

MAMMALIAN RIBOSOMAL PRECURSOR RNA

by

James Searl Robertson

Thesis presented for the degree of  
Doctor of Philosophy,  
The University of Glasgow,

June, 1974.

ProQuest Number: 11018028

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11018028

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

To my parents.

ACKNOWLEDGEMENTS

I am indebted to Dr. B.E.H. Maden, my supervisor for his guidance and advice during the three years of this project and for financial support from M.R.C. through a grant to Dr. Maden.

I am grateful to Professor R.M.S. Smellie for providing the facilities of the Department of Biochemistry.

Also, I wish to thank Mrs. Helen MacDowall for the high standard of typing in this thesis.

ABBREVIATIONS

rRNA	ribosomal RNA
rpre RNA	ribosomal precursor RNA
HnRNA	heterogeneous nuclear RNA
rDNA	ribosomal DNA
RNase	ribonuclease
DNase	deoxyribonuclease
VPDE	venom phosphodiesterase
CMCT	N-cyclohexyl-N'-( $\beta$ -morpholinyl-(4)-ethyl) carbodi-imide-methyl-p-toluene sulphonate
DEAE	diethylaminoethylcellulose
EDTA	ethylenediaminetetra-acetate

CONTENTS

	<u>Page</u>
Title	i
Acknowledgements	iii
Abbreviations	iv
Contents	v
Summary	viii
<u>Chapter 1 INTRODUCTION</u>	
1.1 Mammalian Ribosomal RNA	1
1.2 Ribosomal Precursor RNA	3
1.3 Non-Ribosomal Sequences	4
1.4 5.8S RNA	7
1.5 Aims of Project	7
<u>Chapter 2 MATERIALS AND METHODS</u>	
MATERIALS	
Cells, Media and Solutions	10
Chemicals	11
METHODS	
2.1 Growth, Radioactive Labelling and Harvesting of Cells	13
(i) Cell growth	13
(ii) Radioactive labelling	14
(iii) Cell harvesting	15

	<u>Page</u>
2.2 Purification of RNA	16
(i) Cell fractionation	16
(ii) Nucleolar RNA	18
(iii) Cytoplasmic RNA	19
2.3 RNA Fingerprinting Technique	20
(i) Enzymic digestion of RNA	20
(ii) 2-Dimensional fractionation	22
(iii) Autoradiography	24
2.4 Sequence Analysis	24
(i) Radioactivity estimation	25
(ii) Elution	25
(iii) Base composition analysis	26
(iv) Pancreatic RNase	26
(v) Snake venom phosphodiesterase	27
(vi) Reaction with CMCT	28
(vii) Analysis for pseudouridylic acid	29

### Chapter 3 TRANSCRIBED SPACER SEQUENCES

3.1 Purification of 32S RNA by an "Actinomycin Chase"	30
3.2 Fingerprints of 28S and 32S RNA	31
3.3 Analysis of 32S T <sub>1</sub> Products 1-14	34
3.4 Purification of 45S RNA	38
3.5 Fingerprints of 28S+ 18S and 45S RNA	39
3.6 Analysis of 45S T <sub>1</sub> Products 20-28.	41

	<u>Page</u>
<u>Chapter 4 THE 5.8S SEQUENCE IN RIBOSOMAL PRECURSOR RNA</u>	
4.1 Fingerprints of 5.8S RNA	43
4.2 The 5.8S Sequence in 28S RNA	44
4.3 The 5.8S Sequence in 32S RNA	45
4.4 The 5.8S Sequence in 45S RNA	48
<u>Chapter 5 COMPARISON OF HELA AND L CELL TRANSCRIBED SPACER SEQUENCES</u>	
5.1 Fingerprints of HeLa and L cell 28S and 5.8S RNA	49
5.2 Fingerprints of HeLa and L cell 32S RNA	50
<u>Chapter 6 DISCUSSION</u>	
6.1 Practical Considerations	52
6.2 Inferences on HeLa Cell Transcribed Spacer Regions	56
6.3 The 5.8S Sequence	58
6.4 Precursor RNA from Other Eukaryotes	59
6.5 Precursor RNA from Bacteria	61
6.6 Topological Map of 45S RNA	62
6.7 Secondary Modification of Precursor RNA	65
6.8 Secondary Structure of Precursor RNA	68
6.9 Nucleolar Proteins and Nucleases	71
6.10 Future Experiments	74
<u>REFERENCES</u>	76
<u>APPENDIX</u>	82

SUMMARY

Both 18S and 28S ribosomal RNA in HeLa cells are transcribed in the nucleolus within a single large precursor RNA molecule (45S RNA) which is believed to contain, in addition to the ribosomal sequences, extensive non-ribosomal regions termed "transcribed spacers". Substantial segments of the transcribed spacer region are retained in the various nucleolar intermediates in maturation (41S, 32S and 20S RNA).

Presented in this thesis are the results obtained from a direct analysis of transcribed spacer using RNA fingerprinting techniques with  $^{32}\text{P}$ -labelled RNA.

1. By comparing fingerprints of 28S RNA with its immediate precursor 32S RNA, I identified and characterised some 13 oligonucleotides which are present in 32S RNA only and which thus must be derived from the transcribed spacer region of 32S RNA.
2. Fingerprints of 45S RNA were compared with those of a mixture of 28S + 18S RNA and 9 distinctive transcribed spacer oligonucleotides were identified in addition to those from the 32S molecule.
3. By identifying specific 5.8S oligonucleotides in fingerprints of 32S and 45S RNA I was able to show that 5.8S RNA is transcribed as part of the 45S ribosomal precursor RNA molecule.
4. The primary sequence of transcribed spacer from HeLa ribosomal precursor RNA and L cell ribosomal precursor RNA

was compared and found to be less highly conserved than the ribosomal sequences.

All the above mentioned oligonucleotides are of potential value for further structural analysis and for determining the spacial arrangement of transcribed spacer and ribosomal sequences in 45S ribosomal precursor RNA.

## Chapter 1

### INTRODUCTION

## INTRODUCTION

It is well known that the two principal RNA components of mammalian ribosomes are transcribed initially within a single large precursor molecule in the nucleolus. Maturation of this precursor RNA is a multistep process proceeding through various nucleolar RNA intermediates, and with the apparent loss of substantial segments from the precursor molecules (Maden, 1971, review). This type of maturation process for ribosomal RNA is common to all eukaryotic species so far studied (Grierson et al, 1970). In bacteria there is normally no common precursor molecule to ribosomal RNA, the two major components usually appearing as slightly larger though separate molecules (Adesnik & Levinthal, 1969; Dahlberg & Peacock, 1971). (However this point will be discussed again in the Discussion).

### 1.1 Mammalian Ribosomal RNA

The structure and formation of mammalian ribosomal RNA (rRNA) and ribosomes have been extensively reviewed (Attardi, 1970; Maden, 1971; Burdon, 1971).

A convenient system for the investigation of mammalian ribosomes is provided by HeLa cells in cell culture, and a considerable amount of information has been obtained from their use. This information is broadly in agreement with that obtained from other sources, in both intact animals and other cell culture systems and is likely to be representative, at least in outline, of mammalian systems in general.

HeLa cell rRNA possesses a molecular weight of  $0.70 \times 10^6$  daltons (18S) and  $1.75 \times 10^6$  daltons (28S) (Loening, 1968) (the nomenclature of rRNA and its precursors is based on the approximate sedimentation constant of the RNA). 18S and 28S RNA are methylated (Wagner et al, 1967; Vaughan et al, 1967), contain pseudouridylate residues and have a G+C content of 59% and 67% respectively (Amaldi & Attardi, 1968) (table 1.1). Table 1.2 illustrates the difference in molecular weights of rRNA from different sources. The chemical composition of rRNA varies from 40% G+C in Drosophila rRNA (Rotissa et al, 1966) to the high G+C content of mammalian rRNA.

In addition to the two high molecular weight RNAs, there are two low molecular weight RNAs associated with the large ribosomal subunit. 5S RNA has been identified in a wide variety of organisms including bacteria (Rosset & Monier, 1963; Galibert et al, 1965). A second low molecular weight species can be released from 28S rRNA by treatments which disrupt hydrogen bonding (Pene et al, 1968). This RNA was originally termed 7S RNA but is now recognised as being approximately 5.5S to 5.8S (Weinberg & Penman, 1968; Rubin, 1973). In this thesis it will be described as 5.8S RNA. It is present in eukaryotes generally but not in prokaryotes (Pene et al, 1968) and will be discussed in more detail later.

Table 1.1 Chemical and Physical Properties of HeLa Cell Ribosomal RNA and its Precursors.

	45S	32S	28S	18S	Reference
G+C content (%)	71	70	67	58	(1)
	70	70	67	56	(2)
Ψ content (residues per thousand nucleotides)	8	9	11	15	(1)
Methyl/phosphate ratio	0.363	0.485	0.605	0.890	(3)
Molecular weight x 10 <sup>-6</sup>					
(a) Calculated from sedimentation velocity	4.6	2.2	-	-	(1)
(b) Sedimentation equilibrium	4-4.5	2.4	1.9	0.71	(4)
(c) Acrylamide gel electrophoresis	4.1	2.1	(1.65)*	0.65	(3)

References:

- (1) Amaldi & Attardi, 1968; Jeanteur et al, 1968.
- (2) Willems et al, 1968.
- (3) Weinberg & Penman, 1970.
- (4) McConkey & Hopkins, 1969.

\* Value of Petermann & Pavlovec (1966) used for calibration.

Reproduced from Maden (1971).

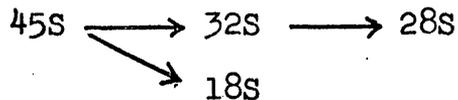
Table 1.2 Molecular Weights of Ribosomal RNA from Various Sources.

Species	Molecular Weight ( $\times 10^{-6}$ )	
<b>Animals</b>		
HeLa	1.75	0.70
Mouse (liver)	1.71	0.70
Rabbit (reticulocytes)	1.72	0.70
<u>Xenopus</u> (liver)	1.54	0.69
<u>Drosophila</u>	1.41	0.73
<b>Plants and protozoa</b>		
Pea, bean, raddish, corn	1.27-1.31	0.70-0.71
<u>Saccharomyces</u>	1.30	0.72
<u>Tetrahymena</u>	1.30	0.69
<b>Prokaryotes</b>		
<u>E. coli</u>	1.07	0.56
<u>Streptomyces</u>	1.11	0.56

Adapted from Loening (1968).

## 1.2 Ribosomal Precursor RNA

On labelling of HeLa cells with radioactive uridine or with methyl-labelled methionine, a flow of radioactivity from nucleus to cytoplasm occurs involving the following RNA species (Scherrer & Darnell, 1962; Scherrer et al, 1963; Penman, 1966; Penman et al, 1966; Girard et al, 1964; Greenberg & Penman, 1966; Zimmerman & Holler, 1967):-



45S and 32S RNA are found exclusively in the nucleolus whilst 18S and 28S RNA pass rapidly into the cytoplasm (Penman, 1966; Penman et al, 1966; Greenberg & Penman, 1966). These observations led to the view that 45S RNA is a precursor to ribosomal RNA and gives rise to the latter as indicated in the above scheme.

Pulse-labelling with methyl-labelled methionine also suggested that methylation of ribosomal RNA takes place almost entirely on the 45S molecule either during or immediately after transcription (Greenberg & Penman, 1966; Zimmerman & Holler, 1967). In support of this precursor-product relationship between nucleolar and ribosomal RNA, it was shown that the 2'-O-methylation pattern of 32S RNA was similar to that of 28S rRNA and that of 45S RNA was similar to that of an equimolar mixture of 18S and 28S rRNA (Wagner et al, 1967).

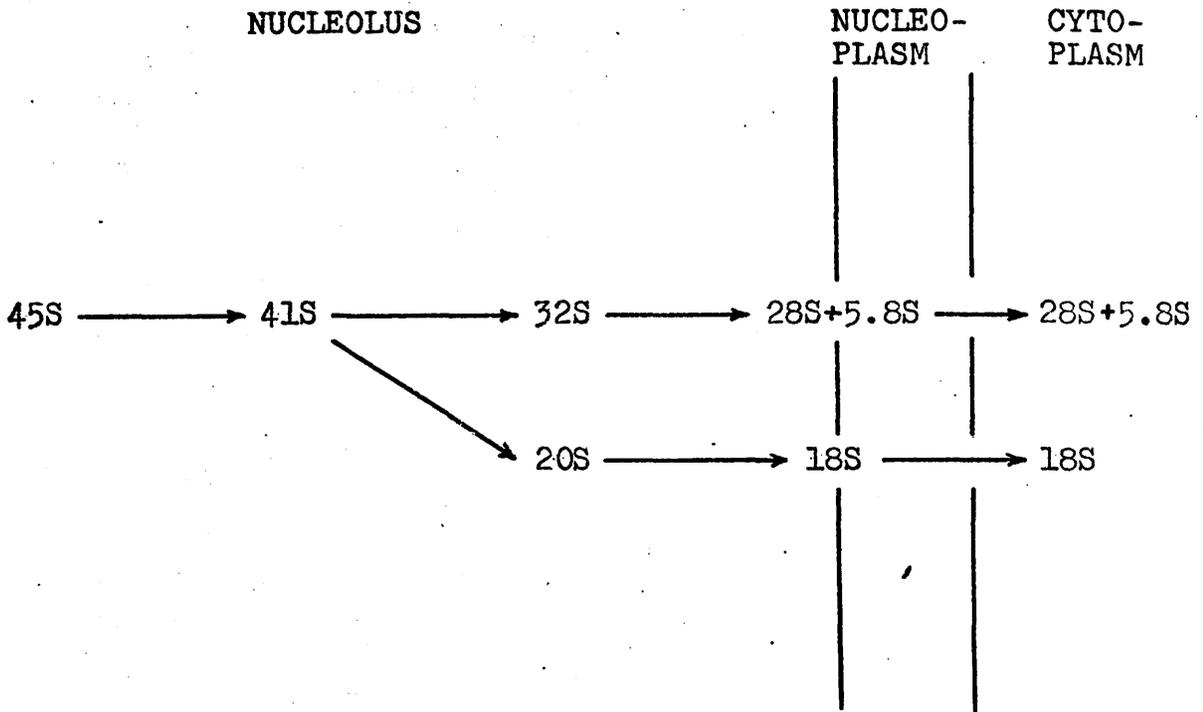


Fig. 1.1. Maturation of 45S rpre RNA.

Analysis of nucleolar RNA by polyacrylamide gel electrophoresis has revealed the presence of two short-lived RNA species, 41S and 20S RNA, in addition to the previously recognised 45S, 32S, 28S and 18S species of nucleolar RNA (Weinberg et al, 1967). Estimates of molecular weights and relative levels of methylation of the nucleolar RNA molecules were used to indicate the probable position of each RNA in a processing scheme from 45S to 18S and 28S rRNA (Weinberg & Penman, 1970). The resulting maturation scheme, illustrated in fig. 1.1, has now been confirmed by "fingerprinting" methyl-labelled nucleolar and ribosomal RNA species (Maden et al, 1972a).

Ribosomal precursor RNA molecules can be extracted from the nucleolus in the form of ribonucleoprotein particles (Warner & Soeiro, 1967; Liao & Perry, 1969; Craig & Perry, 1969; Pederson & Kumar, 1971). These particles appear by polyacrylamide gel electrophoresis of their proteins, and by "fingerprinting" tryptic digests of their proteins after radioactive labelling, to be precursors of cytoplasmic ribosomes (Warner & Soeiro, 1967; Shepherd & Maden, 1972). This implies that ribosome formation takes place in the nucleolus at the level of 45S and 32S RNA.

### 1.3 Non-Ribosomal Sequences (Transcribed Spacers)

Several lines of evidence indicate that about half of the 45S RNA molecule is discarded during maturation to 18S and 28S ribosomal RNA.

(i) Although nearly all (~95%) of the methyl group addition takes place on 45S RNA, the relative level of methylation increases two-fold during the conversion of 45S to 18S and 28S ribosomal RNA (Vaughan et al, 1967; Weinberg et al, 1967; Weinberg & Penman, 1970). This implies that a substantial amount of non-methylated material is removed during maturation.

(ii) The base composition of ribosomal RNA (rRNA) and ribosomal precursor RNA (rpreRNA) was determined by two separate groups, who employed different methods to prepare nucleolar RNA free from contaminating heterogeneous nuclear RNA (Amaldi & Attardi, 1968; Jeanteur et al, 1968; Willems et al, 1968). Both groups found that rRNA (weighted mean of 28S + 18S) possessed a G+C content of 65% and that rpre RNA possessed a G+C content of 70%. This implies the presence of non-ribosomal material of very high (77%) G+C content in 32S and 45S RNA. Attardi's group also established that 32S and 45S RNA remain structurally intact after heating and other types of denaturation, thus strongly suggesting that the polynucleotide chains are covalently continuous (Jeanteur et al, 1968).

(iii) The molecular weights of HeLa rRNA and rpre RNA were estimated by sedimentation velocity (Jeanteur et al, 1968), sedimentation equilibrium (McConkey & Hopkins, 1969) and by polyacrylamide gel electrophoresis

(Weinberg & Penman, 1970) and are in reasonably good agreement with each other (table 1.1). The values obtained with sedimentation velocity are tentative approximations because of the unknown contribution of the conformation of the RNA molecules to the sedimentation rate. The values obtained with polyacrylamide gel electrophoresis are dependent on the accuracy of the previously determined molecular weights of 18S and 28S RNA used as standards. Molecular weight determination by sedimentation equilibrium is reasonably accurate and has indicated that the molecular weight of 32S RNA ( $\sim 2.4 \times 10^6$  daltons) is larger than that of 28S RNA ( $\sim 1.9 \times 10^6$  daltons) and that the molecular weight of 45S RNA ( $\sim 4.4 \times 10^6$  daltons) exceeds by about  $10^6$  daltons the sum of the molecular weights of 32S RNA and 18S RNA ( $\sim 0.7 \times 10^6$  daltons). This implies that about one quarter of the 32S molecule and one third to one half of the 45S molecule are lost during maturation to 18S and 28S rRNA.

(iv) The presence of non-ribosomal sequences in rpre RNA is also implied by competitive RNA-DNA hybridisation experiments (Jeanteur & Attardi, 1969). 45S RNA competed with 28S and 18S rRNA for sites in DNA to the extent expected if about 35% and 13% respectively of the precursor molecule were involved. Similarly, 32S RNA competed with 28S RNA to the extent expected if about 70% of the molecule was involved.

Ribosomal DNA (rDNA) is composed of two repeating nucleotide sequences; a sequence coding for the rpre RNA molecule alternating with a "non-transcribed spacer" sequence of unknown function (Dawid et al, 1970; Reeder & Brown, 1970). As indicated above, rpre RNA is believed to contain non-ribosomal sequences which are discarded during maturation. These presumed sequences have been termed "transcribed spacers" (Brown et al, 1972) to distinguish them from the non-transcribed spacers of rDNA. This term will be used in this thesis.

#### 1.4 5.8S RNA

5.8S RNA is a small (~150 nucleotides) non-covalently bound fragment of RNA which can be released from eukaryotic 28S RNA by briefly heating or otherwise denaturing it. When radioactive uridine is incorporated into growing cells, the 5.8S and major 28S RNA sequences are labelled at the same rate (Pene et al, 1968; Udem & Warner, 1972). Also, 5.8S RNA cannot be released from 32S or 45S RNA by methods which normally disrupt hydrogen bonding (Pene et al, 1968; Udem & Warner, 1972). These experiments imply, but do not prove, that 5.8S RNA is covalently linked within rpre RNA and is generated as a separate entity by nucleolytic cleavage during the final 32S → 28S maturation step.

#### 1.5 Aims of Project

Summarising all of the above findings, HeLa cell 45S RNA appears to contain the sequences of 28S, 18S and possibly 5.8S rRNA together with 40-50% of its length in

the form of transcribed spacers. 32S RNA appears to contain the sequence of 28S rRNA and possibly 5.8S RNA together with 30% of its length as transcribed spacer. However the evidence relating to transcribed spacers and 5.8S RNA is indirect, deriving mainly from differences in the physical and chemical properties of rpre RNA and rRNA. (This evidence implies that HeLa cell transcribed spacer has a high G+C content, is lacking methyl groups and is low or lacking pseudouridylic acid). The evidence for the origin of 5.8S RNA was also indirect, deriving from kinetic data.

This thesis describes firstly the direct identification and characterisation of several distinctive transcribed spacer oligonucleotides derived from HeLa cell 32S and 45S rpre RNA. These oligonucleotides were isolated from the products of complete endonucleolytic digestion of HeLa cell rpre RNA. This was accomplished by comparing "fingerprints" of rRNA with those of rpre RNA and searching for products present only in the rpre RNA fingerprints.

Furthermore by combining the fingerprinting study of 28S rRNA and 32S rpre RNA with that of 5.8S RNA, definitive evidence was obtained that the 5.8S sequence is present as such in HeLa cell rpre RNA and is transcribed as part of the 45S molecule.

Finally, having obtained partial sequence information on HeLa cell transcribed spacers it was of interest to compare this information with that which can be obtained from another species. HeLa cell rpre RNA was compared with mouse

L cell rpre RNA. The results indicated that the transcribed spacer sequences have been less highly conserved than the ribosomal sequences.

## Chapter 2

### MATERIALS AND METHODS

## Chapter 2

### MATERIALS AND METHODS

Table 2.1. Eagle's Minimum Essential Medium (modified)

<u>AMINO ACIDS</u>	mg/1	<u>VITAMINS</u>	mg/1
L-arginine HCl	42.1	D-Calcium pantothenate	2.0
L-cystine	24.0	Chlorine chloride	2.0
L-glutamine	292.0	Folic Acid	2.0
L-Histidine HCl	19.2	i-Inositol	4.0
L-isoleucine	52.5	Nicotinamide	2.0
L-leucine	52.5	Pyridoxal HCl	2.0
L-lysine HCl	73.1	Riboflavin	0.2
L-Methionine	14.9	Thiamin HCl	2.0
L-Phenylalanine	33.0		
L-Threonine	47.6		
L-Tryptophan	8.2		
L-Tyrosine	36.2		
L-Valine	46.9		

INORGANIC SALTS AND OTHER COMPONENTS

CaCl <sub>2</sub> ·2H <sub>2</sub> O	264.0	NaCl	6400.0
Dextrose	4500.0	NaH <sub>2</sub> PO <sub>4</sub> ·2H <sub>2</sub> O	140.0
Fe(NO <sub>3</sub> ) <sub>3</sub> ·9H <sub>2</sub> O	0.1	Phenol Red	17.0
KCl	400.0	NaHCO <sub>3</sub>	2750.0
MgSO <sub>4</sub> ·7H <sub>2</sub> O	200.0		
Penicillin	10 <sup>5</sup> units		
Streptomycin SO <sub>4</sub>	10 <sup>5</sup> µg		

## MATERIALS AND METHODS

### MATERIALS

#### Cells, Media and Solutions.

##### Cells.

HeLa monolayer (Gey et al, 1952) and L929 (Sanford et al, 1948) cell lines were used.

