must be deduced from the absence of any cleavage with the ribonucleases T_1, U_2 and Cl_3 or in other cases when a band is seen in all sequencing tracks (B.A. Roe, E.Y. Chen and P.W. Armstrong, (unpublished observations). Therefore, any nucleotide which is not susceptible to cleavage by these enzymes, for instance due to modification, may show as U even though it may not be derived from U. Further RNA sequence analysis employing the one-hit kinetics method [35,36] has directly confirmed the sequence of nucleotides 4-38 and located the t^6A adjacent to the 3' side of the anticodon. In addition a partial modification of base 35 to m^5C was found.

No differences are found between the bovine DNA and RNA analyses except for the nucleotides at positions 50 and 51. The DNA sequence is T-G but the RNA sequencing gel shows G-U in this position. Whether this discrepancy is real and due to different individual animals is not known. This region is predicted to be base paired as shown in Fig. 1 and, if the difference is real, G-U would result in a less stable structure. It is possible that the difference is due to inversion of the U-G sequence because of secondary structure or base stacking effects persisting in the RNA sequencing gel. No differences are found between the bovine mtDNA sequence and the bovine RNA sequence presented by Arcari and Brownlee (accompanying paper). They have

also located t^6A next to the anticodon but have not been able to identify m^5C. Whether this is due to the use of different tissues is not known.

In order to test whether this RNA can be aminoacylated with serine, purified bovine mitochondria were incubated in the presence of tritiated serine and the other 19 unlabelled amino acids, after which the RNA was extracted. Fig. 2 shows PAGE analysis of this material. Lane A shows charged mt tRNA stained with methylene blue. Lane C is the fluorogram of lane A showing the ^{3}H-serine charged species. Two strongly labelled bands are seen, one (band 1) migrating rapidly and the other (band 3) migrating slowly with a faint band (band 2) migrating with the bulk mt tRNA. The fast band is coincident with the stained fast moving band (lane A) and also with a sample of purified $tRNA^{Ser}_{AGU/C}$ (lane D) from which the RNA sequence was derived. No other mt tRNA migrates in this region. Lanes E and F show 3' ^{32}P-pCp-labelled uncharged mt tRNA and cytoplasmic tRNA respectively. The $tRNA^{Ser}_{AGU/C}$ species is only present in the mt tRNA preparations. Lane B shows a fluorogram of mt tRNA charged with ^{3}H-leucine and was added as a marker; the DNA sequences of the two leucine tRNAs are 70 and 75 nucleotides long.

Prior to deamination and fractionation the purified aminoacylated RNA containing the species charged with ^{3}H-serine was subjected to amino acid

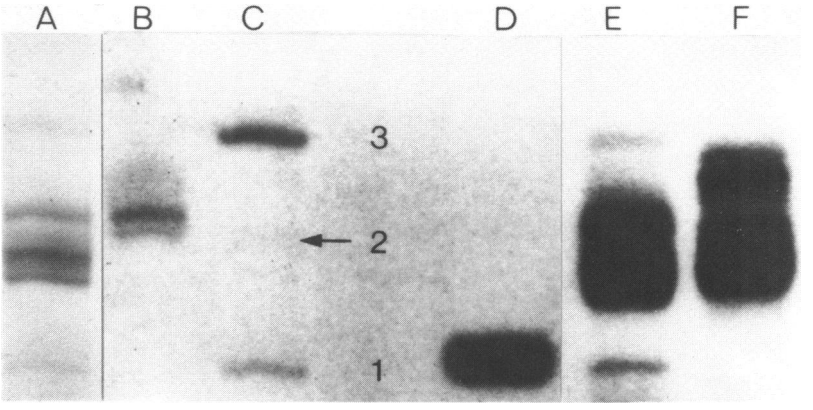

Figure 2. Electrophoretic analysis of aminoacylated bovine mt tRNA. A: mt tRNA charged with ^{3}H-serine and stained with methylene blue. B: fluorogram of mt tRNA charged with ^{3}H-leucine. C: fluorogram of lane A. D: fluorogram of 3' end ^{32}P-pCp-labelled $tRNA^{Ser}(AGU/C)$ used in the ladder sequence analysis. E: fluorogram of 3' end ^{32}P-pCp-labelled uncharged mt tRNA. F: fluorogram of 3' end ^{32}P-pCp-labelled uncharged cytoplasmic tRNA. All lanes originate from the same gel. Lanes B to E contained unlabelled total mt tRNA, the methylene blue stains providing an internal marker. Fluorography ranged from 10 (E,F) to 20 (B,C,D) days.

analysis. This showed that 96% of the labelled amino acid was present as serine, the largest contaminant being threonine at 2%, so that only serine can account for the fast moving band on the fluorogram. Thus there can be no doubt that this RNA can be aminoacylated with serine. We have not yet identified which species the other two bands on the fluorogram (Fig. 2, lane C) correspond to but one of them presumably represents the other $tRNA^{Ser}$ that recognises the codon family UCN [1].

