# TRANSCRIPTION OF THE HeLa CELL MITOCHONDRIAL GENOME

Thesis by

Deanna Ojala Johnson

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1981

(Submitted October 2, 1980)

Dedicated to my parents

Reino and Audrey

**INVICTUS**

*Out of the night that covers me,*
*Black as the Pit from pole to pole,*
*I thank whatever gods may be*
*For my unconquerable soul.*

*In the fell clutch of circumstance*
*I have not winced nor cried aloud.*
*Under the bludgeonings of chance*
*My head is bloody, but unbowed.*

*Beyond this place of wrath and tears*
*Looms but the horror of the shade,*
*And yet the menace of the years*
*Finds, and shall find me, unafraid.*

*It matters not how strait the gate,*
*How charged with punishments the scroll,*
*I am the master of my fate;*
*I am the captain of my soul.*

WILLIAM ERNEST HENLEY

## Acknowledgements

This work was supported in part by a National Research Service Award (T32 GM07616) from the National Institute of General Medical Sciences and also by the John Weigle Memorial Fund.

I am very grateful to many, many people on the staff and faculty here at Caltech who have provided guidance, moral support and friendship over the last several years. In addition, there are those to whom I owe special thanks:

Ralph Yount, who initiated me into the world of science, for which I will always be grateful.

Giuseppe Attardi, my advisor, whose devotion to scientific excellence cannot be overestimated and whose high standards provide an example for those who work with him.

Jack Griffith, whose belief in me and personal and professional support over the last decade means more to me than words can express.

Don Carrigan, a dear friend and scientist, whose genius, humility, and humanity provide a possibility for greatness which few can hope to achieve.

Julio Montoya, whose collaboration during my last year here made possible the completion of this thesis.

Robert P. Johnson, my husband, who has given me the courage and freedom to proceed ahead with life.

## Abstract

The HeLa cell mitochondrial genome has been shown to encode, in addition to a number of tRNAs, two ribosomal RNAs and at least 17 discrete poly(A)-containing RNA species. A variety of techniques, including Southern blots, Northern blots and Berk and Sharp analyses, were utilized to orient the ribosomal and polyadenylated species with respect to a detailed restriction map. Thus, nineteen species were successfully localized to within an accuracy of about 30 to 40 nucleotides, thereby allowing a number of interesting observations to be made concerning the process of transcription of this genome. First, both the ribosomal RNAs and polyadenylated RNA components are transcribed colinearly; no intervening sequences exist. This is in dramatic contrast to the situation observed in mitochondrial DNA of some strains of yeast. Second, there exists no overlapping of the mature polyadenylated species, the ribosomal RNAs and the tRNAs (within the resolution described) and, with the exception of the D-loop region, the mitochondrial DNA sequences appear to be completely utilized for the synthesis of the transcripts. Third, a comparison of the positions of the DNA sequences which encode the polyadenylated transcripts with respect to those encoding the tRNAs shows that the majority of those transcripts are flanked at both 5'- and 3'-ends by a tRNA.

Direct sequencing analyses have been used to investigate the precise relationship of the 5'- and 3'-ends of many of the polyadenylated transcripts (which are presumptive mRNAs) with respect to the ends of the tRNAs. Results obtained have demonstrated that those species which have been shown to be flanked by a tRNA gene at their 5'-end begin immediately following the tRNA 3'-terminal nucleotide. Correspondingly, the 3'-terminal nucleotide of seven of these RNA transcripts has been found to be immediately juxtaposed to the 5'-terminal nucleotide of the flanking tRNA. In one case, the 3'-terminal nucleotide is adjacent to the 5'-terminal nucleotide of the following mRNA. Moreover, an additional striking feature of these RNAs

is that, with one possible exception, they either lack a 5'-end noncoding stretch, or contain an abbreviated version of it. Thus six species begin directly with AUG or AUA (which codes for methionine in human mitochondria), while three species contain those triplets from within one to eight nucleotides of the 5'-end.

A model is proposed whereby structural features or sequences of tRNA specific transcripts play a role in the processing events which form the mature species. Specifically, the tRNA would provide a signal for a precise endonucleolytic cleavage event which would generate the mature species from a larger precursor. The results presented in this thesis, in correspondence with sequencing data (B. Barrell and F. Sanger, personal communication), describe a transcriptional system characterized by simplicity, efficiency and economy.

## TABLE OF CONTENTS


| | | Page |
|---|---|---|
| **Chapter 1** | Molecular Biology of the Animal Cell Mitochondrial Genome: An Introduction . . . . . | 1 |
| **Chapter 2** | Fine Mapping of the Ribosomal RNA Genes of HeLa Cell Mitochondrial DNA . . . . . | 28 |
| **Chapter 3** | The tRNA Genes Punctuate the Reading of Genetic Information in Human Mitochondrial DNA . . . . . | 40 |
| **Chapter 4** | 7S RNA: A Polyadenylated Transcript Which Maps Adjacent to the Origin of DNA Replication in HeLa Cell Mitochondria . . . . . | 85 |
| **Chapter 5** | The Putative mRNA for Subunit II of Cytochrome c Oxidase in Human Mitochondria Starts Directly at the Translation Initiator Codon . . . . . | 109 |
| **Chapter 6** | Distinctive Features of the 5'-Terminal Sequences of the Human Mitochondrial mRNAs . . . . . | 126 |
| **Chapter 7** | Transcription of the HeLa Cell Mitochondrial Genome: The Punctuation Model . . . . . | 150 |

# Chapter 1

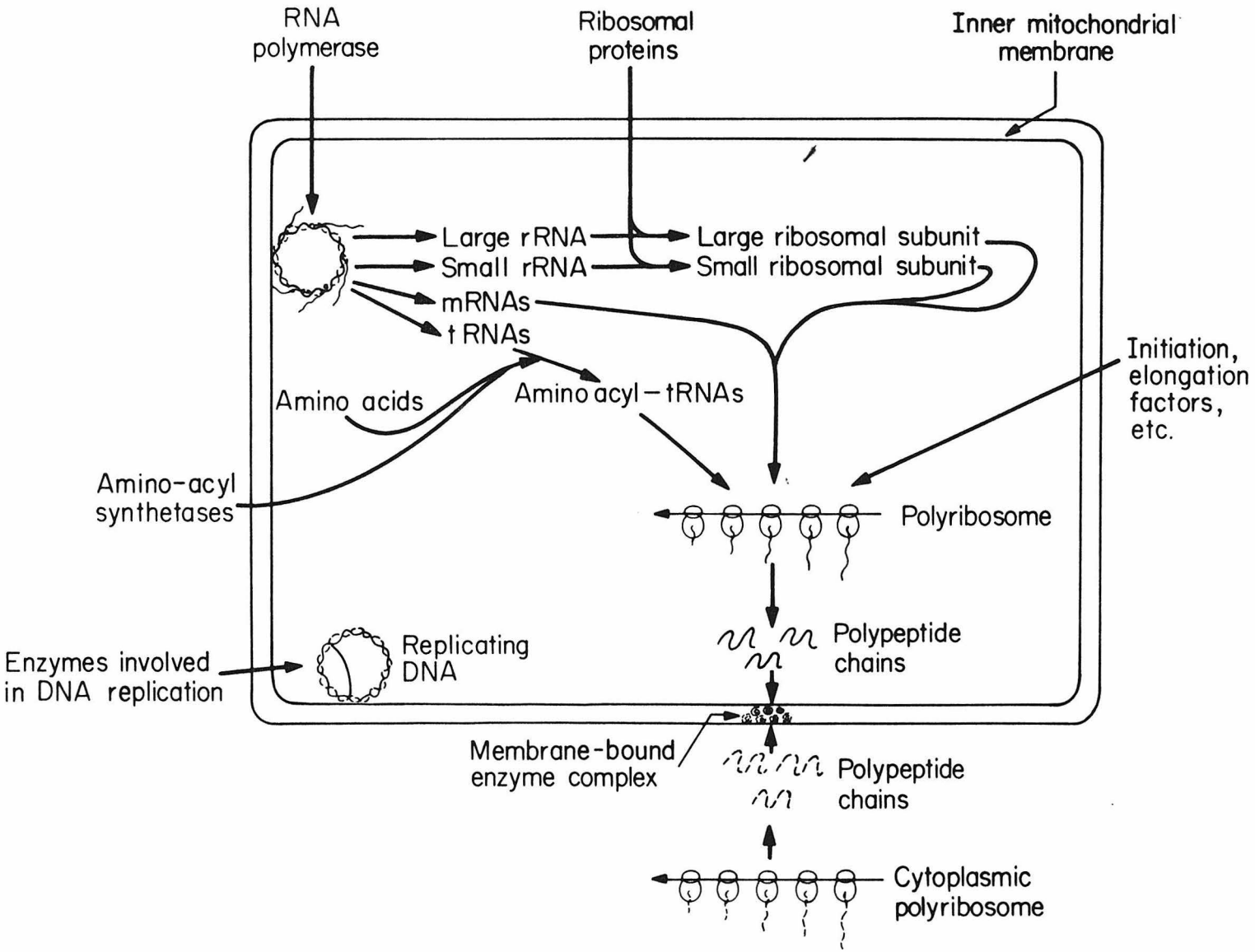
## Molecular Biology of the Animal Cell

## Mitochondrial Genome: An Introduction

## MOLECULAR BIOLOGY OF THE ANIMAL CELL MITOCHONDRIAL GENOME: AN INTRODUCTION

Molecular biology of animal cell mitochondria had its beginnings in the early 1960's with the demonstration that fibrous material contained within mitochondria of certain animal cells showed electron microscopic features and staining reactions characteristic for that of deoxyribonucleic acid (DNA) (Nass and Nass, 1963a,b). These observations were subsequently expanded to include a wide range of phyla (Nass et al., 1965). Shortly thereafter, the existence of a heterogeneous class of cytoplasmic ribonucleic acid (RNA), homologous in sequence to that of mitochondrial DNA, was demonstrated (Attardi and Attardi, 1967, 1968, 1969). Further studies have since revealed that, in fact, the mitochondrial genome represents the smallest replicon known which contains genetic information for the three major classes of RNA (ribosomal, transfer and messenger). Along with these observations came the realization of the tightly coupled relationship existing between the nuclear and cytoplasmic genetic systems. Thus, the differentiation of the organelle into a respiratory-competent mitochondrion requires the expression of both nuclear and mitochondrial genes. Figure 1 (taken from Attardi, 1980) schematically illustrates the generally accepted role of mtDNA in mitochondriogenesis, based on information derived from several systems (Borst and Grivell, 1978; Cummings et al., 1979; Kroon and Saccone, 1980; Borst, 1972; Schatz and Mason, 1974).

The need to maintain, at considerable cost to the cell, two separate genetic systems and the possible mechanisms which regulate the interplay of these systems has intrigued geneticists, biochemists and molecular biologists for many years. Recent rapid technological advances in the isolation and sequencing analyses of nucleic acids and proteins have, however, made it possible for scientists to begin to answer some of these questions.


RNA polymerase
Ribosomal proteins
Inner mitochondrial membrane
Large rRNA
Small rRNA
mRNAs
t RNAs
Large ribosomal subunit
Small ribosomal subunit
Amino acids
Amino acyl – tRNAs
Initiation, elongation factors, etc.
Amino-acyl synthetases
Polyribosome
Polypeptide chains
Enzymes involved in DNA replication
Replicating DNA
Membrane-bound enzyme complex
Polypeptide chains
Cytoplasmic polyribosome

In this introductory chapter, I would like to present a brief history of aspects of mitochondrial molecular biology, concentrating primarily on that of animal cells and to discuss advances made in our knowledge of mitochondrial DNA replication and of its transcriptional and translational gene products. In so doing, I will also summarize the surprising and dramatic recent findings concerning the manner in which the animal cell mitochondrial genome utilizes, in a remarkably efficient and simple way, the genetic information encoded within it.

**Mitochondrial DNA; Physical Characterization and Mode of Replication**

In animal cells, mtDNA consists of a homogeneous population of small duplex circles ranging in size from 4.45 μ in a sea urchin to 5.85 μ in Urechis caupo (a flatworm). As reviewed by Borst in 1972 and 1977, no animal is known that lacks circular mtDNA; the small differences reported in contour length appear to be real, having been obtained in mixing experiments (Nass, 1969). The 5 μ length corresponds to approximately $10 \times 10^6$ daltons or 16,000 nucleotide pairs. The number of mtDNA molecules per cell varies with the organism and cell type examined, ranging from approximately 500 to 2000. A considerable amount of evidence derived from renaturation kinetics (Borst, 1972), electron microscope studies of renatured molecules (Clayton et al., 1970), and restriction enzyme analysis (Potter et al., 1975; Brown and Vinograd, 1974; Klukas and Dawid, 1976; Ojala and Attardi, 1977; Brown and Goodman, 1979; Dawid and Rastl, 1979) indicates that the population of mtDNA molecules within a given cell (and very likely, within the same individual) is substantially homogeneous in sequence. Therefore, within the framework of the eucaryotic genome organization, the population of mtDNA molecules in a cell can be considered to represent one of the many multigene families, which in this case, presides over a differentiation process of the inner mitochondrial membrane. As is the case for other multigene families, the multiplicity of these molecules demands the existence of mechanisms for the maintenance of sequence homogeneity and for keeping constant

the size and number of DNA molecules per cell. As yet, nothing is known about these mechanisms.

As has been found for all closed circular duplex DNAs, isolated mtDNA exists in a negative superhelical form. This property has been used as a basis for the detection and purification of these molecules from contaminating nuclear DNA. Thus, Radloff et al. (1967) demonstrated that the binding of intercalative dyes will cause a partial unwinding of duplex DNA structure. Because the maximum amount of dye that can be bound by a closed circular molecule is smaller than that bound by a linear or nicked circular molecule, the different molecular forms will demonstrate different buoyant densities in salt-dye gradients.

A second interesting property of all animal cell mtDNA studied (and one which also has aided in investigations into the genetic mysteries of this genome) is that, in alkaline CsCl gradients, the DNA will separate into two bands, which in HeLa cells (Attardi et al., 1970) and mouse L-cells (Clayton et al., 1970) (and presumably in the other cell lines also), correspond to the separated complementary strands. In HeLa, this is attributed to the difference in thymine content of the two strands.

Distinctive or unusual bases do not exist in mtDNA, however the presence of ribonucleotides has been observed (Grossman et al., 1973). The origin of the ribonucleotides is unclear; they may result from insertion during replication or repair or perhaps be remnants of RNA primers.

Although, as mentioned above, mtDNA isolated from the same cell (or the same individual) appears to be essentially homogeneous in sequence, there is a substantial amount of evidence that considerable differences exist between mtDNAs of various individuals of a single species; furthermore, the nucleotide sequences studied exhibit a fairly rapid rate of evolution (Dawid et al., 1976; Upholt and Dawid, 1977; Brown et al., 1979, 1980; Dawid and Rastl, 1979; Kroon et al., 1978; Francisco

et al., 1979.) Thus, primarily by the use of heteroduplex mapping and restriction analyses, the rate of evolution has been calculated to exceed that of the single-copy fraction of the nuclear genome (Brown et al., 1979, 1980). This unusual rate of evolution has to be contrasted with the fact that the functions of mtDNA appear to be largely conserved among diverse organisms. Thus, all mtDNAs encode rRNAs, tRNAs, and mRNAs specific for a few polypeptides which appear to be analogous among the diverse species studied.

A sequence anlaysis of human mtDNA has recently been completed (Barrell et al., 1979; Eperon et al., 1980; Sanger et al., 1980; F. Sanger and B. Barrell, personal communication). Comparison of these sequence data with that of mtDNA of other species (which should be forthcoming in the near future) will clarify the nature of the observed heterogeneity.

Replication of the mtDNA of a variety of animal cells occurs by a modified Cairns mode (reviewed by Kasamatsu and Vinograd, 1974). During this process, a replicating intermediate, termed D-loop DNA, is formed by the synthesis of a short stretch of heavy strand (H-strand) DNA [varying in length from about 550 to 680 nucleotides in human cells (Gillum and Clayton, 1978; Brown et al., 1978; Crews et al., 1979)] originating from a defined position on the genome (Ojala and Attardi, 1978; Bogenhagen et al., 1978). This stretch of DNA, referred to as 7S DNA based on its sedimentation coefficient (Kasamatsu et al., 1971) remains hydrogen-bonded to the light strand (L-strand), and as such causes a displacement of the parental H-strand. A large percentage of mtDNA examined contain the D-loop structure (the precise percentage depending on the species) and it has been suggested that synthesis of the 7S DNA stretch results from initiation of replication which proceeds until the positive free energy of superhelix formation is just equal to the negative free energy available from the synthesis reaction (Kasamatsu et al., 1971). Presumably, synthesis of the remainder of the H-strand by repeated nicking and closing cycles

occurs by extension of the 7S mtDNA (although all 7S DNA sequences do not serve as primers because of their relatively rapid turnover rate [Berk and Clayton, 1974]). This asychronous and unidirectional replication continues until approximately 60% of the strand has been copied (Robberson et al., 1972), at which point synthesis of the light strand begins. Replication of both strands then proceeds, with a final separation of fully formed daughter molecules. The nucleotide sequence of the region containing the H-strand replication origin in HeLa cell mtDNA (Crews et al., 1979) and that containing the L-strand replication origin in mouse L-cell mtDNA (Martens and Clayton, 1979) has been determined. Comparisons of the two show no striking correlations, although both show regions of dyad symmetry.

As concerns RNA priming of these replication events, evidence has been presented supporting such a possibility in mouse L-cells (Bogenhagen et al., 1979; Martens and Clayton, 1979). See also Chapter 4 of this thesis.

To elucidate the temporal relationship of the replication event with respect to the cell cycle, two independent studies have been carried out in this laboratory. In HeLa cells (Pico-Mattocia and Attardi, 1972) and in mouse A9 cells (Novitski, 1979) mtDNA synthesis was found to occur predominantly during late S and G2 phases. These results are in apparent conflict with those of Bogenhagen and Clayton (1977), in which a study of unsynchronized A9 cells indicated a constant rate of mtDNA replication throughout the cell cycle. However, it is the belief of our laboratory that the latter results were erroneously interpreted and therefore not in basic disagreement with our own experimental data (refer to Appendix II of Novitski, 1979).

The nature of the controls which exist to couple the nuclear and mitochondrial genomes are still unknown. Models have been proposed in which extra-mitochondrial factors, encoded in nuclear DNA, act to influence (possibly in a negative fashion) the macromolecular synthetic activity of mtDNA. Thus, Barath and Küntzel (1972)

observed, in Neurospora crassa, the synthesis of mitochondrial proteins made by the nucleocytoplasmic system increases when mitochondrial protein synthesis is blocked. They postulated that the nuclear encoded enzymes involved in mitochondrial replication, transcription and translation are coordinately repressed, and that the repressor is a mitochondrial translation product. Attardi et al. (1977), on the basis of investigations of the temporal pattern of mtDNA, RNA and protein synthesis during the cell cycle in HeLa cells synchronized by the selective detachment technique, suggested that labile nuclear signals, produced once per cell cycle and acting stoichiometrically, could be the mechanism whereby the number of mtDNA replication events per cell cycle and therefore the number of mtDNA molecules per cell are kept relatively constant. More recently, Rinaldi et al. (1980) concluded that the cell nucleus negatively controls mitochondrial DNA duplication in sea urchin eggs. This was based on the observation that mtDNA synthesis is stimulated in enucleated and parthenogenetically activated eggs.

The enzymology of mtDNA replication is only slightly better characterized. The presence of a nuclear coded mtDNA specific DNA polymerase ($\gamma$) is well established in vertebrate cells (Scovassi et al., 1979) and also in lower eucaryotes and some invertebrate phyla (Bertazzoni and Scovassi, 1980); furthermore, evidence has been presented for the presence of mitochondrial specific topoisomerases I and II in rat liver extracts (Castora et al., 1980). An association of the region of the genome containing the origin of replication with that of the inner mitochondrial membrane has been postulated based on electron microscopic observations of the existence of a protein complex occurring near or at the origin (Albring et al., 1977) coupled with earlier EM observations of membrane attachment of the genome (Nass, 1969).

**Mitochondrial Transcription**

While studies on mtDNA replication have long been facilitated by techniques

which allowed the isolation of that DNA in relatively pure form, investigations into the nature of mitochondrial transcription and therefore of the transcription products have not been similarly favored. Until recently, investigators were forced to rely on various drug treatments in order to suppress nuclear DNA transcription (treatments which could conceivably affect the validity of the results obtained) (Zylber et al., 1969; Attardi et al., 1970; Perlman and Penman, 1970; Perlman et al., 1973). Additionally, the small amount of mtRNA synthesized (a few micrograms per gram of cells) often required laborious preparative procedures.

However (despite the difficulties), mitochondrial DNA was shown, primarily by hybridization experiments using labeled RNA species, to encode two ribosomal RNAs, a number of tRNAs, and also polyadenylated RNA species (Attardi et al., 1970; Aloni and Attardi, 1971a; Dawid, 1972; Ojala and Attardi, 1972, 1974a,b,c; Lynch and Attardi, 1976). Correspondingly, similar approaches demonstrated that the mitochondrial genome is transcribed not only completely, but also symmetrically (Aloni and Attardi, 1971b,c; Murphy et al., 1975). In addition, concurrent pulse labeling experiments indicated that the rate of in vivo transcription of the H- and L-strands is approximately equal, even though the L-strand transcripts do not accumulate to any great extent (Aloni, 1971c). A more recent approach has suggested that the rate of L-strand transcription may exceed that of the H-strand by a factor of 2 to 3 (Cantatore and Attardi, 1980).

Initial attempts aimed at mapping the sequences encoding the stable transcripts relied primarily on the elegant techniques of electron microscopy, in some cases coupled with the use of the electron-opaque label ferritin. Thus, the ribosomal RNA cistrons in HeLa (Robberson et al., 1972; Wu et al., 1972; Angerer et al., 1976) and in Xenopus laevis (Ramirez and Dawid, 1978), the tRNA cistrons in HeLa (Wu et al., 1972; Angerer et al., 1976) and in Xenopus (Ohi et al., 1978), and those of total RNA transcripts in Xenopus (Rastl and Dawid, 1979) were positioned within

the genome. These initial studies indicated that, by far, the majority of the stable transcripts are encoded by the heavy strand. The asymmetry of the informational content of this DNA is in interesting contrast to that of the transcriptional symmetry.

A second interesting observation produced by the initial mapping studies was the limited number of tRNAs which could be found to hybridize to the mitochondrial genome [17, in HeLa (as determined by hybridization with charge tRNAs) (Lynch and Attardi, 1976) or (as determined by ferritin mapping), 19 in HeLa (Angerer et al., 1976) and 20 in Xenopus laevis (Ohi et al., 1978)]. This is far less than the minimum number of 32 tRNAs required by the wobble hypothesis. Three theories were postulated to account for this discrepancy, i.e., import of nuclear DNA coded tRNAs, restricted codon usage, and different pattern of codon recognition by mitochondrial tRNAs. The latter has proven to be the correct one. The recently completed sequence of human mtDNA (F. Sanger and B. Barrell, personal communication) has revealed the existence of only 22 tRNA coding sequences. More significantly, an analysis of the distribution of the tRNAs coded for by these genes among the codon boxes has revealed a striking pattern: namely, in the eight family boxes with four codons for one amino acid, only one specific tRNA has been found, instead of two as in the universal code; in each case, this single tRNA has a U in the first position of the anticodon. It has been speculated that this tRNA is capable of reading all four codons in the family boxes, probably due to pairing of the U in the first position of the anticodon with all four bases in the third position of the codons in each family box (Barrell et al., 1980). This correlates with similar conclusions reached for yeast (Bonitz et al., 1980) and Neurospora crassa (Heckman et al., 1980) mitochondria. The decoding mechanism proposed for the family boxes cannot obviously operate in the non-family boxes, where two tRNAs are used to read the four codons, because it would lead to misreading. The observation that, in Neurospora crassa, mitochondrial tRNA species for the family boxes have an unmodified U in the first

position of the anticodon, while tRNAs specific for the two codons ending in purines in the non-family boxes have an unknown modified U in the same position, has led to the proposal that this modification may be the mechanism preventing misreading of the two codons ending in pyrimidines in the non-family boxes (Heckman et al., 1980).

The above described codon recognition pattern could well account for the reduced number of tRNAs utilized for mitochondrial protein synthesis. That a further simplification may occur in the mammalian mitochondrial genetic code is suggested by the fact that no tRNA gene for the arginine codons $AG\begin{matrix}A\\G\end{matrix}$ has been found, and that the codons AGA and AGG do not appear to be present in the significant reading frames of human mtDNA (Barrell et al., 1980). Thus, it is possible that protein synthesis can proceed in human mitochondria with only 23 tRNA species. (A tRNA gene for the methionine codons $AU\begin{matrix}A\\G\end{matrix}$, although expected, has not yet been found in human mtDNA.) The proposed codon assignment and codon recognition pattern in mammalian mitochondria are shown in Figure 2 (taken from Attardi, 1980). In each box, the codons (5' → 3') are on the left and the anticodons (3' → 5') are on the right.

Mapping studies of polyadenylated mitochondrial transcripts have proceeded less rapidly than those of the ribosomal and tRNA species, primarily due to the low abundance and difficulties in isolation of purified transcripts. Recent technological advances have, however, made it possible to surpass these problems. Thus, Berk and Sharp (1977, 1978) developed a procedure for the sizing and mapping of transcripts with respect to known restriction sites in which the purification of large amounts of individual transcripts is not required. Advances also in gel technology and auto-radiography increased the possibilities of isolating and detecting purified species.

The Berk and Sharp methodology has been applied, with varying degrees of success, to map the transcripts of mouse L cell (Battey and Clayton, 1978), of Xenopus laevis (Rastl and Dawid, 1979) and of HeLa cell mtDNA (Ojala and Attardi,

| | | | |
|---|---|---|---|
| UUU, UUC } Phe AAG<br>UUA, UUG } Leu AAU | UCU, UCC, UCA, UCG } Ser AGU | UAU, UAC } Tyr AUG<br>UAA, UAG } Ter | UGU, UGC } Cys ACG<br>UGA, UGG } Trp ACU |
| CUU, CUC, CUA, CUG } Leu GAU | CCU, CCC, CCA, CCG } Pro GGU | CAU, CAC } His GUG<br>CAA, CAG } Gln GUU | CGU, CGC, CGA, CGG } Arg GCU |
| AUU, AUC } Ile UAG<br>AUA } Met UAU<br>AUG } F-Met UAC | ACU, ACC, ACA, ACG } Thr UGU | AAU, AAC } Asn UUG<br>AAA, AAG } Lys UUU | AGU, AGC } Ser UCG<br>AGA, AGG } (Arg) (UCU) Ter? |
| GUU, GUC, GUA, GUG } Val CAU | GCU, GCC, GCA, GCG } Ala CGU | GAU, GAC } Asp CUG<br>GAA, GAG } Glu CUU | GGU, GGC, GGA, GGG } Gly CCU |

1980; Ojala et al., 1980a; Chapters 2 and 3 of this thesis). A comparison of these results reveals good similarity between the Xenopus laevis and HeLa cell maps, but little correlation with that of mouse L cell. A detailed discussion of these mapping results, and the conclusions derived from them, is presented in Chapters 2 and 3. Briefly, a major observation in both the Xenopus and HeLa cell mitochondrial systems is that there exists no overlapping of the major stable transcripts. Furthermore, there are no intervening sequences and, with the exception of the region around the origin of replication, the H-strand DNA sequences are saturated with the discrete transcripts encoded by that strand. This tightly-packed arrangement of the transcripts and almost total utilization of the HeLa cell mitochondrial DNA sequences for the production of these transcripts is in dramatic contrast to the situation observed in mitochondria of lower eucaryotes. For example, in yeast, the genes are separated by AT-rich stretches and several of these genes have been shown to contain multiple intervening sequences (Van Ommen et al., 1979; Church et al., 1979).

As concerns the nature of the HeLa cell mitochondrial polyadenylated transcripts, a considerable amount of evidence derived from the detailed mapping, structural, metabolic and sequence studies carried out in our laboratory and correlated with DNA sequence data has provided insight into the function and mode of expression of these species. Recently resolved into 18 discrete components by electrophoresis through agarose slab gels in the presence of methylmercuric hydroxide as a denaturing agent (Amalric et al., 1978), they have been characterized as to their presence in polysomes (Ojala and Attardi, 1972; Amalric et al., 1978), their relative abundance (Attardi et al., 1979; Attardi, 1980), their half-life (Gelfand, 1980), and finally, as to the identity of the DNA sequences from which they are transcribed (Ojala et al., 1980a,b; Cantatore and Attardi, 1980; Montoya, Ojala and Attardi, manuscript in preparation; Ojala, Montoya and Attardi, manuscript in preparation; Chapters 2 to 7 of this thesis). Results obtained strongly support that among these components,

species 5, 7, 9, 11 to 16 and 17 (utilizing the classification of Amalric et al., 1978; refer to Figure 8, Chapter 3) are probably specific mRNAs. In addition, correlation of RNA sequencing data (Montoya, Ojala and Attardi, manuscript in preparation; Chapter 6 of this thesis), or of reading frames in mtDNA (Walker et al., 1980) with protein sequence data (Walker et al., 1980; Chomyn, Hunkapiller and Attardi, manuscipt in preparation; Barrell et al., 1979; A. Tzagoloff, personal communication) further support the mRNA nature of RNAs 9, 11, 14, 15 and 16. Thus, these RNAs appear to be, respectively, the specific messengers for cytochrome c oxidase subunit I, cytochrome b, subunit 6 of the oligomycin sensitive ATPase, and cytochrome c oxidase subunits III and II.

In addition, sequencing analyses of many of the RNA species (detailed discussions are presented in Chapters 5, 6 and 7) reveal further examples of the economic utilization of HeLa cell mitochondrial DNA sequences. In particular, the 5'- and 3'-terminal nucleotides of the majority of the polyadneylated transcripts and those of the ribosomal RNAs are juxtaposed to flanking tRNAs. Furthermore, the available evidence indicates that all of the putative mRNAs coded for by the H-strand contain an initiator codon from within one to eight nucleotides of the 5'-end. The poly(A) addition step, itself, may illustrate a further lesson in economy by providing, for some encoded polypeptides, a terminator codon in the correct reading frame.

Interestingly, while the presence in these species of a poly(A) tail [estimated to consist of stretches of about 55 nucleotides (Hirsch and Penman, 1973; Ojala and Attardi, 1974)] represents a definite eucaryote trait, they also have been found to lack another eucaryotic attribute, i.e., that of the "cap" structure at their 5'-end (Grohmann et al., 1978).

As is the case for mtDNA replication, the enzymology of mitochondrial transcription is not yet well understood. Mitochondrial specific DNA-dependent

RNA polymerases from a variety of systems, including yeast (Levins et al., 1980), Neurospora crassa (Küntzel and Shäfer, 1971), rat liver (Gallerani and Saccone, 1974) and Xenopus oocytes (Wu and Dawid, 1972) have been partially purified and characterized as to their sensitivity to rifampicin and/or α-amanitin. In vitro studies in which mtDNA has been used as the template generally show products which exhibit little similarity to those synthesized in vivo (Gallerani and Saccone, 1974). Additionally, in HeLa cells, a correlation of mtRNA synthesis with that of the cell cycle has been observed. As was found for mtDNA replication, there is a considerable acceleration of mtRNA synthesis during the S and G2 phases of the cell cycle (Pica-Mattocia and Attardi, 1971).

**Mitochondrial Gene Products**

Animal cell mitochondria contain hundreds of polypeptides. Only a small fraction (∿5%), however, are mitochondrially coded and synthesized, and the majority of these (by analogy to lower eucaryotes) appear to be hydrophobic proteins which reside in the inner mitochondrial membrane and which collaborate with extramitochondrially synthesized proteins to form multisubunit enzymatic complexes involved in electron transport and oxidative phosphorylation (Borst and Grivell, 1978). The coordinate synthesis by two different genetic systems and subsequent integration of these proteins into functional complexes remains an intriguing aspect of membrane biology.

The availability of sophisticated genetic tools has allowed rapid progress in the analysis and functional correlation of mitochondrially synthesized polypeptides of lower eucaryotes. Studies in animal cells, however, have had to rely on physical approaches for analysis of mitochondrial gene products. Thus, bidimensional electrophoresis through a combined SDS/8 M urea polyacrylamide slab gel and SDS-polyacrylamide gradient slab gel system has resolved 25 discrete translation products

isolated from HeLa cell mitochondria (Ching, 1980). These products have been defined as mitochondrial in origin on the basis of the emetine resistance and chloramphenicol sensitivity of their labeling. Furthermore, pulse-chase experiments have indicated that they are probably not related in a precursor to product fashion (Ching, 1980).

Unfortunately, there is still little information regarding the nature of these translation products or their correlation with putative mRNAs. Assuming a functional analogy with the mitochondrial genetic systems from lower eucaryotes, human mtDNA would be expected to code for three subunits of cytochrome c oxidase, two to three subunits of the oligomycin sensitive ATPase, one subunit of the cytochome bc complex and one protein associated with mitochondrial ribosomes (Borst and Grivell, 1978; Cummings et al., 1979; Kroon and Saccone, 1980). While genes for five of the above-mentioned products have been identified on the HeLa cell mitochondrial genome on the basis of correlations of RNA and DNA sequence data with protein sequence data, a correlation with the translation products has been made only for the three subunits of cytochrome c oxidase. This is true for human cells (Hare et al., 1980), rat liver cells (Rascati and Parsons, 1979) and Xenopus oocytes (Koch, 1976). Correlations of the polypeptides corresponding to the other genes also identified in human mtDNA should soon be made.

The nature of the remaining polypeptides is still a matter of conjecture. Thirteen reading frames longer than 200 nucleotides have been found in the human mtDNA sequence, 12 on the H-strand and one on the L-strand (Walker et al., 1980). These correlate well with the putative mRNAs mapped on the HeLa cell mitochondrial genome. In addition, it is believed that, in yeast, the genetic loci encoding the seven to eight gene products listed above saturate or nearly saturate the yeast mitochondrial genetic map. Therefore, whatever the function of these unidentified polypeptides may be, investigations into their role in mitochondriogenesis remains an interesting task for the future.

## References

Albring, M., Griffith, J. and Attardi, G. (1977) Association of a protein structure of probable membrane derivation with HeLa cell mitochondrial DNA near its origin of replication. Proc. Nat. Acad. Sci. USA **74**, 1348-1352.

Aloni, Y. and Attardi, G. (1971a) Expression of the mitochondrial genome in HeLa cells. IV. Titration of mitochondrial genes for 16S, 12S and 4S RNA. J. Mol. Biol. **55**, 271-276.

Aloni, Y. and Attardi, G. (1971b) Expression of the mitochondrial genome in HeLa cells. II. Evidence for complete transcription of mitochondrial DNA. J. Mol. Biol. **55**, 251-270.

Aloni, Y. and Attardi, G. (1971c) Expression of the mitochondrial genome in HeLa cells. VII. Symmetric in vivo transcription of mitochondrial DNA in HeLa cells. Proc. Nat. Acad. Sci. USA **68**, 1757-1761.

Amalric, F., Merkel, C., Gelfand, R. and Attardi, G. (1978) Fractionation of mitochondrial RNA from HeLa cells by high-resolution electrophoresis under strongly denaturing conditions. J. Mol. Biol. **118**, 1-25.

Angerer, L., Davidson, N., Murphy, W., Lynch, D. and Attardi, G. (1976) An electron microscope study of the relative positions of the 4S and ribosomal RNA genes in HeLa cell mitochondrial DNA. Cell **9**, 81-90.

Attardi, B. and Attardi, G. (1967) A membrane associated RNA of cytoplasmic origin in HeLa cells. Proc. Nat. Acad. Sci. USA **58**, 1051-1058.

Attardi, B. and Attardi, G. (1969) Sedimentation properties of RNA species homologous to mitochondrial DNA in HeLa cells. Nature **224**, 1079-1083.

Attardi, G. (1980) Organization and expression of the mammalian mitochondrial genome: a lesson in economy. Trends Biochem. Sci., in press.

Attardi, G., Aloni, Y., Attardi, B., Ojala, D., Pica-Mattoccia, L., Robberson, D. L. and Storrie, B. (1970) Transcription of mitochondrial DNA in HeLa cells. Cold Spring Harbor Symp. Quant. Biol. **35**, 599-617.

Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C. and Ojala, D. (1979) The organization of the genes in the human mitochondrial genome and their mode of transcription. In: Extra Chromosomal DNA, Vol. 15. ICN-UCLA Symp. on Molec. and Cellular Biol. Eds. D. J. Cummings, P. Borst, I. B. Dawid, S. M. Weissman and C. F. Fox (Academic Press, New York), pp. 443-469.

Attardi, G., Costantino, P., England, J., Lynch, D., Ojala, D., Posakony, J. and Storrie, B. (1977) Mitochondrial and nuclear gene interaction in HeLa cells. In: Genetic Interaction and Gene Transfer. Brookhaven Symp. in Biol. **29**, 16-35.

Attardi, G. and Attardi, B. (1968) Mitochondrial origin of membrane-associated heterogeneous RNA in HeLa cells. Proc. Nat. Acad. Sci. USA **61**, 261-268.

Barath, Z. and Küntzel, H. (1972) Cooperation of mitochondrial and nuclear genes specifying the mitochondrial genetic apparatus in Neurospora crassa. Proc. Nat. Acad. Sci. USA **69**, 1371-1374.

Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J. Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1980) Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc. Nat. Acad. Sci. USA **77**, 3164-3166.

Barrell, B. G., Bankier, A. T. and Drouin, J. (1979) A different genetic code in human mitochondria. Nature **282**, 189-194.

Battey, J. and Clayton, D. A. (1978) The transcription map of mouse mitochondrial DNA. Cell **14**, 143-156.

Berk, A. J. and Clayton, D. A. (1974) Mechanism of mitochondrial DNA replication in mouse L-cells: asynchronous replication of strands, segregation of circular daughter molecules, aspects of topology and turnover of an initiation sequence. J. Mol. Biol. **86**, 801-824.

Berk, A. J., Sharp, P. A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**, 721-732.

Berk, A. J., Sharp, P. A. (1978) Spliced early mRNAs of simian virus 40. Proc. Nat. Acad. Sci. USA **75**, 1274-1278.

Bertazzoni, U. and Scovassi, A. I. (1980) Mitochondrial DNA polymerase from different eukaryotic cells. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Bogenhagen, D. and Clayton, D. A. (1977) Mouse L-cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. Cell **11**, 719-727.

Bogenhagen, D., Gillum, A. M., Martens, P. A. and Clayton, D. A. (1978) Replication of mouse L-cell mitochondrial DNA. Cold Spring Harbor Symp. Quant. Biol. **43**, 253-262.

Bonitz, S. G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F. G., Nobrega, M. P., Thalenfeld, B. E. and Tzagoloff, A. (1980) Codon recognition rules in yeast mitochondria. Proc. Nat. Acad. Sci. USA **77**, 3167-3170.

Borst, P. (1972) Mitochondrial nucleic acids. Ann. Rev. Biochem. **41**, 333-376.

Borst, P. (1977) Structure and function of mitochondrial DNA. Trends Biochem. Sci. **2**, 31-34.

Borst, P. and Grivell, L. A. (1978) The mitochondrial genome of yeast. Cell **15**, 705-723.

Brown, W. M., Ferris, S. D., George, M. and Wang, A. (1980) The evolution of vertebrate mitochondrial DNA: Frequency and distribution of base substitutions in the mitochondrial genome of mammals. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Brown, W. M., George, M. and Wilson, A. J. (1979) Rapid evolution of animal mitochondrial DNA. Proc. Nat. Acad. Sci. USA **76**, 1967-1971.