##### Eagle's Minimal Essential Medium (MEM)

HeLa and L929 cell lines were grown as monolayers in slightly modified Eagle's MEM (Eagle, 1959) with the following additions:- calf serum to 10% (v/v), penicillin (100 units/ml) and streptomycin sulphate (100 µg/ml). The constituents of this medium are shown in table 2.1.

During labelling with  $^{32}\text{P}_i$ , cells were grown in low phosphate ( $1.8 \times 10^{-4}\text{M}$ ) medium which contained one-fifth the standard phosphate concentration of the above medium.

Amino acids, vitamins and calf serum were purchased from Biocult Laboratories Ltd., Paisley, Scotland.

##### Balanced Salts Solution (BSS : Earle, 1943)

Contained 0.116 M NaCl, 5.4 mM KCl, 1 mM MgSO<sub>4</sub>, 1 mM NaH<sub>2</sub>PO<sub>4</sub>, 1.8 mM CaCl<sub>2</sub> and 0.002% (w/v) phenol red. The pH of this solution was adjusted to 7.0 with 5.6% (w/v) NaHCO<sub>3</sub>.

##### Versene/Trypsin.

Versene 4 parts, trypsin/citrate 1 part (v/v).

Versene.

0.6 mM EDTA in PBS(A).

Phosphate Buffered Saline (A). (PBS(A)).

This is calcium- and magnesium-free PBS. It consists of 0.17 M NaCl, 3.4 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> and 2.4 mM KH<sub>2</sub>PO<sub>4</sub>.

Trypsin/Citrate.

0.25% (w/v) trypsin, 10.5 mM NaCl, 1.0 mM sodium citrate and 0.002% phenol red adjusted to pH 7.8 with NaOH.

RSB (Reticulocyte Salt Buffer)

0.01 M NaCl, 0.0015 M MgCl<sub>2</sub>, 0.01 M tris-HCl, pH 7.4.

HSB (High Salt Buffer)

0.5 M NaCl, 0.05 M MgCl<sub>2</sub>, 0.01 M tris-HCl, pH 7.4.

SUP (Sodium dodecyl sulphate - Urea - Phosphate buffer)

0.5% (w/v) SDS, 0.5 M Urea, 0.01 M Na phosphate, pH 7.0.

LETS (Lithium - EDTA - Tris - Sodium dodecyl sulphate buffer)

0.1 M LiCl, 0.01 M EDTA, 0.01 M tris-HCl, 0.2% SDS, pH 7.4.

Chemicals.

Actinomycin D (200 µg/vial) was obtained from Calbiochem Ltd., London.

<sup>32</sup>P-orthophosphate (<sup>32</sup>Pi; 10 mCi/ml) was supplied in solution in dilute hydrochloric acid by The Radiochemical Centre, Amersham.

Ribonuclease T<sub>1</sub> and pancreatic ribonuclease A were purchased from Calbiochem Ltd., London. Bacterial alkaline

phosphatase and snake venom phosphodiesterase were obtained from the Worthington Biochemical Corporation through Cambrian Chemicals Ltd., London.

The carbodi-imide, N-cyclohexyl-N'-( $\beta$ -morpholinyl-(4)-ethyl) carbodi-imide-methyl-p-toluene sulphonate (CMCT), was a gift from Dr. G.G. Brownlee, Cambridge.

Cellulose acetate electrophoresis strips (25 x 95 cm) were obtained from Oxoid Ltd., London. Whatman papers numbers 52(46 x 57 cm), 3 MM (46 x 57 cm) and DE 81 (DEAE-paper, 46 cm x 50 m) were obtained from H. Reeve-Angel and Co. Ltd., London. Electrophoresis dye mixture contains equal volumes of the following components supplied by G.T. Gurr Ltd., London: 1% xylene cyanol F.F. (blue); 2% orange G (yellow); 1% acid fuschin (pink).

Kodirex KD 54T (35 x 43 cm) X-ray film, DX-80 developer and FX-40 X-ray liquid fixer were obtained from Kodak Ltd.

Liquid scintillation fluid was prepared by dissolving 5 g PPO (2,5-diphenyloxazole) in one litre of toluene.

## METHODS

### 2.1 Growth, Radioactive-labelling and Harvesting of Cells.

#### (1) Cell growth.

HeLa and L929 cell lines were cultured routinely as monolayers in rotating "burlers" (80 oz winchester bottles) according to the method of House and Wildy (1965).  $25 \times 10^6$  cells were seeded with 200 ml of medium in a burler and incubated at  $37^\circ\text{C}$  for 3-4 days in a 5%  $\text{CO}_2$  atmosphere until the cell density had reached approximately  $150\text{-}200 \times 10^6/\text{burler}$ .

For serial passaging, the cells were removed from the glass by trypsin/versene (see later), suspended in medium and dispersed in  $25 \times 10^6$  cell aliquots into sterile burlers containing 200 ml medium. Cell cultures were maintained and contamination checks were carried out by the staff of the Wellcome Cell Culture Unit.

#### Contamination checks

All sterile media and passaged cells were checked regularly for bacterial, fungal or PPLO infection as follows:-

#### Bacterial contamination

Aliquots were grown on blood agar plates and brain-heart infusion broth at  $37^\circ\text{C}$ . Results were considered negative if no growth was seen within 7 days.

### Fungal contamination

Aliquots were added to Sabourand's medium and grown at 32°C. Again, no growth after 7 days was assumed to indicate the absence of mycotic contamination.

### PPLO infection

PPLO agar plates were seeded with passaged cells by piercing the agar surface with a charged Pasteur pipette. The plates were grown in an atmosphere of 5% CO<sub>2</sub> in N<sub>2</sub> at 37°C. To test the efficiency of the method, PPLO infected cultures were prepared as controls. 7 days after inoculation, the plates were examined microscopically for the characteristic "fried egg" appearance of PPLO colonies.

### (ii) Radioactive labelling.

To prepare <sup>32</sup>P-labelled RNA, 45 x 10<sup>6</sup> cells were seeded into 200 ml of medium in a burler and incubated at 37°C. After 30h the cells had settled and reached a density of about 50 x 10<sup>6</sup>/burler. The regular medium was poured off, the cells rinsed once with "low phosphate" medium and 50 ml of "low phosphate" medium were added. Reducing the concentration of unlabelled phosphate during labelling aided the uptake of radioactive phosphate by the cells.

10 mCi of <sup>32</sup>Pi were added and after 18h of exponential growth the cells were harvested having reached a density of approximately 100 x 10<sup>6</sup>/burler.

The rate of synthesis of 45S RNA is affected by metabolic deficiencies (Vaughan et al, 1967; Maden et al, 1969) and better yields were obtained when 10 ml of fresh "low phosphate" medium were added 2-3h prior to harvesting. For the preparation of pure 32S nucleolar RNA, actinomycin D (4 µg/ml) was added 30-40 min before harvesting. This inhibited the synthesis of 45S RNA, allowed existing 45S RNA to mature to smaller RNA species and left none to contaminate 32S RNA during extraction (see Results).

Although nucleolar RNA is labelled within a few hours, labelling was continued for 18h so that <sup>32</sup>P-labelled cytoplasmic RNA could also be prepared from the cells. In addition, after the longer time interval, the specific activity of the intracellular nucleotide pool should be higher. In fact the specific activity of nucleolar RNA extracted after 18h of labelling was much higher than after only 3-6h of labelling.

### (iii) Cell harvesting.

The radioactive medium was carefully poured off and the cells were washed twice with 20 ml of trypsin/versene.

A few ml of trypsin/versene were left in the burler after the second wash and after several minutes at 37°C the cells were seen to dislodge from the glass. They were taken up in 40 ml of ice cold balanced salts solution

containing 10% (v/v) calf serum to neutralise tryptic activity, and transferred to a 50 ml tube on ice. The cells were pelleted at 2,000 rpm for 2 min in an MSE Mistral 4L Centrifuge at 4°C. They were then washed thoroughly (twice) by vigorously resuspending the cell pellet in 25 ml ice cold balanced salts solution and centrifuging as above.

## 2.2 Purification of RNA.

### (1) Cell fractionation.

HeLa cells were fractionated into cytoplasm, nucleoplasm and nucleoli as described by Penman, 1969. All steps were carried out at 4°C.

The cells obtained from one burler (approximately  $10^8$ ), after washing with BSS and pelleting as described above, were carefully drained of residual BSS and suspended in 4.0 ml of hypotonic RSB solution. They were allowed to swell for 10 min and disrupted by homogenising with 20 strokes of a stainless steel Dounce homogeniser (with a clearance of 0.003"). The cell homogenate was transferred to a 15 ml centrifuge tube and centrifuged for 2 min at 2,000 rpm in an MSE Mistral 4L centrifuge.

The cytoplasm (supernatant) was removed with a Pasteur pipette and was used for the preparation of cytoplasmic RNA.

The pelleted nuclei were resuspended in 4.0 ml of RSB, 0.6 ml of a mixed detergent solution was added and the mixture vortexed for 10 sec. The detergent solution

consists of two parts 10% Tween 80 to one part 10% sodium deoxycholate. This step was to remove the outer nuclear membrane and contaminating cytoplasm and will disrupt any remaining intact cells.

The nuclear suspension was centrifuged at 2,000 rpm for 2 min and the supernatant discarded. The surface of the pellet was rinsed with RSB to remove traces of detergent which may inhibit the following enzymic digestion step.

The nuclear pellet was resuspended in 2.0 ml of HSB containing 100  $\mu$ g DNase. At this point, the nuclei burst and the suspension became viscous with DNA clumps. The solution was incubated at 37°C with vigorous pipetting with a wide bore Pasteur pipette until all the DNA clumps were dispersed and the viscosity of the solution was considerably reduced (usually 1 min incubation was sufficient). The digest was quickly cooled in ice and layered on top of a chilled 16 ml 15-30% sucrose-HSB gradient in the small bucket of a Beckman SW 27 rotor and centrifuged at 17,000 rpm for 15 min at 4°C. The nucleoli were deposited on the bottom of the tube and the nucleoplasmic components were distributed through the sucrose gradient.

Fractionation of L929 cells was essentially identical to the above procedure with the following exception. It was necessary to swell L929 cells in RSB diluted 1:3 with distilled water (to make it more

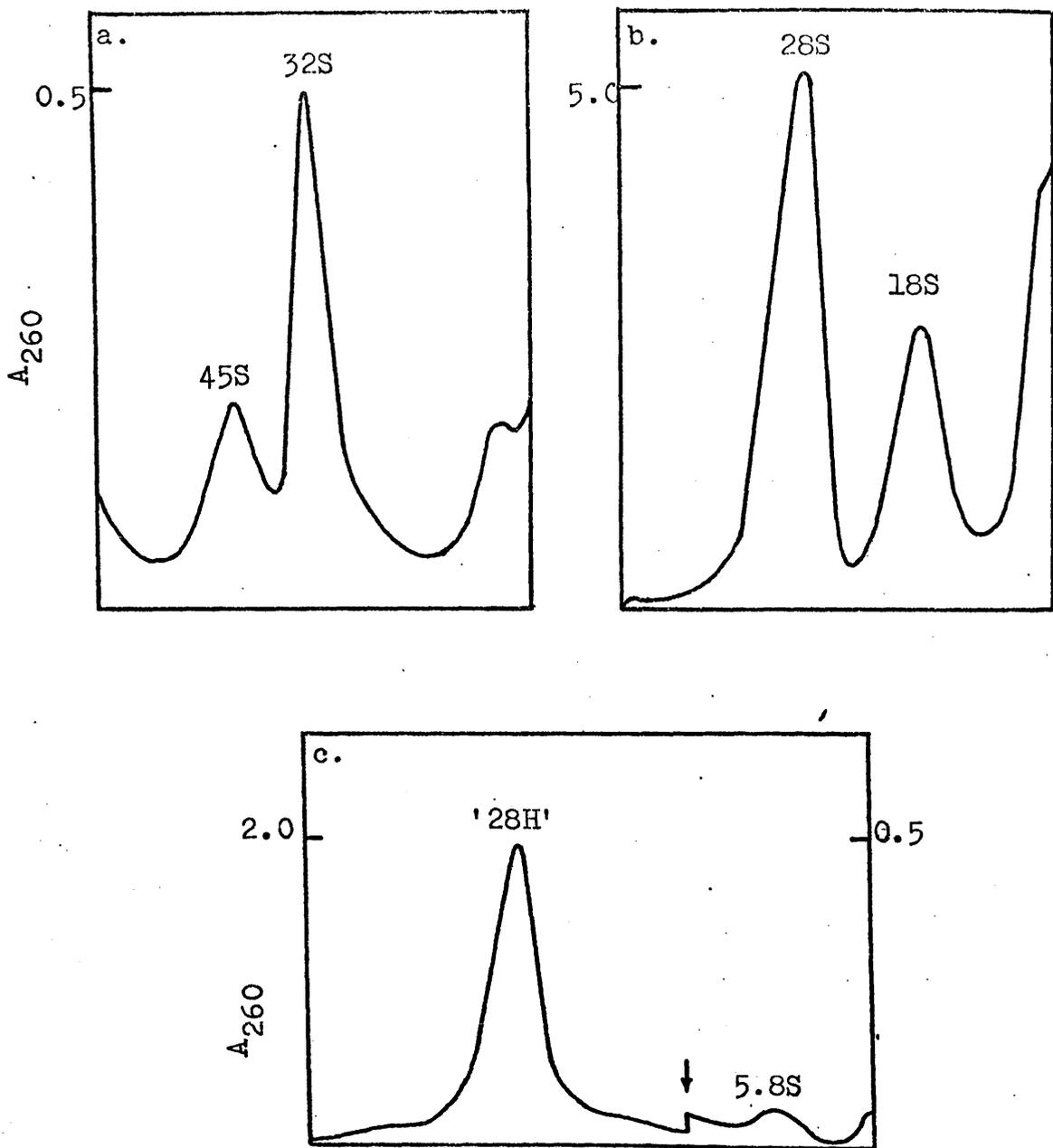


Fig. 2.1. OD profiles of nucleolar RNA, cytoplasmic RNA and heat-treated 28S RNA after centrifugation through sucrose gradients. (a) nucleolar RNA, 15-30% sucrose-LETS, (b) cytoplasmic RNA, 10-25% sucrose-LETS, (c) heat-treated 28S RNA, 10-25% sucrose-LETS. In (c) the sensitivity of the A<sub>260</sub> recording was increased fourfold to the right of the arrow.

hypotonic) before the cells would rupture during homogenisation.

(11) Nucleolar RNA.

Nucleolar RNA was liberated from the nucleolar pellet by resuspending in 1.0 ml of SUP (Soeiro et al, 1968) at room temperature with occasional vortexing and the use of a glass rod. The process took at least half an hour. The nucleolar RNA was layered on a 37 ml 15-30% sucrose-LETS gradient in the large bucket of a Beckman SW27 rotor and centrifuged at 22,000 rpm for 16h at 20°C. It was convenient to centrifuge cytoplasmic RNA on the same run and optimal resolution could be obtained using a 10-25% sucrose-LETS gradient for the latter.

The gradients were monitored at 260 nm through the flow cell of a Gilford spectrophotometer (Fig. 2.1a). The RNA peak fractions were precipitated twice with ethanol to eliminate sodium dodecyl sulphate, redissolved in a small volume of water and the specific activity of the RNA was measured as follows:- the concentration of RNA was measured in the spectrophotometer at 260 nm assuming a relationship of  $E \frac{1 \text{ mg/ml}}{1 \text{ cm}} = 25$ , and the radioactivity was measured by liquid scintillation counting of a 10  $\mu$ l sample dried on a cellulose nitrate filter.  $10^8$  HeLa cells generally yielded 40-50  $\mu$ g of 32S RNA containing 40  $\mu$ Ci  $^{32}\text{P}$ , and 15-20  $\mu$ g of 45S

RNA containing 15  $\mu\text{Ci}$   $^{32}\text{P}$ . L929 cells yielded smaller amounts.

(iii) Cytoplasmic RNA.

Cytoplasmic RNA was obtained by extracting the cytoplasm with phenol and sodium dodecyl sulphate at room temperature.

The cytoplasm obtained from the cell fractionation was transferred from ice to room temperature and sodium dodecyl sulphate was added to 0.5%. An equal volume of phenol was added, the mixture was vortexed for a few minutes and then centrifuged at 3,500 rpm for 20 min in the MSE Mistral 4L centrifuge at room temperature. The aqueous layer was removed and the RNA was precipitated with  $2\frac{1}{2}$  volumes of ethanol.

Cytoplasmic RNA extracted from  $10^8$  cells was dissolved in 2.0 ml of LETS buffer and divided between two 37 ml 10-25% sucrose-LETS gradients and centrifuged in parallel with 15-30% gradients containing nucleolar RNA in the SW27 rotor at 22,000 rpm for 16h at  $20^\circ\text{C}$ . The gradients were monitored at 260 nm through the flow cell of a Gilford spectrophotometer (Fig. 2.1b) and 28S and 18S ribosomal RNA were precipitated twice with ethanol to remove sodium dodecyl sulphate. The specific activities were measured as described above for nucleolar RNA and the RNA was stored at  $-20^\circ\text{C}$  until use.  $10^8$  cells generally yielded about 500  $\mu\text{g}$  of 28S RNA of lower

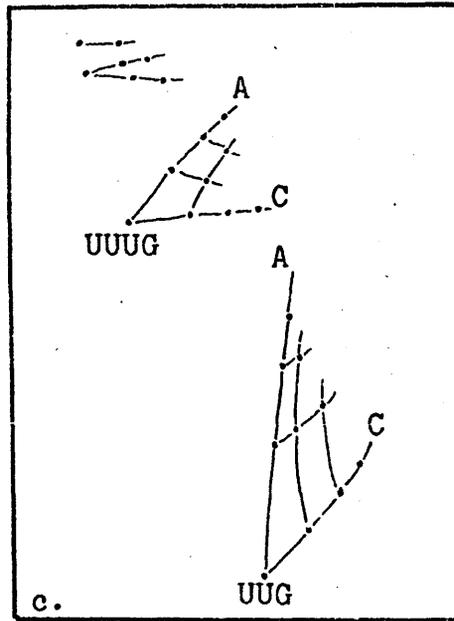
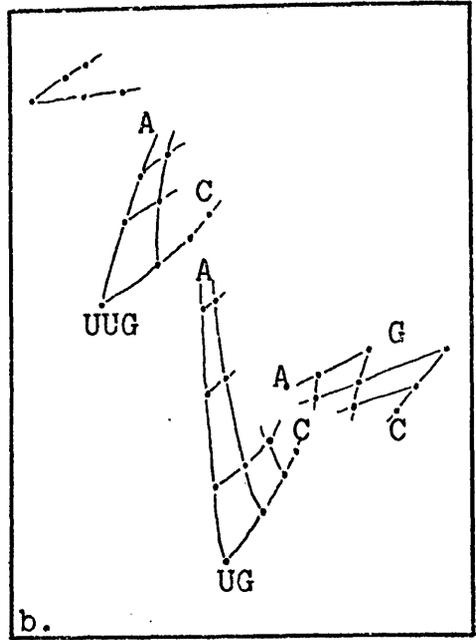
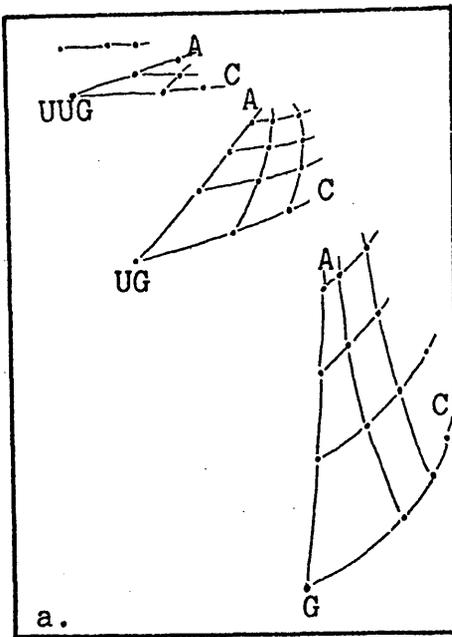


Fig. 2.2. A diagrammatic representation of the fractionation of oligonucleotides according to their composition after  $T_1$  RNase digestion. (a) a standard  $T_1$  fingerprint, (b) a standard  $T_1$  plus phosphatase fingerprint, (c) a long  $T_1$  plus phosphatase fingerprint which has been "stretched out" in the second dimension.

(Adapted from Sanger et al, 1965 and Brownlee & Sanger, 1967).

specific activity than nucleolar RNA and about 200  $\mu\text{g}$  of 18S RNA.

5.8S RNA was prepared from 28S ribosomal RNA. After the first ethanol precipitation, the 28S rRNA was dissolved in 1.0 ml of LETS. It was treated at 60°C for 4 min, cooled quickly to room temperature and layered on a 37 ml 10-25% sucrose-LETS gradient and centrifuged in the SW27 rotor at 22,000 rpm for 16h at 20°C. The gradient was monitored as before (Fig. 2.1c) and RNA from the peaks corresponding to 28S and 5.8S were collected. Carrier RNA was added to the 5.8S sample and the RNA's precipitated twice with ethanol. The specific activities were measured and the RNA stored at -20°C until use.

### 2.3 RNA Fingerprinting Technique.

The RNA fingerprinting and sequencing techniques described in this and the following section have been described very fully by Sanger et al (1965), Brownlee and Sanger (1967) and Brownlee (1972).

#### (i) Enzymic digestion of RNA.

Initially digestion of the RNA was accomplished with  $T_1$  RNase yielding oligonucleotides with a guanylic acid at the 3'-terminus. Fig. 2.2a illustrates the fractionation of the oligonucleotides into "graticules" according to their compositions, as discussed by Sanger et al (1965). The fractionation is influenced in

particular by the number of uridylic acid residues, an increasing uridylic acid content causing reduced mobility in the second dimension of the fractionation system.  $T_1$  RNase digests give rise to a large number of oligonucleotides containing several uridylic acid residues which are poorly resolved in the second dimension.

To improve the fractionation of such components the RNA was digested with a combination of  $T_1$  RNase and bacterial alkaline phosphatase (" $T_1$  plus phosphatase" digests). This resulted in oligonucleotides with a dephosphorylated guanosine moiety at the 3'-terminus; these were found to have an enhanced mobility on DEAE-paper during electrophoresis in the second dimension (Brownlee & Sanger, 1969).

