CANADIAN THESES ON MICROFICHE

ISBN

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliotheque nationale du Canada
Direction du developpement des collections

Service des theses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

RIBOSOMAL RNA STRUCTURE AND EVOLUTION REVEALED BY NUCLEOTIDE SEQUENCE ANALYSIS

by

© Murray N. Schnare

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Dalhousie University

Department of Biochemistry
Dalhousie University
Halifax, Nova Scotia

June, 1984

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Ribosomal RNA sequence comparisons can be used to determine phylogenetic relationships among organisms and should also be useful in evaluating hypotheses concerning the evolutionary origin of organelles. However, no compelling support for either an endosymbiotic or autogenous origin of mitochondria has come from analysis of animal, fungal, and protozoan mitochondrial rRNA sequences. I have therefore utilized newly-developed, rapid RNA sequencing procedures to determine primary structure at the ends of the mitochondrial and cytosol 18S and 26S rRNAs of wheat (a higher plant). These results, along with the results of experiments examining heterologous hybridization between wheat mitochondrial DNA and Escherichia coli 16S and 23S rRNAs, provide strong evidence in support of an endosymbiotic, specifically eubacterial, evolutionary origin of wheat mitochondria.

There is great diversity in both size and base composition of homologous high molecular weight rRNA species from various sources, while certain small rRNAs that are present in some organisms are absent in others. It is hypothesized that much of this diversity may be attributable to differences in the distribution of coding and spacer sequences in rRNA genes, as well as differences in post-transcriptional processing. I have determined the primary sequences of the 5.8S rRNA and four novel small rRNAs from Crithidia fasciculata with the goal of precisely defining spacer sequences in Crithidia rDNA. Knowledge of these rRNA sequences and spacer sequences should be valuable in testing the foregoing hypothesis.

Sequence comparisons are also valuable in defining evolutionarily conserved regions of primary and secondary structure, which presumably represent conserved functional domains of the rRNA molecule. I have determined the 3'-terminal sequence of Crithidia 18S rRNA and, in the light of these new data, I have re-examined certain base-paired interactions that have previously been proposed to be involved in ribosomal subunit association and initiation of translation in eukaryotes.

LIST OF ABBREVIATIONS

A	adenosine
BB	bromophenol blue
bis-acrylamide	N,N'-bis-acrylamide
BSA	bovine serum albumin
C	cytidine
DBAE	N-[N'-(m-dihydroxyborylphenyl)-succinamyl]-aminoethyl
DEAE	diethyl aminoethyl
DMS	dimethyl sulfate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E_1 buffer	30 mM NaH_2PO_4, 36 mM Tris, 1 mM EDTA, pH 7.8
E_2 buffer	15 mM Na_2HPO_4, 15 mM NaH_2PO_4, 3 mM EDTA, pH 6.8
EDTA	ethylenediaminetetracetic acid
EtOH	ethanol
G	guanosine
iRNA	that portion of total cellular RNA which is soluble in 1 M NaCl at 0°C
LSU RNA	the high molecular weight RNA component of the large ribosomal subunit
$Mg(OAc)_2$	magnesium acetate
mRNA	messenger RNA

N	ribonucleoside. The letter p to the right of N indicates 2',(3')-monophosphorylation, while the letter p to the left of N indicates 5'-monophosphorylation. The letter m to the right of N indicates $\underline{O}^{2'}$-methylation. The letter m to the left of N indicates base methylation; a subscript indicates the number of methyl groups, while a superscript indicates the position of the methyl groups.
NaOAc	sodium acetate
NH_4OAc	ammonium acetate
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
S	Svedberg sedimentation unit ($1S = 1 \times 10^{-13}$ sec)
sRNA	that portion of total cellular RNA which is soluble in 1 M NaCl at 0°C
SSC	0.15 M NaCl, 0.015 M sodium citrate
SSU RNA	small subunit ribosomal RNA
TBE	50 mM Tris, 50 mM boric acid, 1 mM EDTA, pH 8.3
TEMED	(N,N,N',N')-tetramethylethylenediamine

Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	uridine
Ψ	pseudouridine
U.V.	ultraviolet
XC	xylene cyanol

ACKNOWLEDGEMENTS

I would like to thank my research supervisor, Dr. Mike Gray, for taking me into his lab at a time when I didn't even know how a pipetman worked. He gave me supervision when I needed it and he allowed me to work independently when I became capable of doing so. I am also very fortunate to be working in the same lab as Dr. Dave Spencer, who has been like a second supervisor. Many of the procedures used, and some of the data presented here, were stolen directly from Dave's desk after he had gone home in the evening.

I would also like to acknowledge the expert technical advice and assistance of Linda Bonen, Ron MacKay, Ford Doolittle, Carl Thomas, T.Y. Huh, Jason Hofman, Susan Forsythe, Bill Walker, David Salgado, Carmen Sapienza, James Collings, Shelley Gray and Rita Breckon. I also thank Reg Lau for pouring ethanol down my back at a time when I really needed it. I am also grateful to Kathy Whalen for her excellent typing of this thesis and for constantly reassuring me that we would make the deadline.

Over the past six years I have been influenced by almost everyone in the Department of Biochemistry and in this final rush I have probably forgotten to acknowledge many important contributions to my research. Hopefully, I will remain here long enough to put back into this lab and this department some of the effort that they have probably wasted on me.

INTRODUCTION

1. The Ribosome: A Macromolecular Complex of Central Biological Importance

Messenger RNA-directed protein synthesis is a very important and extremely complicated process that occurs in all living cells. Since the translation of mRNA (protein biosynthesis) occurs on ribosomes, we cannot fully understand the mechanism of protein synthesis until the structure of the ribosome has been elucidated completely.

Ribosomes from all sources have a small subunit and a large subunit, both of which contain RNA and protein components (51). We have undertaken a study of the RNA components of various ribosomes with a long term goal of understanding the biosynthesis of ribosomes and their constituent rRNAs. Since ribosomes perform essentially the same function in all living things, it follows that rRNA genes must have been established early in evolution and that functionally important regions of rRNA have been evolutionarily conserved. Therefore, comparisons of rRNA genes from different sources should allow us to define these conserved regions, which presumably represent important functional domains of the molecule. Determination of the primary sequence of a rRNA or a rRNA gene is a useful first step in detecting possible base-paired interactions that may be involved in ribosome assembly, mRNA binding, tRNA binding and subunit association.

Primary sequence analysis of a rRNA and of the gene encoding it is valuable in delineating the ends of the gene and spacer

sequences, and should also provide clues about signals for initiation and termination of transcription and for processing of the primary transcript. Comparisons of rRNA sequences are also useful in establishing secondary structure models and in determining evolutionary relationships among organisms (2-50).

2. Structure of Ribosomes and their RNA Components

2.1. Diversity in Ribosome and Ribosomal RNA Structure

Although ribosomes from all sources perform a similar function in protein synthesis, they differ radically in size, RNA to protein ratio and in the number and size of rRNA components (51, Table 1; see also INTRODUCTION, section 3). It can be seen from Table 1 that SSU RNA ranges in size from about 950 nucleotides in animal mitochondria to 1955 nucleotides in wheat mitochondria while LSU RNA ranges in size from about 1560 nucleotides in animal mitochondria to 4718 nucleotides in rat cytoplasm. The typical value for G+C content of rRNA genes is about 55%. This value drops to about 35-45% in animal and ciliate mitochondria and to about 21% in yeast mitochondria.

This great diversity in both size and base composition of homologous rRNA species from different organisms has been useful in defining a universal structural core (323) that is probably responsible for many of the conserved processes in which SSU RNA is involved. However, this same diversity in rRNA structure complicates sequence comparisons that are used to deduce phylogenetic relationships (see DISCUSSION, section 1).

Table 1. High Molecular Weight Ribosomal RNAs of Known Sequence
(a) Sedimentation coefficient, (b) Length (nucleotides), (c) G+C content(%)

	30S subunit			50S subunit		
Prokaryotes	a	b	c	a	b	c
(70S ribosome)						
E. coli (39,41,49)	16S	1542	54	23S	2904	54
A nidulans (192,337)	16S	1487	55	23S	2869	58
P. vulgaris (48)	16S	1544	52			
H volcanii (139)	16S	1472	57			
	40S subunit			**60S subunit**		
Eukaryotic Cytoplasm	a	b	c	a	b	c
(80S ribosome)						
X. laevis (286,361)	18S	1825	54	28S	4110	66
S. cerevisiae (283,111)	18S	1789	45	25S	3392	48
Rat (53,54,329)	18S	1869	56	28S	4718	67
P polycephalum (256),				26S	3788	53
S. carlsbergensis (354)				26S	3393	48
D. discoideum (230)	17S	1875	42			
Mouse (273)	18S	1869	56			
	30S Subunit			**50S Subunit**		
Chloroplasts	a	b	c	a	b	c
(70S ribosome)						
Z. mays (97,298)	16S	1491	56	23S	2903	54
N. tabacum (330,335)	16S	1486	54	23S	2804	55
E. gracilis (120)	16S	1491	47			
C. reinhardii (91)	16S	1475	54			
Mitochondria						
animal	**25-35S Subunit**			**40-45S Subunit**		
(55-60S ribosome)	a	b	c	a	b	c
Mouse (352)	12S	956	35	16S	1582	42
Rat (177,284)	12S	952	39	16S	1559	38
Human (98)	12S	954	45	16S	1559	35
Bovine (7)	12S	955	40	16S	1571	39
fungal	**30-40S Subunit**			**50S Subunit**		
(70-75S ribosome)	a	b	c	a	b	c
Yeast (204,310,311)	15S	1686	22	21S	3273	21
Aspergillus (179,180)	15S	1437	35	21S	2865	29
ciliate	**55S Subunit**			**55S Subunit**		
(80S ribosome)	a	b	c	a	b	c
Paramecium (300)				20S	2204	37
plant	**40S Subunit**			**60S subunit**		
(80S ribosome)	a	b	c	a	b	c
T. aestivum (314)	18S	1955	53			

Note Sedimentation coefficients for ribosomes and ribosomal subunits are from (35,66,187)

2.2. Secondary Structure of LSU and SSU RNAs

Secondary structure models have been proposed for E. coli 16S (224,324,379,380,388) and 23S (37,114,249) rRNAs. Basically, three techniques have been utilized in determination of the secondary structure of rRNA molecules (38). The first involves analysis of single-stranded regions that are accessible to chemical and enzymatic attack, while a second involves analysis of the small base-paired fragments that are produced by partial nuclease digestion of intact rRNAs. Secondary structures developed on the basis of these two methods can be refined by the third method, comparative sequence analysis. This approach assumes that a functionally important element of secondary structure is likely to be conserved between two homologous rRNAs, even where extensive primary sequence divergence has occurred.

The SSU RNA sequences from mammalian (225,323,324,388) and fungal (323,324) mitochondria, the sequences of eukaryotic nuclear-encoded 18S rRNAs (53,252,323,324,388), chloroplast 16S rRNA sequences (323,324,388) and an archaebacterial 16S rRNA sequence (139) can all be drawn to fit secondary structure models that have many features in common with those proposed for E. coli 16S rRNA. Comparison of these models reveals a smaller secondary structure core that is common to all SSU RNAs (323). Similarly, the LSU RNA sequences from chloroplasts (37,114,249), yeast cytoplasm (354) and mitochondria (180,311,323,324,388) can all be accommodated by secondary structure models that closely resemble those proposed for E. coli 23S rRNA.

This conservation of a basic secondary structural core in homologous rRNA species from different sources, which is quite remarkable in view of the great diversity in size and base composition of these molecules (Table 1), implies that secondary structural elements as well as primary sequence are functionally important in protein biosynthesis.

2.3. Modified Nucleoside Components of Ribosomal RNA

In addition to the four major nucleosides present in rRNA populations, minor amounts of modified nucleosides are also present. In rRNA these minor components are usually ribose ($\underline{O}^{2'}$) or base methylated derivatives of the four major nucleosides. Pseudouridine (Ψ), a modified nucleoside in which uracil is linked through its C^5 rather than its N^1 position to the ribose moiety (144), is also found in rRNA, as are N^4-acetylcytidine and 1-methyl-3-(3-amino-3-carboxypropyl)Ψ. Examination of T_1 oligonucleotides that contain modified nucleosides has revealed that rRNAs are homogenous with respect to methylated sequences and that methylation is a highly specific process (103,144). Although the function of modified nucleosides in rRNA is uncertain, there are a few reports which suggest that rRNA methylation may be required for maturation of eukaryotic rRNAs (43,381).

Studies of the modified nucleoside composition of hamster (93,301), human (285,301), mouse (147,301), chicken (301), Xenopus (219,301), Drosophila (220), yeast (176), wheat (196,307) and Crithidia (122) cytosol rRNAs have revealed that most of the modified nucleosides in eukaryotic rRNA are ribose methylated and that many of

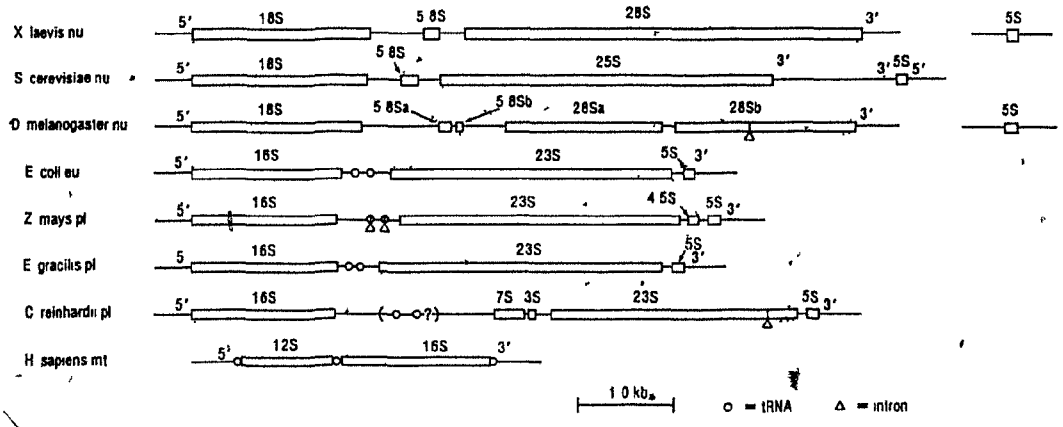


Figure 1. Ribosomal RNA gene organization. eu, eubacterial; nu, nuclear; pl, plastid; mt, mitochondrial. See the text for references.

the methylated sequences are evolutionarily conserved within the eukaryotic kingdom. Analysis of the methylated sequences in E. coli 16S and 23S rRNAs (103) has indicated that most of the modified components are methylated bases and that there is a very low level of $O^{2'}$-methylnucleosides relative to the level observed in eukaryotic rRNAs. The methylation pattern observed in E. coli 16S rRNA is highly conserved within the eubacterial kingdom (378) and is distinct from that found in the archaebacteria (17). The base methylations observed in E. coli 23S rRNA (103) are missing in the 23S rRNAs from Mycoplasma capricolum and Acholeplasma laidlawii (153). However, these two organisms display typically eubacterial 16S rRNA methylation patterns (153).

Hamster mitochondrial LSU RNA contains four ribose methylated nucleosides (16,92), while the SSU RNA contains five base methylated nucleosides (16,93). Yeast (175) and Neurospora (192) mitochondrial LSU RNAs each contain two ribose methylated nucleosides, while the mitochondrial SSU RNAs from these two fungi are completely unmethylated.

3. Evolutionary Origin of Low Molecular Weight Ribosomal RNAs

3.1. Ribosomal RNA Gene Organization (Figures 1 and 2)

3.1.(a) Eubacterial rRNA Genes

In the eubacteria E. coli (40,212,238), Bacillus subtilis (322,387), Bacillis thuringiensis (174) and Anacystis nidulans (336,373), the rRNA gene order is 16S-23S-5S with spacers of variable

lengths separating the three rRNA genes. The rRNA genes in E. coli are co-transcribed in the order 5'-16S-23S-5S-3' from a single promoter at the 16S end (87) and there is evidence for a similar co-transcription of rRNA genes in Bacillus subtilis (387).

There are seven rRNA transcriptional units in E. coli (238), one of which (rrnD) contains two 5S rRNA genes (94). Four of these units have $tRNA^{Glu}$ while three have $tRNA^{Ala}$ and $tRNA^{Ile}$ genes in the transcribed spacer separating the 16S and 23S rRNA genes (238). The complete sequence of the rrnB operon from E. coli has been determined recently (40). There are two rRNA gene clusters in Anacystis nidulans (330,373) and sequence analysis of the 16S-23S spacer of one of these has revealed the presence of both $tRNA^{Ile}$ and $tRNA^{Ala}$ genes (373). Two of the ten rRNA operons in Bacillus subtilis contain $tRNA^{Ala}$ and $tRNA^{Ile}$ genes between the 16S and 23S rRNA genes (211). No tRNA genes are found in the 16S-23S rRNA spacer in rRNA gene clusters from Bacillus thuringiensis (174).

3.1.(b) Archaebacterial rRNA Genes

There is a single copy of the rRNA genes arranged in the order 5'-16S-23S-5S-3' in Halobacterium halobium (152,244), Sulfolobus acidocaldarius (244), Thermoproteus tenax (244), Thermococcus celer (244), Thermophilum pendens, Desulfurococcus mobilis (244) and Desulfurococcus mucosus (244). Methanobacterium thermoautotrophicum (244) and Methanococcus vanielii (160) contain two and four sets of rRNA genes, respectively, arranged in the order 5'-16S-23S-5S-3'. Sulfolobus acidocaldarius (244), Thermococcus celer (244) and

Methanococcus vanielii (160) each contain an extra 5S rRNA gene. The primary sequence has been determined for the spacer separating the 16S and 23S rRNA genes in two of the four rRNA operons in Methanococcus vanielii. One of these spacers contains a gene for $tRNA^{Ala}$, while no tRNA structure could be found in the other spacer (161). There is also a single copy of each rRNA gene in Thermoplasma acidophilum (341). However, these genes appear to be physically and transcriptionally unlinked, with a distance between the 16S and 23S rRNA genes of at least 7.5 kb. The 23S and 5S rRNA genes are at least 6 kb apart and the 16S and 5S rRNA genes are separated by at least 1.5 kb (341).

3.1.(c) Chloroplast rRNA Genes

(i) Euglena gracilis

Genes coding for 16S, 23S and 5S rRNAs are found on each of three tandemly repeated 5.6 kb segments of Euglena chloroplast DNA (134,135,253). The gene order is 5'-16S-23S-5S-3' and the spacer separating the 16S and 23S rRNA genes contains a $tRNA^{Ile}$ and a $tRNA^{Ala}$ gene (119,253,254).

(ii) Chlamydomonas reinhardii

In addition to 16S, 23S and 5S rRNA, Chlamydomonas chloroplast DNA also codes for a 3S and a 7S rRNA (277,278). The genes for these rRNAs are physically (277,278) and probably transcriptionally (277,326) linked in the order 5'-16S-7S-3S-23S-5S-3'. These genes are separated by spacer DNA (277,278) and there is at least one tRNA gene in the 16S-7S rRNA spacer region (223). There are two copies of this

rRNA gene cluster, found in an inverted repeat arrangement in Chlamydomonas chloroplast DNA (278), and both of the 23S rRNA genes contain an intervening sequence near the 3'-terminus (2,278).

(iii) Higher Plants

Two copies of the rRNA genes are located within inverted repeats on the chloroplast DNA of most of the plants that have been examined to date (22,104,166,171,258,259,316,367). However, broad bean (181) and pea (260) chloroplasts contain only one set of rRNA genes. Plant chloroplasts do not contain 7S and 3S rRNAs but they do have a 4.5S rRNA component (33,368) that maps between the 23S and 5S rRNA genes (328). Plant chloroplast rRNA genes are arranged in the order 5'-16S-23S-4.5S-5S-3' (22,96,104,166,328,367) and an 8.2 kb transcript that represents a common precursor of the 16S, 23S and 4.5S rRNAs has been isolated from spinach (145,146) and tobacco (331) chloroplasts.

The nucleotide sequence of a complete rRNA gene cluster has been determined in the case of Zea mays (96,97,178,298,299) and Nicotiana tabacum (329-331,334,335) chloroplast DNAs. Both DNAs contain a $tRNA^{Ile}$ and a $tRNA^{Ala}$ gene in the spacer region separating the 16S and 23S rRNA genes and these spacer-encoded tRNA genes are interrupted by long intervening sequences (178,331).

3.1.(d) Eukaryotic Nuclear rRNA Genes

Typical eukaryotic ribosomes contain four RNA components, designated 25-28S, 17-19S, 5.8S and 5S rRNA (64). Investigation of the high molecular weight rRNAs of various eukaryotes by gel electrophoresis has revealed that these rRNAs have not maintained a

constant size throughout evolutionary history. However, there is no obvious relationship between rRNA size and phylogenetic position (46,208). Eukaryotic rRNA genes are arranged in the order 5'-18S-5.8S-28S-3' in all eukaryotes that have been examined (23,47,63,76,138,209,226,268,274,356), and these genes are co-transcribed from a single promoter to yield a 37S-45S precursor molecule that is processed to produce the mature rRNAs (23,76,102,209,274,267). The 5S rRNA gene is 1.3 kb from the 3'-end of the 25S rRNA gene in yeast (Saccharomyces cerevisiae) (23) and within 7 kb of the 3'-end of the 25S rRNA gene in Dictyostelium discoideum (222). The 5S rRNA gene of Euglena gracilis also maps near the 3'-end of the 25S rRNA gene (73). The Saccharomyces (189,347), Euglena (73) and Dictyostelium (20,121) 5S rRNA genes are not transcriptionally linked to the other rRNA genes, and the 5S rRNA gene of yeast is known to be in the opposite orientation relative to the other rRNA genes (189,347). In all other eukaryotes examined the 5S rRNA genes are not physically or transcriptionally linked to the other rRNA genes (102,209,267). Primary sequence analysis within the rRNA transcriptional units of Saccharomyces cerevisiae (308,309), Saccharomyces carlsbergensis (353,359), Xenopus laevis (143), Xenopus borealis (108), rat (325), mouse (117,234), sea urchin (151) and Physarum polycephalum (256) indicate that there are no tRNA genes in the internal transcribed spacer separating the 18S and 5.8S rRNA genes (ITS1), or in the internal transcribed spacer separating the 5.8S and 26S rRNA genes (ITS2),

The 25S-28S rRNA of many eukaryotes contains a median or nearly median break that is masked by hydrogen bonds between the two halves of the molecule (46). R-Loop mapping of the rDNA reveals a gap (≃100-600 nucleotides long) separating the regions encoding the two halves of the 28S rDNA in Drosophila melanogaster (266,363), Sciara coprophila (276), Acanthamoeba castellanii (74) and Leishmania donovani (203). Similarily, in the dipteran insects, Drosophila melanogaster (262) and Sciara coprophila (165), the 5.8S rRNA homologue is present as two non-covalently associated smaller ("5S" and "2S") rRNAs, and the genes encoding these small rRNAs are separated by short transcribed spacers.

Intervening sequences have been detected in many but not all of the 28S rRNA genes in Drosophila melanogaster (115,363,366), Drosophila virilis (19), other closely related species of Drosophila (209), and in another insect, Calliphora erythrocephala (21). At least in the case of D. melanogaster, the 28S rRNA genes that contain intervening sequences do not appear to be transcribed at a significant level and these genes probably do not give rise to mature, functional 28S rRNA molecules (210).

Two intervening sequences have been found in most if not all of the 26S rRNA genes in Physarum polycephalum (47,138), and there is a single intervening sequence in all of the 26S rRNA genes in some strains of several species of Tetrahymena (82,370). There appear to be two introns in the 25S rRNA genes and one intron in the 17S rRNA genes of Plasmodium lophurae (343).

The excision of the intron sequence and subsequent ligation of the two exon sequences of Tetrahymena thermophila 26S rRNA occurs at the RNA transcript level in the absence of enzymes or other proteins and does not require ATP or GTP hydrolysis (191). This self-splicing ability appears to be an intrinsic property of the intervening sequence and has not yet been observed in other systems.

3.1.(e) Mitochondrial rRNA Genes (Figure 2)

(i) Animals

In human (98), bovine (7) and mouse (352) mitochondrial DNA, the rRNA genes and immediately adjacent tRNA genes are located on the same strand, in the order 5'-$tRNA^{Phe}$-small(12S)rRNA-$tRNA^{Val}$-large(16S) rRNA-$tRNA^{Leu}$-3'. Physical mapping of rat (190), Xenopus (272) and avian (113) mitochondrial DNA also suggests a close linkage of the two rRNA genes. In Drosophila yakuba (61) the rRNA gene order is also 5'-12S-$tRNA^{Val}$-16S-3'; however, in this case $tRNA^{Phe}$ and $tRNA^{Leu}$ genes are not found flanking the rRNA genes. No intervening sequences have been found in animal mitochondrial rRNA (or any other) genes.

(ii) Fungi

In strains of the yeast Saccharomyces, the mitochondrial genes for the LSU and SSU RNAs are far apart, separated by at least 25,000 bp (31,288,289), and these two genes are transcribed from independent promoters (255). The mitochondrial LSU and SSU RNA genes of most other yeasts (58-60,137,364), Neurospora crassa (80,136,150), and Aspergillus nidulans (198) are also separated by large distances on the mitochondrial genome. The mitochondrial DNA of the water mold

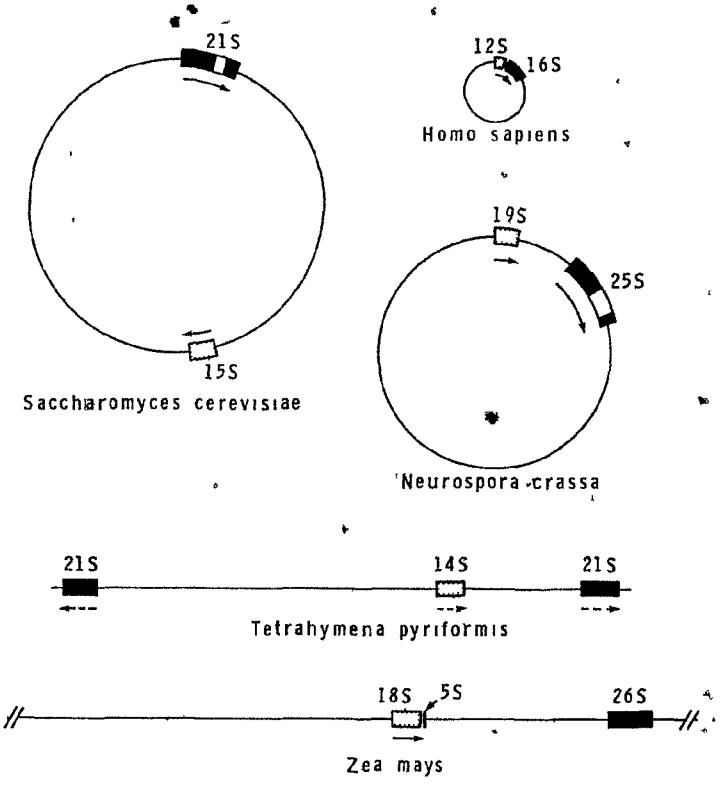


Figure 2. Arrangement of the rRNA genes in diverse mitochondrial genomes. The genomes and their contained rRNA genes are drawn in proportion, except that only a small section of the maize mitochondrial genome is shown. SSU RNA genes (stippled boxes), LSU RNA genes (solid boxes), and introns in the S. cerevisiae 21S and N. crassa 25S rRNA genes (open boxes) are indicated. The continuous arrows show the direction of transcription of the rRNA genes (where known); the dashed arrows indicate the relative orientations of the rRNA genes in Tetrahymena mitochondrial DNA. See the text for references. Reprinted with permission, from M.W. Gray (125).

Achlya has an inverted repeat containing the LSU and SSU rRNA genes (155), an organization reminiscent of the situation in higher plant and Chlamydomonas chloroplasts (see INTRODUCTION, section 3.1.c).

A single intervening sequence has been detected near the 3'-end of the LSU RNA gene in Saccharomyces (32,101), Allomyces macrogynus (30), Neurospora (80,136,142,150) and Aspergillus (198) mitochondria. The LSU RNA gene from Podospora anserina mitochondria contains two intervening sequences (383).

(iii) Ciliates

The mitochondrial genomes of Paramecium aurelia (69) and Tetrahymena pyriformis (116) exist as linear DNA molecules (14 μm and 15 μm, respectively), in contrast to the circular genomes that are characteristic of animal (7) and fungal (31) mitochondria. The Paramecium mitochondrial DNA contains a single copy of the LSU and SSU RNA genes, with the large (21S) rRNA gene near one end of the genome and the small (14S) rRNA gene near the center (69). Tetrahymena mitochondrial DNA also has a 14S rRNA gene near the center of the molecule; however, this genome contains two copies of the 21S rRNA gene that are found in an inverted arrangement, at opposite ends of the linear mitochondrial DNA molecule (116).

(iv) Plants

Plant mitochondrial DNA not only contains a 26S and an 18S rRNA gene, but a 5S rRNA gene as well, which has not been found in animal or fungal mitochondria (200). The 18S and 26S rRNA genes are widely separated on the plant mitochondrial genome (28,156) (15-16 kb in

maize (157,320)), while the 18S and 5S rRNA genes are closely linked (28,156,320). In wheat and maize the mitochondrial 18S and 5S rRNA genes are on the same strand (55,128), with the 5S rRNA gene near the 3'-end of the 18S rRNA gene (55,128,295 and this thesis).

3.2. Structural Homology between Low Molecular Weight rRNAs and the Ends of E. coli 23S rRNA

From the preceeding section it is obvious that there are small rRNAs present in ribosomes from some sources (e.g., 5.8S rRNA in eukaryotes, 4.5S rRNA in higher plant chloroplasts, 3S and 7S rRNA in Chlamydomonas chloroplasts) that are absent in ribosomes from other sources. It has been only recently, with nucleotide sequence analysis of these small rRNAs and of E. coli 23S rRNA, that the structural relationship between these small rRNAs and eubacterial 23S rRNA has become apparent.

The gene coding for the 5.8S rRNA (160 nucleotides) of eukaryotes is separated by a transcribed spacer (ITS2) from the 5'-end of the 25-28S RNA gene and extensive identity has been observed between 5.8S rRNA and the 5'-end of eubacterial 23S rRNA. Interestingly, sequence identity has also been found between the 5'-terminus of eukaryotic 25-28S rRNA and the region beginning about 160 nucleotides in from the 5'-end of eubacterial 23S rRNA. Sequence comparisons of this sort have prompted Nazar (240,241), Jacq (159), Clark and Gerbi (57) and Walker (357) to propose that the 5'-end of 23S rRNA is the structural analogue in prokaryotes of 5.8S rRNA. The 7S and 3S rRNAs of Chlamydomonas chloroplasts also show similarity to sequences at the

5'-termini of eubacterial and other chloroplast 23S rRNAs (277).

Even within the eukaryotes there are examples where excision of spacer sequences results in the production of small rRNAs that are homologous to sections of larger rRNAs. For example, the split 5.8S rRNA in D. melanogaster appears after excision of a short, A-U rich, transcribed spacer, and the split 28S rRNA of some eukaryotes could also possibly be the result of spacer excision (see INTRODUCTION, section 3.1.d).

Comparisons of plant chloroplast and eubacterial rRNA gene sequences have implied that the counterpart of the 3'-end of E. coli 23S rRNA is found as a discrete 4.5S rRNA in chloroplasts (96,215) and is separated from the rest of the 23S rRNA gene by a transcribed spacer (97,330).

The 3'-terminus of animal (7,98,284,352), fungal (180,311), and protozoan (300) mitochondrial LSU RNA genes does not seem to have any obvious sequence similarity to the 3'-end of bacterial 23S rRNA or to chloroplast 4.5S rRNA. We were interested in determining whether the wheat mitochondrial 26S rRNA follows other mitochondrial LSU RNAs in this respect, or whether it contains a 4.5S rRNA homologue at its 3'-terminus. If it does, knowledge of the sequence of this region of the mitochondrial 26S rRNA could shed light on the evolutionary origin of the separate 4.5S rRNA in chloroplasts.

With these objectives in mind, the sequences at the 3'-terminus of wheat mitochondrial and wheat cytosol 26S rRNAs were determined and compared to published sequences of their chloroplast and eubacterial homologues.

3.3 Transcribed Spacers and Variable Regions in rRNA Genes

It has been proposed (57,65) that plant chloroplast 4.5S rRNA and eukaryotic 5.8S rRNA were originally part of a 23S rRNA ancestral gene, and that the transcribed spacers became inserted into the 23S rRNA gene during evolution. Subsequently, a mechanism developed which allowed these sequences to be processed out at the level of the primary transcript, to yield a separate 5.8S rRNA or 4.5S rRNA. It has also been suggested that the spacer separating the 23S and 4.5S rRNA genes in plant chloroplasts is related to bacterial insertion sequences (97) or that the spacer may be homologous to an intron and that the two RNA molecules (23S and 4.5S) remain unligated due to incomplete splicing (330).

Heterologous hybridization experiments between the rDNAs of Xenopus and yeast (118) and between the rDNAs of Xenopus and Neurospora (67) have revealed a pattern in which regions of highly conserved primary structure in eukaryotic rDNA are separated by regions that are highly variable in length and primary structure. These variable regions largely account for the diversity in length and base composition of mature eukaryotic LSU and SSU RNAs (see Table 1). It is also possible that some of the differences in length observed between rRNAs of various organisms are a result of small duplications within the rRNA genes (256).

Comparisons of internal transcribed spacer sequences and mature rRNA sequences in vertebrates (108,117,143,234,325), yeast (308,309,353,355), Sciara coprophila (165) and Drosophila (164,262)

have revealed that while the mature rRNA sequences are highly conserved, there has been rapid sequence divergence within the spacers and that the base composition of the variable regions in rRNA genes is similar to that of the transcribed and excised spacer regions. These observations led Cox and Kelly (65,66) to propose that spacer sequences, introns, and variable regions of rRNA genes may all be evolutionarily related and that they all represent insertions into an ancestral gene.

An alternative interpretation is that the ancient pattern of rRNA gene structure was one of "genes in pieces", as has been suggested in the case of protein coding genes (86). This theory implies that small rRNAs became part of larger rRNA molecules in the course of evolution. It should be noted that acceptance of this theory does not exclude the possibility that some spacers and/or introns are the result of insertions. In either case, it follows that some spacers could be degenerate introns and that some variable regions could be degenerate spacers or introns that are no longer excised.

3.4. The Ribosomal RNAs of Crithidia fasciculata

We have chosen C. fasciculata as a model system for examining the evolution of rDNA spacer sequences and the origin of small rRNAs. The cytoplasmic ribosome of this organism, a trypanosomatid protozoan, is unusual in that the large subunit contains six small rRNAs (designated e, f, g, h, i and j) (122,123), whereas typical eukaryotic ribosomes contain only two low molecular weight rRNA species (5S and 5.8S) (64). Data from other laboratories suggest a very similar or

identical spectrum of rRNA components in Leishmania tarentole (305), Trypanosoma brucei (62) and Euglena gracilis (79). The ribosome of C. fasciculata also contains a split LSU RNA (122,123).

Primary sequence analysis has shown that Crithidia species h is homologous to the 5S rRNA of other eukaryotes (217). As a first step towards gaining an understanding of the biosynthetic and evolutionary origin of Crithidia species e, f, g, i and j, the complete primary structures of these rRNAs have been determined and are presented here.

4. Phylogenetic Relationships and Small Subunit Ribosomal RNAs

4.1. Phylogenetic Classifications

In classifications of organisms into primary groups, the most obvious division is between the Prokaryota, or prokaryotes (bacteria, including cyanobacteria), and Eukaryota, or eukaryotes (all other organisms, unicellular and multicellular) (56,88). The most important distinction between these two groups is that prokaryotic cells lack the nuclear membranes, plastids, and mitochondria that are present in eukaryotic cells (83).

More recently, the eukaryotes have been subdivided into four major kingdoms (369), based primarily on morphological organization and on means of nutrition (photosynthesis, absorption, or ingestion). These kingdoms are Protista (unicellular eukaryotic organisms), Plantae (multicellular green plants and higher algae), Fungi (multinucleate higher fungi), and Animalia (multicellular animals).

Until 1977, theories concerning the evolution of organisms

started with the assumption that prokaryotes evolved first and that eukaryotes evolved directly from a branch of the prokaryotic phylogenetic tree (83,227,228,369). In 1977, a revolutionary new phylogeny, based on T_1 oligonucleotide catalogue analysis of SSU RNA, was proposed by Woese and Fox (377). They showed that the 16S rRNA of one group of prokaryotes, the methanogenic bacteria, is no more closely related to the 16S rRNA of true bacteria (eubacteria) than to the 18S rRNA encoded by the eukaryotic nuclear genome. The halobacteria and thermoacidophiles (106) were also found to be members of this third group, the archaebacteria.

