CONFORMATION OF MAMMALIAN 5.8 S RIBOSOMAL RNA: S_1 NUCLEASE AS A PROBE

M. S. N. KHAN and B. E. H. MADEN
Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, Scotland

Received 14 October 1976

1. Introduction

5.8 S ribosomal RNA is a small RNA species which is characteristic of eukaryotic ribosomes and occurs noncovalently attached to mature 28 S RNA [1]. The nucleotide sequences of yeast and rat hepatoma 5.8 S RNAs are known, and secondary structure models have been proposed on the basis of maximized base-pairing [2,3]. The HeLa cell 5.8 S RNA primary structure was reported to be indistinguishable from that of rat hepatoma, except in the presence of an extra nucleotide at the 3′-end of HeLa 5.8 S RNA [4,5 and this work]. The sequence is shown in fig.1. Here we report on the use of S_1 nuclease as a conformational probe for HeLa cell 5.8 S RNA.

Nuclease S_1 from *Aspergillus oryzae* hydrolyzes single-stranded nucleic acids [6], or single-stranded regions within nucleic acids [7]. At suitable ionic strength and temperature it cleaves tRNA specifically at the anticodon loop and the 3′-end [8]. Under similar conditions to those used for tRNA we find that two internal sequences and the extreme 3′-end of 5.8 S RNA are susceptible to S_1 nuclease.

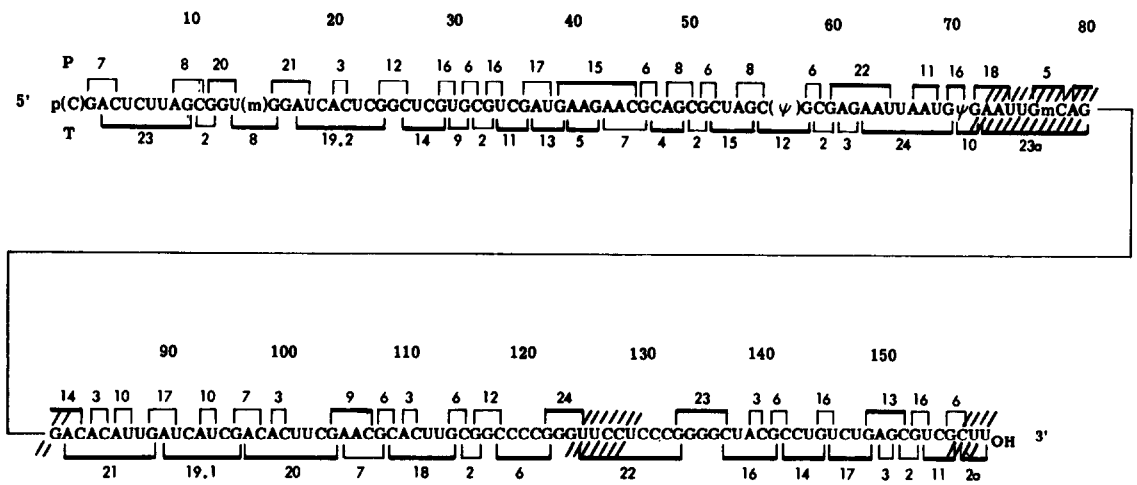

Fig.1. Complete nucleotide sequence of HeLa cell 5.8 S RNA (based on ref. [5]). The sequences of all products of complete pancreatic and T_1 ribonuclease digestion were independently determined in this work and in ref. [4]. Distinctive pancreatic (P) and T_1 (T) digestion products are indicated with heavy bars. Our numbering system differs from that in ref. [5] but for T_1 products is the same as previously used by us [4]. Fractionally chemically modified nucleotides, Um 14 and ψ 57, are in parentheses. Shaded regions represent the sequences susceptible to S_1 nuclease (see table 1).

2. Methods

^{32}P-Labelled 5.8 S RNA was prepared as described [4], and repurified on a second sucrose gradient. Only unlabelled 5.8 S RNA was used as carrier, to prevent protective interactions with other nucleotide sequences. In order to remove sodium dodecyl sulphate and to ensure the presence of sufficient Zn^{2+} for S_1 nuclease digestion, the RNA was reprecipitated at least twice with ethanol from the following 'S_1 digestion buffer': 0.3 M NaCl, 0.03 M sodium acetate, pH 4.6, 0.001 M $ZnSO_4$, 5% glycerol. In one set of experiments the RNA was put through a 'renaturation' protocol [9] followed by recovery of 5.8 S RNA as a unique peak from a sucrose gradient. This was found not to effect the results of subsequent S_1 nuclease digestion.

