Volume 7 Number 4 1979 Nucleic Acids Research

# The nucleotide sequence of human $tRNA^{Gly}$ (anticodon GCC)

Ramesh C.Gupta*, Bruce A.Roe** and K.Randerath*

*Department of Pharmacology, Baylor College of Medicine, Houston, TX 77030, and
**Department of Chemistry, Kent State University, Kent, OH 44242, USA

Received 9 August 1979

## ABSTRACT

The sequence of $tRNA^{Gly}_{GCC}$ from human placenta was determined by recently developed postlabeling techniques. The tRNA was digested completely with RNases T1 and A in the presence of alkaline phosphatase, the oligonucleotides were 3'-terminally ($^3$H)-labeled, mapped on PEI-cellulose thin layers, isolated, and sequenced by methods based on base-specific cleavages. Overlaps were obtained by readout sequencing techniques on polyacrylamide gels and PEI-cellulose thin layers. The thin-layer readout technique was used also to locate and identify modified nucleotides. The primary structure was found to exhibit a large degree of homology (94.6%) with silkworm $tRNA^{Gly}_{GCC}$ but only 67.6% homology with human $tRNA^{Gly}_{CCC}$.

## INTRODUCTION

About 110 tRNA sequences are now known[1] but only relatively few of these are from plants and animals. The development of rapid and sensitive postlabeling techniques[2-7] has facilitated the sequence analysis of mammalian tRNAs which are frequently available only in small amounts. Using such techniques we have determined the primary structure of $tRNA^{Gly}_{GCC}$ from human placenta.

## MATERIALS

Polynucleotide kinase (P-L Biochemicals), nuclease $S_1$ (Miles Laboratories), nuclease $P_1$ (Yamasa Shoyu, Co., Tokyo), RNases $T_2$, $T_1$, $U_2$ (Sankyo), and A (Sigma) were used without further purification. RNase $Phy_1$[8] was generously provided by Dr. J.-P. Bargetzi. Yeast tRNA (Sigma type I) was used as a carrier after deproteinization. ($^3$H)$NaBH_4$ (30-60 Ci/mmole) and ($\gamma$-$^{32}$P)ATP (1500-2500 Ci/mmole) were from Amersham-Searle and

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

ICN, respectively. ($^3$H)NaBH$_4$ was dissolved in 0.1 N KOH and stored in lyophilized form[2,9]. Ultrapure urea was obtained from Schwarz/Mann, and acrylamide and methylene bisacrylamide from Bio-Rad. PEI-cellulose thin layers were prepared and purified as described[10].

## METHODS

Preparation of human placenta $tRNA^{Gly}_{GCC}$.

Partially purified $tRNA^{Gly}$ from human placenta[11] was resolved into 4 peaks by RPC-5 chromatography[12]; peak 2 was used in the present study. The tRNA was further purified by electrophoresis on a 20% polyacrylamide, 7 M urea, pH 8.3, slab gel (40 x 20 x 0.2 cm) and extracted from the gel[13].

Analysis of the fragments in complete RNase $T_1$ and A digests.

(i) Preparation and ($^3$H)-labeling of the digests. tRNA (1 μg/μl) was incubated with 0.1 μg/μl RNase $T_1$ or RNase A, 0.1 μg/μl alkaline phosphatase, and 50 mm bicine, pH 7.8, at 38° for 1½ hr. One μl of 30 mM $NaIO_4$ was added to 4 μl of the digest and the solution kept at 23° for 2 hr. After cooling on ice, 1 μl of 0.1 M potassium phosphate, pH 6.8, was added, followed by 2 μl of 0.3 M ($^3$H)NaBH$_4$ and the solution was kept at 23° for 2 hr. Excess borotritide was destroyed by adding 6 μl of 5N acetic acid and the solution dried in a stream of air. The residue was dissolved in 20 μl of water.