These three primary kingdoms (Eukaryota, Eubacteria and Archaebacteria) appear to be evolutionarily distinct from one another, based not only on rRNA sequence data but on other biochemical features as well (375). The existence of three primary kingdoms, representing discrete and ancient evolutionary lineages, implies not only that archaebacterial, eubacterial and eukaryotic nuclear genomes must have diverged from a common ancestral genome (the "progenote", ref. 376) very early in evolutionary history, but that the eukaryotic nuclear genome must have evolved from a separate lineage (the "urkaryotes", ref. 377), rather than directly from the eubacterial lineage.

4.2 Evolutionary Origin of Mitochondria

The molecular biology of the mitochondrial compartment of eukaryotic cells has been studied quite extensively since Nass and Nass discovered in 1963 that chick embryo mitochondria contain their own complement of nucleic acids (239). Because the appearance of

mitochondria in the eukaryotic cytoplasm was obviously a major event in the evolution of the eukaryotic cell, there has been much debate about the evolutionary origin of these organelles (107).

There are basically two kinds of theories that attempt to explain the origin of the separate, compartmentalized genomes of eukaryotic organelles. Autogenous origin (direct filiation) hypotheses (24,50,172,271,275,345) propose that nuclear and organelle genomes became physically compartmentalized and functionally specialized within a single cell. Xenogenous origin hypotheses (4,83,227,228,374) propose that nuclear genomes and mitochondrial genomes were originally in different sorts of cells. Specifically, the endosymbiont hypothesis suggests that mitochondria originated from prokaryotes that invaded (or were engulfed by) a primitive proto-eukaryote that contained the ancestral eukaryotic genome.

Since many mitochondrial proteins are encoded by the nuclear genome (c.f. ref. 50,127,358), the endosymbiont hypothesis requires gene transfer from the endosymbiont to the nuclear genome. This seemingly difficult concept is not essential to the autogenous origin hypothesis and this has been used as an argument against the endosymbiont hypothesis (271,275,345,346). However, evidence of gene transfer from the mitochondrial to the nuclear genome has recently been obtained in the case of a fungus, Podospora anserina (384), and sequences homologous to mitochondrial genes have been found in the nuclear genomes of Neurospora (351), rat (140), human (340), yeast (100), sea urchin (158), maize (169) and in the insect, Locusta

migratoria (110).

The direct filiation theory implies that the eukaryotic nuclear and mitochondrial genomes shared a common ancestor more recently than either of these did with eubacterial or archaebacterial genomes. In contrast, the endosymbiont hypothesis suggests that organellar and eubacterial genomes shared a common ancestor more recently than did eubacterial and either archaebacterial or eukaryotic nuclear genomes. As data accumulate, it should therefore be possible to distinguish between these two opposing theories.

The endosymbiotic theory of mitochondrial origin has drawn support from analysis of mitochondrial protein components (83,127,162,358,365). Strong homologies have been observed between electron transport components of mitochondria and the aerobic bacteria, Paracoccus denitrificans and Rhodopseudomonas spheroides (162,365). However, the validity of phylogenies based on protein sequence data relating eubacteria and organelles has been brought into question (5,6,346). The problem with this approach is that phylogenies based on cytochrome c sequences do not agree with classical (morphologically based) relationships (5,6), and do not always agree with phylogenetic relationships deduced from other protein sequences (6).

Nuclear, organellar and eubacterial comparisons based on protein sequence data are complicated by the fact that most mitochondrial proteins are encoded by the nuclear genome (112), and by the fact that there are no known nuclear-encoded cytosol homologues for many

bacterial and organellar proteins. In contrast, rRNA molecules appear to be good candidates for phylogenetic comparisons because rRNA can be isolated readily in high yields and because rRNA function in mRNA-directed protein synthesis has been universally conserved. Ribosomal RNA genes are found in the genomes of all self-replicating organisms, as well as mitochondria and chloroplasts, and appear to be evolving at a rate slow enough to allow detection of ancient as well as recent evolutionary events (106,375,377,378).

Ribosomal RNA sequence data have been instrumental in solidifying the view that chloroplasts are of endosymbiotic, specifically eubacterial, origin (91,97,120,127,298,330,335). Because of the great diversity in the size and base composition of homologous rRNA species from different mitochondria (124,127,358, Table 1), no comparably compelling evidence for an endosymbiotic origin of mitochondria has emerged from recently published primary sequence data for mitochondrial LSU and SSU RNAs from animals (7,15,61,98,154,177,284, 352), fungi (179,180,204,310,311) and protozoa (300) and mitochondrial 5S rRNA from plants (132,313). Seemingly strong support for an endosymbiotic origin of mitochondria has come from comparisons of T_1 oligonucleotide catalogues that demonstrate the eubacterial nature of wheat (Triticum aestivum) mitochondrial 18S rRNA (27,72). However, even these data have been given an alternative interpretation (346), one consistent with either an endosymbiotic or autogenous evolutionary origin.

One of the major goals of the research presented in this thesis

was to provide further information on both the primary sequence and modified nucleoside content of the wheat mitochondrial rRNAs. Analysis of the data presented here has provided an additional basis for distinguishing between an endosymbiotic or autogenous origin of wheat mitochondria. The results strongly reinforce the idea that wheat mitochondria had an endosymbiotic, specifically eubacterial, evolutionary origin.

4.3. Structure, Evolution and Function of the 3'-Terminal Sequence of SSU RNA

Until recently it was technically difficult to determine the primary structure of long RNA molecules, and consequently sequence information did not extend beyond the 3'-terminal 20-30 nucleotides of SSU RNA (78,141,302,303). However, the application of rapid sequencing techniques to cloned rDNA (7,41,61,91,98,108,117,121,139,143,151,161,164,170,177,179,204,211,225,230,233,273,283,286,287,298,308,310,335,337-339,352,353) and to SSU RNA itself (11,15,48,49,75,90,154,168,205,349) has now made it possible to determine much longer (> 50 nucleotide) 3'-terminal sequences for SSU RNA. Even so, at the time this research was begun, such sequence information was limited to a few representatives of the eukaryotic animal (11,164,287), plant (11,75), and fungal (308) kingdoms, as well as to a single prokaryote (41,49); as such, it did not encompass a broad range of existing organisms. In particular, no primary sequence information existed for the SSU RNA of any member of the kingdom Protista, which of all the eukaryotic kingdoms appears to be evolutionarily the most ancient, and

phylogenetically the most diverse (369).

In order to gain a more comprehensive view of the function and evolution of the 3'-terminal region of SSU RNA, I have determined the sequence of the first 100 nucleotides at the 3'-end of SSU RNA from the protist, Crithidia fasciculata, a trypanosomatid protozoan. Comparisons of eukaryotic and prokaryotic SSU RNA sequences have indicated that this portion of the molecule is a good phylogenetic indicator, and therefore I have also undertaken an analysis of the 3'-terminal sequences of SSU RNA from wheat mitochondria, wheat cytosol and E. coli in the context of mitochondrial evolution (see INTRODUCTION, section 4.2).

The primary sequence at the immediate 3'-end of SSU RNA is, in fact, highly conserved in both prokaryotes and eukaryotes (11,141,350). In all cases a stable hairpin structure can be formed by intramolecular base-pairing starting about 10-15 residues from the 3'-terminus (11,350). The single-strand loop of this hairpin almost always contains two adjacent N^6,N^6-dimethyladenosine (m_2^6A) residues, situated at an identical position in both prokaryotic and eukaryotic SSU RNA (1,78,291). This remarkable preservation of primary sequence, secondary structure, and post-transcriptional modification argues that the 3'-end of SSU RNA plays a universal role in protein biosynthesis.

Functional interactions, by way of complementary base-pairing, have been postulated between the 3'-end of SSU RNA and both mRNA and 5S rRNA. It has been suggested that interaction between SSU RNA and mRNA may serve in positioning the latter correctly for initiation of

translation (302), as well as in determining the intrinsic capacity of ribosomes to translate a particular cistron (303), while interaction between SSU RNA and 5S rRNA may be involved in the association of large and small ribosomal subunits during protein synthesis (9,13). In fact, specific complexes between prokaryotic 16S rRNA and mRNA (318,319), eukaryotic 18S rRNA and mRNA (10), and eukaryotic 18S rRNA and 5S rRNA (13,14) have been observed, and possible sites of interaction in these complexes have been inferred from sequence data (9-11,141,302,303,318,319,386). However, with the exception of the mRNA binding site in E. coli 16S rRNA (95,332), the 18S rRNA binding site in wheat (246) and Neurospora (168) 5S rRNAs, and the 5S rRNA binding site in Neurospora 18S rRNA (168), there is as yet little direct evidence supporting the particular sequences which are supposed to interact in the various proposed complexes between SSU RNA and other RNAs. In the absence of such information, comparative sequence analysis of SSU RNA may be helpful in evaluating how generally applicable the proposed interactions between it and other RNAs may be (11).

METHODS

1. Preparation of Viable Wheat Embryos

Viable wheat embryos were prepared from pedigree wheat seed (*Triticum aestivum* var. Thatcher) by a modification of the method of Johnston and Stern (163).

One kg lots of wheat seed were ground in a commercial blender for 10 sec at low speed with a Powerstat (Type 116B, Superior Electric Co., Bristol, Conn., U.S.A.) setting of 85. The ground wheat seed was then transferred to a series of Endecott test sieves (Canadian Standard Sieve, W.S. Tyler Co.), arranged from top to bottom in the order #12-, 16-, and 30-mesh, and shaken by hand for 30 sec. The crude embryos found on the #30-mesh sieve were saved for further processing. The intact seeds and large endosperm fragments remaining on the #12-mesh sieve were re-ground for 15 sec and again the crude embryo fraction was collected.

The crude embryos obtained from 10 kg of wheat seed were gently shaken in the #30-mesh sieve while the bran was blown from the mixture with a hair dryer. The wheat embryos were then separated from endosperm fragments by stirring in 1800 ml of a mixture of cyclohexane/carbon tetrachloride (10/25, v/v). The embryos floated to the surface and were recovered by aspiration followed by filtration through Whatman #1 filter paper. This flotation procedure was repeated six times and then the embryos were spread on a glass tray and allowed to dry overnight. The organic solvents could be filtered

through glass wool followed by Whatman #1 paper and then re-used. However, fresh solvent mixtures were always used for the final two flotations.

After blowing off residual bran, the embryos were sieved through a #20-mesh sieve, leaving behind large endosperm fragments. Broken embryos and small endosperm fragments were subsequently removed by sieving through a #24 mesh sieve. The purified embryos, approximately 30-35g, were again blown free of residual bran and stored over Drierite at 4°C.

2. Germination of Wheat Embryos

Wheat embryos were allowed to germinate in 24 g or 48 g batches at room temperature for 24 hr in the dark. The embryos were placed in plastic Petri dishes (13.5 cm diameter, 2.4 g/dish) on Whatman 3MM filter paper in the presence of 16 ml of 1% dextrose.

3. Isolation of Wheat Mitochondrial and Wheat Cytosol Total RNA

Wheat mitochondria and cytosol were prepared from germinating wheat embryos by a modification of the subcellular fractionation procedure of Cunningham and Gray (71). The sucrose solutions were prepared on the day before use; however, 2-mercaptoethanol and BSA were not added to these solutions until 30 min prior to use. Operations were carried out below 5°C using chilled buffers and glassware for all RNA isolation procedures.

Embryos germinated for 24 hr (starting weight, 24 g) were ground

with 100 ml of homogenizing medium [0.44 M sucrose, 50 mM Tris-HCl (pH 8.0), 3 mM EDTA, 1 mM 2-mercaptoethanol, 0.1% BSA (fatty acid free), ref. 182] in a chilled mortar for 5 min and then squeezed through four layers of cheesecloth. This grinding procedure was repeated twice with fresh homogenizing medium (100 ml each time). The combined filtrate was centrifuged for 6 min at 1,000 x g and the supernatant was re-centrifuged for 6 min at 2,000 x g in an IEC centrifuge, model PR-6. This supernatant was centrifuged for 20 min at 24,000 x g in a Sorvall RC2-B centrifuge.

For preparation of cytosol RNA, 10 ml of this post-mitochondrial supernatant ("cytosol") were transferred to a test tube containing 10 ml of 2x detergent mix [2% tri-isopropylnaphthalene sulfonate, 12% sodium 4-aminosalicylate, 0.1 M NaCl, 20 mM Tris-HCl (pH 7.4); ref. 261] and 20 ml of phenol-cresol mix [phenol:m-cresol:H_2O:8-hydroxyquinoline, 500:70:55:0.5 (w:v:v:w), saturated with 10 mM Tris-HCl (pH 7.4); ref. 173]. The sample was vortexed and held on ice.

The crude mitochondrial pellet was suspended in 80 ml of homogenizing medium and the two sequential low speed centrifugation steps were repeated. The crude mitochondria were pelleted by centrifugation for 20 min at 18,000 x g. This mitochondrial pellet was resuspended in homogenizing medium to a final volume of 9 ml and layered on 3 sucrose step gradients (3 ml/gradient containing 7.5 ml of 1.55 M sucrose and 15 ml of 1.15 M sucrose in homogenizing medium) and centrifuged in an SW25.1 rotor at 22,500 rpm for 1 hr in a Beckman model L/HV ultracentrifuge. Most of the upper part of each gradient

was siphoned off and the mitochondrial band was removed from the 1.15-1.55 M sucrose interface using a syringe with a needle having a right angle bend. The purified mitochondrial fraction was slowly diluted with two volumes of 50 mM Tris-HCl (pH 8.0), 3 mM EDTA and recovered by centrifugation for 20 min at 18,000 x g.

The purified mitochondrial pellet was resuspended in 9 ml of extraction buffer [10 mM Tris-HCl (pH 8.5), 50 mM KCl, 10 mM $MgCl_2$, ref. 202] and 1 ml of 20% Triton X-100 in extraction buffer was added. The sample was vortexed for 30 sec and then held on ice for 30 sec. This vortex/chill step was repeated five times and then the lysate was clarified by centrifugation for 10 min at 10,000 x g. The supernatant was transferred to a test tube containing 10 ml of 2x detergent mix and 20 ml of phenol-cresol mix as described above for the cytosol fraction.

Both the mitochondrial and cytosol extracts were shaken for 5 min and the phases were separated by centrifugation for 5 min at 2,000xg. Solid NaCl was dissolved in the aqueous phases to a final concentration of 0.5 M (0.6 g/20 ml) and these aqueous phases were re-extracted twice with an equal volume of phenol-cresol. Two volumes of 95% EtOH were added to the final aqueous phases and the purified wheat mitochondrial and wheat cytosol RNA was precipitated overnight at -20°C.

The RNA samples were pelleted by centrifugation for 10 min at 12,000 x g, washed with 5 ml of 80% EtOH, dried and dissolved in 0.5 ml H_2O.

Aliquots (5 µl) were diluted to 1.0 ml with H_2O and the yield of RNA was determined by measuring the absorbance at 260 nm with a Zeiss spectrophotometer. Fifty µl of 3 M NaOAc were added to the remaining 495 µl of sample and this solution was extracted with an equal volume of phenol-cresol and precipitated with two volumes of 95% EtOH at -70°C. Yields ranged from 30-35 A_{260} units of mitochondrial RNA from 24 g of germinating wheat embryos and 150-160 A_{260} units of cytosol RNA from 10 ml of the post-mitochondrial supernatant. The A_{260}/A_{280} ratio of the RNA samples was usually 1.9-2.0.

4. Isolation of Total RNA from Crithidia fasciculata

The strain of Crithidia fasciculata used as an RNA source in this laboratory originated from Dr. G.E. Kidder, Amherst College. Cultures of C. fasciculata were grown at 25°C in a liquid medium (257) containing 20 g/l Proteose-Peptone (Difco B-120), 10 g/l dextrose, 1.0 mg/l folic acid, and 25 mg/l hemin. A stock solution of folic acid (200 µg/ml) was prepared by suspending 120 mg of folic acid in 500 ml of H_2O and adding 5 M NaOH dropwise until the acid just dissolved (pH<7), after which the solution was diluted to 600 ml with H_2O. This solution was stable at 5°C for two months. A stock solution of hemin (5 mg/ml) was prepared by dissolving 2.5 g of hemin (Sigma equine hemin type III) in 500 ml of 50% (w/v) triethanolamine in H_2O. Cultures were maintained by weekly transfers to 6 ml of fresh medium by S. Gray or R. Breckon in F.B.St.C. Palmer's lab in this department.

A 3 l. Fernbach flask containing 1 l. of autoclaved medium was

inoculated with 0.25 ml of a 4 day maintenance culture and incubated with shaking at 25°C for 3 days. Cells were harvested by centrifugation for 15 min at 800 x g. The cell pellet was resuspended three times in 40 ml of cold 0.85% NaCl and pelleted by centrifugation for 10 min at 2,000 x g. The cells were resuspended in 50 ml of 0.05 M Tris-HCl (pH 8.0) and total RNA was prepared by extracting four times with an equal volume of phenol-cresol. The aqueous phase was made 0.5 M with respect to NaCl after the first extraction and RNA was precipitated from the final aqueous phase at -20°C after addition of two volumes of 95% EtOH. The yield of RNA was about 300 A_{260} units from a 1 l. culture.

Crithidia RNA, uniformly ^{32}P-labelled (4 x 10^5 cpm/μg) in vivo, was prepared from a 50 ml culture containing 5 mCi of [^{32}P]P_i (217).

5. Salt Fractionation of Total RNA

Total RNA was dissolved in H_2O to a concentration of 4-5 mg/ml. Solid NaCl was added to a final concentration of 1 M and the solution was held at 4°C for at least 8 hr. The NaCl-insoluble RNA (iRNA) was collected by centrifugation for 10 min at 3,000 x g and reprecipitated from 1 M NaCl. The combined NaCl-soluble (sRNA) fractions were precipitated with two volumes of 95% EtOH at -20°C and collected by centrifugation for 10 min at 12,000 x g. The iRNA and sRNA pellets were washed twice with 80% EtOH and dissolved in 0.5 ml H_2O for A_{260} measurements. After addition of 50 μl of 3 M NaOAc, the solution was extracted with phenol-cresol and precipitated at -70°C with 1 ml of

95% EtOH.

During high salt fractionation of the wheat mitochondrial and wheat cytosol total RNA preparations, 5S rRNA and tRNA selectively remained in the soluble fraction. In the case of C. fasciculata, the NaCl-soluble fraction contained (in addition to tRNA) small rRNA species f, g, and h (= 5S rRNA), while three other small rRNAs (species e, i (= 5.8S rRNA) and j) were found in the tRNA fraction. Therefore, high salt fractionation of C. faciculata total RNA provided a means of preparing species j (72-73 nucleotides) free of contamination by tRNA.

6. Isolation of Wheat Cytosol Ribosomal Subunits

Ten g of viable, ungerminated wheat embryos were ground in 50 ml of solution A [60 mM KCl, 5 mM $Mg(OAc)_2$, 50 mM Tris-HCl (pH 7.6); ref. 68]. The homogenate was filtered through four layers of cheesecloth and one layer of Miracloth. The filtrate was centrifuged twice for 10 min at 37,000 x g and 20% Triton X-100 in solution A was added to the recovered supernatant to a final concentration of 2%. After vortexing, the Triton X-100 lysate was centrifuged for 90 min at 134,000 x g. The pellet was resuspended in 40 ml of solution A and centrifuged for 10 min at 27,000 x g, and the resulting supernatant then centrifuged for 90 min at 134,000 x g.

The ribosome pellet was resuspended in 2.0 ml of solution C [60 mM KCl, 0.1 mM $Mg(OAc)_2$, 50 mM Tris-HCl (pH 7.6); ref. 68] and clarified by centrifugation for 10 min at 12,000 x g. The supernatant

was diluted with solution C to a ribosome concentration of 40 A_{260} units/ml and 0.5 ml was layered over each of three discontinuous sucrose gradients (4 ml steps of 34, 29, 24, 19, 14 and 10% sucrose; ref. 263) in solution C. Gradients were centrifuged in a Spinco SW25.1 rotor at 20,000 rpm for 16 hr at 5°C and were fractionated by puncturing the bottom of the tubes and manually collecting 30 drop (0.5 ml) fractions. Fractions were diluted with 0.5 ml of solution C and A_{260} measurements were taken. Appropriate fractions were pooled and RNA was extracted from the purified large and small ribosomal subunits by the detergent/phenol-cresol method.

7. Isolation of Wheat Mitochondrial Ribosomal Subunits

Mitochondria were purified from germinating wheat embryos (48 g starting weight) and the clarified Triton X-100-lysate was made 0.3% with respect to Brij 58 (40 ml final volume) and centrifuged for 90 min at 134,000 rpm. The ribosome pellet was resuspended in 2.0 ml of dissociation buffer [0.3 M KCl, 3 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.5); ref. 201] and the solution was clarified by centrifugation for 10 min at 27,000 x g. The supernatant (25-30 A_{260} units) was layered on a discontinuous sucrose gradient containing dissociation buffer. Ribosomal subunits were separated and the rRNA extracted as described for cytosol ribosomal subunits.

8. Isolation of Crithidia fasciculata Ribosomal Subunits

The cell pellet recovered from 6 l. of growth medium was washed

three times with 80 ml of cold 0.85% NaCl and collected by centrifugation for 10 min at 2,000 x g. The cell pellet was resuspended in 40 ml of solution A and centrifuged for 10 min at 2,000 x g. The cells were resuspended in 50 ml of solution A, filtered through glass wool, and passed through a French pressure cell at 2,000 lb in^{-2}. The lysate was centrifuged twice for 15 min at 15,000 x g and 0.1 volume of 20% Triton X-100 in solution A was added to the final supernatant.

This solution was layered over discontinuous sucrose gradients consisting of 2 ml of 1.85 M sucrose and 2 ml of 0.7 M sucrose in solution A, and centrifuged for 20 hr at 134,000 x g. The supernatants were carefully removed by aspiration down to the 0.7 M sucrose step, the tubes were filled up with solution A and then all of the liquid was removed. The sides of each tube were washed with 1.0 ml of solution A, taking care not to disturb the ribosome pellets. The purified ribosomes were dissociated to yield large and small subunits. The subunits were separated and the rRNA extracted as described for wheat cytosol ribosomal subunits.

9. Analytical Polyacrylamide Gel Electrophoresis

As a check of their composition and integrity, isolated RNA samples were subjected to electrophoresis either before or after heat denaturation (60°C for 5 min followed by quick cooling on ice) in polyacrylamide tube gels using a modification of the method of Loening (206).

Plexiglass tubes (0.6 cm internal diameter x 11.5 cm long) were covered at one end with dialysis tubing held in place by a strip of Parafilm. The tubing was covered with Parafilm while the gels were being cast and this Parafilm was removed just prior to electrophoresis. The gel mixture was pipetted into each tube, using a wide-bore pipette, to a height of 10 cm. A level loading surface was obtained by layering H_2O over the acrylamide solution before the gel had polymerized.

2.4% Gels were prepared by mixing 3.6 ml of Acrylamide I solution (15% acrylamide, recrystallized from acetone, 0.75% bis-acrylamide, recrystallized from 95% EtOH, filtered through Whatman #1 paper), 7.5 ml of 3 x E1 buffer (199) and 11.2 ml H_2O. This solution was de-gassed and polymerization was initiated by addition of 18 µl of TEMED and 180 µl of freshly prepared 10% ammonium persulfate.

10% Gels were prepared by mixing 7.5 ml of Acrylamide III solution (30% acrylamide, 0.75% bis-acrylamide), 7.5 ml of 3X E1 buffer and 7.1 ml H_2O. The mixture was de-gassed and 37 µl of TEMED and 370 µl of 10% ammonium persulfate were added.

The 10% and 2.4% gels were pre-run at 5 mA/gel for 30 min at 4°C in E1 buffer containing 0.05% SDS, in a Canalco electrophoresis apparatus (Canalco Industries Corp., Rockville, Md). RNA samples (0.1-0.2 A_{260} units) were applied to the gels in 10-70 µl of E1 buffer containing 5-10% RNase-free sucrose.

Electrophoresis of iRNA was in 2.4% gels at 3 mA/gel for 3.5 hr at 4°C, while sRNA was electrophoresed in 10% gels at 5 mA/gel for 5

hr. After electrophoresis the gels were removed from the tubes, soaked in distilled H_2O for 1-3 hr and then scanned in a Joyce Loebl U.V. scanner attached to a Sargent Model SRLG recorder.

10. Purification of Individual Ribosomal RNAs

10.1. Polyacrylamide Gel Electrophoresis

Preparative gels (20 x 20 x 0.3 cm) were prepared using two glass plates (20 x 20 x 0.4 cm), three spacers (20 x 20 x 0.3 cm) and either a 4- or 6-tooth slot former made from Delrin (Dan-Kar Plastic Products, Reading, Mass.). The two side spacers were shortened to 19 cm and a third 20 cm spacer was used to seal the bottom of the gel. The contacts between the bottom and side spacers and between the glass plates and spacers were sealed with Vaseline. The gel molds were held together with Foldback 1412 binder clamps.

The acrylamide solution in TBE buffer [50 mM Tris, 50 mM boric acid, 1 mM EDTA (final pH 8.3); ref. 85] containing 7 M urea (dissolved by heating under hot tap water) was filtered and de-gassed and after addition of TEMED and ammonium persulfate, was poured between the glass plates (amounts of catalysts used to polymerize 100 ml of acrylamide solution were 0.65 ml (2.5-10%) or 0.5 ml (20%) of 10% ammonium persulfate (freshly prepared) and 20 µl (2.5%) or 10 µl (5-20%) of TEMED). The slot former was inserted and then clamped tightly in place with three size 35 ball and socket clamps. If the top of the gel was not clamped in this way, a thin layer of acrylamide would polymerize between the glass plates and the slot former. When

this occurred it was very difficult to remove the slot former after polymerization and the samples did not run evenly into the gel. The gel was usually allowed to polymerize overnight and then the slot former and the bottom spacer were removed. The bottom of the gel was wiped free of Vaseline and the wells were rinsed with TBE buffer.

Gels were pre-run at 500 V with TBE buffer in a home-made gel support similar to those sold by Dan-Kar Plastic Products. A 500 V power supply (Gelman Instruments Co., Ann Arbor, Michigan) was used. The wells of the gel were filled with TBE buffer and a wick (Eaton-Dikeman chromatography and electrophoresis paper, grade 301-85) was wetted with TBE buffer and used to make the connection between the top of the gel and the top buffer chamber. Wicks were always covered with Handi-Wrap to minimize evaporation. Gels were normally pre-run until the current dropped and remained constant (approx. 50% of the starting current).

(a) Electrophoretic Separation of Low Molecular Weight Ribosomal RNAs

sRNA was electrophoresed in 10% acrylamide gels, while iRNA was electrophoresed in 10% gels that contained a 2.5% stacker gel. These 2.5%/10% composite gels were made by pouring the 10% acrylamide solution (acrylamide:bis-acrylamide, 19:1) down one edge of the gel mold to a height of 15 cm. A flat gel surface was obtained by carefully layering H_2O over the acrylamide solution. After an hour the H_2O was removed and the remainder of the gel mold was filled with the 2.5% acrylamide solution (acrylamide:bis-acrylamide, 9:1).

iRNA and sRNA were dissolved to a concentration of 20 mg/ml in

H_2O, diluted with an equal volume of the loading buffer of Peattie (264) [20 mM Tris-HCl (pH 7.4), 1 mM EDTA, 8 M urea containing 0.05% xylene cyanol (XC) and 0.05% bromophenol blue (BB)], heated to 50-60°C for 5 min, quickly cooled and then loaded on the gel (1.0 mg of RNA (100 µl) was loaded per well). The gels were run in fresh TBE buffer at 500 V until the XC dye migrated to the bottom of the gel. The separated RNAs were visualized by U.V. shadowing (148), and the bands were cut out and stored at -70°C.

(b) Electrophoretic Separation of High Molecular Weight Ribosomal RNAs

High M.W. (16S-28S) rRNAs were separated in 2.5% polyacrylamide gels containing 7 M urea. Gels 1.5 mm thick were used since thicker (3 mm) gels often overheated, and gave poor results.

Samples of iRNA, prepared as described in this thesis, or E. coli 23S + 16S rRNA (Miles Laboratories) were dissolved in H_2O to a concentration of 50 A_{260} units/ml and then diluted with an equal volume of loading buffer. Aliquots (50 µl) of this RNA solution were heated to 60°C for 5 min, quickly cooled and then loaded in each well of a 4 well, 1.5 mm thick 2.5% gel. Electrophoresis was carried out for 5 hr at 500 V.

10.2. Electrophoretic Elution

RNA was eluted from the gel slices by a modification of the electrophoretic procedure described by Allington et al. (3). New sample cups (ISCO) were boiled for 10 min in 0.5 M Na_2EDTA. If this procedure was omitted, subsequent chemical sequencing of labelled RNA

did not produce a good yield of small digestion products. It was not necessary to repeat this boiling procedure when the cups were re-used.

Prior to each use, the sample cups, acrylic rings, and membranes (Spectrapor 3, M.W. cutoff 3,500) were boiled for 10 min in 10% sodium carbonate and then rinsed thoroughly with distilled H_2O. The elution was carried out at 4°C with the external buffer chambers containing TBE and the elution cups containing a 10-fold dilution of this buffer. Electrophoresis was from 45 min to 2 hr at 110 V for RNA sizes ranging from 72-212 nucleotides. After this time the buffer was carefully removed from compartments 1-3 of each elution cup and discarded, while the solution in compartment 4 (200 µl) was collected in a 1.5 ml eppendorf tube, along with a 200 µl rinsing of this compartment with distilled H_2O. The membrane was then removed from this compartment using non-wetting Fisherbrand poly gloves, and RNA bound to this membrane was recovered by squeezing it against the edge of the 1.5 ml eppendorf tube, forcing the last drop of liquid out of the membrane.

Sodium acetate (3 M, 40 µl) was added to the eluted, concentrated RNA, and this solution was extracted with phenol-cresol. The RNA was recovered by precipitation with two volumes of 95% EtOH at -70°C. The RNA was pelleted (10 min in an eppendorf centrifuge) and washed with 80% EtOH. After drying, the precipitate was dissolved in 0.5 ml H_2O and A_{260}/A_{280} measurements were taken. After addition of 50 µl of 3 M NaOAc, the solution was re-extracted with phenol-cresol. The RNA sample was stored at -20°C after addition of 1.0 ml of 95% EtOH.

This electrophoretic elution procedure yielded RNA suitable for rapid sequence analysis, with no carry-over of polyacrylamide. However, the yield decreased as the RNA size increased.

10.3. Homogenization of Gel Slices

An alternative method of recovering RNA from gel slices involved a slight modification of the procedure of Rubin (282). The gel slice was homogenized at 0-5°C in a 14 ml glass homogenizing tube containing 2.0 ml of high salt buffer [0.5 M NaCl, 0.1 M Tris-HCl (pH 9.1), 10 mM EDTA] and 2.0 ml of phenol-cresol. After shaking for 20 min the homogenizing tube was centrifuged for 10 min at 2,000 x g (in a 15 ml tube adapter). The aqueous phase was removed and the phenol-cresol phase was re-extracted with the high salt buffer (5 min). After phase separation, the combined aqueous phases were extracted twice more with phenol-cresol. The RNA was precipitated by addition of two volumes of 95% EtOH, pelleted, washed with 80% EtOH, dried, and dissolved in 0.5 ml of distilled H_2O for A_{260} measurements, as above. After addition of 50 µl of 3 M NaOAc, 2-3 additional phenol-cresol extractions were performed to remove any remaining acrylamide. One ml of 95% EtOH was added to the final aqueous phase and the RNA was stored at -20°C.

This homogenization procedure was not as dependent as the electrophoretic procedure on the size of the RNA. Although there seemed to be some acrylamide carry-over, RNA eluted by this method was suitable for rapid chemical or enzymatic sequence analysis.

11. Preparation of Purified End-Labelled Ribosomal RNAs

11.1. 5'-End-Labelling (Donis-Keller et. al. (85))

A 5 µg pellet of low M.W. rRNA or 25 µg pellet of high M.W. rRNA was washed with 80% EtOH, dried, and dissolved in 20 µl H_2O. After addition of 40 µl of 50 mM Tris-HCl (pH 8.2) the RNA was dephosphorylated with 0.02 units of calf intestinal alkaline phosphatase (Sigma Chemical Co.) for 30 min at 37°C. Six µl of 3 M NaOAc were then added and the solution was extracted twice with phenol-cresol and precipitated with 200 µl of 95% EtOH at -70°C for 1 hr. The pellet was washed with 80% EtOH, dried, and dissolved in 70 µl of 10 mM Tris-HCl (pH 7.4), 1.0 mM spermidine, 0.1 mM EDTA, heated to 50°C for 3 min and chilled on ice. Ten µl of 500 mM Tris-HCl (pH 9.5), 100 mM $MgCl_2$, 0.05 M DTT and 4 units of polynucleotide kinase (P.L. Biochemicals) were added. This mixture was transferred to a tube containing 125 pmoles of [γ-^{32}P]ATP that had been lyophilized and dissolved in 20 µl H_2O. After 30 min at 37°C, 100 µl of 4 M NH_4OAc, 1.0 mM EDTA and 600 µl of 95% EtOH were added and the labelled RNA was precipitated at -70°C.

11.2. 3'-End-Labelling (Peattie (264))

A 5 µg pellet of low M.W. rRNA or 25 µg pellet of high M.W. rRNA was washed with 80% EtOH, dried, and dissolved in 4.73 µl H_2O. An aliquot (11.5 µl) of reaction mix [25 µl of 2x partial mix (100 mM HEPES (pH 7.5), 30 mM $MgCl_2$, 20% DMSO), 3.3 µl 50 mM DTT, 0.5 µl BSA (1.0 mg/ml)] and 0.75 µl of 2 mM ATP were added and this solution

was used to dissolve 125 pmoles [5'-^{32}P]pCp that had been lyophilized. Nuclease-free RNA ligase (3 µl, approx. 7.5 units; P.L. Biochemicals) was added and the reaction was allowed to proceed at 4°C overnight. At the end of the reaction, 160 µl of 10 mM Tris-HCl (pH 7.4), 1.0 mM EDTA, 20 µl of 3 M NaOAc, and 600 µl of 95% EtOH were added and the RNA was precipitated at -70°C.

The 2x partial mix used for 3'-end-labelling was stored at 4°C, while all other solutions used for 3'- or 5'-end-labelling were stored frozen at -20°C. Two additional EtOH precipitations of the end-labelled RNA were usually sufficient to remove most of the remaining unincorporated radioactivity. If eukaryotic cytosol iRNA was to be end-labelled, the 5.8S rRNA was first released from 26S rRNA by a heat denaturation step.

[γ-^{32}P]ATP was synthesized (usually by J. Hofman in W.F. Doolittle's Lab in this department), to a specific activity of 5000 Ci/mmol at a concentration of 20 mCi/ml, as described by Walseth and Johnston (360), and stored in 50% EtOH at -20°C. Three mCi of [γ-^{32}P]ATP were lyophilized and used to synthesize [5'-^{32}P]pCp in a final volume of 50 µl in the presence of 10 units of polynucleotide kinase (360). After heating to 90°C for 3 min, the [5'-^{32}P]pCp was stored at -20°C.

11.3. Purification of End-Labelled RNA

End-labelled RNAs were purified from breakdown products that accumulated during labelling by electrophoresis on 2.5% or 10% polyacrylamide gels (20 x 20 x 0.15 cm) containing 7 M urea. Spacers

and either 8- or 10-tooth slot formers were obtained from Dan-Kar Plastic Products. Samples were loaded (20 µl/well) with drawn-out capillary tubes (1.5-1.8 x 100 mm, Kimax-51). The radioactive RNA was detected by autoradiography and eluted as described previously.

Alternatively, 33 x 40 x 0.05 cm sequencing gels were used to purify end-labelled low M.W. rRNAs. Spacers and slot formers (18 tooth, 1.35 cm each) were made from sheets of Delrin (Dan-Kar Plastic Products). Whatman #1 filter paper wicks were used with these thin gels. Teflon could not be used for spacers, since it was found to inhibit polymerization. Not more than 5 µg of RNA was loaded per well in 5 µl of loading buffer. Polyacrylamide gels (6% or 10%) were pre-run at 1200 V and then run at 1700 V (Dan-Kar high voltage power supply) until the labelled RNA (which was detected by autoradiography) was 5-10 cm from the bottom of the gel.

