Volume 5 Number 2 February 1978 **Nucleic Acids Research**

Isolation from rat liver and sequence of a RNA fragment containing 32 nucleotides from position 5 to 36 from the 3' end of ribosomal 18S RNA

Heidemarie Alberty, Manfred Raba and Hans J. Gross

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, GFR

Received 5 December 1977

ABSTRACT

Crude tRNA isolated from rat liver by the method of Rogg et al. (Biochem. Biophys. Acta 195,13-15 1969) contains N^6-dimethyladenosine (m_2^6A) and was therefore fractionated in order to identify the m_2^6A-containing RNAs. A unique species of RNA was purified which contained all the m_2^6A present in the crude tRNA. Sequence analysis by postlabeling with γ-^{32}P-ATP and polynucleotide kinase revealed that this RNA represents the 32 nucleotides

$$AAGGUUUC\begin{matrix}(C)\\U\end{matrix}GUAGGUGm_2^6Am_2^6ACCUGCGGAAGGAUC$$

from position 5 to 36 of the 3' terminus of ribosomal 18S RNA. The 36 nucleotide long sequence from the 3' end of rat liver 18S rRNA exhibits extensive homology with the corresponding sequence of E. coli 16S rRNA and with the 21 nucleotide long 3' terminal sequence so far known from Saccharomyces carlsbergensis 17S rRNA. A heterogeneity in this sequence provides the first evidence on the molecular level for the existence of (at least) two sets of redundant ribosomal 18S RNA genes in the rat.

INTRODUCTION

Previously we have reported the presence of N^6-dimethyladenosine (m_2^6A) in rat liver 4S RNA, presumably tRNA, and its absence after L-ethionine feeding (1). In an attempt to purify and identify the m_2^6A-containing tRNA(s) we fractionated crude rat liver tRNA and assayed the fractions for m_2^6A. This report shows that a short RNA which copurifies with tRNA is a m_2^6A-containing fragment from the 3'-terminal region of ribosomal 18S RNA.

MATERIALS and METHODS

RNA preparation and fractionation: rat liver tRNA was isolated according to Rogg et al. (2). The crude tRNA was fractionated on DEAE-Sephadex A-50 as described previously for rabbit liver

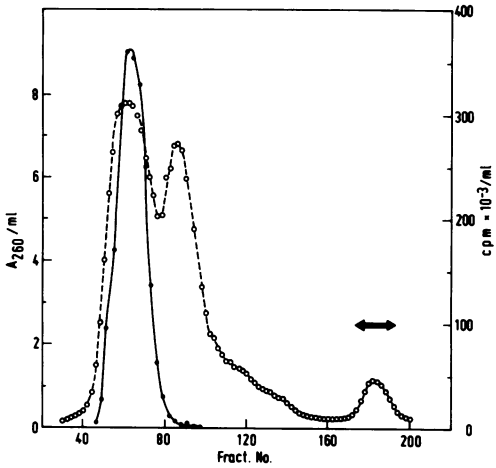

Fig. 1. BD-Cellulose chromatography (conditions as in ref. 3) of 3.560 A_{260} units $tRNA^{Val}_{1}$-fraction (obtained from a DEAE-Sephadex A-50 fractionation of 41.000 A_{260} units of crude rat liver tRNA) yielded 65 A_{260} units of the m^6_2A-containing RNA (arrows) after ethanol precipitation.

