Biochem. J. (1977) **167**, 211–221
Printed in Great Britain

Methylated Nucleotide Sequences in HeLa-Cell Ribosomal Ribonucleic Acid

CORRELATION BETWEEN THE RESULTS FROM 'FINGERPRINTING' HYDROLYSATES OBTAINED BY DIGESTION WITH T_1 RIBONUCLEASE AND WITH T_1 PLUS PANCREATIC RIBONUCLEASE

By B. EDWARD H. MADEN and M. SHAH N. KHAN
Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, U.K.

(*Received* 21 *March* 1977)

The methylated nucleotide sequences in HeLa-cell rRNA were previously characterized after enzymic digestion of the rRNA by T_1 ribonuclease alone or by combined T_1 plus pancreatic ribonucleases. For any methylated product occurring in a T_1-ribonuclease digest there must be one or more corresponding products in a combined T_1-plus-pancreatic-ribonuclease digest. Here we correlate fully the inter-relationship between the methylated products occurring in the two digestion systems. The analysis has led to the resolution of some previous uncertainties and has permitted an almost complete qualitative and quantitative description of the methylated components in HeLa-cell rRNA. The data are compared with those reported by other authors for HeLa-cell rRNA.

Methylated nucleotides in rRNA result from the interaction of methylating enzymes with specific regions in the rRNA sequences during ribosome maturation. These nucleotides become selectively labelled when cells are grown in the presence of methyl-labelled methionine under conditions that suppress non-specific purine-ring labelling. The methylation sites are therefore both biologically interesting and relatively amenable to detailed chemical analysis.

We have used two 'fingerprinting' systems to examine the methylated nucleotide sequences in HeLa-cell rRNA. These involve digesting the rRNA with T_1-ribonuclease alone or with combined T_1 plus pancreatic ribonucleases (Maden & Salim, 1974; Khan & Maden, 1976). In 'fingerprints' of T_1-ribonuclease digests the methylated oligonucleotides range in length from two to about 20 nucleotides long. Most products were recovered approximately once per RNA molecule, but several were recovered in lower yields. For some of the latter it was not clear whether the low recoveries were due to possible artifacts in the 'fingerprinting' system.

In 'fingerprints' obtained after digestion with combined T_1 plus pancreatic ribonuclease the digestion products are shorter than those in the T_1-ribonuclease system (one to seven nucleotides long). In 'fingerprints' of ^{32}P-labelled RNA many methylated products are fully resolved from non-methylated material owing to the presence of enzyme-resistant phosphodiester bonds adjacent to the methylated nucleosides. This feature of the 'fingerprinting' system provided the basis for absolute quantification of the molar frequencies of methylated products in rRNA (Maden *et al.*, 1972; Khan & Maden, 1976). Even for products that are not well resolved in 'fingerprints' of ^{32}P-labelled RNA, quantification relative to other products is generally reliable in 'fingerprints' of [^{14}C]methylated RNA owing to the short sequence lengths of the products.

For any methylated product occurring in a 'fingerprint' of a T_1-ribonuclease digest there must be one, or for certain multiply methylated products more than one, corresponding product in the 'fingerprint' obtained after digestion with combined T_1 plus pancreatic ribonucleases. We commented previously on some inter-relationships between products in the two 'fingerprinting' systems (Maden & Salim, 1974, Appendix II). Here we present a full description of all these inter-relationships. The correlation has led to resolution of a number of uncertainties about the molar yields and the chemical nature of various methylated components. The correlated data should prove useful for further studies, for example on interspecies comparisons and on conformational analysis of the methylated sequences.