Use of these thin preparative gels had two advantages. First, RNA that had 3'- or 5'-terminal length heterogeneity was separated into several discrete species, each suitable for sequencing. Second, when multiple samples had to be eluted from gel slices, the methods presented above became impractical, whereas concentration of RNA in these very thin gel slices allowed the following RNA extraction method to be used.

11.4. Elution of RNA from Sequencing Gels

A modification of the elution procedure of Maxam and Gilbert (229) was used to elute RNA from thin (0.05 cm) gel slices. Gel slices were shaken overnight at 4°C in a 1.5 ml eppendorf tube

containing 400 µl of buffer (0.5 M NH_4OAc, 10 mM $Mg(OAc)_2$, 1.0 mM EDTA) and 400 µl of phenol-cresol. If less than 4-5 µg of RNA was to be eluted, 5 µg of E. coli carrier tRNA was added. After phase separation in an eppendorf centrifuge, 1.0 ml of 95% EtOH was added to the aqueous phase and the RNA was precipitated at -70°C. There was no acrylamide carry-over since there was no need to crush these thin gel slices.

12. Analysis of Terminal Nucleoside Residues

5'-End-labelled RNA (5 µg or less) was completely digested to nucleoside 5'-monophosphates (pN) with snake venom phosphodiesterase (0.25 mg/ml) in 20 µl of 0.125 M ammonium formate (pH 9.2) for 16-24 hr at 37°C. 3'-End-labelled RNA was completely hydrolyzed to nucleoside 2',(3')-monophosphates (Np) in 10 µl of 1.0 M NaOH for 90 hr at room temperature, then neutralized with 1 µl of glacial acetic acid.

The ^{32}P-labelled terminal nucleotides and appropriate unlabelled marker nucleotides were subjected to two dimensional thin layer chromatography (tlc) on cellulose-coated plastic plates (Eastman Chromagram) containing fluorescent indicator. The tlc plates were dipped through a 10% saturated ammonium sulfate solution and thoroughly dried prior to use. Developing solvent in the first dimension was 95% EtOH:H_2O, 4:1 (195), while saturated ammonium sulfate:propan-2-ol, 40:1 (306) was used in the second dimension. Marker nucleotides were visualized under U.V. light and ^{32}P-labelled

terminal nucleotides were detected by autoradiography. Areas of interest were cut from the tlc plate, placed in scintillation vials and soaked in 1.0 ml of 0.6 M NH_4OH for 30 min at room temperature. Ten ml of Aquasol 2 were added and radioactivity was counted in a Nuclear-Chicago, model 6850 scintillation counter.

13. Rapid Sequencing of End-Labelled Ribosomal RNA

13.1. Chemical Sequencing Reactions (264,265,296)

The chemical digestions used to produce random cleavage in RNA sequencing have been developed for use with RNA molecules that have been 3'-end labelled with [5'-^{32}P]pCp (264); they give poor results with 5'-labelled RNA. The protocol described here does not correspond exactly to that used during the initial stages of this research; however, the conditions to be described are those that now give the best and most consistent results.

A suitable aliquot of 3'-end-labelled RNA was pelleted, washed with 80% EtOH and dried. The sample was dissolved in 30 µl H_2O and 6 µl of *E. coli* carrier tRNA (phenol-extracted, 5 mg/ml in H_2O) were added. The sample was then distributed evenly among six 1.5 ml eppendorf tubes labelled G, A, U, C, alkali, and control. The A and G reactions were performed directly on these 6 µl samples, while the remaining 4 samples were frozen and lyophilized.

(a) G Reaction

After addition of 300 µl of 50 mM sodium cacodylate-HCl (pH 5.5), the sample was mixed with the pipette tip and chilled on ice. One µl

of dimethyl sulfate (DMS, 50 µl, equilibrated immediately before use against two volumes of the above buffer at pH 7.0; see below) was added and mixed in with the pipette tip. The sample was incubated at 80°C for 1 min, then chilled to 0°C, and 75 µl of 1.0 M Tris-HOAc (pH 7.5), 1.0 M 2-mercaptoethanol, 1.5 M NaOAc, 0.1 mM EDTA and 900 µl of 95% EtOH were added. The RNA was precipitated at -70°C (30-60 min), pelleted, dissolved in 200 µl of 0.3 M NaOAc, 0.1 mM EDTA, and re-precipitated with 600 µl of 95% EtOH at -70°C. The pellet was washed with 80% EtOH, dried, and dissolved in 10 µl of 1.0 M Tris-HCl (pH 8.2). Ten µl of 0.2 M $NaBH_4$ (7.57 mg/ml, freshly prepared) was added and the borohydride reduction was allowed to proceed at 0°C for 30 min in the dark. The reaction was stopped by addition of 200 µl of 0.6 M NaOAc, 0.6 M HOAc and 600 µl of 95% EtOH, and the RNA was precipitated at -70°C.

When commercially available DMS was used in the published sequencing procedure, the G-reaction (performed at pH 5.5) was not entirely reproducible and sometimes generated non-specific cleavage products, usually after C residues. These poor results were correlated with the presence of an ultraviolet-absorbing contaminant (λ_{max} = 281 nm) in the DMS used. The ultraviolet absorbance disappeared when the DMS was equilibrated against 50 mM sodium cacodylate buffer at pH 7.0, but not when it was equilibrated against the same buffer at pH 5.5. The G-reaction was much more specific and highly reproducible when the reaction was performed at pH 7.0 or when DMS pre-equilibrated at pH 7.0 was used for reaction at pH 5.5.

No additional carrier RNA was added after the borohydride reduction step, since it was not necessary for precipitation of the labelled RNA and would sometimes result in overloading in the G-lane of sequencing gels.

(b) A > G Reaction

Labelled RNA (6 µl) was mixed into 200 µl of 50 mM NaOAc (pH 4.5) and chilled to 0°C. One µl of diethylpyrocarbonate was mixed in and the reaction was allowed to proceed at 80°C for 4 min. The solution was then chilled on ice, 50 µl of 1.5 M NaOAc and 750 µl of 95% EtOH were added and the RNA was precipitated at -70°C. The pellet was dissolved in 200 µl of 0.3 M NaOAc, 0.1 mM EDTA and re-precipitated with 600 µl of 95% EtOH at -70°C.

(c) U Reaction

The dry RNA pellet was dissolved in 10 µl of 50% hydrazine in H_2O and allowed to sit on ice for 8 min. An aliquot (200 µl) of 0.3 M NaOAc, 0.1 mM EDTA was added, followed by 750 µl of 95% EtOH, and the RNA was precipitated at -70°C. The pellet was dissolved in 200 µl of 0.3 M NaOAc, 0.1 mM EDTA and re-precipitated with 600 µl of 95% EtOH at -70°C.

(d) C Reaction

The dry RNA pellet was dissolved in 10 µl of 3 M NaCl in hydrazine (freshly prepared by dissolving 17.5 mg of oven-dried NaCl in 100 µl of hydrazine). After 20 min at 0°C, 0.5 ml of 80% EtOH (-20°C) was added and the RNA was precipitated at -70°C. The pellet

was dissolved in 200 µl of 0.3 M NaOAc, 0.1 mM EDTA and re-precipitated with 600 µl of 95% EtOH at -70°C.

It was important that both the RNA pellet and the NaCl/hydrazine solution be cooled on ice for 10 min before mixing the two, and that the 80% EtOH be kept at -20°C until immediately before use. If these precautions were not taken, a non-specific background ladder would appear in the C-lane of sequencing gels (see Figs. 6 and 10C). Sometimes NaCl precipitated out of solution when the NaCl/hydrazine solution was cooled on ice; however, the reaction worked well as long as the hydrazine remained saturated with NaCl. If anhydrous hydrazine was not used, a U > C or a U + C reaction resulted.

(e) Alternate C Reaction

3-Methyluridine was found to be so highly susceptible to the above C reaction that cleavage at this base was virtually quantitative. An alternative C reaction was therefore necessary to determine C residues beyond this modified nucleoside.

Six µl of 3'-end-labelled RNA were mixed into 300 µl of 50 mM sodium cacodylate-HCl (pH 7.0), 1.0 µl of DMS was added and the sample was incubated at 80°C for 5 min. After two ethanol precipitations and an 80% EtOH wash, the dried RNA was subjected to the U reaction for 4 min at 0°C and precipitated twice as described. 3-Methylcytosine residues generated by the DMS reaction were more susceptible to the U reaction than were uridine residues, and therefore a C or a C>U lane was generated.

(f) Strand Scission Reaction

Pellets of chemically modified RNA were washed with 80% EtOH, dried, and dissolved in a freshly prepared aniline acetate solution (redistilled aniline:glacial acetic acid:H_2O, 1:3:7). After 20 min at 55-65°C in the dark, the samples were frozen and lyophilized, then dissolved in 20 µl H_2O and lyophilized again (2X).

(g) Partial Alkali Digestion (D.F. Spencer, personal communication)

A sample of end-labelled low or high M.W. rRNA containing 5 µg of E. coli carrier tRNA was dissolved in 20 µl of 0.15 M NH_4OH, heated to 90°C for 1 or 3 min respectively, and chilled on ice. The sample was frozen at -70°C and lyophilized, dissolved in 20 µl of H_2O, and lyophilized again. If the carrier tRNA was omitted the digestion times were reduced to 30 sec and 2.5 min, respectively. This procedure offers two advantages over that of Donis-Keller et al. (85). Since the digestion time is very short, it is not necessary to do the reaction in a sealed capillary, and since the NH_4OH can be removed by lyophilization, the sample can be loaded on a sequencing gel in the same loading buffer as the other samples.

The chemically digested RNA, the alkali-digested RNA, and the control were dissolved in 8-10 µl of Peattie's loading buffer and stored at -70°C. All of the reaction buffers and precipitation solutions used in the chemical sequencing reactions were stored at 4°C. Sodium borohydride and oven-dried NaCl were stored under vacuum at room temperature while aniline was stored at -70°C.

13.2. Partial Enzymatic Digestions (84,85,304)

The enzymatic sequencing procedure worked best with RNA labelled at the 5'-end with [γ-^{32}P]ATP. Although RNA labelled at the 3'-end with [5'-^{32}P]pCp could be sequenced using the partial enzymatic digestion method, the gels were harder to interpret, for reasons to be discussed later.

A suitable amount of 3'- or 5'-end-labelled RNA was pelleted, washed with 80% EtOH and dried. The sample was dissolved in 11 µl of H_2O, 2 µl were removed and used for partial alkali digestion and the remaining 9 µl sample was lyophilized. The end-labelled RNA was then dissolved in 45 µl of freshly prepared reaction mixture [70 µl 10 M urea, 2 µl 1.0 M sodium citrate (pH 5.0), 1.0 µl 0.1 M Na_2EDTA (pH 7.0), 2.5 µl marker dye (2% XC + 2% BB), 3 µl carrier RNA (5 mg/ml), 21.5 µl H_2O]. The RNA solution was distributed into nine 0.5 ml eppendorf tubes labelled T_1a, T_1b, U_2a, U_2b, Ma, Mb, Aa, Ab and control (6 µl in the tubes labelled a and 4 µl in the tubes labelled b, leaving 5 µl as the control).

One µl of RNase T_1 (0.01 u/µl) was mixed into the solution in the T_1a tube, and with a fresh pipette tip, 1.5 µl of the contents of the T_1a tube were diluted into the T_1b tube. The same procedure was followed with the other RNases used: RNase U_2 (0.05 u/µl), pancreatic RNase (RNase A, 0.01 u/µl), and RNase Phy M (a 5-fold dilution of a stock solution, obtained from H. Donis-Keller). The amounts of enzyme used for each digestion were not always as described above. The exact amounts usually depended on the results of the

previous experiment. The control and the RNase digestions were incubated at 50°C for 30 min and then the samples were frozen at -70°C. The lyophilized alkali-digested RNA was dissolved in 11 µl of the above reaction mixture and stored frozen at -70°C.

Under the reaction conditions described (6-7 M urea), RNase T_1 was very reliable for defining G residues. RNase U_2 was quite reliable for defining A residues, except that some A residues were cut very faintly. RNase Phy M had an A>U specificity; however, overcutting resulted in G and C cleavage. RNase A did not cut after every pyrimidine, although it clearly defined pyrimidines wherever it did cut.

13.3. Sequencing Gels

(a) Electrophoresis

Sequencing gels (33 x 40 x 0.05 cm) were prepared using Delrin spacers and slot formers (32-, 40- or 56-tooth). End-labelled partial enzymatic or chemical degradation products were resolved on 6%, 10% or 20% polyacrylamide gels (19:1, w:w, acrylamide:bis-acrylamide) containing TBE buffer and 7 M urea. The gels were pre-run at 1200 V and run at 1700 V. Samples were heated to 60°C for 3 min and then loaded in adjacent wells of the sequencing gel in aliquots of 2 µl or less, using drawn capillary tubes.

When short electrophoresis runs in 10% gels were used to obtain the sequence of the first 25-40 nucleotides, it was found that band distortion occurred within a region 10-15 bases from the end of a labelled RNA molecule. It was also observed that with enzymatic

digestions of 3'-end-labelled RNA, the bands corresponding to the first 6-8 positions did not separate in 10% sequencing gels. These problems were not observed in 20% gels that were run until the BB marker dye had migrated 15 cm. [5'-^{32}P]pCp was run as a marker in these 20% gels to ensure that the radioactive band corresponding to the first base was not lost. The order of migration of the pNps was pUp > pCp $\geq$ pAp > pGp.

The sequence from positions ≈ 40 to ≈ 120 could be read from two loadings of a 10% gel, one loading run 4 hr, the other 6 hr. If the RNA was longer than ≈ 120 bases, three loadings of a 6% gel were used (first loading, XC = 40 cm, second loading, XC = 30 cm, third loading, XC = 20 cm). For some RNA molecules (> 170 nucleotides) it proved useful to run 60 cm long, 6% gels until the XC marker dye had migrated 60 cm.

(b) Autoradiography

After electrophoresis, one glass plate was removed from the gel, a used (fixed) X-ray film (35 x 43 cm) was placed over the gel and air bubbles were rolled out from between the film and the gel. Since these acrylamide gels stick to used X-ray film, the gel was removed from the second glass plate by slowly lifting the X-ray film. The gel was covered with Alcoa-Film (Fisher Scientific), placed in a metal cassette (Du Pont) and subjected to autoradiography at -70°C using Kodak XAR-5 or 3M TriLite X-ray film with Dupont Cronex Lightning-Plus intensifying screens. The 3M film was about five times slower than the Kodak film, but gave better resolution of bands.

"Pick-up films" were used because the glass plates would not fit into the metal cassettes. Labelled RNA from the sequencing gels remained on these "pick-up films" even after the gel was removed, and for this reason they could not be re-used until the ^{32}P had decayed. "Pick-up films" were not used with preparative gels, since RNA could not be recovered from gels treated in this way. Gels that contained 7 M urea could be thawed and re-exposed to X-ray film. However, preparative gels that did not contain urea could only be frozen once, since they could not withstand thawing.

(c) Interpretation of Gels

These rapid sequencing procedures rely on the fact that only one discrete end of the RNA molecule is ^{32}P-labelled and that the radioactive digestion products observed on the autoradiogram are separated on the basis of size, so that the position of a radioactive band on the autoradiogram is directly related to the distance of the cleaved base from the labelled end. Thus, by knowing the specificities of the reactions used to produce the adjacent ladders, the sequence can be read directly from the autoradiogram.

Chemical Sequencing Gels - Since the end-labelled digestion products produced by chemical cleavage have 5'-P while those produced by alkali cleavage have 5'-OH, their mobilities are different and so the first ≈ 15-20 fragments produced by chemical cleavage of 3'-end-labelled material do not align with the alkali digestion products on the sequencing gels. For this reason, if any of the chemical sequencing reactions produces an alkali-type breakdown

product within the 3'-terminal ≈ 20 nucleotides, this will appear as an artifactual, extra nucleotide. Analysis of both alkali and enzyme tracks of 3'-end-labelled RNA is therefore essential in detecting these artifacts. Pseudouridine (Ψ) residues are not cleaved by the chemical reactions used and appear as blanks in chemical sequencing gels. The alkali band for Ψ residues is usually fainter than normal on gels of 3'-end-labelled RNA. $O^{2'}$-Methylnucleoside (Nm) residues are cleaved by the chemical sequencing reactions.

The band spacings between alkali digestion products are seldom as uniform as those between bands representing chemical digestion products. A possible explanation for this is that secondary structure interactions in the alkali digestion products are reduced in the chemically digested RNA by production in the latter of modified bases that are not capable of base pairing but that do not result in strand scission during the aniline reaction. For example, m^3C residues produced in the G reaction are not cleaved by the strand scission reaction and do not form stable Watson-Crick base pairs.

Enzyme Sequencing Gels - Regions of very stable secondary structure are not cleaved appreciably by the RNases used, even in the presence of 6-7 M urea. Pseudouridine residues are cleaved, but Nm residues are not.

RNA labelled at the 3'-end with [5'-^{32}P]pCp gave very complicated patterns on enzymatic sequencing. Since enzymatic digestion products have 5'-OH, 3'-P termini, cleavage after the first base releases an unlabelled Cp, and therefore a band corresponding to the first

nucleotide is not normally observed. Since the digestion conditions occasionally generate more than one hit per molecule, it is also possible to cleave off the terminal Cp in a proportion of the molecules and to cleave after other residues as well. If the 3'-terminal nucleotide is an A, for example, this results in double bands in both the U_2 and Phy M lanes of the sequencing gel. The fact that PhyM cleaves after both A and U makes gels of this sort almost impossible to interpret.

Ambiguities observed in sequencing gels were always reproducible, but could usually be resolved by running the gels at higher temperatures (i.e., - increased current), by using higher percentage gels or by sequencing from the other end of the molecule.

14. Modified Nucleoside Analysis

14.1. Unlabelled Marker Nucleotides

pm^3U and m^3Up were prepared by a modification of the procedure of Pochon and Michelson (270). Four mg of poly U (Miles Laboratories) were shaken overnight at 4°C in a 1.5 ml eppendorf tube containing 630 µl of H_2O, 168 µl of tributylamine and 52.5 µl of dimethylsulfate. The mixture was divided in half and after addition of 40 µl of 3 M NaOAc and 1.0 ml of 95% EtOH, the two samples were precipitated at -70°C. The pellets were re-precipitated (from 200 µl H_2O, 20 µl 3 M NaOAc, 600 µl 95% EtOH) once at -20°C overnight and again at -70°C for 2 hr. One pellet was digested with snake venom phosphodiesterase (to produce pm^3U) and stored at -20°C. The second pellet was incubated

for 16 hr at 37°C in 230 µl of 1.0 M HCl (to produce m^3Up) and the resulting solution was lyophilized, dissolved in 600 µl of H_2O and stored at -20°C.

The spectral properties of the two products (as determined using a Beckman DU-8 Spectrophotometer) at various pH values were as expected for m^3U monophosphates (144). The two products could be distinguished by thin layer chromatography (as described for terminal nucleoside analysis), and each migrated as a single U.V.-absorbing spot. After treatment with calf intestinal alkaline phosphatase, the mobilities of both products were altered and the resultant nucleosides comigrated with authentic m^3U.

All other marker N, Np, pNp, Nm-Np and Nm-N used during this research were obtained from commercial suppliers or were prepared as previously described, from wheat embryo iRNA (307) or sRNA (133).

14.2. Analysis of Uniformly-^{32}P-Labelled Small rRNAs

The Np + Nm-Np + pNp products of alkaline (10 µl of 1.0 M NaOH, room temperature, 90 hr) or RNase T_2 (Sankyo; 0.5 unit in 10 µl of 10 mM NH_4OAc (pH 4.5), 37°C, 16 hr) hydrolysis and the pN + pNm products of snake venom phosphodiesterase hydrolysis of uniformly-^{32}P-labelled C. fasciculata small rRNAs were subjected, together with unlabelled markers, to thin layer chromatography (Merck glass-backed cellulose tlc plates containing fluorescent indicator). Developing solvents were isobutyric acid:conc. NH_4OH:H_2O, 49.4:2:27.6, in the first dimension, and propan-2-ol:conc. HCl:H_2O, 70:15:15, in the second (248).

In some experiments, an Nm-Np fraction was isolated from alkaline or T_2 hydrolysates by chromatography on small columns of DEAE-cellulose (196,307). This fraction was subjected to two dimensional tlc either directly or following dephosphorylation (1.0 unit of calf intestinal alkaline phosphatase in 10 µl of 100 mM ammonium formate (pH 9.2), 37°C, 30 min). In the latter case, the first dimension solvent was butan-1-ol:isobutyric acid:conc.NH_4OH:H_2O 75:37.5:2.5:25 (279), 3X, followed in the second dimension by propan-2-ol:conc.HCl:H_2O, 70:15:15, 1X. Individual Nm-N's were eluted from the tlc plate in 200 µl of 0.6 M NH_4OH, lyophilized and hydrolyzed with snake venom phosphodiesterase. The resulting [^{32}P]pN's were identified by two dimensional tlc as described for terminal nucleoside analysis.

In the tlc system described above, pUp cannot be detected because it co-migrates with [^{32}P]P_i; however, the system used for terminal nucleoside analysis (3X in the first dimension, 1X in the second dimension) gives good separation of all pNp and the four standard pN derivatives.

Modified nucleosides identified in this manner can usually be localized in the primary sequence of small rRNA molecules as a result of their abnormal reactivities in the sequencing reactions.

14.3. Modified Nucleosides Near the 3'-Terminus of SSU RNA

In an effort to establish definitively the identities and positions of modified nucleosides in the 3'-terminal region of wheat mitochondrial and *E. coli* small subunit rRNAs, I have performed

preliminary experiments of the following design.

Five µg of 16S/18S rRNA were partially digested with alkali (0.15 M NH_4OH, 90°C, 3 min), after which the resulting fragments were 5'-end-labelled using 1.0 mCi [γ-^{32}P]ATP and polynucleotide kinase. Specific 3'-terminal fragments were then isolated on a boronate column (which binds cis-diols) and fractionated on a 10% polyacrylamide sequencing gel. The 5'-end-labelled, 3'-terminal fragments were eluted from the gel and digested with snake venom phosphodiesterase. The radioactive mononucleotides were identified by tlc in the system used for terminal nucleoside analysis.

N-[N'-(m-dihydroxyborylphenyl)succinamyl]-aminoethyl cellulose (DBAE-cellulose) was prepared essentially as described by Weith et al. (362), except that 3 g of aminoethyl cellulose (0.32 meq/g, Sigma) were used. DBAE-cellulose chromatography (0.6-0.7 ml bed volume in a 1.0 ml tuberculin syringe) was carried out essentially as described by Rosenberg (281), except that the eluted RNA fragments were recovered by ethanol precipitation.

15. Preparation and Detection of Intermolecular RNA:RNA Complexes

Complexes between wheat cytosol 5S and 18S rRNA and between wheat cytosol 5.8S and 26S rRNA were prepared as described by Azad and Lane (14). Wheat cytosol iRNA (26S + 18S + 5.8S) was heat denatured (60°C, 5 min), and an equimolar amount of unlabelled wheat cytosol 5S rRNA

was added. The RNA was recovered by ethanol precipitation and the pellet was washed with 80% EtOH. A small amount (6.5 μg) of this RNA sample and 20,000 c.p.m. of either 5'-end-labelled wheat 5S or 5.8S rRNA (approx. 1 x 10^7 c.p.m./μg) was dissolved in 15 μl of H_2O, and 5 μl of 1.2 M NaCl were added. The RNA solutions in 0.3 M NaCl were heated to 60°C for 3 min and quickly cooled to 0°C.

C. fasciculata 18S rRNA (2.5 μg, isolated from purified small ribosomal subunits) was incubated with equimolar amounts (20,000 c.p.m.) of either C. fasciculata RNA species e, f, g, h or i, in 20 μl of 0.3 M NaCl for 3 min at 60°C, then quickly cooled to 0°C.

Heat-denatured C. fasciculata tRNA (7 μg, depleted of species f, g, h and tRNA by repeated precipitation from 1.0 M NaCl) was supplemented with equimolar amounts of species f, g and h; 20,000 c.p.m. of species e, f, g, h or i (approx. 1 x 10^6 c.p.m./μg) were added and these samples were also incubated in 0.3 M NaCl for 3 min at 60°C, and quickly cooled to 0°C.

Three μl of 7X E_2 buffer (251) containing 30% sucrose were added to each sample, which was then electrophoresed at 4°C for 4 hr at 150 V in 2.5% (20 x 20 x 0.15 cm, 10 well) polyacrylamide gels containing E_2 buffer. These non-denaturing gels were subjected to autoradiography at -70°C to determine whether or not the radioactive low M.W. rRNAs had formed a complex with the high M.W. rRNAs.

16. Preparation of Nitrocellulose Filters Containing Restriction Fragments of Wheat Mitochondrial DNA

Wheat mitochondrial DNA was prepared, digested with restriction

endonucleases and fragments were separated in agarose gels as described (28). The restriction fragments were transferred to nitrocellulose filters by the method of Southern (312) and stored under vacuum at 4°C. All of the nitrocellulose filters containing DNA restriction fragments used in this research were prepared by Dr. T.Y. Huh in this laboratory and the procedures for preparation of wheat mitochondrial DNA and Southern transfer to nitrocellulose have been described in detail by Bonen (26).

17. RNA:DNA Southern Hybridization

The DNA filter strips were placed in Zip-Lock polyethylene bags and wetted with 2.0 ml of hybridization medium (2 x SSC:50% deionized formamide) containing ^{32}P-end-labelled RNA. Formamide (Fisher) was deionized (269) by mixing for 2 hr at 4°C with 2 g of ion-exchange mixed-bed resin (AG501-X8, 20-50 mesh, Bio-Rad, Analytical Grade).

The bag was sealed inside a heavier plastic bag and incubated in a 42°C water bath for 18-24 hr. The filters were rinsed at 42°C with 50 ml of hybridization medium. Filters were washed in a shaking water bath at 42°C in 500 ml of hybridization medium (containing non-deionized, reagent grade formamide) for 45 min, followed by two 45 min washes in 500 ml of 4 x SSC at 42°C and two additional washes in 500 ml of 4 x SSC at room temperature. The filters were dried and subjected to autoradiography.

RESULTS

1. Wheat Mitochondrial and Cytosol RNA

Polyacrylamide gel profiles of wheat mitochondrial and wheat cytosol total RNA are presented in Figure 3. Both preparations show discrete peaks of 26S and 18S rRNA and an unresolved tRNA-5S rRNA peak. A 5.8S rRNA was released from the cytosol 26S rRNA by heat denaturation, while there was no 5.8S rRNA in the mitochondrial preparation. Several minor U.V.-absorbing peaks were reproducibly observed in the mitochondrial RNA preparations (arrows, Figure 3) but were not present in the cytosol RNA preparations (71). The mitochondrial 26S and 5S rRNAs (present in the large ribosomal subunit) and 18S rRNA (small ribosomal subunit) have been shown by T_1 oligonucleotide "fingerprinting" to be distinct in sequence from their cytosol counterparts (70).

2. Crithidia fasciculata Ribosomal RNA

The large ribosomal subunit of C. fasciculata contains a component (designated species a, Figure 4A) analogous to the 25S-28S rRNA of other eukaryotic ribosomes. However, Crithidia species a is heat-labile, dissociating into two high M.W. components (species c and d, Figure 4B) and one low M.W. component (species i). The small ribosomal subunit contains a single high M.W. RNA (species b, Figure 4C) which appears to have a higher molecular mass (0.825×10^6) than typical eukaryotic 18S rRNAs (approx. 0.7×10^6) (122).

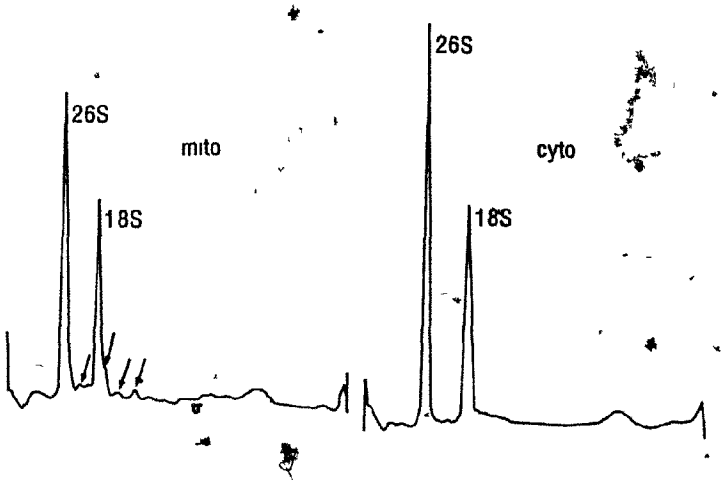


Figure 3. Polyacrylamide gel (2.4%) electrophoretic profiles of total RNA from wheat cytosol (cyto) and wheat mitochondria (mito).

The ribosome of C. fasciculata is unusual in that it contains four novel small rRNAs (designated e, f, g, and j; see Figure 4D) in addition to the 5S (species h) and 5.8S (species i) rRNA species found in other eukaryotes. Their lengths are 212 (e), 183 (f), 135-136 (g), 120 (h), 171-172 (i) and 72-73 (j) (see RESULTS, section 10) nucleotides, with g and j displaying 5'-terminal heterogeneity and i displaying 3'-terminal heterogeneity (296). When ribosomes were dissociated in high potassium (880 mM) buffer, all of the small rRNAs were found in the large ribosomal subunit. However, when subunits were prepared in low magnesium (0.1 mM) buffer (see METHODS), species e and g were released as free RNAs (123).

In non-denaturing polyacrylamide gels (Figure 4D, lanes 1 and 2), wheat 5.8S rRNA (lane 2, 163-164 nucleotides; ref. 218,372) migrates as expected relative to C. fasciculata species e-j (lane 1). However, in denaturing gels containing 7 M urea (Figure 4D, lanes 3 and 4), wheat 5.8S rRNA (lane 3) migrates abnormally slowly compared to the C. fasciculata small rRNAs (lane 4), with an apparent (artifactual) length of 210 nucleotides. The 5.8S rRNA of Acanthamoeba castellanii (162 nucleotides; ref. 216) migrates as expected relative to the Crithidia rRNAs in gels containing 7 M urea (not shown).

3. Intermolecular RNA:RNA Complexes

Azad and Lane (14) have demonstrated that when radioactively labelled wheat 5S rRNA is incubated with unlabelled wheat 18S rRNA in

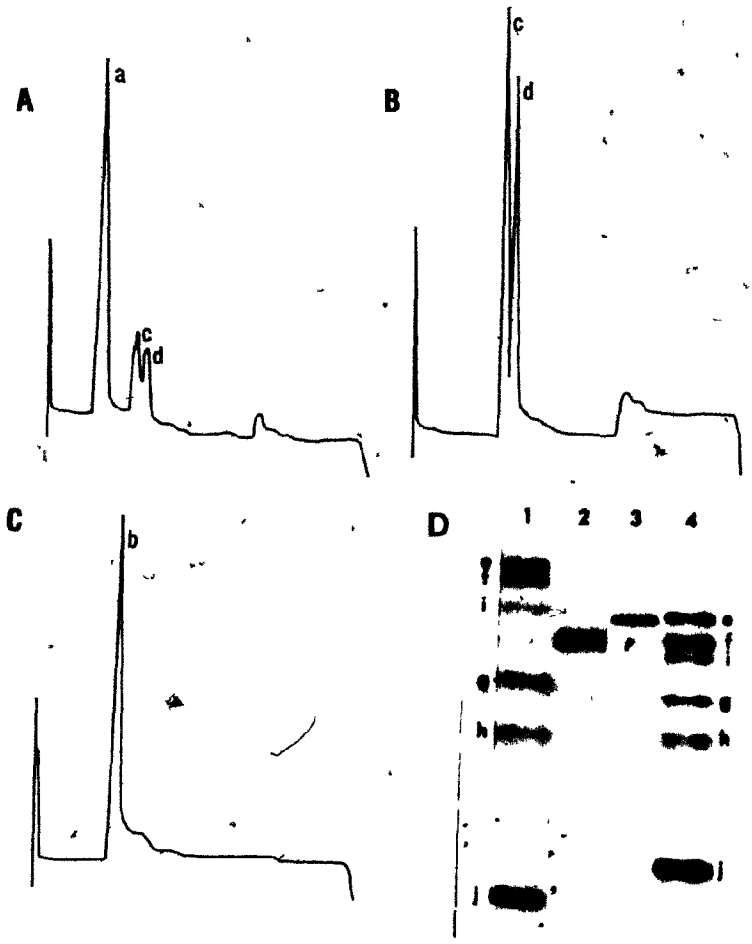


Figure 4. Polyacrylamide gel (2.4%) electrophoretic profiles of: (A) Crithidia 60S ribosomal subunit RNA, unheated; (B) Crithidia 60S ribosomal subunit RNA, heat-denatured; (C) Crithidia 40S ribosomal subunit RNA, unheated. (D) Autoradiograms of 10% polyacrylamide gels of 5'-end labelled Crithidia small rRNAs (lanes 1 and 4) and 5'-end-labelled wheat 5.8S rRNA (lanes 2 and 3). Lanes 1 and 2, non-denaturing gel; lanes 3 and 4, denaturing gel containing 7 M urea.

0.3 M NaCl at 60°C for 3 min, followed by quick cooling on ice, a specific base-paired complex is formed between the two molecules that can be detected by monitoring the transfer of radioactivity from the 5S to the 18S region of a non-denaturing gel.

When 5'-end-labelled Crithidia rRNA species e, f, g ,h (5S) and i (5.8S) were incubated with Crithidia 18S (species b) rRNA under the conditions described above, no intermolecular complexes could be detected (Figure 5A) between Crithidia 18S rRNA and any of these small rRNAs. However, when the 5'-end-labelled Crithidia small rRNAs were incubated in the presence of unlabelled Crithidia high M.W. rRNA (species c and d, the two discrete fragments of Crithidia 28S rRNA, as well as rRNA species b (18S rRNA)), a complex was detected between the 5.8S rRNA (species i) and two of the high M.W. rRNAs (presumably species c and d), as evidenced by the appearance of two bands of radioactivity migrating in the high M.W. region of the gel (Figure 5B).

In each of these experiments a control complex between 5'-end-labelled wheat 5S (5.8S) rRNA and unlabelled wheat 18S (26S) rRNA was detected (not shown). Species j was not included in these experiments because it had not been characterized as a discrete rRNA component at the time these experiments were carried out.

4. 3'-Terminal Sequence of Crithidia fasciculata 18S rRNA

Crithidia rRNA species c and d, the dissociation products of species a, migrate close to species b (18S rRNA) during polyacrylamide

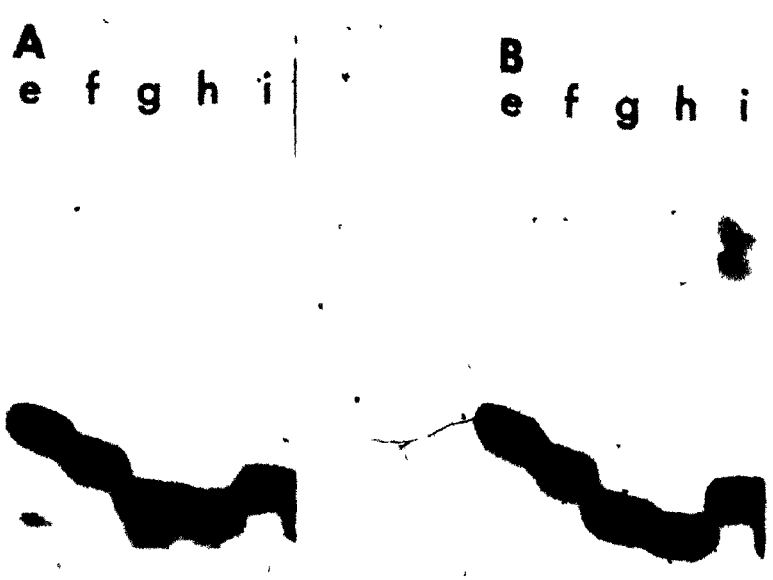


Figure 5. Detection of intermolecular RNA:RNA complexes. 5'-End-labelled Crithidia species e, f, g, h and i were incubated for 3 min in 0.3 M NaCl in the presence of Crithidia species b (A) or species b + c + d (B). Incubation mixtures were resolved in non-denaturing 2.5% polyacrylamide gels, which were subjected to autoradiography.

gel electrophoresis (see Figure 4). Therefore to ensure that c and d did not contaminate b, it was necessary to isolate Crithidia 18S rRNA from purified small ribosomal subunits.