tRNA (3). The m^6_2A-containing fractions, coinciding with $tRNA^{Val}_{1}$, were further separated on benzoylated DEAE (BD)-cellulose as shown in Fig. 1.- RNA fractions were analysed by electrophoresis on 12 % polyacrylamide gels (4).- Nucleotide analyses: RNA fractions were digested by RNAase T2 and the mononucleotides separated and identified by twodimensional thinlayer chromatography (1,3).- Postlabeling and fingerprinting of RNA: digestion of the m^6_2A-containing RNA was performed with pancreatic RNAase or RNAase T1 under standard conditions (5,6). All fragments (Fig. 2 and 3) were sequenced, after controlled exonuclease digestion (3), on DEAE paper (7) and independently by the twodimensional mobility shift method, i.e., 5000 V electrophoresis on cellulose acetate, pH 3.5, and homochromatography in the 2^{nd} dimension (8). Large fragments were obtained by treatment with alkaline phosphatase from E. coli (Boehringer Mannheim) which contained a trace of an unknown endonuclease: 0.5 A_{260} units of RNA were incubated for 2 h at 37°C with 10 mU of alkaline phosphatase in 80 µl 60 mM Tris HCl, pH 8.0. The phosphatase was then inactivated by nitrilotriacetic acid treatment (5,6).- Postlabeling of the m^6_2A-containing RNA and its nuclease digests using 5'-polynucleotide kinase from T4-phage infected E. coli and γ-^{32}P-ATP was carried out as described (6). Purification of the 5'-^{32}P-labeled RNA (sequence a, Fig. 5) or its large fragments (b and c, Fig. 5)

was done on DEAE thinlayer plates (1st dimension: high voltage electrophoresis at pH 3.5 on cellulose acetate, 2nd dimension: homochromatography) (6). The ^{32}P-labeled oligonucleotides were eluted after autoradiography and their sequences determined as follows: (a) complete digestion with nuclease P1 (P-L Biochemicals, Inc., ref. 9), resulting in the ^{32}P-labeled 5'-terminal nucleotides; (b) sequences were obtained as described by controlled digestion with snake venom phosphodiesterase (Worthington) or nuclease P1 followed by electrophoresis on DEAE-paper (7) and (c) independently by twodimensional high voltage electrophoresis/homochromatography on DEAE plates (8) as shown in Fig. 4.

RESULTS

The identification of N^6-dimethyladenosine (m_2^6A) in rat liver 4S RNA (1), presumably tRNA, made it necessary to locate this modified nucleoside in purified tRNAs. We felt that this would

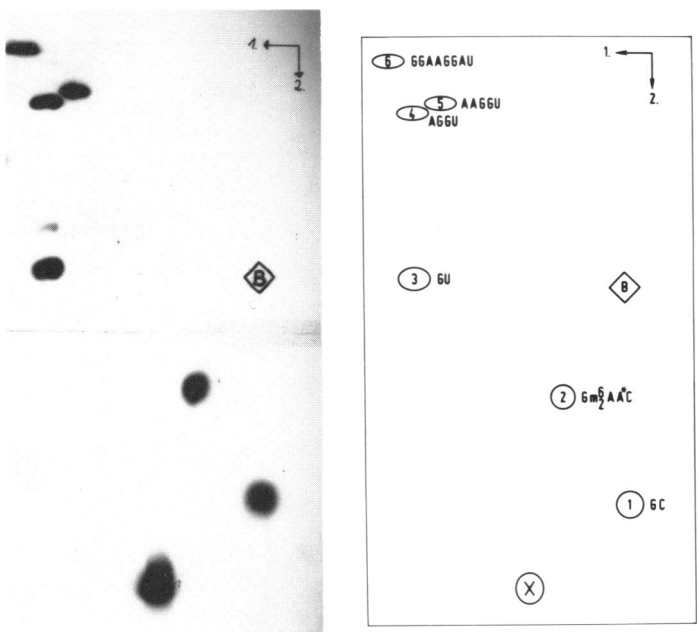

Fig. 2: Polynucleotide kinase fingerprints of a RNAase A digest of the m_2^6A-containing RNA. First dimension: 5000 V, cellulose acetate, pH 3.5, second dimension: DEAE paper electrophoresis in 7 % formic acid. The fragments were obtained in nearly molar ratios. A* = m_2^6A, B = xylene cyanol FF.