The type of fractionation achieved with  $T_1$  plus phosphatase digests is represented diagrammatically in fig 2.2b and has been fully described by Brownlee & Sanger (1967). By "stretching out" the fingerprint in the second dimension (Fig. 2.2c), the combined digestion method has proved particularly useful in investigating the complex mixtures of large oligonucleotides arising from the action of  $T_1$  RNase upon nucleolar and ribosomal RNA. In addition, the system has the advantage that the oligonucleotides produced can be digested directly with snake venom phosphodiesterase.

Long  $T_1$  plus phosphatase fingerprints have been used extensively in this study. On a few occasions pancreatic ribonuclease was used.

The usual procedure was as follows:- samples of RNA were lyophilised in siliconised test tubes and digested with  $T_1$  RNase (enzyme to substrate ratio of 1 to 10, 0.01 M tris-HCl, 0.001 M EDTA, pH 7.4, 30 min) or with a combination of  $T_1$  plus phosphatase (enzyme to substrate ratio of 1 to 10 for  $T_1$  and 1 to 5 for phosphatase, 0.02 M tris-HCl, pH 8.0, 60 min). Conditions for digestion with pancreatic RNase were identical with those for  $T_1$  RNase. In most experiments 20-25 ug of RNA were lyophilised and digested in a volume of 3  $\mu$ l. All digestions were carried out in the tip of a drawn-out capillary at 37°C in a humidified oven and were applied directly to the fractionating system.

(11) 2-dimensional fractionation.

The RNA digestion products were separated in the first dimension by electrophoresis on cellulose acetate at pH 3.5.

Strips of cellulose acetate (2.5 x 95 cm) were wetted with 7M urea, pH 3.5 buffer (5% acetic acid, 7M urea, adjusted to pH 3.5 with pyridine) and excess buffer removed from around the application area. The digest was applied about 10 cm from one end of the strip and marker dye was spotted alongside. The strip was

blotted, quickly immersed in white spirit, placed over an electrophoresis rack as a support and lowered into the electrophoresis tank. Electrophoresis was carried out at 5,000 volts, the strip drawing very little current.

Generally, for a  $T_1$  or pancreatic digest, the first dimension was run until the distance between the slowest pink and the blue markers was about 35 cm. The material between the slowest pink to about 8 cm behind the blue was transferred to DEAE-paper. For a  $T_1$  plus phosphatase digestion, the first dimension was run until the slowest pink had reached the anode buffer. The material from the slowest pink to about 5 cm behind the blue was transferred to DEAE-paper.

The excess white spirit from the electrophoresis tank was allowed to drip off the cellulose acetate and the fractionated oligonucleotides were blotted on to a sheet of DEAE-paper (DE81, 43 x 94 cm) about 10 cm from one end with a pad of 5 strips of Whatman 3MM paper soaked in water. The DEAE-paper was dried and urea which had been carried over from the first dimension buffer was removed by washing with absolute ethanol.

Marker dye was applied and half the paper carefully wetted with 7% formic acid. The sheet was gently draped over a perspex electrophoresis rack and the other half wetted. The rack was carefully lowered into an electrophoresis tank and run at 1,000-1,500 volts.

The second dimension was run until the blue marker had travelled about 40% of the length of the paper for a  $T_1$  or pancreatic digest (16-20h) or until blue had travelled about 90% of the length of the paper for a combined  $T_1$  plus phosphatase digest (36-40h). DEAE-paper draws a high current with resultant heating and the tank cooling system was used.

After electrophoresis the paper was thoroughly dried prior to autoradiography (residual formic acid caused fogging of X-ray film).

#### (iii) Autoradiography.

The DEAE-paper was placed in close contact with Kodirex X-ray film in lead lined folders to locate the radioactive oligonucleotides. Papers were marked with radioactive ink containing  $^{35}\text{S}$ -sulphate, prior to autoradiography, thus enabling the films to be properly aligned with the papers and allowing accurate location of the spots for subsequent excision. The X-ray films were developed in DX-80 developer and fixed in FX-40 fixer. Generally 2  $\mu\text{Ci}$  of applied material gave satisfactory autoradiograms overnight.

## 2.4 Sequence Analyses.

Sequence analysis of oligonucleotides involved frequent incubations of 10  $\mu\text{l}$  volumes at  $37^\circ\text{C}$ . For an incubation period of up to 4h, the samples were contained in the tip of a drawn-out capillary in a humidified oven. For longer

incubation periods, eg. 16h, the 10  $\mu$ l sample was drawn into the capillary and both ends were sealed.

Electrophoresis on DEAE-paper or 52 paper employed the following buffering systems:

7% formic acid,

pH 1.9 buffer (2.5% formic acid, 8.7% acetic acid),

pH 3.5 buffer (5% acetic acid, pyridine to pH 3.5).

Electrophoresis on DEAE-paper was carried out at 1,000-1,500 volts until the blue marker had travelled 30-40 cm (16h).

Electrophoresis on 52 paper was carried out at 4,500 volts until the blue marker had travelled about 10 cm (less than 60').

(i) Radioactivity estimation.

The amount of radioactivity in a spot was determined by excising the spots and counting them directly in a Packard Tri-Carb Liquid Scintillation Spectrometer. It was not important that the areas of the spots were equal although if a spot was exceptionally large it was divided between two vials and the counts were summed. After counting, the pieces of paper were washed by stirring in three changes of toluene and dried in air. No radioactivity was lost from the paper in this way.

(ii) Elution.

Because oligonucleotides are bound ionically to DEAE-paper, a 30% aqueous solution of triethylamine

carbonate, pH 10, which will neutralise the charge on the DEAE, was employed as the eluant. Spots were eluted from Whatman 52 paper with distilled water. An eluate volume of 0.2-0.4 ml was collected on a polyvinyl chloride sheet and evaporated to dryness in an oven at 60°C.

Triethylamine carbonate solution was made by bubbling CO<sub>2</sub> (from solid CO<sub>2</sub>) through a 30% aqueous solution of redistilled triethylamine until the solution became clear. The pH was checked to be within the range 9.9-10.1 (oversaturation with CO<sub>2</sub> lowered the pH).

(iii) Base composition analysis.

The base composition of an oligonucleotide was determined by alkaline hydrolysis with separation of the mononucleotides by electrophoresis.

Oligonucleotides were dissolved in 10 µl of 0.2 M NaOH and sealed inside a capillary tube. After an incubation period of 16-18h at 37°C, the hydrolysate was applied as a 1 cm streak to Whatman 52 paper and electrophoresed at pH 3.5. The four mononucleotides were located by autoradiography and quantitated by scintillation counting.

(iv) Pancreatic RNase.

It was useful to digest T<sub>1</sub> oligonucleotides containing adenylate residues with pancreatic RNase.

Samples were digested in 10 µl of 0.01 M tris-HCl, 0.001 M EDTA, pH 7.4 containing 0.2 mg/ml pancreatic

RNase at 37°C for 30 min. The products were separated on DEAE-paper at pH 1.9 or occasionally on Whatman 52 paper at pH 3.5. The products could be identified either by their mobility or by base composition analysis.

(v) Snake venom phosphodiesterase.

Oligonucleotides from T<sub>1</sub> plus phosphatase fingerprints could be digested directly with venom phosphodiesterase. Other oligonucleotides possessed a 3'-terminal phosphate group and this had to be removed with bacterial alkaline phosphatase before snake venom phosphodiesterase would work.

This was accomplished by dissolving the sample in 10 µl of 0.02 M tris-HCl, pH 8.5 containing 0.5 mg/ml phosphatase. After incubation at 37°C for 60 min the dephosphorylated products were isolated on Whatman 52 paper by electrophoresis at pH 3.5 and eluted with water.

Total venom phosphodiesterase digestion was useful in determining the 5'-terminus nucleotide when combined with data derived from alkaline hydrolysis of the same oligonucleotide.

The dephosphorylated samples were dissolved in 10 µl of 0.025 M tris-HCl, 0.01 M Mg acetate, pH 8.5 containing 0.1 mg/ml of enzyme and incubated for 2h at 37°C. The mononucleotides were separated by electrophoresis at pH 3.5 on Whatman 52 paper and quantitated by

scintillation counting.

Partial venom phosphodiesterase digestion was particularly useful for determining the sequence towards the 3' end of an oligonucleotide. This technique is discussed in more detail in Chapter 3.

The dephosphorylated samples were dissolved in 10-15  $\mu$ l of 0.025 M tris-HCl, 0.01 M Mg acetate, pH 8.5 containing 0.02 mg/ml enzyme. Incubations were carried out at room temperature for times varying from 5-30 min and the partial digestion products were separated on DEAE-paper at pH 3.5 or pH 1.9.

(vi) Reaction with CMCT (a carbodi-imide derivative) (Brownlee et al, 1968).

Carbodi-imide derivatives will react with uridylate residues and to a lesser extent guanylate residues. Digestion of a carbodi-imide treated oligonucleotide with pancreatic RNase, which will cleave only adjacent to unmodified pyrimidines, provide additional sequence analysis.

The sample was dissolved in 20  $\mu$ l of 0.05 M tris-HCl, 0.005 M EDTA, pH 7.4 containing 100 mg/ml of N-cyclohexyl-N'-( $\beta$ -morpholinyl-(4)-ethyl) carbodi-imide-methyl-p-toluene sulphonate (CMCT) and incubated for 16h at 37°C. 10  $\mu$ l of 0.005 M tris-HCl, pH 7.4 containing 2 mg/ml pancreatic RNase is added to the sample and incubated at 37°C for a further 2h. The digest was applied to the middle of a sheet of 52 paper

as both cationic and ionic species were present and electrophoresed at pH 3.5.

(vii) Analysis for pseudouridylic acid.

Pseudouridylic acid is almost indistinguishable from uridylic acid after electrophoresis on 52 paper at pH 3.5. It can be separated by descending paper chromatography using an isopropanol, HCl, water (68:17.6:14.4) solvent (Wyatt, 1951; Fellner, 1969).

The line of spots attributable to uridylic acid after electrophoresis of mononucleotides, was cut out and stitched on to the origin of a fresh sheet of 52 paper. Chromatography was for 36h.

Identification of 340 DNA by its Spacing Sequences

Initially it was decided to analyze the transcribed  
sequences of 340 DNA. This was done for the following  
reasons: (1) the DNA is a large molecule;

**Chapter 3**

**TRANSCRIBED SPACER SEQUENCES**

It was found that these  
sequences are highly conserved  
and are present in all  
strains of 340 DNA. Previously it has been shown that  
the fingerprints of 340 DNA showed evidence  
of contamination by the material (Mader  
et al., 1968). This contamination was believed to arise from  
the material synthesized by 340 DNA in the 340 DNA  
system, possibly by non-specific breakdown of

## TRANSCRIBED SPACER SEQUENCES

### 3.1 Purification of 32S RNA by an "Actinomycin Chase"

Initially it was decided to analyse the transcribed spacer regions of 32S rather than 45S RNA for the following reasons. Although 32S RNA is itself a large molecule containing approximately 6500 nucleotides (McConkey & Hopkins, 1969; Weinberg & Penman, 1970) it is only half the size of the 45S molecule and contains only the 28S ribosomal sequence along with a presumed transcribed spacer region. Thus it was initially simpler to identify transcribed spacer oligonucleotides by comparing 32S fingerprints with those of 28S RNA than by comparing 45S fingerprints with those of 28S+18S RNA. Secondly, 32S RNA is the most abundant HeLa cell rpre RNA species, an important factor for sequence analysis of any transcribed spacer material discovered.

A prerequisite for this analysis was that the 32S RNA sample be pure. Previously it had been noticed that  $^{14}\text{C}$  methyl-labelled fingerprints of 32S RNA showed evidence of weak and variable contamination by 18S material (Maden et al, 1972a). This contamination was presumed to arise from inclusion of small quantities of 45S RNA in the 32S RNA preparation, probably by non-specific breakdown of a small proportion of 45S molecules to "half molecules" during cell fractionation and preparation of nucleoli. This problem was overcome by an "actinomycin chase":- actinomycin D was added to the growing cells 30-40 min before harvesting.

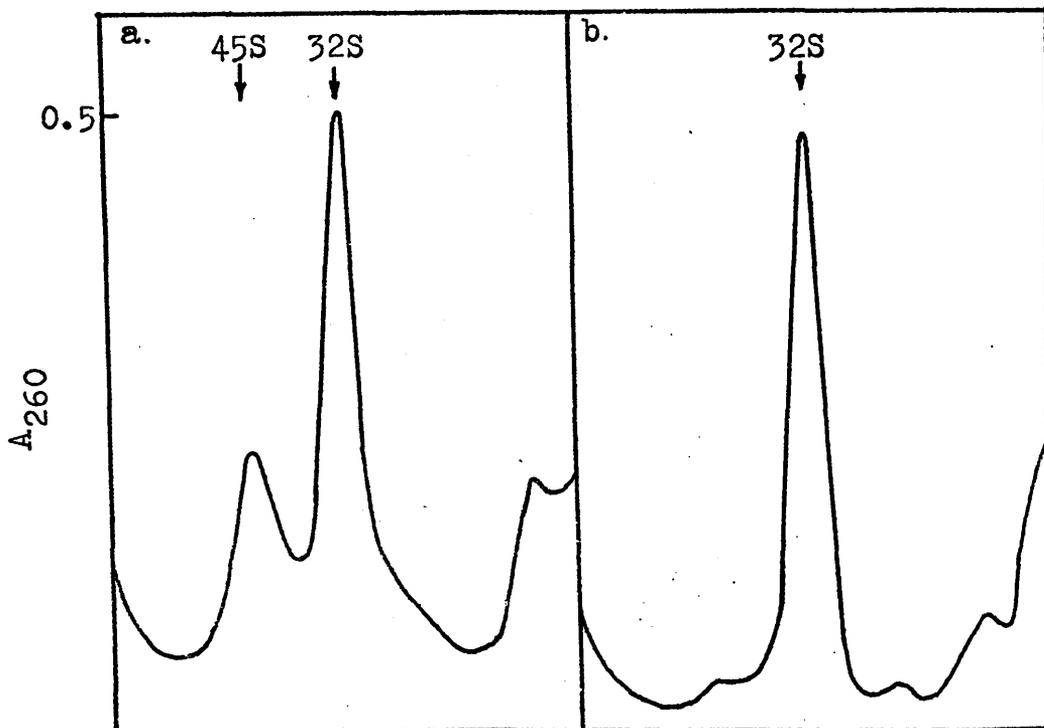


Fig. 3.1. Effect of actinomycin D on nucleolar RNA. Purified nucleoli from  $10^8$  HeLa cells were disrupted in 1 ml "SUP" and subjected to sucrose gradient centrifugation (see Methods). (a) untreated cells, (b) cells harvested 30 min after adding actinomycin D, 4  $\mu$ g/ml. Note almost complete absence of 45S RNA in (b).

Synthesis of 45S RNA quickly ceased, all existing 45S RNA was converted in vivo to lower molecular weight products (including 32S RNA) and none remained to contaminate 32S RNA during fractionation. Fig. 3.1 shows the absorbance profiles of nucleolar RNA extracted before and after treatment of the cells with actinomycin D. 45S RNA is almost completely absent from the 32S preparation after an actinomycin chase. All 32S RNA used in this analysis was therefore prepared from actinomycin treated cells.

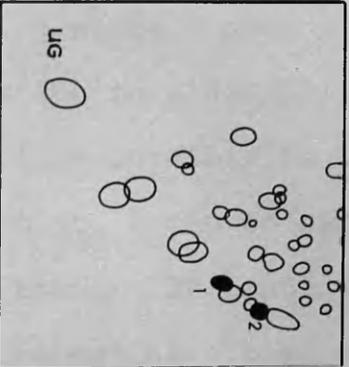
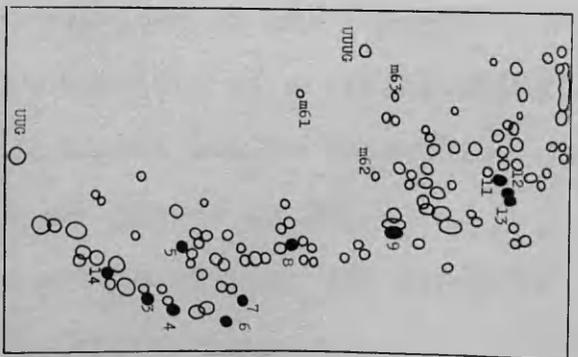
### 3.2 Fingerprints of 28S and 32S RNA.

When this analysis commenced, the following information was available:- (i) all of the methylated sequences of 32S RNA are retained in mature 28S rRNA (published later, Salim & Maden, 1973); (ii) in combined  $T_1$  plus pancreatic ribonuclease digests of RNA there are no qualitatively distinct oligo A tracts in 32S RNA which are not also present in 28S RNA (Maden & Forbes, 1972) (the longest runs of A are four long in both molecules). In order to obtain evidence of the presumed transcribed spacer sequences it was therefore necessary to analyse the complex mixture of products formed after digestion of the RNA's under other conditions.

Fig. 3.2 shows that comparison of  $T_1$  and  $T_1$  plus phosphatase fingerprints of  $^{32}\text{P}$ -labelled 28S and 32S RNA revealed some 13 products which were unique to the 32S fingerprints. Only a few of these extra products were

Fig. 3.2 Extra  $T_1$  products in digests of 32S RNA as compared with 28S RNA. Upper panels, long  $T_1$  plus phosphatase fingerprints of (a) 28S RNA, (b) 32S RNA, (c) 32S key. Lower panels, details of long  $T_1$  fingerprints of (d) 28S RNA, (e) 32S RNA, (f) 32S key. In all cases the first dimension (cellulose acetate, right to left) was run for 5h at 4.7 KV until the blue marker had just reached the top of the electrophoresis rack, and the second dimension (DEAE-paper downwards) was run for 36h at 1.3 KV until the blue marker dye had run 90 cm. All subsequent fingerprints illustrated in this thesis were run under these conditions unless otherwise stated.

All distinctive 32S transcribed spacer products are numbered and marked black in (c) and (f). A streaky product (arrowed in (b), top right hand corner) is reproducibly present in 32S and absent in 28S fingerprints. This is possibly an extra 32S product which breaks down in the second dimension buffer system. Products m61, m62 and m63 are methylated and were used to quantitate the transcribed spacer products (see section 3.3).



clearly identified in  $T_1$  fingerprints even by "stretching out" individual graticules by long separations (Fig. 3.2 d & e). The large number of small products formed on digestion of 28S and 32S RNA decreases the likelihood of detecting a unique sequence in the "no and one uridyate" graticules. It was necessary to digest the RNA with  $T_1$  RNase plus alkaline phosphatase to obtain adequate resolution of the more complex products in the "two or more U" graticules (Fig. 3.2 a & b). For this reason the resolution obtained with long  $T_1$  plus phosphatase fingerprints of the larger, more distinctive sequences has proved most useful throughout this work.

Two of the extra 32S products were found in the "one U" graticule and the remaining were found in the "two or more U" graticules in long  $T_1$  plus phosphatase fingerprints (Fig. 3.2). The extra products are reproducibly present in 32S and absent in 28S fingerprints with the exception of spot no. 4 where a spot sometimes appeared in 28S fingerprints, possibly due to a slight change in mobility of a neighbouring product (due possibly in turn to a slight buffer change). Also spot no. 2 may be present in low yields in 28S fingerprints. By contrast there are no obvious 28S products which were not also present in 32S fingerprints.

It was initially thought that the sequence U(A,U)U-G was an extra 32S product present in submolar yields (Robertson & Maden, 1973). However this is now recognised as being a heterogeneous nuclear RNA contaminant (see section

3.4). Product no. 14 and the streaky product arrowed in Fig. 3.2b have only recently been recognised as being unique 32S products.

Despite good resolution of large oligonucleotides in a  $T_1$  plus phosphatase fingerprint, an appreciable amount of material remains unresolved near the origin of the DEAE-paper. This material can be resolved by the technique of "homochromatography" on thin layers of DEAE-cellulose as the second dimension (Brownlee et al, 1968; Brownlee & Sanger, 1969) instead of electrophoresis on DEAE-paper. The thin layers were developed with "homomixture C" (Brownlee & Sanger, 1969), a concentrated mixture of oligonucleotides. The resolution of the smaller products is poor while fractionation of the largest products (10-15 residues) is generally better than in a  $T_1$  plus phosphatase fingerprint.

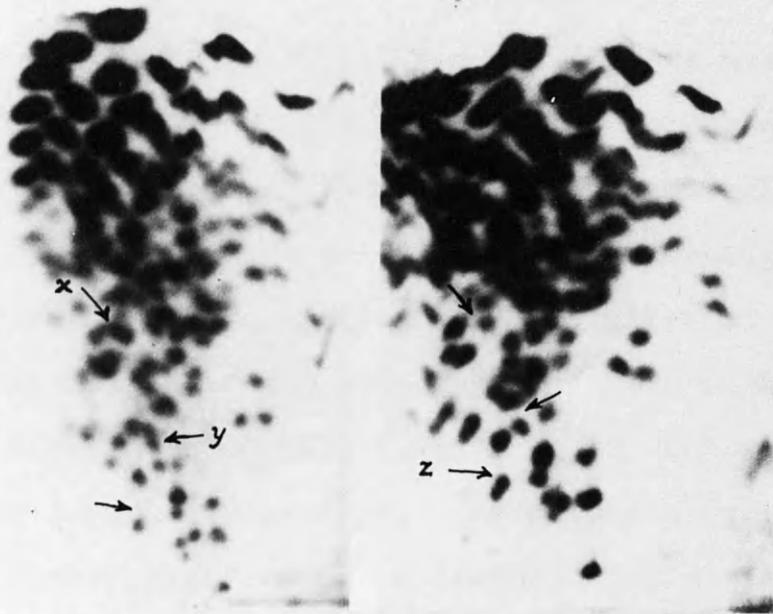
Homochromatograms of  $T_1$  digests of 28S and 32S RNA are shown in Fig. 3.3. They give good separation of the larger products but because the exact patterns are not reproducible (note the slight difference in the relative positions of the spots) it was difficult to identify unique 32S products by comparing the two thin layers. Only two unique 32S products (arrowed x and y in Fig. 3.3a) were detected and these may correspond to one or two of products 1-14. A third spot (arrowed z in Fig. 3.3b) appeared to be present in 28S  $T_1$  homochromatograms and absent in 32S  $T_1$  homochromatograms. This may represent an end-group of the 28S molecule. It remains to be further investigated.

Fig. 3.3 Homochromatography of  $T_1$  digests of 32S and 28S RNA; (a) 32S RNA, (b) 28S RNA. In the first dimension (cellulose acetate, left to right) electrophoresis was run for 1.5h at 4.7 KV and in the second dimension (DEAE-cellulose thin layers, bottom to top) chromatography was run for 20h at 60°C using homomixture C. The exact patterns were not very reproducible, rendering comparison difficult. In (a), x and y indicate two extra 32S spots and in (b), z indicates a spot which is present in 28S RNA and which appears to be absent in 32S RNA.