Brown, W. M. and Goodman, H. M. (1979) Quantitation of intrapopulation variation by restriction endonuclease analysis of human mitochondrial DNA. In: Extrachromosomal DNA, Vol. 15. ICN-UCLA Symp. on Molec. and Cellular Biol. Eds. D. J. Cummings, P. Borst, I. B. Dawid, S. M. Weissman and C. F. Fox (Academic Press, New York), pp. 485-499.

Brown, W. M., Shine, J. and Goodman, H. M. (1978) Human mitochondrial DNA: Analysis of 7S DNA from the origin of replication. Proc. Nat. Acad. Sci. USA **75**, 735-739.

Brown, W. M. and Vinograd, J. (1974) Restriction endonuclease cleavage maps of animal mitochondrial DNAs. Proc. Nat. Acad. Sci. USA **71**, 4617-4621.

Cantatore, P. and Attardi, G. (1980) Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA. Nucl. Acids Res. **8**, 2605-2624.

Castora, F. J., Brown, G. G. and Simpson, M.V. (1980) Two studies on mammalian mtDNA: Evolutionary aspects; enzymology of replication. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Ching, E. P. (1979) Biochemical and functional characterization of the products of mitochondrial protein synthesis in HeLa cells. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Church, G. M., Slonimski, P. P. and Gilbert, W. (1979) Pleiotropic mutations within two yeast mitochondrial cytochrome genes block mRNA processing. Cell **18**, 1209-1215.

Clayton, D. A., Davis, R. W. and Vinograd, J. (1970) Homology and structural relationships between the dimeric and monomeric circular forms of mitochondrial DNA from human leukemic leukocytes. J. Mol. Biol. **47**, 137-153.

Crews, S. Ojala, D., Posakony, J., Nishiguchi, J. and Attardi, G. (1979) Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. Nature **277**, 192-198.

Cummings, D. J., Borst, P., Dawid, I. B., Weissman, S. M. and Fox, C. F. (Eds.) (1979) In: Extrachromosomal DNA, Vol. 15. ICN-UCLA Symp. on Molec. and Cellular Biol. (Academic Press, New York).

Dawid, I. B. (1972) Mitochondrial RNA in Xenopus laevis. I. The expression of the mitochondrial genome. J. Mol. Biol. **63**, 201-216.

Dawid, I. B., Klukas, C. K., Ohi, S., Ramirez, J. L. and Upholt, W. B. (1976) Structure and evolution of animal mitochondrial DNA. In: The Genetic Function of Mitochondrial DNA. Eds. C. Saccone and A. M. Kroon (Elsevier/North-Holland Biomedical Press, Amsterdam, The Netherlands), pp. 3-13.

Dawid, I. B. and Rastl, E. (1979) Structure and Evolution of animal mitochondrial DNA. In: Extrachromosomal DNA. ICN-UCLA Symp. on Molec. and Cellular Biol. Eds. D. J. Cummings, P. Borst, I. B. Dawid, S. M. Weissman and C. F. Fox (Academic Press, New York), pp. 395-407.

Eperon, I. C., Anderson, S. and Nierlich, D. P. (1980) Distinctive sequences of human mitochondrial ribosomal RNA genes. Nature **286**, 460-467.

Francisco, J. F., Brown, G. G. and Simpson, M. V. (1979) Further studies on types A and B rat mtDNAs: Cleavage maps and evidence for cytoplasmic inheritance in mammals. Plasmid **2**, 426-436.

Gallerani, R and Saccone, C. (1974) The DNA-dependent RNA polymerase from rat liver mitochondria. In: The Biogenesis of Mitochondria. Eds. A. M. Kroon and C. Saccone (Academic Press, New York), pp. 59-69.

Gillum, A. and Clayton, D. A. (1978) Displacement-loop replication initiation sequence in animal mitochondrial DNA exists as a family of discrete lengths. Proc. Nat. Acad. Sci. USA **75**, 677-681.

Grohmann, K. Amalric, F., Crews, S. and Attardi, G. (1978) Failure to detect cap structures in mitochondrial DNA-coded poly(A)-containing RNA from HeLa cells. Nucl. Acids Res. **5**, 637-651.

Grossman, L. I., Watson, R. and Vinograd, J. (1973) The presence of ribonucleotides in mature closed-circular mitochondrial DNA. Proc. Nat. Acad. Sci. USA **70**, 3339-3343.

Hare, J. F., Ching, E. and Attardi, G. (1980) Isolation, subunit composition, and site of synthesis of human cytochrome c oxidase. Biochemistry **19**, 2023-2030.

Heckman, J. E., Sarnoff, J., Alzner-DeWeerd, B., Yin, S. and RajBhandary, U. L. (1980) Novel features in the genetic code and codon reading patterns in Neurospora crassa mitochondria based on sequences of six mitochondrial tRNAs. Proc. Nat. Acad. Sci. USA **77**, 3159-3163.

Hirsch, M. and Penman, P. (1973) Mitochondrial polyadenylic acid-containing RNA: localization and characterization. J. Mol. Biol. **80**, 379-391.

Kasamatsu, H., Robberson, D. L. and Vinograd, J. (1971) A novel closed-circular mitochondrial DNA with properties of a replicating intermediate. Proc. Nat. Acad. Sci. USA **68**, 2252-2257.

Kasamatsu, H. and Vinograd, J. (1974) Replication of circular DNA in eukaryotic cells. Ann. Rev. Biochem. **43**, 695-719.

Klukas, C. K. and Dawid, I. B. (1976) Characterization and mapping of mitochondrial ribosomal RNA and mitochondrial DNA in Drosophila melanogaster. Cell **9**, 615-625.

Koch, G. (1976) Synthesis of the mitochondrial inner membrane in cultured Xenopus laevis oocytes. J. Biol. Chem. **251**, 6097-6107.

Kroon, A. M., de Vos, W. M. and Bakker, H. (1978) The heterogeneity of rat liver mitochondrial DNA. Biochim. Biophys. Acta **519**, 269-273.

Kroon, A. M. and Saccone, C. (Eds.) (1980) In: The Organization and Expression of the Mitochondrial Genome (North-Holland, Amsterdam), in press.

Küntzel, H. and Shäfer, K. P. (1971) Mitochondrial RNA polymerase from Neurospora crassa. Nature New Biol. **231**, 265-269.

Levins, D., Lustig, A., Ticho, B., Synenki, R., Merten, S., Locker, J. and Rabinowitz, M. (1980) Transcription and processing of mtDNA. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Lynch, D. C. and Attardi, G. (1976) Amino acid specificity of the transfer RNA species coded for by HeLa cell mitochondrial DNA, J. Mol. Biol. **102**, 125-141.

Martens, P. A. and Clayton, D. A. (1979) Mechanism of mitochondrial DNA relication in mouse L-cells: Localization and sequence of the light-strand origin of replication. J. Mol. Biol. **135**, 327-351.

Murphy, W. I., Attardi, B., Tu, C. and Attardi, G. (1975) Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in HeLa cells. J. Mol. Biol. **99**, 809-814.

Nass, M. M. K. (1969) Mitochondrial DNA. II. Structure and physicochemical properties of isolated DNA. J. Mol. Biol. **42**, 529-545.

Nass, M. M. K. and Nass, S. (1963a) Intramitochondrial fibers with DNA characteristics. I. Fixation and electron staining reactions. J. Cell Biol. **19**, 593-611.

Nass, M. M. K., Nass, S. and Afzelius, B. A. (1965) The general occurrence of mitochondrial DNA. Exptl. Cell Res. **37**, 516-539.

Nass, S. and Nass, M. M. K. (1963b) Intramitochondrial fibers wtih DNA characteristics. II. Enzymatic and other hydrolytic treatments. J. Cell. Biol. **19**, 613-629.

Novitski, C. (1979) Aspects of regulation of mitochondrial DNA replication and transcription in mammalian cells. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Ohi, S., Ramirez, J. L., Upholt, W. D. and Dawid, I. B. (1978) Mapping of mitochondrial 4S RNA genes in Xenopus laevis by electron microscopy. J. Mol. Biol. **121**, 299-310.

Ojala, D. and Attardi, G. (1972) Expression of the mitochondrial genome in HeLa cells. X. Properties of mitochondrial polysomes. J. Mol. Biol. **65**, 273-289.

Ojala, D. and Attardi, G. (1974a) Expression of the mitochondrial genome in HeLa cells. XIX. Occurrence in mitochondria of polyadenylic acid sequences, "free" and covalently linked to mitochondrial DNA-coded RNA. J. Mol. Biol. **82**, 151-174.

Ojala, D. and Attardi, G. (1974b) Identification of discrete polyadenylate-containing RNA components transcribed from HeLa cell mitochondrial DNA. (XX. in the series, Expression of the Mitochondrial Genome in HeLa Cells.) Proc. Nat. Acad. Sci. USA **71**, 563-567.

Ojala, D. and Attardi, G. (1974c) Identification and partial characterization of multiple discrete polyadenylic acid-containing RNA compoments coded for by HeLa cell mitochondrial DNA. J. Mol. Biol. **88**, 205-219.

Ojala, D. and Attardi, G. (1977) A detailed physical map of HeLa cell mitochondrial DNA and its alignment with the positions of known genetic markers. Plasmid **1**, 78-105.

Ojala, D. and Attardi, G. (1978) Precise localization of the origin of replication in a physical map of HeLa cell mitochondrial DNA and isolation of a small fragment that contains it. J. Mol. Biol. **122**, 301-319.

Ojala, D. and Attardi, G. (1980) Fine mapping of the ribosomal RNA genes of HeLa cell mitochondrial DNA. J. Mol. Biol. **138**, 411-420.

Ojala, D., Merkel, C., Gelfand, R. and Attardi, G. (1980a) The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell, in press.

Ojala, D., Montoya, J. and Attardi, G. (1980b) The putative mRNA for subunit II of cytochrome c oxidase in human mitochondria starts directly at the translation initiator codon. Nature, in press.

Perlman, S., Abelson, H. T. and Penman, S. (1973) Mitochondrial protein synthesis: RNA with the properties of eukaryotic messenger RNA. Proc. Nat. Acad. Sci. USA **70**, 350-353.

Perlman, S. and Penman, S. (1970) Protein synthesizing structures associated with mitochondria. Nature **227**, 133-137.

Pica-Mattoccia, L. and Attardi, G. (1971) Expression of the mitochondrial genome in HeLa cells. V. Transcription of mitochondrial DNA in relationship to the cell cycle. J. Mol. Biol. **57**, 615-621.

Pica-Mattoccia, L. and Attardi, G. (1972) Expression of the mitochondrial genome in HeLa cells. IX. Replication on mitochondrial DNA in Relationship to the cell cycle in HeLa cells. J. Mol. Biol. **64**, 465-484.

Potter, S. S., Newbold, J. E., Hutchison, C. A. III and Edgell, M. H. (1975) Specific cleavage analysis of mammalian mitochondrial DNA. Proc. Nat. Acad. Sci. USA **72**, 4496-4500.

Radloff, R., Bauer, W. and Vinograd, J. (1967) A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in HeLa cells. Proc. Nat. Acad. Sci. USA **57**, 1514-1521.

Ramirez, J. L. and Dawid, I. B. (1978) Mapping of mitochondrial DNA in Xenopus laevis and X. borealis: The positions of ribosomal genes and D-loops. J. Mol. Biol. **119**, 133-146.

Rascati, R. J. and Parsons, P. (1979) Biosynthesis of cytochrome c oxidase by isolated rat liver mitochondria. J. Biol. Chem. **254**, 1594-1599.

Rastl, E. and Dawid, I. B. (1979) Expression of the mitochondrial genome in Xenopus laevis: A map of transcripts. Cell **18**, 501-510.

Rinaldi, A. M., Salcher-Cillari, I., Sollazo, M. and Mutolo, V. (1980) Regulation of mitochondrial genomic activity in sea urchin eggs. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Robberson, D., Aloni, Y., Attardi, G. and Davidson, N. (1972) Expression of the mitochondrial genome in HeLa cells. VIII. The relative position of ribosomal RNA genes in mitochondrial DNA. J. Mol. Biol. **64**, 313-317.

Robberson, D. L., Kasamatsu, H. and Vinograd. J. (1972) Replication of mitochondrial DNA. Circular replicative intermediates in mouse L-cells. Proc. Nat. Acad. Sci. USA **69**, 737-741.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol., in press.

Schatz, G. and Mason, T. L. (1974) The biosynthesis of mitochondrial proteins. Ann. Rev. Biochem. **43**, 51-87.

Scovassi, A. I., Wicker, R. and Bertazzoni, A. (1979) A phylogenetic study on vertebrate mitochondrial DNA polymerase. Eur. J. Biochem. **100**, 491-496.

Upholt, W. B. and Dawid, I. B. (1977) Mapping of mitochondrial DNA of individual sheep and goats: Rapid evolution in the D-loop region. Cell **11**, 571-583.

Van Ommen, G.-J. B., Groot, G. S. P. and Grivell, L. A. (1979) Transcription maps of mtDNAs of two strains of saccharomyces: Transcription of strain-specific insertions; complex RNA maturation and splicing. Cell **18**, 511-523.

Walker, J. E.,Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Runswick, M. J., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1980) In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Wu, G.-J. and Dawid, I. B. (1972) Purification and properties of mitochondrial deoxyribonucleic acid dependent ribonucleic acid polymerase from ovaries of Xenopus laevis. Biochemistry **11**, 3589-3595.

Wu, M., Davidson, N., Attardi, G. and Aloni, Y. (1972) Expression of the mitochondrial genome in HeLa cells. XIV. The relative positions of the 4S RNA genes and of the ribosomal RNA genes in mitochondrial DNA. J. Mol. Biol. **71**, 81-93.

Zylber, E., Vesco, C. and Penman, S. (1969) Selective inhibition of the synthesis of mitochondria-associated RNA by ethidium bromide. J. Mol. Biol. **44**, 195-204.

# Chapter 2

## Fine Mapping of the Ribosomal RNA Genes of HeLa Cell Mitochondrial DNA

Reprinted from *J. Mol. Biol.* (1980) **138**, 411–420

# Fine Mapping of the Ribosomal RNA Genes of HeLa Cell Mitochondrial DNA

DEANNA OJALA AND GIUSEPPE ATTARDI

*J. Mol. Biol.* (1980) **138**, 411–420

# Fine Mapping of the Ribosomal RNA Genes of HeLa Cell Mitochondrial DNA

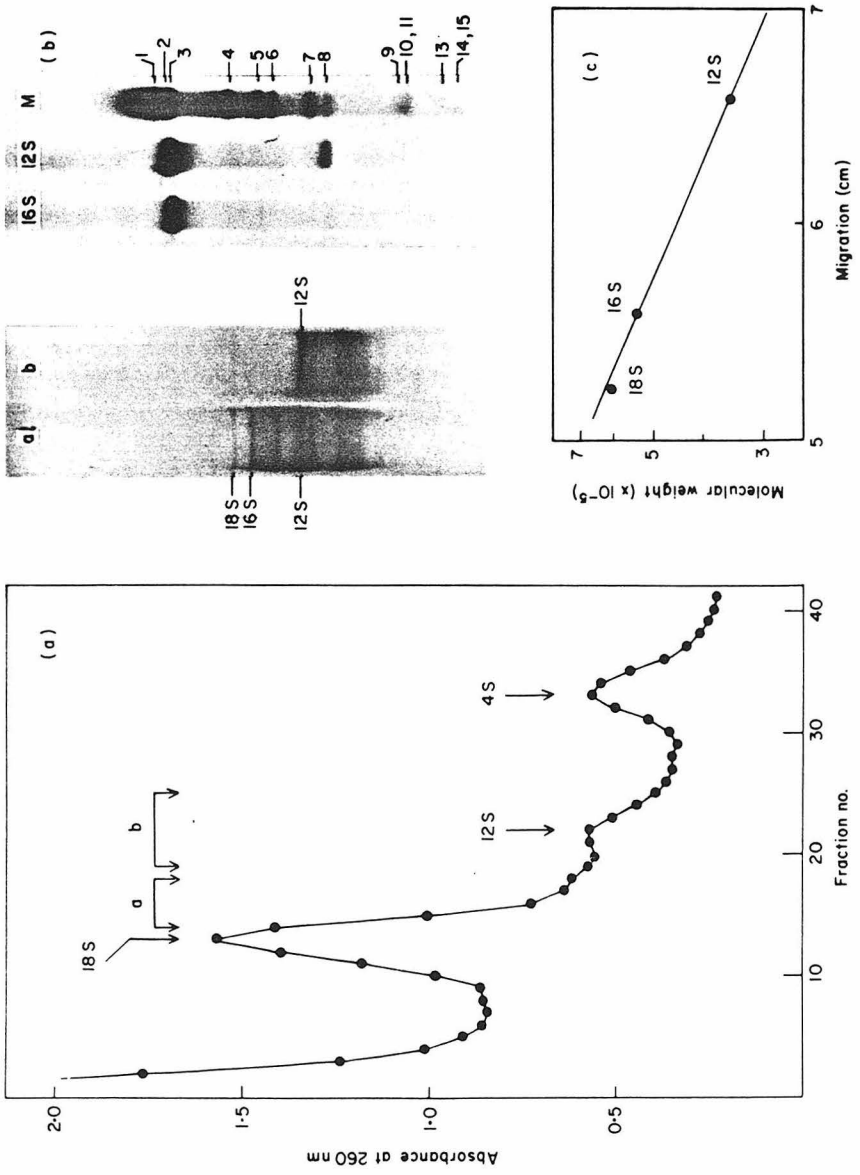
A fine mapping study of the ribosomal RNA region of HeLa cell mitochondrial DNA has been carried out by using as an approach the protection by hybridized 12 S and 16 S rRNA of the complementary sequences in DNA against digestion with the single strand-specific *Aspergillus* nuclease $S_1$ or *Escherichia coli* exonuclease VII. No inserts have been detected in the main body of the 12 S and 16 S rRNA cistrons, in contrast to the situation described in the large mitochondrial ribosomal RNA gene of some strains of yeast and of *Neurospora crassa*. Furthermore, it has been possible to assign more precisely than previously the positions of the 5′ and 3′-ends of the 12 S rRNA and 16 S rRNA genes in the *Hpa*II restriction map of HeLa cell mitochondrial DNA.

Recent work in several laboratories has revealed that the gene for the large ribosomal RNA species in mitochondrial DNA from some strains of yeast (Bos *et al.*, 1978) and from *Neurospora crassa* (Hahn *et al.*, 1979; Heckman & RajBhandary, 1979; Mannella *et al.*, 1979) contains an intervening sequence 1·1 to 1·2 and 2·0 to 2·5 $\times 10^3$ bases long, respectively. These surprising observations have indicated that the mitochondrial system shares with other eukaryotic systems the occurrence of discontinuous genes and the involvement of splicing mechanisms in the processing of the primary transcripts. In animal cell mitochondrial DNA, previous electron microscope studies of RNA–DNA hybrids involving rRNA species (Wu *et al.*, 1972; Angerer *et al.*, 1976; Klukas & Dawid, 1976; Ramirez & Dawid, 1978) had not revealed the presence of obvious discontinuities in the rRNA genes. However, the resolution of this analysis was not such that would have detected discontinuities smaller than 100 nucleotides. And, indeed, there is evidence that some of the known intervening sequences in eukaryotic genes are only 15 to 20 base-pairs long (Goodman *et al.*, 1977; Valenzuela *et al.*, 1978). We have therefore carried out experiments aimed at obtaining information on the possible existence of inserts in the rRNA genes of HeLa cell mitochondrial DNA. For this purpose, we have used the methodology developed by Berk & Sharp (1978): this involves trimming of RNA–DNA hybrids (formed with $^{32}$P-labeled DNA) with the single strand-specific *Aspergillus* nuclease $S_1$ or *Escherichia coli* exonuclease VII and size analysis of the hybridized DNA sequences in denaturing gels. The $S_1$ enzyme is known to have endonuclease activity and is therefore expected to digest any terminal or internal single-stranded DNA segment in the hybrid (Vogt, 1973); in contrast, the exonuclease VII lacks endonuclease activity and thus would only trim non-hybridized DNA segments at the 5′ and 3′-ends of the hybrid (Chase & Richardson, 1974).

In order to isolate the two mitochondrial rRNA species, 12 S and 16 S RNA, the oligo(dT)-cellulose unbound fraction of mitochondrial polysomal RNA (Amalric *et al.*, 1978) was fractionated in a sucrose gradient. As shown in Figure 1(a), the 12 S



0022–2836/80/100411–10 $02.00/0 


(a)
Absorbance at 260 nm
2·0
1·5
1·0
0·5
18S
a
b
12S
4S
10
20
30
40
Fraction no.
(b)
M
12S
16S
a
b
12S
18S
16S
12S
1
2
3
4
5
6
7
8
9
10, 11
13
14, 15
(c)
Molecular weight (x 10⁻⁵)
7
5
3
18S
16S
12S
5
6
7
Migration (cm)

rRNA formed a well recognizable peak in the profile, while the 16 S rRNA was not resolved from the contaminating cytoplasmic 18 S rRNA. The material corresponding to the 16 S and 12 S region of the gradient (indicated by arrows in Fig. 1(a)) was collected by precipitation with ethanol and further fractionated by electrophoresis through an agarose/$CH_3HgOH$ slab gel (Bailey & Davidson, 1976) (Fig. 1(b), left panel). The material from the 12 S cut showed in the ethidium bromide-stained gel a main band presumably corresponding to mitochondrial 12 S rRNA, while the material from the 16 S cut showed, besides a 12 S rRNA band, several slower moving components. Of these, two were tentatively identified as mitochondrial 16 S rRNA and contaminating cytoplasmic 18 S rRNA as the only two which had, relative to 12 S rRNA, the mobility expected from their molecular weight (Fig. 1(c)). To validate this identification, a small portion of the presumptive 12 S and 16 S rRNA components was labeled *in vitro*, after mild alkaline hydrolysis, with [$\gamma$-$^{32}$P]ATP and T4 polynucleotide kinase, and then hybridized with *Hpa*II digests of HeLa mitochondrial DNA transferred on to a nitrocellulose filter (Southern, 1975). As shown in Figure 1(b) (right panel), the presumptive 12 S rRNA hybridized with both *Hpa*II fragments 3 and 8, while the presumptive 16 S rRNA hybridized only with fragment 3. Apart

---

FIG. 1. Isolation of mitochondrial 16 S and 12 S rRNAs. (a) RNA was extracted from mitochondrial polysomal structures of unlabeled HeLa cells, as previously described (Amalric *et al.*, 1978), and passed through an oligo(dT)-cellulose column. The unbound fraction was collected by precipitation with ethanol and run through a 15% to 30% gradient in 0·01 M-Tris buffer (pH 7·0, 25°C), 0·1 M-NaCl, 0·001 M-EDTA, 0·5% sodium dodecyl sulfate. The components in the profile corresponding to the 16 S region of the gradient (labeled a) and to the 12 S peak (labeled b) (as indicated by arrows) were collected by precipitation with ethanol and centrifugation. (b) Left panel: the samples labeled a and b from (a) were run through an agarose (1·4%)/$CH_3HgOH$ gel at 80 V (5 V/cm) for 7·5 h. After soaking the gel for 20 min in 0·02 M-dithiothreitol, and staining it with 2 $\mu$g ethidium bromide/ml, the bands were visualized by longwave ultraviolet irradiation. The bands indicated as 16 S and 12 S were cut out and the RNA was eluted from the gel slices by electrophoresis. Right panel: a small portion of the eluted material from the 16 S and 12 S bands was subjected to mild hydrolysis by heating for 30 min at 65°C in 0·06 M-sodium borate (pH 9·2); after addition of 0·01 M-dithiothreitol and 0·006 M-$MgCl_2$, the 5′-ends of the RNA fragments were labeled by treatment with 10 $\mu$m [$\gamma$-$^{32}$P]ATP (12 mCi/$\mu$mol) and 4 units of T4 polynucleotide kinase (Boehringer & Mannheim) (added twice with a 30-min interval) in a vol. of 50 $\mu$l at 37°C for 60 min. After addition of Tris (pH 7·4) to 0·5 M, the samples were extracted with phenol, and each aqueous phase run through a Sephadex G50 column (0·6 cm × 40 cm) equilibrated with 0·1 M-NaCl, 0·01 M-Tris (pH 7·4), 0·001 M-EDTA. The material in the excluded volume was collected by precipitation with ethanol. Two 1·5-$\mu$g samples of HeLa cell mitochondrial DNA were digested with the *Hpa*II restriction enzyme, run through a 2% agarose slab gel in Tris-acetate buffer (0·01 M-Tris, pH 7·4, 0·05 M-sodium acetate, 0·0025 M-EDTA) and transferred to nitrocellulose paper (Schleicher & Schuell) following the procedure of Southern (1975). The 2 strips were preincubated with 2 ml of 6 × SSC (SSC is 0·15 M-NaCl, 0·015 M-sodium citrate, pH 7·0), 0·02% bovine serum albumin, 0·02% Ficoll (Sigma), 0·5% sodium dodecyl sulfate in the presence of 20 $\mu$g yeast tRNA/ml for 5 h at 68°C; after addition of 80,000 cts/min of the 16 S probe or 280,000 cts/min of the 12 S probe, the two strips were incubated further for 20 h at 68°C. After washing twice for 15 min at 68°C with 6 × SSC, 0·02% bovine serum albumin, 0·02% Ficoll, 0·5% sodium dodecyl sulfate, the filters were dried and autoradiographed using a Du Pont screen intensifier. The lane designated M shows a sample of *in vitro*-labeled mitochondrial DNA *Hpa*II digest (Ojala & Attardi, 1977) run through an agarose gel and transferred to a nitrocellulose filter in parallel with the unlabeled *Hpa*II digests. (c) Plot of molecular sizes of presumptive 12 S, 16 S, and 18 S RNA components in the gel pattern shown in (b) (left panel) *versus* migration. Molecular weights of 6·05 × $10^5$, 5·4 × $10^5$ and 3·5 × $10^5$ were used, respectively, for 18 S cytoplasmic rRNA (Spohr *et al.*, 1976), 16 S and 12 S mitochondrial rRNA (Robberson *et al.*, 1971).

from the absence of hybridization of 16 S rRNA with *Hpa*II fragment 4, which will be discussed further below, these results are in full agreement with the assignment in the *Hpa*II physical map of the sequences coding for the two rRNA species (Fig. 2).

The two purified rRNA species were used for experiments of hybridization with either the heavy (H) strand (which is their coding strand (Aloni & Attardi, 1971)) of total mitochondrial DNA or with restriction fragments of this DNA or their H strands.

Figure 3 shows the autoradiograms, after polyacrylamide/urea gel electrophoresis, of mitochondrial DNA segments protected by mitochondrial 12 S rRNA in the hybrids against digestion with either $S_1$ nuclease or exonuclease VII. In one set of experiments, 12 S rRNA was hybridized either to *Hin*dIII fragment 1 (see map in Fig. 2) at 49°C in high formamide (under conditions favoring RNA–DNA hybridization over DNA–DNA reassociation (Casey & Davidson, 1977)), or to total mitochondrial DNA H strand at 66°C in 0·4 M-salt. Under the two experimental conditions, the hybridized 12 S rRNA protected from both $S_1$ and exonuclease VII digestion a

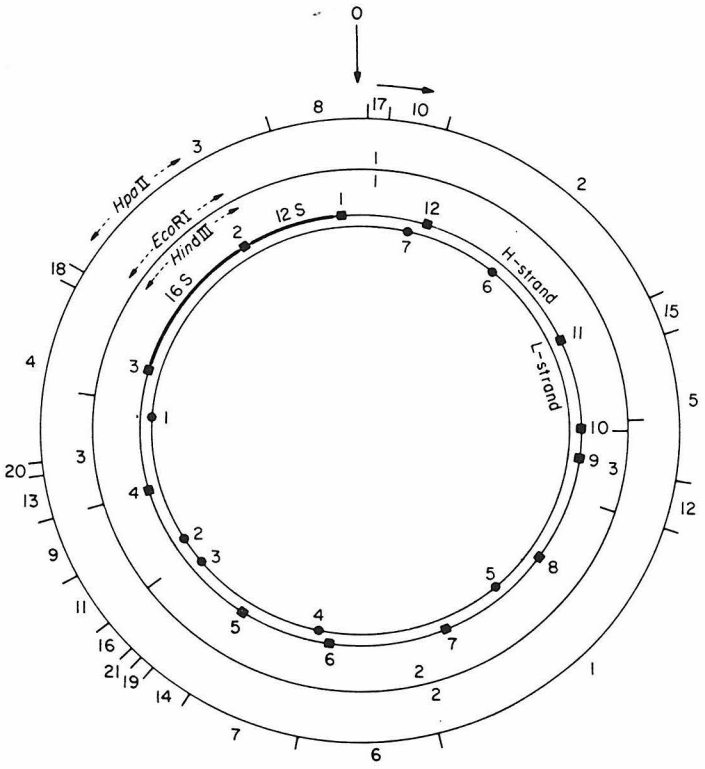
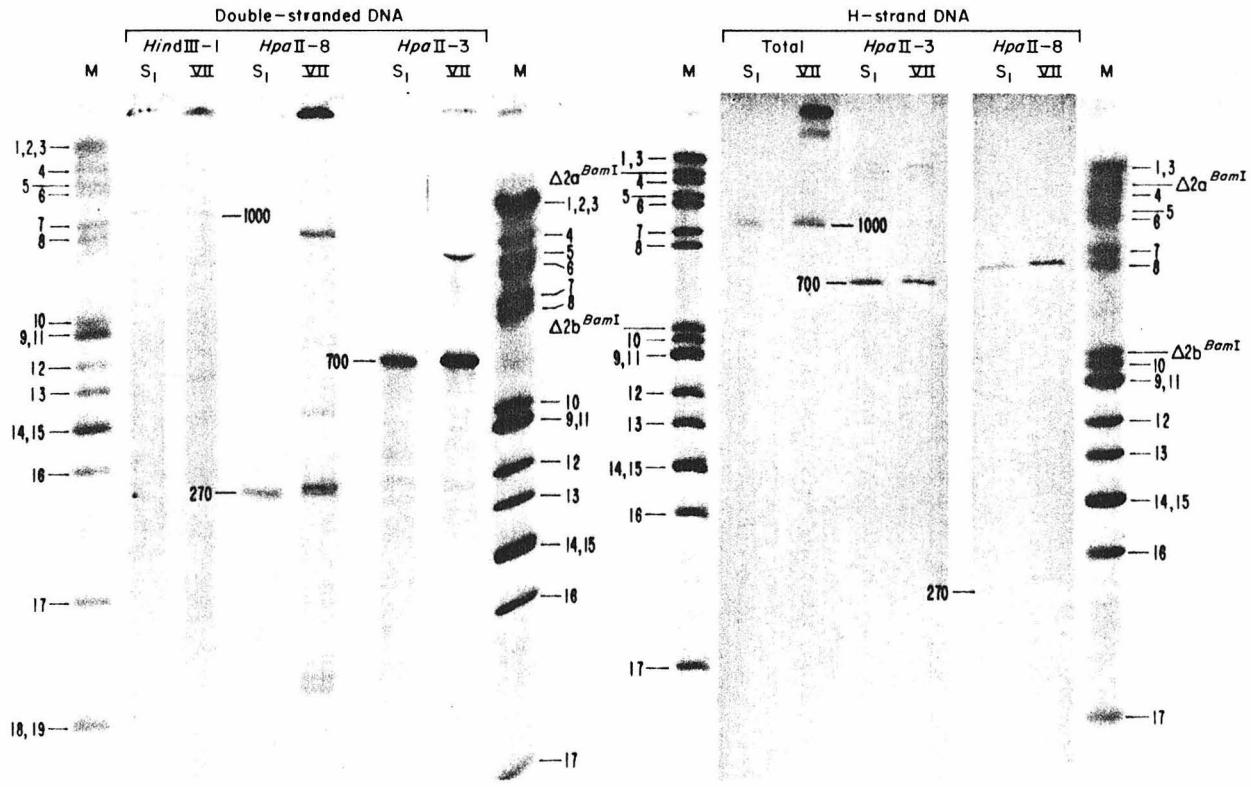



FIG. 2. Physical and genetic maps of HeLa cell mitochondrial DNA. The restriction endonuclease map determined with the enzyme *Hpa*II (Ojala & Attardi, 1977) has been aligned with the *Eco*RI and *Hin*dIII maps (Brown & Vinograd, 1974) and with the map of the positions of the complementary sequences for 12 S and 16 S rRNA on the H strand and for 4 S RNAs on the H (■) and L strands (●) (modified from Ojala & Attardi, 1977). The vertical arrow (marked O) indicates the location of the origin of mitochondrial DNA replication and the rightward arrow the direction of H strand synthesis (Crews *et al.*, 1979).

DNA segment about 1000 nucleotides long. Since the size of 12 S rRNA previously estimated by electron microscopy is about 1100 nucleotides (Robberson *et al.*, 1971), these results indicated that the main body of the 12 S rRNA gene lacks intervening sequences susceptible to $S_1$ digestion. Analysis of the protected DNA segments after hybridization of 12 S rRNA to either the whole *Hpa*II fragment 3 (under the high formamide conditions) or to the H strand of this fragment (at 66°C in 0·4 M-salt) showed a band corresponding to a 700-nucleotide pair long segment after either $S_1$ or exonuclease VII digestion (Fig. 3). A similarity in behavior with respect to $S_1$ and exonuclease VII digestion was likewise observed with hybrids formed, under the high formamide conditions, between 12 S rRNA and *Hpa*II fragment 8. As one can see in Figure 3, after $S_1$ digestion, there is a band corresponding to a 270-nucleotide pair long segment; after exonuclease VII digestion, in addition to this band, there is another one, somewhat more pronounced, moving slightly slower and corresponding to an approximately 275-nucleotide pair long segment. It seems likely that the latter band represents the protected DNA segment incompletely trimmed by the exonuclease VII at one or both ends due to some structural peculiarity. When the same 12 S rRNA preparation was hybridized at 66°C in 0·4 M-salt to the separated H strand of the *Hpa*II fragment 8, a segment of about 270 nucleotides was protected from both $S_1$ and exonuclease VII digestion, with a hint that the protected segment was slightly longer in the case of exonuclease VII. Results identical to those just described, in terms of size of the protected segments, were obtained when hybrids formed, under the high formamide conditions, between 12 S rRNA and *Hpa*II fragment 3 or 8 were analyzed, after $S_1$ digestion, in the native state in a 5% polyacrylamide gel in Tris/borate/$Mg^{2+}$ buffer (Maniatis *et al.*, 1975; not shown).

The results described above are represented diagrammatically in Figure 4. The conclusion of these experiments that the 12 S rRNA coding sequence extends to a position in fragment 8 which is about 270 nucleotide pairs from the border with fragment 3 is in excellent agreement with recent sequencing data, to be reported elsewhere, which have located the 5′-end of 12 S rRNA in correspondence to a residue in the H strand at 286 nucleotides from the *Hpa*II cutting site between fragments 3 and 8 (Crews & Attardi, 1980).

A similar analysis to that described above for 12 S rRNA was performed with 16 S rRNA. As shown in Figure 5, after hybridization with the H strand of total mitochondrial DNA, a fragment about 1600 nucleotide pairs long appears to be protected from both $S_1$ and exonuclease VII digestion. This size is only slightly smaller than that estimated by electron microscopy for 16 S rRNA (about 1700 nucleotides) (Robberson *et al.*, 1971), again pointing to the lack of intervening sequences in the main body of the 16 S rRNA gene. Hybridization of 16 S rRNA to the H strand of *Hpa*II fragment 3 protects a 1350-nucleotide long fragment from both $S_1$ and exonuclease VII digestion. By contrast, no protected fragment was observed after hybridization of 16 S rRNA to the H strand of *Hpa*II fragment 4. In previously reported experiments (Ojala & Attardi, 1977), no or only marginal hybridization of 16 S rRNA with *Hpa*II fragment 4 had been detected; this was contrary to what would have been expected on the basis of the ratio of observed hybridization of 16 S and 12 S rRNA with *Hpa*II fragment 3 and of the electron microscopic estimates of the sizes of 16 S and 12 S rRNA (Robberson *et al.*, 1971) and of *Hpa*II fragment 3


Double-stranded DNA
H-strand DNA
HindIII-1
HpaII-8
HpaII-3
Total
M
S1
VII
1000
700
270
Δ2a BamI
Δ2b BamI
1,2,3
4
5
6
7
8
10
9,11
12
13
14,15
16
17
18,19

(Angerer *et al.*, 1976). For this reason, the previous map assignment of the sequences coding for 16 S rRNA (see Fig. 2) had been given as tentative (Ojala & Attardi, 1977). The presently reported mapping data confirm the lack of detectable hybridization of 16 S rRNA with fragment 4, strongly suggesting that the sequence coding for 16 S rRNA extends into this fragment very little, if at all. From the length of the segment of 16 S rRNA corresponding to *Hpa*II fragment 3 (about 1350 nucleotides) and from the sizes of *Hpa*II fragment 18 (about 150 nucleotide pairs) (Ojala & Attardi, 1977) and of 16 S rRNA (1600 to 1700 nucleotides) one can estimate that a stretch of at most 100 to 200 nucleotides of 16 S rRNA could have its coding sequence in fragment 4. A further implication of these findings is that the size of *Hpa*II fragment 3 previously determined by electron microscopy (Angerer *et al.*, 1976) had been underestimated. The present estimate, based on a length of 1350 and 700 nucleotide pairs for the portions of the 16 S and 12 S rRNA genes, respectively, which are included in this fragment, and on the presence of a 160-nucleotide pair spacer between the two genes (Wu *et al.*, 1972), would be ~2200 nucleotide pairs, i.e. 250 nucleotide pairs larger than the previous value (1950 nucleotide pairs). Recent data of sequential digestion with *Hinc*II of *Hpa*II fragment 3 are in agreement with this conclusion (Ojala, Shaffer & Baskir, unpublished observations). The results described above are summarized in Figure 4.