The sequence of the first ≈ 100 residues at the 3'-end of Crithidia 18S rRNA can be read from the representative autoradiograms shown in Figures 6 and 7. The identity of the 3'-terminal U residue was confirmed by alkaline hydrolysis of [5'-^{32}P]pCp-labelled 18S rRNA, which released most (96%) of the radioactivity as Up.

The band spacings in Figure 6 and 7 provide evidence of modified residues at positions 3, 20, 21, 54 and 65. Residue 3 appears to be a modified U, probably Ψ, since this position was cleaved by RNase PhyM, but was not cleaved strongly in any of the chemical sequencing reactions.

Residues 20 and 21 did not give bands in the chemical or enzymatic sequencing gels, suggesting that, as in the case of residue 3, the 18S rRNA is not efficiently cleaved at these positions. The analogous positions are occupied by two adjacent N^6, N^6-dimethyladenosine (m_2^6A) residues in other SSU RNAs (1,48,49,78,286,291) and by adjacent A residues in the rDNA sequences corresponding to the gene for SSU RNA (41,151,164,233,283,287,325). N^6-Dimethylation is expected to block the A-specific chemical cleavage reaction, and blanks in chemical sequencing gels have been observed for the same residues in other SSU RNAs (11,29,75,167,168,349). Since the dinucleotide m_2^6A-m_2^6Ap has in fact been isolated from Crithidia rRNA (122), the presence of m_2^6A is inferred at positions 20 and 21 of

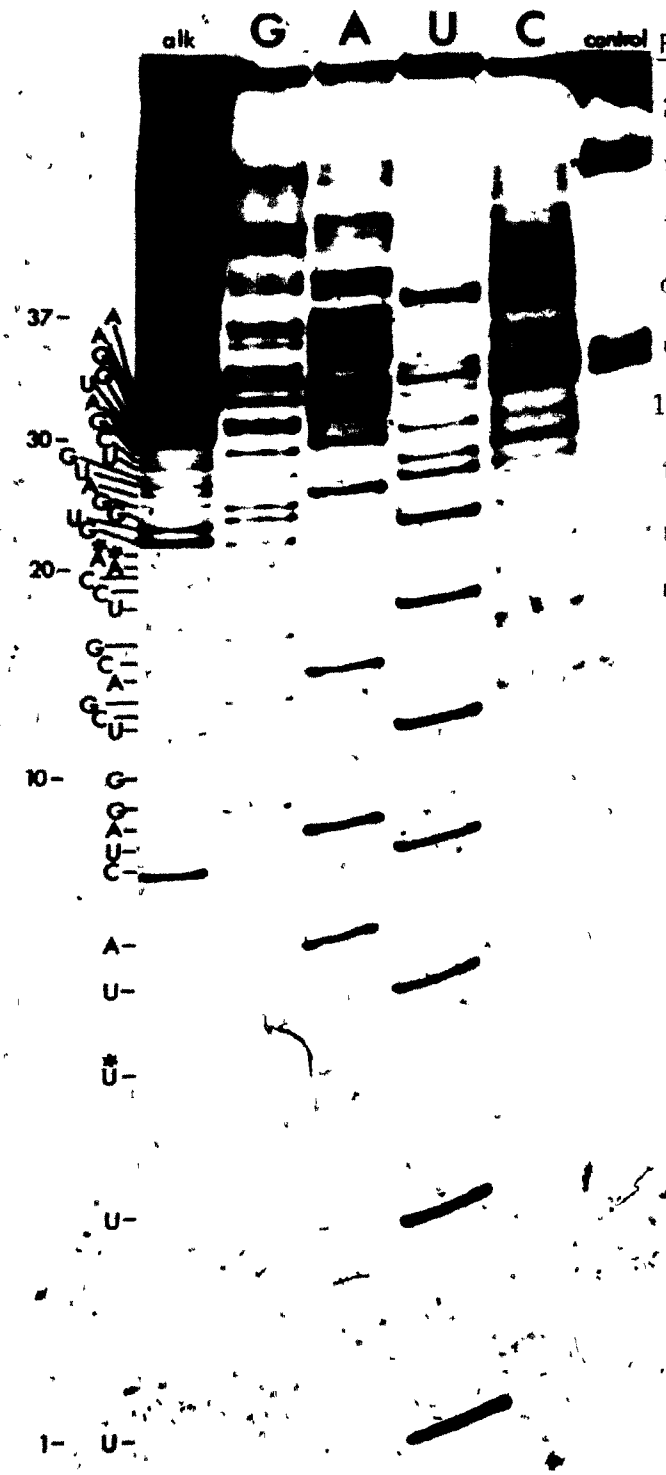


Figure 6. Autoradiogram of 20% polyacrylamide sequencing gel showing the resolution of partial chemical degradation products of 3'-end-labelled Crithidia 18S rRNA. Numbering begins at the 3'-terminus. Asterisks mark probable modified nucleoside constituents, as described in the text.

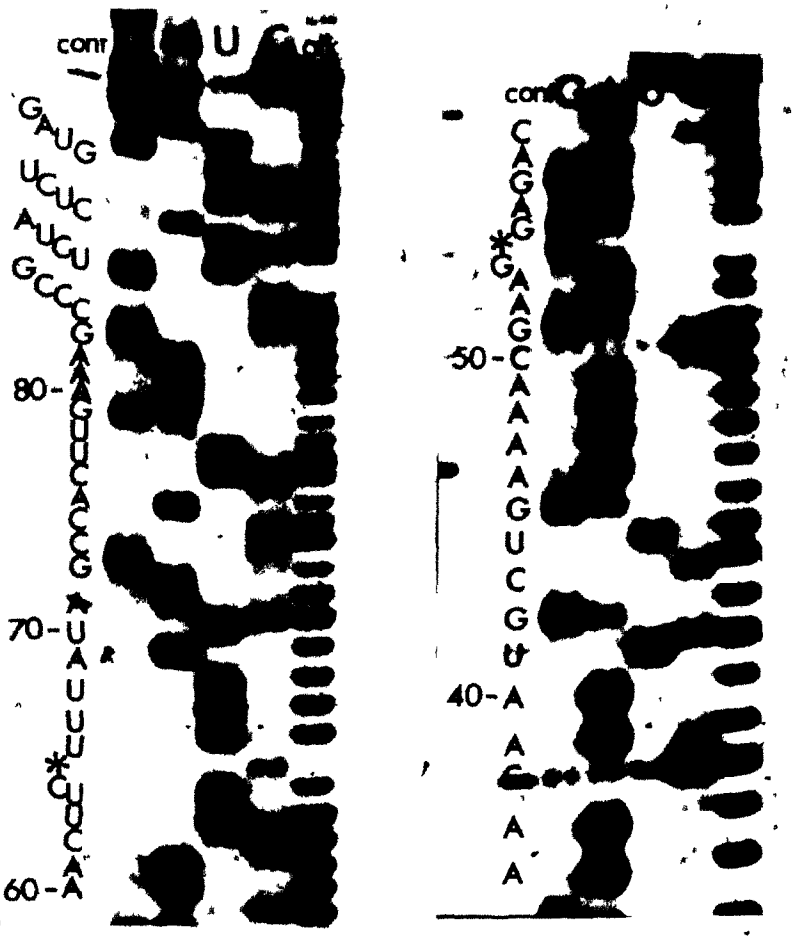


Figure 7. Autoradiograms of 10% polyacrylamide sequencing gels showing the resolution of partial chemical degradation products of 3'-end-labelled Crithidia 18S rRNA. See also Figure 6.

the Crithidia 18S rRNA sequence. The absence of a band corresponding to position 21 in the alkali ladder is consistent with the known alkali resistance of the phosphodiesterase bond joining the m_2^6A residues (245).

Residues 54 and 65 appear as G and C, respectively, in chemical sequencing gels, but there are no corresponding bands in the alkali or enzyme ladders, suggesting that these are alkali-stable, most likely $O^{2'}$-methylnucleosides (Gm and Cm, respectively).

Figure 8 summarizes the data obtained for the 3'-terminal sequence of Crithidia 18S rRNA. As with other SSU RNA sequences, a stable (ΔG = -13.8 kcal/mol; ref. 333) m_2^6A stem and loop structure can be formed close to the 3'-terminus (residues 10-33). The existence of such a hairpin structure in Crithidia 18S rRNA is supported by the fact that its constituent residues are resistant to RNase T_1 and U_2 hydrolysis, even in the presence of 6-7 M urea (data not shown).

5. 3'-Terminal Sequences of Wheat Mitochondrial and Cytosol 18S rRNAs

5.1. Primary Sequences

Partial chemical degradation of [5'-^{32}P]pCp-labelled E. coli 16S rRNA was especially useful for assessing the behavior of known modified nucleosides in this molecule toward the various residue-specific chemical sequencing reactions. For example, residue 45 (a modified U; see below) proved to be highly reactive under the usual conditions of C-specific chemical cleavage (3 M NaCl in

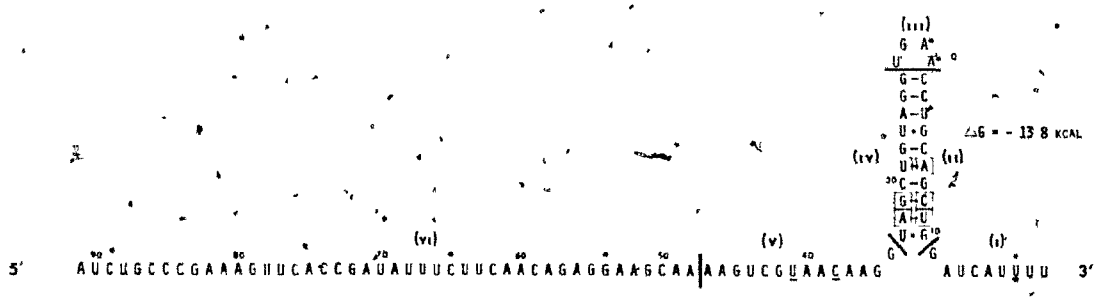


Figure 8. Sequence of the first 91 nucleotide residues at the 3'-end of Crithidia 18S rRNA, indicating the m_2^6A hairpin structure encompassing residues 10-33. The sequence is divided into six sections, defined as follows: (i) extreme 3'-terminal sequence (residues 1-9); (ii) the 3'-half of the m_2^6A stem (residues 10-19); (iii) the single strand loop of the m_2^6A hairpin (residues 20-23); (iv) the 5'-half of the m_2^6A stem (residues 24-33). Section (v) (residue 34-47) is homologous to a single strand region that connects the m_2^6A helix to the next double stranded region in secondary structure models of SSU RNA (90), while section (vi) (residues 48-91) is homologous to the 3'-half of this latter helix. Asterisks denote the positions of probable modified residues, as noted in the text.

hydrazine; ref. 264), as was residue 42 in wheat mitochondrial 18S rRNA; in both cases, cleavage at this position was virtually quantitative (Figure 9; arrows). This made it necessary to use an alternate C reaction (265) to determine C residues beyond this position in the two rRNAs (Figures 9 and 10, C' track). The resulting sequence for the 3'-end of E. coli 16S rRNA is in complete agreement with published data (41,49).

Although end group analysis showed that C was the exclusive 3'-terminal residue of wheat mitochondrial 18S rRNA, length heterogeneity was suggested by the several 3'-terminal fragments produced by C-specific chemical cleavage (Figure 10A). Quantitatively, only the 41- and 42-base fragments were significant, with the latter predominating 4:1. The fragments yielded 91% (41 base) and 88% (42 base) [^{32}P]Cp on end group analysis. This confirms that the mitochondrial 18S rRNA (Figure 25) mostly ends in ...UCG$_{OH}$ (80%) but also in ...UC$_{OH}$ (20%). Enzymatic sequencing of the purified 42-base fragment of the mitochondrial 18S rRNA (Figure 9C) completely confirmed the chemical sequencing results (Figures 9A and 10A).

Terminal heterogeneity of wheat cytosol 18S rRNA was also evident from sequencing gels (appox. 10% of the molecules lacked the ultimate G residue) and 3'-end analysis (which gave 84% G, 11.5% U). These results are consistent with those of Azad and Lane (12) but not with those of Darzynkiewicz et al. (75), who observed some 3'-terminal A residues in their preparations of wheat cytosol 18S rRNA. Except for

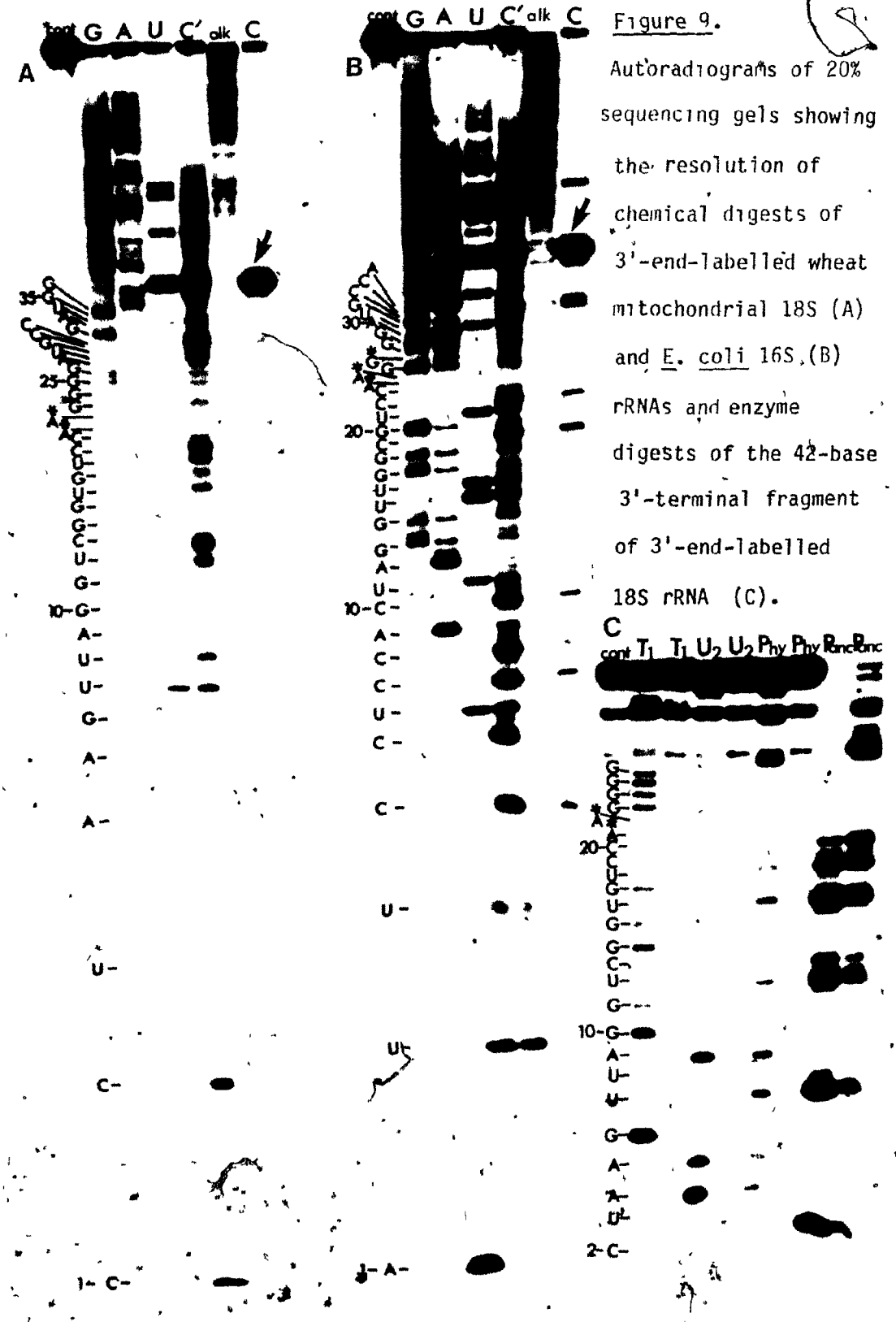


Figure 9. Autoradiograms of 20% sequencing gels showing the resolution of chemical digests of 3'-end-labelled wheat mitochondrial 18S (A) and E. coli 16S (B) rRNAs and enzyme digests of the 42-base 3'-terminal fragment of 3'-end-labelled 18S rRNA (C).

Figure 10. Autoradiograms of 10% sequencing gels showing the resolution of limited chemical digests of 3'-end-labelled wheat mitochondrial 18S (A), E. coli 16S (B), and wheat cytosol 18S (C) rRNAs.

this discrepancy, the primary sequence data obtained here (Figures 10C and 25) agree with, and considerably extend, the 3'-terminal sequence previously reported (75) for wheat cytosol 18S rRNA. The extent of 3'-terminal heterogeneity in wheat mitochondrial and cytosol 18S rRNAs was similar whether they were isolated by direct phenol extraction of intact mitochondria or whole embryos, respectively, or prepared from purified ribosomal subunits, making it unlikely that this heterogeneity was a preparative artifact. Positions 83 to approx. 120 in the wheat cytosol 18S rRNA were almost completely resistant to RNase hydrolysis. This section corresponds to a stable base-paired region in secondary structure models of SSU RNA (90,323).

Figure 11 shows that residues 11-34 of the mitochondrial 18S rRNA sequence can be folded into the hairpin structure characteristic of this region of all SSU RNAs. The base-paired stem differs from that of E. coli 16S rRNA at only three out of twenty positions, affecting two base pairs. Position 24 is occupied by a G residue, as is the homologous position in SSU RNAs from E. coli, chloroplasts, and animal, fungal and protozoan mitochondria. Eukaryotic cytosol 18S rRNA has a U in this position.

5.2. Patterns of Post-Transcriptional Modification

A. Anomalies in the chemical and enzymatic sequencing gels (attributed to altered reactivities of post-transcriptionally modified residues) allowed the following tentative assignments:

(i) 3-Methyluridine (m^3U). Residues 45 in E. coli 16S rRNA, the site of the unusually pronounced C-specific chemical cleavage (Figure

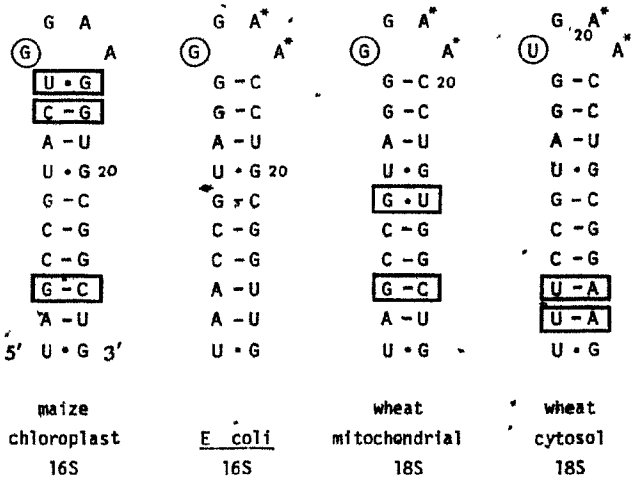


Figure 11. Potential secondary structure in the vicinity of the m^6_2A-m^6_2A (A*-A*) sequence in the indicated SSU RNAs. Rectangles enclose base pairs which differ from the corresponding ones in E. coli 16S rRNA. At the position of the circled residue, all known eubacterial and organellar SSU RNA sequences have a G, whereas all eukaryotic cytosol 18S rRNAs have a U.

9B and 10B), is a methylated U (49) tentatively identified as m^3U (93). The same anomalous C-specific cleavage occurs at position 42 in the wheat mitochondrial 18S rRNA (Figure 9A and 10A), suggesting that (i) the same modified nucleoside (probably m^3U) occupies an identical position in E. coli 16S rRNA and wheat mitochondrial 18S rRNA, and (ii) the modification greatly enhances reactivity in the C-specific chemical cleavage step. There was no cleavage after this residue by RNases Phy M and A, consistent with the known RNase resistance of m^3U (327). In wheat cytosol 18S rRNA, the same position is occupied by unmodified U (Figure 10C).

A similarly-anomalous C-specific cleavage at this position was reported for the 16S rRNA of the cyanobacterium, Anacystis nidulans (29). In this case, cleavage specificity was attributed to secondary structure in the 3'-terminal region, rather than to the presence of an unusually reactive modified residue at the site of cleavage. However, it is unlikely that much secondary structure would persist in 3 M NaCl/hydrazine, and indeed it is for this reason that an alternate C reaction was developed for chemical probing of RNA secondary structure (265). Moreover, wheat, Crithidia and other (11,75,168) eukaryotic 18S rRNAs and hamster mitochondrial SSU RNA (15) show no anomalous cleavage in the C-specific reaction at the position in question, consistent with the known absence of m^3U in hamster cytosol and mitochondrial SSU RNAs (93). A strong C-specific cleavage at this modified uridine has recently been reported for B. stearothermophilus 16S rRNA (90).

Steege et al. (317) have concluded that the modified uridine at the analogous position in Euglena gracilis chloroplast 16S rRNA is $O^{2'}$-methyluridine (Um). Their conclusion was based on the fact that the corresponding nucleoside 3'-monophosphate migrated with Ump on tlc analysis. However, Ump and m^3Up (prepared as described in METHODS) are indistinguishable in the tlc system used by these investigators (data not shown).

Chemical sequencing of 3'-end-labelled E. coli 16S (but not wheat mitochondrial 18S) rRNA revealed an unexpected, fainter band in both the C and U tracks at position 77, above the very intense band corresponding to C-specific cleavage at position 45 (Figures 9B and 10B). Although this might suggest another reactive modified nucleoside at position 77 (known from the gene sequence to be a C (41)), oligonucleotide cataloguing studies (378) have not indicated modification at this position in E. coli 16S rRNA.

(ii) N^6,N^6-Dimethyladenosine (m_2^6A)

E. coli 16S rRNA contains two adjacent m_2^6A residues at positions 24 and 25 (49). These positions in E. coli 16S rRNA and the homologous positions in wheat mitochondrial and cytosol 18S rRNA appear as blanks in chemical and enzymatic sequencing gels. Because m_2^6A residues have been identified in wheat cytosol 18S rRNA (197), in the sequence ...$m_2^6Am_2^6A$... (122), and have been tentatively identified in wheat mitochondrial 18S rRNA (71), the presence of m_2^6A is inferred at positions 21 and 22 in wheat mitochondrial 18S rRNA and at positions 19 and 20 in wheat cytosol 18S rRNA.

(iii) N^2-Methylguanosine (m^2G). $G_{2,}$ in E. coli 16S rRNA was found to be resistant to both RNase T_1 (not shown) and dimethylsulfate (Figure 9B) but to react normally with diethylpyrocarbonate, consistent with published observations on both E. coli (348) and P. vulgaris (48) 16S rRNAs. This residue has been identified as m^2G in E. coli 16S rRNA (348). The analogous residue in wheat mitochondrial 18S rRNA (G_{24}) was also resistant to dimethylsulfate (Figure 9A) but reacted normally with diethylpyrocarbonate and RNase T_1 (Figure 9C).

(iv) O^2-Methyladenosine (Am). In wheat cytosol 18S rRNA A_{51} was resistant to cleavage by both alkali (Figure 10C) and by ribonucleases (not shown), suggesting the presence of Am at this position. No Nm residues were detected in either E. coli 16S rRNA or wheat mitochondrial 18S rRNA within the first 100 nucleotides from the 3'-end.

B. Altered reactivities of modified nucleosides in the sequencing reactions can be very useful in making tentative assignments of the identities of these residues. However, if any significant conclusions are to be based on the modified nucleoside composition of an RNA molecule, the identities of these components must be established definitively.

The methylated uridine at position 45 in E. coli 16S rRNA has only tentatively been identified as m^3U (93), and it has not been formally demonstrated that the m^3U found in complete digests of E. coli 16S rRNA (93) is actually located at position 45. In an effort to resolve this point, 5'-end-labelled 3'-terminal fragments of E.

coli 16S rRNA were isolated by DBAE-cellulose chromatography and separated by polyacrylamide gel electrophoresis (see METHODS, section 14.3). The 5'-end-labelled 45-base fragment was subjected to terminal nucleotide analysis and the resultant autoradiogram, shown in Figure 12A, clearly demonstrates that the modified uridine at position 45 is m^3U. However, there appears to be considerable non-3'-specific binding to the DBAE-cellulose column, which would account for the presence of the four normal [^{32}P]pNs on the autoradiogram. In the same experiment the identities of the m^6_2A residues at position 24 and 25 and the m^2G at position 27 were verified.

In a similar experiment is was found that wheat mitochondrial 18S rRNA does indeed have an m^3U residue at position 42 (Figure 12B) and m^6_2A residues at positions 21 and 22. However, a modified G has not been detected at position 24 in the mitochondrial 18S rRNA.

The non-3'-specific binding to the boronate columns is most likely due either to (a) non-specific attraction between the unreacted, positively charged amino groups on the DBAE-cellulose and the negatively charged phosphates of the RNA fragments, or (b) the 3'-phosphatase activity of the polynucleotide kinase used during 5'-end-labelling of the RNA fragments (45). This problem could probably be overcome or greatly minimized in future experiments by using acetylated DBAE-cellulose (231) and by employing a mutant form of polynucleotide kinase that lacks the 3'-phosphatase activity (44).

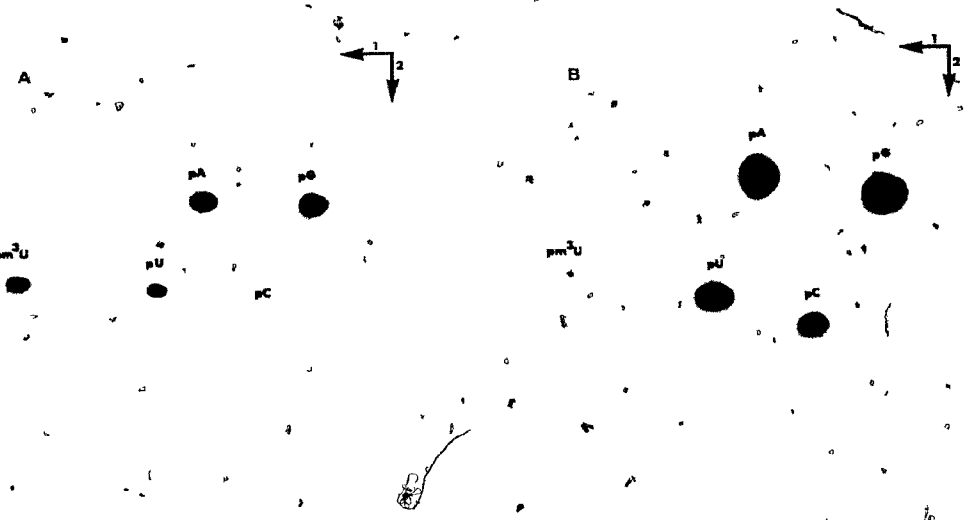


Figure 12. Autoradiograms of thin layer chromatography plates (20 x 10 cm) which establish the presence of pm^3U at analogous positions in *E. coli* 16S (A) and wheat mitochondrial 18S (B) rRNAs. See the text for details.

6. Heterologous rRNA-mtDNA Hybridization

Koncz and Sain (183) demonstrated hybridization between E. coli rRNAs (23S + 16S) and restriction fragments of maize mitochondrial DNA, although these experiments did not directly show that the E. coli rRNA probes were hybridizing specifically to maize mitochondrial rRNA genes. The hybridization of E. coli 16S and 23S rRNAs, individually, to wheat mtDNA restriction fragments known (28) to encode the mitochondrial 26S and 18S rRNAs was therefore examined.

E. coli 23S rRNA hybridized only to those restriction fragments previously reported (28) to contain mitochondrial 26S rRNA coding sequences (Figure 13A), whereas E. coli 16S rRNA hybridized to all of the same fragments (Figure 13B) as wheat mitochondrial 18S rRNA (Figure 13D) in each restriction digest. In addition to the known (28) 18S rRNA-specific Xho I fragments, both E. coli and wheat mitochondrial SSU RNA hybridized weakly to a 5 kbp Xho I fragment that was not scored previously. It is known (28) that there is no significant cross-hybridization between wheat cytosol rRNAs and wheat mitochondrial DNA.

It could be argued that the hybridization of E. coli 16S rRNA to wheat mitochondrial 18S rRNA genes was due largely or solely to the high degree of homology between the 3'-ends of the two SSU RNAs (Figure 25; Table 4), and that the two sequences are otherwise quite divergent. To test this possibility, restriction fragments of wheat mitochondrial DNA were probed with the 45-base, 3'-terminal fragment of E. coli 16S rRNA (prepared by C-specific chemical cleavage of

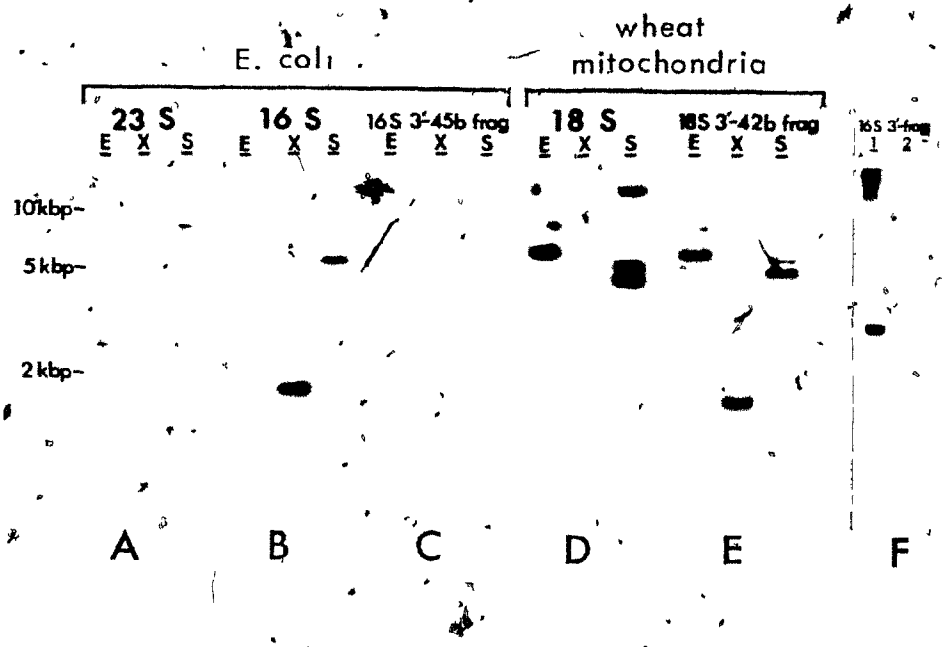


Figure 13. Autoradiograms showing the Southern (312) hybridization patterns obtained with Eco RI (E) Xho I (X), and Sal I (S) restriction fragments of wheat mitochondrial DNA probed with intact E. coli 23S rRNA (A), intact E. coli 16S rRNA (B), the 45-base, 3'-terminal fragment of the 16S rRNA (C), intact wheat mitochondrial 18S rRNA (D), and the 42 base, 3'-terminal fragment of the 18S rRNA (E). Sal I restriction fragments of E. coli DNA (lane 1) and wheat mitochondrial DNA (lane 2) were probed with the 45-base, 3'-terminal fragment of E. coli 16S rRNA (F). All RNA probes were labelled at the 3'-end with [5'-^{32}p]pCp.

3'-end-labelled 16S rRNA). However there was no detectable hybridization (Figure 13C) under conditions in which hybridization of the 42-base, 3'-terminal fragment of wheat mitochondrial 18S rRNA readily occurred (Figure 13E) and where the 45-base, 3'-terminal fragment of E. coli 16S rRNA readily hybridized to Sal I-restricted E. coli DNA (Figure 13F). This indicates that the thermodynamic stability of the heterologous hybrid must be greatly reduced, compared to the homologous one, by the mismatches within the m_2^6A stem and in the region of the Shine-Dalgarno sequence (302,303) in wheat mitochondrial 18S rRNA genes.

7. Number and Orientation of Wheat Mitochondrial 18S and 5S Ribosomal RNA Genes

Genes for wheat mitochondrial 18S and 5S rRNAs are closely linked in the wheat mitochondrial genome (28). In addition, both 18S and 5S rRNAs hybridize with multiple fragments in various restriction digests (28 and Figure 13). The availability of a probe encompassing only the 3'-terminal 41 residues of wheat mitochondrial 18S rRNA has allowed further deductions about the number and arrangement of 18S and 5S rRNA genes in wheat mitochondrial DNA.

Hybridization of the 42-base, 3'-fragment (prepared by C-specific chemical cleavage of 3'-end-labelled wheat mitochondrial 18S rRNA) to Eco RI and Sal I restriction fragments of wheat mitochondrial DNA gave the same pattern as intact 18S rRNA (Figure 13D vs. 13E). This suggests that each of these fragments (4 in the case of Eco

RI-restricted mitochondrial DNA) contains an 18S rRNA gene and therefore that there are at least four copies (or, four different arrangements) of the 18S rRNA gene in wheat mitochondrial DNA. Hybridization of the 42-base, 3'-fragment to Xho I-restricted mitochondrial DNA (Figure 13E) gave only one of the five bands observed when intact mitochondrial 18S rRNA was used as probe (Figure 13D). This result, taken together with the observation that mitochondrial 5S rRNA also hybridizes exclusively to this same band (1.7-1.8 Kb) in Xho I digests of wheat mitochondrial DNA, but gives the same pattern as 18S rRNA when used as a probe against Sal I- and Eco RI-restricted mitochondrial DNA (28), suggests that (i) there are Xho I sites within the 18S rRNA gene; (ii) the wheat mitochondrial 5S rRNA gene is close to the 3'-end of the 18S rRNA rRNA gene; and (iii) there is a single basic structural unit encoding the 5S and 18S rRNAs, but this unit is found at different locations (at least four) within the wheat mitochondrial genome. These conclusions have since been confirmed by mapping and sequencing of cloned restriction fragments containing the mitochondrial 18S and 5S rRNA genes (99,126,128,132).

8. 3'-Terminal Sequences of Wheat Mitochondrial and Cytosol 26S rRNAs

Representative autoradiograms of gels showing the 3'-terminal sequences of wheat mitochondrial and cytosol 26S rRNAs are presented in Figures 14 and 15. The identity of position 31 in the mitochondrial 26S rRNA sequence could not be determined unambigously

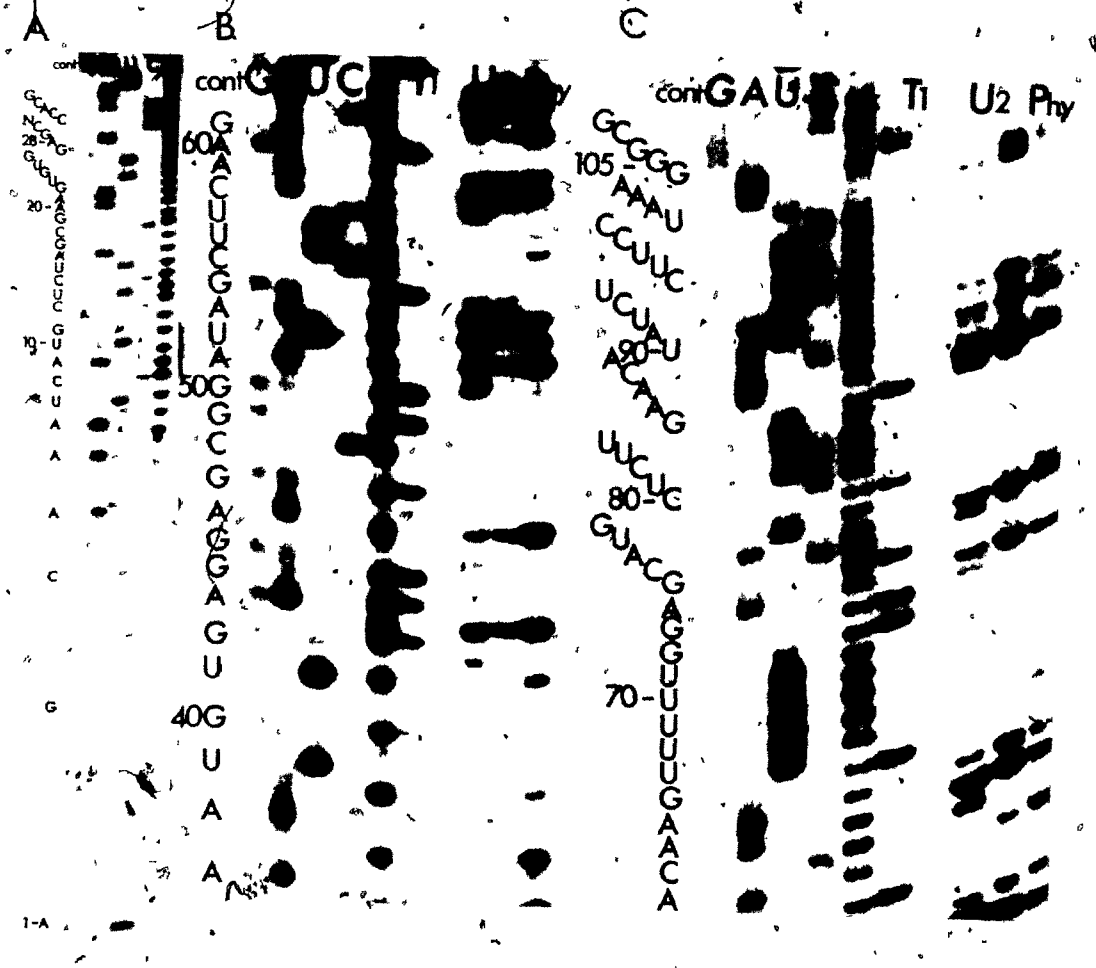


Figure 14. Autoradiograms of 20% (A) and 10% (B) and (C) sequencing gels showing the resolution of partial chemical and enzyme digests of 3'-end-labelled wheat mitochondrial 26S rRNA.