Fig. 3.4 Homochromatography of pancreatic ribonuclease digests of 32S and 28S RNA. (a) 32S RNA, (b) 28S RNA. Running conditions are identical to those described in Fig. 3.3. The arrows indicate a spot which is noticeably more heavily labelled in 32S than in 28S RNA.

32 S

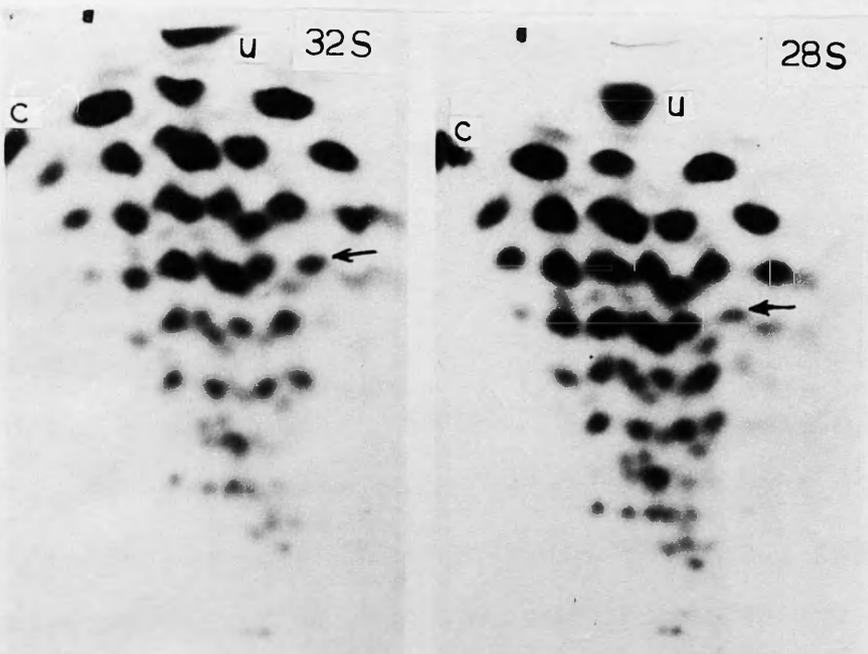
28 S



T<sub>1</sub> HOMO

u 32S

28S



Panc HOMO

Finally pancreatic digests of 28S and 32S RNA were compared for the presence of unique 32S products, using both paper electrophoresis and homochromatography techniques in the second dimension. Pancreatic ribonuclease digestion produces a large number of oligopurines. Because of the insufficient resolution of the many isomers among such products, the likelihood of detecting a 32S product with a unique mobility is greatly reduced. No unique oligopurines were found in pancreatic digests of 32S RNA when compared with those of 28S RNA separated either by electrophoresis on DEAE-paper or by homochromatography (Fig. 3.4).

Homochromatography did reveal one noticeable quantitative difference; an increased intensity of one of the larger 32S products as compared with 28S RNA (arrowed in Fig. 3.4). This product has not been further characterised.

### 3.3 Analysis of 32S T<sub>1</sub> Products 1-14.

The presence of products 1-14 in T<sub>1</sub> digests of 32S RNA but not in 28S RNA indicates that these products are derived from the transcribed spacer region of 32S RNA. These products were further characterised as shown in table 3.2. Because of the high pyrimidine content of these products (determined by base composition analysis) the most useful procedure for sequencing was partial digestion with snake venom phosphodiesterase (VPDE). Venom phosphodiesterase sequentially removed nucleotides from the 3' end of dephosphorylated oligonucleotides. The degradation

products were separated by electrophoresis on DEAE-paper at pH 3.5 (Fig. 3.5). It has been reported that removal of a nucleotide confers a specific change in the mobility of the product dependent on the nature of the nucleotide removed (Brownlee & Sanger, 1967). A value  $M$  was defined as equal to  $x/y$ , where  $y$  is the distance of any oligonucleotide from the origin, and  $x$  is the distance between it and its first degradation product. Thus the nucleotide sequence can in principle be deduced from the relative mobilities of the degradation products, with the reservation that values of  $M$  where the product runs faster than the blue marker dye are unreliable. In addition it is difficult with  $M$  values alone to distinguish weakly labelled genuine degradation products from impurities. In practice, in most cases the nucleotide sequence was determined also by base composition analysis of the degradation products (table 3.1). Generally electrophoresis at pH 3.5 gave good separation of the degradation products although at this pH, removal of an A residue is not clearly recognisable from removal of U by  $M$ -values alone. In a few cases when the yield of the degradation product was low, identification was based on  $M$ -values alone if no A residues were present. In this work, using DEAE-paper at pH 3.5, the  $M$ -value after removal of cytidilic acid was 0.5-0.85, and after removal of uridylic acid, 1.2-1.8.

The structures of products 1-14, derived from partial venom digestion and other data, are shown in table 3.2. Total digestion with VPDE releases the 5' residue as a nucleoside whilst alkaline hydrolysis releases it as a

Fig. 3.5 Fractionation of VPDE partial products of 32S transcribed spacer products 3-12. 32S products 3-12 from T<sub>1</sub> plus phosphatase fingerprints were partially digested with VPDE and the products were separated on DEAE-paper by electrophoresis at pH 3.5 (1.1 KV, 16h) (see Methods and section 3.3). Normally most spots were obtained pure from the fingerprints. Spots 4, 5 and 8 often contained impurities from adjacent oligonucleotides (see Fig. 3.2b). Contaminating digestion products arising from the impurities were identified by their base composition which did not fit in with the sequence of base compositions of the real digestion products. Also, with spots 4 and 5, the faint products situated immediately above the initial products (marked a) must be impurities because the M-value of the first genuine digestion product must be >2 due to the removal of the 3' G residue.



Table 3.1 Sequences of 32S Products 3-13 Deduced by Partial Digestion with VPDE.

Spot No.	VPDE partial product	alkaline hydrolysis			M value	Suggested structure
3	a)	C <sub>4</sub>	U <sub>2</sub>		2.6	(C <sub>2</sub> ,U)CUCG
	b)	C <sub>3</sub>	U <sub>2</sub>		0.5	(C <sub>2</sub> ,U)CUC
	c)	C <sub>3</sub>	U			(C <sub>2</sub> ,U)CU
	d)	C <sub>2</sub>	U			(C <sub>2</sub> ,U)C
4	a)	C <sub>5</sub>	U <sub>2</sub>		2.6	(C <sub>3</sub> ,U)CUCG
	b)	C <sub>4</sub>	U <sub>2</sub>		0.6	(C <sub>3</sub> ,U)CUC
	c)	C <sub>4</sub>	U		1.4	(C <sub>3</sub> ,U)CU
	d)	C <sub>3</sub>	U			(C <sub>3</sub> ,U)C
5	a)	C <sub>2</sub>	U <sub>2</sub>	A	2.4	(U <sub>2</sub> ,A)CCG
	b)	C	U <sub>2</sub>	A	0.7	(U <sub>2</sub> ,A)CC
	c)		U <sub>2</sub>	A	0.5	(U <sub>2</sub> ,A)C
6	a)	C <sub>7</sub>	U <sub>2</sub>		2.6	(C <sub>3</sub> ,U)CCUCCG
	b)	C <sub>6</sub>	U <sub>2</sub>		0.6	(C <sub>3</sub> ,U)CCUCC
	c)	C <sub>5</sub>	U <sub>2</sub>		0.6	(C <sub>3</sub> ,U)CCUC
	d)				1.4 *	(C <sub>3</sub> ,U)CCU
	e)				0.5 *	(C <sub>3</sub> ,U)CC
	f)				0.4 *	(C <sub>3</sub> ,U)C
7	a)	C <sub>5</sub>	U <sub>2</sub>	A <sub>2</sub>	2.3	(C <sub>4</sub> ,U)CUAAG
	b)	C <sub>5</sub>	U <sub>2</sub>	A	1.4	(C <sub>4</sub> ,U)CUAA
	c)	C <sub>5</sub>	U <sub>2</sub>		1.3	(C <sub>4</sub> ,U)CUA
	d)				1.4 *	(C <sub>4</sub> ,U)CU
	e)				0.6 *	(C <sub>4</sub> ,U)C
8	a)	C <sub>2</sub>	U <sub>2</sub>	A <sub>3</sub>	2.2	(A <sub>2</sub> ,C,U)AUCG
	b)	C	U <sub>2</sub>	A <sub>3</sub>	0.7	(A <sub>2</sub> ,C,U)AUC
	c)	C	U	A <sub>3</sub>	1.3	(A <sub>2</sub> ,C,U)AU
	d)	C	U	A <sub>2</sub>	1.8	(A <sub>2</sub> ,C,U)A

Table 3.1 (cont'd).

Spot No.	VPDE partial product	alkaline hydrolysis		M value	Suggested structure
9	a)	C <sub>5</sub>	U <sub>3</sub>	2.5	(C <sub>2</sub> ,U <sub>2</sub> )CCUCG
	b)	C <sub>4</sub>	U <sub>3</sub>	0.5	(C <sub>2</sub> ,U <sub>2</sub> )CCUC
	c)	C <sub>4</sub>	U <sub>2</sub>	1.5	(C <sub>2</sub> ,U <sub>2</sub> )CCU
	d)	C <sub>3</sub>	U <sub>2</sub>	0.5	(C <sub>2</sub> ,U <sub>2</sub> )CC
	e)			0.5 *	(C <sub>2</sub> ,U <sub>2</sub> )C
11	a)	C <sub>5</sub>	U <sub>4</sub>	2.5	(C,U <sub>3</sub> )CCCUCG
	b)	C <sub>4</sub>	U <sub>4</sub>	0.8	(C,U <sub>3</sub> )CCCUC
	c)	C <sub>4</sub>	U <sub>3</sub>	1.3	(C,U <sub>3</sub> )CCCU
	d)	C <sub>3</sub>	U <sub>3</sub>	0.6	(C,U <sub>3</sub> )CCC
	e)			0.5 *	(C,U <sub>3</sub> )CC
	f)			0.5 *	(C,U <sub>3</sub> )C
12	a)	C <sub>5</sub>	U <sub>4</sub>	2.3	(C <sub>3</sub> ,U)CUCUUG
	b)	C <sub>5</sub>	U <sub>3</sub>	1.3	(C <sub>3</sub> ,U)CUCUU
	c)	C <sub>5</sub>	U <sub>2</sub>	1.3	(C <sub>3</sub> ,U)CUCU
	d)	C <sub>4</sub>	U <sub>2</sub>	0.6	(C <sub>3</sub> ,U)CUC
	e)	C <sub>4</sub>	U	1.4	(C <sub>3</sub> ,U)CU
	f)			0.5 *	(C <sub>3</sub> ,U)C
13	a)	C <sub>7</sub>	U <sub>4</sub>	-	(C <sub>6</sub> ,U <sub>4</sub> )CG
	b)	C <sub>6</sub>	U <sub>4</sub>	0.6	(C <sub>6</sub> ,U <sub>4</sub> )C

Nucleotides marked \* were identified by M values only.

Table 3.2 Sequence and Quantitative Data on  $^{32}\text{S}$   $T_1$  Products 1-14.

Spot No.	Alkaline hydrolysis of oligonucleotides (a)				Complete venom phosphodiesterase digestion of oligonucleotides (b)				5' end group deduced	Analysis for $\psi$
	G	U	A	C	G	U	A	C		
1	1.0	1.1	0.0	3.1	1.0	0.4	0.1	2.7	U	None present
2	1.0	1.0	0.1	3.8	1.0	1.0	0.0	2.7	C	"
3	1.0	2.1	0.3	4.0	1.0	2.4	0.1	3.2	C	"
4	-	2.0	0.1	4.8	1.0	2.2	0.1	3.6	C	"
5	-	2.0	1.2	2.0	1.0	2.3	0.1	2.0	A	"
6	-	2.0	0.2	7.0	1.0	2.3	0.2	5.7	C	"
7	-	2.0	1.8	5.0	1.0	1.4	1.9	4.1	C	"
8	-	2.0	2.8	2.4	1.0	2.7	2.5	2.8	?	"
9	-	3.0	0.2	5.1	1.0	2.4	0.1	4.8	U	"
11	-	4.0	(1.1)	5.1	1.0	3.4	0.5	4.9	U	"
12	-	4.0	0.4	5.2	1.0	4.7	0.3	5.4	?	"
13	-	4.0	0.3	7.0	1.0	4.2	0.3	5.9	C	"
14	-	2.0	0.1	2.7	1.0	1.3	0.0	2.7	U	"

(a) Nucleotide compositions of products from  $T_1$  fingerprints were estimated by liquid scintillation counting and are expressed relative to G as one. Compositions of  $T_1$  plus phosphatase products are expressed relative to U, the number of U residues per product being determined by the position of the spot in the fingerprint.

(b) Compositions after venom digestion are expressed relative to G as one. (All base composition data are means of several independent determinations, using material from separate fingerprints.)

(c) Molar yields were determined as described in the text. Results are means from a minimum of seven determinations except for spots 1 and 2 which were from four determinations.

(d) Identified by mobility on DEAE-paper at pH 3.5.

A containing product from pancreatic ribonuclease digestion	Sequence deduced from partial venom digestion	Structure	Molar Yield (c)	
			Observed	Suggested
-	-	UCCCG	5-8	5-8
-	-	C(C <sub>3</sub> ,U)G	3-6	3-6 <sup>(1 in 28S?)</sup>
-	(C <sub>2</sub> ,U)CUCG	C(C,U)CUCG	1.04	1
-	(C <sub>3</sub> ,U)CUCG	C(C <sub>2</sub> ,U)CUCG	1.12	1
AUp (d)	(A,U <sub>2</sub> )CCG	AUCCG	1.47	1-2
-	(C <sub>3</sub> ,U)CCUCCG	C(C <sub>2</sub> ,U)CCUCCG	0.67	1
AAG	(C <sub>4</sub> ,U)CUAAG	C(C <sub>3</sub> ,U)CUAAG	0.85	1
AUp, AAUp	(A <sub>2</sub> ,C,U)AUCG	A(A,U-C)AUCG(e)	0.88	1
-	(U <sub>2</sub> ,C <sub>2</sub> )CCUCCG	U(C <sub>2</sub> ,U)CCUCCG	0.89	1
-	(C,U <sub>3</sub> )CCCUCG	U(C,U <sub>2</sub> )CCCUCG	0.86	1
-	(C <sub>3</sub> ,U)CUCUUG	(C <sub>3</sub> ,U)CUCUUG	0.71	1
-	(C <sub>6</sub> ,U <sub>4</sub> )CG	C(C <sub>5</sub> ,U <sub>4</sub> )CG	0.61	1
-	-	U(C <sub>3</sub> ,U)G	0.99	1

(e) The sequence of spot 8 was further analysed by blocking urydilate and guanylate residues with the carbodi-imide CMCT (see Methods) followed by digestion with pancreatic ribonuclease and base composition analysis of the products. Two main products were formed: (A<sub>2</sub>,U)Cp and (A,U)Cp, but this information did not permit an unambiguous sequence to be assigned to the original product.

(f) The A which appeared in alkaline hydrolysates of spot 11 was later found to be a contaminant from an adjacent adenine-rich oligonucleotide.

nucleoside-2'(3')-phosphate. Thus the 5' residue may be determined from the combined compositional data.

Pancreatic ribonuclease was also used to obtain sequence information on the few A containing products. Chromatography on Up after alkaline hydrolysis (see Methods) indicated that none of the products 1-14 contained  $\psi$ p.

In summary, several of the products are quite large, most are rich in C and possess little A, none of them correspond in mobility to known methylated products and none contain  $\psi$  assayed by chromatography. It is worthy of note that four of the products (3, 4, 9 and 11) end with the sequence -C-U-C-G. The sequence information was incomplete as no reliable method was available for determination of the nucleotide sequence at the 5'-end and also because the final VPDE partial digestion products were obtained in too low yield for base composition analysis. However, having established the size of the oligonucleotides, quantitation was possible.

On the basis of a reliable estimate of the absolute numbers of methyl groups in HeLa cell rRNA (Maden et al, 1972b), most methylated sequences in  $T_1$  fingerprints of 28S and 32S RNA were shown to occur once per molecule (Salim & Maden, 1973; Maden & Salim, 1974). Quantitation of  $T_1$  plus phosphatase products 3-14 was carried out relative to three such methylated sequences m61, m62 and m63 (Fig. 3.2) which occur once per molecule of 28S and 32S RNA. These were previously characterised as U-U-Gm-G, Cm-A-Gm-U-U-G

and mU-U-ψ-A-G respectively and were obtained pure in T<sub>1</sub> plus phosphatase fingerprints of <sup>32</sup>P-labelled 28S and 32S RNA (Salim & Maden, 1973; Maden & Salim, 1974). Quantitation of T<sub>1</sub> products 1 and 2 was similarly carried out relative to two pure methylated products Am-U-G and Um-A-C-G which also occur approximately once per molecule. In both types of fingerprint the mean of the yields of these "reference" products was taken as 1.00.

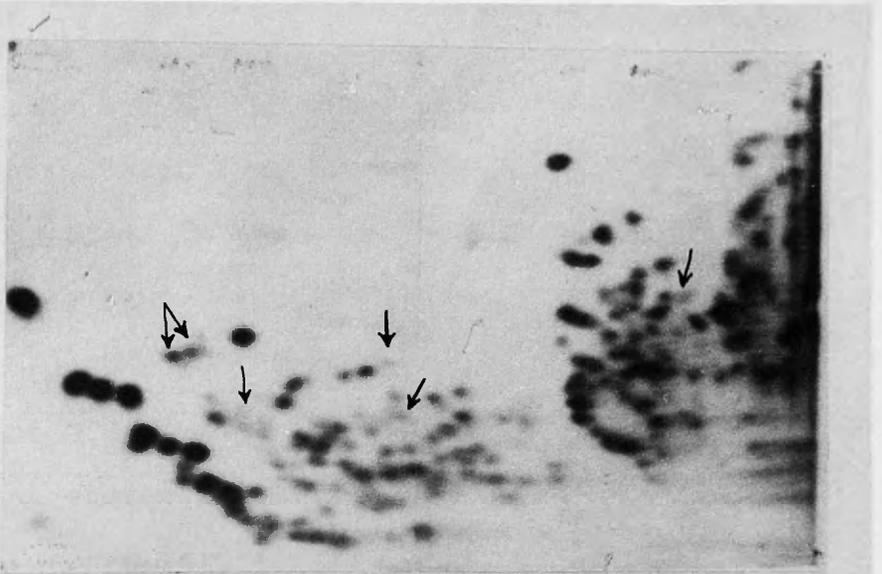
Relative to this value the yields of most of the larger transcribed spacer products fell within the range 0.7-0.9 (table 3.2). Various factors might tend to reduce the yields of large products as compared with small ones in this type of analysis, such as overdigestion with secondary splitting, incomplete transfer from cellulose acetate, or streaking of some products on DEAE-paper. Streaking was troublesome with products 6, 12 and 13. In addition to these factors, small amounts of nucleolar 28S RNA can be co-purified with 32S RNA. The presence of 28S RNA in the 32S RNA preparation will effectively lower the yield of transcribed spacer products. I therefore suggest that most of the recognisable transcribed spacer products occur once per molecule of 32S RNA. Products 1 and 2 clearly occur several times though overlap with neighbouring spots makes precise quantitation difficult. Product 5 is often incompletely resolved from other oligonucleotides and could occur once or twice.

### 3.4 Purification of 45S RNA.

45S RNA contains the sequences of 28S and 18S rRNA together with transcribed spacer regions including (a) the part found in 32S RNA and (b) presumably other extensive regions not present in 32S RNA. Can further transcribed spacer products, from such other regions, be isolated from 45S RNA by comparing fingerprints of 45S RNA with those of 28S+ 18S RNA?

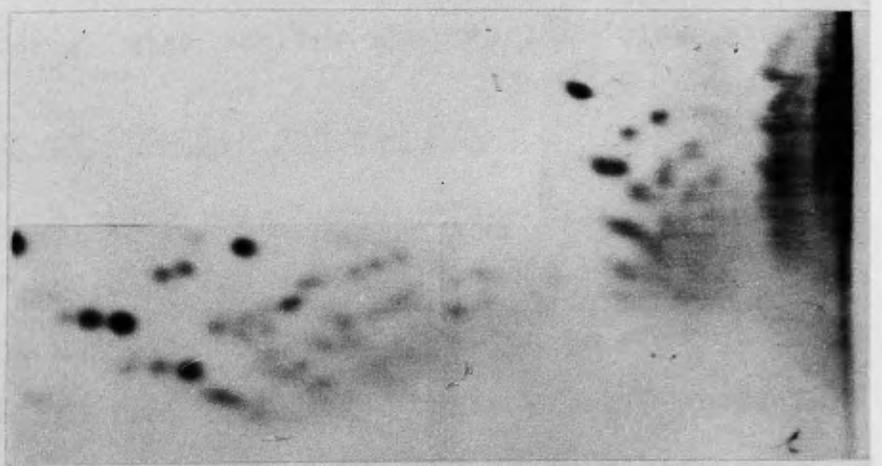
$T_1$  plus phosphatase fingerprints of 18S RNA were very different from those of 28S RNA and fingerprints of a mixture of 28S+ 18S RNA were very complex.  $T_1$  plus phosphatase fingerprints of 45S RNA would thus be expected to be at least equally or even more complex than those of 28S+ 18S RNA. Comparison of 45S fingerprints with those of 28S+ 18S RNA revealed several extra spots in the former, some of which varied in intensity from one preparation to another. Some "shadowy" areas also appeared in the 45S fingerprints. Towards the end of this analysis of 45S RNA it was discovered that the shadowy areas and the extra spots of variable intensity were due to contamination of 45S RNA with heterogeneous nuclear RNA (HnRNA). This was realised on comparison of a  $T_1$  plus phosphatase fingerprint of 45S RNA with one of HnRNA (Fig. 3.6). Although the method used for preparation of 45S RNA was thought to separate HnRNA from the ribosomal RNA precursors (Willems et al, 1968), the close examination of 45S RNA with  $T_1$  plus phosphatase fingerprints indicated the contaminants. A smaller level

Fig. 3.6  $T_1$  plus phosphatase fingerprints of 45S and HnRNA; (a) 45S, (b) HnRNA. HnRNA was prepared in the following way: after sedimenting nucleoli through a 15-30% sucrose-HSB gradient (see Methods), the supernatant was precipitated with 2 volumes of ethanol, redissolved in 3 ml of LETS buffer and extracted with phenol. The aqueous layer was precipitated with ethanol, redissolved in 1 ml of LETS buffer and centrifuged through a 5-20% sucrose-LETS gradient in the SW27 rotor (6h, 25,000 rpm, 20°C). Carrier 28S RNA was added to the RNA sedimenting faster than 45S. This RNA was precipitated with ethanol and recentrifuged as above. RNA sedimenting faster than 45S was collected, carrier RNA added and precipitated twice with ethanol. Conditions of electrophoresis are as in Fig. 3.2. The arrows in (a) denote characteristic contaminating HnRNA spots.



a)

45 S



b)

H<sub>n</sub> RNA

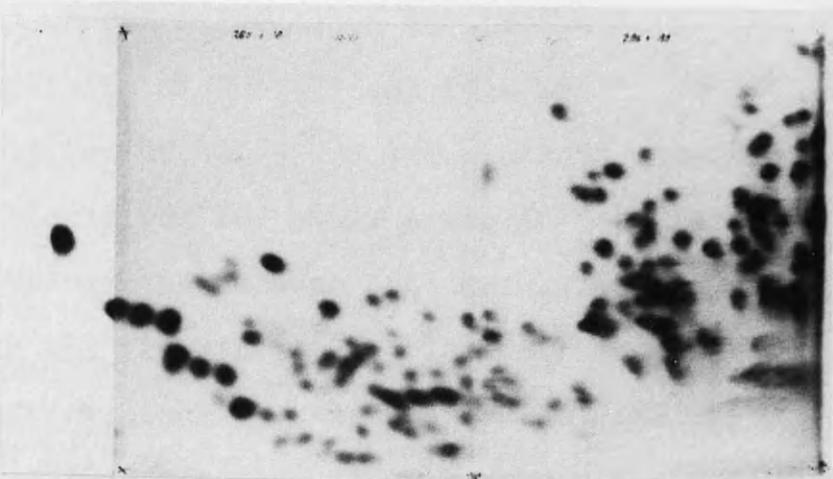
of HnRNA is present in the 32S RNA preparations. This resulted in the incorrect identification of one extra spot as a transcribed spacer product (spot no. 10; Robertson & Maden, 1973).