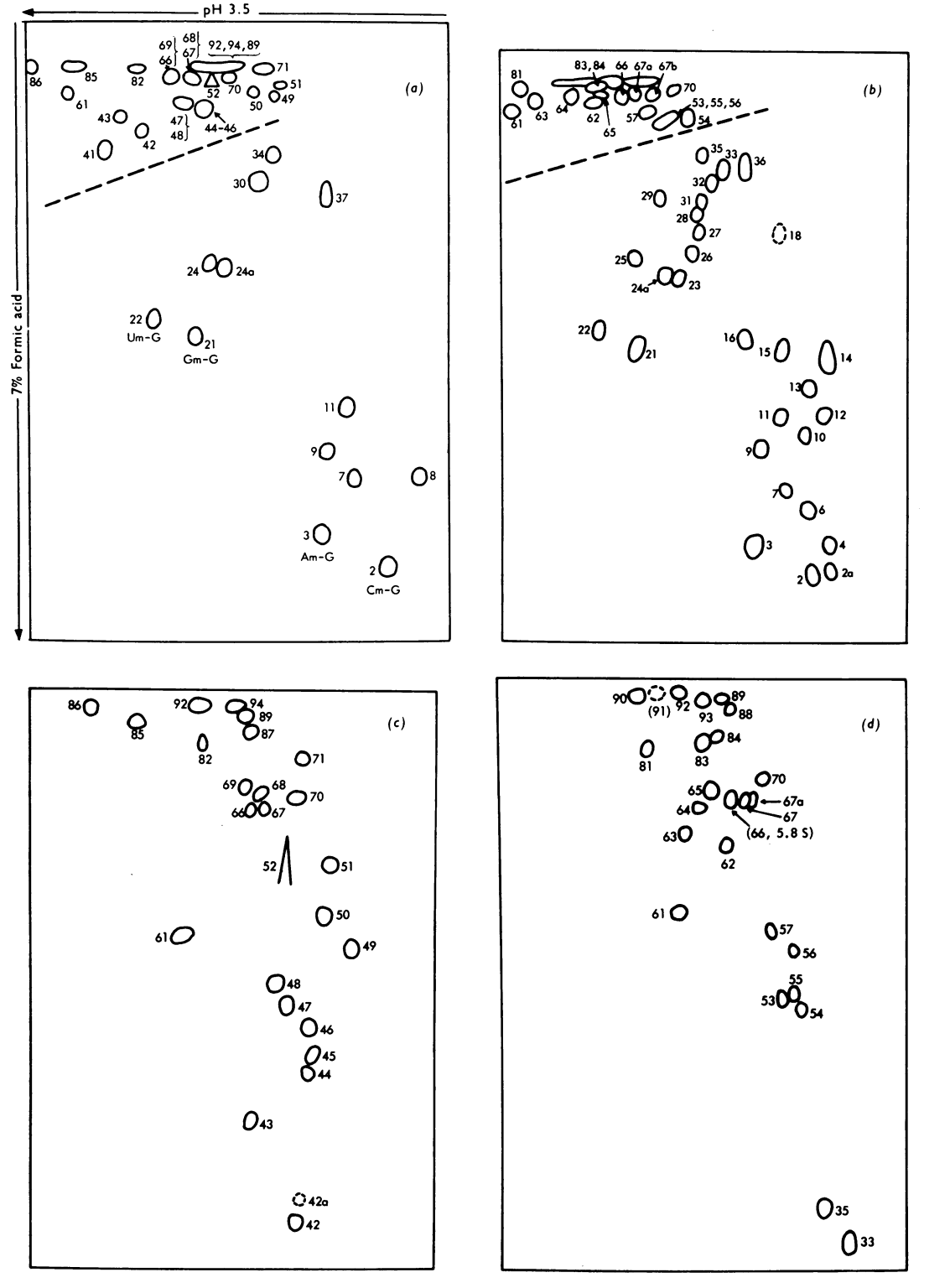
Methods

Preparation of RNA and 'fingerprinting' procedures

These have been described in detail (Maden & Salim, 1974; Khan & Maden, 1976). Other general methods were described by Brownlee (1972).

Numbering systems

For T_1-ribonuclease products the numbering system of Maden & Salim (1974) has been retained, with one change: the T_1-ribonuclease product previously designated 24′ (A-Gm-Gp) is here designated 24a. Since the numbering systems were


pH 3.5
7% Formic acid
(a)
(b)
(c)
(d)
Um-G
Gm-G
Am-G
Cm-G
92,94,89
44-46
83,84
67a
67b
53,55,56
(91)
(66, 5.8 S)

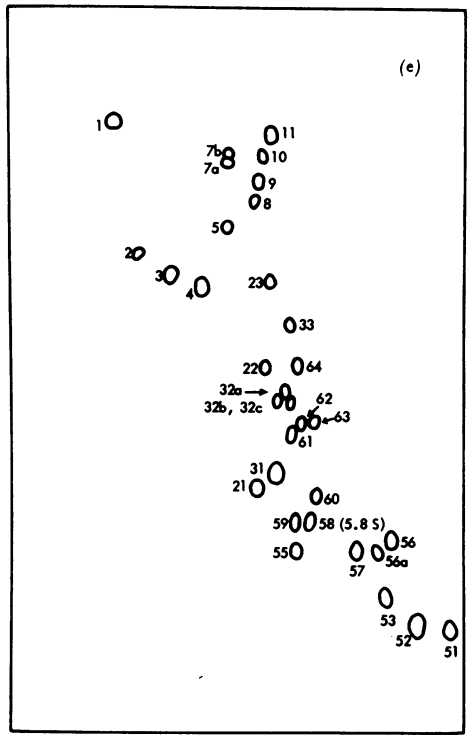


Fig. 1. *Diagram of 'fingerprints' of HeLa-cell methyl-labelled rRNA*

(*a*) 18S rRNA, standard T_1 digest. (*b*) 28S rRNA, standard T_1 digest. (*c*) 18S rRNA, T_1-ribonuclease plus alkaline phosphatase digest, with long separations in both dimensions. (*d*) 28S rRNA, T_1-ribonuclease plus alkaline phosphatase digest, with long separations in both dimensions. Several of the products above the broken lines are poorly resolved in (*a*) and (*b*), but most of these are well resolved in (*c*) and (*d*). The 18S-rRNA product 42a occurs in low and variable yield and is unidentified. The relationship between 28S-rRNA products T90 and T91 is unclear; see the text. (*e*) 28S rRNA, combined T_1-plus-pancreatic-ribonuclease digest. Details of the 'fingerprinting' systems, with electrophoresis conditions, times, voltages and photographs of 'fingerprints', were published previously (Maden & Salim, 1974; Khan & Maden, 1976). The anode is to the left in the first dimension (origin to the right) and at the bottom in the second dimension (origin at the top).

previously shown with 18S and 28S products on the same diagrams, they are reproduced for the two RNA species separately here (Fig. 1). (The keys to 'fingerprints' of combined 18S plus 28S rRNA were useful in relation to 'fingerprints' of 45S ribosomal precursor RNA.) For 'fingerprints' of T_1-plus-pancreatic-ribonuclease digests the previous numbering system (Khan & Maden, 1976) has been retained throughout; the keys (Figs. 1*e* and 2*b*) show the mobilities of the methylated products in this system.

In the present paper ribonuclease-T_1 products are designated T2, T3 etc., and products of combined T_1 plus pancreatic ribonuclease digestion TP1, TP2 etc.

Correlation of T and TP products

In many cases it was obvious from the available sequence data that a particular T_1-ribonuclease product was related to a particular T_1-plus-pancreatic ribonuclease product. In other cases such relationships could not be inferred unambiguously for various reasons. For the longer T_1-ribonuclease products the alkali-stable (2′-*O*-methylated) dinucleotides had been identified, but the analyses had not distinguished between (for example) Am*-Up or A-Am-Up in a particular oligonucleotide. In a few cases there remained uncertainties about the chemical nature of the methylated components in particular digestion products. For these reasons, and to provide a 'consistency check' on all previously reported data, the following procedure was undertaken. All except the smallest T_1-ribonuclease products ([^{14}C]methylated) were eluted from 'fingerprints' and were digested with pancreatic ribonuclease. The products were separated

* Abbreviations: 2′-*O*-ribose-methylated nucleotides are designated Nm, base-methylated nucleotides are designated mN and internal phosphates with a hyphen.

by electrophoresis on DEAE-paper in 7% (v/v) formic acid at 1 kV for 16h. Successive digestion by T_1 and pancreatic ribonuclease should produce the same end products as simultaneous digestion by both enzymes. The mobilities of the products were therefore compared with those of the previously identified T_1-plus-pancreatic-ribonuclease products in the second (DEAE-paper) dimension of the 'fingerprints'. In most cases the mobility, together with previous information on the sequence or alkali-stable components of the T_1-ribonuclease product, was sufficient to establish the product's relationship to a corresponding T_1-plus-pancreatic-ribonuclease product. (Examples are given in the Results and Discussion section, Fig. 2.) In other cases the product or products were re-eluted and digested with alkali or T_2 ribonuclease, followed by electrophoresis on Whatman 52 paper at pH 3.5 (Brownlee, 1972; Maden & Salim, 1974). In a few cases this was followed by dephosphorylation and re-electrophoresis (Maden & Salim, 1974). From the mobilities of the various degradation products the remaining uncertainties were resolved, as described in the Results and Discussion section.