---

FIG. 3. Nuclease $S_1$ and exonuclease VII sizing and mapping of mitochondrial 12 S rRNA. *Hin*dIII or *Hpa*II restriction fragments of HeLa cell mitochondrial DNA were labeled *in vitro* by nick translation according to the procedure of Rigby *et al.* (1977), modified as will be described elsewhere (Cantatore & Attardi, manuscript in preparation). Total H strand mitochondrial DNA was obtained by alkaline cesium chloride centrifugation of closed circular mitochondrial DNA labeled *in vivo* with [$^{32}$P]orthophosphate as previously described (Ojala & Attardi, 1977). Separation of the strands of individual *in vivo*-labeled *Hpa*II fragments was obtained by brief alkali treatment and electrophoresis through a 1·4% agarose gel in Tris/phosphate buffer (in the case of fragment 3) or a 4% polyacrylamide gel in Tris/borate/EDTA buffer (in the case of fragment 8), as will be described in detail elsewhere (Cantatore & Attardi, manuscript in preparation); the H strands of the individual fragments, previously identified by hybridization with separated total H and L strands (cited reference), were eluted from the crushed gel slices by diffusion. Samples of gel-purified 12 S rRNA (~50 ng) were dried down in snap cap vials together with *in vitro*-labeled *Hin*dIII-1 or *Hpa*II-3 or *Hpa*II-8 double-stranded fragments (20 to 70 ng), or with *in vivo*-labeled total H strand (~250 ng), or with the H strand of *in vivo*-labeled *Hpa*II fragment 3 (~25 ng) or fragment 8 (~10 ng). For the hybridizations in high formamide (involving either double-stranded DNA fragments or *Hpa*II-8 H strand), the dried samples were dissolved in 20 μl of 0·04 M-PIPES (pH 7·0), 0·75 M-NaCl, 0·001 M-EDTA, 80% formamide, heated for 5 min at 68°C and incubated for 8 to 20 h at 49°C; the samples were then diluted with 10 vol. of 0·01 M-Tris (pH 7·4), 0·03 M-KCl, 0·01 M-EDTA, 0·01 M-dithiothreitol, and incubated for 60 min at 45°C with *E. coli* exonuclease VII (an amount sufficient to digest completely 25 ng of single-stranded DNA, added twice with a 30-min interval). The samples were then precipitated with ethanol in the presence of 5 μg yeast tRNA, dissolved in 7 M-urea, 0·001 M-Tris (pH 7·4), 0·001 M-EDTA, heated for 3 min at 80°C and fast cooled, and immediately layered over a 4% polyacrylamide (1 : 20 bisacrylamide)/7 M-urea slab gel in Tris/borate/EDTA buffer (Maniatis *et al.*, 1975). Electrophoresis was carried out at 5 V/cm for 12 to 15 h. After rinsing for 15 min in water to remove the urea, the gels were dried and autoradiographed for 2 to 4 days using a screen intensifier. For the hybridizations at 66°C in 0·4 M-salt (involving the H strand of total mitochondrial DNA or of *Hpa*II-3), the dried samples were dissolved in 20 μl of 0·4 M-NaCl, 0·01 M-Tris (pH 8·0), 0·01 M-EDTA, heated for 4 min at 90°C, fast cooled and incubated for 4 h at 66°C; the samples were then diluted with 10 vol. of 0·03 M-sodium acetate (pH 4·6), 0·25 M-NaCl, 0·001 M-$ZnSO_4$ and incubated with 100 units *Aspergillus orizae* $S_1$ nuclease/ml (Sigma) for 30 min at 45°C in the presence of 20 μg denatured HeLa nuclear DNA/ml. Subsequent treatment of these samples was identical to that of the exonuclease VII-digested samples.

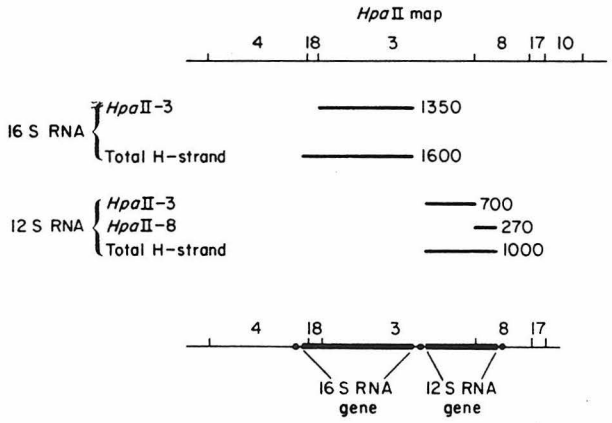



FIG. 4. Diagrammatic representation of the results shown in Figs 3 and 5.

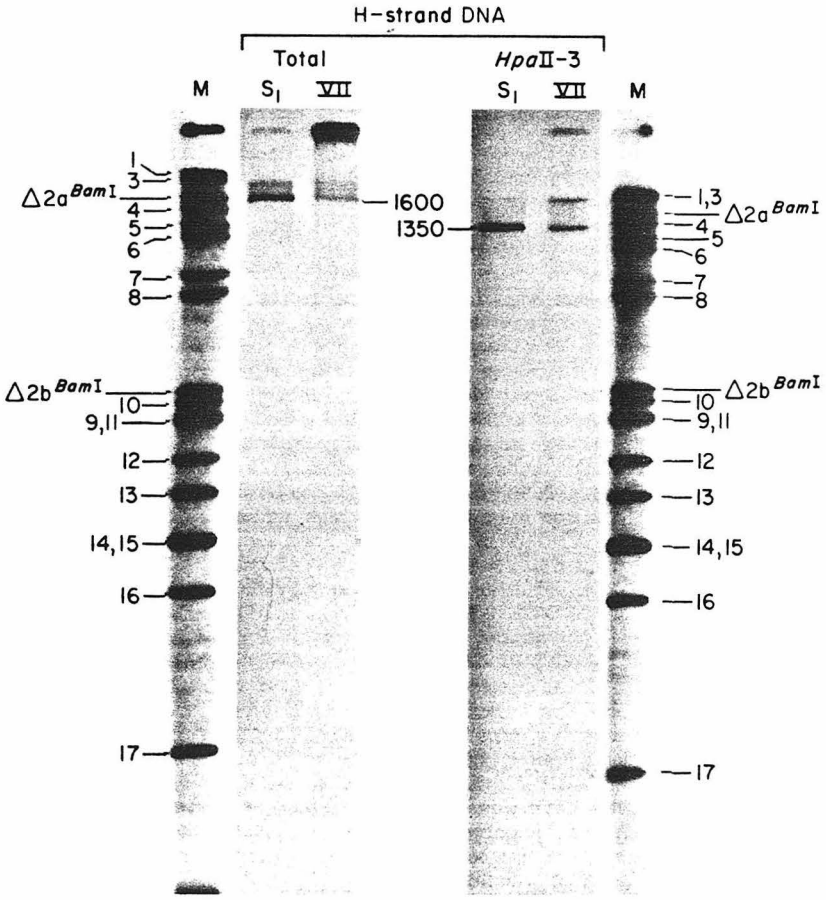



FIG. 5. Nuclease $S_1$ and exonuclease VII sizing and mapping of mitochondrial 16 S rRNA. Experimental details as described in the legend to Fig. 3.

The main conclusion of the experiments described here concerning the absence of inserts in the two mitochondrial rRNA genes of HeLa cells needs to be partially qualified. The experimental approach used here would have not been adequate to reveal the presence of small inserts very near to the ends of the two rRNA genes. In particular, the occurrence of 4 S RNA sites in DNA very close to the 5′-end and 3′-end of both the 12 S and 16 S rRNA genes (Fig. 2) raises the possibility, though unlikely, of the presence in one or both of the two rRNA genes of subterminal DNA inserts, containing a transfer RNA sequence, destined to be excised during maturation of the rRNAs. The only unambiguous approach to analyze the relationship of the 5′-end and 3′-end of the 12 S and 16 S rRNAs to the flanking 4 S RNA coding sequences is the sequencing approach. Recently, by aligning the 5′-end proximal 71 nucleotides of 12 S RNA with the DNA coding sequence in *Hpa*II fragment 8, it has been possible to show that the 12 S RNA sequence and the DNA sequence are co-linear, excluding therefore, in this case, the possibility of a subterminal insert (Crews & Attardi, 1980). A similar type of analysis remains to be done for the 3′-end of the 12 S rRNA and both ends of the 16 S rRNA.

With the above qualifications in mind, the absence of inserts in the large mitochondrial rRNA gene in HeLa cells is in contrast to the well-demonstrated occurrence of an intervening sequence in the homologous gene in some strains of yeast (Bos *et al.*, 1978) and in *Neurospora* (Hahn *et al.*, 1979; Heckman & RajBhandary, 1979; Mannella *et al.*, 1979). This difference supports the evidence which is emerging from studies carried out in various systems, indicating that gene inserts may be present or absent in homologous genes in different organisms, even of the same species, without any obvious effect on their expression. Furthermore, it adds another trait to the previously recognized unique features (complete symmetrical transcription, scattered distribution of tRNA genes), which drastically differentiate the mitochondrial gene organization and expression of human cells (and in general, presumably, of all animal cells) from those of lower eukaryotic cells, in particular yeast and *Neurospora* (Attardi *et al.*, 1979).

These investigations were supported by a grant (GM-11726) from the National Institutes of Health. The authors are grateful to Dr Paul Berg for a gift of *E. coli* exonuclease VII and to Dr Antonio Montalvo for a gift of mitochondrial RNA. The assistance of Ms Arger Drew and Ms Doris Finch is thankfully acknowledged.

Division of Biology
California Institute of Technology
Pasadena, Calif. 91125, U.S.A.

DEANNA OJALA
GIUSEPPE ATTARDI

Received 21 September 1979

## REFERENCES

Aloni, Y. & Attardi, G. (1971). *J. Mol. Biol.* **55**, 271–276.
Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. (1978). *J. Mol. Biol.* **118**, 1–25.
Angerer, L., Davidson, N., Murphy, W., Lynch, D. & Attardi, G. (1976). *Cell*, **9**, 81–90.
Attardi, A., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C. & Ojala, D. (1979). *ICN-UCLA Symposia on Molecular and Cellular Biology, Extrachromosomal DNA* (D. Cummings *et al.*, eds), vol. 15, pp. 443–469, Academic Press, New York.
Bailey, J. M. & Davidson, N. (1976). *Anal. Biochem.* **70**, 75–85.

Berk, A. J. & Sharp, P. A. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 1274–1278.
Bos, J. L., Heyting, C., Borst, P., Arnberg, A. C. & Van Bruggen, E. F. J. (1978). *Nature (London)*, **275**, 336–338.
Brown, W. M. & Vinograd, J. (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 4617–4621.
Casey, J. & Davidson, N. (1977). *Nucl. Acids Res.* **4**, 1539–1552.
Chase, J. W. & Richardson, C. C. (1974). *J. Biol. Chem.* **249**, 4553–4561.
Crews, S. & Attardi, G. (1980). *Cell*, in the press.
Crews, S., Ojala, D., Posakony, J., Nishiguchi, J. & Attardi, G. (1979). *Nature (London)*, **277**, 192–198.
Goodman, H. M., Olson, M. V. & Hall, B. D. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5453–5457.
Hahn, U., Lazarus, C. M., Lunsdorf, H. & Küntzel, H. (1979). *Cell*, **17**, 191–200.
Heckman, J. E. & RajBhandary, U. L. (1979). *Cell*, **17**, 583–595.
Klukas, C. K. & Dawid, I. B. (1976). *Cell*, **9**, 615–625.
Maniatis, T., Jeffrey, A. & van de Sande, H. (1975). *Biochemistry*, **14**, 3787–3794.
Mannella, C. A., Collins, R. A., Green, M. R. & Lambowitz, A. M. (1979). *Proc. Nat. Acad. Sci., U.S.A.* **76**, 2635–2639.
Ojala, D. & Attardi, G. (1977). *Plasmid*, **1**, 78–105.
Ramirez, J. L. & Dawid, I. B. (1978). *J. Mol. Biol.* **119**, 133–146.
Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977). *J. Mol. Biol.* **113**, 237–251.
Robberson, D., Aloni, Y., Attardi, G. & Davidson, N. (1971). *J. Mol. Biol.* **60**, 473–484.
Southern, E. M. (1975). *J. Mol. Biol.* **98**, 503–517.
Spohr, G., Mirault, M.-E., Imaizumi, T. & Scherrer, K. (1976). *Eur. J. Biochem.* **62**, 313–322.
Valenzuela, P., Venegas, A., Weinberg, F. & Bishop, R. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 190–194.
Vogt, V. M. (1973). *Eur. J. Biochem.* **33**, 192–200.
Wu, M., Davidson, N., Attardi, G. & Aloni, Y. (1972). *J. Mol. Biol.* **71**, 81–93.

# Chapter 3

## The tRNA Genes Punctuate the Reading of Genetic Information in Human Mitochondrial DNA

# THE tRNA GENES PUNCTUATE THE READING OF GENETIC INFORMATION IN HUMAN MITOCHONDRIAL DNA

Deanna Ojala, Christian Merkel, Robert Gelfand and Giuseppe Attardi

Division of Biology, California Institute of Technology

Pasadena, California 91125

Running Title: tRNA Gene Punctuation of Mitochondrial DNA Information

Summary

A detailed transcription map of HeLa cell mitochondrial DNA (mtDNA) has been constructed by precisely localizing on the physical map of this DNA by the S1 protection technique the sequences coding for the ribosomal RNA (rRNA) and poly(A)-containing RNA species. This transcription map has been correlated with the positions of the tRNA genes derived from the mtDNA sequence. It has been shown that, with the exception of the D-loop and another small segment near the origin of replication, the mtDNA sequences are completely saturated by the rRNAs, poly(A)-containing RNAs and tRNAs coded for by the two strands. No evidence for intervening sequences has been found. The sequences coding for the individual poly(A)-containing RNA and rRNA species appear to be immediately contiguous on one side, and most frequently on both sides, to tRNA coding sequences. Furthermore, the H-strand sequences coding for the two rRNAs, the poly(A)-containing RNAs and the tRNAs appear to be adjacent to each other, extending from coordinate 2/100 to coordinate 95/100 of the genome, relative to the origin taken as 0/100. The results are consistent with a model of transcription of the H-strand in the form of a single molecule which is processed into mature RNA species by precise endonucleolytic cleavages, occurring almost in all cases immediately before and after a tRNA sequence. In the processing of the primary transcripts, the tRNA sequences may play an important role as recognition signals.

## Introduction

HeLa cell mitochondrial DNA (mtDNA) codes for two ribosomal RNA (rRNA) species (Aloni and Attardi, 1971a), at least twenty-two tRNA species (Lynch and Attardi, 1976; Angerer et al., 1976; Barrell et al., 1980) and many poly(A)-containing RNAs (Ojala and Attardi, 1974; Amalric et al., 1978). Most of the poly(A)-containing RNAs are associated with polysomal structures and are presumably specific mRNAs (Amalric et al., 1978; Ojala et al., 1980; Attardi et al., 1980). In order to gain an insight into the organization and expression of human mtDNA, in the present work, the positions of the sequences complementary to the poly(A)-containing RNAs on the physical map of HeLa cell mtDNA have been precisely determined by the S1 protection technique (Berk and Sharp, 1977, 1978). These mapping data, as well as the previously determined positions of the rRNA genes (Ojala and Attardi, 1980), have been correlated with the positions of the tRNA genes, as derived from the DNA sequence (B. Barrell and F. Sanger, personal communication; Barrell et al., 1979; Crews and Attardi, 1980). The striking result of this analysis is that the sequences coding for the individual rRNA and poly(A)-containing RNA species on the heavy (H) and light (L) strands are immediately contiguous on one side, and most frequently on both sides, to tRNA coding sequences. Furthermore, the H-strand sequences coding for the rRNA species, the poly(A)-containing RNAs and the tRNAs appear to be adjacent to each other, extending from coordinate 2/100 to coordinate 95/100 of the genome, relative to the origin taken as 0/100.

## Results

### Preliminary Mapping of the Poly(A)-Containing RNA Species

Figure 1a shows the autoradiogram, after high resolution $CH_3HgOH$-agarose

slab gel electrophoresis, of the oligo(dT)-cellulose bound RNA fraction from micrococcal nuclease-treated mitochondria of HeLa cells labeled with $^{32}$P-orthophosphate for 2.5 hr in the absence of inhibitors of nuclear RNA synthesis. Treatment of the organelles with the micrococcal nuclease had the purpose of degrading the extramitochondrial nucleic acids and thus permitting the isolation of substantially pure mitochondrial RNA species (Crews and Attardi, 1980; Gelfand, 1980). In Figure 1a, one recognizes all the discrete components, with molecular weights ranging between $9.0 \times 10^4$ daltons (RNA 18) and $3.4 \times 10^6$ daltons (RNA 1), which had been previously described (Amalric et al., 1978), with the exception of RNA 8. It had been shown earlier that RNAs 5 to 17 are coded for by the H-strand, while RNA 18 is coded for by the L-strand (Amalric et al., 1978). In more recent RNA-DNA hybridization experiments, RNA species 1, 2 and 3 have been shown to be complementary to the L-strand (the percentages of radioactivity hybridized to the L- and H-strands for the three species were, respectively, 64 and 12, 82 and 2, and 66 and 16), and RNA species 4, to the H-strand (the percentages of radioactivity hybridized to the L- and H-strands were 25 and 65).

A preliminary localization of the poly(A)-containing RNA species on the mtDNA Hpa II map (Figure 1b) was made by the DNA transfer technique (Southern, 1975). The results of these experiments were in general unambiguous. RNA species 10 mapped as the large rRNA species (16S rRNA) (Ojala and Attardi, 1980), and, therefore, probably represents a fraction of this rRNA component which binds to oligo(dT)-cellulose (see Discussion), in agreement with earlier observations (Amalric et al., 1978); for this reason it was not analyzed further in this work.

RNA transfer experiments (Alwine et al., 1977) confirmed the positions of the poly(A)-containing RNA species on the Hpa II map, and failed to reveal the existence of any additional RNA species which may have not been resolved in the $CH_3HgOH$-agarose gel (Figure 1c). The few ambiguities could be easily explained by contamination of the DNA probes by mtDNA fragments migrating closely in the preparative gels.

Construction of the transcription map

A more precise mapping of the poly(A)-containing RNAs was done by the S1 protection technique (Berk and Sharp, 1977, 1978). In these experiments, separated strands of the restriction fragments of in vivo $^{32}$P-orthophosphate-labeled mtDNA were used. The final specific activity of the DNA obtainable after the lengthy preparative procedure (~25,000 cpm/μg) was in most cases adequate for detecting the protected DNA segments after electrophoresis in polyacrylamide/urea gels. In order to increase the sensitivity of detection, especially when low abundance RNA species were used, advantage was taken of the $^{32}$P-label associated with the poly(A)-containing RNA species by performing the analysis of the protected DNA segments also in the native state (as RNA-DNA hybrids) in Tris-borate-$Mg^{++}$/polyacrylamide gels.

The results of the S1 protection experiments are shown in Figures 2 to 7. In each experiment with a given RNA species and a specific restriction fragment, an RNA control ("-DNA") and a DNA control ("-RNA") were included. The RNA controls often showed S1-resistant bands corresponding to RNA-RNA hybrids formed between the H-strand coded poly(A)-containing RNA species and contaminating L-strand transcripts. The latter, resulting from the complete transcription

of the L-strand (Aloni and Attardi, 1971b; Murphy et al., 1975; Young and Attardi, 1975), were present in the form of heterogeneous background throughout the $CH_3HgOH$-agarose gel tracks (Amalric et al., 1978). The DNA controls run on gels under native conditions often showed bands corresponding to renatured DNA fragments (indicated by asterisks in the Figures), resulting from the presence, in each fragment strand preparation, of contaminating complementary strand. DNA renaturation was observed in the hybridization experiments carried out at 66°C in high salt, but not in those carried out at 49°C in 80% formamide (see Experimental Procedures). The DNA controls run on gels under denaturing conditions usually showed only the denatured DNA strands (likewise indicated by asterisks in the Figures).

Among the S1-resistant bands detected in the RNA-DNA hybridization mixtures, those which could not be accounted for by the contribution due to the RNA alone or the DNA alone were taken to represent either RNA-DNA hybrids (in the gels run under native conditions) or protected DNA segments deriving from these hybrids (in the denaturing gels). The interpretation of the results was, in general, simple; most of the ambiguities could be accounted for by contamination of the RNA species or DNA fragments used for hybridization by RNA species or, respectively, DNA fragments closely migrating in the preparative gels. No evidence for intervening sequences was found in the mtDNA segments coding for the various poly(A)-containing RNAs.

The length of the protected DNA segments, as estimated from their migration in polyacrylamide-urea gels relative to that of single-stranded DNA markers, was generally greater than the size of the protected segments determined under native conditions relative to that of double-stranded DNA

markers. This result may be due to a faster migration of the RNA-DNA hybrids relative to that of DNA duplexes of the same length, a behavior resulting from the difference in their structure. In the present work, the data obtained after electrophoresis of the protected segments in polyacrylamide/urea gels were preferentially used, whenever possible, in the construction of the transcription map, since they appeared to be closer to the true length, as determined, in a few cases, by comparison with the DNA sequence.

Figure 8 represents in a diagrammatic form the mapping data obtained by the S1 protection technique. The rationale followed in the construction of the transcription map is discussed below.

The DNA and RNA transfer experiments had localized the sequences coding for RNA 4 in correspondence of Hpa II-8 and Hpa II-3. The 5'-end of this RNA was placed, on the basis of the S1 protection experiments (Figure 2), at 360 nt from the Hpa II-8/3 site. This position corresponds almost exactly to the 5'-end of the $tRNA^{Phe}$ gene adjoining the 12S rRNA coding sequence (Crews and Attardi, 1980); the protected DNA segment of 280 nt, which corresponds very closely to the portion of Hpa II-8 coding for 12S rRNA (286 nt; Crews and Attardi, 1980), may result from contamination of RNA 4 by 12S rRNA or from the presence of an RNA 4 derivative, from which the $tRNA^{Phe}$ sequence has been cleaved. The 3'-end of RNA 4 was localized in the proximity of the Hpa II-18/4 site on the basis of the length of this RNA species, as estimated from its electrophoretic mobility in $CH_3HgOH$-agarose gel (~2700 nt; Amalric et al., 1978).

Hybridization of RNA 13 with Hpa II-4H and Hind III-1H protected a segment of apparently equal length (880 nt under native conditions, 980 nt in a denaturing gel), indicating that the coding sequence of this RNA is all

comprised within Hpa II-4 (Figure 3); since this sequence would very well fit in the space between the $tRNA^{Leu}$ and $tRNA^{Ile}$ (958 nt, as determined from the mtDNA sequence), it was tentatively mapped in this segment. The 910 nt and 240 nt bands appearing after hybridization with Hind III-1H and, respectively, Hpa II-4H are presumably due to DNA protection by contaminating RNA 12. The 5'-end of RNA 12 was located in Hpa II-4 at 240 nt from the Hpa II-4/20 site, and the 3' end in Hpa II-9 at 260 nt from the Hpa II-13/9 site (Figure 3). The 5'-end of RNA 6 was located in Hind III-1 at 625 nt from the Hind III-1/2 site, the 3'-end in Hpa II-7 at 330 nt from the Hpa II-14/7 site (Figure 4). The 1900 nt band observed after hybridization of RNA 6 with Hind III-1H is probably due to full protection of this RNA by Hind III-1H and contaminating Hind III-2H. The 5'-end of RNA 9 was located in Hind III-1 at 300 nt from the Hind III-1/2 site (position corresponding almost exactly to that mapped on the basis of the 350 nt protected segment of Hpa II-11H), the 3'-end in Hpa II-7 at 310 nt from the Hpa II-14/7 site (Figure 5). The 400 nt band appearing after hybridization of RNA 9 with Hpa II-11H is of uncertain significance: it may result from contamination of RNA 9 by a larger size processing intermediate of RNA 6; further experiments, however, are needed to verify this possibility. The 1640 nt band (1560 nt in a Tris-borate-$Mg^{++}$ buffered gel) observed after hybridization of RNA 9 with Hind III-2H is very probably due to full protection of this RNA by Hind III-2H and contaminating Hind III-1H. The 1320 and 1220 nt bands (1250 and, respectively, 1150 nt in a Tris-borate-$Mg^{++}$ buffered gel) are presumably due to protection of Hind III-2H by RNA 9 and, respectively, contaminating RNA 7. The 5'-end of RNA 16 (Figure 6) was located in Hpa II-7 at 540 nt from the Hpa II-7/22 site; its 3'-end was located in Hpa II-6 at 152 nt from the Hpa II-22/6 site, on the basis of the length of the segment of Hind III-2H protected by the

intact RNA 16 (730 nt) and taking the above indicated 5'-end position as a reference point.

RNA 14 (Figure 6) protected a segment of apparently equal length (790 nt under native conditions, 820 nt in a denaturing gel) of Hind III-2H and Hpa II-6H, indicating that its coding sequence is all comprised within Hpa II-6; since this sequence would very well fit in the Hpa II-6 segment (∿850 nt) between the 3'-end of the $tRNA^{Lys}$ gene and the portion of Hpa II-6 coding for RNA 15 (∿80 nt), it has been mapped in this segment. The 1650 nt band observed after hybridization of RNA 14 with Hind III-2H is probably due to protection of Hind III-2H by RNA 14 and contaminating RNA 15. The 5'-end of RNA 15 (Figure 6) was located in Hpa II-6 at 80 nt from the Hpa II-6/1 site, the 3'-end in Hpa II-1 at 735 nt from the Hpa II-6/1 site. Also after hybridization with RNA 15, Hind III-2H showed a protected segment of 1650 nt (in a denaturing gel, 1700 nt under native conditions), presumably due to the presence of contaminating RNA 14. The thin band moving slower than the 1700 nt band (∿2100 nt) is of uncertain significance.

The coding sequence of RNA 17 (340 nt), all comprised within Hpa II-1, was tentatively mapped in the segment between $tRNA^{Gly}$ and $tRNA^{Arg}$ (346 nt, as determined from the mtDNA sequence) (Figure 2). The 5'-end of RNA 7 (Figure 5) was mapped in Hind III-2 at 1220 nt from the Hind III-2/3 site, the 3'-end in Hind III-3 at 465 nt from the same site. The 1560 nt band (in a Tris-borate-$Mg^{++}$ buffered gel; 1640 nt in a denaturing gel) observed after hybridization of RNA 7 with Hind III-2H is probably due to full protection of contaminating RNA 9 by Hind III-2H and contaminating Hind III-1H. The 5'-end of RNA 5 (Figure 2) was mapped in Hind III-3 at 230 nt from the Hind III-3/1 site, the 3'-end in Hpa II-$\Delta 2a^{Hae}$ at 500 nt from the Bam HI site. The broad diffuse band migrating somewhat faster than the renatured $\Delta 2a^{Hae}$ fragment may be due to a partial nibbling of the fragment

by the S1 enzyme. RNA 11 (Figure 4) protected a segment of apparently equal length (1120 nt under native conditions, 1150 nt in a denaturing gel) of Hind III-1H and $\Delta 2a^{Hae}$ indicating that its coding sequence is all comprised within $\Delta 2a^{Hae}$; since this sequence would fit almost exactly in the mtDNA segment between the $tRNA^{Glu}$ and the $tRNA^{Thr}$ gene (1145 nt, as determined from the mtDNA sequence), it has been mapped in this region.

The DNA transfer experiments had indicated that the L-strand coded RNAs 1, 2 and 3 have overlapping mapping positions in mtDNA. The S1 protection experiments further showed that these RNA species have their 5'-ends at identical or very closely spaced positions, i.e. at 375 nt from the Bam HI site (Figure 7). The 3'-end of RNA 1 was mapped in Hpa II-4 on the basis of its length estimated from its electrophoretic mobility (~10,400 nt, Amalric et al., 1978); the 3'-end of RNA 2 was mapped in Hpa II-7 at 550 nt from the Hpa II-7/22 site, while the 3'-end of RNA 3 was located in Hpa II-1 at 1200 nt from the Hind III-2/3 site. The L-strand coded RNA 18 had been mapped, on the basis of the DNA and RNA transfer experiments, within Hpa II-8 (Figure 1). Its 5'-end has been recently located by RNA sequencing analysis at 219 nt from the origin of replication, while its 3'-end maps at or very near to this origin (Attardi et al., 1980; Crews, Ojala, Gelfand and Attardi, manuscript in preparation).

An inspection of the H-strand transcription map (Figure 8a) reveals that, with the exception of a relatively small region around the origin of replication (~1300 nt), the H-strand is almost completely occupied by the sequences coding for the individual rRNAs and poly(A)-containing RNAs; these sequences are in most cases separated by very short stretches varying in length between approximately 50 and 230 nt.

Correlation of the Transcription Map with the Positions of the tRNA Genes

The two inner circles in Figure 8a and 8b show the positions, relative to the mtDNA Hpa II map, of the genes for the two rRNA species, 16S and 12S RNA, on the H-strand, as determined by the S1 protection technique (Ojala and Attardi, 1980), and those of the 4S RNA sites on the H- and L-strands, as identified by EM visualization of ferritin-4S RNA/DNA hybrids (Angerer et al., 1976; Ojala and Attardi, 1977). The arrows pointing to the H-strand and L-strand indicate the positions on the two strands and identities of the individual tRNA genes, as derived from the recently determined human mtDNA sequence (F. Sanger and B. Barrell, personal communication; Barrell et al., 1979; Crews and Attardi, 1980). It is clear that, although the positions of the tRNA genes identified by sequence analysis show a gross correspondence with the general distribution of the 4S RNA sites recognized by EM, there are significant discrepancies both as concerns the precise location and the number of these genes. These discrepancies prevented the use of the EM data for the purpose of comparison with the mapping positions of the mtDNA transcripts.

In contrast to the ferritin-4S RNA mapping data, the positions of the tRNA genes derived from the DNA sequence could be directly and fairly precisely aligned with the transcription map, since this map had been constructed by measuring the distances of the ends of the transcripts from mtDNA restriction sites, also recognizable in the DNA sequence. The striking result of this alignment is that, in the H-strand, the tRNA genes occupy all the short intervals separating the sequences coding for the individual rRNA and poly(A)-containing RNA species, in such a way that almost every end of these coding sequences appears to be very close to or immediately juxtaposed with the end of a tRNA gene. In the L-strand, the common 5'-end of the three large

transcripts (RNAs 1, 2 and 3) was likewise found to map very close to an L-strand tRNA gene ($tRNA^{Glu}$); the 3'-end of RNA 2 was also localized very near to an L-strand tRNA gene ($tRNA^{Ser}$ (L)), while the 3'-end of RNA 3 appears to map opposite a position in the H-strand adjoining the tRNA gene ($tRNA^{Arg}$).

The Molecular Sizes of the Poly(A)-Containing RNAs

The lengths of the DNA segments protected from S1 digestion by the hybridized RNAs have provided an estimate of the molecular sizes of these RNAs. As shown in Table 1, the sizes estimated in the present work for the H-strand coded RNAs are very close to the lengths of the corresponding inter-tRNA tracts in DNA, as determined from the mtDNA sequence. Such excellent agreement strongly supports the idea that the H-strand sequences coding for the rRNAs, poly(A)-containing RNAs and tRNAs are immediately contiguous to each other.

## Discussion

### The Nature of the Mitochondrial Poly(A)-Containing RNAs

Among the poly(A)-containing RNA species mapped in the present work, components 5, 7, 9, 11 to 16 and 17, because of their relative abundance, their enrichment in partially purified polysomal structures, and their relatively long half-life (Attardi et al., 1979; Gelfand, 1980), are probably specific mRNAs. Correlation of RNA and DNA sequencing data with protein sequence data and with known yeast gene sequences has recently provided strong support for the mRNA nature of several of these species (RNAs 9, 11, 14, 15, 16) (Barrell et al., 1979; Ojala et al., 1980; Attardi et al., 1980; Walker et al., 1980). RNA species 1 to 4 and 6, because of their presence in only marginal amounts in partially purified polysomal structures and their relatively short half-life, are presumably either precursors or intermediates in the pathway of

maturation of the functional species. The mapping data presented here support this conclusion for RNAs 4 and 6. The possible nature of RNA 18 will be discussed below.

The presence of a $4S_E$ (55 nucleotide-long) poly(A) stretch has been directly demonstrated only in the smaller RNA species (# 7 to 18 (Ojala and Attardi, 1974; Amalric et al., 1978)); however, it seems very likely that also the larger species are polyadenylated, because of their retention on oligo(dT)-cellulose after two passages and of the lack of long stretches of A in the mtDNA sequence (B. Barrell and F. Sanger, personal communication). This point will be discussed further below.

General Organization of mtDNA Transcripts

The detailed transcription mapping analysis presented in this paper allows several general conclusions to be drawn. First, with the exception of the D-loop region and of a stretch of about 170 nucleotides between the 5'-end of the H-strand coded $tRNA^{Phe}$ and the 5'-end of the L-strand coded RNA 18, the HeLa cell mtDNA sequences are completely saturated by the discrete poly(A)-containing RNAs, rRNAs and tRNAs coded for by the two strands. The data provided by the sequence analysis of mtDNA (Barrell et al., 1979; Crews and Attardi, 1980; Sanger et al., 1980; B. Barrell and F. Sanger, personal communication) and of mitochondrial RNA (Crews and Attardi, 1980; Ojala et al., 1980; Attardi et al., 1980; Montoya, Ojala and Attardi, manuscript in preparation) indicate that the saturation of the DNA by the RNA sequences extends to the single nucleotide level (see below). The present results confirm, at a finer level of resolution, the earlier observation of complete or almost complete transcription of the HeLa cell mtDNA H-strand (Aloni and Attardi, 1971c). No evidence for intervening sequences has been found in the mtDNA segments coding for the various poly(A)-containing RNAs. This observation extends

the previous similar results concerning the human mitochondrial rRNA cistrons (Ojala and Attardi, 1980) and is in agreement with the evidence presented for the Xenopus laevis mitochondrial genome (Rastl and Dawid, 1979). The transcription map of human mtDNA presented here shows a general similarity to that of Xenopus laevis mtDNA (Rastl and Dawid, 1979) and, somewhat less, to that of mouse mtDNA (Battey and Clayton, 1978).

Another conclusion of the present studies is that, at the level of resolution allowed by this approach, and with the exception of RNA 4 and RNA 6, there is no apparent overlapping, in the H-strand, of the sequences coding for the various poly(A)-containing RNA, rRNA and tRNA species. There are good reasons to believe that the two RNAs representing an exception of the "non-overlapping" rule may be precursors of mature species. In particular, RNA 4 is a possible precursor of the two rRNA species and RNA 6 a precursor of RNA 9. Both the correspondence in mapping position with that of the mature species, and the relatively short half-life of these presumptive precursors (Attardi et al., 1979; Gelfand, 1980) support the above conclusion.

As concerns the relative mapping positions of the H-strand coding sequences and the L-strand sequences which specify what are presumably mature RNA products (i.e., RNA 18 and tRNAs), some of the latter fall within "non-coding" regions of the H-strand (RNA 18, $tRNA^{Glu}$, $tRNA^{Pro}$), others correspond to 5'- or 3'-terminal regions of H-strand coded poly(A)-containing RNAs lying outside of the polypeptide coding stretches of these RNAs, as determined from the DNA sequence (F. Sanger and B. Barrell, personal communication).

The H-Strand Transcripts. The "Punctuation" Model

The most important conclusion of the present work is that the sequences coding for the individual poly(A)-containing RNA species on the H- and L-

strands, as mapped here by the S1 protection technique, are immediately contiguous or very close on one side, and most frequently on both sides, to tRNA coding sequences, as determined from the mtDNA sequence. The present results have been confirmed and refined by recent data of DNA and RNA sequencing analysis: these have shown that the 5'-end of the coding sequences for 12S rRNA and 16S rRNA (Crews and Attardi, 1980) and for most poly(A)-containing RNAs (including RNAs 13, 14, 17 and 11 which had been only tentatively mapped) (Ojala et al., 1980; Attardi et al., 1980; Montoya, Ojala and Attardi, in preparation), as well as the 3'-end of the sequences coding for 12S rRNA (recognized for its homology to the Chinese hamster small mitochondrial rRNA sequence (Dubin and Baer, 1980)) and for a few poly(A)-containing RNA species so far analyzed, are butt-jointed to a tRNA gene, as identified from the DNA sequence (Barrell et al., 1979; Crews and Attardi, 1980; Sanger et al., 1980; B. Barrell and F. Sanger, personal communication). There are a few exceptions to the rule that each end of a coding sequence for an rRNA or poly(A)-containing RNA species is immediately flanked by a tRNA gene. Thus, the RNA 9 coding sequence does not have a tRNA gene on its 5'-side; this RNA, however, is probably not a primary transcription product, but derives, as mentioned above, from RNA 6 by removal of a 5'-terminal stretch containing sequences complementary to four L-strand coded tRNA species. The RNAs 14 and 15 coding sequences are apparently not separated by a tRNA gene; the mapping results presented here and recent sequence data indicate that their coding sequences are immediately contiguous on the H-strand, suggesting the possibility of a common processing intermediate giving rise to the two mature species by a precise endonucleolytic cleavage. A similar situation seems to occur for RNAs 5 and 11.

The mapping data presented here and the sequencing data so far available have indicated that the H-strand sequences coding for the rRNA species, the

poly(A)-containing RNAs and the tRNAs are immediately contiguous to each other, extending continuously from coordinate 2/100 to coordinate 95/100 (relative to the origin taken as 0/100. This arrangement is consistent with a model of transcription of the H-strand in the form of a single molecule which is processed by precise endonucleolytic cleavages before and after each tRNA sequence to yield the mature products or, in some cases, processing intermediates, like the putative precursors of the rRNAs (RNA 4) and of RNA 9 (RNA 6). A recent mapping study on the nascent mitochondrial RNA molecules isolated from transcription complexes of HeLa cell mtDNA (Cantatore and Attardi, 1980) has pointed to the existence of an initiation site for H-strand transcription near the origin of replication: such a site may thus represent the promoter of the single large transcripts postulated here (Attardi et al., 1980).

In the processing of the primary transcripts, the tRNA sequences may play an important role as recognition signals, providing the punctuation in the reading of mtDNA information (Attardi et al., 1980). It is conceivable that the processing enzyme(s) recognizes the cloverleaf structure of the tRNA sequence or some portion of it.

One interesting observation is that all the RNA species deriving from the initial processing of the primary H-strand transcripts are apparently polyadenylated. This suggests that polyadenylation may be linked in some way to the processing event, independently of the functional role of the products. One special comment should be made, in this connection, concerning the polyadenylation of the presumptive rRNA precursor (RNA 4). It seems unlikely that the retention on oligo(dT)-cellulose of this RNA species is due to the presence of internal A-rich stretch(es), since the 12S rRNA and the bulk of 16S rRNA (two species which, on the basis of the available

evidence, share with RNA 4 all sequences except the tRNAs) do not bind to oligo(dT)-cellulose. The small fraction of 16S rRNA which is retained on oligo(dT)-cellulose after two passages (Amalric et al., 1978) may represent incompletely processed 16S rRNA molecules with residual poly(A) tails. Further work is, however, necessary to establish this point conclusively. It should be mentioned that, also in Drosophila melanogaster, the large mitochondrial rRNA component has been shown to have affinity for oligo(dT)-cellulose (Spradling et al., 1977). The occurrence in Chinese hamster cell mitochondria of an apparently polyadenylated $20S_E$ RNA species, behaving in pulse-labeling and RNA-DNA hybridization experiments as an rRNA precursor, has been previously reported (Cleaves et al., 1976).

The L-Strand Transcripts

Although the picture of L-strand transcription is still very preliminary, the available evidence suggests that it may follow the same pattern as the H-strand transcription. In particular, the tRNA sequences may also play a role in the processing of L-strand transcription products. The above mentioned mapping study on the nascent mitochondrial RNA molecules has also pointed to the existence of an initiation site for L-strand transcription near the origin of replication (Cantatore and Attardi, 1980); therefore, it is possible that the common 5'-end of RNAs 1, 2 and 3 represents a processing point of the primary transcripts. These RNAs may result from termination of transcription at alternative fixed points or represent successive steps in a processing scheme. The physiological significance of these large RNAs is not clear. They may be precursors of L-strand coded tRNAs. However, the possibility that these transcripts or their possible derivatives may function as mRNAs, or serve some other function, cannot be excluded.

RNA 18, a species which had been previously described as 7S RNA on the basis of its sedimentation constant (Ojala and Attardi, 1974), may be the mRNA of a small polypeptide or have some other function related to the replication or transcription of mtDNA (Crews, Ojala, Gelfand and Attardi, manuscript in preparation).

## Experimental Procedures

### Materials

Micrococcal nuclease was purchased from Worthington, all restriction enzymes from New England Biolabs, Aspergillus orizae S1 nuclease from Sigma, E. coli DNA polymerase I from Boehringer & Mannheim and methylmercuric hydroxide from Alpha Products.

### Preparation of Mitochondrial RNA

Total oligo(dT)-cellulose bound mitochondrial RNA which was to be used in the RNA transfer experiments (Alwine et al., 1977) was prepared from HeLa cells as previously described (Amalric et al., 1978).