Figure 15. Autoradiograms of 20% (A) and 10% (B) and (C) sequencing gels showing the resolution of partial chemical digests of 3'-end-labelled wheat cytosol 26S rRNA. Asterisks denote positions of probable modified nucleosides.

since no strong cleavage at this position was observed with any of the sequencing reactions employed.

3'-Terminal nucleoside analysis of the mitochondrial and cytosol 26S rRNAs verified the identifies of the 3'-terminal A and U, respectively. However, the presence of minor bands above each band in sequencing gels of the mitochondrial 26S rRNA implies that there is an additional nucleotide (probably A) at the 3'-terminus of a small proportion of the molecules.

It was observed reproducibly that the wheat cytosol 26S rRNA was largely resistant to cleavage by RNases T_1 and U_2 within the 3'-terminal 80 nucleotides, whereas RNase Phy M was found to cleave normally. Blanks in the alkali ladder at positions U_{87} and G_{96} in sequencing gels of the cytosol 26S rRNA suggest that these are alkali-stable (most likely $\underline{O}^2'$-methyl) nucleosides. The faint T_1 cleavage at G_{96} suggests that this nucleoside is only partially $\underline{O}^2'$-methylated.

9. Complete Nucleotide Sequence of Crithidia fasciculata 5.8S rRNA

9.1. End Groups

When 5'-end-labelled C. fasciculata 5.8S rRNA was hydrolyzed with snake venom phosphodiesterase, most (>95%) of the radioactivity was released as pA, thereby establishing the 5'-terminal residue as A. When uniformly-^{32}P-labelled 5.8S rRNA was hydrolyzed with either alkali or T_2 RNase, pAp was released (as the sole pNp derivative), indicating that the 5'-terminus is phosphorylated in vivo.

Alkaline hydrolysis of unresolved 3'-end-labelled 5.8S rRNA released both [^{32}P]Cp (63% of the total radioactivity) and [^{32}P]Ap (34%), indicating 3'-terminal heterogeneity. Two species of 5.8S rRNA could be resolved by electrophoresis in a 6% polyacrylamide gel, and end group analysis of each of these showed that the longer species ended with C while the shorter species ended with A.

9.2. Modified Nucleoside Constituents

Various analyses indicated the presence of Am, Gm, Um, and Ψ, at a level of 2, 1, 0.8, and 1 mol, respectively, per mol of C. fasciculata 5.8S rRNA. The mononucleotide Ψp and the alkali-stable dinucleotides Am-Ap, Am-Gp, Gm-Cp, and Um-Cp were identified among the products of either T_2 RNase hydrolysis (Figure 16) or alkaline hydrolysis (not shown) of uniformly-labelled 5.8S rRNA. The Nm-Np's were also recovered as a separate fraction eluting after the Np's on DEAE-cellulose; they were identified by co-chromatography with authentic Nm-Np markers or with Nm-N markers following dephosphorylation. Each Nm-N was further characterized by venom phosphodiesterase hydrolysis and identification of the resulting [^{32}P]pN. Venom phosphodiesterase hydrolysis of uniformly-^{32}P-labelled 5.8S rRNA yielded pAm, pUm, and pΨ among the ^{32}P-labelled 5'-nucleotide products (pGm was not detected but any that was present would have migrated with pC in the tlc system used).

9.3. Primary Sequence

The sequence of C. fasciculata 5.8S rRNA derived from rapid chemical and enzymatic gel sequencing techniques is shown in Figure

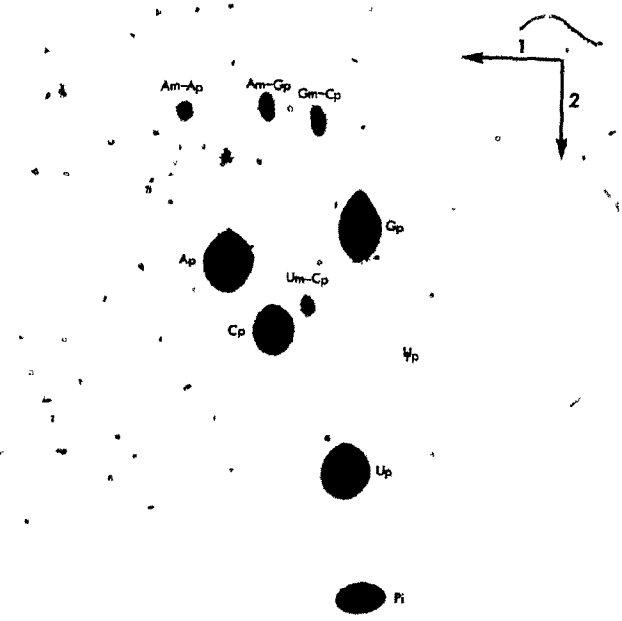


Figure 16. Autoradiogram showing the resolution, by two-dimensional thin layer chromatography (248), of the products of T_2 RNase hydrolysis of uniformly-^{32}P-labelled Crithidia fasciculata 5.8S rRNA. Because it comigrates with Um-Cp in this system, pAp (from the 5'-terminus) was removed by chromatography on DEAE-cellulose (196) before application of the sample (now containing Np and Nm-Np) to the tlc plate.

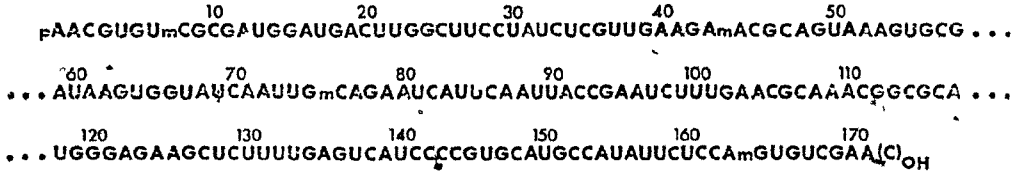


Figure 17. Primary sequence of C. fasciculata 5.8S rRNA. The C ' residue in parentheses at the 3'-end denotes heterogeneity at this terminus (see text).

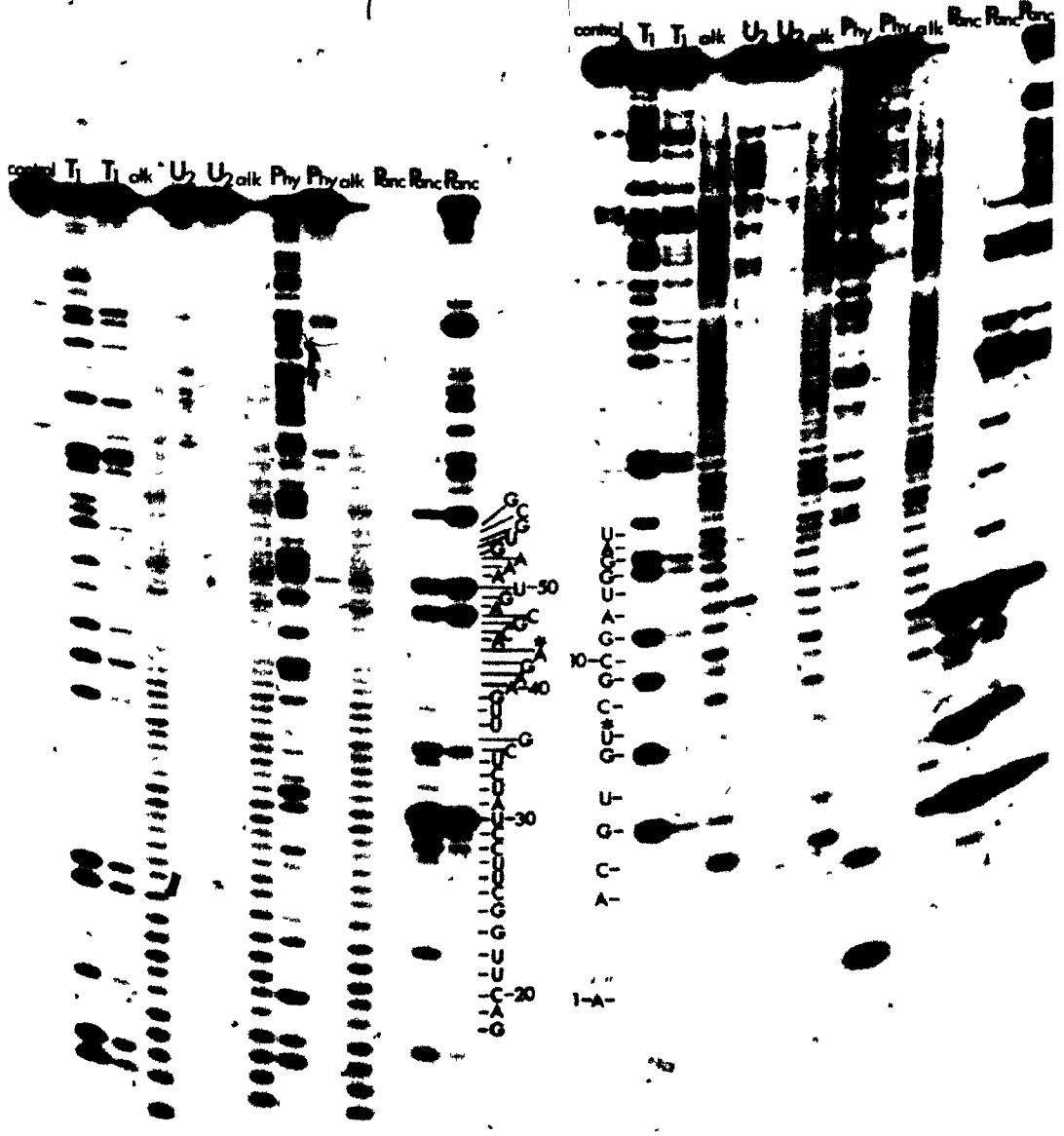


Figure 18. Autoradiograms of sequencing gels of 5'-[^{32}P]-labelled C. fasciculata 5.8S rRNA. The RNA was subjected to partial RNase hydrolysis and the end-labelled products were separated in a 10% polyacrylamide gel.

17. Figure 18 presents a typical enzyme sequencing gel of 5'-end-labelled 5.8S rRNA, from which the first 58 nucleotides of the sequence can be read. Most of the remaining sequence can be read from the representative chemical sequencing gels of 3'-end-labelled 5.8S rRNA shown in Figure 19. Chemical sequencing revealed that the two separated end-labelled species were identical except that the slower-migrating one had an extra C residue at the 3'-end. Residues 70, 82, 86 and 107 appeared reproducibly as both C and G in chemical sequencing gels; however, since these positions were all susceptible to cleavage by pancreatic RNase, they were all confirmed as C's. Residue 161 could not be read from the gel shown in Figure 19 because of the presence of a breakdown product at this position (see control lane). However, this residue was identified as a C in other sequencing gels (not shown). A number of different chemical and enzymatic sequencing gels, using independent preparations of end-labelled 5.8S rRNA, provided confirmation of the sequence presented in Figure 17.

Anomalies in the chemical and alkaline ladders allowed localization of the modified nucleoside constituents identified in hydrolysates of uniformly-^{32}P-labelled 5.8S rRNA. Residue 69 appeared as a U in enzyme gels (not shown) but gave a blank in chemical sequencing gels; also, the alkali band at this position was fainter than normal. These are all features diagnostic of Ψ residues, and therefore the single Ψ residue in C. fasciculata 5.8S rRNA (see above) was placed at position 69. Residues 43, 75, and 163 had no

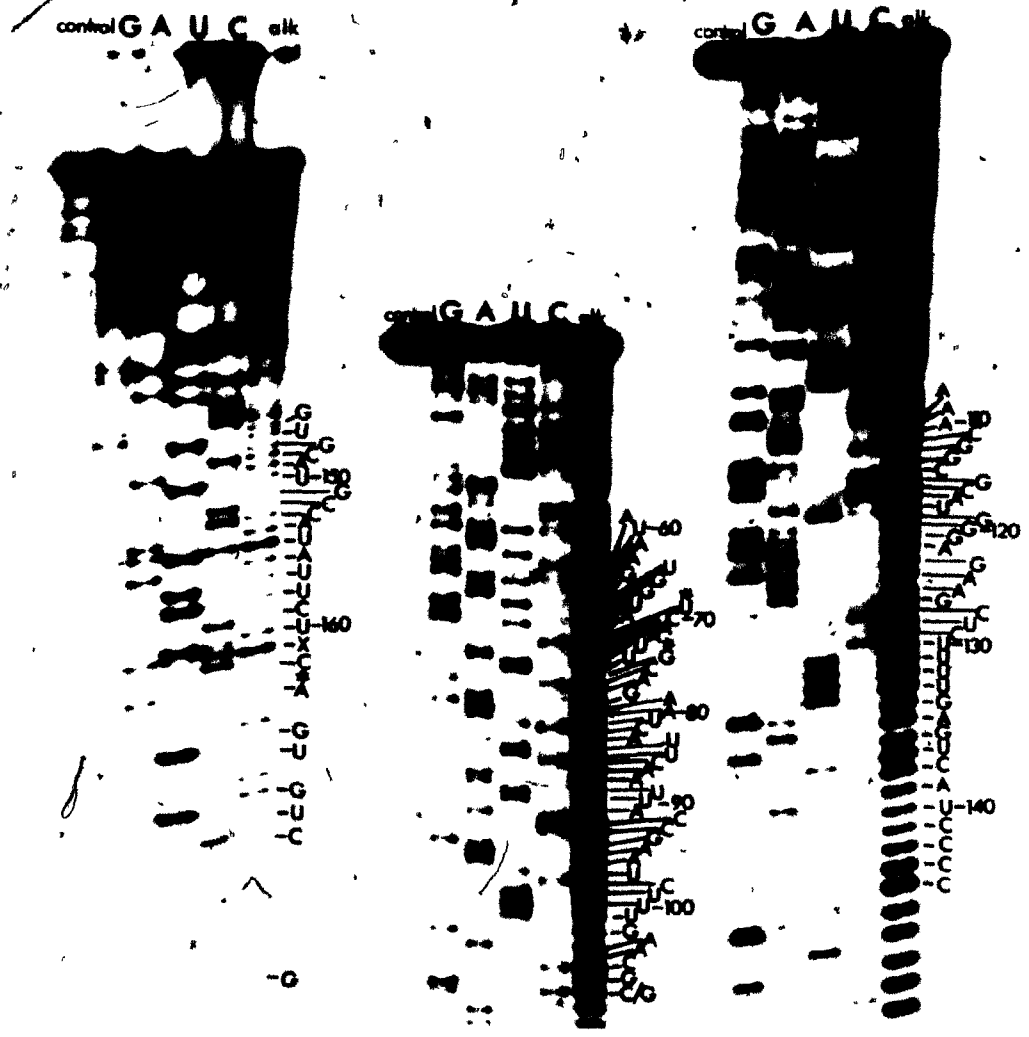


Figure 19. Autoradiograms of sequencing gels of 3'-[^{32}P]pCp-labelled C. fasciculata 5.8S rRNA. The RNA was subjected to partial chemical hydrolysis and the end-labelled products were separated in either 20% (A) or 10% (B) polyacrylamide gels.

corresponding bands in the alkali ladder, indicating the presence of alkali-stable residues at these positions. Consideration of the residues 3' to these positions and the results of Nm-Np analysis (see above) allowed placement of Am at positions 43 (Am-Ap) and 163 (Am-Gp) and Gm at position 75 (Gm-Cp). Residue U, (confirmed as such by chemical sequencing gels of 3'-end-labelled material) gave a much weaker band in the alkali ladder than any of the neighboring residues. Since phosphodiesterase hydrolysis of uniformly-^{32}P-labelled 5.8S rRNA gave approximately 0.8 mol pUm/mol 5.8S rRNA, and RNase T_2 digestion yielded approximately 0.8 mol Um-Cp/mol, the presence of Um was inferred at position 7 and it was concluded that C. fasciculata 5.8S rRNA is incompletely (80%) O^2'-methylated at this position.

10. Primary Structure of Four Novel Small rRNAs from Crithidia fasciculata

The complete primary structures of C. fasciculata rRNA species e, f, g, and j (212, 183, 135-136, and 72-73 nucleotides, respectively) are shown in Figure 20. Autoradiograms of representative sequencing gels, selected to display the entire sequence of each RNA species, are shown in Figures 21-24.

10.1. Primary Sequence

From chemical sequencing gels of 3'-end-labelled material, the nucleotide sequence of each RNA could be read to within one or two residues of the 5'-terminus. Two independent analyses of this type

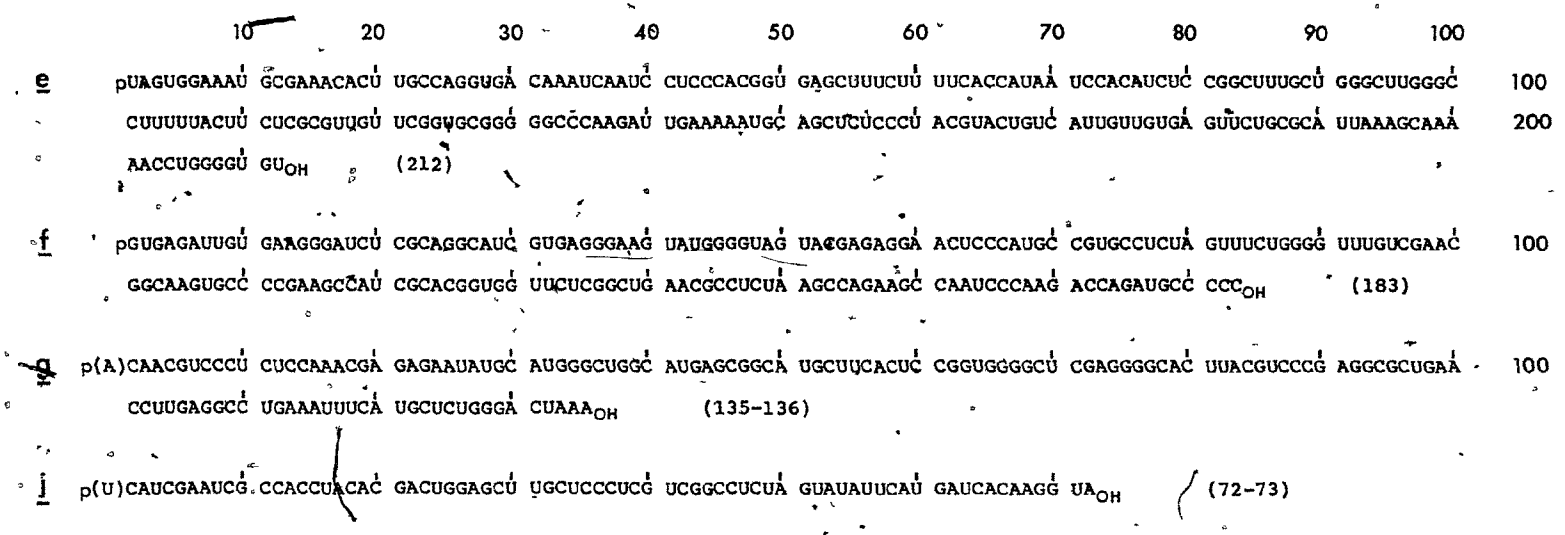


Figure 20. Primary structures of Crithidia rRNA species e, f, g, and j. The residues in parentheses at the 5'-termini of g and j are absent in a proportion of the molecules (see Table 2).

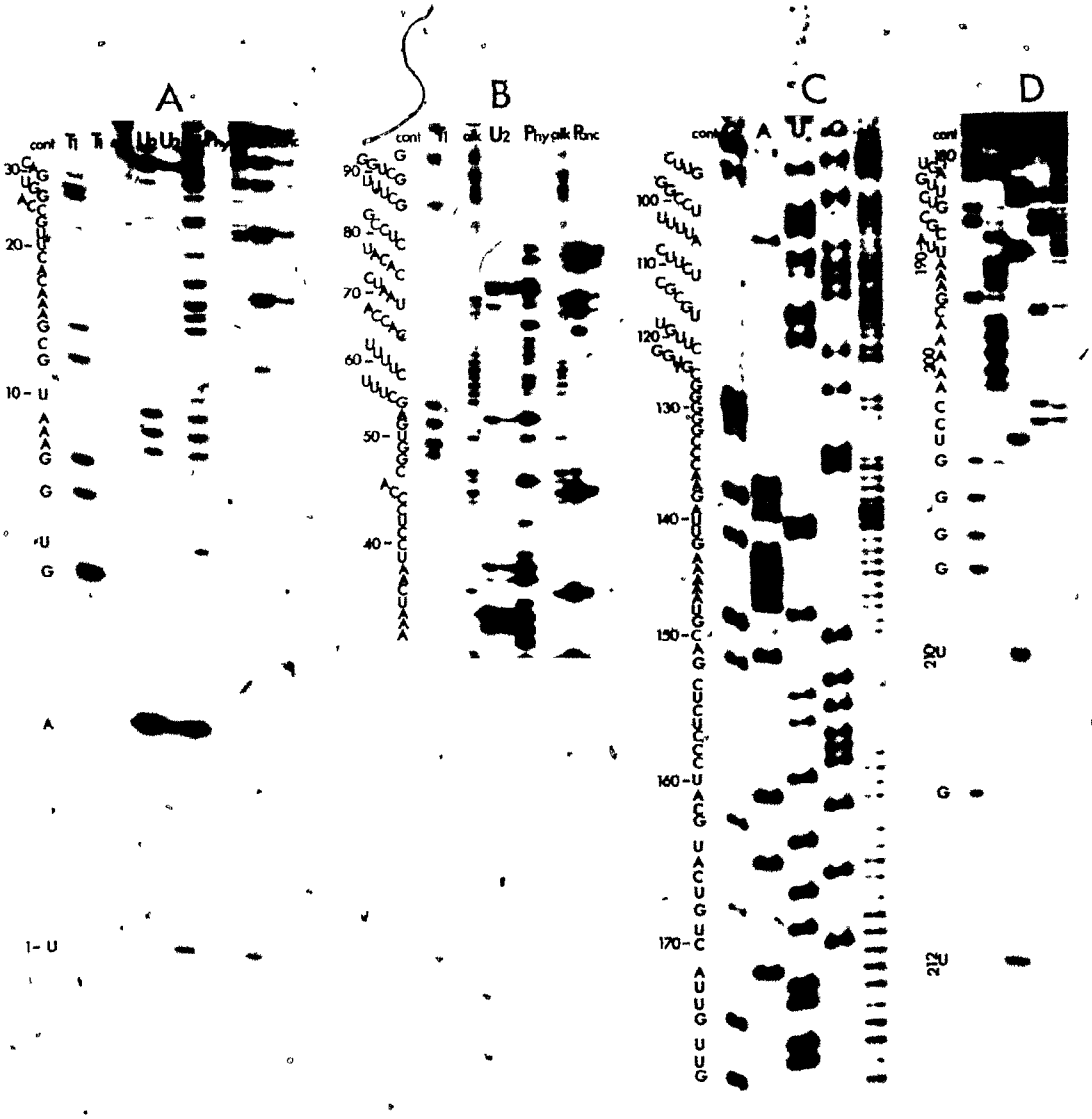


Figure 21. Autoradiograms of polyacrylamide sequencing gels showing the resolution of partial hydrolysis products of end-labelled Crithidia rRNA species e. (A) Enzyme hydrolysis of 5'-labelled RNA, 20% gel; (B) enzymic hydrolysis of 5'-labelled RNA, 6% gel; (C) chemical hydrolysis of 3'-labelled RNA, 6% gel (60 cm long); (D) chemical hydrolysis of 3'-labelled RNA, 20% gel.

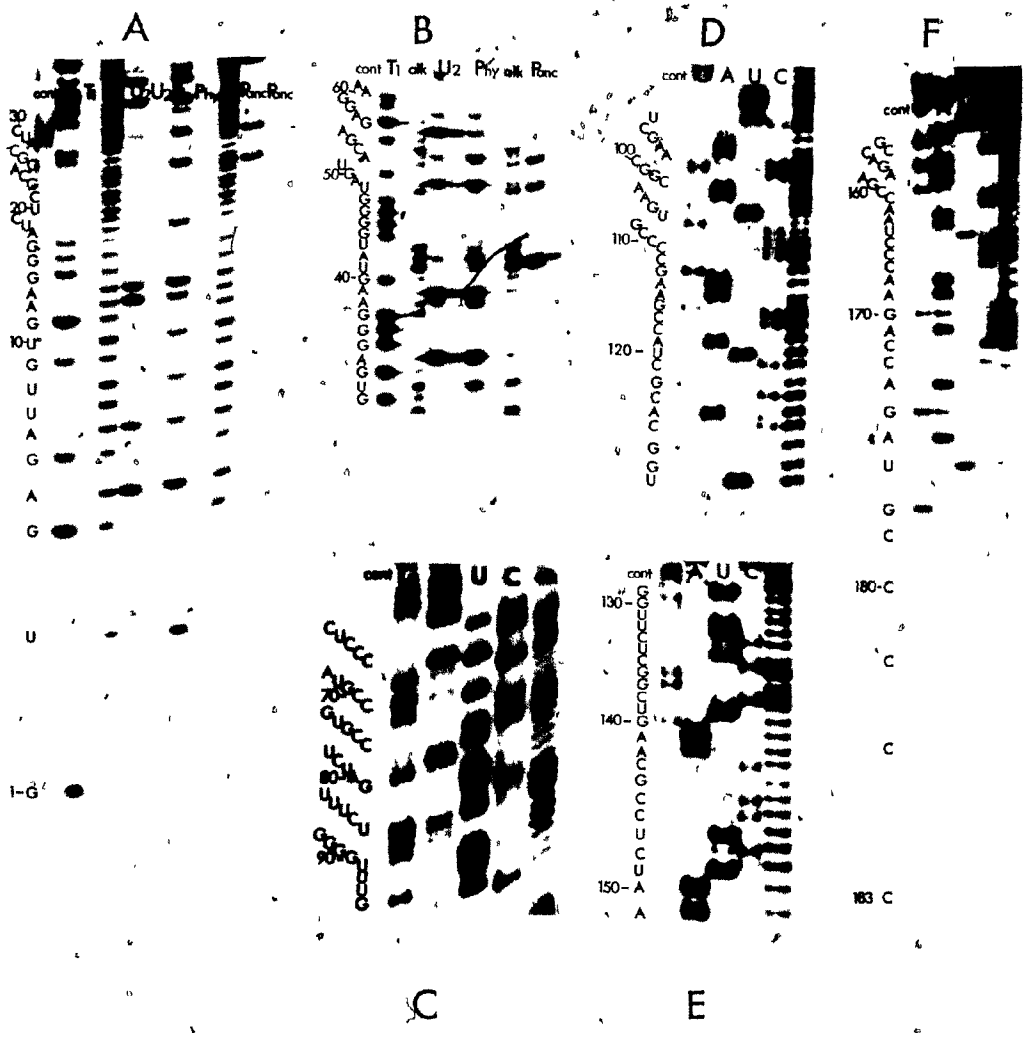


Figure 22. Autoradiograms of polyacrylamide sequencing gels showing the resolution of partial hydrolysis products of end-labelled Crithidia rRNA species f. (A) Enzymic hydrolysis of 5'-labelled RNA, 20% gel; (B) enzymic hydrolysis of 5'-labelled RNA, 6% gel; (C) chemical hydrolysis of 3'-end-labelled RNA, 6% gel (60 cm long); (D) chemical hydrolysis of 3'-labelled RNA, 10% gel; (E,F) chemical hydrolysis of 3'-labelled RNA, 20% gels.

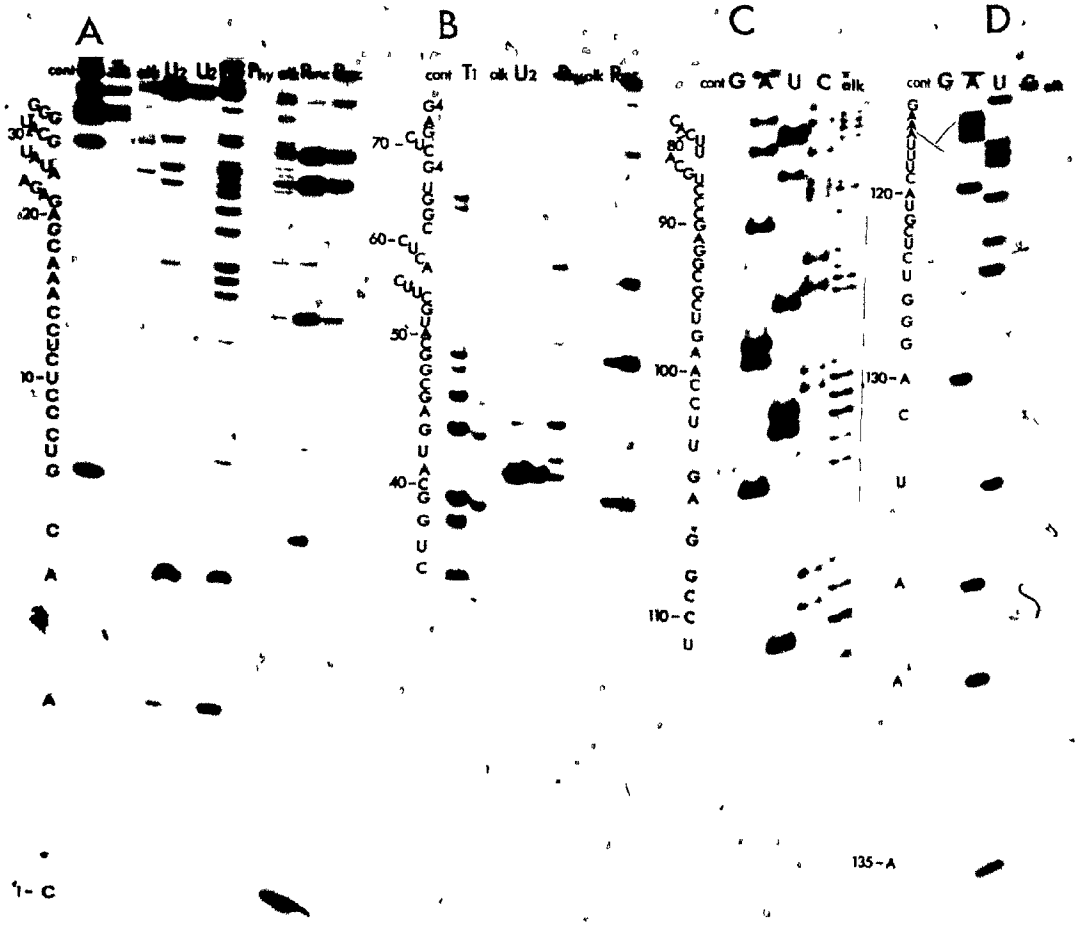


Figure 23. Autoradiograms of polyacrylamide sequencing gels showing the resolution of partial hydrolysis products of end-labelled Crithidia rRNA species g. (A) Enzymic hydrolysis of 5'-labelled RNA, 20% gels; (B) enzymic hydrolysis of 5'-labelled RNA, 6% gel; (C,D) chemical hydrolysis of 3'-labelled RNA, 20% gels.

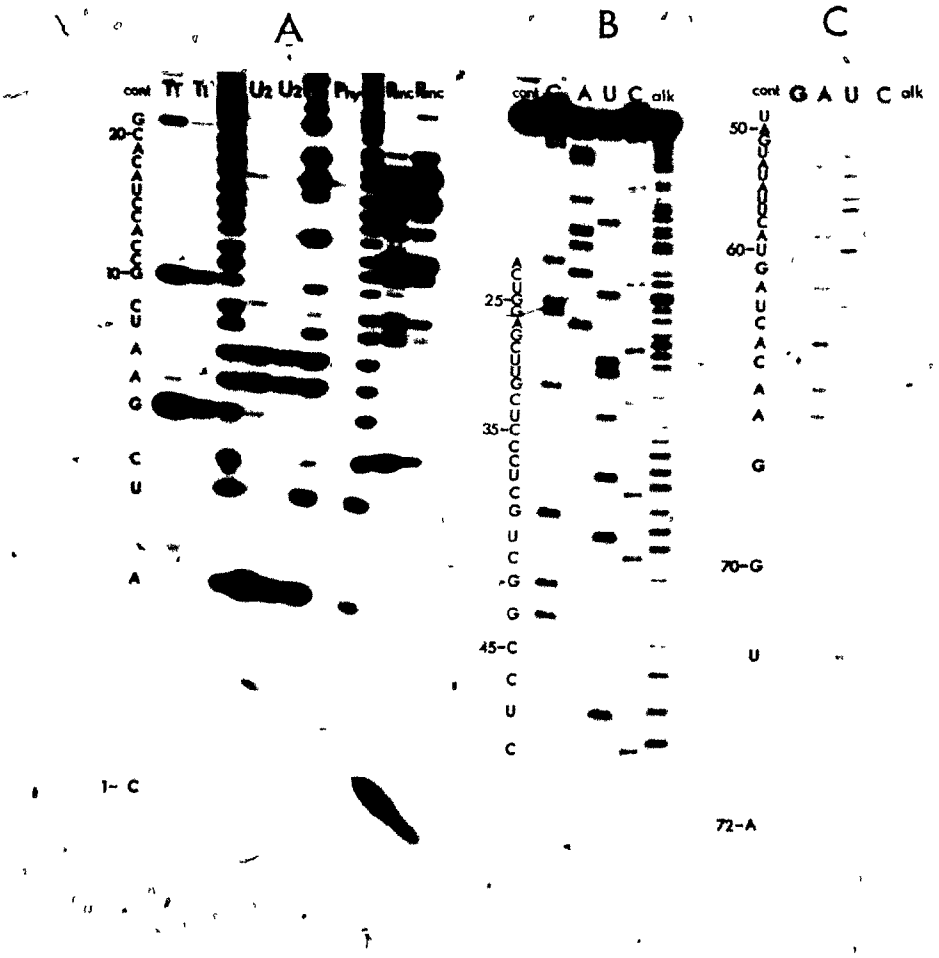


Figure 24. Autoradiograms of polyacrylamide sequencing gels showing the resolution of partial hydrolysis products of end-labelled Crithidia rRNA species 1. (A) Enzymic hydrolysis of 5'-labelled RNA, 20% gel; (B,C) chemical hydrolysis of 3'-labelled RNA, 20% gels.

were carried out for each RNA. In addition, each sequence was determined in its entirety from enzyme sequencing gels of both 5'- and 3'-end-labelled material. These analyses provided substantially overlapping sequence information, in addition to confirming the results of chemical sequence analysis. In this way, all positions in the consensus sequences listed in Figure 20 were multiply verified, with the following exceptions:

species e. At positions 91-93 and 123-124, the number of G residues was ambiguous in both chemical and enzyme gels of 3'-end-labelled material. Gels of 5'-end-labelled material read GGG and GG, respectively. At positions 77-81, 3 C's and 2 U's were evident in 3'-chemical gels, but the order was unclear; 5'-enzyme gels showed UCUCC. Positions 116 and 119 gave strong T_1 bands but also consistently gave weak bands in the U_2 track. It is uncertain whether this is a sequencing artifact or is indicative of cistron heterogeneities. There was no evidence of heterogeneity at these positions in chemical gels.

species f. The order of nucleotides at positions 100-103 and 135-138 was ambiguous in 6% chemical gels of 3'-end-labelled material, but good band separation was obtained in 10% and 20% gels. These latter gels gave the sequence CGGC, as did 6% enzyme gels of 5'-end-labelled material.

species g. At positions 65-68 and 74-77, the number of G residues was ambiguous in 10% chemical and enzyme gels of 3'-end-labelled RNA. However, four G's appeared clearly in 20%

chemical gels of the same material, as well as in 6% enzyme gels of 5'-end-labelled RNA. U_{26} (ambiguous owing to the presence of a breakdown product at this position) was confirmed by chemical sequencing of 3'-end-labelled material, while C_{119} (which in chemical sequencing gels also showed a prominent band in the G track) was confirmed by its susceptibility to pancreatic RNase in the enzyme sequencing procedure. Many positions identified as G also gave bands in the C track during chemical sequencing; these were confirmed as G by susceptibility to diethylpyrocarbonate cleavage (i.e., by a band also appearing in the A track) in chemical sequencing and to RNase T_1 cleavage in enzyme sequencing.

species j. The 3'-end-labelled fragment corresponding to U_{24} ran abnormally fast in 10% chemical and enzyme gels, migrating to the same position as that corresponding to G_{26}; however, it did run properly in 20% gels or when electrophoresis was carried out in 10% gels at a higher temperature. Proper migration was also observed in 10% gels of 5'-end-labelled material. At position 55, there was cleavage by RNase U_2 but there was also a minor band in the T_1 track.