S_1 Digestion was carried out at 20°C for 90 min using one unit of fresh enzyme (Sigma) per microgram of RNA in 'S_1 digestion buffer'. This was followed by phenol extraction at 20°C, except in fig.2(a). Sufficient carrier RNA was added at this step to ensure quantitative recovery of S_1 digest from the aqueous phase in subsequent ethanol precipitation. Complete digestion with pancreatic or T_1 ribonuclease and fingerprinting analysis were carried out by standard methods [10].

3. Results and discussion

Figure 2(a) shows the two-dimensional electrophoretic separation of 5.8 S RNA digested with S_1 nuclease. Some 10% of the material, determined by scintillation counting, was released as mononucleotides together with the 2′-*O*-methylated dinucleotide, pGm–C. The latter must derive from nucleotides 77–78 (fig.1). The 2′-*O*-methyl linkage was resistant to hydrolysis, as noted previously for tRNA [8]. Neither of the pseudodouridines in the sequences was liberated, as indicated by the negative result of a chromatographic analysis [4] of 'pU' for the presence of pψ. The S_1 resistant material migrated rapidly in the first-dimension of the fingerprint (fig.2(a)) but was bound firmly at the origin of the second. This mobility is characteristic of large oligonucleotides which are rich in U and/or G. No short oligonucleotides were released by S_1 nuclease. These results imply that S_1 nuclease hydrolysis 5.8 S RNA only at a few specific sites, which are relatively widely spaced along the primary sequence.

To define the cleavage sites, the large oligonucleotides which were produced by S_1 digestion were recovered, without further fractionation, by phenol extraction and ethanol precipitation. Liberated mononucleotides and Gm–C were not precipitated by ethanol. The ethanol precipitatable material was then digested with pancreatic or T_1 ribonucleases and fingerprints were prepared. The results are shown, together with those for control 5.8 S RNA (no S_1 digestion) in fig.2 (c–h) and in table 1. Interpretation of the data was facilitated by the fact that essentially the whole 5.8 S sequence is spanned by unique pancreatic or T_1 ribonuclease oligonucleotides (fig.1). The molar recoveries of most of these were largely unchanged after S_1 nuclease digestion. Those that were recovered in low yield permitted identification of three S_1 nuclease cleavage sites within 5.8 S RNA (fig.1 and 3).

3.1. *Nucleotides 73–82*

The pancreatic RNAase products, P18 (G–A–A–Up), P5 (Gm–Cp) and P14 (A–G–G–A–Cp), as well as the T_1 product T23a (A–A–U–U–Gm–C–A–Gp) were all recovered in low yield after S_1 nuclease digestion (table 1, and dotted lines in fig.3). T23a is not well resolved from T23 in the fingerprints, but its disappearance after S_1 nuclease digestion was

Fig.2. Fingerprints of HeLa cell 5.8 S RNA before and after S_1 nuclease treatment. (a) represents S_1 nuclease digest fingerprinted directly and (b) is the key. (c), (d) and (e) are pancreatic RNAase fingerprints of 5.8 S RNA. (c) Control, (d) after S_1 nuclease, (e) key. (f), (g) and (h) are T_1 fingerprints of 5.8 S RNA. (f) Control, (g) after S_1, (h) key. Products which were present in markedly diminished yield after S_1 digestion are marked with arrows and are underlined in keys. Pancreatic ribonuclease products 26 and 27 appear occasionally as minor spots due to slightly incomplete digestion. The fingerprints were run under standard conditions [10]. The first-dimension (cellulose acetate) is from right to left and the second-dimension (DEAE-paper) from top to bottom.