be especially important since m_2^6A had not been previously found in an individual tRNA species. We therefore fractionated liver tRNA by column chromatography and assayed the tRNA fractions for m_2^6Ap by RNAase T2 digestion and twodimensional thinlayer chromatography (1). The first fractionation was performed on a DEAE-Sephadex A-50 column. The conditions for this chromatography as well as the resulting tRNA pattern were exactly as described for rabbit liver tRNA (Fig. 1 in ref. 3). The m_2^6A-containing RNA was located under the major peak of valine acceptor activity ($tRNA_1^{Val}$, refs. 3 and 10). These fractions with both $tRNA_1^{Val}$ and the m_2^6A-containing RNA were further fractionated on a column of benzoylated DEAE (BD)-cellulose (Fig. 1). No m_2^6A was detected in the major tRNA peaks containing the valine tRNA fractions. m_2^6A was found exclusively in the RNA which eluted between fractions 173 and 192. Polyacrylamide gel electrophoresis (4) of the m_2^6A-containing RNA revealed that it was less than half the size of tRNA (not shown),

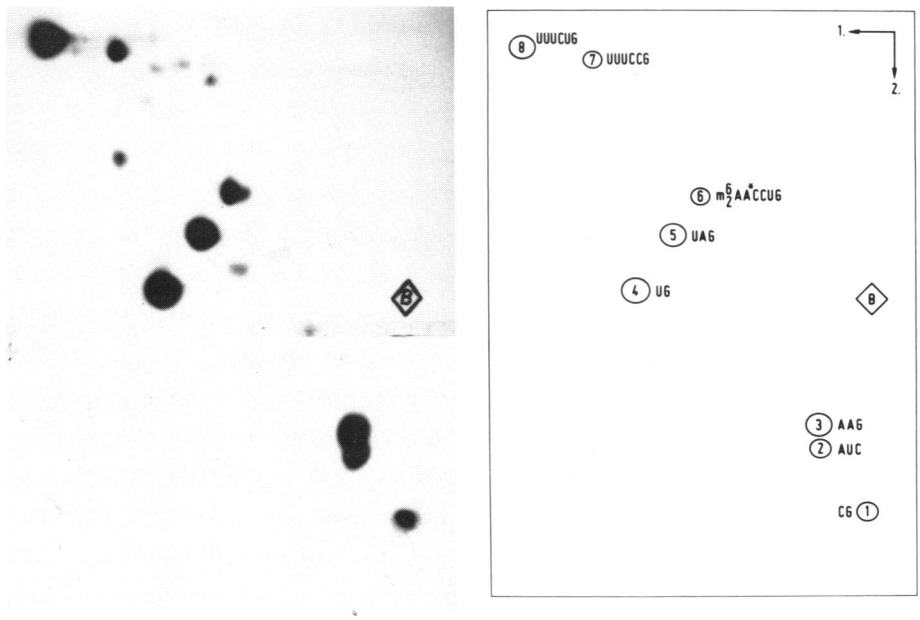

Fig. 3. RNAase T1 fingerprint of the m_2^6A-containing RNA. All fragments were obtained in nearly molar ratios except no. 3 (twofold amount); no. 6 was less efficiently phosphorylated due to the modified 5' terminal nucleoside (m_2^6A); no. 7 (20 % of no. 8) and no. 8 had to be taken together due to sequence heterogeneity (see legend to Fig. 5 and "discussion"). $A^* = m_2^6A$.

and that it was pure enough for sequence analysis by means of in vitro ^{32}P-labeling with 5'-kinase and γ-^{32}P-ATP (3,5,6). Accordingly, m_2^6A was not found any more in crude tRNA which had been purified by preparative polyacrylamide gel electrophoresis.