For use in the DNA transfer (Southern, 1975) and S1 protection experiments (Berk and Sharp. 1977, 1978), individual RNA species were isolated as follows. RNA was extracted from micrococcal nuclease treated mitochondria of HeLa cells labeled for 2.5 hr with $^{32}$P-orthophosphate (Crews and Attardi, 1980). The extracted RNA was passed twice through oligo(dT)-cellulose, with a heat denaturation step (4 min at 65°C in 0.01 M Tris-HCl (pH 7.4), 0.001 M EDTA) performed prior to each passage. The bound fraction was electrophoresed on preparative agarose (1.4%)/methylmercuric hydroxide slab gels and the individual species, identified by autoradiography of the wet gels, were cut out, eluted and ethanol precipitated (Crews and Attardi, 1980).

Preparation of Mitochondrial DNA

Isolation of closed circular mtDNA from either unlabeled cells or cells long term labeled with $^{32}$P-orthophosphate, and subsequent digestion of this DNA with Hpa II - Bam HI or Hind III restriction enzymes were carried out as described (Ojala and Attardi, 1977). The restriction products were fractionated by electrophoresis through 1.5% to 25% polyacrylamide gradient slab gels or 2% agarose slab gels in Tris-acetate-EDTA buffer (Ojala and Attardi, 1977) (Hpa II - Bam HI fragments) or 1.5% agarose gels (Hind III fragments). Bands were visualized by ethidium bromide staining or autoradiography of the wet gel, cut out, electroeluted from the gel slices, and ethanol precipitated. Individual Hpa II - Bam HI restriction fragments were either further purified (for use in the RNA transfer experiments (Alwine et al., 1977)) by rerun on 2% agarose slab gels in Tris-acetate-EDTA buffer, or subjected to strand separation as described by Cantatore and Attardi (1980). The strands of the Hind III restriction fragments were separated by brief alkali treatment (0.02 N NaOH, 2 min at 65°C, followed by rapid chilling in ice), and electrophoresis through 0.7% or 1% agarose slab gels in Tris-phosphate buffer (fragments Hind III-1 and Hind III-2, or Hind III-3, respectively). The strand specificity of the separated strands was determined by hybridization with total H or L strands as described (Cantatore and Attardi, 1980).

Hybridization Experiments by the DNA or RNA Transfer Technique

DNA blot analysis utilized Hpa II - Bam HI restriction digests of mtDNA resolved on 2% agarose slab gels and transferred to nitrocellulose filters (Southern, 1975). Hybridization with individual oligo(dT)-cellulose bound RNA species labeled in vivo with $^{32}$P-orthophosphate was carried out in 2 ml

of 6xSSC, 0.5% SDS, for 24 hr at 68°C. The RNA input varied from 500 to 10,000 cpm. After hybridization, the filter strips were rinsed in 10 ml of hybridization buffer for 2 hr at 68°C and then autoradiographed, after being aligned with an identical strip which had been hybridized to a nick-translated and denatured total mtDNA Hpa II - Bam HI digest.

For the experiments of hybridization by the RNA transfer procedure (Alwine et al., 1977), the oligo(dT)-cellulose-bound mitochondrial RNA was fractionated by electrophoresis on 1.6% agarose-$CH_3HgOH$ slab gels. Preparation of diazobenzyloxymethyl (DBM) paper and transfer of the RNA to the paper were carried out as described (Alwine et al., 1977), except that 0.2 M and 0.025 M phosphate buffer (pH 6.5) were used, instead of borate buffer, in the gel washing steps.

Total mtDNA or purified Hpa II fragments were labeled in vitro by nick translation using the procedure by Rigby et al. (1977), modified as described elsewhere (Cantatore and Attardi, 1980). The specific activity obtained was in the range of $10^7$ - $10^8$ cpm/μg.

Each strip of paper (~1x20 cm), which had bound to it poly(A)-containing mitochondrial RNA from about $2.5x10^8$ cells, was preincubated in 5 ml hybridization buffer, then exposed to $^{32}$P-labeled DNA ($10^6$ - $10^7$ cpm), washed and autoradiographed as described (Alwine et al., 1977). In general, the strips were first hybridized with total $^{32}$P-labeled mtDNA to provide a marker pattern, then hybridized with individual fragments. In order to reuse the RNA-containing strip, the paper was treated with 99% formamide at 65°C: at least six changes of formamide over a period of 24 hr were required to remove the $^{32}$P-labeled DNA probe. Autoradiography was used to ensure that all of the $^{32}$P-labeled DNA probe had been removed prior to reuse of the strip with another probe.

S1 Protection Analysis

Samples of purified RNA species were dried down in snap cap vials together with an appropriate amount of the H-strand or L-strand of the desired fragment (10 to 100 ng). Hybridization at 66°C in 0.4 M NaCl or at 49°C in 80% formamide [the latter conditions favoring RNA-DNA hybridization over DNA-DNA reassociation (Casey and Davidson, 1977)] and treatment of the hybrids with the Aspergillus orizae S1 nuclease were carried out as described elsewhere (Ojala and Attardi, 1980). Electrophoretic analysis of the protected DNA segments was carried out either under native conditions in 4 or 5% polyacrylamide slab gels in Tris-borate-$Mg^{++}$ buffer (TBM) (Maniatis et al., 1975), or under denaturing conditions in 4% polyacrylamide/7 M urea slab gels in Tris-borate-EDTA buffer (TBE) (Maniatis et al., 1975). For the latter purpose, the S1 digested samples were ethanol precipitated in the presence of 5 µg yeast tRNA, denatured by heating for 5 min at 80°C in 0.01 M Tris-HCl (pH 7.4), digested with RNase A (10 µg/ml for 10 min at 20°C in the presence of 0.1 M NaCl), reprecipitated with ethanol in the presence of 5 µg yeast tRNA, dissolved in 7 M urea, 0.001 M Tris-HCl (pH 7.4), 0.001 M EDTA, heated for 3 min at 80°C, fast cooled and immediately layered on the gel. Hpa II - Bam HI digests of HeLa cell mtDNA, in vitro labeled with [$\alpha^{32}$P]-deoxyribonucleoside triphosphates and E. coli DNA polymerase I as previously described (Ojala and Attardi, 1977), were used as markers either directly (TBM gels) or after denaturation as described above (urea gels).

Acknowledgments

These investigations were supported by a grant from NIH (GM-11726). We are very grateful to Drs. F. Sanger and B. Barrell for communicating to us human mtDNA sequence data prior to publication. The technical assistance of Ms. A. Drew is gratefully acknowledged.

References

Aloni, Y., and Attardi, G. (1971a). Expression of the mitochondrial genome in HeLa cells. IV. Titration of mitochondrial genes for 16S, 12S and 4S RNA. J. Mol. Biol. 55, 271-276.

Aloni, Y., and Attardi, G. (1971b). Symmetrical in vivo transcription of mitochondrial DNA in HeLa cells. Proc. Nat. Acad. Sci. USA 68, 1757-1761.

Aloni, Y., and Attardi, G. (1971c). Expression of the mitochondrial genome in HeLa cells. II. Evidence for complete transcription of mitochondrial DNA. J. Mol. Biol. 55, 251-270.

Alwine, J. C., Kemp, D. J., and Stark, G. R. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc. Nat. Acad. Sci. USA 74, 5350-5354.

Amalric, F., Merkel, C., Gelfand, R., and Attardi, G. (1978). Fractionation of mitochondrial RNA from HeLa cells by high-resolution electrophoresis under strongly denaturing conditions. J. Mol. Biol. 118, 1-25.

Angerer, L., Davidson, N., Murphy, W., Lynch, D., and Attardi, G. (1976). An electron microscope study of the relative positions of the 4S and ribosomal RNA genes in HeLa cell mitochondrial DNA. Cell 9, 81-90.

Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C., and Ojala, D. (1979). The organization of the genes in the human mitochondrial genome and their mode of transcription. In Extrachromosomal DNA, ICN-UCLA Symposia on Molecular and Cellular Biology, 15, D. Cummings, P. Borst, I. Dawid, S. Weissman and C. F. Fox, eds. (New York: Academic Press), pp. 443-469.

Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C., Montoya, J., and Ojala, D. (1980). The remarkable features of gene organization and expression of human mitochondrial DNA. In The Organization and Expression of the Mitochondrial Genome, A. M. Kroon and C. Saccone, eds. (Amsterdam: North-Holland), in press.

Barrell, B. G., Bankier, A. T., and Drouin, J. (1979). A different genetic code in human mitochondria. Nature 282, 189-194.

Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1980). Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc. Nat. Acad. Sci. USA 77, 3164-3166.

Battey, J., and Clayton, D. A. (1978). The transcription map of mouse mitochondrial DNA. Cell 14, 143-156.

Berk, A. J., and Sharp, P. A. (1977). Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell 12, 721-732.

Berk, A. J., and Sharp, P. A. (1978). Spliced early mRNAs of simian virus 40. Proc. Nat. Acad. Sci. USA 75, 1274-1278.

Cantatore, P., and Attardi, G. (1980). Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA. Nucl. Acids Res. 8, 2605-2624.

Casey, J., and Davidson, N. (1977). Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucl. Acids Res. 4, 1539-1552.

Cleaves, G. R., Jones, T., and Dubin, D. T. (1976). Properties of a discrete high molecular weight poly(A)-containing mitochondrial RNA. Arch. Biochem. Biophys. 175, 303-311.

Crews, S., and Attardi, G. (1980). The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. Cell 19, 775-784.

Dubin, D. T., and Baer, R. J. (1980). Sequence and structure analysis of mitochondrial ribosomal RNA from hamster cells. In The Organization and Expression of the Mitochondrial Genome, A. M. Kroon and C. Saccone, eds. (Amsterdam: North-Holland), in press.

Gelfand, R. A. (1980). Studies on RNA metabolism in HeLa mitochondria. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Lynch, D. C., and Attardi, G. (1976). Amino acid specificity of the transfer RNA species coded for by HeLa cell mitochondrial DNA. J. Mol. Biol. 102, 125-141.

Maniatis, T., Jeffrey, A., and van de Sande, H. (1975). Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry 14, 3787-3794.

Murphy, W. I., Attardi, B., Tu, C., and Attardi, G. (1975). Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in HeLa cells. J. Mol. Biol. 99, 809-814.

Ojala, D., and Attardi, G. (1974). Identification and partial characterization of multiple discrete polyadenylic acid-containing RNA components coded for by HeLa cell mitochondrial DNA. J. Mol. Biol. 88, 205-219.

Ojala, D., and Attardi, G. (1977). A detailed physical map of HeLa cell mitochondrial DNA and its alignment with the positions of known genetic markers. Plasmid 1, 78-105.

Ojala, D., and Attardi, G. (1978). Precise localization of the origin of replication in a physical map of HeLa cell mitochondrial DNA and isolation of a small fragment that contains it. J. Mol. Biol. 122, 301-319.

Ojala, D., and Attardi, G. (1980). Fine mapping of the ribosomal RNA genes of HeLa cell mitochondrial DNA. J. Mol. Biol. 138, 411-420.

Ojala, D., Montoya, J., and Attardi, G. (1980). The putative mRNA for subunit II of cytochrome c oxidase in human mitochondria starts directly at the translation initiator codon. Nature, in press.

Rastl, E., and Dawid, I. B. (1979). Expression of the mitochondrial genome in Xenopus laevis: a map of transcripts. Cell 18, 501-510.

Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113, 237-251.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol., in press.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Spradling, A., Pardue, M. L., and Penman, S. (1977). Messenger RNA in heat-shocked Drosophila cells. J. Mol. Biol. 109, 559-587.

Walker, J. E., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Runswick, M. J., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1980). The organization of the mammalian mitochondrial genome. In The Organization and Expression of the Mitochondrial Genome, A. M. Kroon and C. Saccone, eds. (Amsterdam: North-Holland), in press.

Young, P. G., and Attardi, G. (1975). Characterization of double-stranded RNA from HeLa cell mitochondria. Biochem. Biophys. Res. Commun. 65, 1201-1207.

Table 1. Molecular size of mitochondrial RNAs

| RNA species | Molecular size (number of nucleotides) | |
|---|---|---|
| | Estimated from S1 protection data | Expected for DNA saturation[a] |
| H transcripts | | |
| 12S | 970 | 954 |
| 16S | 1600 | 1559 |
| # 5 | 2420 | 2410 |
| # 6 | 1955 | 1938 |
| # 7 | 1685 | 1668 |
| # 9 | 1620 | 1617 |
| #11 | 1150 | 1141 |
| #12 | 1030 | 1042 |
| #13 | 980 | 958 |
| #14 | 820 | 842 |
| #15 | 810 | 784 |
| #16 | 730 | 709 |
| #17 | 340 | 346 |
| L transcripts | | |
| #2 | 7070 | |
| #3 | 4155 | |
| #18 (7S) | 215 | |

Footnote to Table I

[a]The lengths listed in this column represent the inter-tRNA distances in mtDNA (as derived from the DNA sequence) corresponding to those RNA species whose coding sequences are flanked by tRNA genes on both sides; for the few RNA species whose coding sequences are flanked only on one side by a tRNA gene, the position of the other RNA terminus in the DNA sequence was identified from the alignment of the RNA and DNA sequences (Attardi et al., 1980; Montoya, Ojala and Attardi, in preparation).

Figure 1

Mapping of the Oligo(dT)-Cellulose Bound Mitochondrial RNA Species by Hybridization with Hpa II mtDNA Fragments After RNA Transfer to DBM-Paper.

(a) Autoradiogram, after electrophoresis through a $CH_3HgOH$-agarose slab gel, of the oligo(dT)-cellulose bound RNA fraction from micrococcal nuclease treated mitochondria of HeLa cells labeled for 2.5 hr with $^{32}P$-orthophosphate.

(b) Physical map of HeLa cell mtDNA indicating the Hpa II, Bam HI, and Hind III restriction sites, and the Hae III restriction site in Hpa II fragment 8. The positions of the genes for the rRNA species, 12S and 16S RNA, as well as the origin (0) and direction of H-strand synthesis, are also shown.

(c) Results of hybridization of total mtDNA or individual purified restriction fragments, labeled with $^{32}P$-dNTPs by nick translation, with the oligo(dT)-cellulose bound mitochondrial RNA fractionated by electrophoresis on $CH_3HgOH$-agarose slab gel and transferred to DBM-paper. In the last lane, a longer exposure of the strip exhibited in the adjacent lane is shown to better visualize the hybridization with RNA 4. Δ8a: subfragment of Hpa II-8 produced by Hae III cleavage (Ojala and Attardi, 1978).

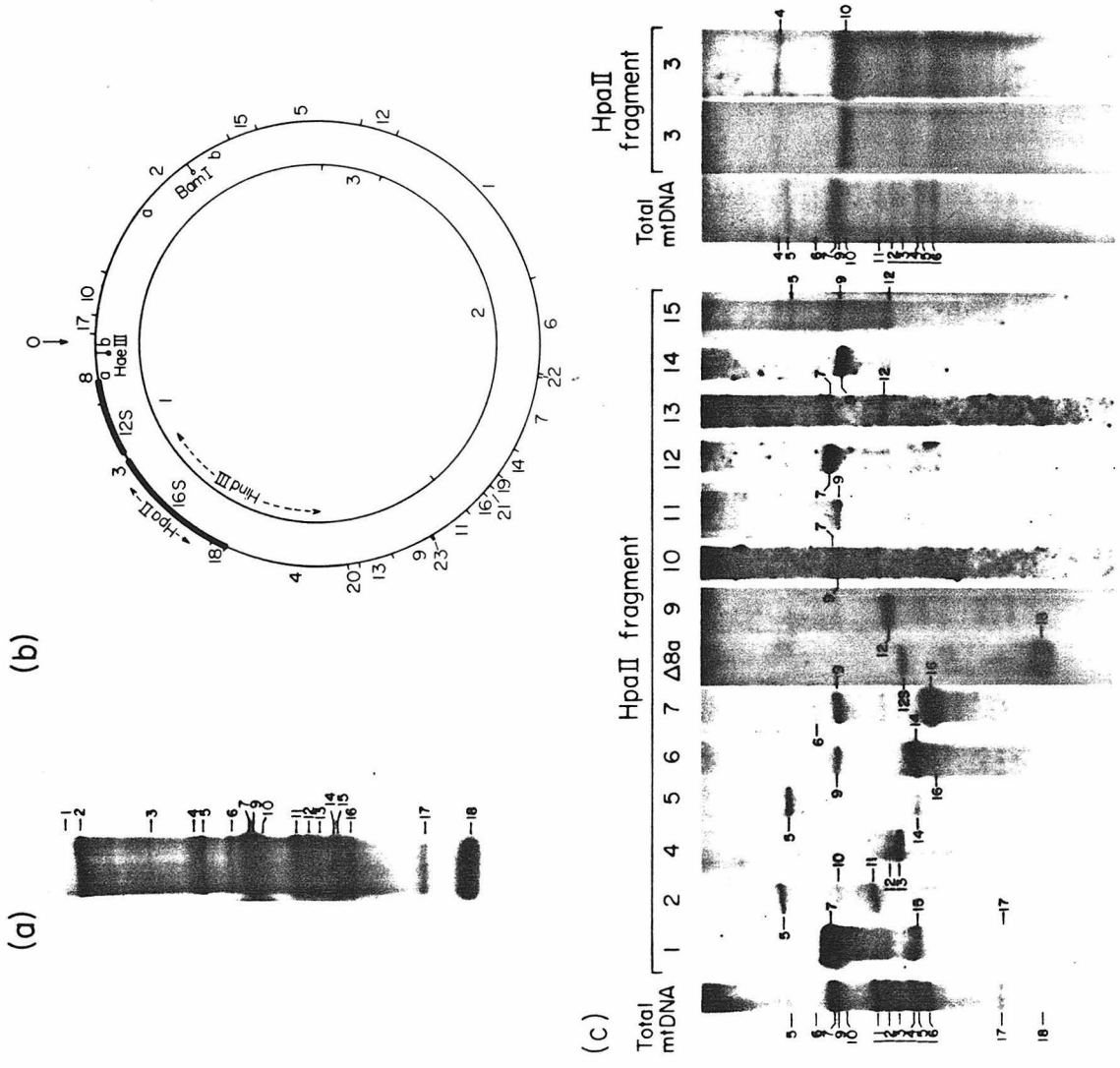

(a)
(b)
(c)
HpaII fragment
Total mtDNA
HpaII fragment
Total mtDNA
HpaII
HaeIII
BamI
HindIII
12S
16S

Figure 2

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 4, 5 and 17, Using the H-Strands of Restriction Fragments.

Hybridization conditions: lanes 1 to 3, 20 hr at 49°C in 80% formamide; all other lanes, 4 hr at 66°C in 0.4 M salt. Electrophoretic analysis: all lanes, 5% polyacrylamide gels in TBM. M: total mtDNA Hpa II + Bam HI digest marker. The asterisks in these and in the following gel patterns indicate renatured DNA fragments (in TBM gels) or denatured DNA strands (in urea gels).

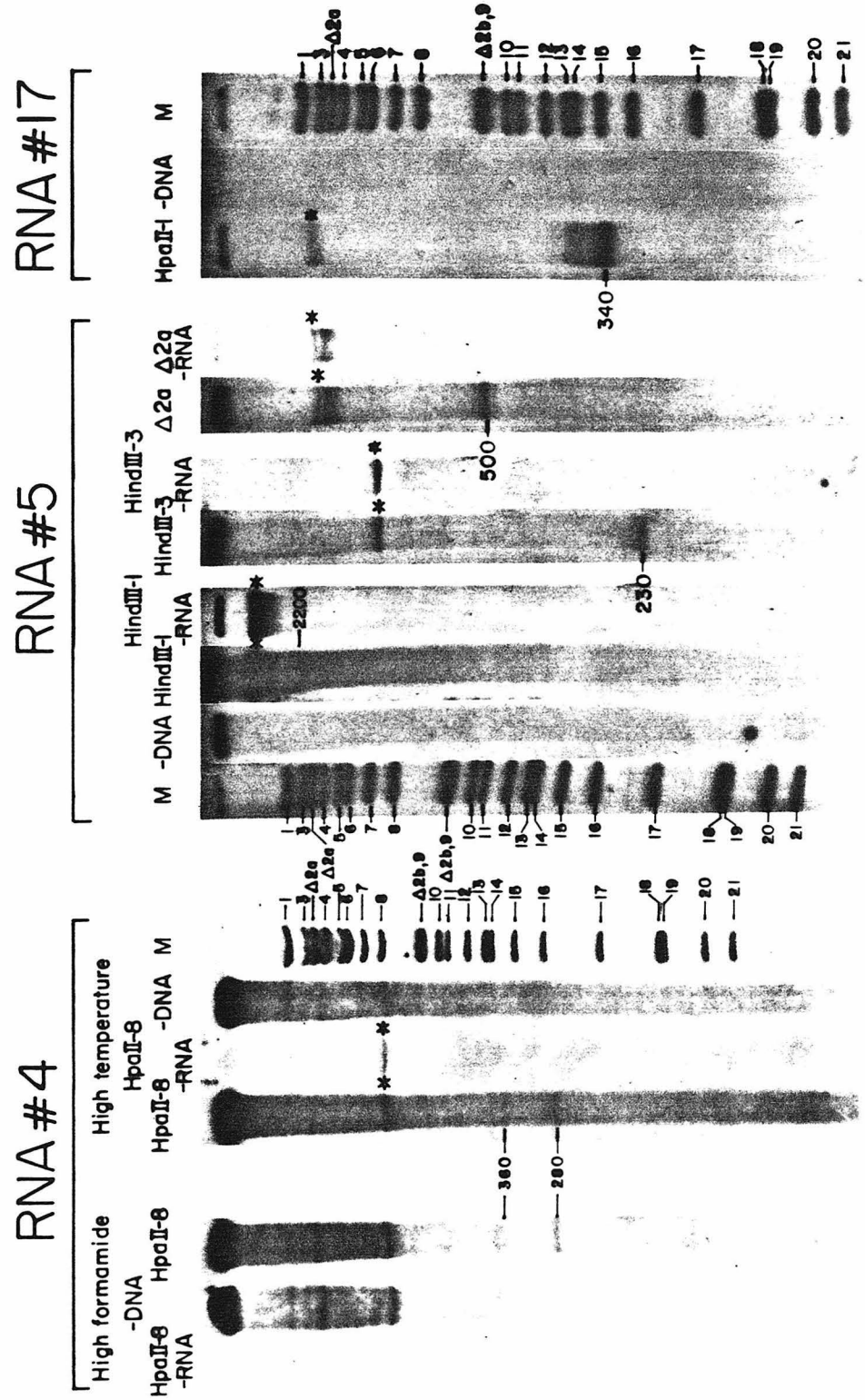

RNA #4
High formamide
-DNA
HpaII-8
-RNA
HpaII-8
High temperature
HpaII-8
-DNA
HpaII-8
-RNA
M
RNA #5
M
-DNA
HindIII-1
-RNA
HindIII-1
HindIII-3
-RNA
HindIII-3
Δ2a
-RNA
Δ2a
Δ2a,b
-RNA
RNA #17
M
HpaII-8
-DNA
340
500
230
2200
380
280

Figure 3

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 12 and 13, Using the H-Strands of Restriction Fragments.

Hybridization conditions: lanes 1, 2, 5, 10, 11 and 12 to 19, 4 hr at 66°C in 0.4 M salt; lanes 6 and 7, 20 hr at 49°C in 80% formamide. Electrophoretic analysis: lanes 1 to 5 and 12 to 16, 5% polyacrylamide gels in TBM; lanes 6 to 8, 4% polyacrylamide gel in TBM; lanes 9 to 11 and 17 to 19, 4% polyacrylamide - 7 M urea gels in TBE. Hpa II fragment 22 is a fragment previously indicated as $X_1$ (Ojala and Attardi, 1978).

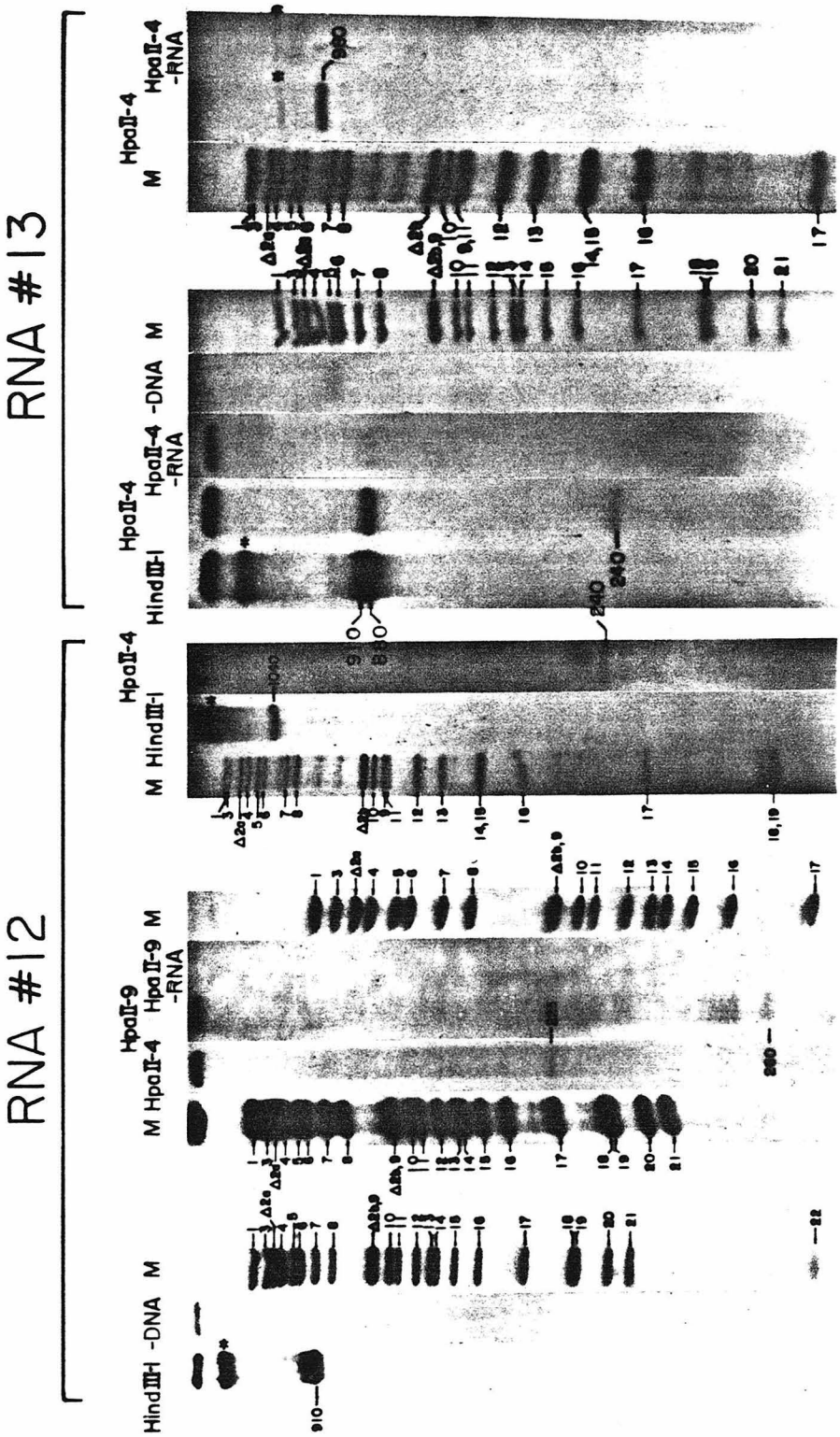

RNA #13
RNA #12
HpaII-4 -RNA
M
HpaII-4
HindIII-I
HpaII-4 -DNA
HpaII-9 -RNA
HpaII-9
HindIII -DNA
240
230
210
—22

Figure 4

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 6 and 11, Using the H-Strands of Restriction Fragments.

Hybridization conditions: all lanes, 4 hr at 66°C in 0.4 M salt. Electrophoretic analysis: lanes 1 to 9 and 17 to 20, 5% polyacrylamide gels in TBM; lanes 10 to 16, 4% polyacrylamide - 7 M urea gel in TBE.

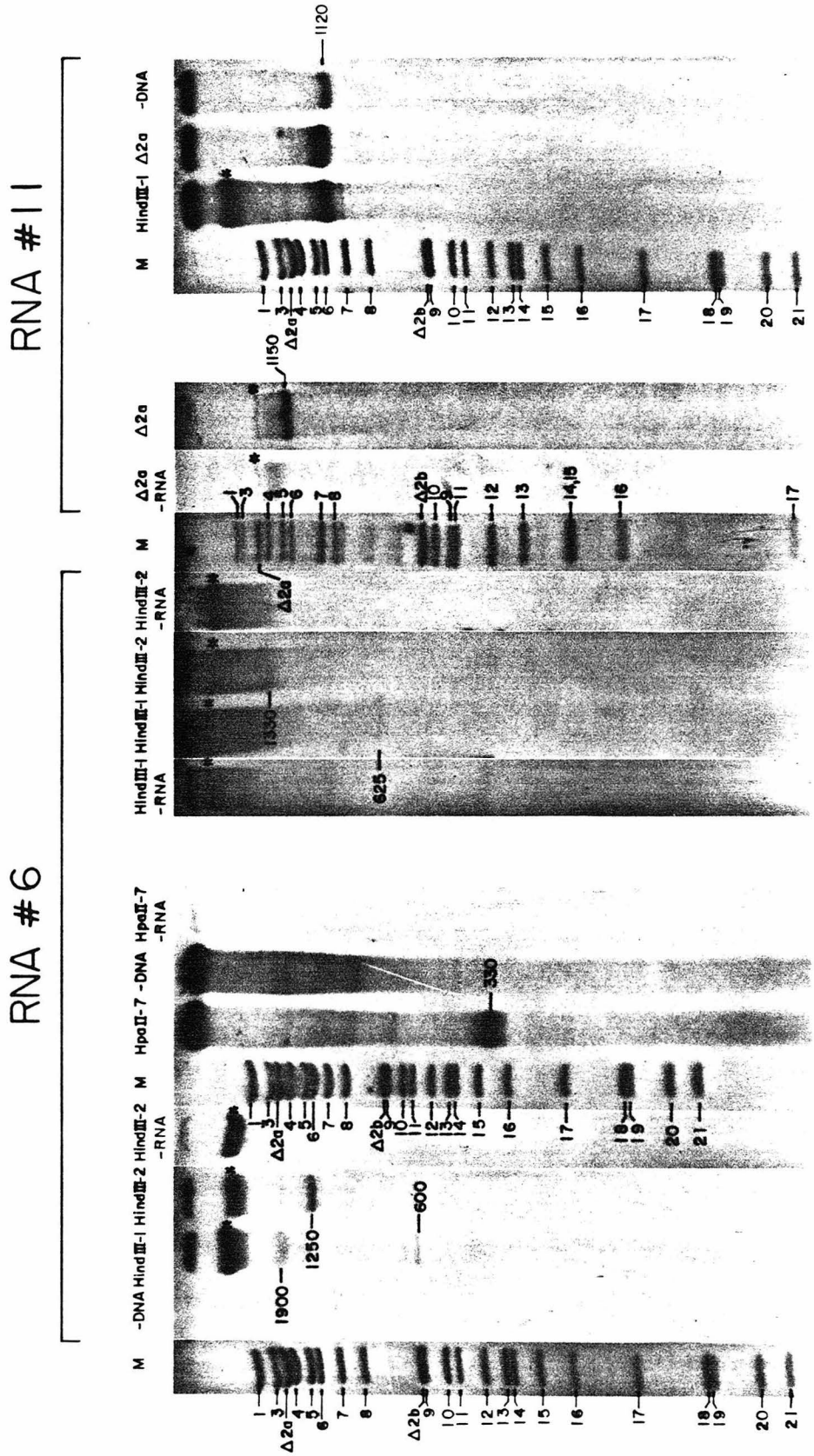

RNA #6
RNA #11
M
-DNA
HindIII-1
HindIII-2
-RNA
HpaII-7
1900
1250
600
330
625
1330
1150
1120
Δ2a
Δ2b

Figure 5

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 7 and 9, Using the H-Strands of Restriction Fragments.

Hybridization conditions: lanes 2 to 4, 9 to 11 and 13 to 15, 4 hr at 66°C in 0.4 M salt; lanes 5, 6, 16, 17 and 19, 20 hr at 49°C in 80% formamide. Electrophoretic analysis: lanes 1 to 7, 12 to 15, 19 and 20, 5% polyacrylamide gels in TBM; lanes 16 to 18, 4% polyacrylamide gel in TBM; lanes 8 to 11, 4% polyacrylamide - 7 M urea gel in TBE.

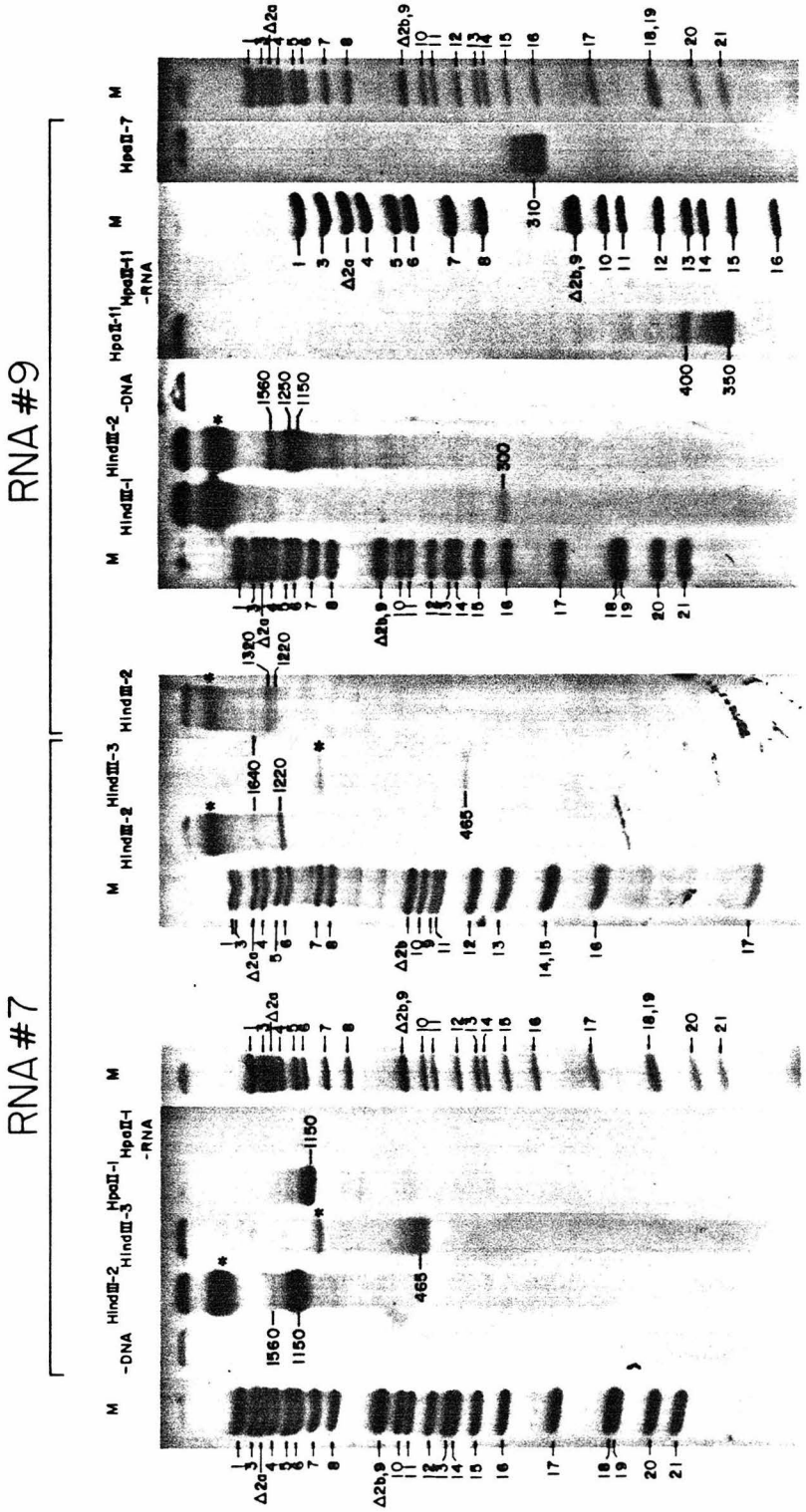

RNA #7
RNA #9
M
-DNA
HindIII-2
HindIII-3
HpaII-1
-RNA
HindIII-1
HpaII-7
HpaII-11
1560
1150
465
1640
1220
1320
1250
300
400
350
310

Figure 6

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 14, 15 and 16 Using H-Strands of Restriction Fragments.

Hybridization conditions: lane 15, 20 hr at 49°C in 80% formamide; all other lanes, 4 hr at 66°C in 0.4 M salt. Electrophoretic analysis: lanes 1 to 5 and 11 to 20, 5% polyacrylamide gels in TBM; lanes 6 to 10 and 21 to 23, 4% polyacrylamide - 7 M urea gel in TBE. The lower asterisk in the last lane indicates a small amount of non-denatured fragment 7, identified by comparison with the migration of native fragment markers.

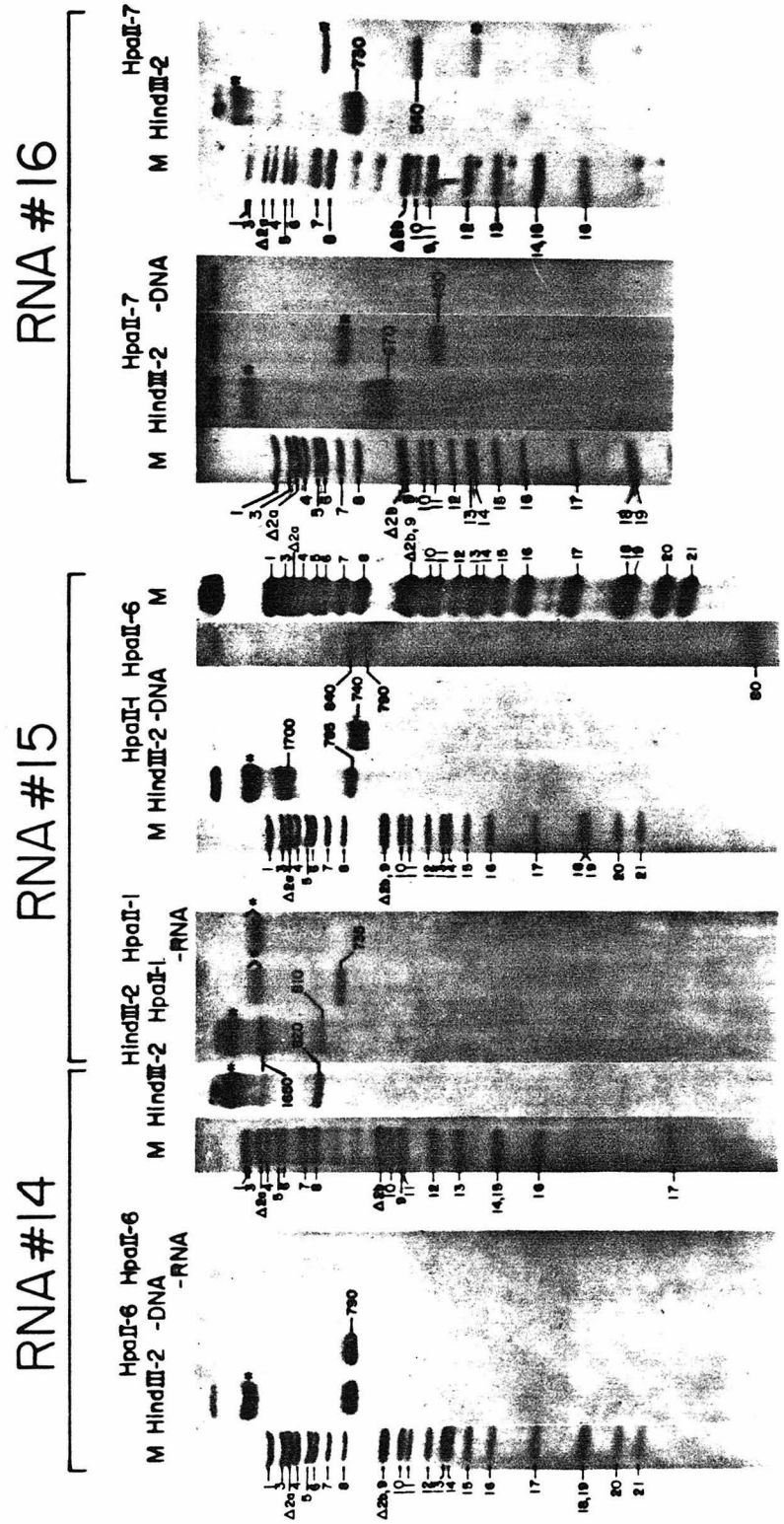

RNA #14
RNA #15
RNA #16
HpaII-6 HpaII-6
M HindIII-2 -DNA
-RNA
HindIII-2 HpaII-1
M HindIII-2 HpaII-1
-RNA
HpaII-1 HpaII-6
M HindIII-2 -DNA
M
HpaII-7
M HindIII-2 -DNA
HpaII-7
M HindIII-2

Figure 7

Nuclease S1 Sizing and Mapping of Mitochondrial RNAs 1, 2 and 3, Using L-Strands of Restriction Fragments.