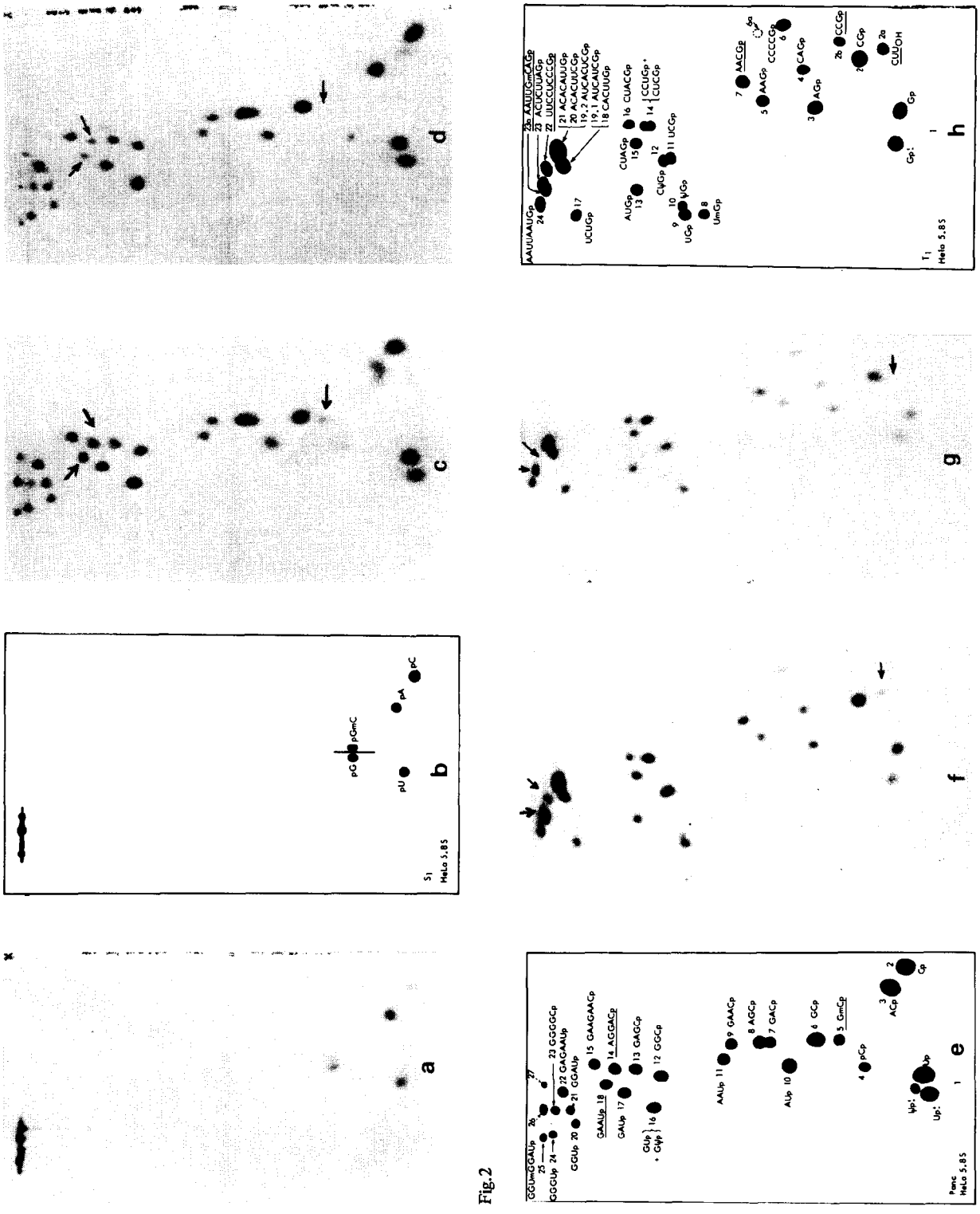
a
b
pG
pGmC
pU
pA
pC
S1
HeLa 5.8S
c
d
e
GGUmGGAUp
GGGUp
GGUp
GAAUp
GAUp
GUp
+ CUp
AAUp
AUp
GGGAAUp
GAGAAUp
GGAUp
GAAGAACp
AGGACp
GAGCp
GGCp
GAACp
AGCp
GACp
GCp
GmCp
pCp
ACp
Cp
Up
Ψp
Ap
Panc
HeLa 5.8S
f
g
h
AAUUAAUGp
AUCUCUUCCCCGp
ACACAUUGp
ACACUUCGp
AUCACUCGp
AUCAUCGp
CACUUGp
UCUGp
AUGp
CUACGp
CUCGp
UCGp
ΨGp
UmGp
AACGp
AAGp
CCCCGp
CAGp
AGp
CCGp
CCp
CUUOH
Gp
Tl
HeLa 5.8S

Fig.2

Table 1
HeLa 5.8 S RNA susceptibility to S_1 nuclease: pancreatic (P) and T_1 RNAase (T) products