The RNAase A and RNAase T1 fingerprints of the m_2^6A-containing RNA are shown in Figs. 2 and 3. The RNAase T1 fragment AUC, which could be connected with the pancreatic fragment GGAAGGAU, has to be located at the 3'-end of the RNA due to the absence of G in the 3'-position. An analysis of the labeled 5'terminal nucleotides of the pancreatic and RNAase T1 fragments showed pm_2^6A to be the only modified nucleotide in a 5'position (fragment 6, Fig. 3). After the sequences of the oligonucleotides had been established by controlled exonuclease treatment,

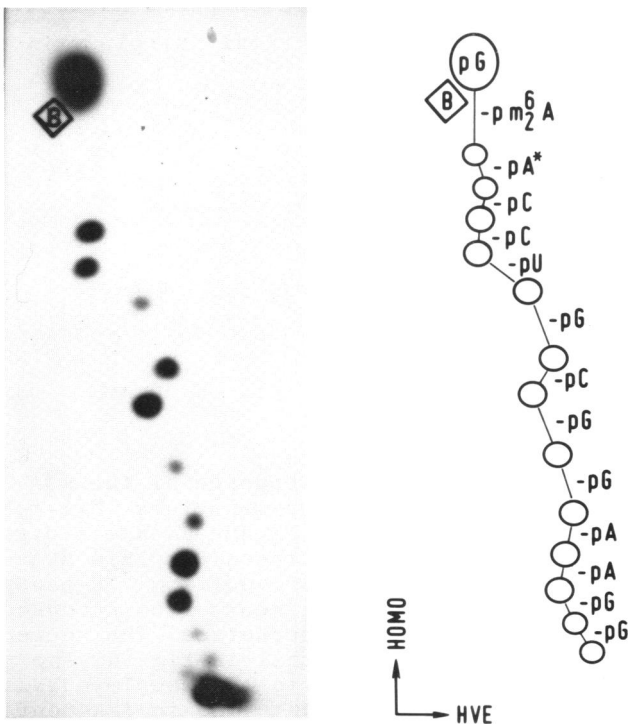

Fig. 4. Sequence analysis of $Gm_2^6AA^*CCUGCGGAAGG$ (fragment c, Fig. 5) obtained from the m_2^6A-containing RNA with an unknown endonuclease from E. coli bacterial alkaline phosphatase. 1st Dimension: 5000 V on cellulose acetate, pH 3.5. 2nd Dimension: homochromatography (6,8). B = Xylene cyanol FF; $A^* = m_2^6A$.

DEAE-paper electrophoresis and autoradiography, overlapping sequences were obtained in two ways:
(a) The whole RNA was phosphorylated at the 5' end and the sequence of the first 12 nucleotides (fragment a, Fig. 5) was determined by the mobility shift method (8); (b) the RNA was treated with bacterial alkaline phosphatase, whereby a trace of an unknown endonuclease cleaved the molecule very specifically. After ^{32}P-phosphorylation, twodimensional fingerprinting (cellulose acetate electrophoresis/homochromatography) (6) and autoradiography (not shown) we isolated, in addition to intact RNA, two fragments (b and c in Fig. 5), the sequences of which were again established by the mobility shift method (8) as shown for fragment c in Fig. 4. The results of these sequence studies allowed the construction of the complete nucleotide sequence. 32 Nucleotides form a symmetrical hairpin with m_2^6A in the loop as shown in Fig. 5.