DISCUSSION

The above results leave no doubt that $3S_E$ RNA is a mt $tRNA^{Ser}_{AGU/C}$ species rather than a "mini-5S" ribosomal RNA as has been suggested previously [8,9]. The latter view was based on the similarity of $3S_E$ RNA to mt rRNA in molar abundance and base ratio and on the failure to detect any methylation; also, its size in relation to the large mitochondrial ribosomal subunit (16S) favoured this view when compared to cytoplasmic 5S and 28S rRNA. Two important questions regarding this tRNA remain: does it function in protein synthesis and, if so, how can it fold up in a tertiary structure that will enable serine to be placed in the correct position on the ribosome to allow chain elongation?

Firstly, there must be a $tRNA^{Ser}$ in the mitochondrion to recognise the codons AGU/C as these codons are found in known mitochondrial genes, e.g. in bovine cytochrome c oxidase subunit II where the complete amino acid sequence [37] and the gene sequence are known. Although we cannot at this stage rigorously rule out the possibility that a single cytoplasmic tRNA specific for these serine codons is imported into the mitochondrion, this seems unlikely as the two major bands charged with ^{3}H-serine (Fig. 2) do not coincide with a cytoplasmic tRNA.

The second crucial question concerns the three-dimensional structure of this tRNA and, in particular, whether it can have a conformation like that found for yeast $tRNA^{Phe}$ [38,39]. This question can be addressed in the wider context of the conformations of the other 21 mt tRNA structures. Although these other tRNAs have the four normal base-paired stems, all except one ($tRNA^{Leu}_{UUA/G}$) are defective in that they lack several of the conserved bases observed to stabilise the yeast $tRNA^{Phe}$ structure. The most common defects involve the lack of G_{18} and G_{19} which hold the dihydrouridine and TψC loops together, but the interactions between the beginning of the dihydrouridine loop (A_{14} and G_{15}) and the connecting region (U_8 and C_{48}) are also defective in a few cases. Apart from these rather specific defects, the structures

appear normal suggesting that they could have a conformation like that of yeast $tRNA^{Phe}$ but stabilised by fewer tertiary interactions.

The $tRNA^{Ser}$ sequences presented here can be regarded as an extreme case of lacking tertiary interactions, now involving the entire dihydrouridine arm (loop and stem). Nevertheless in the mt $tRNA^{Ser}_{AGU/C}$ sequence presented here and in that of rat (K. Koike, personal communication), additional base pairs can be formed extending the anticodon stem to have up to eight base pairs, which could replace the dihydrouridine stem. The extended anticodon stem could then remain perpendicular to the amino acid/TψC stem, so that the relative position and orientation of the aminoacyl stem and anticodon stem could be normal. Thus $tRNA^{Ser}_{AGU/C}$ could participate in protein synthesis, although perhaps with a lower efficiency. This would be acceptable in view of the finding that only 53 AGU/C serine codons are used against 221 UCN serine codons in human mtDNA [2].

The small size of the tRNA may be another example of the very economic use made of mammalian mtDNA. Most genes are butt-jointed or else have very few non-coding bases between them. All tRNAs and rRNAs are small compared to their cytoplasmic counterparts. Also, the use of AUA as a methionine codon and of AGA/G as termination codons means that the genome has to code for two tRNAs less [1-3]. Although not proven, it seems unlikely that the mammalian mitochondrial genome, with its highly economical use of coding capacity, would code for a non-functional tRNA-like structure.

One strange feature of the $tRNA^{Ser}_{AGU/C}$ sequences concerns alternative base pairing between four bases at the end of the anticodon stem and four at the beginning of the TψC stem (Fig. 1). These pairs cannot be incorporated into a cloverleaf secondary structure that has a normal anticodon arm. All the three $tRNA^{Ser}_{AGU/C}$ sequences known at present show this potential for alternative base pairing. Although this may only reflect the high degree of conservation of the specific anticodon and TψC stem sequences, it suggests that alternative base pairing has some functional significance. One effect might be to decrease the level of functional tRNA either by enhancing degradation or by decreasing aminoacylation.

Shortly before this paper was submitted a sequence of $3S_E$ RNA in hamster mitochondria was published [40] and proposed to be a 5S RNA equivalent. The sequence shows high homology to those presented here for human and bovine mitochondria except for an extra G on its 5' terminus. A secondary structure as presented in Fig. 1 can be drawn as an alternative to the proposed structure of a 5S RNA equivalent. The additional evidence supporting 5S RNA

equivalence is a region of complementarity in hamster $3S_E$ RNA to the small mitochondrial ribosomal subunit. Less complementarity is found in the homologous human and bovine systems. In view of the evidence presented here for a tRNA the significance of this complementarity is not known.

ACKNOWLEDGEMENTS

We thank M. Levitt, A. Klug and F.H.C. Crick for valuable advice and discussions. We also thank K.A. Edwards for performing the amino acid analyses. B.A.R. was supported by NIH grants GM-21405 and GM-25962 and is a recipient of NIH Research Career Development Award GM-00124. P.H.S. was supported by the Deutsche Forschungsgemeinschaft through SFB-74 to B. Müller-Hill as well as a short term EMBO fellowship.