(ii) Fractionation of the digests. Postlabeled digest derived from 4-12 μg of tRNA was applied in 5-μl portions to a PEI-cellulose thin-layer sheet (25 x 20 cm) at 2 cm from the left-hand edge and 1.5 cm from the bottom edge. After the origin had been dried, radioactive contaminants were removed by predevelopment of the sheet with water to the origin and then with 4 M lithium formate, 7 M urea, pH 3.5, to 4.5 cm above the origin. The wet sheet was soaked in 300 ml of methanol/conc. ammonia (1000:1, by vol.) for 10 min., dried, and cut at 1 cm above the origin. The lower part of the sheet containing most of the contaminants was discarded. After a Whatman 1 paper wick had been attached to the top of the upper part of the sheet, the chromatogram was developed with water to 2 cm above the lower edge of the chromatogram,

then with 0.05 M LiCl to 4 cm, 0.20 M LiCl to 10 cm, 0.35 M LiCl to 20 cm, and 0.50 M LiCl to 6-7 cm on the wick. After the wick had been cut off, the chromatogram was dried with warm air and soaked in two 300-ml portions of methanol for 10 min each. A wick was attached to the top of the second dimension which was then developed with water to the origin, 0.2 M ammonium formate, pH 2.6, to 4 cm, 1.0 M ammonium formate, pH 2.6, to 9 cm, 2.5 M ammonium formate, pH 2.6, to 16 cm, and 4.0 M ammonium formate, pH 2.6, to 5-6 cm on the wick. The chromatogram was dried, soaked once in 300 ml of methanol for 10 min., and again dried. The ($^{3}$H)-labeled oligonucleotides were located by fluorography[14]. The gradients described were suited to fractionate the components of the RNase $T_1$ digest. For the RNase A digest, the final solvents were 0.35 M LiCl and 2.5 M ammonium formate, pH 2.6, respectively, which resulted in a better resolution of the shorter fragments (chain lengths 3-5).

(iii) Elution of the labeled oligonucleotides. The oligonucleotides were eluted from PEI-cellulose with 4 M pyridinium formate, pH 4.2, in phosphocellulose minicolumns as described[9], except that the lower parts of the columns were drawn out into capillaries (elution time 10-15 min.). The eluates were lyophilized and the residues dissolved in 300 $\mu$l of water. The solutions were lyophilized to remove residual pyridinium formate and the residues finally extracted with two 500-$\mu$l portions of ether. The dry residues were dissolved in 40 $\mu$l of water.

(iv) Sequence analysis of the labeled oligonucleotides. To determine the nucleotide sequence of RNase $T_1$ and A fragments, the base composition, the 3'- and 5'-termini and the sequence of each oligonucleotide-3' dialcohol were analyzed. A scaled-down version[9] of a ($^{3}$H)-derivative method was used for base analysis. Since this involved digestion of the 3'-end-labeled oligonucleotide with RNase $T_2$[9], the 3'-terminus was also obtained by this technique as a ($^{3}$H)-labeled nucleoside trialcohol and identified by thin-layer chromatography[15]. The 5'-terminus was determined by 5'-($^{32}$P)-labeling of the oligonucleotide-3' dialcohol (2-5 pmoles) followed by the enzymatic release of the labeled terminus and subsequent chromatographic identification as a 5'-($^{32}$P)-labeled nucleoside monophosphate[9,16]. For sequence analysis[2,4],

labeled RNase $T_1$ fragments ($4 \times 10^4$-$4 \times 10^5$ dpm; 1-10 pmoles) were partially digested with RNases $U_2$, $Phy_1$, and A, respectively; RNase A fragments were partially digested with RNases $T_1$ and $U_2$, respectively. The products were resolved by size on PEI-cellulose thin layers and aligned by fluorography with products of partial nuclease $S_1$/phosphatase digestion of the labeled oligonucleotide[2,4].

Analysis by readout techniques.

(i) Thin-layer readout technique[7]. 5 μg of the tRNA was heated in 10 μl of double-distilled water at 80° for 6 min. 5'-hydroxyl groups of the generated fragments were labeled with ($\gamma$-$^{32}$P)ATP and polynucleotide kinase, the labeled fragments resolved by size on denaturing polyacrylamide gels and contact-transferred onto a PEI-cellulose thin layer. The radioactive 5'-termini were released by in situ RNase $T_2$ digestion of the labeled fragments on the PEI-cellulose thin layer and identified by thin-layer chromatography[7].