Hybridization conditions: lanes 2 to 6 and 13 to 17, 4 hr at 66°C in o.4 M salt; lanes 7 to 10, 20 hr at 49°C in 80% formamide. Electrophoretic analysis: lanes 1 to 6 and 12 to 17, 4% polyacrylamide gels in TBM; lanes 7 to 11, 5% polyacrylamide gels in TBM.

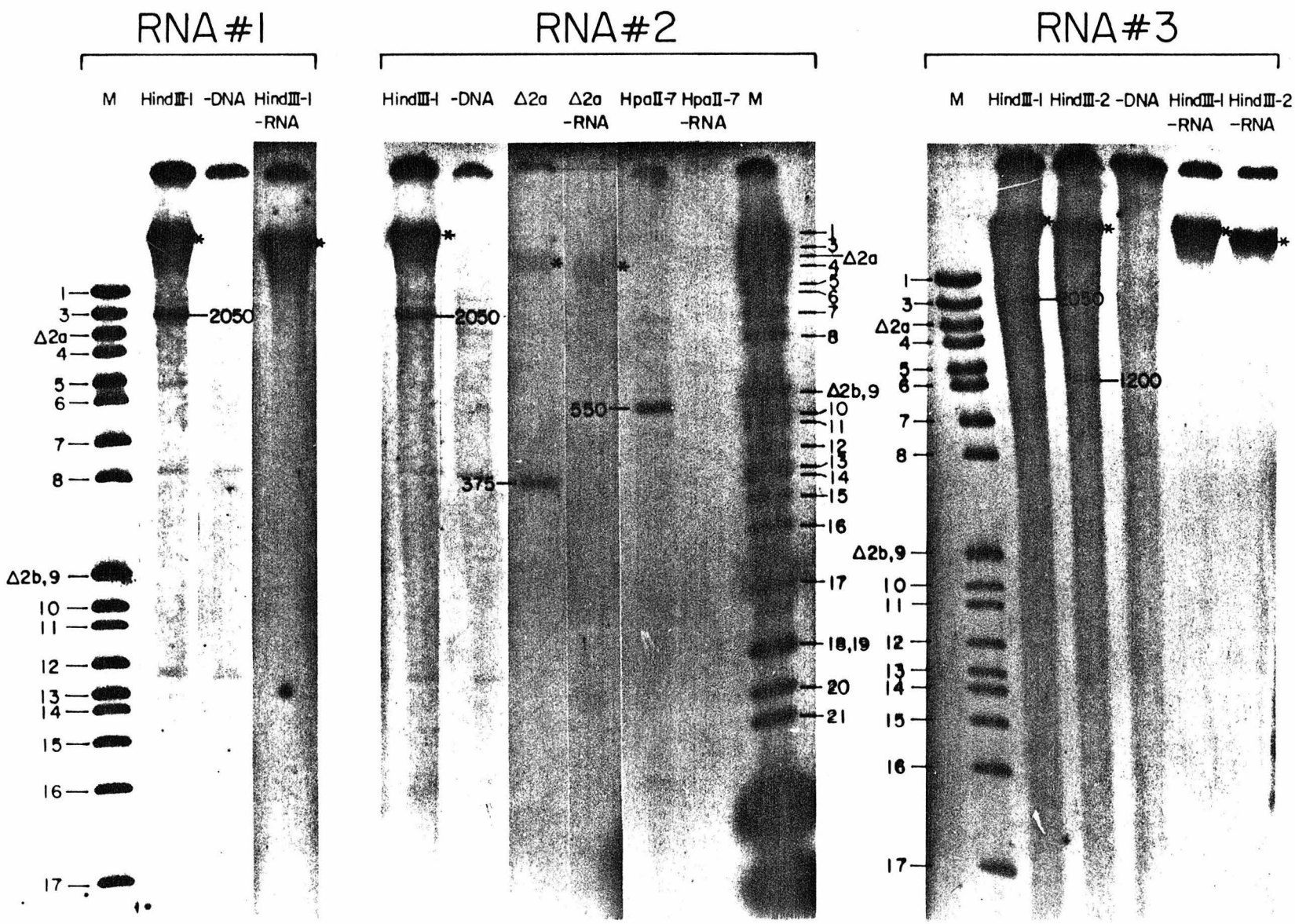
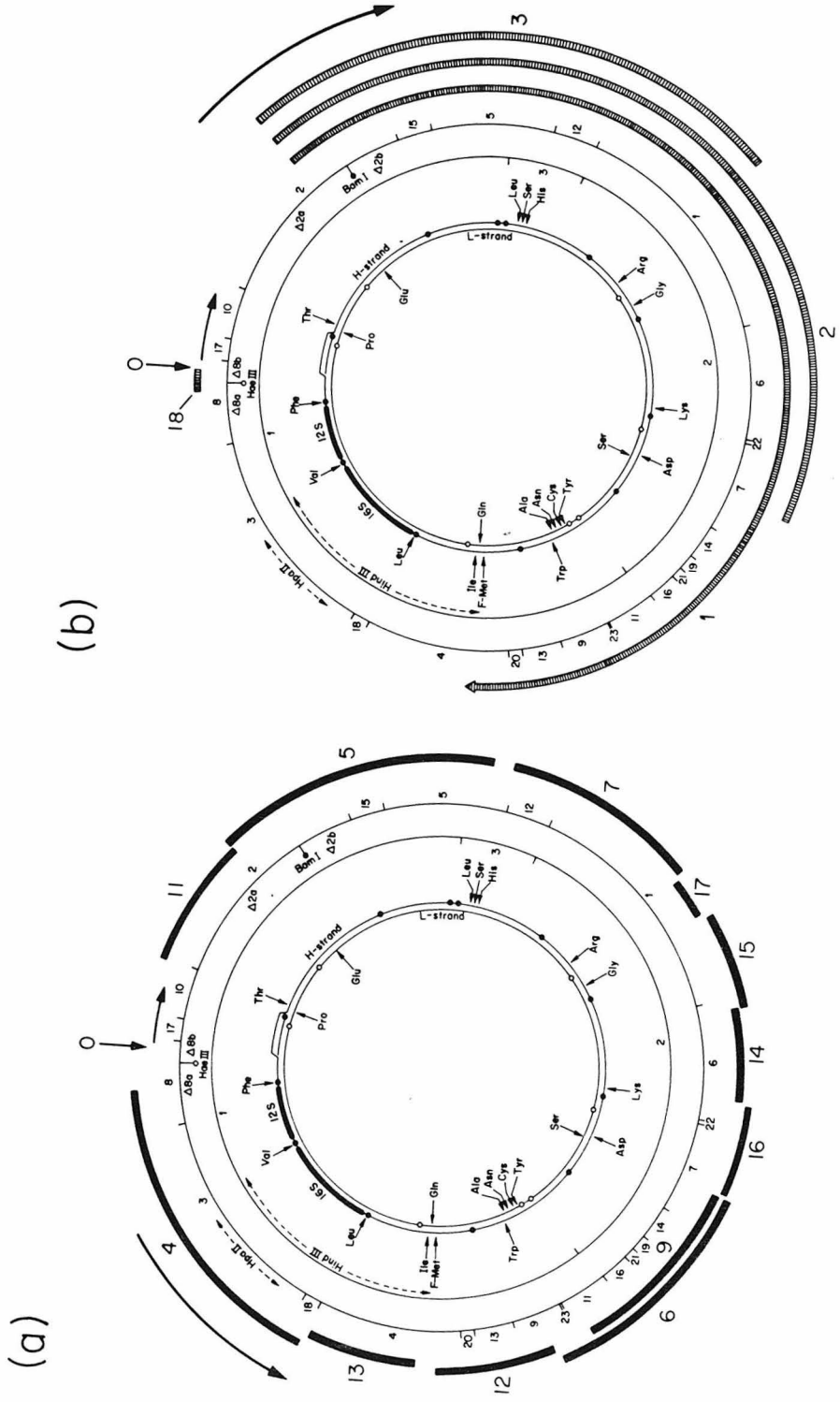

RNA#1
M
HindIII-1
-DNA
HindIII-1 -RNA
2050
1
3
Δ2a
4
5
6
7
8
Δ2b,9
10
11
12
13
14
15
16
17
RNA#2
HindIII-1
-DNA
Δ2a
Δ2a -RNA
HpaII-7
HpaII-7 -RNA
M
2050
550
375
1
3
Δ2a
4
5
6
7
8
Δ2b,9
10
11
12
13
14
15
16
17
18,19
20
21
RNA#3
M
HindIII-1
HindIII-2
-DNA
HindIII-1 -RNA
HindIII-2 -RNA
2050
1200
1
3
Δ2a
4
5
6
7
8
Δ2b,9
10
11
12
13
14
15
16
17

Figure 8

Detailed Physical and Genetic Maps of HeLa Cell mtDNA Illustrating the Positions of the H-Strand Transcription Products (a) and the L-Strand Transcription Products (b).

The two inner circles show the positions of the two rRNA genes (Ojala and Attardi, 1980), those of the 4S RNA genes, as determined by a ferritin mapping technique (Angerer et al., 1976) (●, H-strand 4S RNA genes; o, L-strand 4S RNA genes), and those of the tRNA genes, as derived from the mtDNA sequence (Sanger and Barrell, personal communication; Barrell et al., 1979; Crews and Attardi, 1980) (arrows pointing inward, H-strand genes; arrows pointing outward, L-strand genes). The middle rings indicate the Hpa II, Bam HI, and Hind III restriction sites, and the Hae III site in Hpa II fragment 8. The positions of the transcripts as mapped in the present work are indicated by the wide black or dashed bars. The direction of transcription is shown by the outermost arrows. The origin (O) and direction (small rightward arrow) of H-strand synthesis are also shown.


(a)
(b)
L-strand
H-strand
Leu
Ser
His
Arg
Gly
Lys
Asp
Ser
Trp
Tyr
Cys
Asn
Ala
Gln
Ile
F-Met
Leu
Val
Phe
Thr
Pro
Glu
Bam I
Hae III
Hind III
Hpa II
0
18

# Chapter 4

## 7S RNA: A Polyadenylated Transcript Which Maps Adjacent to the Origin of DNA Replication in HeLa Cell Mitochondria

**Abstract**

7S RNA (RNA 18 in the classification of Amalric et al., 1978) is a small polyadenylated transcript approximately 220 nucleotides in length encoded by the HeLa cell mitochondrial genome. This transcript has been mapped by a combination of Southern and S1 protection experiments and shown to be complementary to the light strand of Hpa II fragment 8. 5'-end sequencing analysis demonstrates that the 5'-terminus of this RNA corresponds to a residue located 219 nucleotides from the origin of DNA replication.

## Introduction

The mitochondrial DNA (mtDNA) of HeLa cells has been shown to be completely and symmetrically transcribed (Aloni and Attardi, 1971a, b; Murphy et al., 1975). Recent pulse labeling experiments have indicated that the rate of transcription of the light (L) strand is two to three times higher than that of the heavy (H) strand (Cantatore and Attardi, 1980); in addition, for the most part the L-strand transcripts are rapidly degraded, producing only a few tRNAs and one small poly(A)-containing RNA (the 7S RNA) as stable products (Ojala and Attardi, 1974b; Angerer et al., 1976). The H-strand, on the other hand, encodes the majority of the tRNAs (Angerer et al., 1976), the two ribosomal RNAs (Attardi and Attardi, 1971) and (with the exception of the 7S RNA) all of the poly(A)-containing RNA species (Amalric et al., 1978) which by virtue of their relative abundance, their enrichment in polysomal structures, and their relatively long half-life (Attardi et al., 1979; Gelfand, 1980) are considered to be specific messenger RNAs. Thus there is an interesting contrast between the informational content of the two strands.

The 7S RNA species, previously designated as such by virtue of its sedimentation rate under denaturing conditions (Ojala and Attardi, 1974a) and more recently referred to as RNA 18, in the classification of Amalric et al. (1978) has long been of interest because of its strand specificity and because it exhibits a relatively high abundance (as compared to other polyadenylated species) in HeLa cells pulse labeled with [$^{32}$P] orthophosphate (in the presence or absence of inhibitors of nuclear RNA synthesis) (Ojala, 1974b; Amalric et al., 1978; Gelfand, 1980) while its half-life is comparable to that of the other species (Gelfand, 1980). In addition, its appearance in polyacrylamide gels under a variety of denaturing conditions is that of a broad and diffuse band, possibly indicative of multiple subcomponents.

To investigate the nature and possible function of this light strand transcript, studies were undertaken to precisely localize the DNA sequences encoding this species

within the HeLa cell mtDNA physical map.

## Materials and Methods

### Isolation and Purification of RNA 18 (7S RNA)

Preparation, by the micrococcal nuclease procedure, of mitochondrial RNA from HeLa cells labeled for 2.5 hr with [$^{32}$P] orthophosphate and subsequent isolation and fractionation of the oligo(dT)-cellulose bound RNA species by 1.4% agarose-$CH_3HgOH$ gel analysis were carried out as described (Crews and Attardi, 1980; Ojala et al., 1980). RNA 18 (7S RNA) was eluted from the gel and used in the experiments to be described below.

### Preliminary Mapping Experiments by the DNA Transfer and S1 Protection Techniques

Procedures for the isolation of closed circular mitochondrial DNA from either unlabeled cells or cells long term labeled with [$^{32}$P] orthophosphate and subsequent digestion of this DNA with Hpa II or Hpa II plus Bam H1 restriction enzymes followed by electrophoretic fractionation of the restriction products on 1.5% to 25% polyacrylamide gradient gels in Tris-acetate-EDTA buffer have been previously published (Ojala and Attardi, 1977). Hpa II fragment 8 was isolated for further use by visualization of the band with ethidium bromide or autoradiography of the wet gel, electroelution from the gel slice and ethanol precipitation.

The DNA transfer technique utilized Hpa II plus Bam H1 restriction digests of mtDNA or a Hae III digest of isolated Hpa II fragment 8 (Ojala and Attardi, 1977) resolved on 2% agarose gels and transferred to nitrocellulose filter strips (Southern, 1975). Hybridization with isolated 7S RNA labeled with [$^{32}$P] orthophosphate in vivo as described above or in vitro (Ojala and Attardi, 1980) followed the procedure of Ojala et al., (1980). A marker pattern was provided by a transferred Hpa II digest hybridized with an end-labeled and denatured total Hpa II digest (Ojala and Attardi, 1977).

For the S1 protection analysis, samples of purified in vivo labeled RNA 18 were dried down in snap cap vials with an excess of the light strand of Hpa II frgament 8 [strand separation of this fragment and determination of the strand specificity were carried out as described by Cantatore and Attardi (1980)]. Conditions for the hybridization reactions were 4 hr at 66°C in 0.4 m NaCl. Treatment of the hybrids with the Aspergillus orizae S1 nuclease and subsequent electrophoretic analysis of the protected DNA segments in 5% polyacrylamide slab gels in Tris-borate-$Mg^{++}$ buffer were carried out as described (Ojala et al., 1980a, b).

**Sequencing Analysis of 7S RNA**

7S RNA, isolated as described above, was labeled at its 5'-end with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase after treatment with bacterial alkaline phosphatase (Crews and Attardi, 1980). The labeled RNA was further purified by electrophoresis on a 5% polyacrylamide/5 M urea gel in Tris-borate-EDTA buffer (Maniatis et al., 1975). The band corresponding to the end labeled RNA was excised and eluted from the gel. Nuclease P1 analysis and RNA sequencing were carried out as previously described (Donis-Keller et al., 1977; Simoncsits et al., 1977; Fujimoto et al., 1974; Crews and Attardi, 1980).

## Results

**Purification and 5'-End Labeling of Mitochondrial 7S RNA**

Figure 1a shows an autoradiogram illustrating the electrophoretic pattern in a 1.4% agarose-$CH_3HgOH$ slab gel of the oligo(dT)-bound mitochondrial RNA isolated by the micrococcal nuclease procedure (Crews and Attardi, 1980) from HeLa cells labeled for 2.5 hr with [$^{32}$P] orthophosphate. The characteristic set of components previously described in HeLa cells (Amalric et al., 1978) are clearly recognizable, with the exception of the three largest species (RNAs 1, 2 and 3), which are extremely short-lived and present in mitochondria in very small amounts

(Attardi et al., 1979; Gelfand, 1980). The numerical designation of the individual components follows that used previously by Amalric et al. (1978). RNA species 4 to 17 have been previously shown to be coded for by the H-strand (Amalric et al., 1978), while RNA species 18 is encoded in the L-strand (Ojala and Attardi, 1974b). The relatively broad and diffuse appearance of the band representing RNA 18 is typical for that RNA species and has been observed in a variety of different denaturing gel systems (Ojala and Attardi, 1974a, b; Amalric et al., 1978; unpublished data).

The material corresponding to band 18 was eluted from the gel and labeled at its 5'-end with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase after treatment with bacterial alkaline phosphatase (Crews and Attardi, 1980). The labeled RNA was further purified by rerun on a 5% polyacrylamide/5 M urea gel (Figure 1b); the band indicated by the brackets was excised, and the material eluted for sequencing and 5'-terminal nucleotide analysis (see below).

**Preliminary Mapping Experiments**

As a first step towards localizing the sequences encoding the 7S RNA within the HeLa cell mtDNA Hpa II physical map (Ojala and Attardi, 1978) hybridization experiments by the DNA transfer technique (Southern, 1975) were carried out. As shown in Figure 1 (lane d), 7S RNA, isolated from material labeled and purified as depicted in lane a and hybridized to an Hpa II plus Bam H1 digest of HeLa cell mtDNA transferred to a nitrocellulose filter strip (Southern, 1975), was found to be complementary to Hpa II fragment 8, as identified by alignment with a marker strip containing an Hpa II digest transferred and hybridized with end-labeled Hpa II fragments (lane c). Furthermore, the sequences encoding this RNA were found to be located in the end of fragment 8 containing the origin of DNA replication (refer to Figure 4) as demonstrated by hybridization of in vitro labeled 7S RNA to both subfragments of Hpa II-8 [$\Delta$8a and $\Delta$8b, produced by the Hae III cleavage of that fragment (Ojala and Attardi,

1978)] resolved on agarose gel and transferred to a nitrocellulose filter (lane e).

A more precise mapping of the sequences within Hpa II fragment 8 encoding 7S RNA was then carried out by using the S1 protection technique (Berk and Sharp, 1977, 1978). In this approach, the L-strand of in vivo labeled Hpa II fragment 8 was hybridized with in vivo labeled 7S RNA. Hybridization conditions in 0.4 m NaCl, and subsequent digestion at 45°C of the hybrids by the Aspergillus orizae S1 nuclease followed by electrophoretic analysis of the hybrids under native conditions have previously been described in detail (Ojala and Attardi, 1980; Ojala et al., 1980b). As shown in lane g, a broad band of 200 to 215 nucleotides in length can be observed over a diffuse background of heterogeneous material. This heterogeneous material is due to the presence, in the hybridization mixture, of contaminating RNA sequences, as shown by a control reaction in which RNA alone was used (lane i). The control reaction in which DNA alone was used is shown in lane h. Size markers were provided by an end labeled Hpa II plus Bam H1 mtDNA digest (lane f). The size estimate for the 7S RNA species determined here (200-215 nucleotides) correlates well with that previously determined from its electrophoretic mobility in denaturing gels (∿230 nucleotides after subtraction of the poly(A) contribution) (Ojala and Attardi, 1974a). The fact that the protected hybrid is reflected by a relatively broad band may be due to the high AT content (∿80%) of the DNA sequence in the regions of the 3'- and 5'-ends of the transcript (refer to Figure 3 and Crews et al., 1979) which would favor "breathing" and therefore, nibbling, of the ends of the hybrid by the S1 nuclease during the 45°C digestion step. An alternative explanation is also possible (see **Discussion**).

**7S RNA Sequence Analysis**

The in vitro 5'-end labeled 7S RNA was first characterized as to its 5'-terminal nucleotide by exhaustive digestion with nuclease P1 and fractionation of the products

on PE1-cellulose TLC plates, as previously described (Crews and Attardi, 1980). The percentage of total [$^{32}$P] radioactivity among the four monophosphates was found to be as follows: AMP, 90.0%; UMP, 4.7%; GMP, 4.1%; CMP, 1.2%. These determinations supported the substantial purity of the RNA preparation.

The remainder of the in vitro labeled 7S RNA was subjected to base specific partial enzymatic hydrolysis according to the procedures of Donis-Keller et al. (1977) and Simoncsits et al. (1977) followed by analysis of the digestion products on a 25% polyacrylamide/7 M urea gel as previously described (Crews and Attardi, 1980). The autoradiograph of the sequencing gel (Figure 2) reveals a complex nucleotide profile in which two, or in some cases, three oligonucleotides are observed at some of the same ladder positions. Similar observations have been made on other mtRNA species sequenced in this laboratory; those results were attributed to the presence, in the sequencing reactions, of multiple species labeled at the 5'-end. Thus, the 12S ribosomal RNA was found to consist of two species differing in molecular size by one nucleotide (Crews and Attardi, 1980), and poly(A)-containing RNA 7, also to consist of two species differing in size from each other by one nucleotide (Montoya, Ojala and Attardi, manuscript in preparation). The presence of multiple species produces, in the sequencing gels, a shadowing effect in which a specific oligonucleotide is found in adjacent ladder positions (the number of adjacent positions reflecting the number of species present). Therefore, when the 7S RNA is considered to consist of four species (two major and two minor) which differ from each other by the successive loss of the 5'-terminal nucleotide, as illustrated in Table 1, a tentative sequence assignment can be made through position 30. Confirmation of this assignment is shown in Figure 3, in which the RNA sequence has been aligned with the previously determined sequence of mtDNA Hpa II fragment 8 (Crews et al., 1979).

In general, the sequencing method used here showed the same characteristics as observed by others; RNases T1 and U2 cleave regularly after Gs and As while

RNase A fails to cleave after some pyrimidines, especially when present in pyrimidine stretches. However, a striking observation in this sequencing gel is the presence of doublets at ladder positions 2 to 8. This was interpreted to reflect the formation, during the enzymatic or chemical cleavage of the RNA, of oligonucleotides with different phosphate end groups (2',3'-cyclic phosphate, 2' and 3' phosphate).

## Discussion

Figure 4 illustrates the organization of the HeLa cell mitochondrial genome in the region of the origin of DNA replication. Depicted in this schematic drawing are the precise positions of the 12S ribosomal RNA gene and of the tRNA genes (Crews and Attardi, 1980; B. Barrell and F. Sanger, personal communication) and that of the 7S RNA coding sequence, as determined in the present work.

Results of the mapping and S1 protection experiments presented here indicate that the 7S RNA is complementary to the light strand of Hpa II fragment 8 and has a length of 200 to 215 nucleotides (which may be an underestimate due to nibbling of the ends of the protected hybrid by S1 nuclease); this size estimate is in good agreement with that previously estimated on the basis of its electrophoretic mobility in denaturing gels (230 nucleotides). Sequencing analysis show that the 5'-terminus of this RNA corresponds to a residue in the L-strand of Hpa II fragment 8 at 219 nucleotides from the origin of DNA replication. The 3'-end would thus be placed at, or very close to, the replication origin.

In spite of the intriguing possiblities suggested by this mapped position, the actual function of the 7S RNA remains unknown. Three posibilities exist. It could encode a small polypeptide or function, in some manner, as a primer for replication of the H-strand; thirdly, it could be involved in the initiation of transcription of the L-strand. Support for the first hypothesis is provided by the fact that there is an open reading frame starting at 65 nucleotides and ending at 199 nucleotides

from the 5'-end of the 7S RNA, which would code for a polypeptide 45 amino acids long. However, this frame contains two arginine codons (AGA and AGG) which have been suggested to represent termination codons in human mitochondria (Barrell et al., 1980). If this codon assignment is correct, then the only open reading frame remaining in 7S RNA would be that starting at nucleotide 128 or 131 and ending at nucleotide 199 from the 5'-end, which would code for a polypeptide 24 or 23 amino acids long, respectively. Both the large and the small hypothetical polypeptides would have a considerable proportion of hydrophobic amino acids. It should be emphasized that, although small polypeptides are synthesized in HeLa cell mitochondria (Ching, 1980), there is no evidence that the 7S RNA functions as mRNA.

The second possible function of this RNA species, that of priming H-strand DNA synthesis, receives support from analogous systems in which evidence has been reported suggesting that RNA priming is utilized for initiation of DNA synthesis. Thus, in mouse L-cells, 5'-end group analysis of the 7S mtDNA sequence (which primes H-strand replication) was found to contain ribonucleotides (representing ∿20% of the total 5'-ends) at a few specific positions (Bogenhagen et al., 1978). It was proposed that these ribonucleotides are residual portions of primer RNA used to initiate DNA synthesis, and as such, are the first evidence for RNA priming during origin initiation of DNA synthesis in higher cells. Secondly, Itoh and Tomizawa (1980) have proposed that the replication of the closed circular DNA of plasmid Col El of Escherichia coli involves a mechanism by which RNA polymerase synthesizes a transcript that is processed by RNase H and then used as a primer by DNA polymerase I. In this case transcription of the primer frequently extends past the origin and terminates at a number of downstream sites.

Evidence for the third possibility, in which this RNA may function as an initiator of L-strand transcription, has been provided by mapping studies in which nascent RNA molecules were annealed with the L-strands of mtDNA Hpa II fragments

(Cantatore and Attardi, 1980). As analyzed by the S1 protection technique, the protected hybrids involved fragments located in the quadrant region adjacent to the origin of replication in the direction of L-strand transcription. These results support the possibility that a promoter for L-strand transcription is located in the region of the genome near the origin of DNA replication.

As previously mentioned, the HeLa cell mt 7S RNA exhibits a broad band in all gel systems in which it has been analyzed. Because the sequence analysis described in the present work reveals that this RNA is presumably transcribed from a unique position on the L-strand (assuming the observed presence of the four molecular species results from a subsequent successive loss of the 5'-terminal nucleotide by an in vivo processing mechanism or perhaps by degradation during the isolation procedure), the broad band exhibited in the gels (encompassing a range of ∿35 nucleotides) may result from variation in the length of the poly(A) stretch, or from variation in its point of attachment. In an effort to determine the precise position(s) of the 3'-terminus of the 7S RNA, experiments using phased oligo(dT) primers in conjunction with cDNA synthesis by reverse transcriptase are now being carried out. [This methodology has been successful in mapping the 3'-end of many of the H-strand coded poly(A)-containing RNA species (see Chapter 7).] Preliminary results obtained for the 7S RNA support the possibility that termination of this transcript occurs at multiple positions on the genome. Hopefully, further analyses will refine these results and define more precisely the function of this unusual RNA species.

**Acknowledgements**

The work described in this chapter was carried out in collaboration with Steve Crews and Robert Gelfand. In addition, experiments are now in progress by Dr. Julio Montoya to further characterize the 3'-end of the 7S RNA.

## References

Aloni, Y. and Attardi, G. (1971a) Expression of the mitochondrial genome in HeLa cells. II. Evidence for complete transcription of mitochondrial DNA. J. Mol. Biol. **55**, 251-270.

Aloni, Y. and Attardi, G. (1971b) Expression of the mitochondrial genome in HeLa cells. VII. Symmetric in vivo transcription of mitochondrial DNA in HeLa cells. Proc. Nat. Acad. Sci. USA **68**, 1757-1761.

Amalric, F., Merkel, C., Gelfand, R. and Attardi, G. (1978) Fractionation of mitochondrial RNA from HeLa cells by high-resolution electrophoresis under strongly denaturing conditions. J. Mol. Biol. **118**, 1-25.

Angerer, L., Davidson, N., Murphy, W., Lynch, D. and Attardi, G. (1976) An electron microscope study of the relative positions of the 4S and ribosomal RNA genes in HeLa cell mitochondrial DNA. Cell **9**, 81-90.

Attardi, B. and Attardi, G. (1971) Expression of the mitochondrial genome in HeLa cells. I. Properties of the discrete RNA components from the mitochondrial fraction. J. Mol. Biol. **55**, 231.

Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C. and Ojala, D. (1979) The organization of the genes in the human mitochondrial genome and their mode of transcription. In: Extrachromosomal DNA, Vol. 15. ICN-UCLA Symp. on Molec. and Cellular Biol. Eds. D. Cummings, P. Borst, I. Dawid, S. Weissman and C. F. Fox (Academic Press, New York), pp. 443-469.

Berk, A. J. and Sharp, P. A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**, 721-732.

Berk, A. J. and Sharp, P. A. (1978) Spliced early mRNAs of simian virus 40. Proc. Nat. Acad. Sci. USA **75**, 1274-1278.

Bogenhagen, D., Gillum, A. M., Martens, P. A. and Clayten, D. A. (1978) Replication of mouse L-cell mitochondrial DNA. Cold Spring Harbor Symp. Quant. Biol. **43**, 253-262.

Ching, E. P. (1979) Biochemical and functional characterization of the products of mitochondrial protein synthesis in HeLa cells. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Crews, S. and Attardi, G. (1980) The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. Cell **19**, 775-784.

Crews, S., Ojala, D., Posakony, J., Nishiguchi, J. and Attardi, G. (1979) Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. Nature **277**, 192-198.

Cantatore, P. and Attardi, G. (1980) Mapping of nascent light- and heavy-strand transcripts on the physical map of HeLa cell mitochondrial DNA. Nucl. Acids Res. **8**, 2605-2624.

Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) Mapping adenines, guanines, and pyrimidines in RNA. Nucl. Acids Res. **4**, 2527-2538.

Fujimoto, M., Kuninaka, A. and Yoshino, H. (1974) Substrate specificity of nuclease P1. Agr. Biol. Chem. **38**, 1555-1561.

Gelfand, R. A. (1980) Studies on RNA metabolism in HeLa mitochondria. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Itoh, T. and Tomizawa, J.-I. (1980) Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc. Nat. Acad. Sci. USA **77**, 2450-2454.

Maniatis, T., Jeffrey, A. and van de Sande, H. (1975) Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry **14**, 3787-3794.

Murphy, W. I., Attardi, B., Tu, C. and Attardi, G. (1975) Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in HeLa cells. J. Mol. Biol. **99**, 809-814.

Ojala, D. and Attardi, G. (1974a) Identification of discrete polyadenylate-containing RNA components transcribed from HeLa cell mitochondrial DNA. (XX. in the series, Expression of the Mitochondrial Genome in HeLa Cells.) Proc. Nat. Acad. Sci. USA **71**, 563-567.

Ojala, D. and Attardi, G. (1974b) Identification and partial characterization of multiple discrete polyadenylic acid-containing RNA components coded for by HeLa cell mitochondrial DNA. J. Mol. Biol. **88**, 205-219.

Ojala, D. and Attardi, G. (1977) A detailed physical map of HeLa cell mitochondrial DNA and its alignment with the positions of known genetic markers. Plasmid **1**, 78-105.

Ojala, D. and Attardi, G. (1978) Precise localization of the origin of replication in a physical map of HeLa cell mitochondrial DNA and isolation of a small fragment that contains it. J. Mol. Biol. **122**, 301-319.

Ojala, D. and Attardi, G. (1980) Fine mapping of the ribosomal RNA genes of HeLa cell mitochondrial DNA. J. Mol. Biol. **138**, 411-420.

Ojala, D., Merkel, C., Gelfand, R. and Attardi, G. (1980b) The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell, in press.

Ojala, D., Montoya, J. and Attardi, G. (1980a) The putative mRNA for subunit II of cytochrome c oxidase in human mitochondria starts directly at the translation initiator codon. Nature **287**, 79-82.

Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. and Guilley, H. (1977) New rapid gel sequencing method for RNA. Nature **269**, 833-836.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**, 503-517.

**Table 1**

Interpretation of the RNA Sequencing Gel

| | Ladder Position Number[a] | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Molecular Species[b] | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | A | A | A | G | A | Y | A | A | A | A | Y | Y | Y | G | A | A | A | Y | Y | Y | G | G | Y | Y | A | G | G | Y | Y | G |
| 2 | A | A | G | A | Y | A | A | A | A | Y | Y | Y | G | A | A | A | Y | Y | Y | G | G | Y | Y | A | G | G | Y | Y | G | |
| 3 | A | G | A | Y | A | A | A | A | Y | Y | Y | G | A | A | A | Y | | | | | | | | | | | | | | |
| 4 | G | A | Y | A | A | A | A | Y | Y | Y | G | A | A | A | | | | | | | | | | | | | | | | |

[a]Denotes the ladder position in the RNA sequencing gel (Figure 3) as indicated in that figure by a dashed line or bracket.

[b]Refers to molecular species of the 7S RNA differing at their 5'-end by the successive loss of the 5'-terminal nucleotide and correlated with the observed bands in the sequencing gel.

**Figure 1**

Isolation, 5'-end labeling and preliminary mapping of mitochondrial 7S RNA. Lane a: autoradiogram of the oligo(dT)-bound fraction of mitochondrial RNA isolated by the micrococcal nuclease procedure from HeLa cells labeled for 2.5 hr with [$^{32}$P] orthophosphate and fractionated by electrophoresis through a 1.4% agarose-$CH_3HgOH$ gel. Lane b: rerun, through a 5% polyacrylamide/5 M urea gel, of 7S RNA (RNA 18) after elution from the gel track shown in lane a and 5'-end labeling with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase. Lanes c, d and e: hybridization by the DNA transfer technique. 7S RNA was labeled in vivo and hybridized with a transferred Hpa II plus Bam H1 digest (lane d), or labeled in vitro and hybridized with a transferred Hae III digest of Hpa II fragment 8 (lane e). Lane c is a marker pattern of an end labeled and denatured Hpa II digest. Lanes f to i: nuclease S1 sizing of 7S RNA. The L-strand of Hpa II fragment 8 was hybridized to 7S RNA for 4 hr at 66°C in 0.4 m NaCl, treated with S1 nuclease and the protected hybrid analyzed by electrophoresis through a 5% polyacrylamide gel in Tris-borate-$Mg^{++}$ buffer (lane g). Control reactions in which the RNA or DNA were omitted are shown in lanes h and i, respectively. Lane f: size markers provided by end labeled Hpa II plus Bam H1 digest.

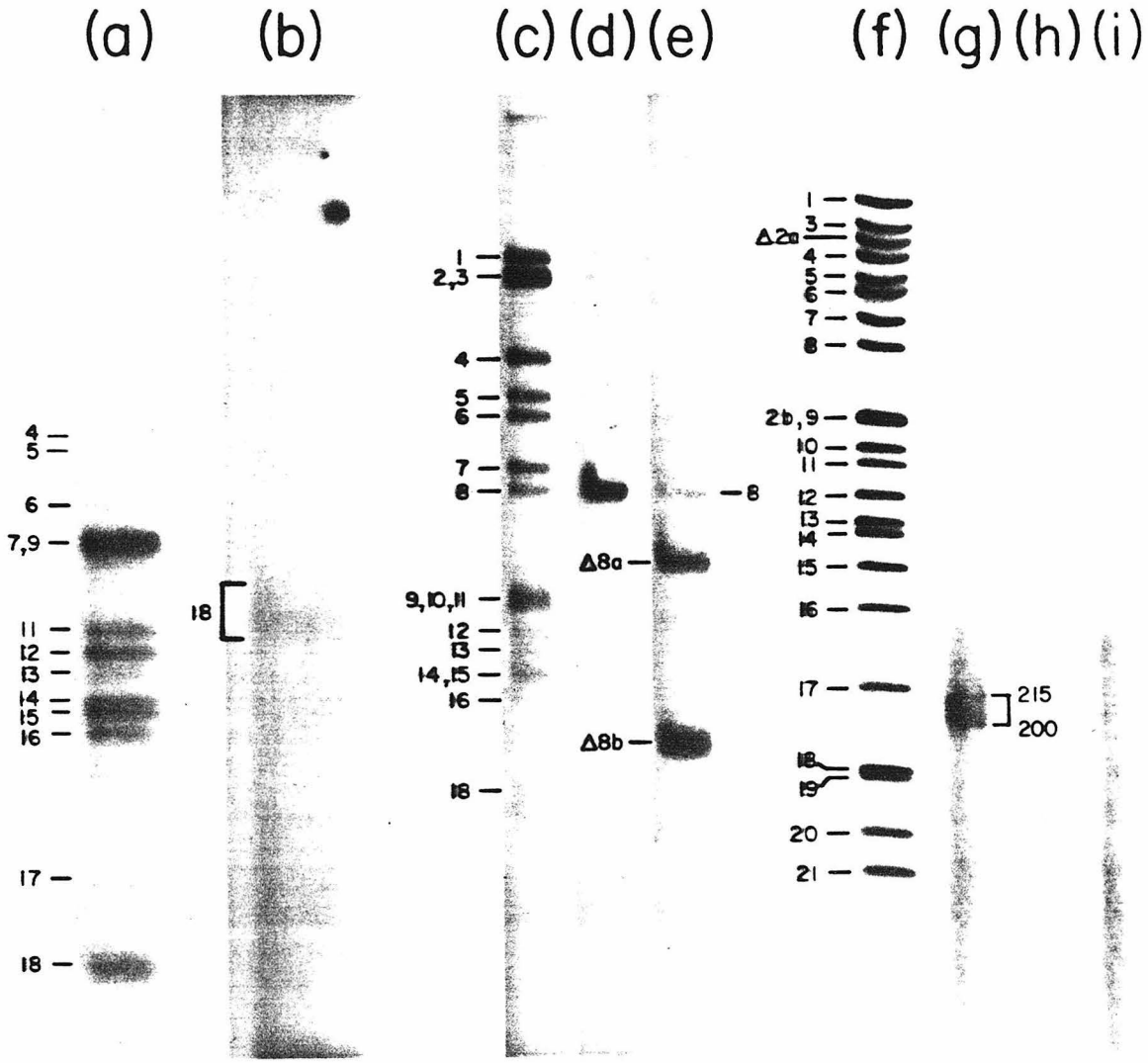

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
4
5
6
7,9
11
12
13
14
15
16
17
18
18
1
2,3
4
5
6
7
8
9,10,11
12
13
14,15
16
18
8
Δ8a
Δ8b
1
3
Δ2a
4
5
6
7
8
2b,9
10
11
12
13
14
15
16
17
18
19
20
21
215
200

**Figure 2**

Sequencing gel of mitochondrial 7S RNA. Autoradiogram of 25% polyacrylamide/urea gel of partial enzymatic digestions of 5'-end labeled 7S RNA. Shown at the top of the gel are the cleavage specificities: RNase A (C + U), RNase U2 (A), RNase T1 (G), formamide ladder ($L_F$) and RNA incubated without enzyme (-Enz). Numbers indicate designated ladder positions (brackets at positions 2 to 8 indicate doublets at those positions resulting, presumably, from the formation during the enzymatic or chemical cleavage steps of oligonucleotides with different phosphate end-groups).