An attempt to obtain 45S RNA free from HnRNA was made by recentrifuging the 45S RNA in low ionic strength buffer (Warner et al, 1966; Attardi et al, 1966; Jeanteur et al, 1968). 45S RNA prepared in the usual way was centrifuged through a sucrose-low ionic strength buffer gradient ( $10^{-3}$  M tris-HCl;  $0.25 \times 10^{-3}$  M EDTA, pH 7.0) after a heat denaturation step ( $80^{\circ}\text{C}$ , 3 min) to promote complete dissociation of any possible aggregates of shorter RNA chains sedimenting in the 45S region. However, fingerprints of 45S RNA prepared in this way showed no less contaminating HnRNA material when compared with earlier fingerprints of 45S RNA.

### 3.5 Fingerprints of 28S+ 18S and 45S RNA.

Comparison of  $T_1$  plus phosphatase fingerprints of 45S RNA with those of 28S+ 18S RNA revealed some 9 products which were unique to 45S RNA (Fig. 3.7) in addition to products 1-14 already characterised from 32S RNA. These extra 9 products must be derived from that part of the 45S transcribed spacer region which does not appear in 32S RNA. This region will hereafter be termed the 45S unique transcribed spacer region, the total transcribed spacer regions of 45S RNA consisting of 45S unique transcribed spacer and the 32S transcribed spacer region.

Fig. 3.7 Extra products in  $T_1$  plus phosphatase fingerprints of 45S RNA as compared with 28S+ 18S fingerprints. (a) 28s+ 18S, (b) 45S, (c) 45S key. Conditions of electrophoresis are as in Fig. 3.2. The products which are found only in 45S fingerprints are numbered and marked black. The shaded products are 32S transcribed spacer products 3-14 described earlier. Spot no. 30 is present in the 28s+ 18S fingerprint and absent in the 45S fingerprint. The arrows indicate the two spots used for quantitation.



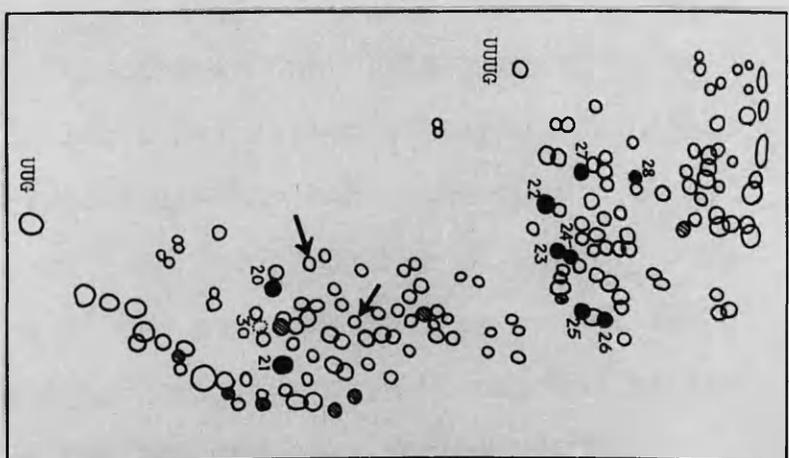
d)

28+18 S



b)

45 S



c)

45 KEY

The likelihood of detecting extra 45S products is reduced by the complexity of the 45S and 28S+ 18S fingerprints. Only two of the extra 45S products are found in the "two U" graticule and the other seven in the "three U" graticule. (32S transcribed spacer products 11-13 did not resolve well in 45S fingerprints. This was probably due to a slight buffer change causing changes in mobility of these and neighbouring products as they were not clearly resolved in 32S fingerprints made at the same time).

In addition to the qualitative differences between 45S and 28S+ 18S fingerprints, many quantitative differences were detected. These were particularly evident along the "C rich" or right hand edge of the "two U" graticule of the 45S fingerprint. This indicated the presence of many C rich products in 45S transcribed spacer, in agreement with the indirect evidence that the transcribed spacer has a high C content (Jeanteur et al, 1968; Willems et al, 1968).

One 18S product was not present in the 45S fingerprint. This spot is termed no. 30 (Fig. 3.7c). The most likely explanation for this finding is that this product is an end group of the 18S molecule. From alkaline hydrolysis and total venom phosphodiesterase digestion (see Methods) the product appears to have the structure  $U(AUC_2)G$ . The presence of G at the 3'-end suggests that it is not the 3'-end group of 18S RNA. pUp has been identified as the 5'-end nucleotide of 18S RNA (Maden & Forbes, 1972). Therefore  $U(AUC_2)G$  may be the 5'-end group. However as this

Table 3.3 Quantitation of  $^{32}\text{S}$  Transcribed Spacer Products in 45S  $T_1$  plus Phosphatase Fingerprints.

Spot No.	Structure	Molar yield
4	C(C,U)CUCG	1.19
6	C(C <sub>2</sub> ,U)CUCG	0.88
7	C(C <sub>3</sub> ,U)CUAAG	0.94
8	A(A,U-C)AUCG	0.67
9	U(C <sub>2</sub> ,U)CCUCG	1.30
14	U(C <sub>3</sub> ,U)G	0.99

Products 3, 5 and 11-13 were poorly resolved from neighbouring oligonucleotides. Products 3 and 5 appeared to be present once per mole. Quantitation was relative to two products (arrowed in fig. 3.7c) referred to in section 3.6.

Table 3.4 Sequence and Quantitative Data on 45S T<sub>1</sub> Products 20-28.

Spot No.	Alkaline hydrolysis of oligonucleotides (a)			Complete venom phosphodiesterase digestion of oligonucleotides (b)			5' end group deduced	Analysis A containing product from pancreatic ribonuclease digestion	Structure	Molar yield (c)		
	G	U	A	G	U	A					None present	AUP, ACP
	G	U	A	C	A	C						
20	-	2.0	1.2	1.2	1.0	2.5	0.2	1.0	A	None present	AUUG	1
21	-	2.0	0.9	3.2	1.0	2.3	1.0	2.3	C	"	AC	2.51
22	-	3.0	0.3	1.8	1.0	2.4	0.1	1.9	U	"	-	1.30
23	-	3.0	0.2	3.0	1.0	2.5	0.1	3.0	U	"	-	2.21
24	-	3.0	0.3	3.3	1.0	3.3	0.3	2.3	C	"	-	1.21
25	-	3.0	0.5	5.6	-	-	-	-	-	"	-	1.28
26	-	3.0	0.6	6.4	-	-	-	-	-	"	-	2.13 (d)
27	-	3.0	0.9	1.3	-	-	-	-	-	"	-	1.00
28	-	3.0	2.2	0.6	-	-	-	-	-	"	-	0.98

(a) Compositions after alkaline hydrolysis are expressed relative to U, the number of U residues per product being determined from the position of the spot in the fingerprint. At least three determinations were made except for products 20, 25 and 28 for which only two determinations were made.

(b) Compositions after venom digestion are expressed relative to G as one. Two determinations were made of products 22-24 and only one determination was made of products 20 and 21.

(c) Molar yields were determined as described in the text. Results are means from a minimum of seven determinations.

(d) Products 25 and 26 are not well resolved from neighbouring 28s+ 18S spots. Their molar yields were determined by counting the complex of spots in both 45S and 28s+ 18S fingerprints.

product was isolated from a  $T_1$  plus phosphatase fingerprint the 5'-phosphate, if originally present, will have been removed. A search is currently under way to find a corresponding phosphorylated product in  $T_1$  fingerprints.

Products 3-14 described previously as present in 32S and absent in 28S RNA digests appear by inspection to be present in 45S and absent in 28S+ 18S fingerprints. They occur in approximately the same molar yields in 45S fingerprints as they did in 32S fingerprints (table 3.3). Thus they do not occur in 18S RNA and are completely unique to the transcribed spacer region of 32S RNA.

### 3.6 Analysis of 45S $T_1$ Products 20-28.

Products 20-28 were subjected to the standard techniques for determination of nucleotide sequences (table 3.4). However the amounts of material available for analysis were small, and degradation products after partial digestion with venom phosphodiesterase were obtained in too low yield to be of any use in sequence determination. In a few cases, the 5'-end nucleotide was deduced from a comparison of the compositional data after alkaline hydrolysis and total venom digestion. The results indicated that products 20-28 are rich in C and low in A, none correspond in mobility to known methylated products and none contain  $\psi$  assayed by chromatography. As in the case of 32S products, the sequence information was sufficient to establish the sizes of the 45S products and thereby permit quantitation.

Quantitation of products 20-28 was carried out initially

relative to the three methylated products m61, m62 and m63 which were obtained pure and occur once per molecule of 28S and 32S RNA. Product m61 also occurs in 18S RNA and thus occurs twice in 45S RNA. The results obtained for products 20-28 and for several other pure spots chosen at random gave molar yields which were higher than anticipated and which did not approximate to an integral value. Two such spots which were well resolved and which consistently appeared to be present in unimolar yield were therefore chosen as new reference spots for quantitation. The mean of these "reference" products (arrowed in fig. 3.7c) was taken as 1.00 and the results obtained for products 20-28 are shown in table 3.4. The values are slightly greater than whole numbers probably due to HnRNA contamination.

On this basis quantitation of the methylated products m61, m62 and m63 and several other 45S methylated sequences (Salim & Maden, 1973) gave submolar yields of about 0.6. Thus the 45S RNA appears to be submethylated possibly through the combined effects of the treatment of the cells during labelling and harvesting and a metabolic deficiency in the growth medium of, for example, methionine. Such a deficiency is known to cause submethylation of 45S RNA.

## Chapter 4

### THE 5.8S SEQUENCE IN RIBOSOMAL PRECURSOR RNA

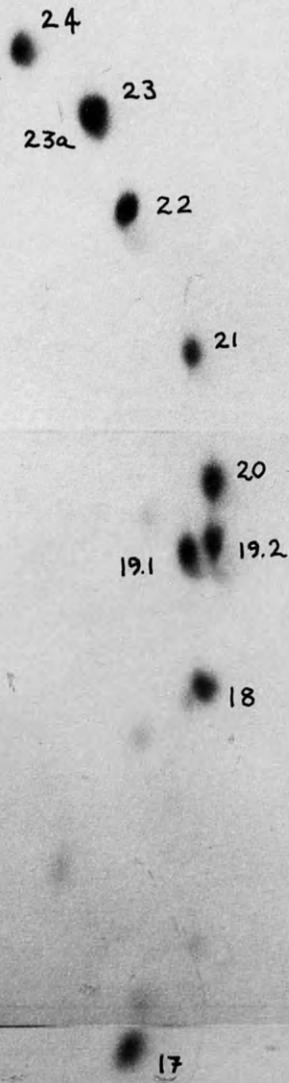
## THE 5.8S SEQUENCE IN RIBOSOMAL PRECURSOR RNA

### 4.1 Fingerprints of 5.8S RNA.

So far, this report has demonstrated the possibility of detecting and analysing unique oligonucleotides from such large molecules as 32S and 45S RNA. While the above analysis was proceeding Dr. Maden (of this laboratory) was examining 5.8S RNA for modified nucleotides, and several large oligonucleotides (constituting ~45% of the molecule) were characterised from "long"  $T_1$  plus phosphatase fingerprints of 5.8S RNA (Fig. 4.1). At this point it was realised that if 5.8S RNA is present as such in ribosomal precursor RNA then some or all of these large  $T_1$  plus phosphatase products should be identifiable in the "long"  $T_1$  plus phosphatase fingerprints of 32S RNA, as it is likely that any large (hexa - or larger) 32S oligonucleotide will be unique (see Appendix).

In his search for modified nucleotides, Dr. Maden discovered two pseudouridylate residues (products 10 and 12, see Fig. 5.2) and one methylated sequence (product 23a). The alkali-stable dinucleotide of product 23a (GmC) had previously been identified as the methylated component of a 28S product in this position of the 28S fingerprint (methylated product 66; Salim & Maden, 1973). A 3'-OH end group was isolated (product 2a,  $CUU_{OH}$ ; not shown) but so far no 5'-phosphate end group has been discovered.

Fig. 4.1  $T_1$  plus phosphatase fingerprint of 5.8S RNA. Conditions of electrophoresis are as in Fig. 3.2. Spots 23 and 23a generally do not separate well in 5.8S fingerprints. A long  $T_1$  fingerprint of 5.8S RNA is shown in Fig. 5.2 in the next chapter.



5.8 S

Table 4.1 Sequence and Quantitative Data on 5.8S  
T<sub>1</sub> Products 10, 12 and 17-24.

Spot No.	Structure <sup>1</sup>	Molar yield <sup>2</sup>
10	ψG	~ 1
12	CψG	~ 1
17	UCUG	1.0
18	CACUUG	0.7
19.1	AUCAUG	0.6
19.2	AUCACUG	0.6
20	ACACUUG	0.5
21	ACACAUUG	0.5
22	UCCUCCCG	0.9
23	ACUCUUG	} 1.3
23a	AAUUGmCAG	
24	AAUUAUG	0.6

1. The sequences were deduced mainly by Dr. Maden using partial VPDE digestion and were completed by myself.

Pancreatic ribonuclease was used to complete the sequences of spots 19.1-21 and to confirm the sequences of spots 23, 23a and 24. Determination of the 5'-end group by total VPDE digestion (see Methods) completed the sequences of spots 17 and 18 and confirmed the sequences of spots 19.1, 20, 21, 23 and 24.

2. The low molar yields of most of the T<sub>1</sub> plus phosphatase products are believed to be due to overdigestion of isolated 5.8S RNA (see section 4.1).

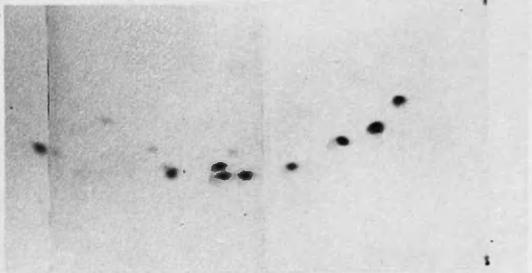
The sequences of the two pseudouridylate containing products and of the large products 17-24 are shown in table 4.1. The low molar yield of several of the products is almost certainly caused by overdigestion. Some of the breakdown products appear as faint extra products which become heavily labelled in extreme cases of overdigestion. The results in the next section indicate that the 5.8S products occur in almost unimolar yield in 28S RNA.

#### 4.2 The 5.8S Sequence in 28S RNA.

To assist in determining whether the 5.8S oligonucleotides are present in 32S fingerprints, products 18-24 were first located in 28S  $T_1$  plus phosphatase fingerprints. 28S RNA prepared by cold phenol extraction retains the 5.8S fragment whereas "28H" RNA (heat-treated 28S) lacks the fragment.  $T_1$  plus phosphatase fingerprints of 5.8S, 28S and "28H" RNA are shown in Fig. 4.2.

The approximate location of 5.8S products 18-24 in the 28S fingerprint was determined by superposition of a 5.8S fingerprint over a 28S fingerprint. Visual inspection of a 28S and "28H" fingerprint revealed their exact location as follows. Spots 18, 19.1 and 21 were clearly correlated in 28S fingerprints with 3 spots that were missing in "28H" fingerprints. Spots 19.2 and 20 were detected at the bottom and top respectively of a cluster of spots in the 28S fingerprint and were absent in the "28H" fingerprint. Spots 22, 23, 23a and 24 were correlated with four spots

Fig. 4.2  $T_1$  plus phosphatase fingerprints of (a) 5.8S, (b) 28S, (c) "28H", and (e) 32S RNA. (d) and (f) are keys to 5.8S and 32S RNA respectively. Conditions of electrophoresis are as in Fig. 3.2. The degree of separation of products 23 and 23a varied and was considerably better in 28S and 32S fingerprints than in 5.8S fingerprints. The arrows signify the positions of the 5.8S products in 28S and 32S fingerprints and the absence or diminished intensities of corresponding spots in "28H" fingerprints. 5.8S spots or mixtures of spots containing 5.8S material are numbered and marked black in (f) whilst products from 32S transcribed spacer are shaded. Products m61, m62 and m63 are pure methylated products that were used for quantitation of other products (table 4.2).



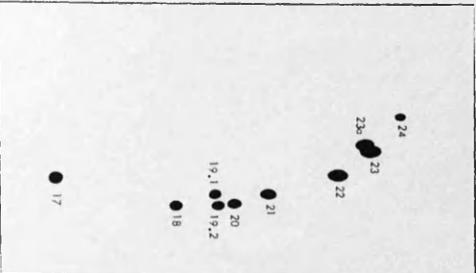
d) 5.8 S



b) 28 S



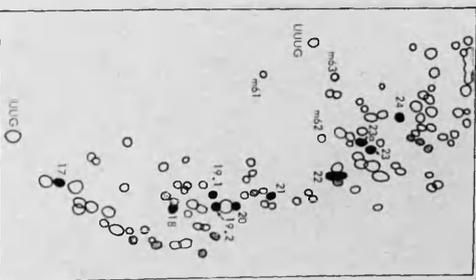
c) 28 H



d) 5.8 KEY



e) 32 S



f) 32 KEY

Table 4.2 Quantitation of 5.8S Products 18-24 in T<sub>1</sub> plus Phosphatase Fingerprints of 28S and 32S RNA.

Spot No.	Molar yields observed 28S "28H"	Deduced 5.8S in 28S (28S-"28H")	Observed 32S	Deduced 5.8S in 32S (32S-"28H")	Suggested frequency in 32S
18	0.93	0.93	1.25	1.25	~ 1
19.1	0.85	0.85	1.09	1.09	~ 1
19.2 ↓ 20	3.52	1.77	4.14	2.39	~ 1
21	0.70	0.70	0.93	0.93	~ 1
22	3.11	0.74	3.55	1.18	~ 1
23	1.68	0.79	1.95	1.06	~ 1
23a	1.60	0.71	1.81	1.10	~ 1
24	1.88	0.93	1.92	0.99	~ 1

Spots corresponding in mobility to 5.8S spots 18-24 were excised for scintillation counting. The spots were quantitated on the basis of their sequences and relative to three methylated products m61, m62 and m63 (see section 3.3). To determine the yield of 5.8S products in 28S and 32S fingerprints it was necessary, except for spots 18, 19.1 and 21, to determine the difference between 28S and "28H" fingerprints or between 32S and "28H" fingerprints respectively (third and fifth columns). The overlapping "28H" products may be chemically different from the 5.8S ones but interfere with quantitation. The 32S complex representing spot 22 may include partial contamination by a previously characterised 32S transcribed spacer product (see section 3.2).

in the 28S fingerprint which were less strongly labelled in the "28H" fingerprint. These four products do not possess unique mobilities but occupy positions to which other oligonucleotides from the "28H" molecule also run. The 5.8S products are indicated in the 28S fingerprint by arrows (see Fig. 4.2b). The arrows in the "28H" fingerprint (Fig. 4.2c) indicate the absence of products 18-21 and the positions of the "28H" oligonucleotides which occupy the same positions in the fingerprint as products 22-24.

The presence of products 18-24 in 28S fingerprints when compared with "28H" fingerprints was confirmed by scintillation counting (table 4.2). Spots 18, 19.1 and 21 occur approximately once per mole of 28S RNA and not at all in "28H" RNA. For the remaining spots, the difference in radioactivity between 28S and "28H" fingerprints indicated the loss of approximately one mole of each presumed 5.8S product from "28H" RNA.

#### 4.3 The 5.8S Sequence in 32S RNA.

The 5.8S products 18-24 were found in  $T_1$  plus phosphatase fingerprints of 32S RNA as determined by visual inspection (arrows in Fig. 4.2e) and by scintillation counting (table 4.2). 28S RNA and 32S RNA show similar quantitative data, both RNA's differing from "28H" RNA by approximately one mole of each of the 5.8S products.