(ii) Gel readout technique[2-5]. The tRNA was 5'-($^{32}$P)-labeled under conditions described by Silberklang et al[17], except that the tRNA concentration was 1.5 μg/μl. The labeled RNA was purified on a 15% polyacrylamide, 7 M urea, pH 8.3, slab gel (40 x 20 x 0.2 cm), located by autoradiography and extracted from the gel[13] in the presence of 15 μg of yeast tRNA as a carrier. The RNA was precipitated at -18° overnight after adding 0.1 vol. of 20% potassium acetate, pH 5.0, and 3 vol. of acetonitrile/ethanol (4:1, by vol.). For 5'-terminal analysis, part of the labeled RNA (about 5,000 dpm) was incubated with RNase $T_2$ (0.3 unit/μl) in 10 μl of 50 mM sodium acetate, 1 mM EDTA, pH 4.5, at 38° for 1 hr. The released radioactive terminal nucleotide was identified chromatographically[7]. Partial digestions of the labeled RNA (0.5 μg, about $2 \times 10^5$ dpm) were conducted in a reaction volume of 20 μl. 4.5 μg of yeast tRNA was added as a carrier except in the incubation with RNase $Phy_1$ which contained 1 μg of carrier RNA. Bromophenol blue and xylene cyanol FF (4 μg each) were added as tracking dyes prior to incubation. Digestions with alkali and RNases $T_1$ and $U_2$ were performed similarly as described by Donis-Keller et al[3]. RNase A digestion was performed in 70% dimethylsulfoxide similarly as described by Simoncsits et al[5].

For digestion with RNase $Phy_1$, the reaction mixture contained 7 M urea, 1 mM EDTA, 50 mM sodium citrate, pH 5.0, 40% glycerol, and 68 milliunits of RNase $Phy_1$. Incubations were performed at 38° and 50°, respectively, for 10 min. These denaturing conditions, which are different from those described by us for oligonucleotides[4], were required to cleave the susceptible residues in the helical regions of the tRNA. All digests were kept frozen at -80° until use. Digests (4-6 µl each) were electrophoresed on denaturing polyacrylamide gels under published conditions[3], except that the gels were 0.06 cm thick.

## RESULTS

Products of complete RNase $T_1$ and RNase A digestions. Figures 1A and 1B depict fractionations of radioactive oligonucleotide-3' dialcohols obtained by RNase $T_1$ and A digestions, respectively, and subsequent ($^3$H)-postlabeling. The products in either digest are well resolved except C-A-Um-U-m$^2$G'/U-U-C-A-G' (Fig. 1A) and A-C'/A-m$^5$C' and G-G-D'/m$^2$G-G-U' (Fig. 1B). These tri- and pentanucleotide pairs were further resolved after contact transfer[15] onto a fresh PEI-cellulose sheet by rechromatography in 0.2 and 0.4 M LiCl, respectively. The dinucleotides were analyzed directly. The sequence of smaller oligonucleotides (chain lengths $\leq$ 3) was determined by base composition and terminal analysis. The sequence of most of the larger fragments was established by readout sequence analysis[2,4]. The positions of modified residues in the nucleotide chain were usually determined by base composition analysis in conjunction with 3'- and 5'- terminal analysis.

A-Um-U', C-A-Um-U-m$^2$G', and C-C-A-m$^5$C-G' required additional techniques to place the modified residues. Base analysis[9] of A-Um-U' showed a single A and no pyrimidine nucleoside trialcohol was obtained from the 3'-terminus. Because RNase $T_2$ did not release the 3'-terminus as a nucleoside trialcohol the nucleotide adjacent to the 3'-terminus was ribose-methylated. The 3'-terminus was, however, released by nuclease $P_1$/phosphatase digestion[2] and identified chromatographically as U'. Because the dinucleotide dialcohol, obtained by nuclease $S_1$/phosphatase de-