10.2. End Groups

Sequence analysis confirmed that the length heterogeneity observed during purification of species g and j was due to the presence of an extra nucleotide at the 5'-end in a proportion of the molecules (see Figure 20 and Table 2). Alkaline hydrolysis of rRNAs uniformly-^{32}P-labelled in vivo yielded easily identifiable pNp derivatives, which in each case were those predicted by 5'-end group

Table 2. End group analysis of Crithidia fasciculata small rRNAs

	5'-End analysis % radioactivity released[a] as:				3'-End analysis % radioactivity released[b] as:			
Species	pA	pC	pG	pU	Ap	Cp[c]	Gp	Up[c]
e	4.5	0.4	1.5	93.6	0.4	2.1	1.3	96.2
f	2.3	1.0	91.5	5.2	6.1	91.5	1.3	1.1
g	14.5	83.9	1.0	0.6	88.3	5.7	3.2	2.8
h	4.4	3.5	89.7	2.4	1.8	94.6	0.3	3.3
i	88.1	3.2	3.1	5.6	33.8	63.1	0.5	2.6
j	5.0	60.4	2.0	32.6	93.8	4.5	1.2	0.5

[a] By phosphodiesterase hydrolysis of [5'-^{32}P]-labelled RNA.

[b] By alkaline hydrolysis of [5'-^{32}P]pCp-labelled RNA.

[c] Corrected for 6.3% deamination of C residues occurring under the conditions of alkaline hydrolysis (129).

analysis (Table 2). These results indicate that each of the four novel small rRNAs from C. fasciculata is 5'-monophosphorylated in vivo. In contrast to species i (5.8S) but like species h (5S), species e, f, g, and j were found to have homogenous 3'-termini (Table 2). All these small RNAs were readily labelled to high specific activity with [5-^{32}P]pCp in the RNA ligase reaction, without prior phosphatase treatment, indicating the presence of an unphosphorylated 3'-end in each case.

10.3. Modified Nucleosides

Submolar amounts of radioactivity migrating as Ψp (0.4-0.5 mol/mol RNA) were detected in alkaline hydrolysates of species f, g, and j, whereas close to a unimolar quantity (0.93 mol/mol) of Ψp was measured in the case of species e. Position 125 in species e appeared as a U in enzyme sequencing gels, but gave a blank in chemical sequencing gels. Since Ψ is known to be refractory to the U-specific chemical cleavage reaction, it appears that there is a single Ψ residue at position 125 of C. fasciculata species e. There was no evidence of similarly-behaving residues in sequencing gels of species f, g and j. It is not clear whether the "[^{32}P]Ψp" detected in these cases is attributable to partial U→Ψ modification at particular positions in these RNAs, or whether it represents spurious radioactivity co-migrating with marker (unlabelled) Ψp in these analyses. It was previously concluded that Crithidia 5S (217) and 5.8S rRNAs contain 0 and one Ψ residue/molecule, respectively. These RNAs yielded 0.48 (5S; species h) and 0.99 (5.8S; species i) molar

equivalent of pΨ upon phosphodiesterase hydrolysis of uniformly-^{32}P-labelled material.

As in the case of Crithidia 5S rRNA (217), but in contrast to 5.8S rRNA, no alkali-stable dinucleotides (Nm-Np) were detected in the four small rRNAs analyzed here.

DISCUSSION

1. Evolutionary Relationships Based on Ribosomal RNA Sequence Comparisons

For the purpose of determining phylogenetic relationships among eukaryotes, eubacteria, archaebacteria and organelles (specifically mitochondria), we have chosen the approach of comparative analysis of rRNA sequences. A complete rRNA sequence can be obtained by cloning the relevant rRNA gene followed by subsequent sequencing of the cloned DNA either by the chemical method of Maxam and Gilbert (229) or by the dideoxy-chain termination method of Sanger (232,290). However, determination of complete rRNA gene sequences is not very practical in terms of surveying a wide range of organisms, since considerable time is required to clone and analyze any particular rRNA gene.

Although the gene cloning step can be avoided by direct sequencing of the rRNA itself, there are as yet no convenient methods for direct determination of complete sequences of large (16S-28S) rRNAs. Partial rRNA sequence data can be obtained by T_1 oligonucleotide cataloguing of rRNA molecules uniformly labelled with [^{32}P]Pi *in vivo* (342). The problem with this approach, especially in the case of organelles, is that it is often difficult to obtain adequate amounts of rRNAs, labelled to sufficiently high specific activity. For example, a starting amount of 125 mCi of [^{32}P]Pi was required to produce the quantities of wheat mitochondrial 18S [^{32}P]rRNA needed to generate a T_1 oligonucleotide catalogue of this

molecule. Even then, the sequence of certain oligomers could not be completely elucidated (27).

More recently, techniques have been developed (84,85,264,304) which, after end-labelling of the RNA *in vitro*, allow direct determination of the sequence of about 100-200 nucleotides from the labelled end. Since RNA uniformly labelled *in vivo* is not required for these new rapid RNA sequencing techniques, organelle rRNAs can be analyzed as rapidly and as accurately as eukaryotic or prokaryotic rRNAs. For these reasons we have chosen to use the partial chemical degradation method of Peattie (264) and the partial enzymatic hydrolysis methods of Donis-Keller *et al.* (84,85) and Simoncsits *et al.* (304) in determining rRNA sequences from a variety of sources.

While the enzymatic sequencing procedure works well with both 3'- and 5'-end-labelled rRNA, the chemical sequencing procedure is only applicable to 3'-end-labelled RNA. A sequence cannot be determined with complete confidence using the enzymatic method alone, and we are therefore limited in the final analysis to comparisons of the complete sequences of small (e.g., 5S) rRNA sequences or 3'-terminal sequences of large rRNAs.

Although 5S rRNA sequences have been widely used for constructing phylogenetic trees, this RNA species has been detected only in the mitochondria of plants (200), but not those of animals, fungi, or protozoa, and therefore these latter organelles cannot be included in phylogenies based on 5S rRNA sequences. In the case of LSU RNA, the sequence at the 3'-end has diverged to such an extent that it is not

possible to align the 3'-terminal sequences of mitochondrial LSU RNAs with each other or with those of eukaryotes and prokaryotes. Indeed, the 3'-terminal sequence of eukaryotic LSU RNA has diverged to such an extent that the rat and yeast sequences appear to be no more closely related than would be expected by chance (52). For these reasons, the 3'-terminus of LSU RNA is not a good phylogenetic indicator.

In contrast, the primary sequence (11,141,350) and secondary structure (90) at the 3'-end of SSU RNA is highly conserved, and post-transcriptional base methylations appear to be clustered near the 3'-end of this molecule (219). It therefore seems likely that this region of SSU RNA has a crucial (probably universal) function in protein synthesis. Primary sequence analysis of the 3'-ends of SSU RNAs should therefore be very useful in phylogenetic comparisons among eukaryotes, prokaryotes and organelles. Such analyses should also give information about post-transcriptional modification, which is valuable since an RNA modification conserved between two organisms would imply that the genes encoding proteins involved in rRNA modification have also been conserved. In order to obtain a more diverse collection of eukaryotic cytosol sequences, I have determined the sequence at the 3'-end of the SSU RNA from the protist Crithidia fasciculata, a trypanosomatid protozoan. I have also determined the 3'-terminal sequences of wheat mitochondrial and wheat cytosol SSU RNAs. Since this work was completed, I have sequenced the 3'-ends of Acanthamoeba castellanii cytosol SSU RNA and Tetrahymena pyriformis mitochondrial SSU RNA and these sequences are also included in the

comparisons presented below.

The available 3'-terminal SSU RNA sequences are presented in the alignment shown in Figure 25. Numbering from the 3'-end, all of the sequences can be aligned without any insertions or deletions from E. coli positions 9 to 60. The section of SSU RNA immediately following this region (corresponding to E. coli positions 61-124) is highly variable in both length and primary sequence (323) and has been almost completely deleted in the SSU RNA of Drosophila and mosquito mitochondria (61,154). An alignment of the other animal, fungal and protozoan mitochondrial sequences past this point would be uncertain at best, and therefore useless in phylogenetic comparisons.

All of the eukaryotic sequences are obviously homologous up to E. coli position 85, and an alignment based on complete SSU RNA sequences (139) shows that eukaryotic, eubacterial and archaebacterial sequences can all be aligned up to this point without any additions or deletions, except for the Shine-Dalgarno sequence. Therefore, the tables of homology (Tables 3 and 4) comparing eubacterial, eukaryotic, archaebacterial, chloroplast, and plant mitochondrial SSU RNA are based on this 77 nucleotide (E. coli positions 9-85) sequence alignment. Comparisons which include animal, fungal and protozoan mitochondrial sequences (Table 5) are necessarily based on the shorter, 52 nucleotide (E. coli positions 9-60) alignment (Figure 25).

The Neurospora crassa 3'-terminal 18S rRNA sequence was determined by Kelly and Cox (168), using the chemical sequencing method of Peattie (264). The sequence they report has an extra G

Figure 25 Alignment of 3'-terminal SSU RNA sequences. To facilitate comparison, the sequences have been divided into 6 sections [(i)-(vi)], as defined in Figure 8. N, eukaryotic nuclear-encoded, E, eubacterial; A, archaebacterial; C, plastid, M, mitochondrial. The sequences are from Crithidia fasciculata, CRI, Dictyostelium discoideum, DIC; Acanthamoeba castellani, ACA, Triticum aestivum (wheat), WHT, Neurospora crassa, NCR, Saccharomyces cerevisiae, SCE, rat, rabbit and mouse, MAM, Xenopus laevis and Xenopus borealis, XEN, Drosophila melanogaster, DRO, Bombyx mori, BMO, Halobacterium volcanii, HVO, Escherichia coli, ECO; Proteus vulgaris, PVU, Bacillus subtilis, BSU, Bacillus brevis, BBR; Bacillus stearothermophilus, BST; Anacystis nidulans, ANI, Euglena gracilis, EGR; Chlamydomonas reinhardii, CRE, Zea mays, MAI; mosquito, MOS, hamster, HAM, human, HUM, bovine, BOV, rat, RAT, mouse, MOU; Asperigillus nidulans, ALS; Tetrahymena pyriformis, TPY.

between positions 7 and 10 relative to other eukaryotic SSU RNA sequences. The presence of an "artifactual" extra G at this position on chemical sequencing gels of E. coli 16S rRNA (350) and other eukaryotic 18S rRNAs (11) has been reported and I have assumed that this extra G in the Neurospora sequence is also an artifact.

The Bombyx mori 3'-terminal 18S rRNA sequence (287) was determined at the DNA level and has a one nucleotide deletion (position 22) relative to all other SSU RNAs. Due to the highly conserved nature of the 3'-terminus of SSU RNA, this deletion is most likely a sequencing error, and so the presence of the U residue characteristic of eukaryotic SSU RNAs has been assumed for the sequence alignment of Figure 25. Van Charldorp and Van Knippenberg (350) also assume the presence of a U in this position of B. mori 18S rRNA.

The 3'-terminal sequence of rat 18S rRNA, determined at the DNA level by Subrahmanyam et al. (325), has two deletions compared to the same sequence reported by Torzynski et al. (338,339) and Chan et al. (53). In their sequence alignment, Subrahmanyam et al. (325) mistakenly leave out the same two nucleotides in all of the other sequences. I have chosen to use the sequence reported by Torczynski et al. (338,339) and Chan et al. (53), which is identical to the mouse sequence (117,233).

The 3'-terminal sequence of rabbit 18S rRNA reported by Lockard et al. (205) differs from that reported by Darzynkiewicz et al. (75) in one position. The nucleotide in question (position 37) is in a

region that is highly conserved in eukaryotes and prokaryotes and therefore the sequence of Darzynkiewicz et al. (75), which has a C instead of a G in the position, is most likely correct.

Azad and Deacon (11) have reported the 3'-terminal sequences of rat, mouse, rabbit, chicken and barley 18S rRNA. Their sequences of rat, mouse, and rabbit 18S rRNA, based on chemical sequencing of 3'-end-labelled RNA, do not agree with the sequences shown in Figure 25. Most of the errors probably lie in their interpretation of the data, since the correct sequence can be read from the sequencing gels shown in their paper (11).

The DNA sequence corresponding to the 3'-end of Anacystis nidulans 16S rRNA reported by Williamson and Doolittle (373) does not agree with that reported by Tomioka and Sugiura (337). Williamson (89) has since revised the original sequence, so that it now agrees with that of Tomioka and Sugiura (337).

The 3'-terminal sequence of the yeast mitochondrial 15S rRNA gene reported by Sor and Fukuhara (310) differs from that of Li and Tzagaloff (204). The sequence reported by Li and Tzagaloff has been used in Figure 25 since they have confirmed their DNA sequence by chemical sequence analysis of 3'-end-labelled rRNA (204).

2. Structure and Evolution of the 3'-Terminus of Eukaryotic SSU RNA

Table 3 shows the extent of sequence identity between the 3'-terminal sequences of Crithidia and wheat 18S rRNAs and the 3'-terminal sequences of other eukaryotic 18S rRNAs, as well as

A i-vi

	DIC	ACA	WHT	NCR	SCV	MAM	XEN	DRO	BMO	EUB	HVO
CRI	74	71.4	74	72.7	70.1	72.7	72.7	79.2	76.6	60	62.3
DIC		85.7	87	81.8	77.9	79.2	80.5	83.1	83.1	59.7	61
ACA			94.8	76.6	79.2	85.7	85.7	85.7	85.7	58 2	61
WHT				77.9	81.8	85.7	85.7	87	88.3	59.5	61
NCR					84.4	80 5	81.8	87	88.3	57.4	62 3
SCE						85.7	85 7	89.6	92.2	57.1	61
MAM							97.4	92.2	89.6	60	61
XEN								92.2	92 2	57.9	59.7
DRO									97	58.7	59.7
BMO										57.1	59.7
EUB										82.5	65.2

B i-v

	DIC	ACA	WHT	NCR	SCE	MAM	XEN	DRO	BMO	EUB	HVO
CRI	90.7	86	86	86	86	86	86	86	86	83.3	90.7
DIC		95.3	95.3	95.3	95.3	95.3	95.3	95.3	95.3	83.6	90.7
ACA			95.3	90.7	95.3	95.3	95.3	95.3	95.3	81.4	86
WHT				90.7	100	100	100	100	100	81.4	86
NCR					90.7	90.7	90.7	90.7	90.7	81.4	86
SCE						100	100	100	100	81.4	86
MAM							100	100	100	81.4	86
XEN								100	100	81.4	86
DRO									100	81.4	86
BMO										81.4	86
EUB										91.4	89.8

C vi

	DIC	ACA	WHT	NCR	SCE	MAM	XEN	DRO	BMO	EUB	HVO
CRI	52.9	59.2	58.8	55.9	50	55.9	55.9	70.6	64.7	30.6	32.2
DIC		73.5	76.5	64.7	55.9	58.8	61.8	67.6	67.6	27.1	23.5
ACA			94.1	58.8	58.8	73.5	73.5	73.5	73.5	28.8	29.4
WHT				61.8	58.8	67.6	67.6	70.6	73.5	31.8	29.4
NCR					76.5	67.6	70.6	82.3	85.3	27.1	32.3
SCE						67.6	67.6	76.5	82.3	26.5	29.4
MAM							94.1	82.3	76.5	32.9	29.4
XEN								82.3	82.3	34.1	26.5
DRO									94.1	30	26.5
BMO										26.5	26.5
EUB										71.2	21.8

Table 3. Percentage identity between the 3'-terminal sequences of eukaryotic SSU RNAs. Calculations are based on. (A) 77 positions, E. coli positions 9-85, SSU RNA sections (i-vi), (B) 43 positions, E. coli positions 9-51, SSU RNA sections (i-v), (c) 34 positions, E. coli positions 52-85, SSU RNA section (vi). EUB, average % identity with the eubacterial sequences listed in Figure 25. See Figure 25 for other abbreviations.

prokaryotic 16S rRNAs. To facilitate this comparison, each SSU RNA sequence has been divided into six sections, as defined in Figure 8. Crithidia and wheat 18S rRNAs are typically eukaryotic in that they specifically lack the polypyrimidine sequence (CCUCC in E. coli 16S rRNA; positions 4-8) which in eubacteria is postulated to interact with the 5'-leader sequence in mRNA (302,303). In addition, both wheat and Crithidia 18S rRNAs have a U at the same position (22 in wheat, 23 in Crithidia) as in all other eukaryotic 18S rRNAs so far sequenced. Prokaryotic 16S rRNAs have a G at the analogous position. However, Crithidia 18S rRNA differs from wheat 18S rRNA at 6 of 20 positions that constitute the m_2^6A stem. It is striking that the three substitutions in section (ii) of the Crithidia sequence are accompanied by reciprocal changes in section (iv) so that base pairing (and therefore overall secondary structure) is strictly conserved throughout the m_2^6A stem (Figure 8) in spite of the divergence in primary sequence. It is interesting that reciprocal changes in the m_2^6A stem also occur in several of the other SSU RNA sequences presented in Figure 25.

The immediate 3'-terminal sequence of eukaryotic SSU RNA is ...UUG_{OH} in wheat, barley (11), and Euglena (317), ...$UUUU_{OH}$ in Crithidia and Dictyostelium (230), and ...UUA_{OH} in all other eukaryotes so far examined, as well as in E. coli 16S rRNA (Figure 25). A Ψ residue has been tentatively identified at position 2 in Acanthamoeba 18S rRNA (M.N. Schnare, unpublished) and at position 3 in Crithidia 18S rRNA (293; see RESULTS, section 4). Euglena

18S rRNA is known to have a Ψ at the same position (317).

It is evident from Figure 25 and Table 3 that the primary sequence of sections (i)-(v) at the 3'-end of SSU RNA is highly conserved, not only among eukaryotes but between eukaryotes and prokaryotes as well. Particularly notable is the absolute conservation of primary sequence of section (v) in all of the eukaryotic and prokaryotic SSU RNAs, which according to secondary structure models (90,323) is a single-stranded region connecting two base-paired helices in SSU RNA. This suggests that in this region, primary structure per se participates in and is crucial for the same specific function(s) in both eukaryotic and prokaryotic SSU RNAs. However, primary structure in this region is not absolutely conserved in the mitochondrial SSU RNAs of animals, fungi and protozoa (Figure 25).

Figure 25 presents clear evidence for retention of a eukaryotic-like pattern of sequence in section (vi) of Crithidia 18S rRNA, although quantitatively the data in Table 3 (whether based on the entire sequence (A) or on section (vi) alone (C)) suggest that Crithidia 18S rRNA is not as closely related to other 18S rRNAs as these are to each other. The data presented in Table 3 suggest that Crithidia represents an early branching from the eukaryotic line of descent. The high degree of divergence of the Crithidia sequence relative to those of other eukaryotes cannot be explained by a "fast evolutionary clock" (230) in Crithidia, since the Crithidia sequence has diverged no further away from the prokaryotic sequences than have

those of other eukaryotes (Table 3).

McCarroll et al. (230) have constructed a phylogenetic tree based on SSU RNA sequences and have concluded that the point of divergence of Dictyostelium from other eukaryotes was shortly after the point at which the mitochondrial line diverged from the rest of the eubacteria. They further concluded that "eukaryotes that represent substantially deeper branches than D. discoidium will not have mitochondria as we know them". With this in mind, it is interesting that Crithidia, which appears from these data (Table 3) to branch earlier than D. discoideum, contains a kinetoplast instead of typical mitochondria (359).

As noted earlier, Crithidia 18S rRNA seems to be unusually large: in polyacrylamide gels, for example, it migrates with an apparent molecular weight of 0.83-0.84 x 10^6 daltons (109,122,237,315), compared with a typical value of 0.7 x 10^6 daltons for other eukaryotic 18S rRNAs (8). The data presented here along with data for the 5'-terminal sequence of Crithidia 18S rRNA (M.N. Schnare, unpublished) indicate that this apparent size difference cannot be accounted for by extensions at either the 3'- or 5'-end of Crithidia 18S rRNA.

The SSU RNA of other kinetoplastidae (Leishmania, Trypanosoma) also appears to be atypically large (236), as does the SSU RNA of several other protists (149,207,321). On the other hand, the SSU RNA of the ciliated protozoa (Paramecium, Tetrahymena) has the typical eukaryotic size of 0.7 x 10^6 daltons (8,207). In the context of

assessing the evolution of multicellular higher organisms from primitive unicellular organisms, it is intriguing that the 3'-terminal sequences of wheat and Acanthamoeba 18S rRNA share such a high degree of homology (Figure 25; Table 3).

3. Functions of the 3'-Terminus of Eukaryotic SSU RNA

Knowledge of the 3'-terminal sequence of Crithidia 18S rRNA allows us to ask whether this molecule could participate in specific base-paired interactions that have been postulated to involve the 3'-end of other eukaryotic 18S rRNAs. Like other 18S rRNAs, Crithidia 18S rRNA lacks the prokaryotic polypyrimidine sequence that has been implicated in mRNA binding in bacteria (302,303,318). Hagenbuchle et al. (141) have suggested a different, although analogous, interaction between a conserved purine-rich sequence (GCGGAAGG) near the 3'-terminus of eukaryotic 18S rRNA and a complementary sequence within the 5'-leader region of eukaryotic mRNA. However, this suggestion has been criticized (18,81), and Kozak (184-187) has proposed an alternative "scanning" model, in which the 5'-terminus per se, rather than a specific sequence of nucleotides in the 5'-leader region, is the primary determinant for binding eukaryotic mRNAs to small ribosomal subunits.

In the position of the putative mRNA binding site (141) in 18S rRNA, Crithidia 18S rRNA has the sequence GCAGCUGG (residues 9 to 16): a difference of 3 out of 8 residues. The more recently determined sequences of Acanthamoeba (M.N. Schnare, unpublished) and

Dictyostelium (230) 18S rRNA each have a single base change, while Neurospora 18S rRNA (168) has two base changes in this conserved region. Divergence to this extent would not be expected if this sequence had a universal function in binding eukaryotic mRNAs. In fact, the base changes in the Crithidia sequence either entirely preclude pairing between it and postulated 18S rRNA binding sites in eukaryotic mRNAs (18,141,188) or render such pairings much less likely, since the hybrid structures which can be written have substantially reduced thermodynamic stability. One might argue that the 18S rRNA binding sites within the 5'-leader region of Crithidia mRNA have undergone compensating base changes that allow an analogous base-paired interaction to occur. However, this seems unlikely, since Crithidia mRNA is efficiently translated in a wheat germ cell-free system (280). Thus, the mRNA-18S rRNA binding model of Hagenbuchle et al. (141) does not appear to be applicable to Crithidia. More recently, Sargan et al. (292) have proposed a base-paired interaction between two non-contiguous sequences at the base of the m_2^6A hairpin of eukaryotic 18S rRNA and the 5'-leader sequences of mRNA. This sequence (GGU...GGAU) is completely conserved in all of the available eukaryotic SSU RNA sequences (Figure 25, residues 7-10 and 33-35 in Crithidia 18S rRNA).

It has also been postulated that interaction between 5S rRNA and the 3'-end of eukaryotic 18S rRNA may play a role in the association of large and small ribosomal subunits during initiation of protein synthesis (9,13). However, it should be noted that removal of 5S rRNA

from rabbit reticulocyte ribosomes does not induce subunit dissociation (168). Formation of a specific complex *in vitro* has been demonstrated between wheat 18S and 5S rRNAs (13,14), and the 18S rRNA binding site in wheat 5S rRNA has been identified and sequenced (246). In *Neurospora* (168) an interaction has been detected between 5S rRNA and a 3'-terminal fragment from 18S rRNA and between 18S rRNA and a specific 3'-terminal fragment from 5S rRNA. The base-paired structure that can be written (Figure 26) between the known binding site in 5S rRNA and the postulated one in 18S rRNA is thermodynamically very stable in the case of wheat, in agreement with the observed thermal stability of the complex formed *in vitro* (14). Azad and Deacon (11) have shown that the proposed 5S rRNA binding site is totally conserved in the 18S rRNA of various vertebrate animals and of barley embryos, as would be expected if this sequence is generally important for interaction with a conserved sequence in 5S rRNA.

Since the sequence of *Crithidia* 5S rRNA has been determined (217), we examined the relevant regions of *Crithidia* 5S and 18S rRNAs to ascertain whether a stable complex, analogous to that between wheat 18S and 5S rRNAs, could be formed. As shown in Figure 26, base changes occur in both the 18S rRNA and 5S rRNA of *Crithidia* in the regions homologous to the proposed sites of interaction in wheat 18S and 5S rRNAs. These sequence changes are non-reciprocal in the sense that a change in the 18S sequence is not accompanied by a corresponding change in the 5S sequence that would allow retention of base-pairing, and the same is true for the changes in the 5S rRNA

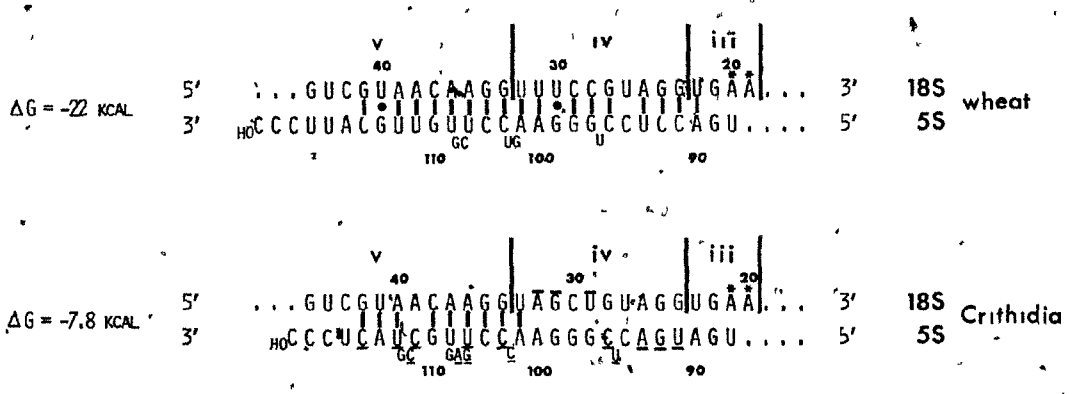


Figure 26. Potential base-paired complexes between 18S and 5S rRNAs from wheat and Crithidia. The 18S rRNA binding site in wheat 5S rRNA has been identified by sequence analysis (246), whereas the 5S rRNA binding site in wheat 18S rRNA is inferred (9,11). The wheat 5S rRNA sequence shown is the corrected version determined in our laboratory (218), so that the wheat 5S-18S complex as written is slightly, but not significantly, different from that postulated in (9,11). The Crithidia 5S-18S complex has been constructed using the strictly-homologous sections of Crithidia 5S rRNA (217) and Crithidia 18S rRNA (see Figure 25). Residues that differ between Crithidia and wheat are overlined (18S rRNA) and underlined (5S rRNA) in the Crithidia complex.

sequence. As a result, the possibilities for base-pairing between Crithidia 18S and 5S rRNAs are greatly reduced in this region, and the complex which can be formed is much less stable thermodynamically than the wheat complex. In addition, base pairing cannot extend into region (iv) (Figure 8) of Crithidia 18S rRNA, so that disruption of the m_2^6A stem, an important feature of the Azad model (9,11), is not possible. In examining the sequence of Crithidia 5S rRNA, we have noted another region (residues 57-74) that could potentially form a stable duplex ($\Delta G = -17.6$ kcal; ref. 333) with the 3'-end of Crithidia 18S rRNA (residues 22-41), effectively disrupting the m_2^6A stem in the process. However, since we have not been able to observe a complex between Crithidia 5S and 18S rRNAs in vitro (see RESULTS, section 3), under conditions (14) in which a control complex between wheat 5S and 18S rRNAs readily forms, the significance of this alternative potential interaction is unclear.

In summary, the sequence obtained for the 3'-end of Crithidia 18S rRNA does not lend support to functional models that involve complementary base-pairing between eukaryotic 18S rRNA and either mRNA (141) or 5S rRNA (9). However, it should be remembered that the ribosome of Crithidia has some unusual structural characteristics (see INTRODUCTION, section 3.4 and RESULTS, section 2), including the fact that it possesses four novel, low molecular weight rRNAs (in addition to 5S and 5.8S rRNAs) (122,296), all localized in the large ribosomal subunit (123). This raises the possibility that functional interactions between Crithidia 18S rRNA and other RNAs might differ in

certain respects from those postulated for other eukaryotes. It is conceivable, for example, that one or more of the additional small rRNA species present in Crithidia large subunits has assumed the proposed role of 5S rRNA in binding to 18S rRNA in small subunits, although so far we have not been successful in demonstrating complex formation between any of these and the 18S rRNA (see RESULTS, section 3).

4. The Evolutionary Origin of Wheat Mitochondria

4.1. Evidence Supporting a Eubacterial Origin

(a) Evidence Prior to This Work

The use of T_1 oligonucleotide fingerprinting techniques allowed Cunningham et al. (70) to conclude that the 26S, 18S and 5S rRNA species isolated from wheat mitochondria are distinct from their cytosol counterparts. These rRNAs were later shown by Bonen and Gray (28) to hybridize to specific fragments (generated by restriction endonuclease digestion) of wheat mitochondrial DNA under conditions in which the cytosol rRNAs did not hybridize.

The T_1 oligonucleotide catalogue of wheat mitochondrial 18S rRNA strongly resembles T_1 oligonucleotide catalogues of eubacterial and chloroplast 16S rRNAs, but lacks detectable homology with the T_1 oligonucleotide catalogue of wheat cytosol 18S rRNA (27). In addition, among those T_1 oligonucleotides which the mitochondrial 18S rRNA shares with eubacterial and chloroplast 16S rRNAs, there is an especially high proportion of ones identified (378) as conserved in

the evolution of eubacterial 16S but not eukaryotic 18S rRNA species (72). Bonen et al. (27,72) therefore concluded that the mitochondrial 18S rRNA is eubacterial in nature, a finding that supports an endosymbiotic (77,227) origin of mitochondria. However, Uzzell and Spolsky (346) have interpreted these same data as being consistent with either an endosymbiotic or an autogenous (271,345) origin of mitochondria, a view which is also held by Mahler (221).

(b) Evidence Presented in This Thesis

Extensive primary sequence identity between E. coli and wheat mitochondrial rRNAs is indicated by the fact that the bacterial 16S and 23S rRNAs hybridize specifically to those restriction fragments containing wheat mitochondrial 18S and 26S rRNA genes, respectively. Although attempts to hybridize E. coli rRNAs to the mitochondrial DNA of other eukaryotes have not been reported, it is notable that no hybridization was observed between spinach chloroplast rRNAs and Acanthamoeba castellanii mitochondrial DNA under conditions in which the former hybridized extensively to E. coli rDNA (25).

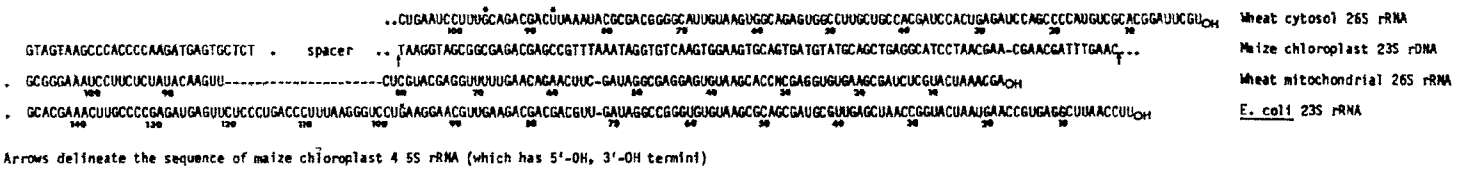
There is a substantially greater degree of primary sequence identity between the 3'-terminal sequences of wheat mitochondrial 18S rRNA and E. coli 16S rRNA (82%) than between either wheat mitochondrial and cytosol 18S rRNAs (58%) or wheat cytosol 18S and E. coli 16S rRNAs (61%) (see Figure 25, Tables 3 and 4). In fact the 3'-terminal sequence of E. coli 16S rRNA is more closely related to the wheat mitochondrial sequence than it is to other known eubacterial and chloroplast 16S rRNA sequences (Table 4).

	BSU (E)	BBR (E)	BST (E)	ANI (E)	EGR (C)	CRE (C)	MAI (C)	WHT (M)	EUK (N)	HVO (A)	WHT (N)
ECO (E)	71.4	72.7	70.1	72.7	70.1	59.7	70.1	81.8	61.7	70.1	61.3
BSU (E)		92.2	93.5	85.7	75.3	68.8	75.3	70.1	57.1	63.6	58.4
BBR (E)			92.2	87	76.6	70.1	76.6	70.1	59.9	62.3	61.3
BST (E)				87	79.2	75.3	80.5	72.7	57.3	66.2	58.4
ANI (E)					84.4	81.8	84.4	75.3	56.9	63.6	58.4
EGR (C)						76.6	87	70.1	53	62.3	53.2
CRE (C)							79.2	62.3	49	59.7	49.3
MAI (C)								72.7	56.8	67.5	58.4
WHT (M)									57.9	66.2	58.4
EUK (N)									83.5	60.9	84.7
HVO (A)											61.3

Table 4. Percentage identity between the 3'-terminal sequences of eubacterial and organellar SSU RNAs. Calculations are based on 77 positions, E. coli positions 9-85. EUK, average % identity with the eukaryotic nuclear-encoded sequences listed in Figure 25. See Figure 25 for other abbreviations.

At a position occupied by m^3U in E. coli 16S rRNA, the same modified nucleoside is present in wheat mitochondrial 18S rRNA but not in wheat cytosol 18S rRNA (see RESULTS, sections 5). Interestingly, m^3U is present in a "universal" T_1 oligonucleotide (U*AACAAGp of (378).) in eubacterial 16S rRNA. Two other 16S universals, CCm^7GCGp and $m^4CmCCGp$, also appear in the wheat mitochondrial 18S rRNA T_1 catalogue (27,72). Therefore, in both primary sequence and post-transcriptional modification pattern, wheat mitochondrial 18S rRNA is obviously much more closely related to its eubacterial homologue than it is to its counterpart in the cytoplasmic compartment of the same cell.

Based on the alignment presented in Figure 27, the 3'-terminal sequence of wheat mitochondrial 26S rRNA (positions 1-80) shows a clear relationship to the E. coli 23S rRNA sequence (58% primary sequence identity) and to that of maize chloroplast 4.5S rRNA (48% identity), while it is only distantly related to the wheat cytosol 26S rRNA sequence (29%). Note the three stretches of identity between the mitochondrial and bacterial sequences at mitochondrial positions 5-11, 35-42, and 48-53. Using strictly homologous positions, the wheat mitochondrial and chloroplast sequences can be folded to fit the secondary structure model proposed for this region of the bacterial 23S rRNA (214) (see Figure 28). Even though there is considerable primary sequence divergence, secondary structure has been conserved through compensating base changes. The wheat cytosol sequence cannot be fitted to this secondary structure model. These observations,



Arrows delineate the sequence of maize chloroplast 4 5S rRNA (which has 5'-OH, 3'-OH termini)

Figure 27. Alignment of the 3'-terminal sequences of wheat cytosol and mitochondrial 26S rRNAs, E. coli 23S rRNA (39) and the corresponding region of maize chloroplast rDNA (97).