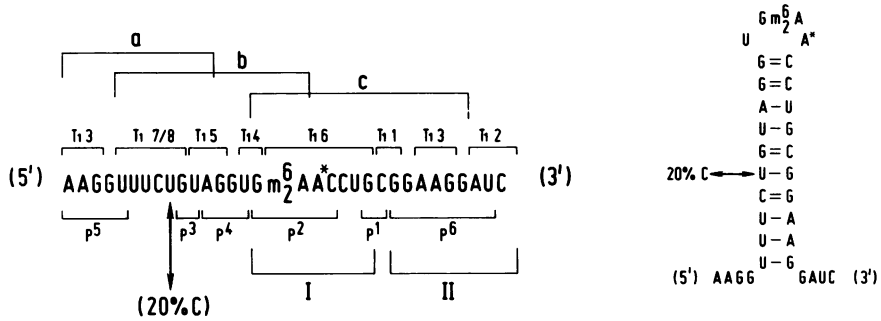

Fig. 5. Derivation of the nucleotide sequence of the m_2^6A-containing RNA, and its possible secondary structure. $T_1$1-$T_1$8; p1-p6: complete fragments from RNAase T1 and RNAase A digestion (Figs. 2 and 3). I and II: overlaps between complete RNAase T1 and RNAase A fragments. a: sequence derived from ^{32}P-phosphorylation of the intact RNA. b and c: fragments from an unknown endonuclease in bacterial alkaline phosphatase. The sequence analysis of fragment c is shown in detail in Fig. 4. The presence of C (arrow) was not only found in the complete RNAase T1 fingerprint (fragment 7, Fig. 3), but also in fragments a and b.

DISCUSSION

After having found that crude rat liver tRNA contains m_2^6A which disappears upon L-ethionine feeding (1), we isolated and sequenced the m_2^6A-containing RNA. This RNA is not at all related

to tRNA, but appears to be a 3'-terminal fragment of ribosomal 18S RNA. The decrease of m_2^6A in crude 4S RNA might hence have two reasons: methylation of A to m_2^6A may be inhibited in liver 18S rRNA, or 18S rRNA is fully methylated inspite of L-ethionine feeding, however, the nuclease which cuts this small fragment out of 18S rRNA may be inhibited by L-ethionine feeding (extremely unlikely). Although we purified 18S rRNA from rat liver ribosomes and identified m_2^6A by the tritium labeling technique of Randerath (11), we were not too successful in applying this method on 18S rRNA from L-ethionine fed rats. Due to the low m_2^6A content in 18S rRNA (2 among ca. 1900 nucleotides), and a background of radioactivity present in the m_2^6A region of the twodimensional thinlayer chromatogram, and possibly because of lack of extensive experience with this method we felt that the m_2^6A decrease observed in 18S rRNA from L-ethionine fed rats was not reproducible and clearcut enough. It therefore remains open whether L-ethionine, in contrast to D-ethionine, causes a reproducible decrease of m_2^6A in 18S rRNA or, unlikely, prevents this fragment from being cut out of 18S rRNA.-

The 32 nucleotide long m_2^6A-containing RNA purified from the crude tRNA fraction of liver tRNA was easily identified as the fragment from position 5 to 36 of the 3' terminus of 18S rRNA for the following reasons:

(a) it contains the universal $m_2^6Am_2^6A$-sequence in the $m_2^6Am_2^6ACCUG$ hexanucleotide (Fig. 3), whereby we only have full evidence for the left m_2^6A, since it is obtained ^{32}P-labeled in this RNAase T1 fragment. The evidence for the second m_2^6A is indirect:

(i) controlled RNAase P1 digestion of fragment c (Fig. 5) shows that the phosphodiester bonds between m_2^6A, A^* and C in the m_2^6A-A^*-C sequence are highly resistant to this nuclease, indicating that A^* is a modified adenosine (most likely m_2^6A);

(ii) the twodimensional nucleotide analysis (1,3) showed only m_2^6Ap in addition to Ap, Cp, Gp and Up (not shown).

(b) this RNA shows extensive homology with the corresponding sequences of *E. coli* 16S (12) and yeast 17S (13) rRNA (Fig. 6).-

It should be mentioned here that until now there is no evidence for a biological significance of this rRNA fragment, which may

well be an artefact, i.e., caused by nuclease degradation during the process of tRNA preparation (2). This rRNA fragment is certainly not related to ribosomal "2S" RNA isolated by Jordan (14) from cultured Drosophila cells, since the fingerprints are completely different; also, a relation to a "translational control RNA" (tcRNA) observed in ribosomal wash fractions seems unlikely (15, 16).-

A comparative presentation of E. coli, yeast and rat 3'-terminal 16S/17S/18S rRNA sequences reveals several features of interest (Fig. 6):