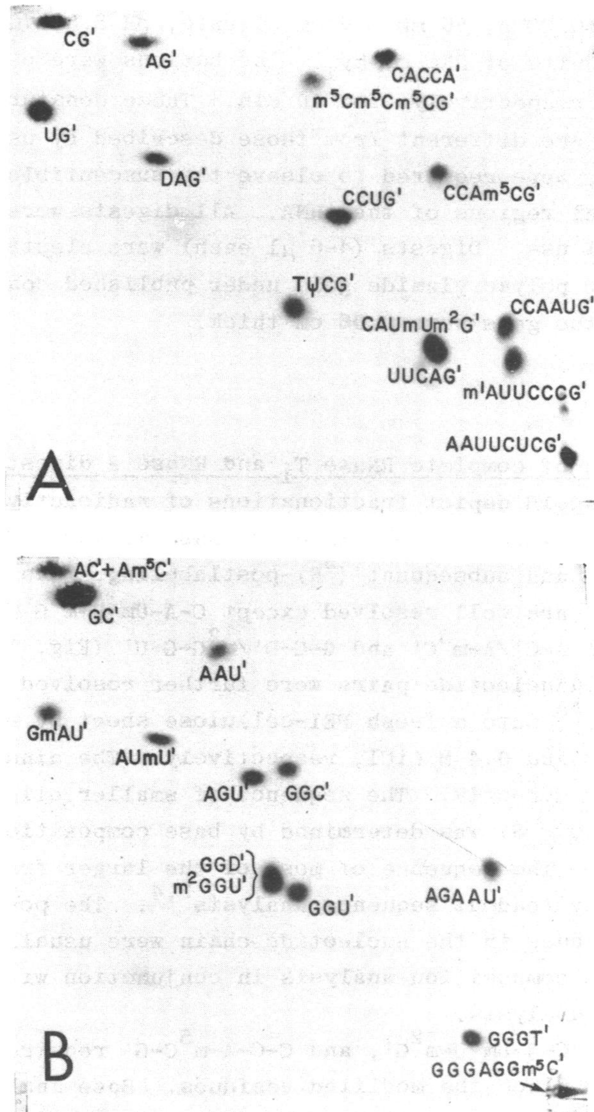


**Fig. 1**

PEI-cellulose map of 3'-($^3$H)-end-labeled oligonucleotide dialcohols in a RNase $T_1$ digest (A) and a RNase A digest (B) of human $tRNA^{Gly}_{GCC}$. First dimension (LiCl gradient), from right to left; second dimension (ammonium formate, pH 2.6, gradient), from bottom to top. Detection by fluorography. CG', UG', etc., oligonucleotide-3' dialcohol derivatives of CpG, UpG, etc. CAUmUm$^2$G'/UUCAG' (A) and GGD'/m$^2$ GGU' (B) were resolved by rechromatography (see text).

gradation[9] of A-Um-U', migrated faster than a marker of U-U' on PEI-cellulose thin layers in ammonium formate, pH 2.6, the ribose-methylated constituent appeared to be Um. Direct evidence for this was obtained from the thin-layer readout sequence analysis of the tRNA (see below).

3'-terminal analysis indicated the presence of 27% C-A-Um-U-G' in the C-A-Um-U-$m^2$G' spot; the ribose-methylated residue was identified as Um because A-Um-U' was obtained by RNase A digestion and ($^3$H)-labeling. The ribose-methylated nucleotide could not be identified by a previously described procedure[18] involving 5'-($^{32}$P)-labeling of the alkali-resistant dinucleotide (Um-U) as no labeling of this particular compound was obtained under the described conditions.

In C-C-A-$m^5$C-G', $m^5$C was placed adjacent to the 3'-terminus because A-$m^5$C' was present in the RNase A digest (Fig. 1B).

Sequence derived by readout techniques. Readout sequencing techniques on PEI-cellulose thin layers and polyacrylamide gels provided the necessary overlaps to derive the sequence of the tRNA.

The thin-layer readout technique enabled us to read the sequence from position 2 to position 73 except for positions 24-29 (counted 3' to 5', see legend of Fig. 2); fragments 24-29 were not sufficiently resolved due to tight secondary structure, resulting in band compression. Figure 2 exemplifies the analysis of major and modified nucleotides between positions 49-73 of the tRNA. All fragments (Fig. 2A) showed distinct single termini except for fragment 69 which gave two termini, $pm^2$Gp and pGp (Fig. 2B), indicating partial methylation at this position (see above). The jump between fragments 69 and 71 (Fig. 2A) is due to the presence of a ribose-methylated residue at position 71. The 5'-terminal nucleotide of fragment 71 was found to be pUm-Up (Fig. 2B), indicating positions 71 and 70 to be Um and U, respectively.