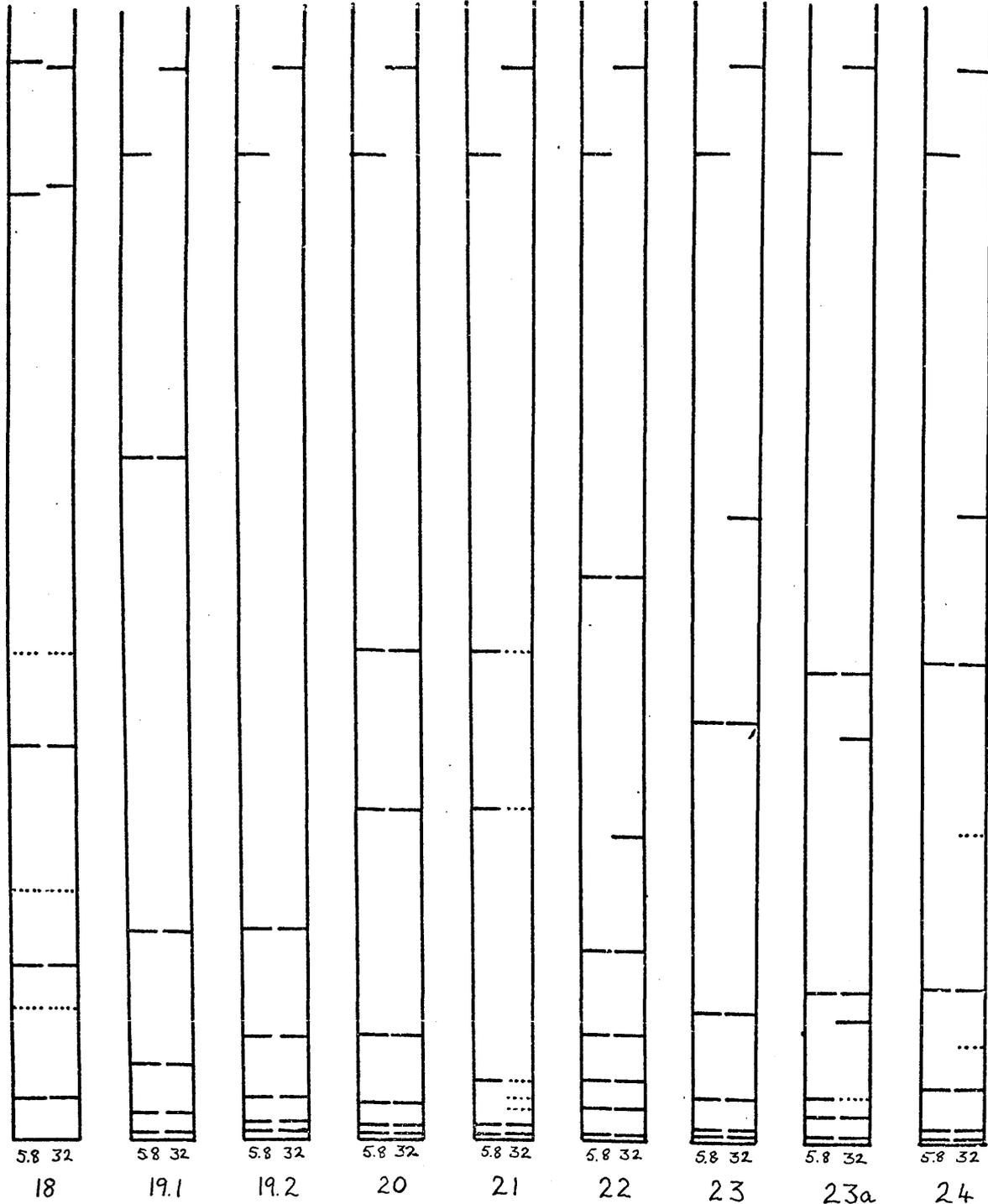


Fig. 4.3 Line diagram of the fractionation of VPDE partial products of 5.8S spots 18-24 from 5.8S and 32S fingerprints. The partial products were separated on DEAE-paper at pH 3.5 (1.1 KV, 16h).

Oligonucleotides from the 5.8S fingerprint were digested and fractionated on a separate occasion from those from the 32S fingerprint. Thus the difference in mobility of G is due to a slight difference in the pH of the buffer used. Dashed lines indicate faint products. Spot 21 is obtained reasonably pure from the 32S fingerprint and the faint extra products produced on partial digestion are probably artefactual. Spots 22-24 from the 32S fingerprint contain extra degradation products due to the contaminating "28H" oligonucleotides.

Product 17 (UCUG) is too small to serve as a distinctive 5.8S "marker" oligonucleotide. The data in table 4.2 suggest that it also occurs (twice) in the transcribed spacer region of 32S RNA. Product 22 lies in a complex cluster of spots which is further complicated in the 32S fingerprint by an overlapping transcribed spacer oligonucleotide. Quantitation suggests that the 5.8S product is present in this group.

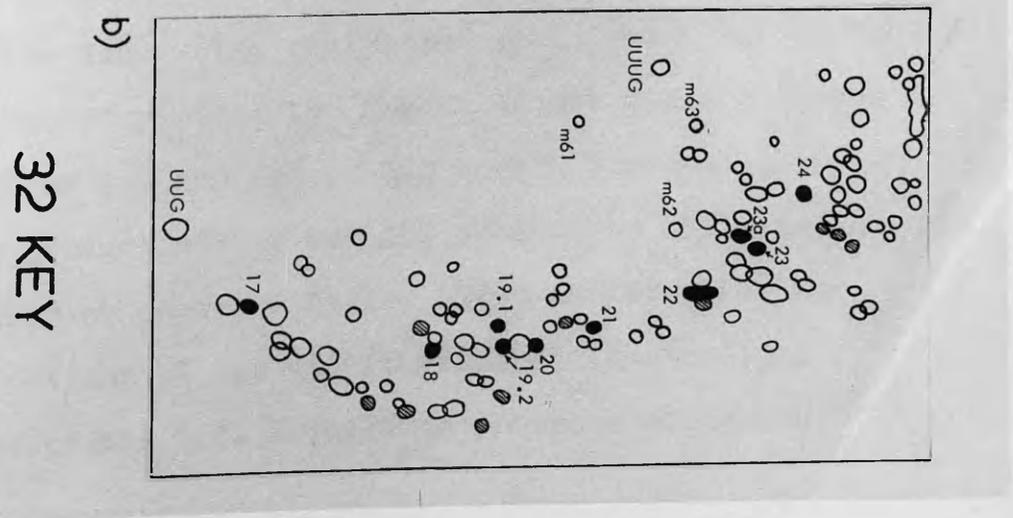
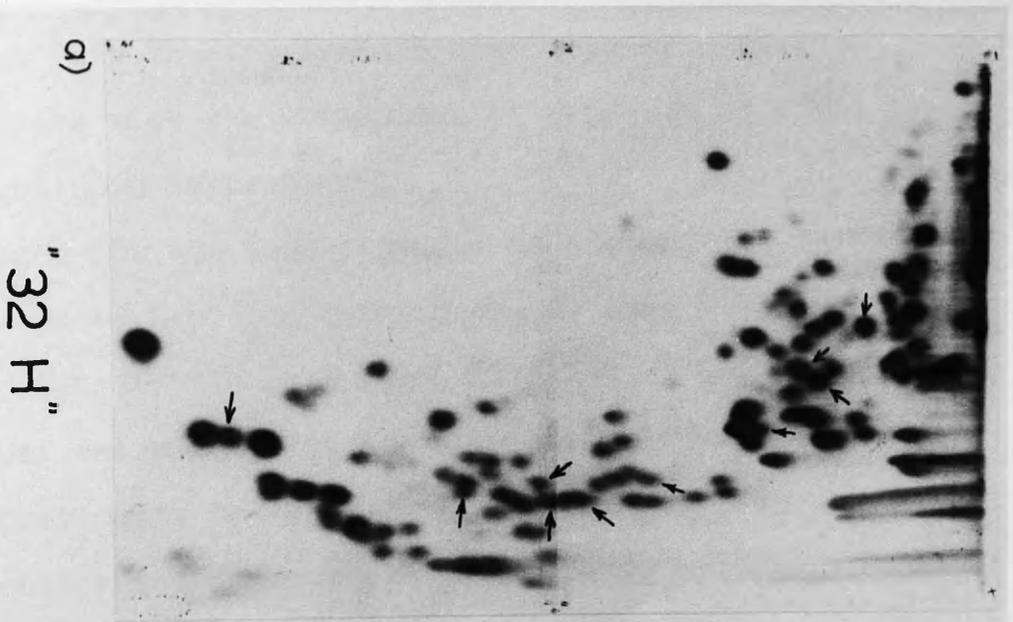
Further proof was obtained that the products, identified in 32S fingerprints as 5.8S products 18-24, possess identical sequences to the 5.8S products and were not merely of the same mobility. This was accomplished by comparing venom phosphodiesterase (VPDE) partial digestion patterns of the 32S oligonucleotides with those of the 5.8S products (Fig. 4.3). These patterns are representative of the nucleotide sequence (Brownlee & Sanger, 1967; see also section 3.3). The results indicate that the products identified in the 32S fingerprint possess identical sequences to the 5.8S products 18-24. The oligonucleotides which could be obtained in reasonably pure form from 32S RNA fingerprints (nos 18, 19.1, 19.2, 20 and 21) yielded an array of degradation products identical to those obtained for the oligonucleotides from purified 5.8S RNA (Fig. 4.3). Those spots (nos. 22, 23, 23a and 24) which could not be obtained pure from the 32S fingerprint yielded a pattern which contained a few extra products in addition to the

5.8S degradation products. These extra degradation products arose from the contaminating spots of the 32S fingerprint. In addition to this evidence, spot 23a from 32S and 28S fingerprints yielded the alkali-stable dinucleotide GmC after alkaline hydrolysis. Thus by the combined criteria of mobility, quantitation with respect to "28H" RNA and further degradation steps, all of the 5.8S products 18-24 are present in  $T_1$  plus phosphatase fingerprints of 32S RNA.

It has been shown that 5.8S RNA could not be released from 32S RNA by treatment which normally disrupts hydrogen bonding (Pene et al, 1968). Fingerprints of heat-treated 32S RNA (60°C, 4 min) still retained the 5.8S products as determined by visual inspection and quantitation (Fig. 4.4 and table 4.2). The results in table 4.2 were in fact obtained from fingerprints of heat-treated 32S RNA. This indicates that the 5.8S sequence is firmly attached to the 32S molecule and implies (though does not rigorously prove) that the linkage is covalent.

It is interesting to note the derivation, from within the 32S molecule, of some oligonucleotides which are clustered together in the fingerprint. For example, in the cluster of 3 spots of which 5.8S product no. 18 is one, one oligonucleotide is from the 5.8S region, one is from the "28H" region and the third is from the transcribed spacer region.

Fig. 4.4  $T_1$  plus phosphatase fingerprint of heat-treated 32S RNA. (a) heat-treated 32S RNA, (b) key. Conditions of electrophoresis are as in Fig. 3.2. The arrows indicate the 5.8S products 17-24. A few faint extra products are present due to HnRNA contamination (see section 3.4).



#### 4.4 The 5.8S Sequence in 45S RNA.

T<sub>1</sub> plus phosphatase fingerprints of 45S RNA are very complex (see section 3.5) and were not subjected to the type of analysis described above for 32S RNA. Well resolved spots with the mobilities of products 19.1 and 21 were evident upon inspection.

Evidence for the 5.8S sequence in 45S RNA was derived from the work of Drs. Salim and Maden in this laboratory on fingerprints of <sup>14</sup>C methyl-labelled RNA. 5.8S product 23a, which contains the alkali-stable dinucleotide GmC, was previously recognised as a 28S methylated oligonucleotide (spot no. 66, Maden & Salim, 1974). In methyl-labelled 28S+ 18S fingerprints, product 23a overlaps in mobility an 18S methylated product containing the dinucleotide AmU. The equivalent spot from a methyl-labelled 45S fingerprint yielded two alkali-stable dinucleotides, identified as GmC and AmU. GmC does not occur in any other 28S product nor in any 18S product in this part of the fingerprint (Maden & Salim, 1974) and its presence in this position of the 45S fingerprint is therefore diagnostic of the 5.8S methylated sequence AAUUGmCAG.

## Chapter 5

### COMPARISON OF HELA AND L CELL TRANSCRIBED SPACER SEQUENCES

COMPARISON OF HELA AND L CELL TRANSCRIBED  
SPACER SEQUENCES

5.1 Fingerprints of HeLa and L cell 28S and 5.8S RNA.

A further possible application of the type of fingerprinting analysis described in this thesis is for interspecies comparisons. It has been shown for two related species Xenopus laevis and Xenopus mulleri, that the ribosomal RNA's are similar with respect to size, base composition and nucleotide sequence as determined by RNA-DNA hybridisation (Brown et al, 1972). On the other hand the RNA precursor molecules from the two species are of the same size but differ in their nucleotide sequence as determined by RNA-DNA hybridisation. These differences must be in the transcribed spacer regions of the precursor molecule. I decided to compare ribosomal and ribosomal precursor RNA from HeLa cells with the RNA from a related species using the fingerprinting technique. Mouse L cell RNA was chosen because the nucleolar RNA could be extracted with minimal modification of the technique which was currently in use for the preparation of HeLa cell rpre RNA (see Methods).

$T_1$  plus phosphatase fingerprints of HeLa and L cell 28S RNA appear to be almost identical (Fig. 5.1a & b). In the "two U" graticule only 2 spots are unique to the HeLa 28S fingerprint and only 4 spots are unique to the L 28S fingerprint, whereas 46 spots are common to both fingerprints. In the "three U" graticule, approximately

Fig. 5.1  $T_1$  plus phosphatase fingerprints of 28S RNA from (a) HeLa cells and (b) L cells. Conditions of electrophoresis are as in Fig. 3.2. The arrows indicate spots which are unique to (a) the HeLa 28S fingerprint and (b) the L 28S fingerprint.

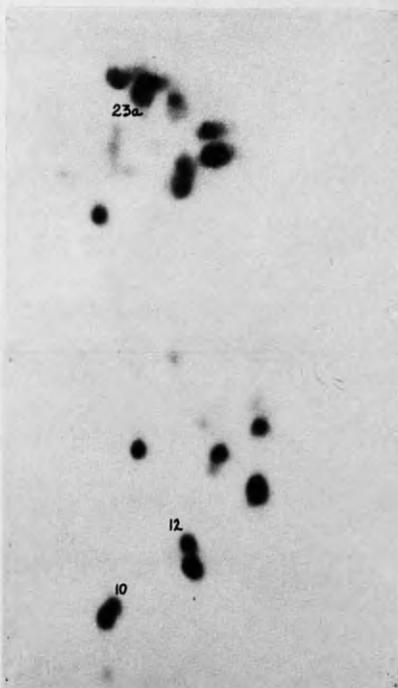
Fig. 5.2 Long  $T_1$  fingerprints of 5.8S RNA from (a) HeLa cells and (b) L cells. Conditions of electrophoresis are as in Fig. 3.2. In the L 5.8S fingerprint, products 10 and 12 contain  $\psi$  and product 23a contains the alkali-stable dinucleotide GmC as do the three respective products in the HeLa 5.8S fingerprint (see section 4.1).



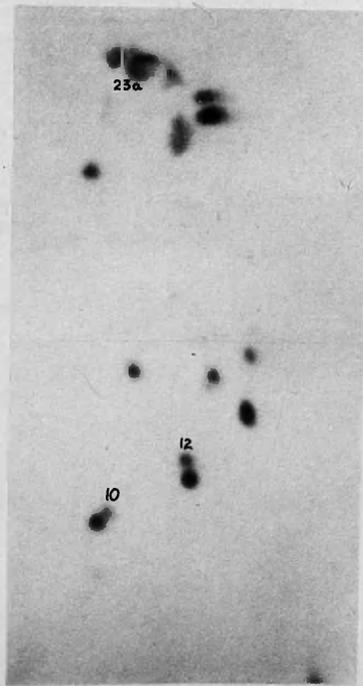
a) HELA 28 S



b) L 28 S



a) HELA 5.8 S



b) L 5.8 S

6 spots are unique to the HeLa 28S fingerprint and approximately 3 spots are unique to the L 28S fingerprint, whereas approximately 26 spots appear to be common to both. The differences in the "two and three U" graticules of these fingerprints are indicated by arrows in Fig. 5.1.

The results presented in Chapter 4 indicate that the 5.8S RNA sequence is separate from the 28S sequence and is situated within the transcribed spacer region of 32S RNA. Therefore it was interesting that long  $T_1$  fingerprints of 5.8S RNA from HeLa and L cells appear to be identical (Fig. 5.2). The two pseudouridylate containing products of HeLa 5.8S RNA, spots 10 and 12, are present in the L 5.8S fingerprint and were also shown to contain pseudouridylic acid. In addition, spot 23a for L cell 5.8S RNA contains the alkali-stable dinucleotide GmC as does the equivalent HeLa 5.8S product.

In as much as the positions of the spots in the fingerprints reflect the actual nucleotide sequence, the 28S and 5.8S RNA sequences from HeLa and L cells appear to have been highly conserved through evolution. Have the primary structures of the transcribed spacer regions also been highly conserved?

## 5.2 Fingerprints of HeLa and L Cell 32S RNA.

$T_1$  plus phosphatase fingerprints of HeLa 32S RNA used in this analysis resolved 9 of the previously characterised transcribed spacer products. Spots 12 and

Fig. 5.3  $T_1$  plus phosphatase fingerprints of HeLa 32S and L 32S RNA. (a) HeLa 32S, (b) L 32S, (c) HeLa 32S key, (d) HeLa + L 32S key, (e) L32S key. Conditions of electrophoresis are as described in Fig. 3.2. In the HeLa 32S fingerprints run in parallel with L 32S fingerprints for this comparison, products 12 and 13 did not resolve well and although they are visible in the HeLa 32S fingerprint shown in this figure, they have not been included in the keys.

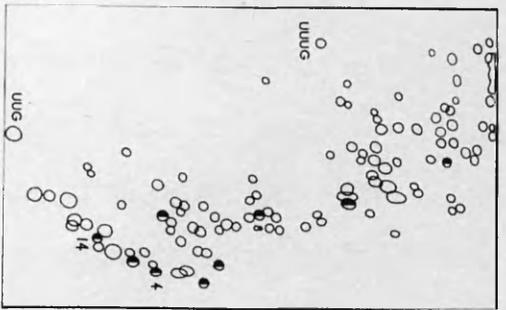
Symbols:            ①      HeLa transcribed spacer products  
                         ②      L transcribed spacer products  
                         ③      HeLa and L transcribed spacer products.



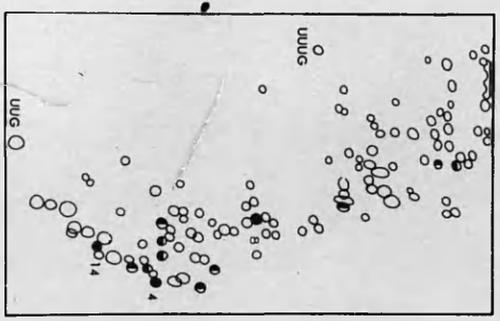
d) HELA 32S



b) L 32S



c) HELA KEY



d) HELA + L KEY



e) L KEY

13 did not appear as unique products in the 32S fingerprint as mentioned earlier (section 3.5). Comparison of  $T_1$  plus phosphatase fingerprints of L cell 28S and 32S RNA revealed some 7 products which were unique to the 32S fingerprint (Figs. 5.1b and 5.3b & e). These extra products must be derived from the transcribed spacer region of L cell 32S RNA. Only 3 of the transcribed spacer products (nos. 4, 8 and 14) are common to both HeLa 32S RNA and L 32S RNA, illustrated in the combined key to both HeLa and L 32S RNA (Fig. 5.3d). Six transcribed spacer products are unique to HeLa 32S RNA and 4 are unique to L 32S RNA.

No sequence information is available to date on the unique transcribed spacer products from L 32S RNA and the extent of the difference of their sequences from the HeLa products is unknown. One nucleotide substitution may account for the difference in mobility observed between some of the HeLa and L transcribed spacer products. However, the results indicate that the transcribed spacer sequences appear to have been less highly conserved than the ribosomal sequences (see also Discussion).

## Chapter 6

### DISCUSSION

## DISCUSSION

### 6.1 Practical Considerations.

#### RNA

A prerequisite for an analysis of high molecular weight RNA by fingerprinting was that the RNA sample be pure. In section 3.1 I have described a method for preparing 32S RNA of high purity (the actinomycin chase procedure) and in section 3.4 I have described the difficulties encountered when an RNA sample is not pure. The attempts to further purify 45S RNA by centrifugation in low ionic strength buffer failed. It may be possible to separate 45S from HnRNA on basis of their different GC contents, for example by equilibrium centrifugation in  $\text{Cs}_2\text{SO}_4$  gradients.

#### Radioactivity.

A major problem in RNA sequence analysis is the quantity of radioactivity available in the RNA species. This is determined by the quantity of RNA obtainable and its specific activity. One burler of HeLa cells ( $\sim 10^8$  cells) yields 40-50  $\mu\text{g}$  of 32S RNA. On addition of 10 mCi  $^{32}\text{P}_i$  to a burler, nucleolar RNA could be extracted with a specific activity of 1 mCi/mg. This value could be increased slightly by the addition of 20 mCi  $^{32}\text{P}_i$  to a burler, however in order to double the yield of radioactivity it was more efficient to label two burlers of cells with 10 mCi each rather than label one burler with 20 mCi. The specific activity of

nucleolar RNA is high because it is a kinetic intermediate in rRNA biosynthesis, which turns over fairly rapidly. However it constitutes only ~5% of the cellular RNA and only small amounts are obtained, especially of the 45S species, relative to rRNA. Ribosomal RNA is obtained with a lower specific activity, but the total amount of radioactivity is greater because of the large quantity of rRNA in the cell.

### Fingerprinting.

It is apparent from this report that "long"  $T_1$  plus phosphatase fingerprints are very useful in analysing high molecular weight RNA. This was first realised during the analysis of the 32S transcribed spacer region. On  $T_1$  digestion of a large RNA molecule the majority of the products are likely to be short oligonucleotides and few of these are likely to be unique. For a large RNA containing 35% G, it can be calculated (see Appendix) that after complete  $T_1$  digestion, approximately 11% of the products should be hexanucleotides or larger products and in a molecule the size of 32S RNA, the number of such products will be a little over 200. Thus in "long"  $T_1$  plus phosphatase fingerprints, in which most of the products will be hexanucleotides or larger, many of them should occur once per molecule and many, but not all, should be resolved from neighbouring products (isomers and other related products) in good quality fingerprints.

Normally  $T_1$  plus phosphatase fingerprints were reproducible. However occasionally streaky overdigestion

products were present due to undesirable endonuclease activity in the alkaline phosphatase, and in some digests action of phosphatase was incomplete. Because relatively little nucleolar RNA is obtained from the cells, loss of the RNA in bad fingerprints was costly in terms of the time spent in preparing more RNA. The alkaline phosphatase used was the purest commercially available, the purity of the enzyme varying between batches.

Large  $T_1$  products could also be resolved on thin layers of DEAE-cellulose though this method was less reproducible and depended to a large extent on the quality of the commercially obtained thin layers.

#### Sequence Analysis.

Sequence analysis of the transcribed spacer products which were isolated was difficult due to their high pyrimidine content. The number of cytidine residues in some of the products was difficult to determine accurately by base composition analysis only and these may be confirmed by other methods (eg by analysis of partial venom digestion products). However in a few cases (see below) the number of cytidine residues remains ambiguous, eg. after separation of the partial venom products of 32S spot no. 11 (see Fig. 3.5) an additional degradation product may be present unresolved from the unchanged oligonucleotide. This would indicate the presence of an additional residue (a cytidine residue) in the original oligonucleotide.