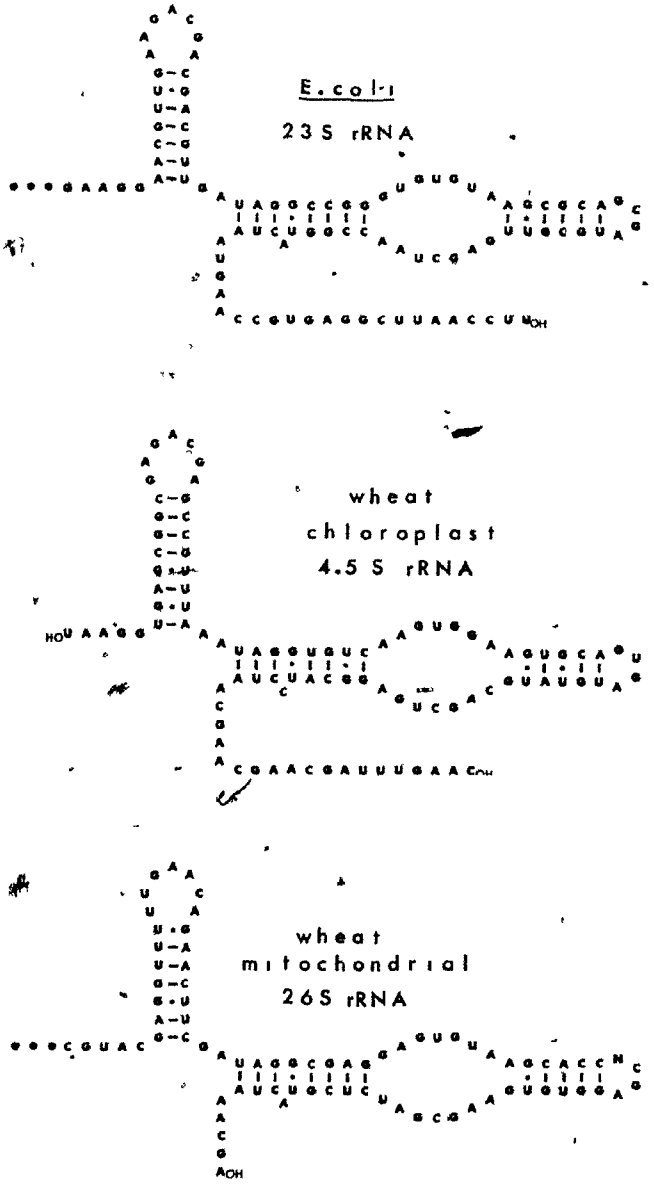


Figure 28. Potential secondary structure at the 3'-terminus of wheat mitochondrial 26S rRNA, E. coli 23S rRNA (214) and wheat chloroplast 4.5S rRNA (370).

along with the results of heterologous hybridization experiments (see above) suggest an endosymbiotic, eubacterial origin of wheat mitochondrial 26S rRNA genes.

(c) Recent Evidence

Restriction fragments of wheat mitochondrial DNA, which contain the genes encoding 18S and 26S rRNA, have been cloned in pBR322 by F. Quetier and B. Lejeune (Orsay, France). David Spencer in our laboratory has now determined the complete sequences of both of these rRNA genes (314, and unpublished). Remarkable similarities in primary sequence and potential secondary structure between the wheat mitochondrial 18S/26S rRNA and *E. coli* 16S/23S rRNA provide compelling evidence in favor of an evolutionary origin of wheat mitochondria from eubacteria-like endosymbionts. The primary sequence of the wheat mitochondrial $tRNA_f^{Met}$ gene (132) also displays a high degree of sequence similarity to its eubacterial counterpart.

4.2. Non-Eubacterial Characteristics of Wheat Mitochondrial rRNA Genes

(a) Ribosomal RNA Gene Organization

The genes for wheat mitochondrial 26S and 18S rRNAs are far apart on the mitochondrial genome, while the genes for 18S and 5S rRNA are closely linked (28). Experiments presented here (295; RESULTS, section 7) establish that the 5S rRNA gene is close to the 3'-end of the 18S rRNA gene. These observations are in contrast to the closely linked 5'-16S-23S-5S-3' rRNA gene arrangement found in eubacteria (see INTRODUCTION, section 3.1). However, this rRNA gene organization is

equally non-eukaryotic and therefore supports neither the endosymbiotic nor autogenous theories of mitochondrial origin.

(b) Wheat Mitochondrial 5S rRNA

Spencer et al. (313) have determined the complete primary sequence of wheat mitochondrial 5S rRNA. In view of the striking primary sequence similarity detected between the wheat mitochondrial large rRNAs and their eubacterial counterparts, it is rather surprising that the 5S rRNA sequence is neither obviously prokaryotic nor eukaryotic in nature, but shows characteristics of both classes of 5S rRNA as well as some unique features (131,313).

(c) Initiation of Translation

Although the 3'-terminal sequences of all SSU RNAs are highly similar (Figure 25, SSU RNA sections (i)-(v)), there appears to be an insertion of a five nucleotide sequence (CCUCC) into eubacterial 16S rRNAs relative to eukaryotic 18S rRNAs (or a deletion in the eukaryotic 18S rRNAs). This "Shine-Dalgarno" sequence, which functions during intiation of protein synthesis in eubacteria (302,303), is also present in archaebacterial (139,161,167), Chlamydomonas chloroplast (91), and plant chloroplast (170,298,335) 16S rRNAs. Euglena chloroplast 16S rRNA has a single nucleotide difference in this region (119,254,317). There is no obvious "Shine-Dalgarno" sequence in animal mitochondrial SSU RNA, and at the position in fungal and protozoan mitochondrial SSU RNAs (Figure 25) corresponding to the position of the "Shine-Dalgarno" sequence in E. coli 16S rRNA, primary sequence has not been evolutionarily conserved.

If one compares the immediate 3'-terminal sequence of E. coli 16S rRNA (...CCUCCUUA$_{OH}$) with the immediate 3'-terminal sequence (RESULTS, section 5) and flanking region (D.F. Spencer, personal communication) of a wheat mitochondrial 18S rRNA gene (...AAUCCUUC...), it is obvious that the "Shine-Dalgarno" sequence has not been deleted in the wheat mitochondrial 18S rRNA gene. However, the fact that primary sequence has not been completely conserved (note the two C to A substitutions), and that length heterogeneity results in the loss of the 3'-terminal C of the "Shine-Dalgarno" sequence in 20% of the mitochondrial 18S rRNA molecules (RESULTS, section 5), suggests that this sequence is not involved in initiation of protein synthesis in wheat mitochondria. It is important to note that there is as yet no evidence for or against a eukaryotic type of "scanning" mechanism (see DISCUSSION, section 3) of initiation of protein synthesis for plant mitochondrial ribosomes.

(d) Ribosomal RNA Processing

Wheat mitochondrial 26S rRNA lacks the sequences corresponding to the 3'-terminal 17 nucleotides (Figures 14 and 28) and 5'-terminal 11 nucleotides (D.F. Spencer, M.N. Schnare and M.W. Gray, unpublished) of E. coli 23S rRNA. It has been proposed that these regions of E. coli 23S rRNA are base-paired with each other as part of a double-stranded structure necessary for processing of E. coli rRNA precursors (36, 385). The absence of these sequences at the ends of wheat mitochondrial 26S rRNA implies that an alternate rRNA processing mechanism operates in wheat mitochondria. Processing in animal mitochondria

involves precise excision of tRNA sequences found immediately adjacent to the rRNA sequences in the precursor (98,352). In this regard it is interesting that a tRNA gene is situated immediately adjacent to the 5'-terminus of the wheat mitochondrial 18S rRNA gene (132). However, sequence analysis of regions of wheat mitochondrial DNA that flank the 26S rRNA gene have not revealed any tRNA genes or other obvious processing signals (D.F. Spencer, personal communication).

The mechanism of rRNA processing in eukaryotes, although obviously different from that found in E. coli, is not clearly understood (34). We are therefore uncertain as to the relationship between wheat mitochondrial and eukaryotic mechanisms of rRNA processing.

4.3 Conclusions About Mitochondrial Evolution

In addressing the question, "Where did mitochondria come from?", there are essentially two possibilities. Mitochondria either arose autogenously within a single sort of cell or as a result of an endosymbiosis between two different sorts of cells (see INTRODUCTION, section 4.2).

Even though the wheat mitochondrial genome has some non-eubacterial characteristics (DISCUSSION, section 4.2), it is important to point out that these data do not argue for or against either an endosymbiotic or autogenous origin of wheat mitochondria. These characteristics, which are not obviously prokaryotic or eukaryotic, can be taken as evidence that the mitochondrion of wheat has had a significant period of independent evolution, whether its origin was endosymbiotic or autogenous. In fact, there is evidence, presented

and discussed here (DISCUSSION, section 4.1), that strongly favors an endosymbiotic, specifically eubacterial origin of wheat mitochondria.

All theories concerning mitochondrial evolution must concede that organelle, eukaryotic and eubacterial genomes shared a single common ancestor at some point in evolution. With this in mind, even with the knowledge that wheat mitochondrial 18S and 26S rRNAs are obviously eubacterial in nature, it is still possible to argue against an endosymbiotic origin. The argument is as follows: (i) Mitochondria arose autogenously within the "protoeukaryotes" early in evolution, while these ancestral eukaryotes were still very similar to eubacteria. (ii) There was a slow rate of mutation fixation during the evolution of eubacteria and wheat mitochondria, which would account for the high degree of rRNA sequence similarity evident between the two. Thus, the similarities between present day mitochondria and eubacteria are a reflection of "retention of primitive traits" that were present in the last common ancestor of mitochondria, eukaryotes and eubacteria. The differences between mitochondria and eubacteria represent "derived traits" that occurred during the separate evolution of the two genomes. (iii) After the event(s) that culminated in the autogenous origin of mitochondria, the eukaryotic nuclear genome underwent rapid evolution away from the common ancestor of mitochondria, eukaryotes and eubacteria. The nuclear genome retained fewer "primitive traits", which accounts for the low level of similarity observed between wheat cytosol and either the wheat mitochondrial or eubacterial rRNAs.

If the above scenario is correct, we can make the following predictions: (1) It follows from the assumption of a fast rate of mutation fixation in eukaryotes relative to eubacteria that the degree of rRNA sequence conservation among eukaryotes should be much lower than among eubacteria. (2) If wheat mitochondria had an autogenous origin, its rRNAs should be related to eubacterial rRNAs, but not specifically to the rRNAs of a particular subgroup of the eubacterial lineage.

The 3'-terminal SSU RNA sequences shown in Figure 25 represent a broad range of both eukaryotes and eubacteria and therefore constitute a good data set for testing the above predictions. The data presented in Tables 3 and 4 show that the level of rRNA sequence conservation within the eukaryotes (average eukaryote vs. eukaryote = 83.5%, lowest = 70.1% identity) is as great as within the eubacteria (average eubacterium vs. eubacterium = 82.5%, lowest = 70.1% identity) which is contrary to prediction (1), above. The 3'-terminal sequence of wheat mitochondrial 18S rRNA appears to be more closely related to the 16S rRNA of E. coli (a representative gram-negative bacterium) than to Anacystis nidulans (cyanobacterium) or Bacillus (gram-positive) 16S rRNAs. This finding is contrary to prediction (2), above. We are therefore left with no alternative but to conclude that wheat mitochondria is of an endosymbiotic, specifically eubacterial, evolutionary origin.

No firm support for an endosymbiotic, eubacterial origin of mitochondria in general has emerged from primary sequence data for

mitochondrial rRNAs from animals (7,15,61,98,154,177,284,352), fungi (179,180,204,310,311), and protozoa (300). In these cases, even though the mitochondrial rRNAs may appear from sequence comparisons to be somewhat more closely related to their eubacterial than to their eukaryotic cytosol homologues (see Table 5, ref. 125,130,180,193,230, 314), the degree of identity is no greater (and usually less) than between homologous eubacterial and eukaryotic rRNAs.

Nevertheless, there are striking structural and biochemical similarities in diverse mitochondria, and mitochondrial DNA encodes basically the same set of respiratory chain polypeptides in all eukaryotes (124). Therefore, in view of the essentially proven endosymbiotic, eubacterial evolutionary origin of wheat mitochondria (see above), and presumably other higher plant mitochondria (55,156), it would seem likely that animal, fungal and protozoan mitochondria have also had an endosymbiotic, eubacterial evolutionary origin. The very high rate of mutation fixation in mammalian mitochondrial DNA (42,235) provides an explanation as to why their rRNA genes show a relatively low degree of sequence similarity to their eubacterial counterparts.

5. Low Molecular Weight rRNAs and Their Analogues

5.1. The 3'-Terminus of LSU RNA

Comparisons of the sequences of plant chloroplast and eubacterial rRNA genes have strongly suggested that the chloroplast counterpart of the 3'-end of *E. coli* 23S rRNA is a discrete 4.5S rRNA (96,215), whose

% Identity

	WHT N	ECO E	HVO, A
MAI C	71.2	86.5	82.7
WHT M	75	90.4	84,6
MOS M	50	51.9	55.8
DRO M	53.8	53.8	61.5
HAM M	61.5	71.2	69.2
HUM M	59.6	69.2	65.4
BOV M	59.6	71.2	67.3
RAT M , MOU M	61.5	71.2	67.3
SCE M	71.2	73.1	61.5
ALS M	50	57.7	55.8
TPY M	61.5	71.2	65.4
WHT N		80.8	79
ECO E			86.5

Table 5. 3'-Terminal sequences of mitochondrial SSU RNA compared to the 3'-terminal sequences of wheat cytosol 18S rRNA, E. coli 16S rRNA and H. volcanii 16S rRNA. Calculations are based on E. coli positions 9-60. See also Figure 25.

gene is separated from the rest of the 23S rRNA gene by a spacer ranging from about 78 nucleotides in maize (97) to 101 nucleotides in tobacco (330).

The 3'-terminus of animal (7,98,284,352), fungal (180,311) and protozoan (300) mitochondrial LSU RNA does not seem to share any obvious sequence similarity with the 3'-end of bacterial 23S rRNA or with plant chloroplast 4.5S rRNA. However, wheat mitochondrial 26S rRNA does contain a 4.5S rRNA homologue as it 3'-terminus (DISCUSSION, section 4.1, Figures 27 and 28). We therefore examined the sequence of this region of wheat mitochondrial 26S rRNA in the hope that it would shed light on the evolutionary origin of the separate 4.5S rRNA in plant chloroplasts.

The region of E. coli 23S rRNA corresponding to the position of the transcribed spacer separating the 23S and 4.5S rRNA genes in plant chloroplast DNA has been shown by sequence analysis (36,39,214) to contain five cistron heterogeneities, while the wheat mitochonodrial 26S rRNA has a 21 nucleotide deletion relative to E. coli 23S rRNA in this region (Figure 27). These observations suggest that E. coli positions 100-120 may be a transcribed but unexcised relic of the chloroplast spacer, and that this spacer may have been deleted altogether in the wheat mitochondrial 26S rRNA genes.

Alternatively, since it has been suggested that the chloroplast spacer sequence may be analogous to an intron and that the two RNA molecules (23S and 4.5S) remain unligated due to incomplete splicing (330), it seemed possible that wheat mitochondrial 26S rRNA genes have

retained an intron at this position, which would imply that intron excision and splicing occur during the biosynthesis of wheat mitochondrial 26S rRNA. However, DNA sequence analysis (D.F. Spencer, personal communication) has established that there is no intron in this region of wheat mitochondrial 26S rRNA genes.

5.2. Crithidia fasciculata 5.8S rRNA

Alignment with published 5.8S rRNA sequences (Figure 29) clearly shows that Crithidia species i is a 5.8S rRNA, despite its exceptional length (171-172 nucleotides). Excluding the "G+C rich hairpin" (a feature of secondary structure models of 5.8S rRNA (105,213,243,282)), there is obvious primary sequence identity within the regions encompassing nucleotides 1-114 and 149-172 of the Crithidia sequence. Particularly notable are two moderately long stretches (residues 38-49 and 102-108) that appear to be conserved in almost all 5.8S rRNAs; (344) the former contains the GAAC tetranucleotide that has been proposed to interact with the common GTΨC sequence in tRNA (247,382). The conserved GGAU sequence that may be involved in the interaction of 5.8S rRNA with LSU RNA (242) is also present in Crithidia 5.8S rRNA (positions 14-17).

Figure 30 illustrates the fact that neither length nor primary sequence is very highly conserved in the G+C rich hairpin region, although the potential for a substantial degree of base-pairing is preserved. It is obvious from this figure, and from the sequence alignment shown in Figure 29, that the extra nucleotides in Crithidia 5.8S rRNA are located in this hairpin: its length is 34 nucleotides in

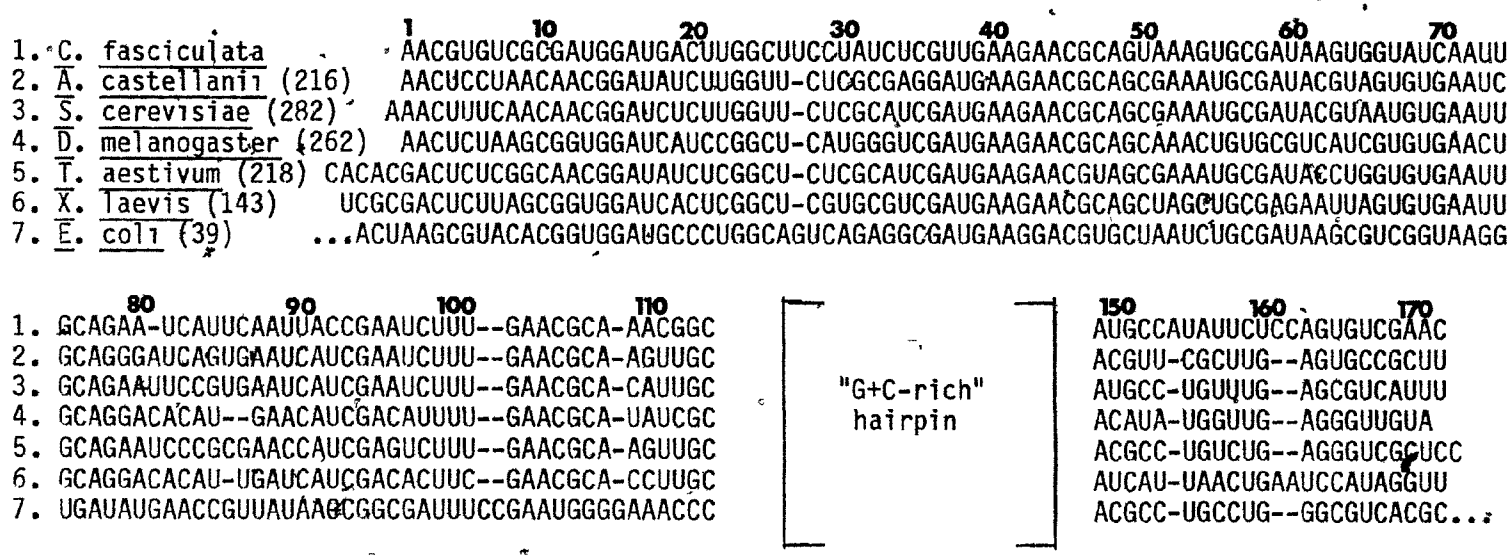


Figure 29. Alignment of representative 5.8S rRNA sequences (references as indicated in the figure), including that determined here for C. fasciculata. Sequences were arranged for maximal homology in pairwise comparisons, with a minimal number of assumed additions and deletions. The 5'-terminal sequence of E. coli 23S rRNA (39) is also included in this alignment. Note that the highly variable region of primary sequence encompassing the "G+C rich hairpin" (see Figure 30) is not included in this alignment.

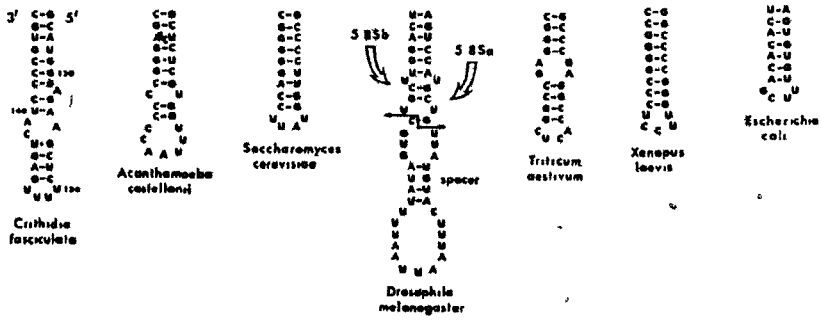


Figure 30. Potential secondary structure in the "G+C-rich hairpin" region of the 5.8S rRNA sequences listed in Figure 29, and in the analogous region of E. coli 23S rRNA. The arrows in the Drosophila 5.8S sequence indicate the cleavages that occur during 5.8S rRNA maturation and which result in the removal of the transcribed spacer region. In the case of wheat, initial uncertainties in the sequence of this region (218) have been resolved (372), and the revised version of the G+C-rich hairpin is shown.

Crithidia but only 22-27 nucleotides in other eukaryotes. In this regard, it is noteworthy that in D. melanogaster (262) and S. coprophila (165), the 5.8S rRNA consists of two fragments (5.8Sa and 5.8Sb) that are joined noncovalently at the "G+C rich" hairpin. This arrangement results from the excision (without subsequent splicing) of a short transcribed spacer that separates the 5.8Sa and 5.8Sb regions in the primary transcript. Pavlakis et al. (262) have suggested that either the spacer sequence was inserted into Drosophila 5.8S rDNA at a time following the divergence of the vertebrates and invertebrates, or that the spacer existed in ancestral 5.8S rRNA genes but was eliminated during evolution from species that do not now contain it. Since the extra nucleotides in Crithidia 5.8S rRNA occur in the same region as the Drosophila and Sciara spacers, it is plausible to suggest that C. fasciculata has also retained a spacer in its 5.8S rRNA genes, but has lost (or not acquired) the mechanism to excise this spacer from the 5.8S primary transcript. It should be noted, however, that the putative Crithidia spacer is not obviously homologous with either the Drosophila or Sciara spacer sequences, being neither as long nor as A-U-rich.

Our alignment of 5.8S rRNAs (Figure 29) also includes the 5'-end of E. coli 23S rRNA, which Nazar (240) has postulated is the structural analogue in prokaryotes of 5.8S rRNA. The alignment confirms that there is extensive similarity between the 5'-end of E. coli 23S rRNA and a range of eukaryotic 5.8S rRNAs, although quantitatively the degree of similarity (58-66 identities in pairwise

comparisons) is significantly less than between any two 5.8S rRNAs (81-111 identities). Interestingly, a stable hairpin loop can be formed with that part of the E. coli 23S rRNA sequence that is analogous to the "G+C rich" hairpin of 5.8S rRNA (Figure 30). However, this hairpin region has been largely deleted in Anacystis nidulans and chloroplast 23S rRNAs (192).

5.3 Novel Small rRNAs from Crithidia fasciculata

Previous work from this laboratory (123) ruled out the possibility that the novel small rRNAs described in this thesis are generated by quantitative and highly specific (albeit artifactual) RNase cleavage of large rRNAs during isolation. The implication was, therefore, that these small RNAs are native components of the Crithidia ribosome. The primary structure analysis reported here supports this conclusion, since it is highly unlikely that spurious degradation could account for the discrete nature of the 5'- and 3'-termini represented in species e, f, g and j. Analysis of uniformly-^{32}P-labelled rRNAs has indicated that the four novel small rRNAs are each 5'-monophosphorylated in vivo, while the ability to label the RNAs to high specific activity with [5'-^{32}P]pCp in the RNA ligase reaction suggests a nonphosphorylated 3'-terminus in each case. Such ends (5'-P, 3'-OH) are characteristic of biosynthetic termini, whereas degradation products (at least those produced by a cyclizing type RNase activity such as pancreatic RNase) would be expected to have 5'-OH and 3'-P ends. In their basic structural features, therefore, Crithidia species e, f, g and j are

indistinguishable from those RNAs (such as 5S and 5.8S) that are recognized structural components of eukaryotic ribosomes.

Although the data are entirely consistent with the view that e, f, g and j are bona fide constituents of the Crithidia ribosome, they do not at this point allow us to draw definitive conclusions about the biosynthetic origin of these rRNAs. While preliminary experiments have indicated that genes for all the RNA species of the Crithidia ribosome are physically linked, it remains to be determined whether there is complete transcriptional linkage as well, or whether some of the novel small rRNAs, like eukaryotic 5S rRNA, are products of a transcriptional pathway separate from that which gives rise to the 18S, 5.8S and 28S rRNAs. If, in fact, species e, f, g and j are derived from the same large precursor rRNA as the 18S, 5.8S and 28S rRNA species, it must be established whether post-transcriptional processing involves removal of spacer sequences delineating the coding regions of the novel small rRNAs, as in the case of eukaryotic 5.8S rRNA, or whether the small rRNAs are generated by phosphodiester bond cleavages between completely contiguous coding sequences (not separated by spacers). The primary sequence information presented here, together with detailed restriction mapping and nucleotide sequence analysis of cloned Crithidia rDNA, should clearly distinguish among these possibilities.

6. Concluding Remarks and Future Prospects

In the course of the work presented in this thesis I have touched on many aspects of rRNA research. The most important conclusion that

can be drawn from the data presented here (3'-terminal SSU RNA sequences, the results of heterologous hybridization experiments, and preliminary analysis of modified nucleoside composition near the 3'-terminus of wheat mitochondrial 18S rRNA) is that plant mitochondria are of an endosymbiotic, specifically eubacterial, origin. Some of the supporting data have been published in scientific journals (295,314). For the sake of completion, I am planning to analyze the wheat mitochondrial 18S and 26S rRNAs in terms of content and position of modified nucleoside components.

Since wheat mitochondrial and E. coli SSU RNAs share a very similar methylation pattern, at least within the 3'-terminal 100 nucleotides, it should be possible to clone the genes for specific E. coli methylases (e.g., those responsible for production of m_2^6A and m^3U) and then to use these cloned genes to detect the genes responsible for mitochondrial rRNA modification. It would be interesting, for example, to determine whether genes responsible for methylation of wheat mitochondrial 18S rRNA are encoded by the nuclear or mitochondrial genome. It would also be interesting to determine whether the same methylase is responsible for production of m_2^6A in both wheat mitochondrial and wheat cytosol 18S rRNAs.

Our published observations on the rRNAs of Crithidia fasciculata may be summarized as follows: (i) Within its 3'-terminal 100 nucleotides, Crithidia 18S rRNA is the most divergent eukaryotic SSU RNA so far analyzed (Schnare and Gray, ref. 293). The sequence in this 3'-terminal region that has been proposed to be conserved in eukaryotic SSU RNA and to interact with 5S rRNA and mRNA is not

conserved in Crithidia 18S RNA (293). (ii) The Crithidia 5.8S rRNA is at least 10 nucleotides longer than typical eukaryotic 5.8S rRNAs and these extra nucleotides are in the same region as the spacer separating the 5.8Sa and 5.8Sb genes in Drosophila rDNA (Schnare and Gray, ref. 294). (iii) The four extra small rRNAs in C. fasciculata contain 5'-P and 3'-OH and they each have a unique sequence (Schnare et al., ref 296). Knowledge of these rRNA sequences will be valuable in precisely defining spacer sequences in Crithidia rDNA.

Many of the questions raised in this thesis about the origin and subsequent evolution of small rRNAs still await a final answer. However, we have completed most of the preliminary work on the Crithidia rRNAs necessary to allow a resolution of most of these questions in the near future. I will predict, based partially on subunit localization (123) and partially on preliminary sequence data for the ends of the two 26S rRNA fragments in Crithidia (M.N. Schnare, unpublished), that the four novel small rRNAs will map either within the gap separating the two halves of Crithidia 26S rRNA or at the 3'-end of the 26S rRNA gene. I would also predict that there will be spacer sequences separating these small rRNAs and neighboring large rRNAs and that these spacers will represent highly variable regions in other known 25S-28S rRNA sequences. Decisions about whether the appearance of spacer sequences (which result in the production of small rRNAs) was an ancient or relatively recent event in the evolution of eukaryotic rRNA genes will rest ultimately on any data we can obtain about the phylogenetic position of Crithidia. In this regard, I believe it will be necessary to sequence a complete

Crithidia 18S rRNA gene, which should verify that *Crithidia* represents one of the earliest branchings within the eukaryotes.

BIBLIOGRAPHY

1. Alberty, H., Raba, M. and Gross, H.J. (1978) Nucleic Acids Res. 5, 425-434.

2. Allet, B. and Rochaix, J.D. (1979) Cell 18, 55-60.

3. Allington, W.B., Cordry, A.L, McCullough, G.A., Mitchell, D.E. and Nelson, J.W. (1978) Anal. Biochem. 85, 188-196.

4. Altman, R. (1890) Die Elementarorganismen und ihre Beziehungen zu den Zellen. Verlag von Veit and Co. Leipzig.

5. Ambler, R.P., Daniel, M., Hermoso, J., Meyer, T.E., Bartsch, R.G., and Kamen, M.D. (1979) Nature 278, 659-660.

6. Ambler, R.P., Meyer, T.E. and Kamen, M.D. (1979) Nature 278, 661-662.

7. Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F., and Young, I.G. (1982) J. Mol. Biol. 156, 683-717.

8. Attardi, G. and Amaldi, F. (1970) Ann. Rev. Biochem. 39, 183-226.

9. Azad, A.A. (1979) Nucleic Acids Res. 7, 1913-1929.

10. Azad, A.A. and Deacon, N.J. (1979) Biochem. Biophys. Res. Comm. 86, 568-576.

11. Azad, A.A. and Deacon, N.J. (1980) Nucleic Acids Res. 8, 4365-4376.

12. Azad, A.A. and Lane, B.G. (1973) Can. J. Biochem. 51, 1195-1202.

13. Azad, A.A. and Lane, B.G. (1973) Can. J. Biochem. 51, 1669-1672.

14. Azad, A.A. and Lane, B.G. (1975) Can. J. Biochem. 53, 320-327.

15. Baer, R. and Dubin, D.T. (1980) Nucleic Acids Res. 8, 4927-4941.

16. Baer, R.J. and Dubin, D.T. (1981) Nucleic Acids Res. 9, 323-337.

17. Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R. and Wolfe, R.S. (1979) Microbiol. Rev. 43, 260-296.

18. Baralle, F.E. and Brownlee, G.G. (1978) Nature 274, 84-87.

19. Barnett, T. and Rae, P.M.M. (1979) Cell 16, 763-775.

20. Batts-Young, B. and Lodish, H.F. (1978) Proc. Natl. Acad. Sci. USA 75, 740-744.

21. Beckingham, K. and White, R. (1980) J. Mol. Biol. 137, 349-373.

22. Bedbrook, J.R., Kolodner, R. and Bogorad, L. (1977) Cell 11, 739-749.

23. Bell, G.I., DeGennaro, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P. and Rutter, W.J. (1977) J. Biol. Chem. 252, 8118-8125.

24. Bogorad, L. (1975) Science 188, 891-898.

25. Bohnert, H.J., Gordon, K.H.J. and Crouse, E.J. (1980) Mol. Gen. Genet. 179, 539-545.

26. Bonen, L. (1980) Ph.D. Thesis; Dalhousie University.

27. Bonen, L., Cunningham, R.S., Gray, M.W. and Doolittle, W.F. (1977) Nucleic Acids Res. 4, 663-671.

28. Bonen, L. and Gray, M.W. (1980) Nucleic Acids Res. 8, 319-335.

29. Borbely, G. and Simoncsits, A. (1981) Biochem. Biophys. Res. Comm. 101, 846-852.

30. Borkhardt, B. and Delius, H. (1983) Curr. Genet. 7, 327-333.

31. Borst, P. and Grivell, L.A. (1978) Cell 15, 705-723.

32. Bos, J.L., Heyting, C., Borst, P., Arnberg, A.C. and Van Bruggen, E.F.J. (1978) Nature 275, 336-338.

33. Bowman, C.M. and Dyer, T.A. (1979) Biochem. Journal 183, 605-613.

34. Bowman, L.H., Goldman, W.E., Goldberg, G.I., Hebert, M.B. and Schlessinger, D.S. (1983) Mol. Cell. Biol. 3, 1501-1510.

35. Boynton, J.E., Gillham, N.W. and Lambowitz, A.M. (1980) In Chamblis, G., Craven, G.R., Davies, J., Davis, K., Kahan, L. and Nomura, M. (ed.). Ribosomes. Structure, Function and Genetics. University Park Press, Baltimore. pp. 903-950.

36. Bram, R.J., Young, R.A. and Steitz, J. (1980) Cell 19, 393-401.

37. Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.-P., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 4303-4324.

38. Brimacombe, R. (1982) Biochem. Soc. Symp. 47, 49-60.

39. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Natl. Acad. Sci. USA 77, 201-204.

40. Brosius, J., Dull, T.J., Sleeter, D.D. and Noller, H.F. (1981) J. Mol. Biol. 148, 107-127.

41. Brosius, J., Palmer, M.L., Kennedy, P.J., and Noller, H.F. (1978) Proc. Natl. Acad. Sci. USA 75, 4801-4805.

42. Brown, W.M., George, M. Jr., Wilson, A.C. (1979) Proc. Natl. Acad. Sci. USA 76, 1967-1971.

43. Caboche, M. and Bachellerie, J.-P. (1977) Eur. J. Biochem. 74, 19-29.

44. Cameron, V., Soltis, D., and Uhlenbeck, O.C. (1978) Nucleic Acids Res. 5, 825-833.

45. Cameron, V. and Uhlenbeck, O.C. (1977) Biochemistry 16, 5120-5126.

46. Cammarano, P., Londei, P., Mazzei, F. and Biagini, R. (1982) Comp. Biochem. Physiol. 73B, 435-449.

47. Campbell, G.R., Littau, V.C., Melera, P.W., Allfrey, V.G. and Johnson, E.M. (1979) Nucleic Acids Res. 6, 1433-1447.

48. Carbon, P., Ebel, J.P. and Ehresmann, C. (1981) Nucleic Acids Res. 9, 2325-2333.

49. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.-P. (1979) Eur. J. Biochem. 100, 399-410.

50. Cavalier-Smith, T. (1975) Nature 256, 463-468.

51. Chamblis, G., Craven, G.R., Davies, J., Davis, K., Kahan, L. and Nomura, M. (ed.) (1980) Ribosomes. Structure, Function and Genetics. University Park Press, Baltimore.

52. Chan, Y.-L., Endo, Y. and Wool, I.G. (1983) J. Biol. Chem. 258, 12768-12770.

53. Chan, Y.-L., Gutell, R., Noller, H.F. and Wool, I.G. (1984) J. Biol. Chem. 259, 224-230.

54. Chan, Y.-L., Olvera, J. and Wool, I.G. (1983) Nucleic Acids Res. 11, 7819-7831.

55. Chao, S., Sederoff, R.R. and Levings, C.S. III (1983) Plant Physiol. 71, 190-193.

56. Chatton, E. (1938) Titres et Travaux Scientifiques. E. Sottano, Sète.

57. Clarke, C.G. and Gerbi, S.A. (1982) J. Mol. Evol. 18, 329-336.

58. Clark-Walker, G.D., McArthur, C.R. and Sriprakash, K.S. (1981) J. Mol. Biol. 147, 399-415.

59. Clark-Walker, G.D., McArthur, C.R. and Sriprakash, K.S. (1983) J. Mol. Evol. 19, 333-341.

60. Clark-Walker, G.D., Sriprakash, K.S., McArthur, C.R. and Azad, A.A. (1980) Curr. Genet. 1, 209-217.

61. Clary, D.O., Goddard, J.M., Martin, N.C., Fauron, C.M.-R. and Wolstenholme, D.R. (1982) Nucleic Acids Res. 10, 6619-6637.

62. Cordingley, J.S. and Turner, M.J. (1980) Mol. Biochem. Parasitol. 1, 91-96.

63. Cory, S. and Adams, J.M. (1977) Cell 11, 795-805.

64. Cox, R.A. (1977) Prog. Biophys. Mol. Biol. 32, 193-231.

65. Cox, R.A. and Kelly, J.M. (1981) FEBS Lett. 130, 1-6.

66. Cox, R.A. and Kelly, J.M. (1982) Biochem. Soc. Symp. 47, 11-48.

67. Cox, R.A. and Thompson, R.D. (1980) Biochem. Journal 187, 75-90.

68. Cross, G.A.M. (1970) Biochim. Biophys. Acta 204, 470-477.

69. Cummings, D.J., Maki, R.A., Conlon, P.J. and Laping, J. (1980) Mol. Gen. Genet. 178, 499-510.

70. Cunningham, R.S., Bonen, L., Doolittle, W.F. and Gray, M.W. (1976) FEBS Lett. 69, 116-122.

71. Cunningham, R.S. and Gray, M.W. (1977) Biochim. Biophys. Acta 475, 476-491.

72. Cunningham, R.S., Gray, M.W., Doolittle, W.F. and Bonen, L. (1977) in Acides Nucléiques et Synthèse des Protéines Chez les Végétaux, Bogorad, L. and Weil, J.H. (ed.). Centre National de la Recherche Scientifique, Paris, pp. 243-248.