The gel readout technique enabled us to read positions 1-55 and 66-71, except for modified nucleotides. The pyrimidines were distinguished by digestion with RNase $Phy_1$[8], providing a "minus C" pattern[4,5]. Since RNases $Phy_1$ and A produced little or no cleavage between position 42 and 55, the pyrimidines could not be

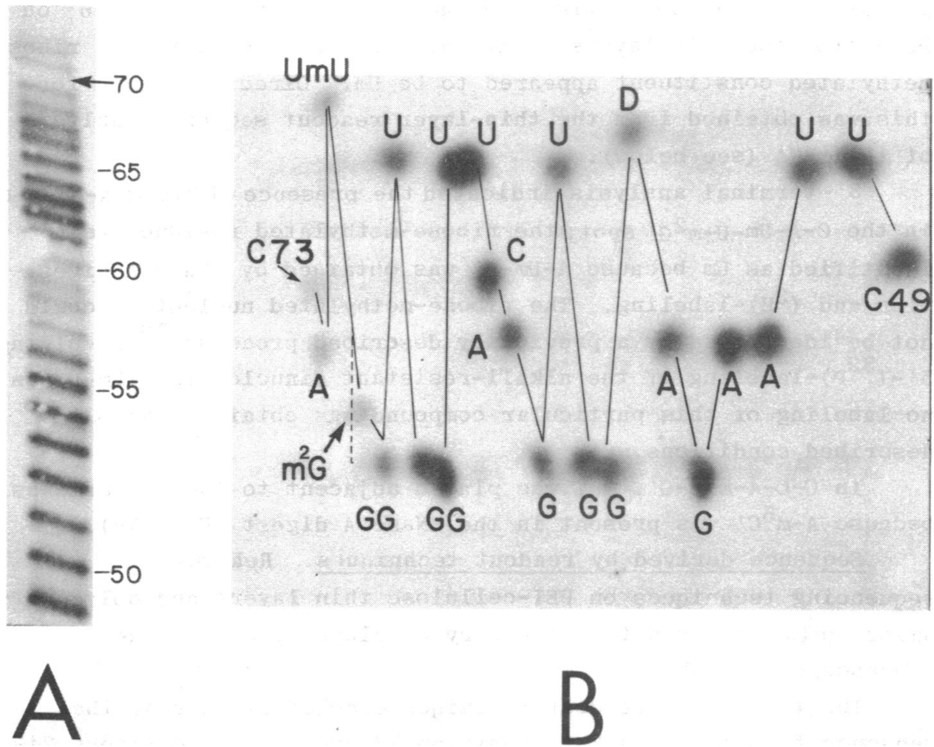


**Fig. 2.**

PEI-cellulose thin-layer readout[7] of positions 49-73 of human $tRNA^{Gly}_{GCC}$. For reasons given elsewhere[7] the positions were numbered 3' to 5'. The sequence shown corresponds to positions 2-26 when counted from the 5'-end of the tRNA (Fig. 3). A, print ladder obtained by brief heating of the RNA in water, 5'-($^{32}P$)-labeling of the fragments, gel electrophoresis, contact transfer of the gel ladder to a PEI-cellulose thin layer, and autoradiography. B, analysis of the 5'-termini of the print ladder by in situ RNase $T_2$ digestion, PEI-cellulose thin-layer chromatography in 0.55 M ammonium sulfate, and autoradiography. U63 and U64 are not clearly resolved; a shorter exposure (not shown) indicated two spots. In addition, the ladder (A) shows 2 bands corresponding to these positions.

identified in this region. This appears to be due to tight secondary structure, an effect also noticed in the RNase $T_1$ track for the same region.

Information obtained by combining the 3 different techniques

enabled us to deduce an unambiguous sequence of the tRNA as shown in Fig. 3.

## DISCUSSION

The sequence of human $tRNA^{Gly}_{GCC}$ was determined by recently developed radioactive derivative procedures[2-5,7] which have been reviewed elsewhere[19,20] and therefore will not be discussed further here.