Attempts to sequence the 5' end of the oligonucleotides by partial digestion with spleen phosphodiesterase failed,

probably due to the inhibitory action of cytidine residues on spleen phosphodiesterase. At the time of sequencing these transcribed spacer products the technique of digestion with  $U_2$  RNase was introduced to this laboratory, however initial attempts with this enzyme failed. Digestion with  $U_2$  RNase would have been useful with the few A containing products from the 45S transcribed spacer region. In the later stages of this work, identification of the 5' end nucleotide by comparing alkaline hydrolysis and total venom digestion data was not very accurate. This was due to phosphatase activity in recent batches of venom phosphodiesterase and affected only the sequencing of the 45S transcribed spacer products. Endonuclease activity has also been detected in the venom phosphodiesterase (McLennan & Lane, 1968). This could cause secondary splitting of a product during partial digestion with venom phosphodiesterase. The faint product lying between partial products 8c and 8d in Fig. 3.5 is likely to be an example of this.

I consider the partial sequence analyses which I have obtained for the 32S and 45S transcribed spacer products to be accurate with the exception that the number of cytidine residues in 32S spots 11, 12 and 13 may be ambiguous for reasons stated above.

For the purposes of what follows, certain general conclusions are more important than the actual nucleotide

sequences. The products contain no modified bases and their existence provides direct proof of the transcribed spacer regions. The sequence data were sufficient to establish the size of the products and thereby permit quantitation.

## 6.2 Inferences on HeLa Cell Transcribed Spacer Regions.

It is interesting to compare the properties of the distinctive products of the 32S transcribed spacer region and of the 45S unique transcribed spacer region (ie that part of the 45S transcribed spacer region which does not appear in 32S RNA) with properties which may be inferred, from indirect evidence, as characterising these regions as a whole.

The available parameters are summarised in table 6.1. If, as is likely, the data in the first four columns of the table are correct then the following information may be deduced. The 32S transcribed spacer region comprises some 1,350-1,650 nucleotides, is very rich in G+C and is substantially chemically unmodified. Similarly the 45S unique transcribed spacer region comprises some 4,000 nucleotides, is very rich in G+C and is substantially chemically unmodified. The 32S products (nos. 1-14) and the 45S products (nos. 20-28) comprise some 154 and 77 nucleotides respectively and, as I have shown, are rich in C ( $T_1$  digestion products contain only one G residue of course) and lack modified components. The products range in size from pentanucleotides to dodecanucleotides and several are

Table 6.1 Inferred Properties of the Transcribed Spacer Regions and Composition of Material Recovered in Products 1-14 and 20-28.

	45S	32S	28S	18S
molecular wt. <sup>1</sup>	4.1x10 <sup>6</sup>	2.1x10 <sup>6</sup>	1.65x10 <sup>6</sup>	0.65x10 <sup>6</sup>
Approx. number of nucleotides <sup>2</sup>	12,300	6,300	4,950	1,950
G <sup>3</sup>	34.2%	35.3%	35.0%	30.5%
A	12.5%	13.6%	16.0%	20.2%
C	36.7%	34.8%	32.3%	27.3%
U	15.7%	15.5%	15.6%	20.5%
ψ	0.85%	0.90%	1.10%	1.52%
Number of methyl groups <sup>4</sup>	~ 111	~ 71	~ 71	46

- Weinberg & Penman, 1970.
- The number of nucleotides in an RNA may be estimated approximately by assuming 3 nucleotides per 1000 daltons.
- The base composition data of 45S, 32S, 28S and 18S RNA are from Jeanteur et al (1968). Values for 32S transcribed spacer and 45S unique transcribed spacer were calculated using the molecular weights shown above. The base composition values differ from those obtained by Jeanteur et al (1968) on account of the different values of the molecular weights used.
- Salim & Maden, 1973; Maden & Salim, 1974.

32S transcribed spacer		Products 1-14 <u>Nucleotides recovered</u>		45S unique transcribed spacer		Products 20-28 <u>Nucleotides recovered</u>	
32S-28S		Number	% of total	45S-(32S+18S)		Number	% of total
0.45x10 <sup>6</sup>				1.35x10 <sup>6</sup>			
	1,350	153	11%		4,050	77	2%
36.4%	491	21	4%	34.3%	1389	11	0.8%
4.6%	62	6	10%	7.1%	288	6	2%
44.2%	597	85	14%	44.2%	1790	30	2%
15.1%	204	40	20%	13.7%	555	30	5.5%
0.14%	0-2	0	-	0.45%	0-20	0	-
	0	0			0	0	

recovered in yields which suggest that they in fact occur once per molecule of rpre RNA.

If the 32S transcribed spacer region is assumed to comprise 1,350 nucleotides (table 6.1) about 30% of them (or 400 nucleotides) might be expected to be liberated as hexanucleotides or larger products (see Appendix) and many of these should be resolved from other material in the fingerprints. Similarly, if the 45S unique transcribed spacer region is assumed to comprise 4,000 nucleotides (Table 6.1) about 30% of them (or 1,500) might be expected to be liberated in the form of hexanucleotides or larger products. I have resolved some 123 nucleotides from the 32S transcribed spacer region ( $\sim 8\%$  of the presumed total) and some 67 nucleotides from the 45S unique transcribed spacer region ( $\sim 1.7\%$  of this region) in the form of hexanucleotides or larger products. The likelihood of smaller products giving rise to distinctive spots in the fingerprints is much lower. However the fact that the pentanucleotide UCCCG occurs several times per mole of 32S transcribed spacer material, but not in the 28S sequence itself, is remarkable. Many of the transcribed spacer products will be found as quantitative rather than qualitative differences between the rpre RNA and the rRNA fingerprints. These are particularly evident along the "C rich" or right hand edge of the "two U" graticule of the 45S fingerprint (see fig. 3.7).

In so far as it is justifiable to draw conclusions

from evidence derived from limited parts of the molecule, the molar yields of products 1-14 and 20-28 suggest a fairly high degree of sequence homogeneity in the transcribed spacer regions. Animal cells possess several hundred copies of the genes for rRNA (1,100 in HeLa cells, Jeanteur & Attardi, 1969), the presence of so many undoubtedly being related to the rate at which rapidly dividing cells must produce ribosomes ( $\sim 10^2$  per second per HeLa cell). These multiple genetic units for rRNA are believed to consist of identical or closely similar sequences on the basis of renaturation kinetics of Xenopus rDNA (Birnstiel et al, 1969). My results are consistent with the view that this homogeneity extends to the transcribed spacer regions of rpre RNA.

Models for the evolution of such tandem genes are discussed by Brown et al (1972). Such models must explain the maintenance of homogeneity of these genes within a species but permit their divergence between species (see also section 6.4). The "magnification" or disproportionate replication process (Tartof, 1971) and the "master-slave" hypothesis (Callan, 1967) allow the horizontal spread of a mutation through adjacent genes only. A third method involves amplification of one or more of a limited number of chromosomal gene copies followed by recombination with chromosomal rDNA. This would allow the parallel evolution of unlinked clusters of rDNA gene copies.

### 6.3 The 5.8S Sequence.

Products 18-24 from  $T_1$  plus phosphatase fingerprints of 5.8S RNA (see chapter 4) range in size from a hexanucleotide to a nonanucleotide and represent  $\sim 45\%$

of the molecule ( $\sim 155$  nucleotides in the total molecule, Maden & Robertson, 1974). It was calculated (see Appendix) that about 60 ( $\sim 11\%$ ) of the  $T_1$  digestion products of the 32S transcribed spacer region should be hexanucleotides or larger products. It is extremely unlikely that the nine 5.8S oligonucleotides (18-24) would all occur "by chance" in 32S fingerprints by being scattered at random through the transcribed spacer region of the molecule. This is especially so when one of the longest products is methylated in the same position in both 5.8S and 32S RNA. It is concluded that the 5.8S sequence is present as such in 32S RNA.

The data are less complete for 45S RNA but identification of GmC in the correct position of  $^{14}\text{C}$  methyl fingerprints as well as spot 19.1 which is particularly well resolved in  $^{32}\text{P}$  fingerprints, indicates strongly that the 5.8S sequence is also present within 45S RNA and is transcribed as part of the 45S molecule. Location of the 5.8S sequence within the precursor will be discussed later in section 6.6.

#### 6.4 Precursor RNA from Other Eukaryotes.

Determination of the extent and kinetics of hybridisation of rRNA to DNA from a variety of species indicated that there is a great deal of variability in rRNA sequences among widely different organisms whilst the rRNA sequence of related species is evidently highly similar (Birnstiel & Grunstein, 1972). To what extent have the transcribed spacer sequences in ribosomal precursor

RNA of related species also been conserved? Presented in chapter 5 are the preliminary results of an investigation comparing HeLa cell 32S rpre RNA with mouse L cell 32S rpre RNA.

Before interpreting these results it should be pointed out that a comparison of  $T_1$  plus phosphatase fingerprints will provide only a crude comparison of the nucleotide sequence of the RNA as an equivalence in position of any two spots does not, in the absence of further evidence, conclusively indicate an identical nucleotide sequence. Also, when considering spots which have differing mobilities in a fingerprint it should be noted that a substantial change in mobility of any spot can be caused by merely one base change within the oligonucleotide. For example, a change of base from C $\rightarrow$ U would involve a change in mobility of the spot from one graticule to another.

Notwithstanding these limitations, the fingerprints presented in chapter 5 appear to indicate that the nucleotide sequences of HeLa and L cell 28S and 5.8S RNA are very similar (section 5.1). The transcribed spacer regions of HeLa and L cell 32S RNA, in comparison with the ribosomal sequences, have been less highly conserved during evolution (section 5.2). This conclusion agrees with the results obtained from a comparison of X. laevis and X. mulleri rRNA and rpre RNA by RNA-DNA hybridisation (Brown et al, 1972; Honjo & Reeder, 1973). Brown et al. (1972), who also show that the precursor molecule from the two species was of the same size, suggest that the length

Table 6.2 Molecular Weights of the "Ribosomal Transcription Unit" and of rRNA from Various Organisms.

	Molecular Wt. x 10 <sup>-6</sup>		% of transcription unit as transcribed spacer
	Transcription Unit	rRNA (28S+18S)	
Mouse	4.19	2.35	44
Potoroo (Marsupial)	4.19	2.35	44
Chicken	3.92	2.24	43
Iguana	2.74	2.13	22
Frog	2.76	2.19	21
Trout	2.70	2.20	19
<u>Drosophila</u>	2.85	2.05	28
Tobacco	2.76	1.95	29

Adapted from Perry et al (1970).

of the transcribed spacer region may be more important for an organism's survival than its exact nucleotide sequence.

Previous work has shown that whilst the size of rRNA has changed only slightly in size through evolution (Loening, 1968; and Table 1.2 facing p.3), the size of the precursor increases markedly between reptiles and birds/mammals (Perry et al, 1970; Grierson et al, 1970). Thus the amount of transcribed spacer in rpre RNA is about 45% in mammals and birds and about 20-30% in lower animals (table 6.2).

It has been suggested that this increase in size of the transcribed spacer region is the result of a shift in the initiation or termination sites of transcription and not in the length of the DNA cistron (Perry et al, 1970; Grierson et al, 1970). This predicts that the amount of the non-transcribed spacer DNA in the ribosomal cistrons in mammals is much less than that in lower forms. The extra transcribed spacer region in mammals will appear at the 3' (18S) end of the precursor molecule (see Fig. 6.1 and Grierson et al, 1970).

#### 6.5 Precursor RNA from Bacteria.

The existence in E. coli of a precursor to 16S rRNA and of a precursor to 23S rRNA has been demonstrated by pulse-labelling experiments with radioactive uracil (Adesnik & Levinthal, 1969; Dahlberg & Peacock, 1971). The relative mobilities of the 16S RNA species in polyacrylamide gels suggest that 16S rpre RNA contains approximately 200 extra nucleotides. 23S rpre RNA does not appear to contain

any extra nucleotides, conversion of 23S rpre RNA to mature 23S rRNA possibly involving alterations to the secondary and tertiary structure only (Dahlberg & Peacock, 1971).

The presence of extra nucleotides in 16S rpre RNA has been demonstrated by RNA fingerprinting (Sogin et al, 1971; Brownlee & Cartwright, 1971; Lowry & Dahlberg, 1971; Hayes et al, 1971).

Previously it was thought that there was no common precursor in bacteria though recently a 30S rpre RNA species was discovered in a mutant of E. coli deficient in RNase III (Nikolaev et al, 1973). This 30S RNA molecule was shown by two different techniques, RNA-DNA competitive hybridisation and molecular weight determination by polyacrylamide gel electrophoresis, to contain 27% of its length as transcribed spacer in addition to the 16S and 23S rRNA sequences. When incubated with highly purified RNase III, pure 30S rpre RNA is cleaved into two RNA species which comigrate in gels and sucrose gradients with 23S and 16S rpre RNA's seen in vivo. Thus bacteria appear to make rpre RNA similar in some respects to that observed in eukaryotes, the size of the initial transcript in E. coli being only slightly smaller than that of plants and lower eukaryotes (Nikolaev et al, 1973; Perry et al, 1970; Grierson et al, 1970). In normal E. coli cells the rpre RNA is ordinarily cleaved endonucleolytically during its formation.

## 6.6 Topological Map of 45S RNA.

What is known about the spatial arrangement of the ribosomal sequences, the transcribed spacer regions and the

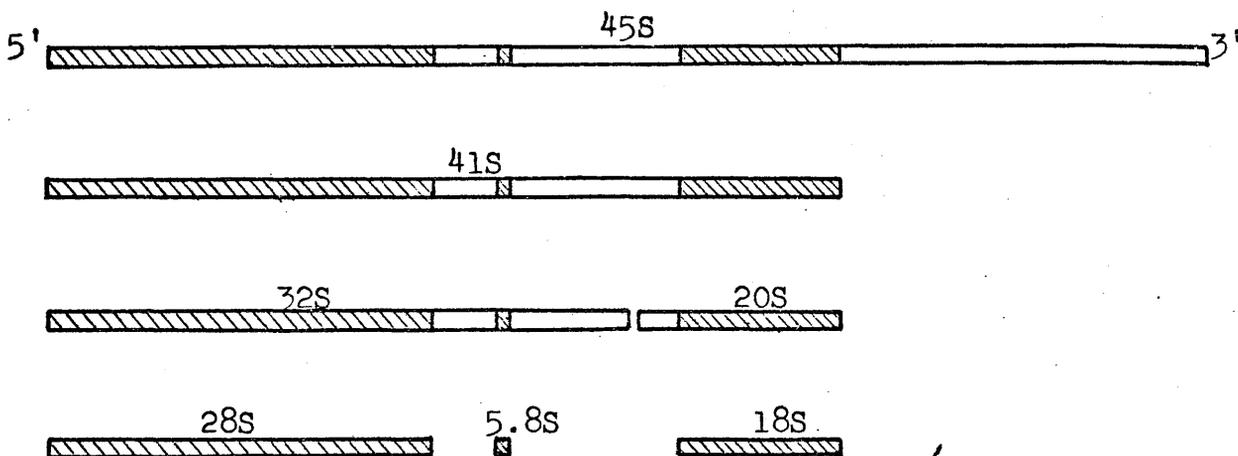


Fig. 6.1 Topological map and maturation scheme for 45S RNA. Since the transcribed spacer regions are not found unattached to rpre RNA, it is possible that they are removed by exonuclease action following preliminary cleavage of rpre RNA where necessary by an endonuclease. The 5.8S sequence is found hydrogen bonded to the 28S molecule and a loop of transcribed spacer may be envisaged between these sequences.

(Adapted from Wellauer & Dawid, 1973).

5.8S sequence within the 45S RNA molecule? A topological model (depicting the spatial arrangement) and maturation scheme for mammalian 45S RNA are shown in Fig. 6.1. The main features of this model are based on secondary structure maps of HeLa rRNA and rpre RNA (Wellauer & Dawid, 1973). The RNA molecules displayed a highly reproducible secondary structure of hairpin loops after they were spread and examined in an electron microscope. These double-stranded regions are likely to be related to the high GC content of the RNA molecules. Absence of further secondary structure is not indicative of their absence in vivo. The polarity of the molecules was established by Wellauer & Dawid by partially digesting 28S or 45S RNA with a 3'-OH exonuclease (Perry & Kelley, 1972). Electron micrographs of the product revealed that the molecules had lost portions from one end only. This end was thus the 3' end. This result relies heavily on the specificity of the exonuclease. In their own experiments, Perry & Kelley (1972) used this exonuclease to demonstrate the presence of long methyl deficient segments at the 3' end of both 32S and 45S RNA. The findings of Wellauer & Dawid (1973) do not bear directly on the position of the 5.8S sequence. Recently evidence was obtained from RNA-DNA hybridisation experiments with Xenopus laevis (Spiers & Birnstiel, 1974) that the 5.8S sequence is situated between the 28S and 18S sequences within ribosomal DNA.

Knowledge of end groups would be useful in determining

the spatial arrangement of the ribosomal sequences within the precursor molecules. Few end groups have been identified during this study for rRNA and rpre RNA. If spot no. 30 (see section 3.5) is in fact the 5' end group of 18S rRNA then its absence in 45S fingerprints supports this model of 45S RNA with the 18S sequence located inside the 45S molecule. In his analysis of 5.8S RNA, Dr. Maden identified the 3' end group of 5.8S RNA as  $\text{CUU}_{\text{OH}}$ . A preliminary investigation revealed the absence of this end group in 32S fingerprints indicating that the 5.8S sequence is not situated at the 3' end of the 32S molecule. End group analysis was performed on Xenopus 40S rpre RNA and 28S and 18S rRNA species by searching alkaline hydrolysates of the RNA's for 5' triphosphates and by labelling in vitro the 5' ends of the RNA's using polynucleotide kinase. 40S rpre RNA and 28S rRNA were found to have different 5'-end groups, this being inconsistent with the topological model of mammalian precursor RNA depicted in Fig. 6.1. (Slack & Loening, 1974)

A major difference in the formation of rRNA between prokaryotes and eukaryotes is the arrangement of the ribosomal sequences within the precursor molecule. The evidence described above indicates that in eukaryotes the larger ribosomal RNA species (28S) is situated at or near the 5' end of the precursor molecule. In E. coli incorporation of label into rRNA after inhibition of RNA synthesis with rifampicin indicates that the small ribosomal RNA species (16S) is synthesised first, i.e. the smaller rRNA is situated at the 5' end of the precursor molecule (Bremer & Berry, 1971).

## 6.7 Secondary Modification of Precursor RNA.

Methylation and possibly pseudouridylate formation are restricted to the ribosomal and 5.8S sequences (Salim & Maden, 1973; Jeanteur et al, 1968; Maden & Forbes, 1972; Maden & Robertson, 1974). This suggests that they may be involved in the preservation of these sequences during maturation. Pseudouridylate formation takes place in the nucleolus at the level of 45S RNA (Maden & Forbes, 1972) though very little is known regarding its presence in precursor RNA.

Both eukaryotic and prokaryotic rRNA are methylated although the pattern of methylation in each type is different. In prokaryotes, methylation occurs mainly on the bases whilst in eukaryotes 95% of the methylation occurs on the 2'-O-ribose position (Fellner, 1969; Wagner et al, 1967; Vaughan et al, 1967).

Methylation is believed to occur on mammalian precursor RNA close to the point of transcription (Greenberg & Penman, 1966) but if the submethylation of 45S RNA reported in section 3.6 is real and not artefactual, ie. the 45S preparation contained 45S molecules which were still to undergo methylation, then methylation must occur a short time after transcription of 45S RNA is complete.

It is interesting that 5.8S RNA from HeLa and L cells contains the methylated dinucleotide GmC and that this dinucleotide is probably present in plant 5.8S RNA (unidentified product X of Woledge et al, 1974). However GmC was not found in yeast 5.8S RNA (Rubin, 1973).

Results from the use of mammalian cells suggest the involvement of methylation in the final stages of rRNA maturation. This evidence derives from methionine starvation of HeLa cells which causes a cessation of ribosome formation (Vaughan et al, 1967). Under these conditions nucleolar rpre RNA is severely deficient in 2'-O-methyl groups. 45S rpre RNA continues to be synthesised and to be converted to 32S RNA in the nucleolus. However the submethylated 32S RNA must be quickly degraded as no mature rRNA is formed. A correct level of methylation therefore appears to be required for some late step in rpre RNA maturation.

The high level of 2'-O-ribose methylation of mammalian rRNA may play some protective role during ribosome formation. RNA hydrolysis normally occurs via a 2'-3' cyclic intermediate. The presence of a 2'-O-methyl group confers resistance to the adjacent phosphodiester linkage to hydrolysis via a 2'-3' intermediate. From the high turnover rate of HnRNA, the nucleoplasm must contain high nuclease activity. It has been suggested that 2'-O-methylation protects exposed parts of the RNA in the newly formed ribosome from nucleases as the particle passes through the nucleoplasm on its way to the cytoplasm (Maden, 1971). Prokaryotes do not possess a nucleoplasm nor the associated high level of nucleases and thus do not require a high level of 2'-O-ribose methylation.

In mammalian systems, some methylation also occurs after maturation of rpre RNA is complete. This methylation occurs in the cytoplasm involving mainly the small (18S)

rRNA species, 6 methyl groups being added to 18S rRNA at this point (Salim & Maden, 1973). These 6 methylations are all base methylations, four of them being found in the sequence  $m_2^6Am_2^6ACUG$  and it is significant that a similar sequence containing  $m_2^6Am_2^6A$  is found in two unrelated organisms, yeast (Klootwijk et al, 1972) and E. coli (Fellner, 1969). This sequence is therefore probably widespread and may play some general role. The presence of a methylation process in the cytoplasm of eukaryotes indicates that these methylations are not involved in rpre RNA maturation.

In prokaryotes, methylation of 16S rRNA appears to occur relatively late in the ribosome maturation process (Lowry & Dahlberg, 1971). Ribosomal proteins bind initially to the 5' end of the 16S rpre RNA molecule (possibly whilst transcription is still taking place) and methylation occurs mainly in the 3' terminal 25% of the molecule (Lowry & Dahlberg, 1971; Fellner et al, 1972).

Another mutant of E. coli, a  $met^-$  strain, forms ribosomes from submethylated RNA (Beaud & Hayes, 1971). The absence of the missing methyl groups does not prevent ribosome assembly but does affect the ability of the ribosomes to function correctly. Further evidence that the doubly methylated sequence  $m_2^6Am_2^6A$  is not concerned with rRNA maturation arises from a mutant of E. coli which shows resistance to the drug kasugamycin (Helser et al, 1971, 1972). The site of action of this drug is the 30S ribosomal

subunit. The only detectable difference between the 30S subunit of kasugamycin sensitive and kasugamycin resistant strains was the absence of the methyl groups in the sequence  $m_2^6Am_2^6ACCUG$  in the resistant strain. This has been shown to be due to a mutation in a gene which codes for an RNA methylase (Helser et al, 1972). Ribosome assembly will take place on RNA which lacks these methylations. These experiments with mutants of E. coli suggest that methylation is not involved or at least not directly involved in the processing of the precursor RNA.