An important technical point was the following. Several T_1-ribonuclease products give rise, on digestion with pancreatic ribonuclease, to methylated components terminated by G. Some of these T_1-ribonuclease products are long and are therefore best separated in 'fingerprints' of 'T_1-ribonuclease plus alkaline phosphatase' digests (see Figs, 1*c* and 1*d*). However, dephosphorylation affects the mobility on DEAE-paper, not only of the original T_1-ribonuclease product, but also of its pancreatic-ribonuclease degradation product. Although these effects on mobility were generally interpretable, nevertheless, as an additional check we carried out pancreatic-ribonuclease digestion on all products from 'standard T_1-ribonuclease fingerprints', as well as from 'fingerprints' of T_1-ribonuclease plus alkaline phosphatase digests. Many long products remained close to the origin in the standard T_1-ribonuclease 'fingerprints', but it was possible to excise and carry out confirmatory analyses on areas that contained, at the most, two or three unresolved products (Figs. 1*a* and 1*b*), on which sufficient partial sequence data were already available to permit the relationships between the individual T_1- and T_1-plus-pancreatic-ribonuclease products to be deduced. Examples are described in the first part of the Results and Discussion section.

Results and Discussion

Product inter-relationships

Fig. 2(*a*) shows the pancreatic-ribonuclease degradation products of several consecutively numbered T_1-ribonuclease oligonucleotides from 18S rRNA. The relationships to TP products were deduced from these mobilities and from additional criteria as follows.

Oligonucleotide T85 is a single T_1-ribonuclease product that has previously been characterized as containing two methylated components Um-Cp and Um-Gp (Maden & Salim, 1974). The mobilities of the pancreatic-ribonuclease degradation products are consistent with these designations (Fig. 2*b*, products TP3 and TP55; see also Table 1). Moreover, Um-G was released from the 'fingerprint' of a T_1-ribonuclease-plus-alkaline-phosphatase digest as the dinucleoside monophosphate, Um-G_{OH}, which migrates much more rapidly than Um-Gp (and just behind Um-Cp). The isomer Gm-Up would have been released unchanged as the dinucleotide diphosphate. Um-Cp (TP55) is directly distinguishable from its isomer Cm-Up (TP59) by its mobility on DEAE-paper. Thus the present findings are fully consistent with the previously reported product designations.

Oligonucleotide T82 was known to contain Am-Up. Fig. 2 shows that this is present within the pancreatic-ribonuclease product, A-Am-Up. Therefore oligonucleotide T82 corresponds uniquely to oligonucleotide TP22, since each is present once per 18S molecule (Table 1).

Oligonucleotide T71 was known to contain Am-Ap. Of the various TP products that co-migrate with the pancreatic-ribonuclease-digestion product of oligonucleotide T71 (Fig. 2*a*), only one contains Am-A, namely Am-A-A-Up in oligonucleotide TP23 (Fig. 2*b*, Table 1). Product T71 therefore contains the partial sequence, Am-A-A-Up.

Oligonucleotide T70 was known to contain A-Gm-Cp from an analysis of ^{32}P-labelled material (Salim, 1972). The ^{14}C-labelled pancreatic-ribonuclease-digestion product migrates to the same position as oligonucleotide TP62. This contains the isomers A-Gm-Cp and A-Cm-Gp. Since A-Cm-Gp corresponds to product T7 (Table 1), the result confirms that product T70 contains A-Gm-Cp.

Oligonucleotides T66 plus T69 are separated in the second dimension of 'fingerprints' obtained after digestion with T_1 ribonuclease plus alkaline phosphatase (Fig. 1*c*), but co-migrate in 'fingerprints' obtained after digestion with T_1 ribonuclease alone (Fig. 1*a*). Oligonucleotide T66 was known to contain Am-UP, and this is confirmed by the mobility of the more rapidly migrating product, which corresponds to that of TP21 (Am-Up). Oligonucleotide T69 was incorrectly reported to contain Um-Ap (Maden & Salim, 1974), but has now been found to contain Gm-Gp, and this is consistent with the quantitative data on product TP4 and the three corresponding T_1-ribonuclease products in Table 1. The previous error was probably due to either or both of the following difficulties: all four products T66–T69 migrate