* Present address: Institut für Genetik der Universität, 5 Koln 41, Weyertal 121, West Germany.

REFERENCES

1. Barrell, B.G., Anderson, S., Bankier, A.T., de Bruijn, M.H.L., Chen, E., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith A.J.H., Staden, R. and Young, I.G. (1980). Proc. Natl. Acad. Sci. USA 77, 3164-3166.
2. Barrell, B.G., Anderson, S., Bankier, A.T., de Bruijn, M.H.L., Chen, E., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1980). In Proceedings of the 31st Mosbach Colloquium (Springer-Verlag), in press.
3. Barrell, B.G., Bankier, A.T. and Drouin, J. (1979). Nature 282, 189-194.
4. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980). Nature 286, 460-467.
5. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980). J. Mol. Biol., in press.
6. Crews, S. and Attardi, G. (1980). Cell 19, 775-784.
7. Gauss, D.H., Gruter, F. and Sprinzl, M. (1979). Nucleic Acids Res. 6, r1-r19.
8. Dubin, D.T., Jones, T.H. and Cleaves, G.R. (1974). Biochem. Biophys. Res. Comm. 56, 401-406.
9. Dubin, D.T. and Friend, D.A. (1972). J. Mol. Biol. 71, 163-175.
10. Drouin, J. (1980). J. Mol. Biol. 140, 15-34.
11. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.
12. Sanger, F. and Coulson, A.R. (1978). FEBS Letters 87, 107-110.
13. Smith, A.J.H. (1979). Nucleic Acids Res. 6, 831-848.
14. Gronenborn, B. and Messing, J. (1978). Nature 272, 375-377.
15. Schreier, P.H. and Cortese, R. (1979). J. Mol. Biol. 129, 169-172.
16. Anderson, S., Gait, M.J., Mayol, L. and Young, I.G. (1980). Nucleic Acids Res. 8, 1731-1743.
17. Staden, R. (1980). Nucleic Acids Res. 8, 3673-3694.
18. Lynch, D. and Attardi, G. (1976). J. Mol. Biol. 102, 125-141.

19. Almaric, F., Merkel, C., Gelfand, R. and Attardi, G. (1978). J. Mol. Biol. 118, 1-25.
20. Roe, B.A. (1975). Nucleic Acids Res. 2, 21-42.
21. Bruce, A.G. and Uhlenbeck, O.C. (1978). Nucleic Acids Res. 5, 3665-3677.
22. Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977). Nature 269, 833-836.
23. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977). Nucleic Acids Res. 4, 2527-2538.
24. Chen, E.Y. and Roe, B.A. (1978). Biochem. Biophys. Res. Comm. 82, 235-246.
25. Lockard, R.E., Alzner-Deweerd, B., Heckman, J.E., MacGee, J., Tabor, M.W. and RajBhandary, U.L. (1978). Nucleic Acids Res. 5, 37-56.
26. Levy, C.C. and Karpetsky, T.P. (1980). J. Biol. Chem. 255, 2153-2159.
27. Randerath, K., Gupta, R.C. and Randerath, E. (1979). In Transfer RNA: Structure, Properties and Recognition. Cold Spring Harbor Monograph Series.
28. Smith, A.L. (1967). Methods in Enzymology X, 81-86.
29. Wallace, R.B. and Freeman, K.B. (1974). Biochem. Biophys. Res. Comm. 60, 1440-1445.
30. Clarkson, S.G., Birnstiel, M.L. and Serra, V. (1973). J. Mol. Biol. 79, 391-410.
31. Blattler, D.O., Garner, F., Van Slyke, K. and Bradley, A. (1972). J. Chromatogr. 64, 147-155.
32. Laskey, R.A. (1980). Methods in Enzymology 65, 363-371.
33. Mendez, E. and Lai, C.Y. (1975). Analyt. Biochem. 68, 47-53.
34. Walker, J.E., Shaw, D.C., Northrop, F.D. and Horsnell, T. (1977). INSERM Symposium No. 5 (Previero, A. and Coletti-Previero, M.-A., eds), Elsevier/North-Holland Biomedical Press 277-285.
35. Stanley, J. and Vassilenko, S. (1978). Nature 274, 87-89.
36. Gupta, R.C., Roe, B.A. and Randerath, K. (1980). Biochemistry 19, 1699-1705.
37. Steffens, G.J. and Buse, G. (1979). Hoppe-Seyler's Z. Physiol. Chem. 360, 613-619.
38. Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, A.H.J., Seeman, N.C. and Rich, A. (1974). Science 185, 435-440.
39. Ladner, J.E., Jack, A., Robertus, J.D., Brown, R.S., Rhodes, D., Clark, B.F.C. and Klug, A. (1975). Proc. Nat. Acad. Sci. USA 72, 4414-4418.
40. Baer, R.J. and Dubin, D.T. (1980). Nucleic Acids Res. 8, 3603-3610.