In summary, early methylation of eukaryotic precursor RNA takes place in the nucleolus and is predominantly 2'-O-ribose methylation. These methylations may play some protective role. Some late methylations take place in the cytoplasm. These are base methylations and are most likely to be involved in ribosome function. Methylation of bacterial rRNA is predominantly base methylation and the evidence suggests their involvement in ribosome function rather than formation.

## 6.8 Secondary Structure of Precursor RNA.

The existence of double stranded regions in the transcribed spacer of ribosomal precursor RNA has been demonstrated by electron-microscopy (Wellauer & Dawid, 1973) and by studies on the hyperchromicity of the RNA (Hadjiolov & Cox, 1973). The electron-micrographs were discussed earlier in section 6.6 and reveal extensive regions of hairpin loops in the transcribed spacer regions. In

mammalian rpre RNA, the G+C content of the transcribed spacer regions is very high (> 70%) (Jeanteur et al, 1968). This value is much higher than for the ribosomal regions (~ 60-65%). The conservation of double-stranded regions in the transcribed spacer during spreading for electron microscopy is probably related to the high G+C content of these regions (G-C base pairs are stronger than A-U base pairs). 18S rRNA which has <60% G+C shows no base pairing under the conditions of spreading.

The spectrophotometric analysis of rpre RNA and rRNA of Krebs ascites cells (Hadjiolov & Cox, 1973) revealed that the secondary structure content of 45S rpre RNA is high and is similar to that of 28S and 18S rRNA. The RNA's appear to consist of numerous hairpin loops with double helical segments containing an average of 14-16 base paired nucleotides.

What evidence exists to suggest that features of secondary structure of rpre RNA are involved in the specificity of RNA maturation? Much of this evidence derives from the prokaryotic system, from the use of a mutant of E. coli. As was mentioned earlier, prokaryotic rRNA is synthesised as a single large (30S) precursor molecule which was first identified in an RNase III deficient mutant of E. coli. It thus seems likely that RNase III is responsible for at least the initial cleavage of the 30S rpre RNA molecule. This enzyme acts endonucleolytically and is specific for double-stranded RNA (Robertson et al, 1968). This indicates the involvement of double-stranded regions in the processing of rpre RNA.

What evidence is there to suggest that double-stranded regions are involved in the processing of eukaryotic rpre RNA? Snyder et al (1971) have shown that the intercalating dyes proflavin, ellipticine and ethidium interfere with the processing of nucleolar rpre RNA in L 1210 mouse lymphoma cells. In the presence of one of these drugs, processing of 45S rpre RNA is reduced to a very low level. This effect is exerted on normally formed 45S RNA and is not secondary to inhibition of RNA or protein synthesis. These compounds are known to intercalate between base pairs in helical regions of DNA and this is probably true also for RNA. Such double-stranded regions must exist at the specific cleavage sites of 45S RNA.

More recently it has been shown that 5'-azacytidine interferes with the processing of 45S RNA in HeLa cells (Reichman et al, 1973). In the presence of 5'-azacytidine, 45S RNA continued to be synthesised but little 32S RNA was formed, most of the 45S RNA being degraded. It is very likely that this aberrant processing is due to the incorporation of the pyrimidine analogue into the precursor molecule and the consequent alteration of the precursor as a proper substrate either for the processing enzymes or for assembly with ribosomal proteins. This may be due to altered base pairing ability of 5'-azacytidine causing changes in the secondary structure of the precursor molecule.

Summarising, there are three separate lines of evidence that suggest the participation of double-stranded regions in the maturation of rpre RNA. (1) Double-stranded regions do in fact exist in the transcribed spacer regions, (2) the

identification of an endonuclease (RNase III) which is specific for double-stranded regions and which is involved in processing rpre RNA in *E. coli*, and (3) the inhibition of processing of rpre RNA by compounds which interfere with base-pairing.

## 6.9 Nucleolar Proteins and Nucleases.

### Nucleolar Proteins.

Ribosomal precursor RNA assembles, shortly after transcription, with ribosomal proteins and with nucleolar non-ribosomal proteins to form nascent ribosomes within the nucleolus (Kumar & Warner, 1972). Two ribonucleoprotein particles have been isolated from HeLa nucleoli, (1) an 80S particle which contains 45S rpre RNA, 5S RNA, the larger ribosomal subunit proteins and a few of the smaller ribosomal subunit proteins, and (2) a 55S particle which contains 32S rpre RNA, 5S RNA and the larger ribosomal subunit proteins (Warner & Soeiro, 1967; Shepherd & Maden, 1972). In addition, both particles contain a considerable amount of nucleolar non-ribosomal protein presumably associated with the transcribed spacer regions of the rpre RNA (Kumar & Warner, 1972). These proteins do not resemble ribosomal proteins and are not found in mature ribosomes. They remain stable within the nucleolus and are possibly reutilised in the assembly of nascent ribosomes. No function has as yet been attributed to these nucleolar proteins and it is quite likely that they contribute in some way to the specificity of processing of rpre RNA.

When protein synthesis is inhibited either rapidly with cycloheximide (Willems et al, 1969; Craig & Perry, 1970) or gradually by starvation for essential amino acids (Maden et al, 1969), both the synthesis and the maturation of rpre RNA are affected. Treatment with cycloheximide usually results in promptly decreased rates of both the synthesis and the maturation of rpre RNA whilst starvation for an essential amino acid produces a greater effect on the processing of the RNA than on its synthesis.

Similar effects are encountered with several temperature sensitive mutants of yeast in which there is a marked diminution in the rate of synthesis of rpre RNA and an almost complete inhibition of proper maturation at the restrictive temperature (Warner & Udem, 1972). The incorporation of proteins into nascent ribosomes is also greatly reduced. These mutants are presumed to carry genetic lesions affecting various proteins involved in the ribosome maturation process.

The inhibition of RNA maturation when protein synthesis is arrested or reduced lends support to the role of the nucleolar proteins in the processing of rpre RNA. However, an effect on the processing enzymes themselves or on some other regulatory proteins cannot be ruled out. In addition it is possible that the supply of some of these proteins may be rate limiting for rRNA synthesis or some other process rather than being involved directly in the maturation process itself.

#### Nucleases.

As yet there is no definitive evidence concerning the

enzymes responsible for rpre RNA processing in eukaryotes. In E. coli, as already discussed, RNase III has been implicated in the cleavage of 30S rpre RNA. This enzyme cleaves endonucleolytically and is specific for double-stranded RNA (Robertson et al, 1968). Another enzyme which has been extracted from E. coli, RNase II, acts both endonucleolytically and exonucleolytically (Spahr, 1964) and will promote the transformation of 17S rpre RNA to mature 16S rRNA (Corte et al, 1971). In a mutant of E. coli, temperature sensitive in RNase II, the 17S rpre RNA species accumulate at the restrictive temperature. This suggests RNase II is a processing enzyme.

In the eukaryotic system, a 3'-OH specific exoribonuclease, also mentioned above, has been isolated from the nuclei of mouse L cells (Perry & Kelley, 1972) and this enzyme will remove long methyl-deficient segments from the 3' end of 32S and 45S rpre RNA (see section 6.6). However this enzyme is found predominantly in the nucleoplasm and thus is unlikely to be specifically involved in rRNA maturation.

An enzyme has been isolated from HeLa cell nucleoli and will convert in vitro isolated nascent ribosomes containing 45S rpre RNA into particles containing RNA molecules which correspond in size to the intermediate precursor RNA's normally found in vivo (Mirault & Scherrer, 1972). This enzyme is reported to have endonucleolytic activity. However from a recent report, this endonucleolytic activity is apparently the effect of several enzymatic

activities of weak specificity (Kwan et al, 1974). There is so far little evidence to correlate these enzymes with the rRNA maturation process.

#### 6.10 Future Experiments.

Processing of ribosomal precursor RNA probably occurs at specific regions of secondary structure. An examination of the secondary structure of the transcribed spacer region especially around the cleavage points would be of considerable interest and may help us in our understanding of the relationship of the transcribed spacer region to the ribosomal sequences and the role of these regions in ribosome maturation.

Because of the size of the precursor molecules, such an examination by sequence analysis should be performed on fragments of the RNA and not on the intact molecules themselves. These fragments could be obtained by partial digestion of rpre RNA with a specific endonuclease, separation of the products on polyacrylamide gels or sucrose gradients and identification of fragments of the transcribed spacer region by using the previously characterised transcribed spacer products nos. 1-14 and 20-28 as markers. In this respect, RNase III which specifically cleaves at double-stranded regions would prove a useful tool for introducing specific nicks into a precursor molecule. A thorough search for end groups of the various rpre RNA intermediates and of rRNA would also be very useful.

Knowledge of the distance, in terms of primary structure between the ribosomal, transcribed spacer and 5.8S sequences

may suggest models for folding and processing of the precursor during maturation. If an extensive region of transcribed spacer separates the 28S and 5.8S sequences within the 32S molecule then one would envisage a considerable amount of secondary structure within that transcribed spacer region to provide specific alignment of the 5.8S sequence with that part of the 28S sequence with which it forms hydrogen bonds. As already mentioned, Perry & Kelley (1972) have isolated an exonuclease which will hydrolyse large segments from the 3'-OH end of 45S and 32S RNA. Fingerprints of 45S and 32S RNA after prior digestion by such an exonuclease would provide direct evidence of the transcribed spacer sequences at the 3' end of 45S and 32S RNA. Also, by using the previously characterised 32S transcribed spacer and 5.8S oligonucleotides as markers, it should be possible to determine how far along from the 3' end of the 32S molecule the 5.8S sequence is located.

## REFERENCES

REFERENCES

- Adesnik, M. & Levinthal, C. (1969) *J. Mol. Biol.* 46, 281-303.
- Amaldi, F. & Attardi, G. (1968) *J. Mol. Biol.* 33, 737-755.
- Attardi, G. & Amaldi, F. (1970) *Ann. Rev. Biochem.* 39,  
183-226.
- Attardi, G., Parnas, H., Huang, M-I.H. & Attardi, B.  
(1966) *J. Mol. Biol.* 20, 145-182.
- Beaud, G. & Hayes, D.H. (1971) *Eur. J. Biochem.* 19, 323-339.
- Birnstiel, M. & Grunstein, M. (1972) *FEBS Symp.* 23, 349-365.
- Birnstiel, M., Grunstein, M. & Speirs, J. (1969)  
*Nature* 223, 1265-1267.
- Bremer, H. & Berry, L. (1971) *Nature New Biol.* 234, 81-83.
- Brown, D.D., Wensink, P.C. & Jordan, E. (1972) *J. Mol. Biol.*  
63, 57-73.
- Brownlee, G.G. (1972) *Lab. Techniques in Biochem. and  
Mol. Biol. - Determination of Sequences in RNA  
(Work, T.S. & Work, E. Eds) North Holland Publishing  
Co., Netherlands.*
- Brownlee, G.G. & Cartwright, E. (1971) *Nature New Biol.*  
232, 50-52.
- Brownlee, G.G. & Sanger, F. (1967) *J. Mol. Biol.* 23, 337-353.
- Brownlee, G.G. & Sanger, F. (1969) *Eur. J. Biochem.* 11,  
395-399.
- Brownlee, G.G., Sanger, F. & Barrell, B.G. (1968)  
*J. Mol. Biol.* 34, 379-412.
- Burdon, R.H. (1971) in "Progress in Nucleic Acid Research  
and Molecular Biology" 11, 33-79, (Davidson, J.N.  
& Cohn, W.E. Eds) Academic Press, New York.

- Callan, H.G. (1967) *J. Cell Sci.* 2, 1-7.
- Corte, G., Schlessinger, D., Longo, D. & Venkov, P.  
(1971) *J. Mol. Biol.* 60, 325-338.
- Craig, N.C. & Perry, R.P. (1969) *J. Cell Biol.* 45, 554-564.
- Dahlberg, A.E. & Peacock, A.C. (1971) *J. Mol. Biol.* 55, 61-74.
- Dawid, I.B., Brown, D.D. & Reeder, R.H. (1970) *J. Mol. Biol.*  
51, 341-360.
- Eagle, H. (1959) *Science* 130, 432-437.
- Earle, W.R. (1943) *J. Natl. Cancer Inst.* 4, 165-212.
- Fellner, P. (1969) *Eur. J. Biochem.* 11, 12-27.
- Fellner, P., Ehresmann, C., Stiegler, P. & Ebel, J.P.  
(1972) *Nature New Biol.* 239, 1-5.
- Galibert, F., Larsen, C.J., Lelong, J.C. & Boiron, M.  
(1965) *Nature* 207, 1039-1041.
- Gey, G.O., Coffman, W.D. & Kubicek, M.T. (1952)  
*Cancer Res.* 12, 264-265.
- Girard, M., Penman, S. & Darnell, J.E. (1964) *Proc. Natl.  
Acad. Sci. U.S.A.* 51, 205-211.
- Greenberg, H. & Penman, S. (1966) *J. Mol. Biol.* 21, 527-535.
- Grierson, D., Rogers, M.E., Sartirana, M.L. & Loening, U.E.  
(1970) *Cold. Spr. Harb. Symp. Quant. Biol.* 35, 589-598.
- Hadjiolov, A.A. & Cox, R.A. (1973) *Biochem. J.* 135, 349-351.
- Hayes, F., Hayes, D., Fellner, P. & Ehresmann, C. (1972)  
*Nature New Biol.* 232, 54-55.
- Helser, T.L., Davies, J.E. & Dahlberg, J.E. (1971) *Nature  
New Biol.* 233, 12-14.
- Helser, T.L., Davies, J.E. & Dahlberg, J.E. (1972) *Nature  
New Biol.* 235, 6-9.

- Honjo, T. & Reeder, R.H. (1973) *J. Mol. Biol.* 80, 217-228.
- House, W. & Wildy, P. (1965) *Lab. Practice* 14, 594.
- Jeanteur, P., Amaldi, F. & Attardi, G. (1968) *J. Mol. Biol.* 33, 757-775.
- Jeanteur, P. & Attardi, G. (1969) *J. Mol. Biol.* 45, 305-324.
- Klootwijk, J., Van Den Bos, R.C. & Planta, R.J. (1972) *FEBS Letts.* 27, 102-106.
- Kumar, A. & Warner, J.R. (1972) *J. Mol. Biol.* 63, 233-246.
- Kwan, C.N., Gotoh, S. & Schlessinger, D. (1974) *Biochim. Biophys. Acta* 349, 428-441.
- Liau, M.C. & Perry, R.P. (1969) *J. Cell Biol.* 42, 272-283.
- Loening, U.E. (1968) *J. Mol. Biol.* 38, 355-365.
- Lowry, C.V. & Dahlberg, J.E. (1971) *Nature New Biol.* 232, 52-54.
- Maden, B.E.H. (1971) *Progress in Biophysics and Molecular Biology* 22, 127-177.
- Maden, B.E.H. & Forbes, J. (1972) *FEBS Letts.* 28, 289-292.
- Maden, B.E.H., Lees, C.D. & Salim, M. (1972b) *FEBS Letts.* 28, 293-296.
- Maden, B.E.H. & Robertson, J.S. (1974) *J. Mol. Biol.* 87, 227-235.
- Maden, B.E.H. & Salim, M. (1974) *J. Mol. Biol.* in press.
- Maden, B.E.H., Salim, M. & Summers, D.F. (1972a) *Nature New Biol.* 237, 5-9.
- Maden, B.E.H., Vaughan, M.H., Warner, J.R. & Darnell, J.E. (1969) *J. Mol. Biol.* 45, 265-275.
- McConkey, E.H. & Hopkins, J.W. (1969) *J. Mol. Biol.* 39, 545-550.

- McLennan, B.D. & Lane, B.G. (1968) *Can. J. Biochem.* 46, 81-92.
- Mirault, M.E. & Scherrer, K. (1972) *FEBS Letts.* 20, 233-238.
- Nikolaev, N., Silengo, L. & Schlessinger, D. (1973)  
*Proc. Natl. Acad. Sci. U.S.A.* 70, 3361-3365.
- Pederson, T. & Kumar, A. (1971) *J. Mol. Biol.* 61, 655-668.
- Pene, J.J., Knight, E. & Darnell, J.E. (1968) *J. Mol. Biol.*  
33, 609-623.
- Penman, S. (1966) *J. Mol. Biol.* 17, 117-130.
- Penman, S. (1969) *Fundamental Techniques in Virology* (Habel, K. & Salzman, N.P. Eds) Chapter 5, Academic Press, New York.
- Penman, S., Smith, I. & Holtzman, E. (1966) *Science* 154, 786-789.
- Perry, R.P., Cheng, T-Y., Freed, J.J., Greenberg, J.R., Kelley, D.E. & Tartof, K.D. (1970) *Proc. Natl. Acad. Sci. U.S.A.* 65, 609-616.
- Perry, R.P. & Kelley, D.E. (1972) *J. Mol. Biol.* 70, 265-279.
- Petermann, M.L. & Pavlovec, A. (1966) *Biochim. Biophys. Acta* 114, 264-276.
- Reeder, R.H. & Brown, D.D. (1970) *J. Mol. Biol.* 51, 361-377.
- Reichman, M., Karlan, D. & Penman, S. (1973) *Biochim. Biophys. Acta* 299, 173-175.
- Ritossa, F.M., Atwood, K.C. & Spiegelman, S. (1966) *Genetics* 54, 687-693.
- Robertson, H.D., Webster, R.D. & Zinder, N.D. (1968) *J. Biol. Chem.* 243, 82-91.
- Robertson, J.S. & Maden, B.E.H. (1973) *Biochim. Biophys. Acta* 331, 61-70.

- Rosset, R. & Monier, R. (1963) *Biochim. Biophys. Acta* 68, 653-656.
- Rubin, G. (1973) *J. Biol. Chem.* 248, 3860-3875.
- Salim, M. & Maden, B.E.H. (1973) *Nature* 244, 334-336.
- Sanford, K.K., Earle, W.R. & Likley, G.D. (1948) *J. Natl. Cancer Inst.* 2, 229-246.
- Sanger, F., Brownlee, G.G. & Barrell, B.G. (1965) *J. Mol. Biol.* 13, 373-398.
- Scherrer, K. & Darnell, J.E. (1962) *Biochem. Biophys. Res. Commun.* 7, 486-490.
- Scherrer, K., Latham, H. & Darnell, J.E. (1963) *Proc. Natl. Acad. Sci. U.S.A.* 49, 240-248.
- Shepherd, J. & Maden, B.E.H. (1972) *Nature* 236, 211-214.
- Slack, J.M.W. & Loening, U.E. (1974) *Eur. J. Biochem.* 43, 59-67.
- Snyder, A.L., Kann, H.E. & Kohn, K.W. (1971) *J. Mol. Biol.* 58, 555-565.
- Soeiro, R., Vaughan, M.H. & Darnell, J.E. (1968) *J. Cell Biol.* 36, 91-101.
- Sogin, M., Pace, B., Pace, N.R. & Woese, C.R. (1971) *Nature New Biol.* 232, 48-49.
- Spahr, P.F. (1964) *J. Biol. Chem.* 239, 3716-3726.
- Speirs, J. & Birnstiel, M. (1974) *J. Mol. Biol.* 87.
- Tartof, K.D. (1971) *Science* 171, 294-297.
- Udem, S.A. & Warner, J.R. (1972) *J. Mol. Biol.* 65, 227-242.
- Vaughan, M.H., Soeiro, R., Warner, J.R. & Darnell, J.E. (1967) *Proc. Natl. Acad. Sci. U.S.A.* 58, 1527-1534.
- Wagner, E.K., Penman, S. & Ingram, V.M. (1967) *J. Mol. Biol.* 29, 371-387.

- Warner, J.R. & Soeiro, R. (1967) Proc. Natl. Acad. Sci. U.S.A. 58, 1984-1990.
- Warner, J.R., Soeiro, R., Birnboim, H.C., Girard, M. & Darnell, J.E. (1966) J. Mol. Biol. 19, 349-361.
- Warner, J.R. & Udem, S.A. (1972) J. Mol. Biol. 65, 243-258.
- Weinberg, R.A., Loening, U.E., Willems, M. & Penman, S. (1967) Proc. Natl. Acad. Sci. U.S.A. 58, 1088-1095.
- Weinberg, R.A. & Penman, S. (1968) J. Mol. Biol. 38, 289-304.
- Weinberg, R.A. & Penman, S. (1970) J. Mol. Biol. 47, 169-178.
- Wellauer, P.K. & Dawid, I.B. (1973) Proc. Natl. Acad. Sci. U.S.A. 70, 2827-2831.
- Willems, M., Wagner, E., Laing, R. & Penman, S. (1968) J. Mol. Biol. 32, 211-220.
- Woledge, J., Corry, M.J. & Payne, P.I. (1974) Biochim. Biophys. Acta 349, 339-350.
- Wyatt, G.R. (1951) Biochem. J. 48, 584-590.
- Zimmerman, E.F. & Holler, R.W. (1967) J. Mol. Biol. 23, 149-161.

**APPENDIX**

APPENDIX

The Theoretical Distribution of Nucleotides Among the Products  
of  $T_1$  Digestion.

If  $x\%$  of an RNA molecule consists of G and the distribution of G within the sequence is essentially random, then the probability of G occurring next to another G is  $x\%$ . Hence on complete  $T_1$  digestion  $x\%$  of the products should be free G and  $(1-x)\%$  as larger products (...XG). Of these larger products,  $x\%$  should be di-nucleotides (GXG) and  $(1-x)\%$  again as larger products (...XXG). Extending this argument the fraction of the products of  $n$  or more nucleotides long will be  $(1-x)^{n-1}\%$  and the fraction of the products of  $n$  nucleotides long will be  $x(1-x)^{n-1}\%$ .

For an RNA molecule  $y$  nucleotides long, the total number of products will be  $xy$ ; thus the number of products of  $n$  nucleotides long will be  $xy \cdot x(1-x)^{n-1}$  or  $x^2y(1-x)^{n-1}$ . Knowing the total number of nucleotides appearing in mono-, di-, trinucleotides etc., the number of nucleotides appearing in products of  $n$  or more nucleotides can be determined.

For the 32S transcribed spacer region (36% G) the fraction of the products of  $n$  or more nucleotides long will be 10.7%. If this region is assumed to comprise 1350 nucleotides then 1050 of them will appear in mono- up to pentanucleotides; thus 30% of the molecule will appear as hexanucleotides or larger products.