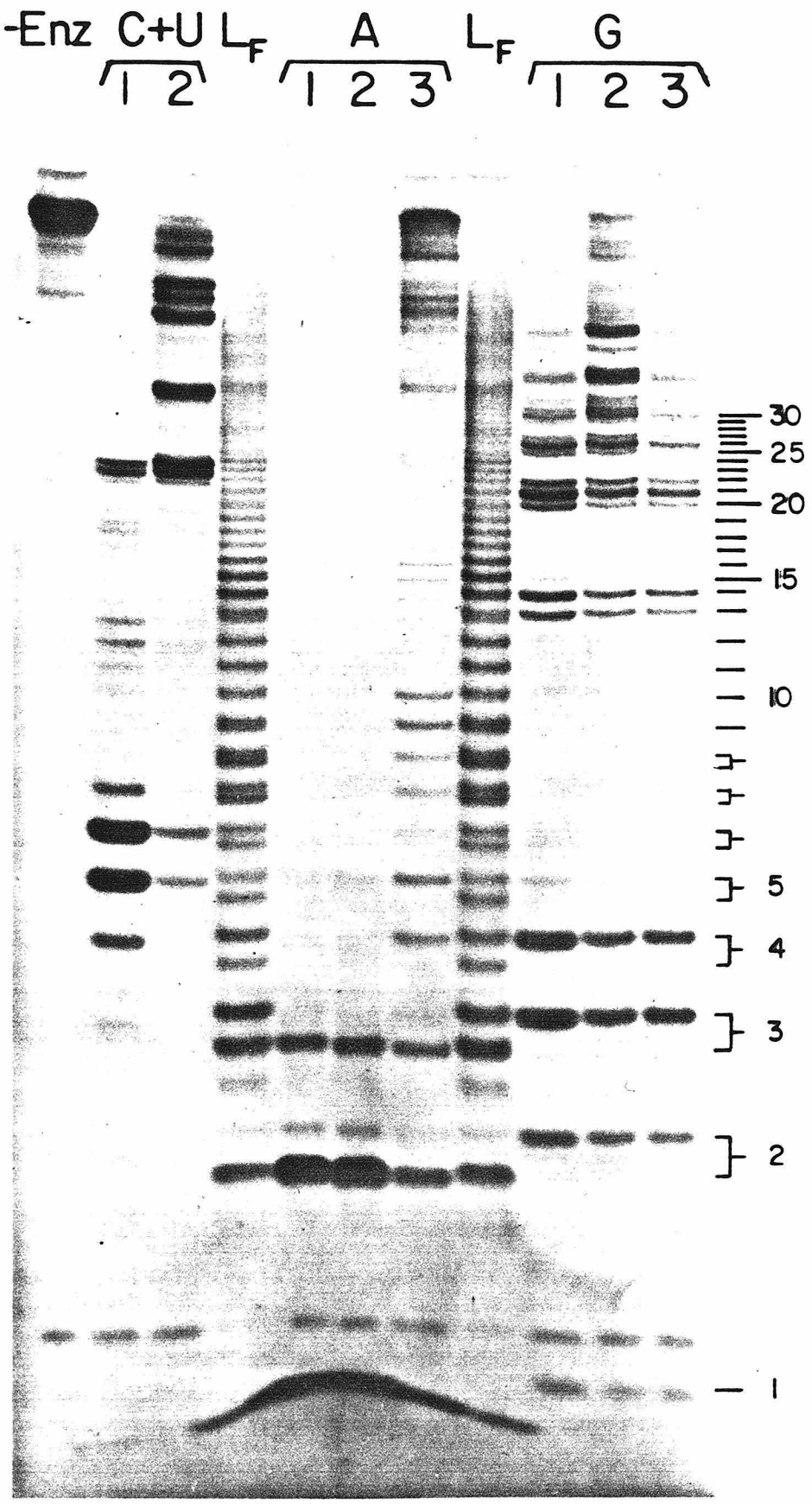

-Enz
C+U
1 2
$L_F$
A
1 2 3
$L_F$
G
1 2 3
30
25
20
15
10
5
4
3
2
1

**Figure 3**

Alignment of the nucleotide sequence of the 5'-end region of 7S RNA with the DNA sequence of Hpa II fragment 8. The sequence of the 5'-terminal segment of 7S DNA and its alignment with the Hpa II-8 sequence is also shown.

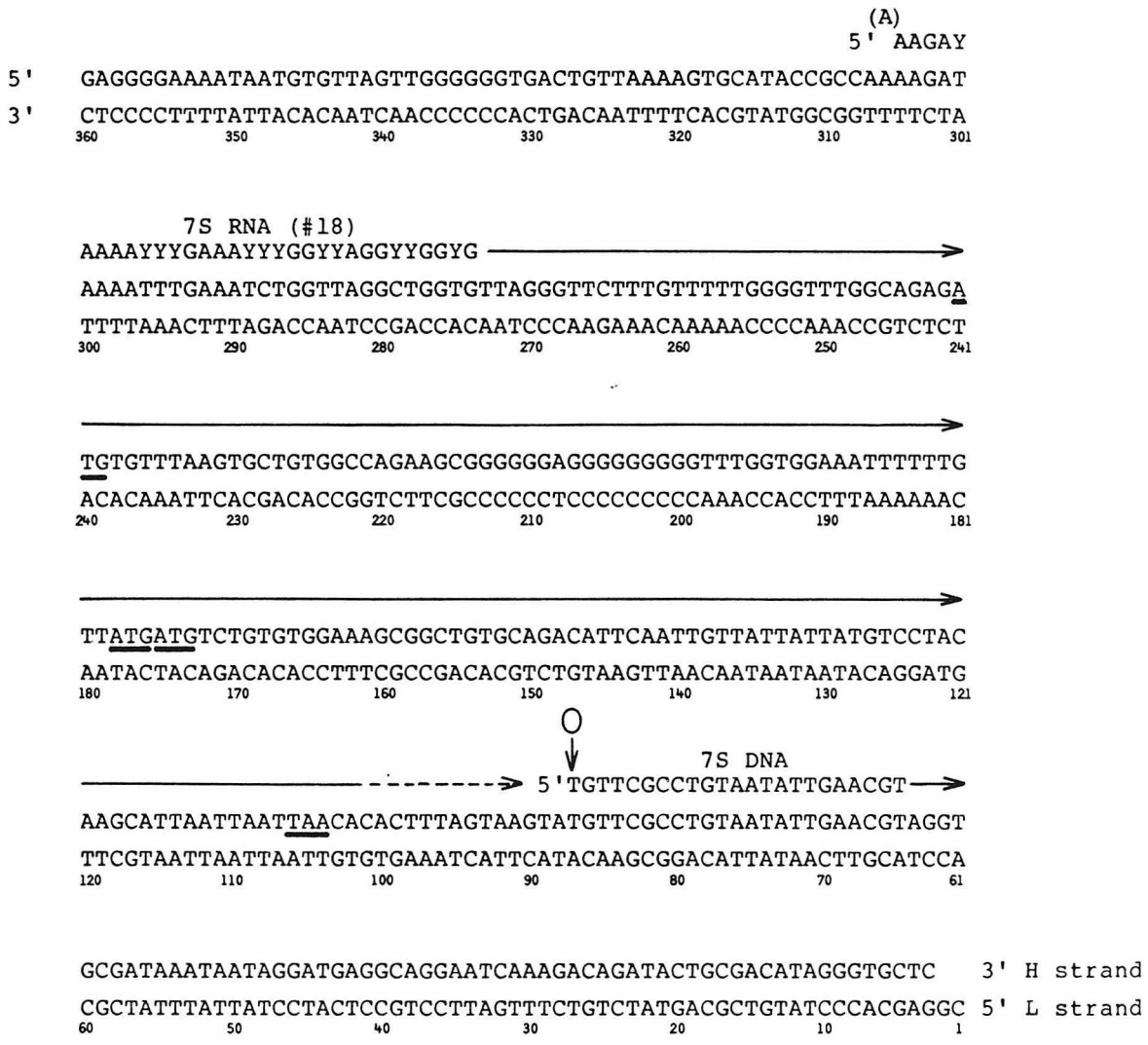
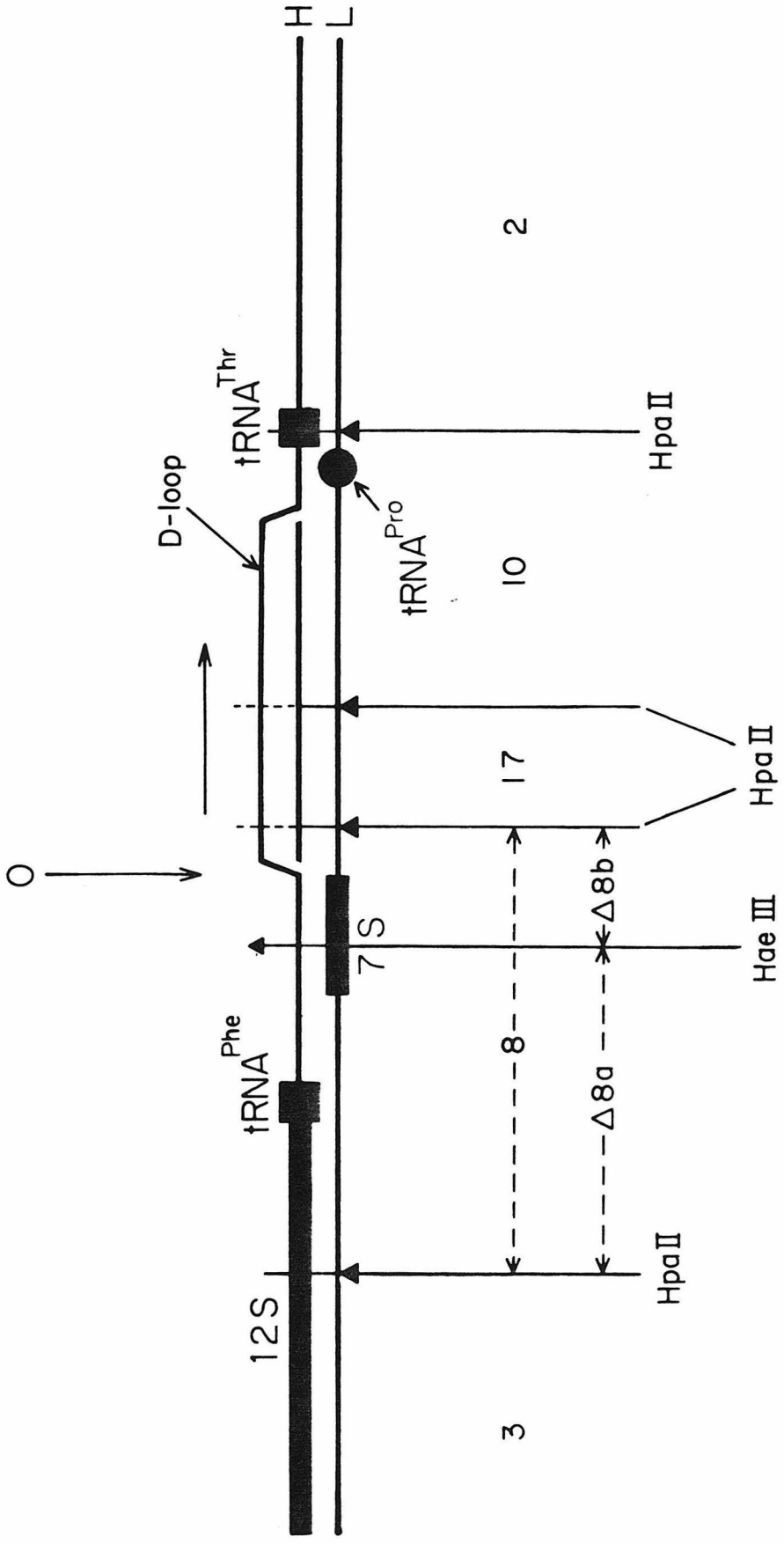

(A)
5' AAGAY
5' GAGGGGAAAATAATGTGTTAGTTGGGGGGTGACTGTTAAAAGTGCATACCGCCAAAAGAT
3' CTCCCCTTTTATTACACAATCAACCCCCCACTGACAATTTTCACGTATGGCGGTTTTCTA
360 350 340 330 320 310 301
7S RNA (#18)
AAAAYYYGAAAYYYGGYYAGGYYGGYG
AAAATTTGAAATCTGGTTAGGCTGGTGTTAGGGTTCTTTGTTTTTGGGGTTTGGCAGAGA
TTTTAAACTTTAGACCAATCCGACCACAATCCCAAGAAACAAAAACCCCAAACCGTCTCT
300 290 280 270 260 250 241
TGTGTTTAAGTGCTGTGGCCAGAAGCGGGGGGAGGGGGGGGGTTTGGTGGAAATTTTTTG
ACACAAATTCACGACACCGGTCTTCGCCCCCCTCCCCCCCCCAAACCACCTTTAAAAAAC
240 230 220 210 200 190 181
TTATGATGTCTGTGTGGAAAGCGGCTGTGCAGACATTCAATTGTTATTATTATGTCCTAC
AATACTACAGACACACCTTTCGCCGACACGTCTGTAAGTTAACAATAATAATACAGGATG
180 170 160 150 140 130 121
7S DNA
5'TGTTCGCCTGTAATATTGAACGT
AAGCATTAATTAATTAACACACTTTAGTAAGTATGTTCGCCTGTAATATTGAACGTAGGT
TTCGTAATTAATTAATTGTGTGAAATCATTCATACAAGCGGACATTATAACTTGCATCCA
120 110 100 90 80 70 61
GCGATAAATAATAGGATGAGGCAGGAATCAAAGACAGATACTGCGACATAGGGTGCTC 3' H strand
CGCTATTTATTATCCTACTCCGTCCTTAGTTTCTGTCTATGACGCTGTATCCCACGAGGC 5' L strand
60 50 40 30 20 10 1

**Figure 4**

Region of the HeLa cell mtDNA physical map illustrating the gene organization around the origin of DNA replication (O). The precise positions of the 12S rRNA gene and of the tRNA genes on the H-strand (■) and L-strand (●) as derived from mtDNA sequence data (Crews and Attardi, 1980; F. Sanger and B. Burrell, personal communication) and that of the 7S RNA coding sequence on the L-strand, as determined in the present work, are shown. The rightward arrow above the D-loop indicates the direction of H-strand synthesis as well as the direction of transcription of the 7S RNA. The vertical solid arrows indicate Hpa II and Hae III cleavage sites.


H
L
2
tRNA^Thr
HpaII
D-loop
tRNA^Pro
10
17
HpaII
O
Δ8b
7S
HaeIII
8
tRNA^Phe
Δ8a
12S
HpaII
3

# Chapter 5

## The Putative mRNA for Subunit II of Cytochrome c Oxidase in Human Mitochondria Starts Directly at the Translation Initiator Codon

It has been recently shown that the gene for subunit II of cytochrome c oxidase (COII) in human mitochondrial DNA (mtDNA) is immediately contiguous, on its 5'-end side, to a $tRNA^{Asp}$ gene[1]. Since all eukaryotic mRNAs so far analyzed have been shown to have on the 5'-side of the coding sequence a noncoding stretch, which is presumably used for ribosome attachment[2,3], it was reasonable to ask whether, in the case of the human COII mRNA, the above function is performed by the $tRNA^{Asp}$ sequence or a portion of it, or whether on the contrary this mRNA lacks completely a 5' noncoding stretch. In the present study, a 5'-end proximal segment of about 30 nucleotides of the putative COII mRNA from HeLa cells has been sequenced and aligned with the COII coding sequence in human mtDNA. The results clearly show that this RNA starts directly at the AUG initiator codon.

Figure 1 shows a region of the HpaII restriction map of HeLa cell mtDNA[4] with the precise positions of the COII gene and of the adjacent putative tRNA genes, as determined from DNA sequencing analysis[1]. Preliminary RNA mapping experiments using the Southern[5] and Alwine et al.[6] methods had shown that a single mtDNA heavy (H) strand coded poly(A)-containing RNA (RNA species 16, in the classification of Amalric et al.[7]) maps in correspondence to the COII gene. This RNA, on account of its size (about 800 nucleotides), its abundance, its presence in polysomes and its relatively long half-life, is very probably the mRNA for subunit II. Recently, by using the S1 protection technique[8,9], it has been possible to localize the 5'-end of RNA 16 more precisely, i.e., at about 540 nucleotides from the HpaII site between fragments 7 and 22; in the same experiments, the 3'-end of the non-poly(A) segment of this RNA has been mapped in the region of the $tRNA^{Lys}$ gene, with indications that it may overlap, in part, or possibly completely the $tRNA^{Lys}$ coding sequence (Ojala, Merkel, Gelfand & Attardi, manuscript in preparation) (Fig. 1).

In order to prepare RNA 16 of high purity and in adequate amounts for sequencing analysis, the micrococcal nuclease procedure for purification of mitochondrial

RNA[10] was scaled up as described in the legend for Fig. 2. Lane (a) in this figure shows the electrophoretic pattern in a $CH_3HgOH$-agarose slab gel, after ethidium bromide staining, of the oligo(dT)-bound fraction of mitochondrial RNA derived from about $2.5 \times 10^9$ HeLa cells. The typical pattern of the mtDNA-coded poly(A)-containing RNA species previously described in HeLa cells[7] is clearly recognizable, with the exception of the species known to be relatively short-lived and present in very small amounts (RNA species 1-4 and 6)[11,12]. RNA species 16 was eluted from the gel track shown in Fig. 2a and an identical track, treated with RNase-free DNase (to eliminate any contaminating DNA fragments), and rerun on a gel under the same conditions. A sharp band was observed at the expected position (Fig. 2b) and the RNA in this band was eluted. Material from about $8 \times 10^9$ cells, purified as described above, was labeled at its 5'-end with $\gamma$-$^{32}$P-ATP and polynucleotide kinase after dephosphorylation with bacterial alkaline phosphatase. The labeled RNA was run on a $CH_3HgOH$-agarose gel, a band corresponding to intact RNA 16 was observed (Fig. 2c) and the RNA in this band was eluted for the sequencing reactions. Mitochondrial 12S rRNA, to be used as a control for the in vitro labeling and sequencing technology used in the present work, was purified from the oligo(dT)-unbound RNA fraction of micrococcal nuclease treated mitochondria and labeled in vitro at its 5'-end as described above.

In order to characterize the 5'-terminus of RNA 16, the in vitro 5'-end labeled RNA was subjected to a complete P1[13] digestion and fractionation on a PEI-cellulose TLC plate. As a control, 5'-end-labeled 12S rRNA was subjected to the same analysis. As shown in Table 1, almost 92% of the radioactivity in the 12S rRNA was found to be associated with pA, the remainder being distributed among the three other nucleotides, in perfect agreement with previous work[10]. In RNA 16, the predominant 5' terminal nucleotide ($\sim$84%) was pA. This result indicated the substantial purity of the RNA 16 used for sequencing in the present work.

Sequencing of the 5' end proximal segment of RNA 16 and of 12S rRNA was performed by the methods of Donis-Keller et al.[14] and Simoncsits et al.[15], utilizing base-specific enzymatic cleavage of the RNA followed by electrophoresis on polyacrylamide/urea gels. The RNases T1 (Sigma) (G-specific), A (Sigma) (C+U-specific) and U2 (Sankyo) (A-specific) were used for this analysis. Since the purpose of this work was the alignment of the RNA sequence with the known DNA sequence, the precise identification of the pyrimidines was deemed to be unnecessary.

The 12S rRNA gel patterns (not shown) yielded a 5'-end-proximal sequence identical to that previously shown for this RNA species[10]. The gel patterns for RNA 16 (Fig. 3) gave a 5' end proximal sequence which could be read for 29 nucleotides (Fig. 3). The band corresponding to the second nucleotide (a pyrimidine) was very faint in the RNA sample digested with 0.25 ng RNase A (Fig. 3, left panel), but was clearly visible in the sample digested with higher enzyme levels (Fig. 3, right panel). There were a few extra bands appearing in the region corresponding to small oligonucleotides after RNase U2 or T1 digestion (and also after treatment of the sample with the higher levels of RNase A). These bands were in most cases less strong than the bona fide specific RNase cleavage products; because of this and because they did not correspond in migration to bands visible in the ladder or corresponded to obviously abnormal ladder steps (like the fairly strong band corresponding to a pyrimidine in the second lane of the right panel), they could be in general recognized as spurious bands. The nature of these extra bands is not clear; they may reflect the presence in the RNA 16 preparation, during the 5'-end labeling reaction, of contaminating oligonucleotides, or may be due to the formation, during the enzymatic or chemical degradation of the RNA, of oligonucleotides with different phosphate end-groups (2',3'-cyclic phosphate, 2' and 3' phosphate). Similar bands migrating in abnormal position have been observed, previously[10] and in the present work, in the sequencing gel patterns for 12S rRNA. The reading of the first few nucleotides

of the sequence near the 5'-end was made somewhat difficult by the presence of these extra bands; beyond the fifth nucleotide, however, the sequence could be read unambiguously. The sequence obtained is shown in Fig. 4, aligned with the DNA sequence of the COII gene region. The entire 29 nucleotide stretch of RNA 16 which has been sequenced appears to be colinear with the DNA sequence. The striking result is that the 5'-end of the RNA corresponds precisely to the first nucleotide of the COII coding sequence.

The observation that the COII mRNA starts directly at the initiator codon raises interesting questions about the mechanism whereby the mitochondrial ribosomes attach to this messenger. In all eukaryotic mRNAs analyzed so far there is a stretch of variable length which precedes the initiator codon and which is assumed to be involved in ribosome attachment[2,3]. In E. coli mRNAs there is likewise a 5'-noncoding segment containing a ribosome binding site[16,17]. A single exception to this rule has been described, namely that of the lambda repressor mRNA involved in repressor maintenance[18]. In this case, it was argued that the lack of a strong ribosome binding site could account for the low efficiency of translation of this mRNA. In the case of the mitochondria from human cells (and probably from other animal cells) it is reasonable to think that the special features of their ribosomes make them suitable for binding directly to the initiator codon. It will be interesting to see whether the lack of a 5'-noncoding stretch is a general feature of mitochondrial mRNAs in HeLa cells. Preliminary observations indicate that another mitochondrial poly(A)-containing RNA in these cells, RNA 15, starts with an AUG, which may be the initiator codon for the polypeptide coded by this RNA (Montoya, Ojala & Attardi, unpublished observations).

This work was supported by NIH grant GM-11726. We thank S. Crews for valuable discussions and A. Drew for technical assistance.

DEANNA K. OJALA

JULIO MONTOYA

GIUSEPPE ATTARDI

Division of Biology

California Institute of Technology

Pasadena, California 91125

**References**


1. Barrell, B. G., Bankier, A. T. & Drouin, J. Nature **282**, 189-194 (1979).
2. Hagenbüchle, O., Santer, M., Steitz, J. A. & Mans, R. J. Cell **13**, 551-563 (1978).
3. Kozak, M. Cell **15**, 1109-1123 (1978).
4. Ojala, D. & Attardi, G. Plasmid **1**, 78-105 (1977).
5. Southern, E. M. J. molec. Biol. **98**, 503-517 (1975).
6. Alwine, J. C., Kemp, D. J. & Stark, G. R. Proc. natn. Acad. Sci. U.S.A. **74**, 5350-5354 (1977).
7. Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. J. molec. Biol. **118**, 1-25 (1978).
8. Berk, A. J. & Sharp, P. A. Cell **12**, 721-732 (1977).
9. Berk, A. J. & Sharp. P. A. Proc. natn. Acad. Sci. U.S.A. **75**, 1274-1278 (1978).
10. Crews, S. & Attardi, G. Cell **19**, 775-784 (1980).
11. Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C. & Ojala, D. in Extrachromosomal DNA. ICN-UCLA Symposia on Molecular and Cellular Biology, **15** (eds Cummings, D., et al.) 443-469 (Academic Press, New York, 1979).
12. Gelfand, R. Ph.D. Thesis, California Institute of Technology, Pasadena, California, U.S.A.
13. Fujimoto, M., Kuninaka, A. & Yoshino, H. Agr. Biol. Chem. **38**, 1555-1561.
14. Donis-Keller, H., Maxam, A. M. & Gilbert, W. Nucleic Acids Res. **4**, 2527-2538 (1977).
15. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guilley, H. Nature **269**, 833-836 (1977).
16. Shine, J. & Dalgarno, L. Biochem. J. **141**, 609-615 (1974).
17. Steitz, J. A. & Jakes, K. Proc. natn. Acad. Sci. U.S.A. **72**, 4734-4738 (1975).
18. Ptashne, M., Backman, K., Humayun, M. Z., Jeffrey, A., Mauer, R., Meyer, B. & Sauer, R. T. Science **194**, 156-161 (1976).

**Table 1** 5'-terminal nucleotide analysis of in vitro labeled mitochondrial poly(A)-containing RNA 16 and 12S rRNA

| | Percentage of total $^{32}P$ radioactivity* | | | |
|---|---|---|---|---|
| RNA species | AMP | CMP | GMP | UMP |
| 12S rRNA | 91.8 | 1.9 | 2.0 | 4.2 |
| Poly(A)-containing RNA 16 | 83.6 | 1.3 | 8.2 | 7.0 |

The 5'-terminal nucleotide was determined by exhaustive digestion of the in vitro labeled RNAs with nuclease P1[13] (Calbiochem) and fractionation of the products on PEI-cellulose TLC plates, as previously described[10].

*The values refer to the radioactivity which migrated from the origin. The % of the input radioactivity recovered from the origin was, respectively, 1.0 and 1.6% for 12S rRNA and RNA 16.

**Figure 1**

Region of the HpaII restriction map of HeLa cell mtDNA[4] showing the precise positions of the COII gene and of the adjacent tRNA genes. The positions of the genes were derived from the published DNA sequencing analysis[1]. The mapping position of poly(A)-containing RNA 16 (Ojala, Merkel, Gelfand & Attardi, manuscript in preparation) is shown in the lower part of the figure; the dashed portion indicates the limits of uncertainty in this determination.

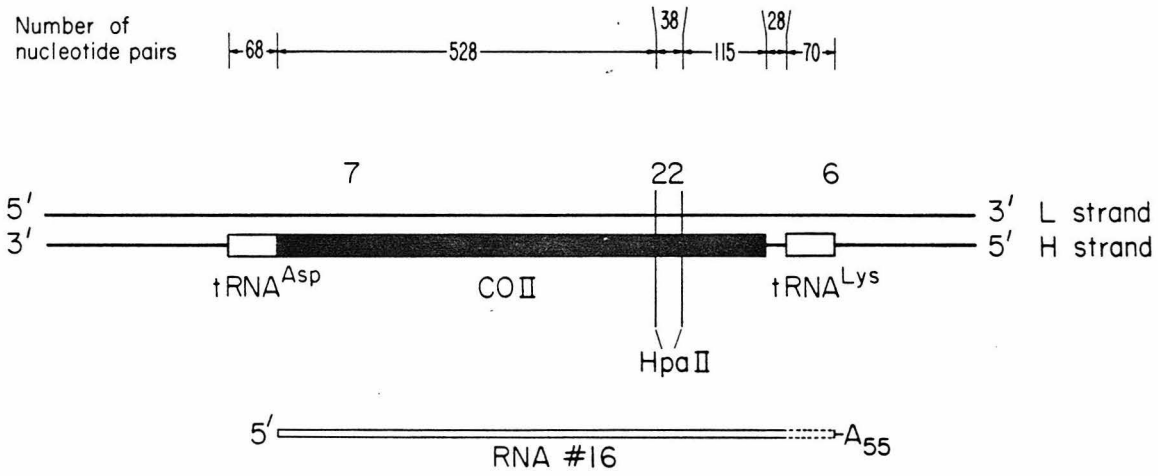

Number of
nucleotide pairs
68
528
38
115
28
70
7
22
6
5′
3′
L strand
3′
5′
H strand
tRNA^Asp
COII
tRNA^Lys
HpaII
5′
A55
RNA #16

Fig. 1

**Figure 2**

Isolation and 5'-end labeling of mitochondrial poly(A)-containing RNA 16. (a) The mitochondrial fraction from 5 x $10^9$ HeLa cells (~30 ml packed cells) was resuspended in 0.25 M sucrose, 0.01 M Tris-HCl, pH 8.0 (25°C), 0.001 M $CaCl_2$ (1 ml/1.5 x $10^8$ cells), and treated with 100 units/ml of micrococcal nuclease (Millipore Corp., Freehold, N.J.) at 2°C for 30 min. EGTA was then added to 2.5 mM and, after 5 min, the suspension was diluted with 0.25 M sucrose, 0.01 M TrisHCl, pH 6.7, 0.01 M EDTA (10 ml/1.5 x $10^8$ cells), and the mitochondria were pelleted by centrifugation at 10,000 rpm in a SS-34 rotor. The pellet was resuspended in 0.01 M Tris-HCl, pH 7.4, 0.15 M NaCl, 0.001 M EDTA (1 ml/1.5 x $10^8$ cells), incubated for 5 min at room temperature with 100 µg pronase/ml, then lysed with 1% SDS and incubated for an additional 30 min. RNA was extracted by phenol/chloroform/isoamyl alcohol (25:25:1) as previously described[7], recovered by ethanol precipitation, denatured by heating at 63°C for 4 min in $10^{-2}$ M Tris-HCl, pH 7.4, $10^{-3}$ M EDTA and passed through an oligo(dT)-cellulose column. The bound fraction was eluted, heat-denatured again as described above and rerun through the oligo(dT)-cellulose column. The final bound fraction was collected by ethanol precipitation and run through two tracks of a $CH_3HgOH$-agarose (1.4%) slab gel[7]. (b) The band in (a) which corresponds to poly(A)-containing RNA 16 was identified by ethidium bromide staining and eluted as previously described[10]. After ethanol precipitation, the RNA was dissolved in 250 µl 0.05 M Tris-HCl, pH 6.7, 0.0025 M $MgCl_2$, 0.025 M KCl, incubated with 20 µg/ml RNase-free DNase (Boehringer and Mannheim), pronase-SDS-phenol extracted and rerun through a $CH_3HgOH$-agarose slab gel. (c) The RNA 16 band in (b) was eluted, ethanol precipitated, dissolved in 100 µl of 0.01 M Tris-HCl, pH 8.0, and incubated at 65°C for 15 min with 100 units of bacterial alkaline phosphatase (B.R.L.) per µg

RNA (an amount of about 0.17 μg RNA 16 was roughly estimated to be recovered from ~5 x $10^9$ cells on the basis of its ethidium bromide staining, as compared to that of a known amount of 12S rRNA). Nitrilotriacetic acid was added to 10 mM and, after 15 min at room temperature, 2 μg of closed-circular pBR322 DNA were added as a carrier, and the nucleic acids were extracted with phenol/chloroform/isoamyl alcohol, ethanol-precipitated and dissolved in 15 μl kinase buffer (0.05 M Tris-HCl, pH 8.0, 0.01 M $MgCl_2$, 0.005 M DTT) containing 0.5 nmole of $\gamma$-$^{32}$P-ATP (8,000 Ci/mmole). After addition of 2 units of polynucleotide kinase (Boehringer and Mannheim), the reaction was carried out at 37°C for 15 min and then stopped by addition of 1.0 M ammonium acetate. Carrier tRNA was then added and the mixture ethanol precipitated and run on a $CH_3HgOH$-agarose (1.4%) slab gel.

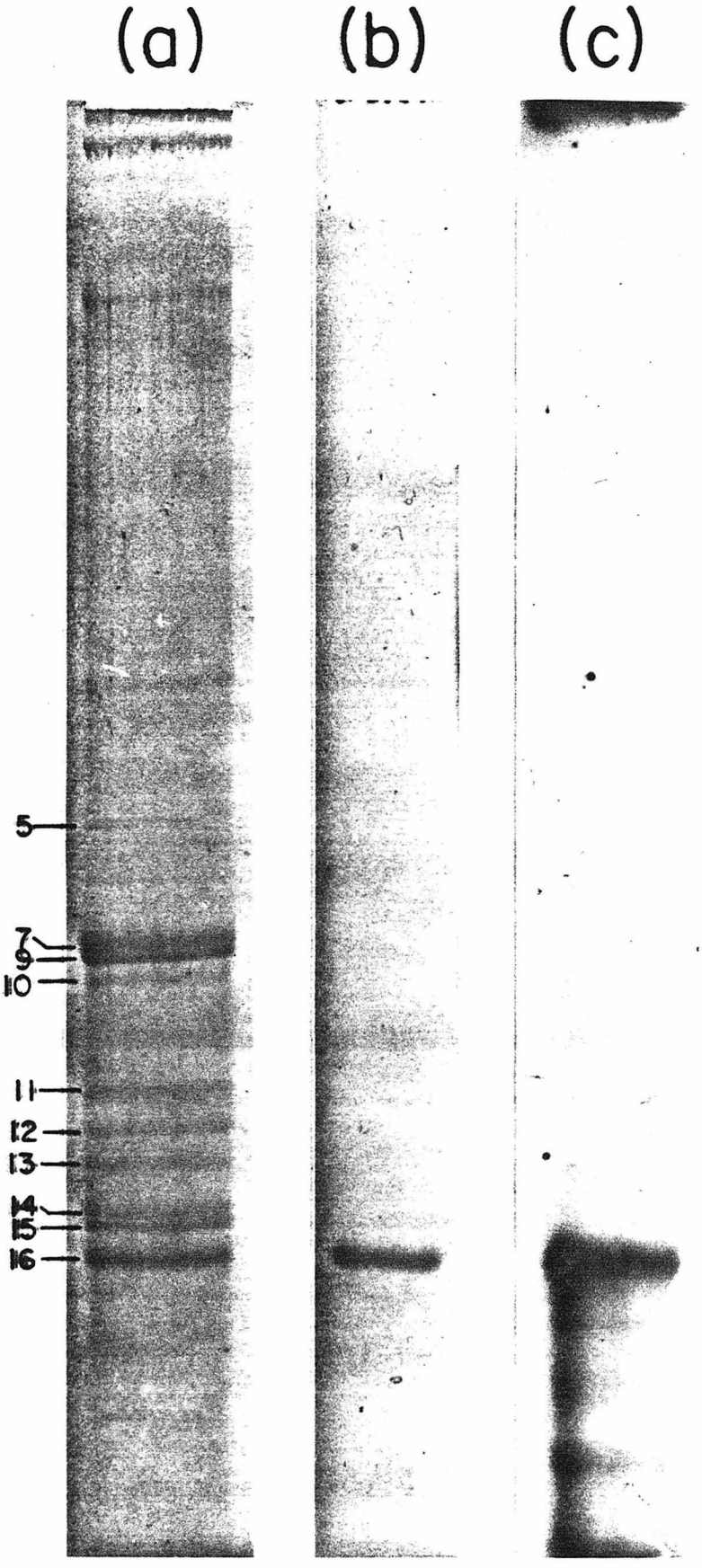

(a)
(b)
(c)
5
7
9
10
11
12
13
14
15
16

**Figure 3**

RNA sequencing of mitochondrial poly(A)-containing RNA 16. Autoradiograms of 25% sequencing gels (showing nucleotides 1-29) of partial enzymatic digests of 5'-end labeled RNA 16. Shown at the top of the gel are the cleavage specificities: RNase U2 (A), RNase T1 (G), RNase A (C+U), alkali ladder ($L_A$), formamide ladder ($L_F$). The reaction conditions described by Donis-Keller et al.[14] were used. The amounts of enzymes used (per ~30 ng RNA 16 and 3 μg yeast tRNA carrier) were, in the left panel: A 0.25 ng; T1 (1,2): 0.10, 0.05 units; U2 (1,2): 1.0, 0.5 units; in the right panel: A (1,2): 0.50, 0.75 ng; T1: 0.02 units; U2: 0.15, 0.5 units. The alkali and hot formamide degradations for the formation of the ladder were performed as described by Donis-Keller et al.[14] and by Simoncsits et al.[15], respectively. Electrophoresis was carried out on thin (0.5 mm) 25% polyacrylamide (30:1 bisacrylamide)/7 M urea gels in Tris-borate-EDTA buffer at 1,000 V for 6 hr.

