NUCLEOTIDE SEQUENCE OF YEAST 5 S RIBOSOMAL RNA

J. HINDLEY and S.M. PAGE

Molecular Enzymology Laboratory, Department of Biochemistry, University of Bristol, Bristol BS8 1TD, England

Received 6 July 1972

1. Introduction

The low molecular weight 5 S ribosomal RNA has been found in the larger ribosomal subunit of all species investigated. While its function in ribosome structure is obscure, it would be anticipated that it, at least, fulfills a similar role in the structure and function of the larger subunit of all different species. The primary sequence of the *E. coli* [1], *P. fluorescens* [2] and KB-cell [3] 5 S RNA's have been determined. Physico-chemical studies [4, 5] have given estimates of 60–80% base pairing in *E. coli* 5 S RNA and the variable susceptibility of particular regions of the molecule to digestion with ribonuclease T_1 [6, 7], modification by carbodiimide [8] and oligonucleotide binding [9] all lend support to the idea of a particular, though as yet undefined, conformation. Knowledge of the sequence of a third unrelated 5 S RNA species might be expected to guide structural analysis if a similar conformation for the different 5 S species is assumed.

This communication reports the primary sequence of *Saccharomyces carlsbergensis* 5 S RNA. A novel feature of this RNA, which distinguishes it from the other 5 S RNA's studied, is the occurrence of a pppG residue at the 5′-terminus. Unless the existence of kinases in the cell, which subsequently phosphorylate the 5′-terminus is assumed, the presence of a 5′-terminal triphosphate suggests the 5 S RNA to be a primary gene product rather than a post-transcriptional product derived from precursor RNA with an extended sequence at the 5′-end.

2. Materials and methods

The yeast was grown in a labelled medium containing [^{32}P]phosphate, as described previously [10, 11]. After harvesting, RNA was extracted by resuspending the cells in 0.05 M Tris-HCl, pH 7.0, 0.01 M EDTA, 2% SDS and shaking with phenol previously equilibrated with 0.05 M Tris-HCl, pH 7.0. Under these conditions the yeast cell wall remains impermeable to high molecular weight RNA, essentially only 5 S RNA and tRNA being released into the medium. 5 S RNA was separated from tRNA and traces of other higher molecular weight species by electrophoresis through a polyacrylamide gel [12]. The position of the 5 S band was located by autoradiography, excised, and the RNA extracted in the presence of 50–100 μg unlabelled carrier RNA. After precipitation, the final product (including the carrier RNA) had a specific activity of about 4×10^6 cpm per μg RNA.

Complete ribonuclease T_1 and pancreatic ribonuclease digests were prepared and the products separated and sequenced by established methods [10, 1, 13]. Partial ribonuclease T_1 digests were prepared using enzyme to substrate ratios between 1:200 and 1:1,500 (w/w) in 0.02 M Tris-HCl, pH 7.2, 0.02 M $MgCl_2$ and digesting for ½–1 hr at 0°. The partial products were separated as described previously [14]. Between 15 and 40 well resolved spots were usually selected for further study. The corresponding oligonucleotides were eluted with 30% triethylamine carbonate, pH 10, and identified by quantitative analysis of their ribonuclease T_1 and pancreatic ribonuclease digestion products [14].

Table 1
Relative molar yields of oligonucleotides from digests of yeast 5 S RNA.

I. Ribonuclease T_1 derived oligonucleotides.

Spot No.	Sequence	Average molar yield* (4 experiments)	Corrected molar yield†
1	G	6.62	6
2	CG	2.44	2
3	AG	2.59	2
4	ACCG	1.04	1
5	CACCG	0.93	1
6	$CAAUCU_{OH}$	1.00	1
7	AAAG	0.93	1
8	UG	4.73	4
9	CUG	2.28	2
10	UAG	3.17	3
11	UCCG	2.22	1
12	CCUG		1
13	UAAG	1.01	1
14	ACCAUACG	1.72	1
15	AAACCUAG		1
16	UUG	1.06	1
17	UUAAG	0.93	1
18	AUAACCUG	0.88	1
19	UUCUCCG	1.02	1
20	CCAUACCAUCUAG	0.60	1
21	ppG	0.51	1
22	pppG	0.34	
		Total	121

* Expressed relative to average of UAAG, UUG and UUAAG taken as 1.00.