Like all other eukaryotic glycine tRNAs sequenced to date[21-25],

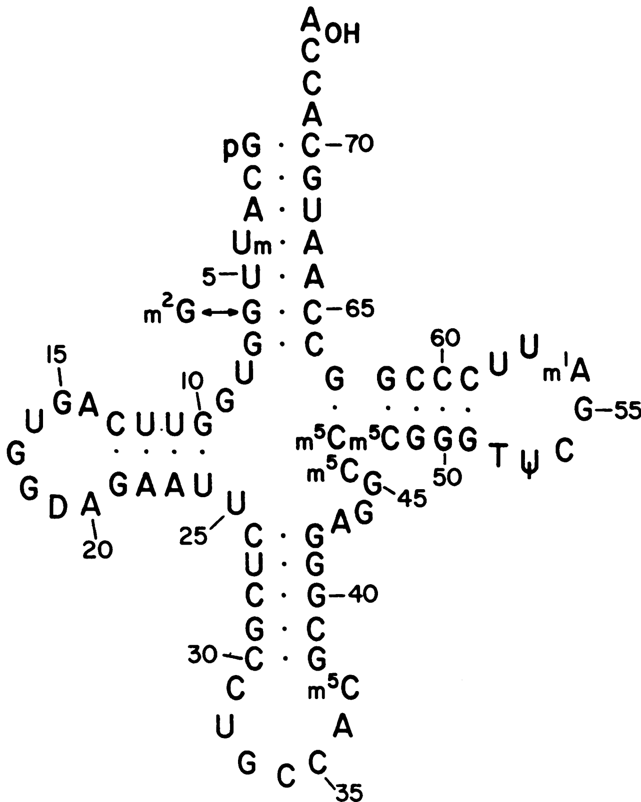


Fig. 3.

The primary structure of human $tRNA^{Gly}_{GCC}$ in cloverleaf form. The structure is homologous to that of silkworm $tRNA^{Gly}_{GCC}$[23,24], except that the latter tRNA contains C5, G6, $m^1G9$, G25, C43 and G66.

the human $tRNA^{Gly}_{GCC}$ has a ribose-methylated pyrimidine nucleoside in position 4. In addition, a partial $N^2$-methylation of G which is not present in glycine tRNAs from other eukaryotes was found in position 6. This is the only tRNA known to have 2 methylated nucleosides in the amino acid acceptor stem. A noteworthy feature shared with other eukaryotic glycine tRNAs[22-24] is the presence of 4 $m^5C$ residues in human $tRNA^{Gly}_{GCC}$.

As pointed out previously[20], all glycine tRNAs from higher eukaryotes sequenced to date exhibit the same TΨ arm sequence. This evolutionary conservation indicates important, as yet undetermined functions of this particular sequence in eukaryotic glycine tRNAs.

In human $tRNA^{Gly}_{GCC}$, U25 appears capable of forming a secondary base pair with A43, a feature shared with all other animal glycine tRNAs sequenced thus far[23-25], although U25 = A43 is replaced by G = C in human $tRNA^{Gly}_{CCC}$ (R. C. Gupta, B. A. Roe and K. Randerath, in preparation) and silkworm $tRNA^{Gly}_{GCC}$[23,24], and by A = U in silkworm $tRNA^{Gly}_{NCC}$ (where N is a derivative of U)[25].

The sequence of human $tRNA^{Gly}_{GCC}$ exhibits a remarkable homology of 94.6% (if differences in posttranscriptional modifications are excluded) with $tRNA^{Gly}_{GCC}$ isolated from the posterior silk gland of <u>Bombyx</u> <u>mori</u> which differs from the human glycine tRNA only by the presence of C5, G25, C43, and G66[23,24]. Such a high degree of homology has not been found for other low molecular weight RNAs (e.g., 5S rRNA) of insects and mammals. Unfortunately, as no sequences of other isoaccepting tRNAs from insects and mammals have been published it is not clear at this time whether the high degree of homology observed for the insect and mammalian glycine tRNAs decoding the same codon is a general phenomenon. Interestingly, human $tRNA^{Gly}_{GCC}$ exhibits only 67.6% homology with human $tRNA^{Gly}_{CCC}$ (R. C. Gupta, B. A. Roe and K. Randerath, in preparation). Similarly, silkworm $tRNA^{Gly}_{GCC}$ exhibits only 75.7% homology with silkworm $tRNA^{Gly}_{NCC}$[25]. Thus, in this particular case, intraspecies variations between isoaccepting animal tRNAs decoding different codons are much larger than differences between tRNAs decoding the same codon from different phyla of the animal kingdom. Further sequence studies on isoaccepting animal tRNAs are required to determine whether the observations regarding the animal

glycine tRNAs apply to other animal tRNA isoacceptors as well.