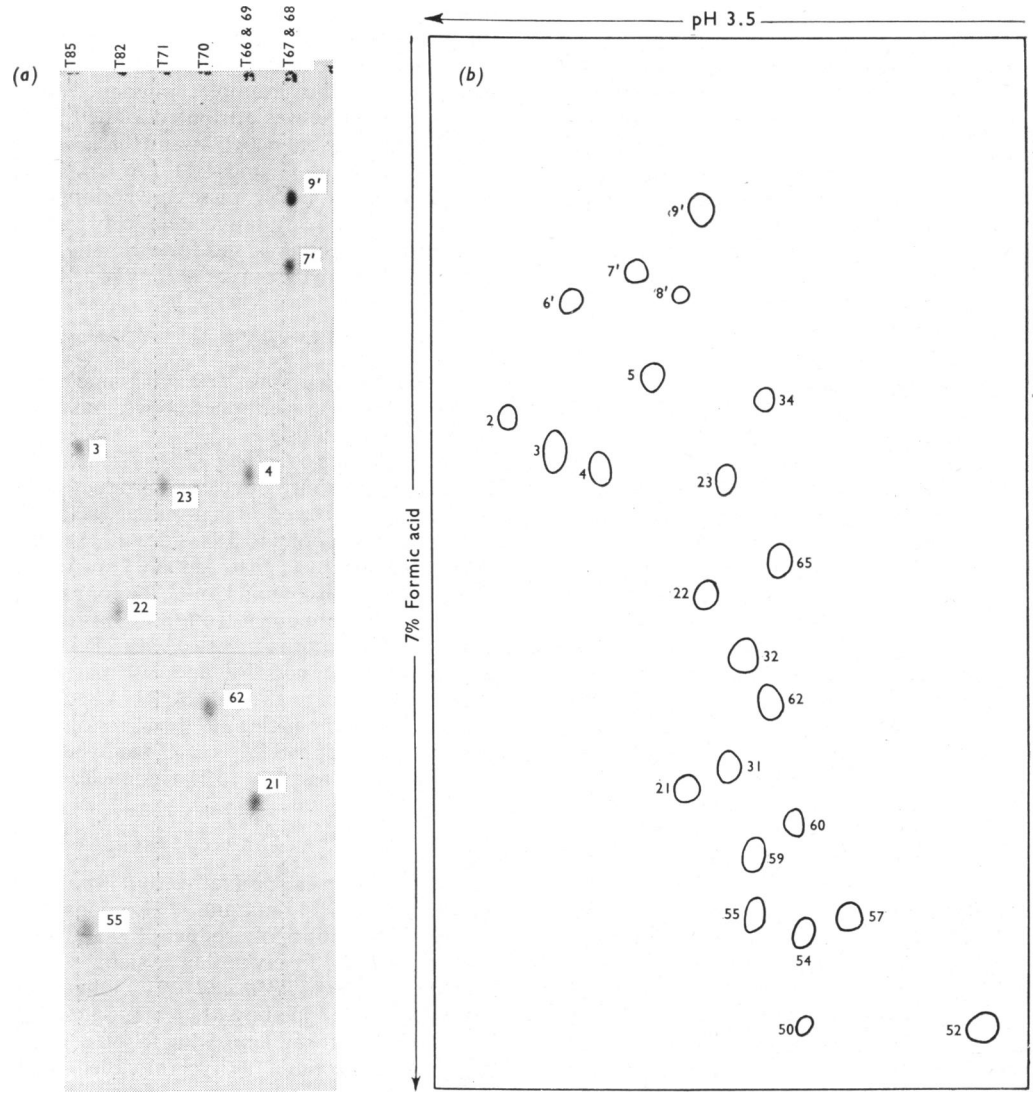


Fig. 2. (*a*) *Electrophoretic separation on DEAE-paper (in 7% formic acid) of pancreatic-ribonuclease digests of oligonucleotides from a 'fingerprint' T_1-ribonuclease digest of HeLa-cell methyl-labelled 18S rRNA*; (*b*) *diagram of combined T_1-plus-pancreatic-ribonuclease digest of HeLa-cell methyl-labelled 18S rRNA, drawn from the 'fingerprint' in Fig. 3 of Khan & Maden* (1976), *and shown with the second dimension (DEAE-paper, in 7% formic acid) on the same scale as in* (*a*)
The separations were for 16h at 1 kV in the second dimension (downwards). The anode is at the bottom.

rather close to each other in the 'fingerprint' (Fig. 1*c*) and they also give rise to alkali-stable components that migrate rather similarly on electrophoresis on Whatman 52 paper at pH 3.5

Oligonucleotides T67 plus T68, like oligonucleotides T66 and T69, are separated in 'fingerprints' obtained after digestion with T_1 ribonuclease plus alkaline phosphatase (Fig. 1*c*), but co-migrate in 'fingerprints' obtained after digestion with T_1 ribonuclease alone (Fig. 1*a*). Oligonucleotide T67 was known to contain Um-Ap, and this can now be identified with Um-A-A-Up, TP7′. Oligonucleotide T68 was known to contain two alkali-stable components, Um-Ap and Cm-Ap. These can now be identified with the single doubly methylated sequence, TP9′, A-(A,Um-A)-Cm-A-Up.

Tables 1 *and* 2

From an extension of this analysis and these types of reasoning it was possible to collate the information shown in Tables 1 and 2. Certain general features of the Tables may be noted here. Most of the sequence data were obtained previously (Maden & Salim, 1974; Khan & Maden, 1976), but in several cases the present analysis permitted slight extensions to the previously published sequences of T_1-ribonuclease digestion products, as in the examples described above. In a few cases amendments have been made. These are discussed under 'Specific points' below.

Three types of sequence relationship occur in the Tables. First, several T_1-ribonuclease products may correspond to a single T_1-plus-pancreatic-ribonuclease 'spot', for example TP spots 3 or 4. The T_1-plus-pancreatic-ribonuclease products are listed on the left of the Tables to facilitate the display of quantitative aspects of this type of relationship. Alternatively a single T_1-ribonuclease product may correspond to a single T_1-plus-pancreatic-ribonuclease product, for example oligonucleotide T82 to TP22, mentioned above. Finally a single doubly methylated T_1 ribonuclease product may correspond to two 'derivative' T_1-plus-pancreatic ribonuclease products, for example oligonucleotide T85 to TP3 and TP55, also mentioned above.

The molar yields of the products are from 'fingerprints' of [^{14}C]methylated RNA. They are based essentially on those published previously (Maden & Salim, 1974; Khan & Maden, 1976), but have been refined slightly by inclusion of data from additional 'fingerprints'. For 'fingerprints' obtained after digestion with combined T_1 plus pancreatic ribonuclease the absolute molar yields from ^{32}P-labelled RNA were very similar to those shown here for most of the well-resolved TP spots 2–11. However, spots TP2 and TP6 gave slightly higher yields in 'fingerprints' of ^{32}P-labelled RNA (Khan & Maden, 1976), probably owing to the presence of traces of incompletely hydrolysed non-methylated material. The fact that certain products give slightly higher than integral molar-yield values in 'fingerprints' of [^{14}C]methylated RNA results from expressing the individual yields relative to means of several products in each 'fingerprinting' system.

It is important to note the conventions used for expressing the yields of multiply methylated products. In general we have expressed these yields in terms of numbers of methyl groups, and not in terms of the molar frequencies of the oligonucleotides. Thus 18S-rRNA product TP54 occurs once per molecule but contains four methyl groups, and is therefore attributed a value of 4 in Table 1. This convention corresponds to that of Maden & Salim (1974) and differs from that of Khan & Maden (1976) in which data from 'fingerprints' of ^{32}P-labelled RNA were also included. However, where a single T_1-ribonuclease product corresponds to two T_1-plus-pancreatic-ribonuclease products the yield was treated as in this example: oligonucleotide T85, mentioned above, was attributed a 'methyl' yield of 2.22 by Maden & Salim (1974), and this is here divided into 1.11 for Um-G and 1.11 for Um-C. As will become apparent below, these conventions are useful for comparing quantitative data between the two 'fingerprinting' systems and for estimating the total numbers of methyl groups.