(a) The four 3'-terminal nucleotides of our 32 nucleotide long RNA (position 5 to 36) overlap with the octanucleotide $GAUCAUUA_{OH}$ known to be the 3'-terminal sequence of yeast, Drosophila, rabbit (17, 18) and possibly all eukaryotic 17S/18S rRNAs. This universal "eukaryotic" octanucleotide, the first six nucleotides of which earlier had been thought to be "terminator anticodons" (17, 18), contains four to six nucleotides complementary to the initiator regions of several eukaryotic mRNAs including Brome Mosaic Virus RNA (19) SV40-VP_1 RNA (20) and rabbit ß-globin mRNA (21, 22). This led to the proposal that eukaryotic initiation of protein synthesis may involve base-pairing between the initiator sequence of mRNA and the 3' terminus of the small ribosomal unit's 17S/18S RNA (19-22) as earlier shown for E. coli (17, 18, 23-26).

(b) All 16S/17S/18S rRNAs seem to contain the "universal" sequence of unknown function $m_2^6Am_2^6A$ in a loop near the 3' end.

(c) The extensive homology between all three sequences suggests the possibility of a universal function of at least the 36 (E. coli: 41) nucleotides from the 3' end of eukaryotic (prokaryotic) 17S/18S (16S) rRNAs.

(d) Heterogeneity of the ribosomal genes in mammalia is well documented (summarized in refs. 27 and 28) and was located in nontranscribed spacer regions. The presence of 20 % C in place of U_{28} in rat 18S RNA provides the very first evidence, on the molecular level, for the existence of heterogeneity in transcribed regions, i.e., for (at least) two sets of redundant 18S RNA genes in the rat.

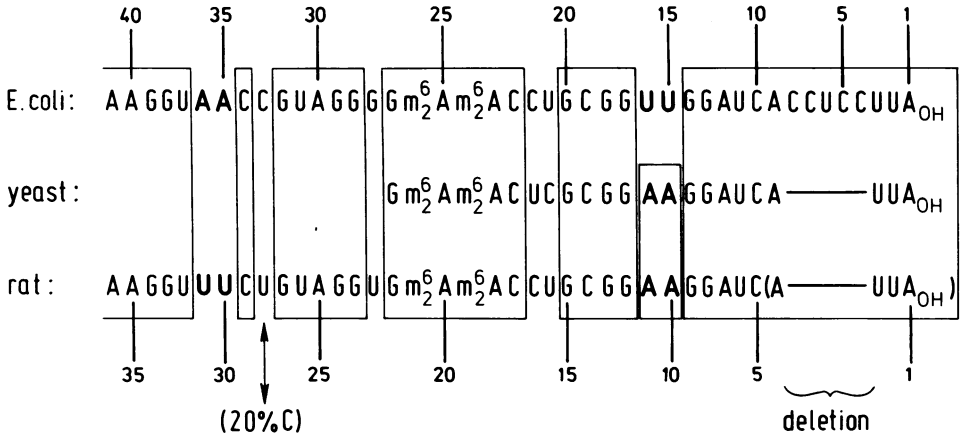

Fig. 6. Sequences of 5' ends of E. coli 16S (12), yeast 17S (13) and rat 18S rRNAs. Identical nucleotides or sequences are boxed. A deletion of 5 nucleotides and two pairs of coupled transversions (darker letters) near the 3' termini seem to reflect the evolution from prokaryotes to eukaryotes. The rRNA fragment isolated and sequenced in this work is represented by nucleotides no. 5 to 36; the nucleotides no. 1 to 4 are shown in brackets since they are derived from an overlap with the known 3' terminal octanucleotide GAUCAUUA$_{OH}$ (17, 18).