Spot No.	Sequence	Molar yield		
		Control		After S_1
S_1 nuclease susceptible sequences				
P5	Gm–Cp	0.92	(1)	0.17
P14	A–G–G–A–Cp	0.97	(1)	0.42
P18	G–A–A–Up	1.11	(1)	0.46
T2a	C–U–U_{OH}	0.78	(1)	0.21
T22	U–U–C–C–U–C–C–C–Gp	0.65	(~1)	0.15
T23a	A–A–U–U–Gm–C–A–Gp	1.02	(1)	0.00
Important protected features				
P9	G–A–A–Cp	0.92	(1)	1.00
P15	G–A–A–G–A–A–Cp	0.86	(1)	0.85
T8	Um–Gp	0.19	(0.2)	0.19
T11	U–C–Gp	2.95	(2)	2.48
T12	C–ψ–Gp		(~1)	
T15	C–U–A–Gp	1.15	(1)	0.99
T19.2	A–U–C–A–C–U–C–Gp	0.98	(1)	0.92
T20	A–C–A–C–U–U–C–Gp	0.98	(1)	0.92

Molar yields of complete pancreatic and T_1 RNAase digestion products of 5.8 S RNA, control and after S_1, are compared (see also fig.2(c–h)). Control pancreatic values are the means of four determinations from fingerprints such as shown in fig.2(c). Control T_1 values are the means of three standard runs (as shown in fig.2(f)) and two determinations with long separations in both dimensions to improve the resolution of products 18–24. S_1 digest values are the means of two pancreatic fingerprints, two standard T_1 fingerprints and two long T_1 fingerprints.

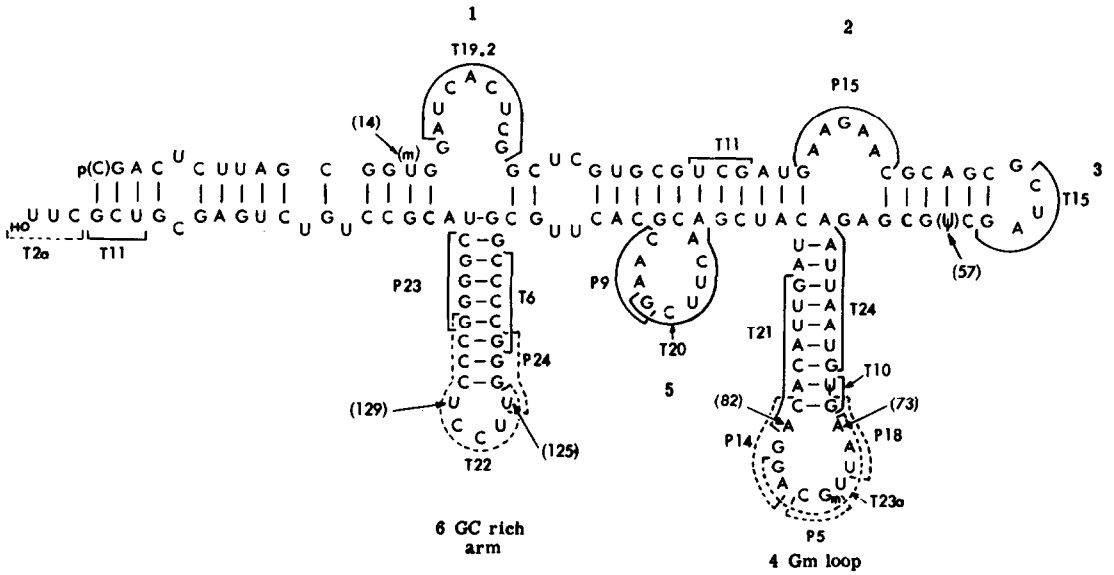

Fig.3

demonstrated by complete absence of the alkali-stable dinucleotide, Gm–Cp, among the alkaline hydrolysis products of the region representing spots 23 and 23a. There were no obvious 'clipped ends' corresponding to P14 or P18. However there was little or no diminution of the adjacent oligonucleotides, T21 (A–C–A–C–U–U–A–Gp), T24 (A–A–U–U–A–A–U–Gp) or T10 (ψ–Gp). These results are consistent with the proposed [2] secondary structure for nucleotides 64–91 (see fig.3), A28 in product T21 being at a loop–stem transition. The Tinoco [11] stability number for this 'arm' is +3. The very low yield of Gm–Cp from the ethanol precipitate after S_1 nuclease digestion suggests that the methylation site is the primary point of attack on the loop.

3.2. *Nucleotides 125–129*

Product T22 (U–U–C–C–U–C–C–C–Gp) was recovered in low yield after S_1 nuclease digestion. The nucleotides, U–U–C–C–U, are depicted in the proposed secondary structure (see fig.3) as a hairpin loop attached to a G–C rich stem. The Tinoco stability number for the proposed arm is +10. Oligonucleotides from the proposed stem were undiminished after S_1 digestion, except for a slight diminution of P24 (G–G–G–Up), which overlaps T22 at the stem–loop transition.