Specific points: 18*S rRNA*

*Oligonucleotides TP*4, *T*69. Oligonucleotide T69 contains Gm-G, as mentioned above, and not Um-A, as reported previously.

*Oligonucleotide TP*8′ *and TP*55, *T*94 *and T*89. The T_1-plus-pancreatic-ribonuclease product TP8′ was persistently recovered in half-molar yield. Product TP8′ must correspond to T94, since these are the only 18S-rRNA products that contain Gm-A. Products T94 and T89 both contain Um-C, the combined molar yields of this component from both T_1-ribonuclease products being approx. 1 (listed with TP 55). It seems possible that products T89 and T94 may be derived from the same part of the 18S-rRNA sequence, but that in 50% of the molecules the G residue, which is terminal in product T89, is 2′-*O*-methylated in T94, and is therefore resistant to T_1 ribonuclease:

T89 . . . . . Um-C . . . A-G
T94 . . . . . Um-C . . . A-Gm-A-G

This explanation is speculative, but would account for the molar-yield data and is also consistent with the related mobilities of products T89 and T94 in the T_1-ribonuclease 'fingerprinting' system.

*Oligonucleotides TP*50 *and TP*37. Oligonucleotide T37 was not identified by Maden & Salim (1974). It is now known to be approximately ten nucleotides long and to contain a hypermodified nucleotide 3-(3-amino-3-carboxypropyl)-1-methylpseudouridine (Saponara & Enger, 1974; Maden *et al.*, 1975). The low recovery of the T_1-ribonuclease product is an artifact and is due to streaking and, possibly, poor transfer from the first to the second dimension of the 'fingerprint'. The corresponding oligonucleotide TP50 is recovered in almost unimolar yield.

*Oligonucleotides TP*2 *and T*86. The 2′-*O*-methyl group is located on the first U residue of product T86, as shown here, and not on the second, as reported previously (Maden & Salim, 1974). This conclusion arises from an analysis (B. E. H. Maden, unpublished work) of 'fingerprints' obtained after digestion of methyl-labelled RNA with pancreatic ribonuclease only, from which it was possible to relate a unique product, . . . G-Um-U, with product T86.

Table 1. *Relationship between methylated products in the two 'fingerprinting' systems of HeLa-cell 18S rRNA*

The molar yields of combined T_1-plus-pancreatic-ribonuclease products are means from five 'fingerprints' of methyl-labelled rRNA. Molar yields of T_1-ribonuclease products are means from five 'fingerprints' of methyl-labelled rRNA after digestion with T_1 ribonuclease alone for products 2–43 (Fig. 1*a*) and four 'fingerprints' after digestion with combined T_1-ribonuclease plus alkaline phosphatase for products 41–94 (Fig. 1*c*). (*a*) See the text for comments on oligonucleotides T69 and T86. (b) One mol of Um-Gp in TP3 and 1 mol of Um-Cp in TP55 are each derived from T85; see the text and Fig. 2(*a*). (*c*) The location of the 2′-*O*-methyl group on the first rather than the second adenosine is tentative. (*d*) Product T52 always streaked, as shown in Fig. 1(*c*) and in Plate 2 of Maden & Salim (1974). However, the yield of this product was higher than that of the known corresponding TP34. We have noticed an extra weakly labelled TP product, 33b, which migrates slightly faster in the second dimension than product TP34, and which may possibly also be related to product T52. (*e*) The methylated component co-migrates with m^6Ap on Whatman 52 paper at pH3.5 after hydrolysis of the oligonucleotide by either alkali or T_2 ribonuclease. (*f*) A product corresponding in mobility to TP65, and designated m^7G-A-A-Up, also occurs in the 'fingerprint' of yeast 17S rRNA after digestion with combined T_1 plus pancreatic ribonucleases (Klootwijk & Planta, 1974). When product T49 is eluted with 30% triethylamine carbonate, pH10, and digested with pancreatic ribonuclease, two products are seen on electrophoresis on DEAE-paper in 7% formic acid. The more rapidly migrating product corresponds in mobility to TP65 and is presumably m^7G-A-A-Up. The more slowly migrating product corresponds approximately in mobility to (G,A_2,U). This is presumed to contain 4-amino-5-*N*-methyl-formamidoisocytidine, the alkali-degradation product of m^7G (Brownlee, 1972), produced by elution with triethylamine. The nucleotide containing this product behaves approximately like Gp on electrophoresis (Brownlee, 1972; B. E. H. Maden, unpublished work).