73. Curtis, S.E. and Rawson, J.R.Y. (1981) Gene 15, 237-247.

74. D'Alessio, J.M., Harris, G.H., Perna, P.J. and Paule, M.R. (1981) Biochemistry 20, 3822-3827.

75. Darzynkiewicz, E., Nakashima, K., and Shatkin, A.J. (1980) J. Biol. Chem. 255, 4973-4975.

76. Dawid, I.B. and Wellauer, P.K. (1976) Cell 8, 443-448.

77. Dayhoff, M.O. and Schwartz, R.M. (1981) Ann. N.Y. Acad. Sci. 361, 92-103.

78. de Jonge, P., Klootwijk, J. and Planta, R.J. (1977) Nucleic Acids Res. 4, 3655-3663.

79. Delihas, N., Andersen, J., Andresini, W., Kaufman, L. and Lyman, H. (1981) Nucleic Acids Res. 9, 6627-6633.

80. de Vries, H., de Jonge, J.C., Bakker, H., Meurs, H. and Kroon, A. (1979) Nucleic Acids Res. 6, 1791-1803.

81. De Wachter, R. (1979) Nucleic Acids Res. 7, 2045-2054.

82. Din, N. and Engberg, J. (1979) J. Mol. Biol. 134, 555-574.

83. Dodson, E.O. (1979) Can. J. Microbiol. 25, 651-674.

84. Donis-Keller, H. (1980) Nucleic Acids Res. 8, 3133-3142.

85. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527-2538.

86. Doolittle, W.F. (1978) Nature 272, 581-582.

87. Doolittle, W.F. and Pace, N.R. (1971) Proc. Natl. Acad. Sci. USA. 68, 1786-1790.

88. Dougherty, E.C. (1957) J. Protozool. 4 (Suppl.), 14. Reprinted in Symp. Soc. Gen. Microbiol. 20, 433.

89. Douglas, S.E. (1983) Ph.D. Thesis. Dalhousie University.

90. Douthwaite, S., Christensen, A. and Garret, R.A. (1983) J. Mol. Biol. 169, 249-279.

91. Dron, M., Rahire, M. and Rochaix, J.-D. (1982) Nucleic Acids Res. 10, 7609-7620.

92. Dubin, D.T. and Taylor, R.H. (1978) J. Mol. Biol. 121, 523-540.

93. Dubin, D.T., Taylor, R.H. and Davenport, L.W. (1978) Nucleic Acids Res. 5, 4385-4397.

94. Duester, G.L. and Holmes, W.M. (1980) Nucleic Acids Res. 8, 3793-3807.

95. Eckhardt, H., and Lührmann, R. (1979) J. Biol. Chem. 254, 11185-11188.

96. Edwards, K., Bedbrook, J., Dyer, T. and Kössel, H. (1981) Biochem. Int. 2, 533-538.

97. Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 2853-2869.

98. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467.

99. Falconet, D., Lejeune, B., Quetier, F. and Gray, M.W. (1984) EMBO Journal 3, 297-302.

100. Farrelly, F. and Butow, R. (1983) Nature 301, 296-301.

101. Faye, G., Dennebouy, N., Kujawa, C. and Jacq, C. (1979) Mol. Gen. Genet. 168, 101-109.

102. Federoff, N.V. (1979) Cell 16, 697-710.

103. Fellner, P. and Sanger, F. (1968) Nature 219, 236-238.

104. Fluhr, R. and Edelman, M. (1981) Mol. Gen. Genet. 181, 484-490.

105. Ford, P.J. and Mathieson, T. (1978) Eur. J. Biochem. 87, 199-214.

106. Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N. and Woese, C.R. (1980) Science 209, 457-463.

107. Frederick, J.F. (ed.) (1981) Ann. N.Y. Acad. Sci. 361.

108. Furlong, J.C. and Maden, B.E.H. (1983) EMBO Journal 2, 443-448.

109. Garvin, R.T., Hill, R.C. and Weber, M.M. (1978) Arch. Biochem. Biophys. 191, 774-781.

110. Gellissen, G., Bradfield, J.Y., White, B.N. and Wyatt, G.R. (1983) Nature 301, 631-634.

111. Georgiev, O.I., Nikolaev, N., Hadjiolov, A.A., Skryabin, K.G., Zakharyev, V.M. and Bayev, A.A. (1981) Nucleic Acids Res. 9, 6953-6958.

112. Gillham, N.W. (1978) Organelle Heredity. Raven Press, New York.

113. Glaus, K.R., Zassenhaus, H.P., Fechheimer, N.S. and Perlman, P.S. (1980). In Kroon, A.M. and Saccone, C. (ed.). The Organization and Expression of the Mitochondrial Genome. Elsevier/North-Holland Biomedical Press, Amsterdam. p. 131-135.

114. Glotz, C., Zwieb, C., Brimacombe, R., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 3287-3306.

115. Glover, D.M. and Hogness, D.S. (1977) Cell 10, 167-176.

116. Goldbach, R.W., Borst, P., Bollen-De Boer, J.E. and Van Bruggen, E.F.J. (1978) Biochim. Biophys. Acta 521, 169-186.

117. Goldman, W.E., Goldberg, G., Bowman, L.H., Steinmetz, D. and Schlessinger, D. (1983) Mol. Cell. Biol. 3, 1488-1500.

118. Gourse, R.L. and Gerbi, S.A. (1980) J. Mol. Biol. 140, 321-339.

119. Graf, L., Kossel, H. and Stutz, E. (1980) Nature 286, 908-910.

120. Graf, L., Roux, E., Stutz, E. and Kössel, H. (1982) Nucleic Acids Res. 10, 6369-6381.

121. Grainger, R.M. and Maizels, N. (1980) Cell 20, 619-623.

122. Gray, M.W. (1979) Can. J. Biochem. 57, 914-926.

123. Gray, M.W. (1981) Mol. Cell Biol. 1, 347-357.

124. Gray, M.W. (1982) Can. J. Biochem. 60, 157-171.

125. Gray, M.W. (1983) BioScience 33, 693-699.

126. Gray, M.W., Bonen, L., Falconet, D., Huh, T.Y., Schnare, M.N. and Spencer, D.F. (1982). In Slonimski, P., Borst, P. and Attardi G. (ed.) Mitochondrial Genes. Cold Spring Harbor Laboratory, N.Y. pp. 483-488.

127. Gray, M.W. and Doolittle, W.F. (1982) Microbiol. Rev. 46, 1-42.

128. Gray, M.W., Huh, T.Y., Schnare, M.N., Spencer, D.F. and Falconet,

D. (1983) In Ciferri, O. and Dure, L., III (ed). Structure and Function of Plant Genomes. Plenum Press, New York/London. pp. 373-380.

129. Gray, M.W. and Lane, B.G. (1967) Biochim. Biophys. Acta 134, 243-257.

130. Gray, M.W., Sankoff, D. and Cedergren, R.J. (1984) Nucleic Acids Res. 12, 5837-5852.

131. Gray, M.W. and Spencer, D.F. (1981) Nucleic Acids Res. 9, 3523-3529.

132. Gray, M.W. and Spencer, D.F. (1983) FEBS Lett. 161, 323-327.

133. Gray, M.W. and Thomas, C.A. (1978) Can. J. Biochem. 56, 500-507.

134. Gray, P.W. and Hallick, R.B. (1978) Biochemistry 17, 284-289.

135. Gray, P.W. and Hallick, R.B. (1979) Biochemistry 18, 1820-1825.

136. Green, M.R., Grimm, M.F., Goewert, R.R., Collins, R.A., Cole, M.D., Lambowitz, A.M., Heckman, J.E., Yin, S. and RajBhandary, U.L. (1981) J. Biol. Chem. 256, 2027-2034.

137. Groot, G.S.P. and Van Harten-Loosbroek, N. (1980) Curr. Genet. 1, 133-135.

138. Gubler, U., Wyler, T. and Braun, R. (1979) FEBS Lett. 100, 347-350.

139. Gupta, R., Lanter, J.M. and Woese, C.R. (1983) Science 221, 656-659.

140. Hadler, H.I., Dimitrijevic, B. and Mahalingam, R. (1983) Proc. Natl. Acad. Sci. USA 80, 6495-6499.

141. Hagenbüchle, O., Santer, M., Steitz, J.A. and Mans, R.J. (1978) Cell 13, 551-563.

142. Hahn, U., Lazarus, C.M., Lünsdorf, H. and Küntzel, H. (1979) Cell 17, 191-200.

143. Hall, L.M.C. and Maden, B.E.H. (1980) Nucleic Acids Res. 8, 5993-6005.

144. Hall, R.H. (1971) The Modified Nucleosides in Nucleic Acids. Columbia University Press, New York/London.

145. Hartley, M.R. (1979) Eur. J. Biochem. 96, 311-320.

146. Hartley, M.R. and Head, C. (1979) Eur. J. Biochem. 96, 301-309.

147. Hashimoto, S., Sakai, M. and Muramatsu, M. (1975) Biochemistry 14, 1956-1964.

148. Hassur, S.M., and Whitlock, H.W. Jr. (1974) Anal: Biochem. 59, 162-164.

149. Hawkins, S.E. and Hughes, J.M. (1973) Nature New Biol. 246, 199-200.

150. Heckman, J.E. and RajBhandary, U.L. (1979) Cell 17, 583-595.

151. Hindenach, B.R. and Stafford, D.W. (1984) Nucleic Acids Res. 12, 1737-1747.

152. Hofman, J.D., Lau, R.H. and Doolittle, W.F. (1979) Nucleic Acids Res. 7, 1321-1333.

153. HsuChen, C.-C. and Dubin, D.T. (1980) J. Bacteriol. 144, 991-998.

154. Dubin, D.T. and HsuChen, C.-C. (1983) Plasmid 9, 307-320.

155. Hudspeth, M.E.S., Shumard, D.S., Bradford, C.J.R. and Grossman, L.I. (1983) Proc. Natl. Acad. Sci. USA 80, 142-146.

156. Huh, T.Y. and Gray, M.W. (1982) Plant Mol. Biol. 1, 245-249.

157. Iams, K.P. and Sinclair, J.H. (1982) Proc. Natl. Acad. Sci. USA 79, 5926-5929.

158. Jacobs, H.T., Pasokony, J.W., Grula, J.W., Roberts, J.W., Xin, J.-H., Britten, R.J. and Davidson, E.H. (1983) J. Mol. Biol. 165, 609-632.

159. Jacq, B. (1981) Nucleic Acids Res. 9, 2913-2932.

160. Jarsch, M., Altenbuchner, J. and Böck, A. (1983) Mol. Gen. Genet. 189, 41-47.

161. Jarsch, M. and Böck, A. (1983) Nucleic Acids Res. 11, 7537-7544.

162. John, P., and Whatley, F.R. (1975) Nature 254, 495-498.

163. Johnston, F.B. and Stern, H. (1957) Nature 179, 160-161.

164. Jordan, B.R., Latil-Damotte, M. and Jourdan, R. (1980) FEBS Lett. 117, 227-231.

165. Jordan, B.R., Latil-Damotte, M. and Jourdan, R. (1980) Nucleic Acids Res. 8, 3565-3573.

166. Jurgenson, J.E. and Bourque, D.P. (1980) Nucleic Acids Res. 8, 3505-3516.

167. Kagramanova, V.K., Mankin, A.S., Baratova, L.A. and Bogdanov, A.A. (1982) FEBS Lett. 144, 177-180.

168. Kelly, J.M. and Cox, R.A. (1982) Nucleic Acids Res. 10, 6733-6745.

169. Kemble, R.J., Mans, R.J., Gabay-Laughnan, S. and Laughnan, J.R. (1983) Nature 304, 744-747.

170. Keus, R.J.A., Dekker, A.F., van Roon, M.A. and Groot, G.S.P. (1983) Nucleic Acids Res. 11, 6465-6474.

171. Keus, R.J.A., Roovers, D.J., van Heerikhuizen, H. and Groot, G.S.P. (1983) Curr. Genet. 7, 7-12.

172. Keyhani, E. (1981) Ann. N.Y. Acad. Sci. 361, 376-396.

173. Kirby, K.S. (1965) Biochem. Journal 96, 266-269.

174. Klier, A.F., Kunst, F. and Rapoport, G. (1979) Nucleic Acids Res. 7, 997-1010.

175. Klootwijk, J., Klein, I. and Grivell, L.A. (1975) J. Mol. Biol. 97, 337-350.

176. Klootwijk, J. and Planta, R.J. (1973) Eur. J. Biochem. 39, 325-333.

177. Kobayashi, M., Seki, T., Yaginuma, K. and Koike, K. (1981) Gene 16, 297-307.

178. Koch, W., Edwards, K., and Kössel, H. (1981) Cell 25, 203-213.

179. Köchel, H.G. and Küntzel, H. (1981) Nucleic Acids Res. 9, 5689-5696.

180. Köchel, H.G. and Küntzel, H. (1982) Nucleic Acids Res. 10, 4795-4801.

181. Koller, B. and Delius, H. (1980) Mol. Gen. Genet. 178, 261-269.

182. Kölodner, R. and Tewari, K.K. (1975) Biochim. Biophys. Acta 402, 372-390.

183. Koncz, Cs. and Sain, B. (1980) FEBS Lett. 109, 141-144.

184. Kozak, M. (1978) Cell 15, 1109-1123.

185. Kozak, M. (1980) Cell 22, 7-8.

186. Kozak, M. (1982) Biochem. Soc. Symp. 47, 113-128.

187. Kozak, M. (1983) Microbiol. Rev. 47, 1-45.

188. Kozak, M. (1984) Nucleic Acids Res. 12, 857-872.

189. Kramer, R.A., Philippsen, P. and Davis, R.W. (1978) J. Mol. Biol. 123, 405-416.

190. Kroon, A.M., Pepe, G., Bakker, H., Holtrop, M., Bollen, J.E., Van Bruggen, E.F.J., Cantatore, P., Terpstra, P. and Saccone, C. (1977) Biochim. Biophys. Acta 478, 128-145.

191. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Cell 31, 147-157.

192. Kumano, M., Tomioka, N. and Sugiura, M. (1983) Gene 24, 219-225.

193. Küntzel, H. and Köchel, H.G. (1981) Nature 293, 751-755.

194. Lambowitz, A.M. and Luck, D.J.L. (1976) J. Biol. Chem. 251, 3081-3095.

195. Lane, B.G. (1963) Biochim. Biophys. Acta 72, 110-112.

196. Lane, B.G. (1965) Biochemistry 4, 212-219.

197. Lau, R.Y., Kennedy, T.D. and Lane, B.G. (1974) Can. J. Biochem. 52, 1110-1123.

198. Lazarus, C.M., Lünsdorf, H., Hahn, U., Stępień, P.P., and Küntzel, H. (1980) Mol. Gen. Genet. 177, 389-397.

199. Leaver, C.J. (1973) Biochem. Journal 135, 237-240.

200. Leaver, C.J. and Gray, M.W. (1982) Ann. Rev. Plant Physiol. 33, 373-402.

201. Leaver, C.J. and Harmey, M.A. (1973) Biochem. Soc. Symp. 38, 175-193.

202. Leaver, C.J. and Harmey, M.A. (1976) Biochem. Journal 157, 275-277.

203. Leon, W., Fouts, D.L. and Manning, J. (1978) Nucleic Acids Res. 5, 491-504.

204. Li, M., Tzagoloff, A., Underbrink-Lyon, K. and Martin, N.C. (1982) J. Biol. Chem. 257, 5921-5928.

205. Lockard, R.E., Connaughton, J.F. and Kumar, A. (1982) Nucleic Acids Res. 10, 3445-3457.

206. Loening, U.E. (1967) Biochem. Journal 102, 251-257.

207. Loening, U.E. (1968) J. Mol. Biol. 38, 355-365.

208. Londei, P., Cammarano, P., Mazzei, F. and Romeo, A. (1982) Comp. Biochem. Physiol. 73B, 423-434.

209. Long, E.O. and Dawid, I.B. (1980) Ann. Rev. Biochem. 49, 727-764.

210. Long, E.O. and Dawid, I.B. (1979) Cell 18, 1185-1196.

211. Loughney, K., Lund, E. and Dahlberg, J.E. (1982) Nucleic Acids Res. 10, 1607-1624.

212. Lund, E., Dahlberg, J.E., Lindahl, L., Jaskunas, S.R., Dennis, P.P. and Nomura, M. (1976) Cell 7, 165-177.

213. Luoma, G.A. and Marshall, A.G. (1978) Proc. Natl. Acad. Sci. USA 75, 4901-4905.

214. Machatt, M.A., Ebel, J.-P. and Branlant, C. (1981) Nucleic Acids Res. 9, 1533-1549.

215. MacKay, R.M. (1981) FEBS Lett. 123, 17-18.

216. MacKay, R.M. and Doolittle, W.F. (1981) Nucleic Acids Res. 9, 3321-3334.

217. MacKay, R.M., Gray, M.W. and Doolittle, W.F. (1980) Nucleic Acids Res. 8, 4911-4917.

218. MacKay, R.M., Spencer, D.F., Doolittle, W.F. and Gray, M.W. (1980) Eur. J. Biochem. 112, 561-576.

219. Maden, B.E.H. (1980) Nature 288, 293-296.

220. Maden, B.E.H. and Tartof, K. (1974) J. Mol. Biol. 90, 51-64.

221. Mahler, H.R. (1983) Int. Rev. Cytol. 82, 1-98.

222. Maizels, N. (1976) Cell 9, 431-438.

223. Malnoë, P.M. and Rochaix, J.-D. (1978) Mol. Gen. Genet. 166, 269-275.

224. Maly, P. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 7263-7286.

225. Mankin, A.S. and Kopylov, A.M. (1981) Biochem. Int. 3, 587-593.

226. Marco, Y. and Rochaix, J.D. (1980) Mol. Gen. Genet. 177, 715-723.

227. Margulis, L. (1970) Origin of Eukaryotic Cells, Yale University Press, New Haven.

228. Margulis, L. (1981) Symbiosis in Cell Evolution. W.H. Freeman and Company, San Francisco.

229. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.

230. McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R., and Sogin, M.L. (1983) Biochemistry 22, 5858-5868.

231. McCutchan, T.F., Gilham, P.T. and Söll, D. (1975) Nucleic Acids Res. 2, 853-864.

232. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.

233. Michot, B., Bachellerie, J.-P., Raynal, F. and Renalier, M.-H. (1982) FEBS Lett. 142, 260-266.

234. Michot, B., Bachellerie, J.-P. and Raynal, F. (1983) Nucleic Acids Res. 11, 3375-3391.

235. Miyata, T., Hayashida, H., Kikuno, R., Hasegawa, M., Kobayashi, M. and Koike, K. (1982) J. Mol. Evol. 19, 28-35.

236. Morales, N.M. and Roberts, J.F. (1978) Comp. Biochem. Physiol. 59B, 1-4.

237. Morales, N.M. and Roberts, J.F. (1978) J. Protozool. 25, 140-144.

238. Morgan, E.A., Ikemura, T. and Nomura, M. (1977) Proc. Natl. Acad. Sci. USA 74, 2710-2714.

239. Nass, M.M.K. and Nass, S. (1963) J. Cell. Biol. 19, 593-611.

240. Nazar, R.N. (1980) FEBS Lett. 119, 212-214.

241. Nazar, R.N. (1982) FEBS Lett. 143, 161-162.

242. Nazar, R.N. and Seitz, T.O. (1980) FEBS Lett. 115, 71-76.

243. Nazar, R.N., Sitz, T.O. and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597.

244. Neumann, H., Gierl, A., Tu, J., Leibrock, J., Staiger, D. and Zillig, W. (1983) Mol. Gen. Genet. 192, 66-72.

245. Nichols, J.L., and Lane, B.G. (1966) Biochim. Biophys. Acta 119, 649-651.

246. Nichols, J.L., and Wijesinghe, W. (1978) Can. J. Biochem. 56, 760-764.

247. Nishikawa, K. and Takemura, S. (1974) FEBS Lett 40, 106-109.

248. Nishimura, S., Harada, F., Narushima, U., and Seno, T. (1967) Biochim. Biophys. Acta. 142, 133-148.

249. Noller, H.F., Kop, J., Wheaton, V., Brosius, J., Gutell, R.R., Kopylov, A.M., Dohme, F., Herr, W., Stahl, D.A., Gupta, R. and Woese, C.R. (1981) Nucleic Acids Res. 9, 6167-6189.

250. Noller, H.F. and Woese, C.R. (1981) Science 212, 403-411.

251. Olver, K.M. and Lane, B.G. (1970) Can. J. Biochem. 48, 1113-1121.

252. Olsen, G.J., McCarroll, R. and Sogin, M.L. (1983) Nucleic Acids Res. 11, 8037-8049.

253. Orozco, E.M., Jr., Gray, P.W. and Hallick, R.B. (1980) J. Biol. Chem. 255, 10991-10996.

254. Orozco, E.M. Jr., Rushlow, K.E., Dodd, J.R., and Hallick, R.B. (1980) J. Biol. Chem. 255, 10997-11003.

255. Osinga, K.A. and Tabak, H.F. (1982) Nucleic Acids Res. 10, 3617-3626.

256. Otsuka, T., Nomiyama, H., Yoshida, H., Kukita, T., Kuhara, S. and Sakaki, Y. (1983) Proc. Natl. Acad. Sci. USA 80, 3163-3167.

257. Palmer, F.B.St.C. (1973) Biochim. Biophys. Acta 316, 296-304.

258. Palmer, J.D., Singh, G.P. and Pillay, D.T.N. (1983) Mol. Gen. Genet. 190, 13-19.

259. Palmer, J.D. and Stein, D.B. (1982) Curr. Genet. 5, 165-170.

260. Palmer, J.D. and Thompson, W.F. (1981) Proc. Natl. Acad. Sci. USA 78, 5533-5537.

261. Parish, J.H. and Kirby, K.S. (1966) Biochim. Biophys. Acta 129, 554-562.

262. Pavlakis, G.N., Jordan, B.R., Wurst, R.M. and Vournakis, J.N. (1979) Nucleic Acids Res. 7, 2213-2238.

263. Payne, P.I. and Loening, U.E. (1970) Biochim. Biophys. Acta 224, 128-135.

264. Peattie, D.A. (1979) Proc. Natl. Acad. Sci. USA 76, 1760-1764.

265. Peattie, D.A. and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA 77, 4679-4682.

266. Pellegrini, M., Manning, J.E. and Davidson, N. (1977) Cell 10, 213-224.

267. Perry, R.P. (1976) Ann. Rev. Biochem. 45, 605-629.

268. Philippsen, P., Thomas, M., Kramer, R.A. and Davis, R.W. (1978) J. Mol. Biol. 123, 387-404.

269. Pinder, J.C., Staynov, D.Z. and Gratzer, W.B. (1974) Biochemistry 13, 5373-5378.

270. Pochon, F. and Michelson, A.M. (1967) Biochim. Biophys. Acta 149, 99-106.

271. Raff, R.A. and Mahler, H.R. (1972) Science 177, 575-582.

272. Ramirez, J.L. and Dawid, I.B. (1978) J. Mol. Biol. 119, 133-146.

273. Raynal, F., Michot, B., Bachellerie, J.P. (1984) FEBS Lett. 167, 263-268.

274. Reeder, R.H., Higashinakagawa, T. and Miller, O. Jr. (1976) Cell 8, 449-454.

275. Reijnders, L. (1975) J. Mol. Evol. 5, 167-176.

276. Renkawitz, R., Gerbi, S.A. and Glätzer, K.H. (1979) Mol. Gen. Genet. 173, 1-13.

277. Rochaix, J.D. and Darlix, J.L. (1982) J. Mol. Biol. 159, 383-395.

278. Rochaix, J.D. and Malnoe, P. (1978) Cell 15, 661-670.

279. Rogg, H., Brambilla, R., Keith, G. and Staehelin, M. (1976) Nucleic Acids Res. 3, 285-295.

280. Rondinelli, E., Soares, C.M., De Castro, J.F. and De Castro, F.T. (1981) Arch. Biochem. Biophys. 209, 349-355.

281. Rosenberg, M. (1974) Nucleic Acids Res. 1, 653-671.

282. Rubin, G.M. (1973) J. Biol. Chem. 248, 3860-3875.

283. Rubtsov, P.M., Musakhanov, M.M., Zakharyev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980) Nucleic Acids Res. 8, 5779-5794.

284. Saccone, C., Cantatore, P., Gadaleta, G., Gallerani, R., Lanave, C., Pepe, G. and Kroon, A.M. (1981) Nucleic Acids Res. 9, 4139-4148.

285. Salim, M. and Maden, B.E.H. (1973) Nature 244, 334-336.

286. Salim, M. and Maden, B.E.H. (1981) Nature 291, 205-208.

287. Samols, D.R., Hagenbuchle, O. and Gage, L.P. (1979) Nucleic Acids Res. 7, 1109-1119.

288. Sanders, J.P.M., Heyting, C. and Borst, P. (1975) Biochem. Biophys. Res. Comm. 65, 699-707.

289. Sanders, J.P.M., Heyting, C., Verbeet, M., Ph., Meijlink, F.C.P.W. and Borst, P. (1977) Mol. Gen. Genet. 157, 239-261.

290. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

291. Santer, M., Chung, S.-C., Harmon, G., Estner, M., Hendrick, J.P., Hopper, J., Brecht, C., Pandhi, P. (1979) J. Bacteriol. 140, 131-140.

292. Sargan, D.R., Gregory, S.P. and Butterworth, P.H.W. (1982) FEBS Lett. 147, 133-136.

293. Schnare, M.N. and Gray, M.W. (1981) FEBS Lett. 128, 298-304.

294. Schnare, M.N. and Gray, M.W. (1982) Nucleic Acids Res. 10, 2085-2092.

295. Schnare, M.N. and Gray, M.W. (1982) Nucleic Acids Res. 10, 3921-3932.

296. Schnare, M.N., Spencer, D.F. and Gray, M.W. (1983) Can. J. Biochem. Cell. Biol. 61, 38-45.

297. Schwartz, R.M. and Dayhoff, M.O. (1978) Science 199, 395-403.

298. Schwarz, Zs. and Kössel, H. (1980) Nature 283, 739-742.

299. Schwarz, Z., Kossel, H., Schwarz, E. and Bogorad, L. (1981) Proc. Natl. Acad. Sci. USA 78, 4748-4752.

300. Seilhamer, J.J. and Cummings, D.J. (1981) Nucleic Acids Res. 9, 6391-6406.

301. Shah, M., Khan, N., Salim, M. and Maden, B.E.H. (1978) Biochem. Journal 169, 531-542.

302. Shine, J., and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342-1346.

303. Shine, J., and Dalgarno, L. (1975) Nature 254, 34-38.

304. Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977) Nature 269, 833-836.

305. Simpson, L. and Simpson, A.M. (1978) Cell 14, 169-178.

306. Singh, H. and Lane, B.G. (1964) Can. J. Biochem. 42, 87-93.

307. Singh, H. and Lane, B.G. (1964) Can. J. Biochem. 42, 1011-1021.

308. Skryabin, K.G., Kraev, A.S., Rubstov, P.M. and Bayev, A.A. (1979) Dokl. Akad. Nauk. SSSR 247, 761-765.

309. Skryabin, K.G., Zakharyev, V.M., Rubstov, P.M. and Bayev, A.A. (1979) Dokl. Akad. Nauk. SSSR, 247, 1275-1277.

310. Sor, F. and Fukuhara, H. (1980) C.R. Hebd. Seances Sor. D. Acad. Sci. 291, 933-936.

311. Sor, F. and Fukuhara, H. (1983) Nucleic Acids Res. 11, 339-348.

312. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.

313. Spencer, D.F., Bonen, L. and Gray, M.W. (1981) Biochemistry 20, 4022-4029.

314. Spencer, D.F., Schnare, M.N. and Gray, M.W. (1984) Proc. Natl. Acad. Sci. USA 81, 493-497.

315. Spencer, R. and Cross, G.A.M. (1976) J. Gen. Microbiol. 93, 82-88.

316. Spielmann, A., Ortiz, W. and Stutz, E. (1983) Mol. Gen. Genet. 190, 5-12.

317. Steege, D.A., Graves, M.C. and Spremulli, L.L. (1982) J. Biol. Chem. 257, 10430-10439.

318. Steitz, J.A. and Jakes, K. (1975) Proc. Natl. Acad. Sci. USA 72, 4734-4738.

319. Steitz, J.A. and Steege, D.A. (1977) J. Mol. Biol. 114, 545-558.

320. Stern, D.B., Dyer, T.A. and Lonsdale, D.M. (1982) Nucleic Acids Res. 10, 3333-3340.

321. Stevens, A.R. and Pachler, P.F. (1972) J. Mol. Biol. 66, 225-237.

322. Stewart, G.C., Wilson, F.E. and Bott, K.F. (1982) Gene 19, 153-162.

323. Stiegler, P., Carbon, P., Ebel, J.-P. and Ehresmann, C. (1981) Eur. J. Biochem. 120, 487-495.

324. Stiegler, P., Carbon, P., Zuker, M., Ebel, J.-P. and Ehresmann, C. (1981) Nucleic Acids Res. 9, 2153-2172.

325. Subrahmanyam, C.S., Cassidy, B., Busch, H. and Rothblum, L.I. (1982) Nucleic Acids Res. 10, 3667-3680.

326. Surzycki, S.J. and Rochaix, J.D. (1971) J. Mol. Biol. 62, 89-109.

327. Szer, W. and Shugar, D. (1960) Acta Biochem. Polon. 7, 491-504.

328. Takaiwa, F. and Sugiura, M. (1980) Gene 10, 95-103.

329. Takaiwa, F. and Sugiura, M. (1980) Mol. Gen. Genet. 180, 1-4.

330. Takaiwa, F. and Sugiura, M. (1982) Eur. J. Biochem. 124, 13-19.

331. Takaiwa, F. and Sugiura, M. (1982) Nucleic Acids Res. 10, 2665-2676.

332. Taniguchi, T. and Weissmann, C. (1978) Nature 275, 770-772.

333. Tinoco, I., Jr., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40-41.

334. Tohdoh, N., Shinozaki, K. and Sugiura, M. (1981) Nucleic Acids Res. 9, 5399-5406.

335. Tohdoh, N. and Sugiura, M. (1982) Gene 17, 213-218.

336. Tomioka, N., Shinozaki, K. and Sugiura, M. (1981) Mol. Gen. Genet. 184, 359-363.

337. Tomioka, N. and Sugiura, M. (1983) Mol. Gen. Genet. 191, 46-50.

338. Torczynski, R., Bollon, A.P. and Fuke, M. (1981) Mol. Gen. Genet. 184, 557-559.

339. Torczynski, R., Bollon, A.P. and Fuke, M. (1983) Nucleic Acids Res. 11, 4879-4890.

340. Tsuzuki, T., Nomiyama, H., Setoyama, C., Maeda, S. and Shimada, K. (1983) Gene 25, 223-229.

341. Tu, J. and Zillig, W. (1982) Nucleic Acids Res. 10, 7231-7245.

342. Uchida, T., Bonen, L., Schaup, H.W., Lewis, B.J., Zablen, L. and Woese, C. (1974) J. Mol. Evol. 3, 63-77.

343. Unnasch, T.R. and Wirth, D.F. (1983) Nucleic Acids Res. 11, 8461-8472.

344. Ursi, D., Vandenberghe, A. and De Wachter, R. (1983) Nucleic Acids Res. 11, 8111-8120.

345. Uzzell, T., and Spolsky, C. (1974) Am. Sci. 62, 334-343.

346. Uzzell, T. and Spolsky, C. (1981) Ann. N.Y. Acad. Sci. 361, 481-499.

347. Valenzuela, P., Bell, G.I., Venegas, A., Sewell, E.T., Masiarz, F.R., DeGennaro, L.J., Weinberg, F. and Rutter, W.J. (1977) J. Biol. Chem. 252, 8126-8135.

348. Van Charldorp, R., Heus, H.A. and Van Knippenberg, P.H. (1981) Nucleic Acids Res. 9, 2717-2725.

349. Van Charldorp, R., Van Kimmenade, A.M.A., Van Knippenberg, P.H. (1981) Nucleic Acids Res. 9, 4909-4917.

350. Van Charldorp, R. and Van Knippenberg, P.H. (1982) Nucleic Acids Res. 10, 1149-1158.

351. van den Boogaart, P., Samallo, J. and Agsteribbe, E. (1982) Nature 298, 187-189.

352. Van Etten, R.A., Walberg, M.W. and Clayton, D.A. (1980) Cell 22, 157-170.

353. Veldman, G.M., Brand, R.C., Klootwijk, J. and Planta, R.J. (1980) Nucleic Acids Res. 8, 2907-2920.

354. Veldman, G.M., Klootwijk, J., deRegt, V.C.H.F., Planta, R.J., Branlant, C., Krol, A. and Ebel, J.-P. (1981) Nucleic Acids Res. 9, 6935-6952.

355. Veldman, G.M., Klootwijk, J., van Heerikhuizen, H. and Planta, R.J. (1981) Nucleic Acids Res. 9, 4847-4862.

356. Walker, T.A. and Pace, N.R. (1977) Nucleic Acids Res. 4, 595-601.

357. Walker, W.F. (1981) FEBS Lett. 126, 150-151.

358. Wallace, D.C. (1982) Microbiol. Rev. 46, 208-240.

359. Wallace, F.G. (1979) Biology of the Kinetoplastida. Vol. 2, Lumsden, W.H. and Evans, D.A. (ed.) Academic Press. N.Y.

360. Walseth, T.F. and Johnson, R.A. (1979) Biochim. Biophys. Acta. 562, 11-31.

361. Ware, V.C., Tague, B.W., Clark, C.G., Gourse, R.L., Brand, R.C. and Gerbi, S.A. (1983) Nucleic Acids Res. 11, 7795-7817.

362. Weith, H.L., Wiebers, J.L. and Gilham, P.T. (1970) Biochemistry 9, 4396-4401.

363. Wellauer, P.K. and Dawid, I.B. (1977) Cell 10, 193-212.

364. Wesolowski, M., Algeri, A. and Fukuhara, H. (1981) Curr. Genet. 3, 157-162.

365. Whatley, F.R. (1981) Ann. N.Y. Acad. Sci. 361, 330-340.

366. White, R.L. and Hogness, D.S. (1977) Cell 10, 177-192.

367. Whitfeld, P.R., Herrmann, R.G. and Bottomley, W. (1978) Nucleic Acids Res. 5, 1741-1751.

368. Whitfeld, P.R., Leaver, C.J., Bottomley, W. and Atchison, B.A. (1978) Biochem. J. 175, 1103-1112.

369. Whittaker, R.H. (1969) Science 163, 150-160.

370. Wild, M.A. and Gall, J.G. (1979) Cell 16, 565-573.

371. Wildeman, A.G. and Nazar, R.N. (1980) J. Biol. Chem. 255, 11896-11900.

372. Wildeman, A.G. and Nazar, R.N. (1982) Eur. J. Biochem. 121, 357-363.

373. Williamson, S.E. and Doolittle, W.F. (1983) Nucleic Acids Res. 11, 225-235.

374. Woese, C.R. (1977) J. Mol. Evol. 10, 93-96.

375. Woese, C.R. (1981) Sci. Am. 244, 98-122.

376. Woese, C.R. and Fox, G.E. (1977) J. Mol. Evol. 10, 1-6.

377. Woese, C.R. and Fox, G.E. (1977) Proc. Natl. Acad. Sci. USA 74, 5088-5090.

378. Woese, C.R., Fox, G.E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B.J. and Stahl, D. (1975) Nature 254, 83-86.

379. Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983) Microbiol. Rev. 47, 621-669.

380. Woese, C.R., Magrum, L.J., Gupta, R., Siegel, R.B., Stahl, D.A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J.J. and Noller, H.F. (1980) Nucleic Acids Res. 8, 2275-2293.

381. Wolf, S.F. and Schlessinger, D. (1977) Biochemistry 16, 2783-2791.

382. Wrede, P. and Erdmann, V.A. (1977) Proc. Natl. Acad. Sci. USA 74, 2706-2709.

383. Wright, R.M. and Cummings, D.J. (1983) Curr. Genet. 7, 151-157.

384. Wright, R.M. and Cummings, D.J. (1983) Nature 302, 86-88.

385. Young, R.A. and Steitz, J.A. (1978) Proc. Natl. Acad. Sci. USA 75, 3593-3597.

386. Ziff, E.B. and Evans, R.M. (1978) Cell 15, 1463-1475.

387. Zingales, B. and Colli, W. (1977) Biochim. Biophys. Acta 474, 562-577.

388. Zwieb, C., Glotz, C. and Brimacombe, R. (1981) Nucleic Acids Res. 9, 3621-3640.