(e) The differences between E. coli and rat 18S rRNA as shown in Fig. 6 deserve special attention. There is a transversion mutation from E. coli G_{27} to rat U_{22}, however more striking, two pairs of transversions, E. coli A_{35}-A_{36} to rat U_{30}-U_{31}, and E. coli U_{15}-U_{16} to rat A_{10}-A_{11}, in symmetrical position to the phosphate linking E. coli $m_2^6A_{25}$ and G_{26}, and rat $m_2^6A_{20}$ and G_{21}, respectively. The interesting point is that these two pairs of transversions must have been coupled during evolution, possibly in order to maintain two (A:U) basepairs near the basis of the hairpin structure (Fig. 5) Since two A:U pairs in this position do not significantly contribute to the thermodynamic stability of the m_2^6A stem (Fig. 5), their maintenance during evolution might be related to a universal function of the m_2^6A stem and loop, e.g., processing of rRNA, initiation of protein synthesis, or some other purpose.

REFERENCES

1 Wildenauer, D. and Gross, H.J. (1974) Nucleic Acids Res. 1, 279-288
2 Rogg, H., Wehrli, W. and Staehelin, M. (1969) Biochim. Biophys. Acta 195, 13-15
3 Jank, P., Shindo-Okada, N., Nishimura, S. and Gross, H.J. (1977) Nucleic Acids Res. 4, 1999-2008
4 Philipps, G.R. (1971) Analyt. Biochem. 44, 345-357
5 Simsek, M., Ziegenmeyer, J. Heckman, J. and RajBhandary, U.L. (1973) Proc. Natl. Acad. Sci. (USA) 70, 1041-1045
6 Gross, H.J., Domdey, H. and Sänger, H.L. (1977) Nucleic Acids Res. 4, 2021-2028
7 Sanger, F., Brownlee, G.G. and Barrell, B.G. (1965) J. Mol. Biol. 13, 373-398
8 Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) Nucleic Acids Res. 1, 331-353
9 Silberklang, M., Prochiantz, A., Haenni, A.-L. and RajBhandary, U.L. (1977) J. Eur. Biochem. 72, 465-478
10 Mayr, U., Bermayer, H.-P., Weidinger, G., Jungwirth, Ch., Gross, H.J. and Bodo, G. (1977) Eur. J. Biochem. 76, 541-551
11 Randerath, E., Yu, C.-T. and Randerath, K. (1972) Analyt. Biochem. 48, 172-198
12 Ehresmann, C., Stiegler, P., Mackie, G.A., Zimmermann, R.A., Ebel, J.P. and Fellner, P. (1975) Nucleic Acids Res. 2, 265-278
13 Jonge, P. de, Klootwijk, J. and Planta, R.J. (1977) Nucleic Acids Res. 4, 3655-3663
14 Jordan, B.R. (1974) FEBS Lett. 44, 39-42
15 Raymondjean, M., Bogdanovsky, D., Bachner, L., Kneip, B. and Schapira, G. (1977) FEBS Lett. 76, 311-315
16 Heywood, S.M., Kennedy, D.S. and Bester, A.J. (1974) Proc. Natl. Acad. Sci. USA 71, 2428-2431
17 Dalgarno, L. and Shine, J. (1973) Nature New Biology 245, 261-262
18 Shine, J. and Dalgarno, L. (1974) Biochem. J. 141, 609-615
19 Dasgupta, R., Shih, D.S., Saris, C. and Kaesberg, P. (1975) Nature 256, 624-628
20 Voorde, A. van de, Contreras, R., Rogiers, R. and Fiers, W. (1976) Cell 9, 117-120
21 Legon, S. (1976) J. Mol. Biol. 106, 37-53
22 Baralle, F.E. (1977) Cell 10, 549-558
23 Dahlberg, A.E. and Dahlberg, J.E. (1975) Proc. Natl. Acad. Sci. USA 72, 2940-2944
24 Steitz, J.A. and Jakes, K. (1975) Proc. Natl. Acad. Sci. USA 72, 4734-4738
25 Steitz, J.A. and Steege, D.A. (1977) J. Mol. Biol. 114 545-558
26 Steege, D.A. (1977) Proc. Natl. Acad. Sci. USA 74, 4163-4167
27 Arnheim, N. and Southern, E.M. (1977) Cell 11, 363-370
28 Cory, S. and Adams, J.M. (1977) Cell 11, 795-805.