Combined T_1-plus-pancreatic-ribonuclease products			T_1-ribonuclease products					
Spot no.	Sequence	Yield (CH_3)	Suggested (both systems)	Yield (CH_3)		Spot no.	Sequence	Notes
TP2	Um-Up	1.05	(1)	1.11		T86	(G)-Um-U-U-A-C-U-U-U-Gp	(*a*)
TP3	Um-Gp +Gm-Up	4.28	(4.5)	4.55	0.95	T22	(G)-Um-Gp	
					0.51	T42	(G)-C-U-Um-Gp	
					0.84	T47	(G)-A-A-C-U-Um-Gp	
					1.11	T85	.... Um-Gp	(*b*)
					1.14	T41	(G)-Gm-U-Gp	
TP4	Gm-Gp	2.89	(3)	3.03	1.05	T21	(G)-Gm-Gp	
					1.12	T61	(G)-ψ-U-Gm-Gp	
					0.86	T69	... Gm-Gp	(*a*)
TP5	A-Gm-Gp	2.69	(2.5–3)	2.47	1.71	T24a	(G)-A-Gm-Gp	
					0.76	T50	... A-Gm-Gp	
TP6′	A-Um-Up	0.97	(1)	1.01		T43	(G)-A-Um-U-Gp	
TP7′	Um-A-A-Up	1.00	(1)	0.93		T67	... Um-A-A-Up ...	
TP8′	A-Gm-A-Gp	0.51	(0.5)	0.54		T94	... A-Gm-A-Gp ...	
TP9′	A-(A,Um-A)-Cm-A-Up	1.72	(2)	1.82		T68	... A-(A,Um-A)-Cm-A-Up ...	
TP21	Am-Up	2.95	(3)	2.75	1.03	T46	... Am-U-Gp	
					0.92	T66	... Am-Up ...	
					0.80	T87	... Am-Up ...	
TP22	A-Am-Up	1.12	(1)	0.91		T82	... A-Am-Up ...	
TP23	Am-A-A-Up	1.00			0.73	T71	... Am-A-A-Up ...	(*c*)
	+A-A-Am-Up	0.67	(1.6)	1.39	0.66	T92	... A-A-Am-Up ...	
TP31	Am-Gp	2.25	(2)	2.00	1.07	T3	(G)-Am-Gp	
					0.93	T24	(G)-U-Am-Gp	
TP32	Am-A-Gp	2.12	(2)	1.87	0.91	T9	(G)-Am-A-Gp	
					0.96	T48	(G)-A-U-U-Am-A-Gp	
TP34	(A_3,Am-A)Gp	0.74	(1)	1.00		T52	... (A_3,Am-A)-Gp	(*d*)
TP50	amΨ-Cp	0.94	(1)	0.50		T37	... amΨ-Cp ...	
TP52	Cm-Cp	2.89	(2.5)	2.14	1.55	T8	(G)-(C,Cm-C)-Gp	
					0.59	T51	... Cm-Cp ...	
TP54	m_2^6A-m_2^6A-Cp	4.28	(4)	3.75		T30	(G)-m_2^6A-m_2^6A-Cp ...	
TP55	Um-Cp	3.01	(3)	3.10	0.99	T44	(G)-A-U-Um-C-C-Gp	
					1.11	T85	... Um-Cp ...	(*b*)
					0.47	T89	... Um-Cp ...	
					0.53	T94	... Um-Cp ...	
TP57	A-Am-Cp +(m^6A,A)Cp	3.07	(3)	2.73	1.91	T11	(G)-A-Am-C-Gp	
					0.82	T34	... (m^6A,A)Cp ...	(*e*)
TP59	Cm-Up	1.06	(1)	1.04		T45	(G)-(A,C,U)-Cm-U-Gp	
TP60	Cm-Gp	1.11	(1)	1.06		T2	(G)-Cm-Gp	
TP62	A-Cm-Gp +A-Gm-Cp	2.14	(2)	1.88	0.95	T7	(G)-A-Cm-Gp	
					0.93	T70	... A-Gm-Cp	
TP65	m^7G-A-A-Up	1.04	(1)	1.01		T49	... m^7G-A-A-Up	(*f*)
Total CH_3 groups			(45)					

Table 2. *Inter-relationship between methylated products of combined T_1 plus pancreatic ribonuclease digestion and T_1 ribonuclease digestion of HeLa-cell 28 S rRNA*

Molar yields of T_1-plus-pancreatic-ribonuclease products are means from three 'fingerprints' of methyl-labelled RNA. T_1 molar yields are from five 'fingerprints' obtained from digestion with T_1 ribonuclease only (products T1 to T36) and four 'fingerprints' obtained after digestion with combined T_1 ribonuclease plus alkaline phosphatase (products T53 to T93). (*a*) Product T92b is poorly resolved, even in 'fingerprints' obtained after digestion with T_1 ribonuclease plus alkaline phosphatase and its yield is uncertain. (*b*) In addition to product TP11 there was a small quantity of an extra product. The latter corresponded in mobility to a product previously numbered TP31a in chick 28 S RNA, Am-Gm-Cp (Khan & Maden, 1976). It is possible that in a small proportion of molecules, T93 is not methylated at C in HeLa cells. (*c*) See comment on oligonucleotides T67 and T67a in the text. (*d*) Product T18, a low-yield 'fractional' product, may also contain Am-Gp and contribute approx. 0.2 mol to oligonucleotide TP31.

Combined T_1-plus-pancreatic-ribonuclease products				T_1-ribonuclease products			
Spot no.	Sequence	Yield (CH_3)	Suggested (both systems)	Yield (CH_3)	Spot no.	Sequence	Notes
TP1	Um-Gm-Up +Um-Gm-ψp	4.25	(4)	3.76	1.93 T83	.. Um-Gm-Up ...	
					1.83 T92a	... Um-Gm-ψp	
TP2	m^3U-Up	1.19	(1)	1.18	T63	(G)-m^3U-U-ψ-A-Gp	
TP3	Um-Gp +Gm-Up	5.42	(5.5)	5.10	1.05 T22	(G)-Um-Gp	
					0.89 T53	(G)-Gm-U-A_{2-3}-C-Gp	
					1.06 T81	(G)-U-A-U-Gm-U-Gp	
					0.76 T84	... Gm-Up ...	
					0.84 T90	... Gm-Up ...	
					0.5(?) T92b	... Gm-Up ...	(*a*)
TP4	Gm-Gp	8.93	(9)	8.95	0.79 T16	(G)-C-C-Gm-Gp	
					6.97 T21	(G)-Gm-Gp	
					1.19 T61	(G)-U-U-Gm-Gp	
TP5	A-Gm-Gp	0.96	(1)	1.09	T24a	(G)-A-Gm-Gp	
TP7a	A-A-Um-Gp				0.93 T35	(G)-A-A-C-A-A-Um-Gp	
TP7b	A-A-ψm-Gp	1.68	(2)	1.83	0.90 T29	(G)-A-A-ψm-Gp	
TP8	Cm-A-Gm-Up	2.06	(2)	2.21	T62	(G)-Cm-A-Gm-U-U-Gp	
TP9	Gm-A-Am-A-Gp	1.64	(2)	1.80	T36	(G)-(A,C_{3-5})-Gm-A-Am-A-Gp	
TP10	A-(A,Gm-A)Gp	0.84	(1)	0.93	T31	(G)-A-(A,Gm-A)Gp	
TP11	Am-Gm-Cm-A-A-A-Up	2.22	(2–3)	1.63	T93	... Am-Gm-Cm-A-A-A-Up	(*b*)
TP21	Am-Up	2.56	(2.5)	2.68	1.03 T25	(G)-Am-U-Gp	
					1 T67a?	... Am-Up ...	(*c*)
					0.65 T70	... Am-Up ...	
TP22	A-Am-Up	0.42	(0.5)	0.5	T67?	... A-Am-Up ...	(*c*)
TP23	A-A-Am-Up	0.99	(1)	1.10	T33	... A-A-Am-Up ...	
TP31	Am-Gp	6.56	(6.5)	6.42	4.42 T3	(G)-Am-Gp	(*d*)
					1.15 T6	(G)-C-Am-Gp	
					0.85 T15	(G)-A-A-C-Am-Gp	
TP32a	A-Am-Gp				1.15 T9	(G)-A-Am-Gp	
TP32b	A-Cm-Up				0.80 T88	... A-Cm-Up ...	
TP32c	Cm-A-Up	3.00	(3)	2.96	1.01 T27	(G)-Cm-A-U-Gp	
TP33	A-Am-A-Gp	1.02	(1)	1.03	T32	(G)-(C,U)-A-Am-A-Gp	
TP51	mCp	2.29	(2)	2.3	0.2 T2a	(G)-mC-Gp	
					1.10 T33	... mCp ...	
					1.0 T67	... mCp ...	(*c*)
TP52	Cm-Cp	2.83	(3)	2.64	0.93 T13	(G)-Cm-C-A-A-Gp	
					0.90 T54	... Cm-Cp ...	
					0.80 T89	... Cm-Cp ...	
TP53	Am-Cp	1.07	(1)	0.83	T14	... Am-Cp ...	
TP55	Um-Cp	1.89	(2)	1.79	0.95 T55	... Um-Cp ...	
					0.84 T90	... Um-Cp ...	
TP56	(m^1A,A)-Cp	1.13	(1)	0.83	T14	... (m^1A,A)-Cp ...	
TP56a	A-Cm-Cp	1.01	(1)	0.95	T10	(G)-A-Cm-C-Gp	
TP57	A-Am-Cp +(m^6A,A)Cp	4.28	(4)	3.99	1.05 T11	(G)-A-Am-C-Gp	
					0.98 T57	(G)-A-Am-C-ψ-A-U-Gp	
					0.98 T65	... A-Am-Cp ...	
					0.98 T28	... (m^6A,A)Cp ...	