II. Pancreatic ribonuclease derived oligonucleotides.

Spot No.	Sequence	Average molar yield** (4 experiments)	Corrected molar yield†
1	U	9.86	10
2	C	13.06	15
3	AC	3.09	3
4	AAC	1.06	1
5	GC	4.00	4
6	AU	3.17	3
7	GAC	2.14	2
8	AAGC	1.11	1
9	AAU	1.03	1
10	GAAC	0.80	1
11	GU	3.95	4
12	GGC	0.97	1
13	GAU	0.99	1
14	AGU	2.99	3
15	AAGAGC	0.66	1
16	AGAAAGC	0.62	1
17	GGU	0.97	1
18	GAGU	1.82	1
19	AGGU		1
20	GGGU	0.90	1
21	pppGGU + ppGGU	0.79	1
		Total	120 ($+U_{OH}$ = 121)

** Expressed relative to average of AAU, GGC and GGU taken as 1.00.

† The corrected molar yields of free U and C were deduced using the known sequence and yields of the T_1-oligonucleotides. The number of free G's was deduced from the number of GG sequences in the pancreatic oligonucleotides.

3. Results and discussion

Table 1 lists the sequences and molar ratios of the oligonucleotides from the different nuclease digests. Apart from the heterogeneity observed in the 5′-terminal oligonucleotide, discussed below, a total of 21 different pancreatic oligonucleotides and 20 different T_1 oligonucleotides were obtained. With the exception of the T_1 oligonucleotides nos. 11, 12 and 14, 15, (table 1) and the pancreatic oligonucleotides nos. 18, 19, the resolution on the fingerprints [15] was adequate for the separate quantitation and subsequent analysis of all other fragments. Longer electrophoretic runs in the second dimension, in which the free mononucleotides were run off the end of the sheet, satisfactorily resolved these slower moving products.

(p)ppGGUUGCGGCCAUACCAUCUAGAAAGCACCGUUCUCCGUCC-

-GAUAACCUGUAGUUAAGCUGGUAAGAGCCUGACCGAGUAG-

-UGUAGUGGGUGACCAUACGCGAAACCUAGGUGCUGCAAUCU$_{OH}$

Fig. 1. Nucleotide sequence of yeast 5 S RNA. The wavy line indicates a region where the sequence is uncertain.

3.1. *Characterization of 3′-terminal oligonucleotide*

Spot no. 6 from the T_1 digest (table 1) gave the composition $C_2A_2U_1$ upon alkaline hydrolysis. The absence of a G residue suggested this product might arise from the 3′-end of the molecule. Treatment with alkaline phosphatase failed to release any free phosphate and the mobility of the spot was unchanged on fingerprinting, suggesting the absence of a 3′-phosphorylated end [16]. Complete digestion with snake venom phosphodiesterase gave the composition $C_1A_2U_2$. The 3′-terminal residue is therefore ...pU_{OH} and the 5′-nucleotide is Cp... . From analysis of pancreatic RNAase digestion products before and after modification with CMCT* [1] the sequence $CAAUCU_{OH}$ was deduced.

3.2. *Characterization of 5′-terminal residue*

An alkaline hydrolysate of 5 S RNA (0.3 N KOH, 18 hr at 37°) was neutralized with perchloric acid and the supernatant fractionated by two-dimensional chromatography on PEI thin layers [17]. In addition to the major mononucleotides two minor components were found corresponding to ppGp and pppGp. Quantitatively 0.62–0.75 moles of ppGp + pppGp were found per mole of 5 S RNA. Under identical conditions, a hydrolysate of uniformly labelled Qβ RNA, which possesses a 5′-terminal pppG group [18] gave 0.69 mole ppGp + pppGp per mole of RNA. Pancreatic RNAase digestion of the 5 S RNA yielded a component which, on fingerprinting, moved very rapidly in the first dimension but remained at the origin in the second dimension. Alkaline hydrolysis gave Gp and Up in equimolar amounts together with a mixture of ppGp and pppGp. Ribonuclease T_1 digests yielded two anomalous components which, after elution, chromatographed with the mobilities of ppGp and pppGp [19]. In none of these experiments was any pGp detected.