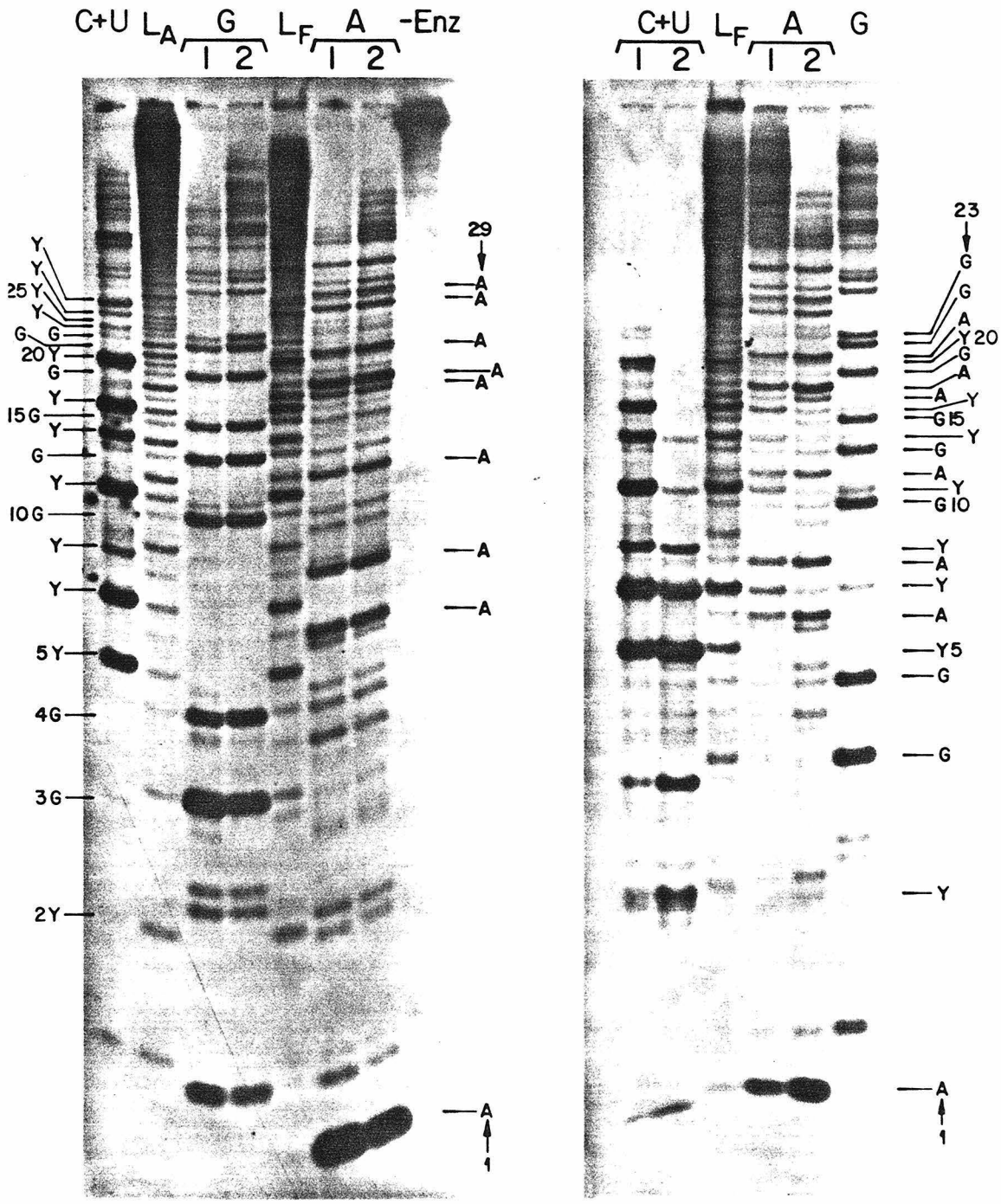
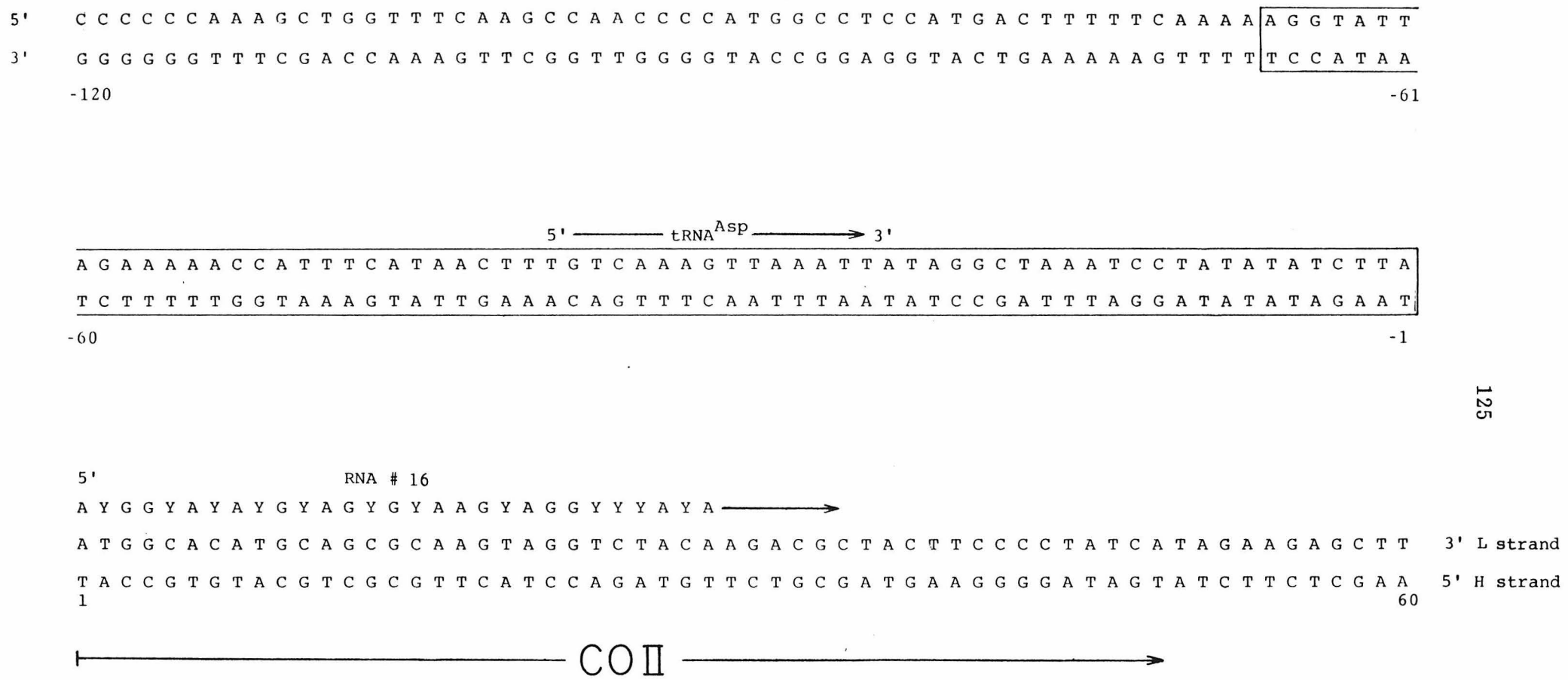

C+U
L_A
G
1 2
L_F
A
1 2
-Enz
29
25
20
15
10
5
4
3
2
1
C+U
L_F
A
1 2
G
23
20
15
10
5
1

**Figure 4**

Alignment of the nucleotide sequence of the 5'-end region of poly(A)-containing RNA 16 with the DNA sequence of the COII gene. The RNA sequence was derived from Fig. 3. The sequence of the 5'-end-proximal region of the COII gene and of the adjacent DNA segment on its 5'-side is reproduced from Barrell et al.[1]. Y: pyrimidine.


5' C C C C C C A A A G C T G G T T T C A A G C C A A C C C C A T G G C C T C C A T G A C T T T T T C A A A A A G G T A T T
3' G G G G G G T T T C G A C C A A A G T T C G G T T G G G G T A C C G G A G G T A C T G A A A A A G T T T T T C C A T A A
-120
-61
5' tRNA Asp 3'
A G A A A A A C C A T T T C A T A A C T T T G T C A A A G T T A A A T T A T A G G C T A A A T C C T A T A T A T C T T A
T C T T T T T G G T A A A G T A T T G A A A C A G T T T C A A T T T A A T A T C C G A T T T A G G A T A T A T A G A A T
-60
-1
125
5' RNA # 16
A Y G G Y A Y A Y G Y A G Y G Y A A G Y A G G Y Y Y A Y A
A T G G C A C A T G C A G C G C A A G T A G G T C T A C A A G A C G C T A C T T C C C C T A T C A T A G A A G A G C T T 3' L strand
T A C C G T G T A C G T C G C G T T C A T C C A G A T G T T C T G C G A T G A A G G G G A T A G T A T C T T C T C G A A 5' H strand
1
60
COII

# Chapter 6

## Distinctive Features of the 5'-terminal Sequences of the Human Mitochondrial mRNAs

**Abstract**

The 5'-end proximal sequences of all the putative mRNAs coded for by the heavy strand of HeLa cell mitochondrial DNA (mtDNA) have been determined and aligned with the DNA sequence. All these mRNAs start directly at, or very near to an AUG or AUA triplet, with the exception of one, which starts at an AUU. The available evidence indicates that the terminal or subterminal AUGs and AUAs, and possibly also the terminal AUU, are initiator codons for the corresponding polypeptides. In most cases, the individual mRNA coding sequences are flanked on their 5' side by a tRNA gene, without any intervening nucleotide.

## Introduction

A recently completed sequence analysis of human mitochondrial DNA (mtDNA) has revealed an extremely high degree of packing of genetic information in this DNA (ref. 1 and F. Sanger and B. Barrell, personal communication). In agreement with these observations, a detailed transcription mapping study of HeLa cell mtDNA by the S1 protection technique has shown that the sequences of the heavy (H) strand are almost completely saturated by the rRNAs, poly(A)-containing RNAs and tRNAs coded for by this strand[2,3]. A particularly intriguing feature of the human mitochondrial gene organization is the frequent juxtaposition, or close proximity, of the individual protein coding sequences, on their 5'-side, with a tRNA gene or another reading frame. Thus, the gene for subunit II of cytochrome c oxidase (COII) has been shown to be immediately contiguous to a $tRNA^{Asp}$ gene without any intervening nucleotide[4]. This unusual gene organization has raised the question of whether these putative mitochondrial mRNAs have, as all eucaryotic mRNAs so far analyzed[5,6], a 5' non-coding stretch, which may thus overlap the contiguous tRNA gene or reading frame, or whether they lack completely a leader sequence on their 5'-side. As a first approach to this question, the 5'-end proximal region of the putative COII mRNA has recently been sequenced and the sequence thus determined, aligned with the COII gene sequence[7]. The results clearly showed that the 5'-end of the COII mRNA corresponds precisely to the first nucleotide of the COII coding sequence. These experiments left open the question of whether the complete lack of 5' non-coding nucleotides is a general feature of human mitochondrial mRNAs, reflecting stringent constraints in the initiation of mitochondrial protein synthesis or specific rules of RNA processing, or whether, on the contrary, the position of the initiator codon relative to the 5'-end may vary in different mRNAs, possibly depending upon the position of the adjacent tRNA gene or reading frame. In the present work, we have obtained information bearing on these questions by sequencing a 5'-end proximal segment of all the

H-strand coded polysome-associated poly(A)-containing RNAs, which are presumably specific mRNAs, and by aligning the RNA sequences with the mtDNA sequence.

## Results

### Isolation and 5'-end labeling of mitochondrial mRNAs

In order to isolate HeLa cell mitochondrial poly(A)containing RNAs in adequate amounts for sequencing analysis, the micrococcal nuclease procedure for the purification of mitochondrial RNA[8] was scaled up as previously described[7]. Figure 1a shows the electrophoretic pattern in a 1.4% agarose-$CH_3HgOH$ slab gel, after ethidium bromide staining, of a large scale preparation of mitochondrial oligo(dT)-bound RNA from micrococcal nuclease-treated organelles (from ∿30 gm of cells). One clearly recognizes the characteristic set of components previously detected in [$^{32}$P] labeled preparations from HeLa cells[9], with the exception of RNA species 1, which is extremely short-lived and present in mitochondria in very small amounts[10,11]. Among these components, RNA species 5, 7, 9 and 11 to 17, because of their relative abundance, their enrichment in partially purified polysomal structures[9], and their relatively long half-life[10,11], are very probably specific mitochondrial mRNAs. Correlation of RNA sequence data, or of reading frames in mtDNA, with protein sequence data and with genetically characterized yeast gene sequences has provided strong support for the mRNA nature of several of these species[1,2]. However, no direct in vitro demonstration of the coding capacity of these RNAs is available as yet. Each of the putative mRNAs was eluted from the gel, treated with RNase-free DNase (to eliminate any contaminating DNA fragments deriving from the micrococcal nuclease treatment of the crude mitochondrial fraction), and rerun on a gel under the same conditions. As shown in Fig. 1b for RNA 13, a sharp band was observed at the expected position, and the RNA in this band was eluted. Good resolution of RNAs 7 and 9,

and 14 and 15 was obtained in a similar rerun on a 2% agarose-$CH_3HgOH$ gel (Fig. 1d and e). Material purified, as described above, from 40 to 250 gm of cells, depending on the RNA species, was finally labeled at its 5'-end with $\gamma$-$^{32}$P-ATP and polynucleotide kinase after dephosphorylation with bacterial alkaline phosphatase, rerun on a $CH_3HgOH$-agarose gel (Fig. 1c) and eluted for the sequencing reactions.

**RNA sequence analysis**

The isolated in vitro labeled RNA species were first characterized in their 5' termini. As shown in Table 1, the distribution of radioactivity among the 5'-mononucleotides indicated the adequate degree of purity of all the RNA species, with the exception of RNAs 5 and 7. RNA 5, which was the least abundant RNA species analyzed, had a little more than half of its radioactivity associated with pA, while RNA 7 seemed to consist of two components approximately equal in amount, one terminating in pA and the other in pU.

Sequencing of the 5'-end proximal segments of the poly(A)-containing RNAs was performed by the methods of Donis-Keller et al.[13] and Simoncsits et al.[14], utilizing cleavage of the RNAs by RNases $T_1$ (G-specific), A (C+U-specific) and $U_2$ (A-specific), followed by electrophoresis on polyacrylamide/urea gels[8]. The essential features of these RNA sequencing methods, which had been previously reported by others[13,14] and ourselves[7,8], were also observed here (Figs. 2, 3 and 4). RNases $T_1$ and $U_2$ were found to cleave regularly and exclusively after Gs and As, although with varying efficiency probably depending on the surrounding sequence. RNase A cleaved exclusively after pyrimidines, but, as previously described, failed to cleave after some pyrimidines, especially when these were part of pyrimidine stretches. Accordingly, the absence of a band in any given position, after digestion with RNase A or $U_2$ or $T_1$, was tentatively interpreted to reflect the presence at that position of a pyrimidine, if the data indicated that this putative pyrimidine

was part of a stretch of pyrimidines (two or more); typically, in such a situation, the slowest moving among the pyrimidine-terminated oligonucleotides corresponding to the pyrimidine stretch was always clearly visible, with one or more of the other products being sometimes visible as fainter bands.

In most cases, a single band was observed at any given ladder position in the gel track, in agreement with the idea that a predominant RNA species was being analyzed. In some cases, however, a definitely weaker band was observed at the same position as the predominant band, and disregarded as a contaminant. (In many such instances, such weaker bands were also present in the untreated RNA control.) In a few cases, no band was seen in any of the lanes at a position corresponding to a ladder step, or two bands of comparable intensity were observed at the same position; in these cases, no conclusive identification was made.

A complication in the interpretation of the results was introduced by the presence of extra bands in the gel pattern, especially in the region corresponding to the smaller oligonucleotides. These extra bands were in general less strong than the bona fide specific RNase cleavage products. Because of this and because they did not correspond to bands visible in the ladder or corresponded to obviously abnormal steps of the ladder, these extra bands could in general be recognized as spurious. The nature of these extra bands is not clear. They may be due to the presence in the RNA preparation, during the 5' end-labeling reaction, of contaminating RNA species or DNA fragments (deriving from DNA degradation during the DNase treatment step), or they could reflect the formation, during the enzymatic or chemical cleavage of the RNA, of oligonucleotides with different phosphate end-groups (2',3'-cyclic phosphate, 2' and 3' phosphate). The latter explanation seems the most likely to account for the fact that, in some sequencing analyses (RNAs 5, 7 and 17), both the specific RNase cleavage products and the chemical cleavage products gave doublets of bands instead of single bands.

In each experiment, an untreated RNA sample was run in parallel with the RNase or chemically treated samples. This control ("-Enz") in most cases did not show bands or exhibited bands only in the region of the larger oligonucleotides.

The interpretation of the sequence of RNA 7 (Fig. 4) offered some difficulties because of the repetition of each band at the next position in the ladder. A similar observation had been previously made for 12S rRNA[8] and interpreted to reflect the existence of two species of 12S rRNA differing in molecular size by one nucleotide at the 5'-end. In the case of RNA 7, in agreement with the $P_1$ digestion data (Table 1), analysis of the sequencing gel showed that it could be interpreted as resulting from the electrophoretic separation of the products of RNase digestion of two species, one starting with AUG and the other UG; a comparison with the DNA sequence (see below) confirmed the validity of this interpretation. RNA 5 is present in HeLa cells in extremely minute amounts, and only a partial purification could be obtained even after two electrophoretic runs through $CH_3HgOH$-agarose slab gels, as revealed by the $P_1$ digestion data (Table 1). Therefore, only partial sequence data could be obtained for this RNA species. However, it was possible to identify with reasonable confidence the first two nucleotides (AY), the sixth (Y) and the ninth (G) (not shown).

**Most mRNAs start at or very near an AUG or AUA triplet**

Using the criteria described above, a stretch of 10 to 54 nucleotides, depending upon the RNA species, could be read in the sequence of all the poly(A)-containing RNAs analyzed here (except RNA 5). The sequences thus obtained are shown in Fig. 5; for RNA species 9, 12, 13, and 16, the corresponding DNA sequences which have been published or are in press (refs. 4 and 15) are also reported to show their perfect alignment with the RNA sequences. The remaining RNA species 5, 7, 11, 14, 15, and 17 have also been compared with DNA sequence data (F. Sanger and

B. Barrell, personal communication) and, likewise, were found to be in agreement. The comparison of the RNA and DNA sequences clearly indicates that all the AYG and AYA triplets underlined in the RNA sequences are indeed AUG and AUA codons. The partial sequence data for RNA 5 are consistent with the first triplet being an AUA. Thus, the striking results of this analysis are that, with one possible exception (RNA 12), all the putative human mitochondrial mRNAs start directly with an AUG or AUA triplet (which is a methionine codon in human mitochondria[1]) or have a few nucleotides (1 to 8) preceding the AUG or AUA. Sequencing data of human[16] and bovine cytochrome c oxidase subunits[1,4] and comparison with the yeast cytochrome b gene sequence (A. Tzagoloff, personal communication) have indicated that the first AUGs of RNA 9 (COI mRNA), RNA 16 (COII mRNA), RNA 15 (COIII mRNA) and RNA 11 (cyt. b mRNA) are initiator codons for the corresponding polypeptides. Thus, it seems reasonable to extrapolate from these results and interpret the 5' proximal AUGs of the other mRNAs likewise as initiator codons. The mtDNA sequencing data suggest that AUA may function as an initiator codon in human mitochondria[1], and the data presented here do support this possibility. If the first AUA of RNA 13 is indeed the initiator codon, it would mean that, with the exception of RNA 12, all putative mitochondrial mRNAs have the initiator codon within the first six nucleotides. There is also the possibility that RNA 12 too may follow this rule. In fact, in bovine mtDNA, the first triplet of the reading frame corresponding to RNA 12 is AUG, instead of AUU as in human mtDNA; this observation has suggested the possibility that AUU may function as initiator codon for the polypeptide coded for by RNA 12, being read either by the $tRNA^{Ile}$ or the $tRNA^{F\text{-}Met}$[1].

## Discussion

### How do ribosomes attach to mitochondrial mRNAs?

The finding that most mitochondrial mRNAs either start directly at the

initiator codon or have only a few nucleotides (1 to 8) preceding this codon poses interesting questions concerning the mechanism whereby mitochondrial ribosomes attach to these messengers. Both in eukaryotic and prokaryotic mRNAs, there is a stretch of variable length preceding the initiator codon[5,6]. In prokaryotic mRNAs there is good evidence that this stretch contains a ribosome-binding site, which includes a properly positioned sequence complementary to the 3'-end of 16S rRNA[17,18]. A single exception to this rule has been described, namely that of the lambda repressor mRNA involved in repressor maintenance, which starts directly at the translation initiator codon[20]. In this case, the low efficiency of translation of the mRNA was attributed to the lack of a strong ribosome binding site. The observation that, in the 5' non-coding region of some eukaryotic mRNA, there is a considerable degree of complementarity with a purine-rich sequence near the 3'-end of 18S rRNA[5] has suggested that specific mRNA-rRNA base pairing may also help the positioning of eukaryotic ribosomes on the mRNAs; however, the variability in the position of the pyrimidine-rich sequence within the mRNA molecules and even the lack of these sequences in some mRNAs has raised questions about the generality of this mechanism[6].

A common feature which has emerged from the sequencing analysis of a variety of eukaryotic (cellular and viral) mRNAs is that in most, if not all these mRNAs, translation starts at the AUG closest to the 5'-end, apparently independently of the length of the 5' non-coding region (varying in the known cases between 11 and over 200 nucleotides) and of its nucleotide sequence[6,19]. No minimum distance between initiator codon and 5'-end of the mRNA, which would allow proper initiation, has yet been defined for cellular or viral mRNAs; however, in the case of Sindbis virus RNA, there is evidence suggesting that the AUG immediately adjacent to the cap[20] is not used for initiation[6]; this would imply that cytoplasmic ribosomes are unable to recognize an initiator codon in such a location. To explain why initiation

of translation is limited to the most 5'-proximal AUG, a "scanning" mechanism has been proposed[6], whereby a 40S ribosomal subunit would attach initially to the 5'-end of an RNA chain and subsequently migrate toward the interior of the RNA, stopping when the first AUG is encountered; at this point the 60S subunit would join and translation start. The cap structure, in this model, would enhance the binding of the 40S subunits to the 5'-end of the mRNA, but would not be required. It is interesting that HeLa cell and in general, presumably, mammalian mitochondrial mRNAs are not capped[21].

The terminal or subterminal position of the initiator codon in human mitochondrial mRNAs may eliminate the need for a scanning process, at least an extensive one. A plausible mechanism would involve attachment of the ribosome at or near the 5'-end of the mRNA with recognition of the initiator codon either directly or after a fine adjustment. It is conceivable that the secondary structure of mitochondrial mRNA may be such that it would exclude all internal sites containing AUG or AUA codons, exposing only the terminal or subterminal one. It is also conceivable that the special features of mammalian mitochondrial ribosomes or some initiation factor would make the ribosomes suitable for recognizing the terminal or subterminal initiator codons. Analysis of the potential secondary structures of human mitochondrial mRNAs and binding studies with mitochondrial or other ribosomes should provide in the future some clues as to mechanisms operating in the initiation of translation in human mitochondria.

**Most mRNA coding sequences are immediately adjacent at their 5'-end to tRNA genes**

The complete absence or minimal length of the 5' non-coding stretch in human mitochondrial mRNAs may represent the vestige of a primitive organization, or may reflect, with other traits (for example, the relatively small size of the rRNA

and tRNA species), an evolutionary trend to simplicity and economy of the mitochondrial genome[22]. In any case, the degree of proximity of the initiator codon to the 5'-end of the mitochondrial mRNAs seems to be dictated primarily by the features of gene organization of this genome, in particular, by the positions of the tRNA genes. In fact, whatever the location of the initiator codon is relative to the 5'-end of the mitochondrial mRNAs, it appears from the data presented here (Fig. 5) that in every case where a tRNA gene exists on the 5' side of the mRNA coding sequence, the latter always starts immediately after the tRNA gene. This observation confirms and refines the results of a recent detailed transcription analysis of HeLa cell mtDNA[3]. These mapping and sequencing data strongly support the idea that the tRNA sequences represent the recognition signals for a putative processing enzyme which, by precise endonucleolytic cleavages, would release the 5'-ends of the mature mRNAs[2,3]. As will be discussed elsewhere[23], recent experimental evidence indicates that the tRNA sequences may perform a similar function in the processing steps which release the 3'-ends of the human mitochondrial mRNAs. These striking features of the organization of the human mitochondrial genome underlie the crucial role that the tRNA sequences probably play in mitochondrial RNA processing in human cells[2,3].

**Acknowledgements**

This work was accomplished through the combined efforts of Dr. Julio Montoya and myself. In addition, we are very grateful to Drs. F. Sanger and B. Barrell for communicating to us the human mtDNA sequence prior to publication.

## References


1. Walker, J. E., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Runswick, M. J., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. in The Organization and Expression of the Mitocondrial Genome (eds Kroon, A. M. & Saccone, C.) (North-Holland, Amsterdam, in the press).
2. Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C., Montoya, J. & Ojala, D. in The Organization and Expression of the Mitochondrial Genome (eds Kroon, A. M. & Saccone, C.) (North-Holland, Amsterdam, in the press).
3. Ojala, D., Merkel, C., Gelfand, R. & Attardi, G. Cell (in the press).
4. Barrell, B. G., Bankier, A. T. & Drouin, J. Nature **282**, 189-194 (1979).
5. Hagenbüchle, O., Santer, M., Steitz, J. A. & Mans, R. J. Cell **13**, 551-563 (1978).
6. Kozak, M. Cell **15**, 1109-1123 (1978).
7. Ojala, D., Montoya, J. & Attardi, G. Nature (in the press).
8. Crews, S. & Attardi, G. Cell **19**, 775-784 (1980).
9. Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. J. molec. Biol. **118**, 1-25 (1978).
10. Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C. & Ojala, D. in Proc. ICN-UCLA Symp. on molec. cell Biol. **15**, 443-469 (1979).
11. Gelfand, R. Ph.D. Thesis, California Institute of Technology, Pasadena, California (1980).
12. Fujimoto, M., Kuninaka, A. & Yoshino, H. Agr. Biol. Chem. **38**, 1555-1561 (1974).
13. Donis-Keller, H., Maxam, A. M. & Gilbert, W. Nucleic Acids Res. **4**, 2527-2538 (1977).

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

15. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. J. molec. Biol. (in the press).

16. Chomyn, A., Hunkapiller, M. & Attardi, G. (in preparation).

17. Shine, J. & Dalgarno, L. Biochem. J. **141**, 609-615 (1974).

18. Steitz, J. A. & Jakes, K. Proc. natn. Acad. Sci. U.S.A. **72**, 4734-4738 (1975).

19. Sherman, F., Stewart, J. W. & Schweingruber, A. M. Cell **20**, 215-222 (1980).

20. Hefti, E., Bishop, D. H. L., Dubin, D. T. & Stollar, V. J. Virol. **17**, 149-159 (1976).

21. Grohmann, K., Amalric, F., Crews, S. & Attardi, G. Nucleic Acids Res. **5**, 637-651 (1978).

22. Attardi, G. Trends in Biochem. Sci. (in the press).

23. Ojala, D., Montoya, J. & Attardi, G. (unpublished results).

**Table 1** 5'-Terminal nucleotide analysis of in vitro labeled mitochondrial poly(A)-containing RNA species

| Poly(A)-containing RNA species | Percentage of total $^{32}P$ radioactivity* | | | |
|---|---|---|---|---|
| | AMP | CMP | GMP | UMP |
| 5 | 52.5 | 8.1 | 27.6 | 11.7 |
| 7 | 43.8 | 2.1 | 7.8 | 47.1 |
| 9 | 12.4 | 80.9 | 2.0 | 4.5 |
| 11 | 81.1 | 3.0 | 10.2 | 5.5 |
| 12 | 78.9 | 2.3 | 12.7 | 5.9 |
| 13 | 81.4 | 2.4 | 10.0 | 6.0 |
| 14 | 91.0 | 3.1 | 2.2 | 3.5 |
| 15 | 95.9 | 0.3 | 1.6 | 2.1 |
| 17 | 73.7 | 2.5 | 6.4 | 17.2 |

The 5'-terminal nucleotide was determined by exhaustive digestion of the in vitro labeled RNAs with nuclease P1[12] (Calbiochem) and fractionation of the products on PEI-cellulose TLC plates, as previously described[8].

*The values refer to the radioactivity which migrated from the origin. The % of the input radioactivity recovered from the origin varied between 0.1 and 5%.

**Figure 1**

Isolation and 5'-end labeling of mitochondrial poly(A)-containing RNA species. Lane a, Electrophoretic pattern, after ethidium bromide staining, of the oligo(dT)-bound mitochondrial RNA fraction from ∿30 gm of HeLa cells run through a 1.4% agarose-$CH_3HgOH$ slab gel. Lanes b and c, RNA species 13 was eluted from several gel tracks as that shown in lane a, treated with RNase-free DNase, and rerun on a gel under the same conditions (b), then eluted again, labeled at its 5'-end with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase after bacterial alkaline phosphatase treatment, and rerun (c). Lanes d and e, Rerun of RNA species 14 plus 15 (d) or 7 plus 9 (e) on 2% agarose-$CH_3HgOH$ gels after elution from the gel track shown in lane a and treatment with DNase. Reference is made to previous work for the large-scale preparation of mitochondrial RNA by the micrococcal nuclease procedure[7,8], subsequent isolation and fractionation by 1.4% agarose-$CH_3HgOH$ slab gel electrophoresis of the oligo(dT)-cellulose bound RNA, elution, DNase treatment and rerun on agarose-$CH_3HgOH$ slab gels of the individual components[7], and 5'-end labeling of the purified RNA species[8].

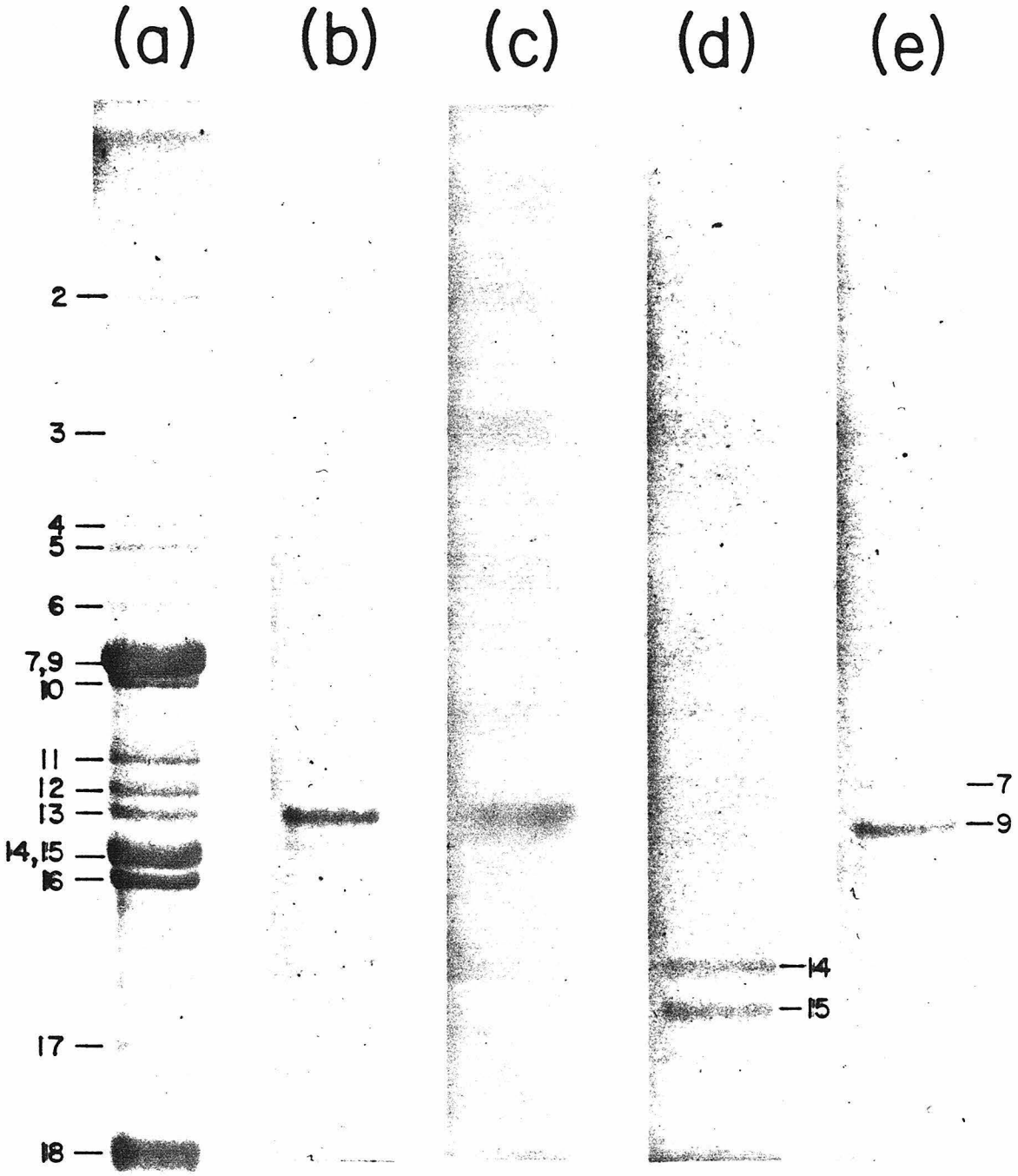

(a)
(b)
(c)
(d)
(e)
2
3
4
5
6
7,9
10
11
12
13
14,15
16
17
18
7
9
14
15

**Figure 2**

Sequencing gels of mitochondrial poly(A)-containing RNA species 13, 12 and 9. Lanes 1-6, RNA 13, 25% polyacrylamide/urea gel. Shown at the top of the gel are the cleavage specificities: RNase A (C + U), alkaline ladder ($L_A$), RNase $U_2$ (A), RNase T1 (G), and RNA incubated without enzyme (-Enz). Y indicates a pyrimidine band. Lanes 7-13, RNA 12, digested and analyzed as indicated for RNA 13. The two left lanes show the digests obtained with decreasing amounts of RNase A. $L_F$ indicates a formamide ladder. The horizontal ticks without a letter designation indicate ladder positions for which no nucleotide assignment has been made. Lanes 14-24, RNA 9, digested as indicated for RNA 13 and analyzed on a 25% (lanes 14-20) or, for the sequencing of residues from ∿20 to 60, a 10% (lanes 21-24) polyacrylamide/urea gel. Reference is made to previous work for the details of the base specific enzymatic cleavage reactions[7,13,14] and for the conditions of electrophoretic analysis[7,8]. In some experiments, in order to visualize a pyrimidine band in the first or second position of the sequence, the amount of RNase A used (per ∿30 ng RNA and 3 μg yeast tRNA carrier) was increased to 20 ng.

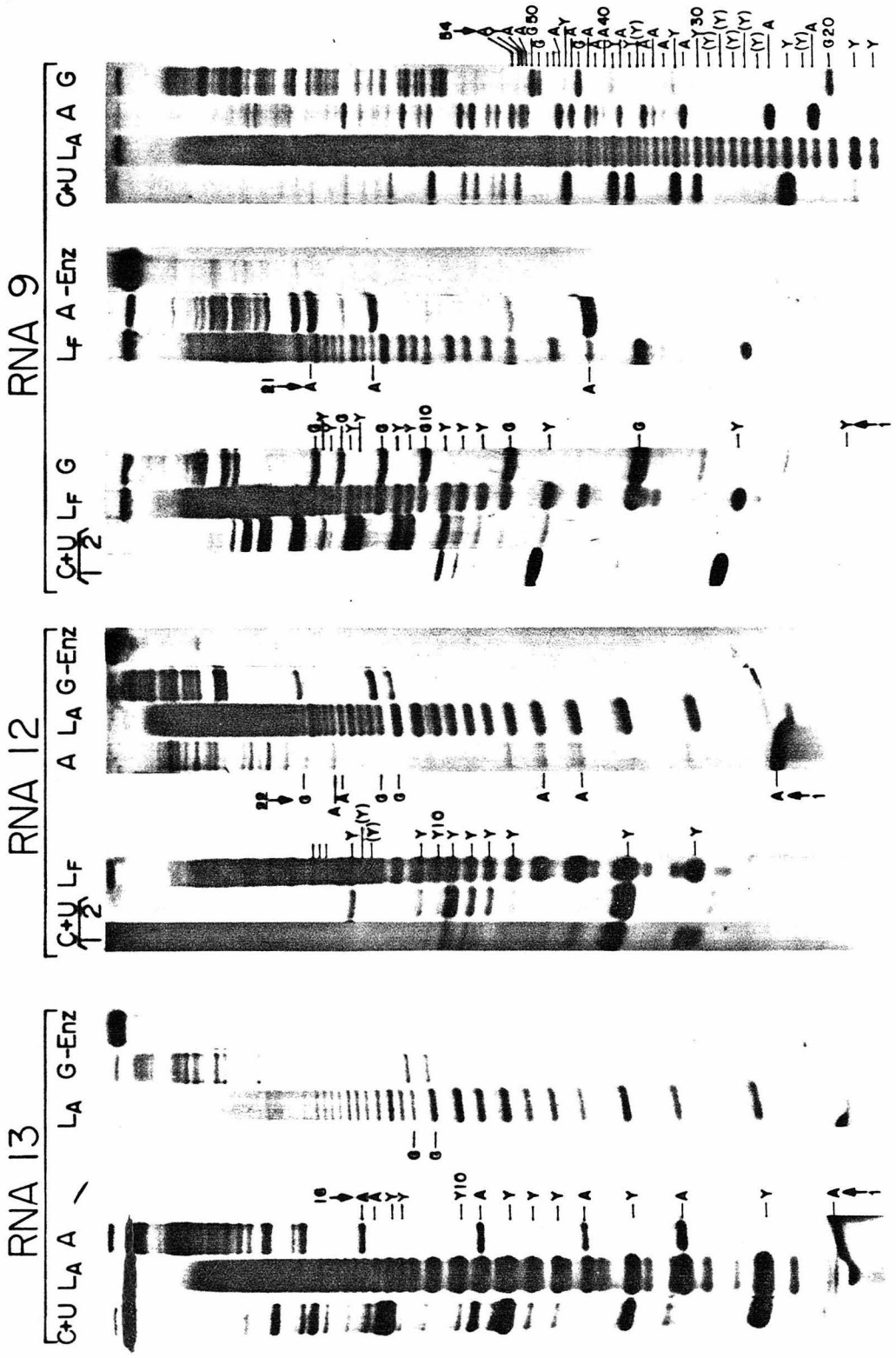

RNA 9
C+U LA A G
LF A-Enz
C+U LF G
1 2
RNA 12
A LA G-Enz
C+U LF
1 2
RNA 13
LA G-Enz
C+U LA A

**Figure 3**

Sequencing gels of mitochondrial poly(A)-containing RNA species 14, 15 and 17. Lanes 1-6, RNA 14, 25% polyacrylamide/urea gel. Lanes 7-16, RNA 15, digested as shown for RNA 14, and analyzed on a 25% (lanes 7-11) or a 10% (lanes 12-16) polyacrylamide/urea gel. Lanes 17-23, RNA 17, digested and analyzed as described for RNA 14.

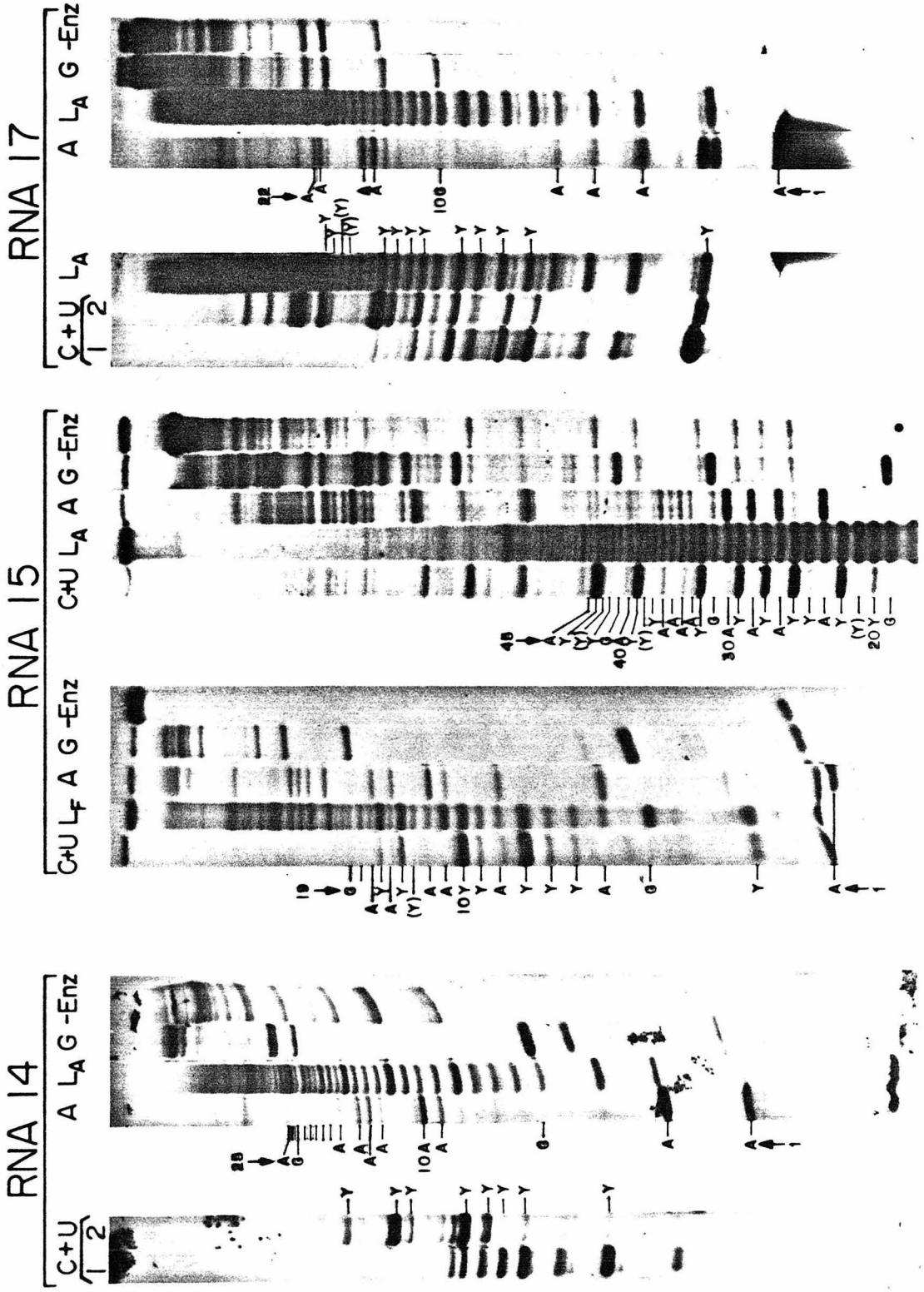

RNA 17
A L_A G -Enz
C+U L_A
1 2
RNA 15
C+U L_A A G -Enz
C+U L_F A G -Enz
RNA 14
A L_A G -Enz
C+U
1 2

**Figure 4**

Sequencing gels of mitochondrial poly(A)-containing RNA species 7 and 11. Lanes 1-6, RNA 7, 25% polyacrylamide/urea gel. The capital and lower case letters indicate, respectively, the nucleotide assignments of the presumptively longer RNA species and of a species differing from this one by the lack of the 5'-terminal nucleotide at the time of labeling. Lanes 7-18, RNA 11, digested as indicated for RNA 7 and analyzed on a 25% (lanes 11-13) or a 10% (lanes 14-18) polyacrylamide/urea gel.