Table 2—*continued*

Combined T_1-plus-pancreatic-ribonuclease products			T_1-ribonuclease products				
Spot no.	Sequence	Yield (CH_3)	Suggested (both systems)	Yield (CH_3)	Spot no.	Sequence	Notes
TP59	Cm-Up	3.34	(3)	2.71	0.85 T23	(G)-Cm-U-Gp	
					0.78 T56	... Cm-Up	
					1.08 T64	(G)-A-U-Cm-U-U-Gp	
TP60	Cm-Gp	1.83	(2)	2.01	1.08 T2	(G)-Cm-Gp	
					0.93 T4	(G)-C-Cm-Gp	
TP61	Um-A-Cp	1.12	(1)	0.99	T26	(G)-Um-A-C-Gp	
TP62	A-Cm-Gp	0.99	(1)	1.14	T7	(G)-A-Cm-Gp	
TP63	Cm-A-Gp	0.93	(1)	1.03	T12	(G)-C-C-Cm-A-Gp	
TP64	(Gm-A,A)Cp	0.91	(1)	0.84	T90	... Gm-A ...	
Total CH_3 groups, 28S RNA			(67–68)				
5.8S RNA:							
TP3	Um-Gp	0.21	(0.2)	0.2	T22	(G)-Um-Gp	
TP58	Gm-Cp	0.91	(1)	0.84	T66	(G)-A-A-U-U-Gm-C-A-Gp	
Total CH_3 groups, 28S plus 5.8S			(69)				

Specific points: 28 *S-rRNA*

*Oligonucleotides TP*3, *T*22. The T_1-ribonuclease product Um-G (T22) was previously attributed the rather high yield of 1.26 (Maden & Salim, 1974). This was due to the presence of 0.2 mol of Um-G from a fractionally methylated site on 5.8S rRNA (Nazar *et al.*, 1975; Khan & Maden, 1976), unrecognized at that time.

Products containing Am-C, A-Am-C, A-Cm-C, (m^1A, A)-C and (m^6A, A)-C. The inter-relationship between the various T_1-ribonuclease and T_1-plus-pancreatic-ribonuclease products containing these sequences was at first confusing, but has now been clarified. The crucial observations concerned oligonucleotide T14 (listed opposite TP53 and TP56 in Table 2), which was known to contain a base-methylated A and Am-C (Maden & Salim, 1974). Digestion of oligonucleotide T14 with pancreatic ribonuclease seemed to produce three products. The most rapidly migrating one was identified as Am-Cp. The other two products were only marginally separated, forming a split band whose mobility was roughly coincident on DEAE-paper with that of A-Am-Cp (TP57). The major component of this band was identified as containing m^1A by the following criteria. Alkaline hydrolysis yielded a product whose mobility on Whatman paper at pH 3.5 was similar to that of Ap or m^6Ap. However, hydrolysis with T_2 ribonuclease yielded mainly a product of practically zero mobility at pH 3.5, indicating that the base contained an integral positive charge which fully neutralized the negative charge of the phosphate group. In addition, a small amount of the presumed m^6Ap was present in the T_2-ribonuclease hydrolysate. 1-Methyladenosine is converted into 6-methyladenosine by hydroxyl ions [Dimroth rearrangement; Macon & Wolfenden, (1968) and references therein]. Kinetic studies have shown that the rate of conversion is appreciable at pH 10 (Macon & Wolfenden, 1968), the pH at which oligonucleotides are eluted from DEAE-paper by triethylamine carbonate before further digestion procedures. We conclude that the methylated base in oligonucleotides T14 and TP56 is m^1A, and that this is partly converted into m^6A on elution of the respective oligonucleotides with triethylamine carbonate. Analyses of ^{32}P-labelled material from oligonucleotide TP56 indicated the composition as (m^1A,A)Cp. It was demonstrated by similar procedures that oligonucleotide T28 (and also 18S-rRNA spot T34) contain (m^6A,A)-Cp. This, unlike (m^1A,A)-Cp, possesses no unusual charge properties and co-migrates with A-Am-Cp (TP57) in both dimensions of 'fingerprints' obtained after digestion with combined T_1 plus pancreatic ribonucleases. A-Am-Cp is present in three products from 28S rRNA. T_1-ribonuclease product A-Cm-Cp (TP56a) is present in oligonucleotide T10.

*Oligonucleotides TP*2, *T*63. The phosphodiester bond in mU-U is resistant to pancreatic ribonuclease but is cleaved by alkali. As this is the only base-methylated U in 28S rRNA, we infer it to be m^3U (Klagsbrun, 1973).

*Oligonucleotide T*67. This can be partly resolved into two components, designated T67 and T67a. The distribution of Am-Up and A-Am-Up between components T67 and T67a has not been established.