The heterogeneity of the 5′-terminal group could be real or apparent; the sequence pppGGU... may have suffered partial loss of the γ-phosphate either *in vivo* or during subsequent analysis. The observation that, under the same conditions, Qβ RNA also yields both ppGp and pppGp suggests that the formation of the former molecule may be an artifact of the isolation or work-up procedure.

* Abbreviation used: CMCT, *N*-cyclohexyl-*N*′-(β-morpholinyl-(4)-ethyl) carbodiimide-methyl-*p*-toluene sulphonate.

3.3. *Primary sequence of 5 S RNA*

A total of 126 partial ribonuclease T_1 digestion products were analysed varying from 4 to 41 nucleotides in length. This data permitted a unique sequence of 121 residues to be deduced for the 5 S RNA molecule (fig. 1). As with the other 5 S RNA's sequenced, an extended base paired structure (9 base pairs) can be proposed for the 5′- and 3′-terminal regions [1–3]. Comparison of this sequence with the *E. coli* and KB-cell 5 S RNA sequences show several regions of homology, marked similarities being observed to the KB-cell sequence. The implications of this as regards models for their secondary structure will be discussed elsewhere. The sequence GAAC, found in both *E. coli* and KB-cell 5 S RNA and considered as a candidate for recognition of the GTψC sequence in tRNA, is not found in the yeast sequence. As pointed out previously [1], however, there are likely to be many other GAAC sequences in the larger ribosomal RNA.

Acknowledgement

This work was supported in part by S.R.C. grant B/SR/3230 to J.H.

References

[1] G.G. Brownlee, F. Sanger and B.G. Barrell, J. Mol. Biol. 34 (1968) 379.
[2] B. DuBuy and S.M. Weissman, J. Biol. Chem. 246 (1971) 747.
[3] B.G. Forget and S.M. Weissman, Science 158 (1967) 1695.
[4] C.R. Cantor, Proc. Natl. Acad. Sci. U.S. 59 (1968) 478.
[5] H. Boedtker and D.G. Kelling, Biochem. Biophys. Res. Commun. 29 (1967) 758.
[6] R. Vigne and B.R. Jordan, Biochimie 53 (1971) 981.
[7] B.R. Jordan, J. Mol. Biol. 55 (1971) 423.
[8] J.C. Lee and V.M. Ingram, J. Mol. Biol. 41 (1969) 431.
[9] J.B. Lewis and P. Doty, Nature 225 (1970) 5110.
[10] F. Sanger, G.G. Brownlee and B.G. Barrell, J. Mol. Biol. 13 (1965) 373.
[11] J. Hindley, Chromatographic and Electrophoretic Techniques, Vol. I (Wm. Heinemann, 1969) pp. 979–982.
[12] H. Weber, M.A. Billeter, S. Kahane, C. Weissmann, J. Hindley and A. Porter, Nature New Biology 237 (1972) 166.
[13] G.G. Brownlee and F. Sanger, J. Mol. Biol. 23 (1967) 337.

[14] G.G. Brownlee and F. Sanger, European J. Biochem. 11 (1969) 395.
[15] J. Hindley, J. Mol. Biol. 30 (1967) 125.
[16] D.M. Kettlewell, Biochem. J. 106 (1968) 11P.
[17] K. Randerath and E. Randerath, Methods in Enzymology, Vol. XII, Part A (Acad. Press, 1967) pp. 330–343.
[18] J.E. Dahlberg, Nature 220 (1968) 548.
[19] E. Messens and M. Van Montaguqe, FEBS Letters 1 (1968) 326.