ACKNOWLEDGEMENTS

This work was supported by the USPHS (grants CA-13591 to K. R. and GM-21405 to B. A. R.; Research Career Development Award K04-00178 to B. A. R.) and the American Cancer Society (grant NP-230 to B. A. R.). Drs. M. P. J. S. Anandaraj, E. Y. Chen, and H. Y. Tsen assisted during the early phase of this work. We thank Dr. J.-P. Bargetzi for a sample of RNase $Phy_1$ and Dr. E. Randerath for discussions and suggestions.

REFERENCES

1. Sprinzl, M., Grüter, F. and Gauss, D. H. (1979) Nucl. Acids Res., 6, r1-r19.
2. Gupta, R. C. and Randerath, K. (1977) Nucl. Acids Res., 4, 1957-1978.
3. Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) Nucl. Acids Res., 4, 2527-2538.
4. Gupta, R. C. and Randerath, K. (1977) Nucl. Acids Res., 4, 3441-3454.
5. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. and Guilley, H. (1977) Nature, 269, 833-836.
6. 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.
7. Gupta, R. C. and Randerath, K. (1979) Nucl. Acids Res., in press.
8. Pilly, D., Niemeyer, A., Schmidt, M. and Bargetzi, J.-P. (1978) J. Biol. Chem., 253, 437-445.
9. Gupta, R. C., Randerath, E. and Randerath, K. (1976) Nucl. Acids Res., 3, 2895-2914.
10. Randerath, K. and Randerath, E. (1967) in Methods in Enzymology, Vol. 12A "Nucleic Acids", pp. 323-347, Academic Press, New York.
11. Anandaraj, M. P. J. S. and Roe, B. A. (1975) Biochemistry, 14, 5068-5073.
12. Roe, B. A., Chen, E. Y. and Tsen, H. Y. (1976) Biochem. Biophys. Res. Commun., 68, 1339-1347.
13. Chia, L. S. Y., Randerath, K. and Randerath, E. (1973) Anal. Biochem., 55, 102-113.
14. Randerath, K. (1970) Anal. Biochem., 34, 188-205.
15. Randerath, K. Randerath, E., Chia, L. S. Y., Gupta, R. C. and Sivarajan, M. (1974) Nucl. Acids Res., 1, 1121-1141.
16. Gupta, R. C., Randerath, E. and Randerath, K. (1976) Nucl. Acids Res., 3, 2915-2921.
17. Silberklang, M., Prochiantz, A., Haenni, A.-L. and RajBhandary, U. L. (1977) Eur. J. Biochem., 72, 465-478.
18. Gupta, R. C., Randerath, K. and Randerath, E. (1976) Anal. Biochem., 76, 269-280.
19. Randerath, E., Gupta, R. C. and Randerath, K. (1979) in Methods in Enzymology, Vol. 65 "Nucleic Acids", in press, Academic Press, New York.

20. Randerath, K., Gupta, R. C., Randerath, E. (1979) in "Transfer RNA: Structure, Properties and Recognition", Cold Spring Harbor Laboratory, Monograph, in press.
21. Yoshida, M. (1973) Biochem. Biophys. Res. Commun., 50, 779-784.
22. Marcu, K. B., Mignery, R. E. and Dudock, B. S. (1977) Biochemistry, 16, 797-806.
23. Garel, J. P. and Keith, G. (1977) Nature, 269, 350-352.
24. Zúñiga, M. C. and Steitz, J. A. (1977) Nucl. Acids Res., 4, 4175-4196.
25. Kawakami, M., Nishio, K. and Takemura, S. (1978) FEBS Lett., 87, 288-290.