*Oligonucleotides T*90 *and T*91. On hydrolysis with pancreatic ribonuclease, oligonucleotide T90 yields three methylated components: Gm-Up, corresponding to TP3, Um-Cp, corresponding to TP55, and (A,Gm-A)-Cp, corresponding to TP64. Oligonucleo-

tide TP55 is liberated only twice per molecule of 28S rRNA. The T_1-ribonuclease product, T55, also contains Um-Cp. It was reported that a third T_1-ribonuclease product, T91, contained Um-Cp (Maden & Salim, 1974). An alternative possibility, which would be consistent with the observed molar yield of product TP55, is that oligonucleotide T90 is a complex spot, oligonucleotide T91 being the Um-Cp-containing component, and that this is sometimes partially resolved as a separate spot. Because of this possibility we have omitted T91 as a separate product from Table 2.

Quantitative aspects of the data

The molar yields of the T_1-plus-pancreatic-ribonuclease products should correspond to those of the respective T_1-ribonuclease products. In most cases the agreement is good. The few discrepancies were discussed above or are mentioned in the legends to the tables. From the data it is possible to suggest numbers of methyl groups corresponding jointly to each TP product and the related T_1-ribonuclease product(s) (Tables 1 and 2). The sum of these suggested numbers should give an accurate estimate of the total numbers of methyl groups in each rRNA species, since the individual numbers are themselves the results of cross-checking between the two systems. The total numbers of methyl groups estimated in this way are shown at the foot of the Tables. The number for 18S rRNA is 45, compared with a previous estimate of 46 (Maden & Salim, 1974), and for 28S rRNA is 68, compared with a previous estimate of 70 (already quoted).

These estimates involve simplifications with respect to the actual numbers of methylation sites. The 18S rRNA would be more accurately described as possessing 39 sites with singly fully methylated nucleotides, four sites with singly fractionally methylated nucleotides (T8, T42, T92 and T94) and one site with two adjacent dimethyladenosines (T30). In this description the doubly methylated sequences, T68 and T85, are regarded as each possessing two nearby but separate methylation sites. Similarly 28S rRNA possesses at least three sites where the methylation of a nucleotide is fractional (T2a, T18 and T67) and several sites where two or more methyl groups are clustered. The 5.8S rRNA possesses a fractionally methylated Um-Gp (Nazar *et al.*, 1975; Khan & Maden, 1976). Although the term 'numbers of methyl groups per RNA molecule' is only an approximation to the numbers of methylation sites, it is a useful descriptive term.

Comparison with other analyses of HeLa-cell rRNA

We discussed previously the quantitative relationship of our data to those of Wagner *et al.* (1967); see page 163 of Maden & Salim (1974). Iwanami & Brown (1968) reported on the methylated components in HeLa-cell rRNA after an acid-hydrolysis procedure. They failed to identify any component as being qualitatively unique to 18S or 28S rRNA, possibly owing to incomplete resolution of the two RNA species on sucrose gradients, as suggested by Klagsbrun (1973). They did, however, identify m^4C in HeLa-cell rRNA, and it is therefore probable that our unidentified base-methylated C in oligonucleotide T33 and/or T67 is m^4C. Klagsbrun (1973) also analysed the methylated components in HeLa-cell rRNA after an acid-hydrolysis procedure (trifluoroacetic acid for 30min at 170°C). Several methylated bases were clearly attributable to 18S or 28S RNA: m_2^6A and m^7G were present predominantly in 18S rRNA, m^1A and m^3U predominantly in 28S rRNA, and m^6A in both RNA species, as reported here. In the studies by Iwanami & Brown (1968) and by Klagsbrun (1973) the proportion of methylation on ribose to that on the bases was considerably lower (about 40% of the total) than that reported here (about 90% of the total). Since the present findings are based on the results of mild enzymic-hydrolysis procedures we have confidence in our numerical data. It seems likely that acid hydrolysis, which degrades 2′-*O*-methylated compounds to nucleosides or free 2′-*O*-methylribose, leads in some way to underestimation of the amount of ribose methylation. Nevertheless acid hydrolysis, like T_2-ribonuclease hydrolysis, conserves the various alkali-labile *N*-methylated purines and pyrimidines, and Klagsbrun's (1973) main findings on these compounds are in qualitative agreement with ours.

Concluding remarks

Because the data from the two 'fingerprinting' systems are essentially in complete agreement, we are confident that the information in Tables 1 and 2 gives an accurate summary of the methylated components in HeLa-cell rRNA. Comparing the two sets of data served to focus attention on the remaining uncertainties, particularly those concerning methylated bases, and most of these uncertainties appear now to have been resolved. We believe that complete knowledge of the inter-relationships between the various T_1-ribonuclease and combined T_1-plus-pancreatic-ribonuclease digestion products will be useful for further studies on the methylated sequences in rRNA, for example in comparative studies and in conformational analyses of the methylated sequences, since both 'fingerprinting' systems are useful for such studies in our laboratory.

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

References

Brownlee, G. G. (1972) *Determination of Sequences in RNA*, North-Holland/Elsevier, Amsterdam

Iwanami, Y. & Brown, G. M. (1968) *Arch. Biochem. Biophys.* **126**, 8–15

Khan, M. S. N. & Maden, B. E. H. (1976) *J. Mol. Biol.* **101**, 235–254

Klagsbrun, M. (1973) *J. Biol. Chem.* **248**, 2612–2620

Klootwijk, J. & Planta, R. J. (1974) *Mol. Biol. Rep.* **1**, 187–191

Macon, J. B. & Wolfenden, R. (1968) *Biochemistry* **7**, 3453–3458

Maden, B. E. H. & Salim, M. (1974) *J. Mol. Biol.* **88**, 133–164

Maden, B. E. H., Lees, C. D. & Salim, M. (1972) *FEBS Lett.* **28**, 293–296

Maden, B. E. H., Forbes, J., de Jonge, P. & Klootwijk, J. (1975) *FEBS Lett.* **59**, 60–63

Nazar, R. N., Sitz, T. O. & Busch, H. (1975) *FEBS Lett.* **59**, 83–87

Salim, M. (1972) Ph.D. Thesis, University of Glasgow

Saponara, A. G. & Enger, M. D. (1974) *Biochim. Biophys. Acta* **349**, 61–77

Wagner, E. K., Penman, S. & Ingram, V. I. (1967) *J. Mol. Biol.* **29**, 371–387