3.3. *The 3′-end*

The 3′-terminal oligonucleotide, C–U–U_{OH}, was recovered in low yield after S_1 digestion, indicating that the 3′-end is exposed. The adjacent oligonucleotide, T11 (G)–U–C–G, occurs twice in the sequence. Its yield was only slightly diminished after S_1 digestion, indicating that both GUCG sequences are protected.

All other pancreatic and T_1 oligonucleotides show little or no susceptibility to S_1 nuclease. In particular T19.2, P15, T15, and T20 plus P9, from the proposed loops 1, 2, 3 and 5, were recovered in good yields (table 1). So also were the various 'stem' oligonucleotides from the Gm arm and the G–C rich arm (see fig.3). So, finally, were the fractionally modified sequences T8, containg Um, and T12, containing ψ (table 1).

Failure of material representing the proposed loops 1, 2, 3 and 5 to undergo significant cleavage by S_1 nuclease indicates that these regions are protected, perhaps by tertiary interactions. Possible analogies might be the T–ψ–C loop and the dihydro-U U loop of tRNA, both of which are resistant to S_1 nuclease [8] and are involved in tertiary interactions in the three-dimensional structure [12].

We conclude that in 5.8 S RNA the Gm loop (73–81), loop (125–129) of the G–C rich stem and the extreme 3′-end (157–159) are single-stranded and accessible to S_1 nuclease. Since this test can distinguish a loop of as few as 5 nucleotides (125–129), any other single-stranded features, if they exist, must be shorter than this. Detection of such features may be facilitated by chemical probes for unpaired nucleotides, now in use in our laboratory [13].

King and Gould [9] proposed, on the basis of hyperchromicity measurements, that there are twice as many G–C pairs as A–U pairs in the mammalian 5.8 S structure (whose sequence was then unknown). The structure in fig.3 possesses 32 G–C pairs, 14 A–U pairs (of which two are A–ψ pairs, whose spectral behaviour would probably be similar to A–U pairs) and three G–U pairs. Thus the agreement between base pairing in the proposed secondary structure and that predicted by King and Gould is good.

Acknowledgements

This work was supported by the Medical Research Council. We thank J. Forbes for skilled technical assistance.

Fig.3. Secondary structure of 5.8 S RNA, proposed by Nazar et al. [2]. The proposed structure was for rat hepatoma. Here the HeLa sequence is shown in the same arrangement. Oligonucleotides corresponding to the sites of attack by S_1 nuclease are indicated by dotted lines. No other oligonucleotides were significantly attacked; important protected ones are shown in solid lines (locations of all unique oligonucleotides are shown in fig.1). Numbers 1–6 are the proposed loops in the Nazar [2] structure; 1, 2, 3 and 5 appear to be protected. Numbers in parentheses indicate nucleotides which are specifically referred to in the text.

References

[1] Pene, J. J., Knight, E. and Darnell, J. E. (1968) J. Mol. Biol. 33, 609–623.
[2] Nazar, R. N., Sitz, T. O. and Busch, H. (1975) J. Biol. Chem. 250, 8591–8597.
[3] Rubin, G. M. (1973) J. Biol. Chem. 248, 3860–3875.
[4] Maden, B. E. H. and Robertson, J. S. (1974) J. Mol. Biol. 87, 227–235.
[5] Nazar, R. N., Sitz, T. O. and Busch, H. (1976) Biochemistry 15, 505–508.
[6] Ando, T. (1966) Biochim. Biophys. Acta 114, 158–168.
[7] Rushizky, G. W., Shaternikov, V. A., Mozejko, J. H. and Sober, H. A. (1975) Biochemistry 14, 4221–4226.
[8] Harada, F. and Dahlberg, J. E. (1975) Nucleic Acids Research 2, 865–871.
[9] King, H. W. S. and Gould, H. (1970) J. Mol. Biol. 51, 687–702.
[10] Brownlee, G. G. (1972) in: Determination of sequences in RNA. North-Holland/Elsevier, Amsterdam.
[11] Tinoco, I., Uhlenbecke, O. C. and Levine, M. D. Nature 230, 362–367.
[12] Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C. and Klug, A. (1974) Nature 250.
[13] Goddard, J. P. and Maden, B. E. H. (1976) Nucleic Acid Research 3, 431–440.