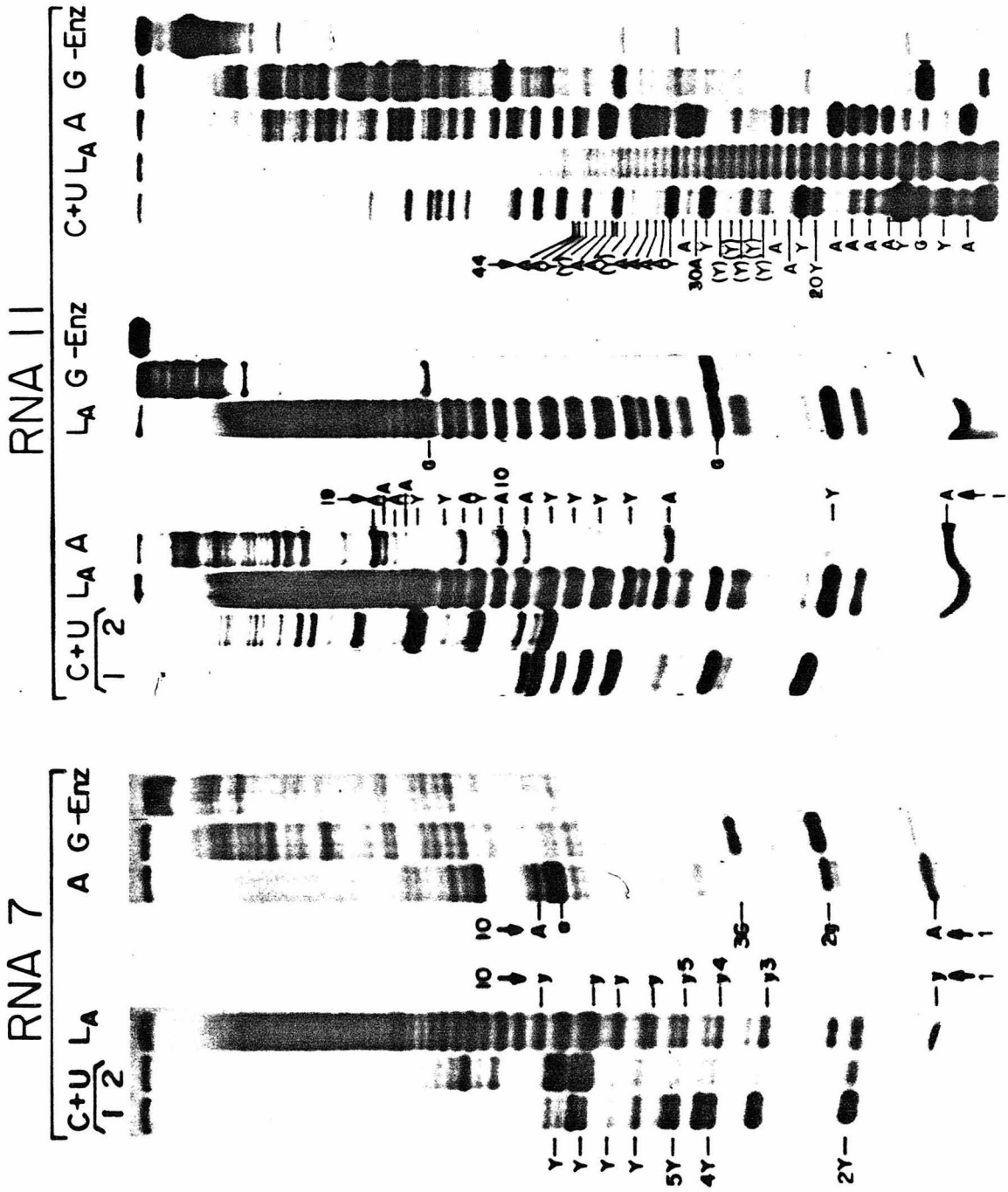

RNA 11
C+U
1 2
LA A
LA G-Enz
C+U LA A G-Enz
RNA 7
C+U
1 2
LA
A G-Enz

**Figure 5**

The sequences of the 5'-end proximal regions of the mitochondrial poly(A)-containing RNAs determined in the present work are shown, aligned with the corresponding published mtDNA sequences. Y indicates a pyrimidine (the brackets denote a tentative identification, as explained in the text), a dash (-), an unidentified base; at position 38 in the RNA 11 sequence two possible assignments are indicated.


RNA 5
5' A Y - - - Y - - G.....
1
RNA 7
5' A Y G Y Y Y Y Y Y A.....
1 10
RNA 9
5' Y Y G A Y G Y Y Y G Y Y G A Y Y G Y Y G A(Y)Y A(Y)(Y)(Y)(Y)(Y)(Y)Y A Y A A A(Y)Y A Y A A A G A Y A - - G G A A Y A.....
1 10 20 30 40
C T G A T G T T C G C C G A C C G T T G A C T A T T C T C T A C A A A C C A C A A A G A C A T T G G A A C A.....
1 10 20 30 40
RNA 11
5' A Y G A Y Y Y Y A A Y A Y G Y A A A A Y Y A A(Y)(Y)(Y)(Y)(Y)(Y)Y A A Y A A A A(Y)Y A A(Y)Y A A.....
1 10 20 30 G 40
RNA 12
5' A Y Y A A Y Y Y Y Y Y G G(Y)(Y)Y A A.....
1 10
tRNA F-Met
C C C G T A C T A A T T A A T C C C C T G G C C C A A.....
1 10
RNA 13
5' A Y A Y A Y Y Y A Y G G Y Y A A .....
1 10
tRNA Leu
T T C T T A A C A A C A T A C C C A T G G C C A A.....
1 10
RNA 14
5' A A Y G Y Y Y Y A A Y Y A A A Y A.....
1 10
RNA 15
5' A Y G A Y Y Y A Y Y A A(Y)Y A Y A - G Y(Y)Y A Y Y A Y A Y A G Y A A A A Y(Y)Y A G Y(Y)Y A.....
1 10 20 30 40
RNA 16
5' A Y G G Y A Y A Y G Y A G Y G Y A A G Y A G G Y Y Y A Y A.....
1 10 20
tRNA Asp
T A T A T C T T A A T G G C A C A T G C A G C G C A A G T A G G T C T A C A.....
1 10 20
RNA 17
5' A Y A A A Y Y Y Y G Y Y Y Y A A(Y)(Y)Y Y A A.....
1 10 20

# Chapter 7

## Transcription of the HeLa Cell Mitochondrial Genome: The Punctuation Model

**Abstract**

The partial nucleotide sequence of the 3'-end of eight poly(A)-containing RNA species (presumed to be messenger RNAs) encoded by HeLa cell mitochondrial DNA has been determined and aligned with its corresponding DNA sequence. The 3'-terminal nucleotide of all of these RNAs has been found to be juxtaposed to the 5'-terminal nucleotide of a flanking tRNA or in one case, juxtaposed to the 5'-terminal nucleotide of a flanking poly(A)-containing RNA. These results offer further support for a mode of transcription of the HeLa cell mitochondrial genome which has been termed the "punctuation model".

## Introduction

As discussed previously in this thesis, the sequence analysis of human mitochondrial DNA has revealed a high degree of packing of genetic information encoded within this genome (F. Sanger and B. Barrell, personal communication). In addition (as shown in Figure 1), a detailed transcription map of HeLa cell mitochondrial DNA has indicated that the H-strand sequences coding for the rRNA species, the poly(A)-containing RNAs and the tRNAs appear to be contiguous to each other, extending from coordinate 2/100 to coordinate 95/100 (relative to the origin taken as 0/100) (Ojala et al., 1980b; Chapter 3 of this thesis). More recently, the development of a procedure for the isolation of the mitochondrial RNA transcripts in a pure form and relatively large amounts (Crews and Attardi, 1980; Ojala et al., 1980a) has allowed us to extend the mapping studies to the nucleotide level by carrying out sequencing analyses of the 5'-end of the majority of the encoded transcripts. Thus the 5'-end nucleotide sequence of RNA species 13, 12, 9, 16, 14, 15, 17, 7, 5 and 11 has been determined and aligned with the corresponding DNA sequence (Montoya, Ojala and Attardi, manuscript in preparation; Chapter 6 of this thesis). These results have revealed the fact that the 5'-terminal nucleotide of seven out of ten of these species is immediately juxtaposed to the 3'-terminal nucleotide of a flanking tRNA. This extends previous observations made for the 12S and 16S rRNA genes in which the 5'-terminal nucleotide of these cistrons is also juxtaposed to a flanking tRNA gene (Crews and Attardi, 1980; Eperon et al., 1980). In addition, these studies support the idea that the tRNA sequences may function as recognition signals for a putative processing enzyme which, by precise endonucleolytic cleavages, would release the 5'-ends of the mature mRNAs.

To determine if the 3'-ends of these RNAs are similarly juxtaposed to flanking tRNAs, a rapid and sensitive methodology has been developed which allows the alignment

of a partially determined nucleotide sequence with that of known DNA sequence data. Results thus obtained, and presented in this chapter, for a number of these species offers further support for the above described mode of transcription of the HeLa cell mitochondrial genome. This mode of transcription has been termed "the punctuation model."

## Materials and Methods

### Materials

High specific activity [$\gamma$-$^{32}$P]ATP (8000 Ci/mmole) was purchased from ICN, oligo(dT) primers [$p(dT)_8$-dA, $p(dT)_8$-dC and $p(dT)_8$-dG] from Collaborative Research, Inc., bacterial alkaline phosphatase (BAP) from Bethesda Research Laboratories, Inc., and polynucleotide kinase from Boehringer Mannheim. Reverse transcriptase isolated from BAI strain A avian myeloblastosis virus was a gift from Dr. John Beard.

### Methods

#### Isolation and Purification of Mitochondrial Poly(A)-Containing RNA Species

The large scale preparation by the micrococcal nuclease procedure of mitochondrial RNA from HeLa cells, subsequent isolation and fractionation by 1.4% agarose-$CH_3HgOH$ slab gel electrophoresis of the oligo(dT)-bound fraction, elution, DNase treatment (to eliminate contaminating DNA fragments deriving from the micrococcal nuclease treatment of the crude mitochondrial fraction) and rerun on agarose-$CH_3HgOH$ slab gels of the individual components have previously been described in detail (Ojala et al., 1980; Chapter 5 of this thesis).

#### 5'-End Labeling and Gel Purification of the Oligonucleotide Primers

5 µg ($\sim$1.6 nmoles) each of $p(dT)_8$-dA, $p(dT)_8$-dC and $p(dT)_8$-dG, resuspended in 50 µl 0.01 M Tris, pH 8.0 (at 25°C), were dephosphorylated by incubation with

bacterial alkaline phosphatase (35 units BAP per nmole primer) for 30 min at 37°C. Reaction volumes were adjusted to 100 µl with $H_2O$ and the mixtures extracted twice with an equal volume of phenol-chloroform-isoamylic alcohol (25:25:1). The phenol phases were reextracted with 100 µl $H_2O$, the aqueous phases combined and extracted twice with ethyl ether (to remove remaining phenol). The mixtures were lyophilized to dryness and each resuspended in 10 µl of kinase buffer (0.05 M Tris, pH 8.0, 0.01 M $MgCl_2$, 0.005 M dithiothreitol) containing [$\gamma$-$^{32}$P]ATP (2 mCi, 0.25 nmole). Final ATP concentration was 25 µM. Two units of polynucleotide kinase were added and the mixtures incubated 30 min at 37°C. The reactions were stopped by addition of 10 µl of 10 M urea containing 0.5% each of bromophenol blue and xylene cyanol FF dyes. The samples were electrophoresed on a 15% polyacrylamide (20:1 bis-acrylamide) gel in Tris-borate-EDTA buffer (Maxam and Gilbert, 1977). The gel was subjected to autoradiography and the bands corresponding to the intact primers excised and the oligonucleotides eluted by two serial washes of 0.5 ml $H_2O$. The washes were combined, filtered through glass wool (to eliminate gel particles), lyophilized to dryness, and resuspended in 100 µl $H_2O$. Final specific activity of the labeled primers was 5-10 x $10^7$ counts/µg.

**Determination of the Nucleotide Adjacent to the Poly(A) Tail**

Each polyadenylated RNA species (∿0.05 to 0.10 pmoles) was mixed (at 4°C) with a 100-fold molar excess of each of the three 5'-end labeled primers and 50 µM of each of the four deoxyribonucleoside triphosphates. Buffer conditions were adjusted to 50 mM Tris, pH 8.0, 50 mM KCl, 5 mM $MgCl_2$ and 10 mM dithiothreitol. Four units of reverse transcriptase were added and the mixtures incubated for 3 min at 39°C. In some experiments, the reaction mixtures were heated for 5 min at 75°C and then cooled for 15 min at 4°C prior to addition of the enzyme, in order to reduce possible secondary structure formation of the template RNA and thus facilitate

hybridization of the primer with the poly(A) tails. No differences in the results were observed, however, by this modification. cDNA synthesis was terminated by placing the mixtures in ice for ∿15 sec, and then adding EDTA to 0.01 M, followed by addition of NaCl to 0.3 M, 2 volumes of ethanol and 5 μg carrier yeast tRNA. [Care was taken to keep the mixtures at a temperature of 2°C or less to insure that the shorter complementary transcripts would not melt from the template and thereby be lost during the ethanol precipitation steps.] The products were collected by centrifugation, washed with cold ethanol (to remove excess labeled primer), recentrifuged, dried and resuspended in 10 μl of $10^{-3}$ M Tris-EDTA, followed by 10 μl of 10 M urea containing 0.5% each of bromophenol blue and xylene cyanol FF dyes, heated 2 min at 90°C, then layered on a thin (0.5 mm) 10% polyacrylamide/7 M urea sequencing gel in Tris-borate-EDTA buffer (Maxam and Gilbert, 1977). Electrophoresis was carried out for 3.5 hr at 1000 volts after a prerun of 1 hr at 800 volts. The gels were then subjected to autoradiography at -70°C with the aid of an intensifying screen.

**Sequencing Reactions**

The methodology applied, in the present work, for the determination of the identity of nucleotides at specific positions in the proximity of the poly(A) tail is schematically diagrammed in Figure 2 (in which RNA 16 is used as an example). Following this example, 0.05 to 0.10 pmoles of each RNA species was similarly incubated with its complementary 5'-end labeled oligo(dT) primer (identified as described in the **Results** section) and subjected to cDNA synthesis by reverse transcriptase in the presence of three of the four deoxyribonucleoside triphosphates. Under these conditions cDNA synthesis will proceed until the polymerase reaches the position in the template for which no complementary base is present in the mixture. Subsequent analysis, on a sequencing gel, of the extended primer will then indicate the position

in the nucleotide sequence (as numbered from the 3'-end base adjacent to the poly(A) tail) of the particular base missing in the reaction.

Incubations were carried out as described in the preceding section for the determination of the nucleotide adjacent to the poly(A) tail with the exception that the reaction time was 5 min; four reactions were performed for each RNA species, each reaction lacking one of the four dNTPs. In addition, for each RNA species, a "ladder" reaction was performed, in which all four bases were present.

Transcription was terminated by quick cooling the mixtures in ice-$H_2O$ and adding EDTA to 0.01 M. Excess labeled primer was separated from the hybrid by passage of the mixture (after addition of 5$\gamma$ yeast tRNA carrier) through a Sephadex G-50 column (1 x 25 cm) equilibrated with 0.3 M NaCl, 0.01 M Tris, pH 7.4, 0.01 M EDTA. As before, this and subsequent steps were carried out at 2°C in order to reduce melting of the primer from the template. The void volume was collected, ethanol precipitated, resuspended in 10 $\mu$l $10^{-3}$ M Tris-EDTA followed by 10 $\mu$l 10 M urea/dye solution, heated for 2 min at 90°C, then layered on a 10% polyacrylamide/7 M urea sequencing gel. Results were visualized by autoradiography.

## Results

The first step in the general approach followed here was the 5'-end labeling and subsequent purification of the oligo(dT) primers to be used in the reverse transcriptase reactions. It was considered necessary to obtain high specific activity labeled primers for two reasons. First, HeLa cell mitochondiral polyadenylated RNA species are available in limited amounts [estimated to vary from 0.003 to 0.015 pmoles per gram of cells (depending on the particular RNA species analyzed here)]. Secondly, the possible presence of contaminating short oligonucleotides resulting from the DNase treatment (which could function as primers), eliminated the possibility of using [$\alpha$-$^{32}$P]dNTPs as precursors in the transcriptase reaction. Therefore, each primer

was subjected to 5'-end labeling with high specific activity [$\gamma$-$^{32}$P]ATP as described in **Materials and Methods** and the intact primers isolated by electrophoresis on a polyacrylamide/urea gel. As shown in Figure 3a for p(dT)$_8$-dC, the majority of the counts migrated as a single band at the position expected for an oligonucleotide nine bases in length. The faint bands migrating above and below the main band are due to contaminating oligonucleotides differing in size by one or two bases. The band indicated by the brackets was excised and the material eluted from the gel.

The correct primer to be used with each RNA species (and, thereby the identity of the nucleotide adjacent to the poly(A) tail) was then determined in the following manner. Incubations were carried out in which each primer was tested individually for its ability to prime cDNA synthesis using the RNA as a template. As has been shown previously by others (Cheng et al., 1976; Proudfoot, 1977; Brownlee and Cartwright, 1977; Symons, 1979), the presence of a specific nucleotide at the 3'-end of the oligo(dT) stretch insures that the primer will anneal to the beginning of the poly(A) tail, thereby enabling the reverse transcriptase to initiate cDNA synthesis efficiently and accurately at a specific position which lies 2 bases upstream from the first A residue of the poly(A) tail. Figure 3b illustrates the results of such a test. In this case, RNA 16 was incubated with p(dT)$_8$-dA (dA), p(dT)$_8$-dC (dC), or p(dT)$_8$-dG (dG) in the presence of reverse transcriptase, as described in **Materials and Methods**, and the resulting products analyzed on a 10% polyacrylamide/7 M urea sequencing gel. As can be seen, the mixtures in which p(dT)$_8$-dA or p(dT)$_8$-dC were tested show little extension synthesis, while p(dT)$_8$-dG produced a prominent band pattern extending near the top of the gel. The band pattern thus produced, with the irregular spacing and highly variable intensity of the individual bands was typical and very reproducible for the different RNAs; each RNA produced a characteristic pattern and, as a result, this pattern was considered diagnostic of a positive result.

In some cases, in these and in the following sequencing experiments, a background ladder (in which the bands were fainter and more uniform in their intensity and position) was also observed. This was interpreted to result from reverse transcription of contaminating RNA species.

Figures 4 and 5 illustrate the results obtained when individual RNA species were incubated with the correct oligo(dT) primer (i.e., that in which the 3' terminal nucleotide is complementary to the nucleotide adjacent to the poly(A) tail) and subjected to cDNA synthesis by reverse transcriptase in the presence of three of the four deoxyribonucleoside triphosphates. The omitted nucleotide is indicated at the top of each lane; "L" indicates a control ladder reaction. In general, for each "minus" mixture, a prominent band can be observed which correlates with a specific ladder position. This band reflects the length of the oligo(dT) primer extended by the transcriptase up to a position at which the following nucleotide is missing in the reaction mixture. For example, in RNA 13, the "-A" lane shows a band at ladder position 15, the "-C" lane, a band at ladder position 25, and the "-T" lane, a band at ladder position 11. Therefore, the identity of the nucleotides at positions 16, 26 and 12 in the cDNA can be read as A, C and T, respectively. Correspondingly, the identity of the nucleotides of those positions in the RNA template is U, G and A. The "-G" lane showed no prominent band and, in this case, could be interpreted as being due to the presence of a C residue at position 10 in the RNA sequence. The remainder of the RNA species showed results similar to that of RNA 13, although in some cases, as will be discussed below, an unambiguous assignment of the positions of all four nucleotides could not be made. Thus, some of the RNA species (in particular RNAs 12, 9, 14 and 11) did not show a visible band at position 10 either in the ladder or experimental lane. The probable cause of this is the melting of the short extended primer (10 bases in length) from the RNA template during the chromatography step. In these cases, a positive identification was not made for positions 10 or 11. In some

experiments, multiple prominent bands were observed in an individual lane; when these correlated with prominent ladder bands (for example, refer to the "-C" lane of RNA 12) they were, in general, interpreted to reflect the variable ability of the transcriptase to copy different portions of the RNA template [as has been observed by others (Getz et al., 1974; Verma and Baltimore, 1974; Brownlee and Cartwright, 1977)] and were thus disregarded. In some cases, a faint second band was observed at a position higher up in the ladder than that of the first band (for example, the "-A" lane of RNA 14, the "-G" lane of RNA 12, and the "-A" lane of RNA 16). These were tentatively interpreted to be due to read-through past the first position (perhaps due to the presence of contaminating nucleotides in the reaction mixture). The very intense band in all lanes at ladder position 9 resulted from incomplete separation of unhybridized primer from the reaction product during the chromatography step. Experiments are now in progress to improve this part of the procedure.

In summary, the methodology described above was designed to offer a rapid and sensitive technique for the determination of the precise position of the 3'-end of a polyadenylated transcript encoded within a DNA stretch for which the nucleotide sequence is known. As the results described here demonstrate, the identity of the nucleotides at a number of positions in the proximity of the poly(A) tail can be determined. Subsequent alignment of those determinations with the known DNA sequence then allows the unambiguous mapping of the 3'-end of the transcript. Such an alignment is shown in Figure 6. Here, the nucleotide positions determined for RNA species 13, 12, 9 and 16 have been correlated with the previously determined DNA sequences (Barrell et al., 1979; Sanger, et al., 1980).

## Discussion

The goal of the experiments described in this chapter was to precisely position the 3'-ends of the HeLa cell mitochondrial poly(A)-containing RNA species

within the genetic and physical maps of the HeLa cell mitochondrial genome. As exemplified in Figure 2 for RNA 16, and further illustrated in Figure 6 for RNAs 13, 12, 9 and 16, the approach followed was the alignment of the partially determined nucleotide sequence of the 3'-end of each of the RNA species with that of their respective DNA sequences. The 3'-ends of these transcripts had previously been positioned [within an accuracy of 20-40 nucleotides (Ojala et al., 1980b)] with respect to Hpa II restriction sites; furthermore, the nucleotide sequence of human mitochondrial DNA and, therefore, the positions of the tRNA genes, as reflected by that sequence, has recently been completed (Barrell et al., 1979; Eperon et al., 1980; Sanger et al., 1980; Crews et al., 1979; Crews and Attardi, 1980; F. Sanger and B. Barrell, personal communication). Therefore, the precise positions of the sequences encoding many of the polyadenylated RNA species can be precisely correlated with those of the tRNA genes.

Figure 7 presents a schematic diagram summarizing the relationship of the 3'-end coding sequences of many of the mitochondrial RNA transcripts with those of flanking tRNA genes. Thus, the 12S ribosomal RNA 3'-end coding sequence [recognized for its homology to the Chinese hamster small mitochondrial ribosomal RNA sequence (Dubin and Baer, 1980)] and the poly(A)-containing RNA 3'-end coding sequences [data for RNAs 13, 12, 9 and 16 are taken from Figure 6; data for RNAs 14, 15, 7 and 11 are taken from Figure 5 and aligned with DNA sequence data (B. Barrell and F. Sanger, personal communication)] are shown aligned with their respective tRNA genes. The significant observation illustrated in this figure (and indicated by the arrows) is that the 3'-end coding sequence for each RNA is immediately adjacent to the 5'-end of a tRNA gene, or in the case of RNA 14, is immediately adjacent to the coding sequence of RNA 15 [and thus to the initiator codon of subunit III of cytochrome c oxidase (A. Tzagoloff, personal communication)]. In some cases, the RNA species are shown as terminating with one to three A residues (i.e., RNAs 9,

13 and 14). Obviously, in these instances, it is not possible to distinguish between an A which is part of the poly(A) tail from that which is part of the message. By analogy with the remaining RNA species, we have indicated the 3'-terminal nucleotide of RNAs 9, 13 and 14 as that being the one prior to the 5' terminal nucleotide of the flanking tRNA gene or coding stretch.

The striking relationship of the mitochondrial transcripts described above, i.e., that of the immediate juxtaposition of 3'-end coding sequences with those of flanking tRNA genes, has also been observed for the 5'-ends of the same transcripts. This observation is summarized in Figure 8. Shown in this figure, in their mapped order, are the 12S and 16S ribosomal RNAs and all of the polyadenylated RNA transcripts coded for by the H-strand [with the exception of RNA 4, which is presumed to be a precursor of the two ribosomal RNAs on the basis of its mapped position (Ojala et al., 1980b) and relatively short half-life (Gelfand, 1980)], and their relationship with flanking tRNA genes on their 5'-side. Thus, the 5'-end coding sequences for ten of the thirteen rRNA and polyadenylated RNA species shown here are immediately adjacent to the 3'-terminal nucleotide of a tRNA gene (Crews and Attardi, 1980; Eperon et al., 1980; Ojala et al., 1980; Montoya, Ojala and Attardi, manuscript in preparaton; Chapter 6 of this thesis). The three exceptions are RNA 9, RNA 15 and RNA 11. RNA 9 is probably not a primary transcription product, but derives from RNA 6 by removal of a 5'-terminal stretch containing sequences complementary to four L-strand coded tRNA species. The coding sequences of RNAs 14 and 15 are apparently not separated by a tRNA gene; instead, as previously discussed, the 3'-terminal nucleotide of RNA 14 is juxtaposed to the 5'-terminal nucleotide of RNA 15. A similar situation may occur for RNAs 5 and 11; however, in this case, a tRNA glutamine gene is encoded in the L-strand very close to the 5'-terminus of RNA 11.

The recently completed sequence analysis of human mitochondrial DNA has revealed a high degree of packing of genetic information (F. Sanger and

B. Barrell, personal communication). The transcriptional mapping studies described in Chapter 3 and the 5'-end and 3'-end sequence analyses presented in Chapter 6 and schematically diagrammed here (Figures 7 and 8) corroborate and extend those observations. Thus, RNAs 5, 7, 9 and 11 to 17 [considered to function as mitochondrial specific mRNAs because of their relative abundance, enrichment in polysomal structures, and, additionally, by correlation of the RNA sequence data, or of reading frames in mitochondrial DNA (Walker et al., 1980), with protein sequence data (Walker et al., 1980; Chomyn, Hunkapiller and Attardi, manuscript in preparation; Barrell et al., 1979) and with genetically characterized yeast gene segments (A. Tzagoloff, personal communication)] either start directly at the initiator codon or have only a few nucleotides (1 to 8) preceding this codon. In fact, if the first AUA of RNA 13 functions as the initiator codon, and if the first AUU of RNA 12 also functions in that capacity (as suggested by Walker et al., 1980), then all putative mitochondrial mRNAs would contain the initiator codon within the first six nucleotides.

An examination of the 3'-end of these RNA species reveals a second example of economical and efficient use of genetic information. Analysis of the polypeptide reading frames contained within the RNA coding sequences (references cited above) combined with the unusual codon recognition pattern observed in mitochondria from mammalian cells [UGA codes for tryptophan and AUA is read as methionine rather than as isoleucine (Barrell et al., 1979); no tRNA gene has been found for the arginine codons $AG\begin{matrix}A\\G\end{matrix}$, leading to the suggestion that these codons (in addition to UAA and UAG) may function as terminators (Barrell et al., 1980)] leads to the observation that RNAs 13, 12, 15 and 11 contain no termination codons. It has been proposed that these codons may be provided by the addition of the poly(A) tail (Walker et al., 1980). In fact, examination of the sequence data for RNAs 13 and 12 (Figure 6) show that these species terminate with a UA and a U, respectively. Addition of the poly(A) tail would thus create UA(A) and U(A)(A) triplets, which exist in the

correct reading frame to function as terminator codons. Likewise the gel patterns of RNAs 15 and 11 (Figure 5) indicate that these species both terminate in a U. Correlation with the reading frames of these RNAs (A. Tzagoloff, personal communication) suggests that poly(A) addition will in these cases also provide a terminator codon in the correct reading frame.

As concerns the overall sequence organization of the HeLa cell mitochondrial genome, the mapping data presented in Chapter 3, and the sequencing data presented in Chapters 5, 6 and here indicate that the H-strand sequences coding for the rRNA species, the poly(A)-containing RNAs and the tRNAs are immediately contiguous to each other, extending from coordinate 2/100 to coordinate 95/100 (relative to the origin taken as 0/100). [It is considered likely that the short H-strand segment between $tRNA^{Ile}$ and $tRNA^{F\text{-}met}$, which is complementary to the L-strand $tRNA^{Glu}$, is transcribed and then eliminated in the processing of the primary transcripts.] Therefore, with the exception of the region around the origin of H-strand DNA replication ($\sim$1200 nucleotides), the H-strand DNA sequences are completely saturated by discrete poly(A)-containing RNAs, rRNAs and tRNAs. This arrangement is consistent with a model of transcription of the H-strand in the form of a single molecule which is rapidly processed by precise endonucleolytic cleavages before and after each tRNA sequence to yield the mature products or, in some cases, processing intermediates, like the putative precursors of the rRNAs (RNA 4) and of RNA 9 (RNA 6). Such a model is presented in Figure 9, in which the H and L strands are shown, in addition to the positions of the tRNA genes encoded by the H-strand (F. Sanger and B. Barrell, personal communication). Direction of transcription is from left to right. Previous work from this laboratory (Cantatore and Attardi, 1980) has indicated the existence of a single initiation site for H-strand transcription near the origin of replication. Thus, the 5'-terminus of the large transcript is depicted as originating near or at the 5'-end of RNA 4 (presumed to be a precursor to the ribosomal RNAs plus the

phenylalanine and valine tRNAs). Processing of the nascent transcript by precise endonucleolytic cleavages before and after each tRNA would occur rapidly during transcription (no evidence has been found for discrete H-strand transcripts larger than that of RNA 4). The processing events would not necessarily occur in the sequential order shown.

During these processing events, the tRNA sequences may play an important role as recognition signals, providing the "punctuation" in the reading of mtDNA information (Attardi et al., 1980). Thus, the processing enzyme(s) may recognize the cloverleaf strucure of the tRNA sequence or some portion of it. Exceptions to this model occur at the 5'-end of RNA 9, and at the border between RNAs 14 and 15, and RNAs 5 and 11. However, the DNA segments immediately preceding RNA 9 and RNA 5 contain sequences complementary to an L-strand coded tRNA and could therefore form a cloverleaf-like structure. As to the border between RNAs 14 and 15, a computer analysis of the nucleotide sequence in that region of the genome is now underway to search for the possibility of a sequence or secondary structure formation which could provide a similar function.

As concerns the possible nature of the enzyme(s) involved, considerable analogies can be found in the enzymatic system which processes the rRNA and tRNA transcripts in E. coli. Thus, RNases III, E, F and P are endonucleases which, in a concerted manner, cleave at precise positions in the large primary transcripts to release, respectively, the 16S and 23S rRNAs, the 5S RNA and the 3' and 5' ends of the tRNAs (Apirion et al., 1980). It is believed that these enzymes recognize sites which are involved in secondary structure arrangements. Secondly, genetic studies of the mitochondria of yeast have resulted in the identification of a protein, encoded within an intron of cytochrome b, which catalytically controls the selective splicing of the intron (Jacq et al., 1980). This enzyme has been termed "messenger RNA-maturase" and acts by both cis and trans mechanisms. It is conceivable that

a similar protein could function in the mitochondria of higher eucaryotes, such that the RNA processing events would be under the control of the mitochondrial genome.

If the proposed model is true, then the observed large excess of the ribosomal RNA species over the mRNAs and the quantitative differences among the latter may be accounted for by premature termination of transcription past the ribosomal RNA cistrons [as has been documented for the adenovirus 2 system (Evans et al., 1979; Fraser et al., 1979)] and by differences in metabolic stability of the various RNA species (Gelfand, 1980).

Transcription studies of the L-strand are still in a preliminary state; the majority of the transcripts produced by this strand are very short-lived and therefore difficult to isolate in amounts adequate for sequencing studies. However, the mapping experiments presented in Chapter 3 support the possibility that transcription of this strand may also follow the punctuation model. Thus, the common 5'-end of the large transcripts 1, 2 and 3 encoded by the L-strand have been mapped very close to a tRNA gene on that strand (Ojala et al., 1980; Chapter 3 of this thesis; F. Sanger and B. Barrell, personal communication); additionally the 3'-end of RNA 2 has been localized near an L-strand tRNA gene, and the 3'-end of RNA 3 mapped adjacent to a sequence complementary to an H-strand tRNA gene.

In summary, the results discussed in this and preceding chapters of this thesis describe a transcriptional system which is strikingly characterized by simplicity, efficiency and economy. It is possible that these characteristics of the human cell mitochondrial genome represent a highly evolved form which has emerged in response to particular selective pressures.

**Acknowledgements**

The experiments described in this chapter were carried out in collaboration with Dr. Julio Montoya.

## References

Apirion, D., Ghora, B. K., Plautz, G., Misra, T. K. and Gegenheimer, P. (1980) Processing of rRNA and tRNA in Escherichia coli: Cooperation between processing enzymes. In: Transfer RNA: Biological Aspects. Eds. D. Soll, J. Abelson and P. Schemmel (Cold Spring Harbor Laboratory).

Attardi, G., Cantatore, P., Ching, E., Crews, S., Gelfand, R., Merkel, C., Montoya, J. and Ojala, D. (1980) The remarkable features of gene organization and expression of human mitochondrial DNA. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1980) Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc. Nat. Acad. Sci. USA **77**, 3164-3166.

Barrell, B. G., Bankier, A. T. and Drouin, J. (1979) A different genetic code in human mitochondria. Nature **282**, 189-194.

Brownlee, G. G. and Cartwright, E. M. (1977) Rapid gel sequencing of RNA by primed synthesis with reverse transcriptase. J. Mol. Biol. **114**, 93-117.

Cantatore, P. and Attardi, G. (1980) Mapping of nascent light and heavy strand transcripts on the physical map of HeLa cell mitochondrial DNA. Nucl. Acids Res. **8**, 2605-2624.

Cheng, C. C., Brownlee, G. G., Carey, N. H., Doel, M. T., Gillam, S. and Smith, M. (1976) The 3' terminal sequence of chicken ovalbumin messenger RNA and its comparison with other messenger RNA molecules. J. Mol. Biol. **107**, 527-547.

Crews, S. and Attardi, G. (1980) The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. Cell **19**, 775-784.

Crews, S., Ojala, D., Posakony, J., Nishiguchi, J. and Attardi, G. (1979) Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. Nature **277**, 192-198.

Dubin, D. T. and Baer, R. J. (1980) Sequence and structure analysis of mitochondrial ribosomal RNA from hamster cells. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Eperon, I. C., Anderson, S. and Nierlich, D. P. (1980) Distinctive sequences of human mitochondrial ribosomal RNA genes. Nature **286**, 460-467.

Evans, R., Weber, J., Zeff, E. and Darnell, J. E. (1979) Premature termination during adenovirus transcription. Nature **278**, 367-370.

Fraser, N. W., Sehgal, P. B. and Darnell, J. E. (1979) Multiple discrete sites for premature RNA chain termination late in adenovirus 2 infection: Enhancement by 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). Proc. Nat. Acad. Sci. USA **76**, 2571-2575.

Gelfand, R. A. (1980) Studies on RNA metabolism in HeLa mitochondria. Ph.D. Thesis, California Institute of Technology, Pasadena, California.

Getz, M. J., Birnie, G. D. and Paul, J. (1974) Transcription of in vitro polyadenylated ribonucleic acid with reverse transcriptase. Biochemistry **13**, 2235-2240.

Jacq, C., Lazowska, J. and Slonimski, P. P. (1980) Cytochrome b messenger maturase encoded in the intron regulates the expression of the split gene: I. Physical location and base sequence of intron mutations. In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

Maxam, A. M. and Gilbert, W. (1977) A new method for sequencing DNA. Proc. Nat. Acad. Sci. USA **74**, 560-564.

Ojala, D. and Attardi, G. (1980) Fine mapping of the ribosomal RNA genes of HeLa cell mitochondrial DNA. J. Mol. Biol. **138**, 411-420.

Ojala, D., Merkel, C., Gelfand, R. and Attardi, G. (1980b) The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell, in press.

Ojala, D., Montoya, J. and Attardi, G. (1980b) The putative mRNA for subunit II of cytochrome c oxidase in human mitochondria starts directly at the translation initiator codon. Nature, in press.

Proudfoot, N. J. (1977) Complete 3' noncoding region sequences of rabbit and human β-globin messenger RNAs. Cell **10**, 559-570.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol., in press.

Symons, R. (1979) Extensive sequence homology at the 3'-termini of the four RNAs of cucumber mosaic virus. Nucl. Acids Res. **7**, 825-837.

Verma, I. N. and Baltimore, D. (1974) Purification of the RNA-directed DNA polymerase from avian myeloplastosis virus and its assay with polynucleotide templates. In: Methods in Enzymology, Vol. 29. Eds. L. Grossman and K. Moldave (Academic Press, New York and London), pp. 125-130.

Walker, J. E., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Runswick, M. J., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1980) In: The Organization and Expression of the Mitochondrial Genome. Eds. A. M. Kroon and C. Saccone (North-Holland, Amsterdam), in press.

**Figure 1**

Detailed genetic map of the HeLa cell mitochondrial genome illustrating the positions of the H-strand transcription products.

The two inner circles show the positions of the two rRNA genes (Ojala and Attardi, 1980), and those of the tRNA genes, as derived from the mtDNA sequence (F. Sanger and B. Barrell, personal communication; Barrell et al., 1979; Crews and Attardi, 1980). The positions of the polyadenylated transcripts (Ojala et al., 1980) are indicated by the wide black bars. The direction of transcription is shown by the outermost arrow. The origin (O) and direction (small rightward arrow) of the H-strand synthesis are also shown.

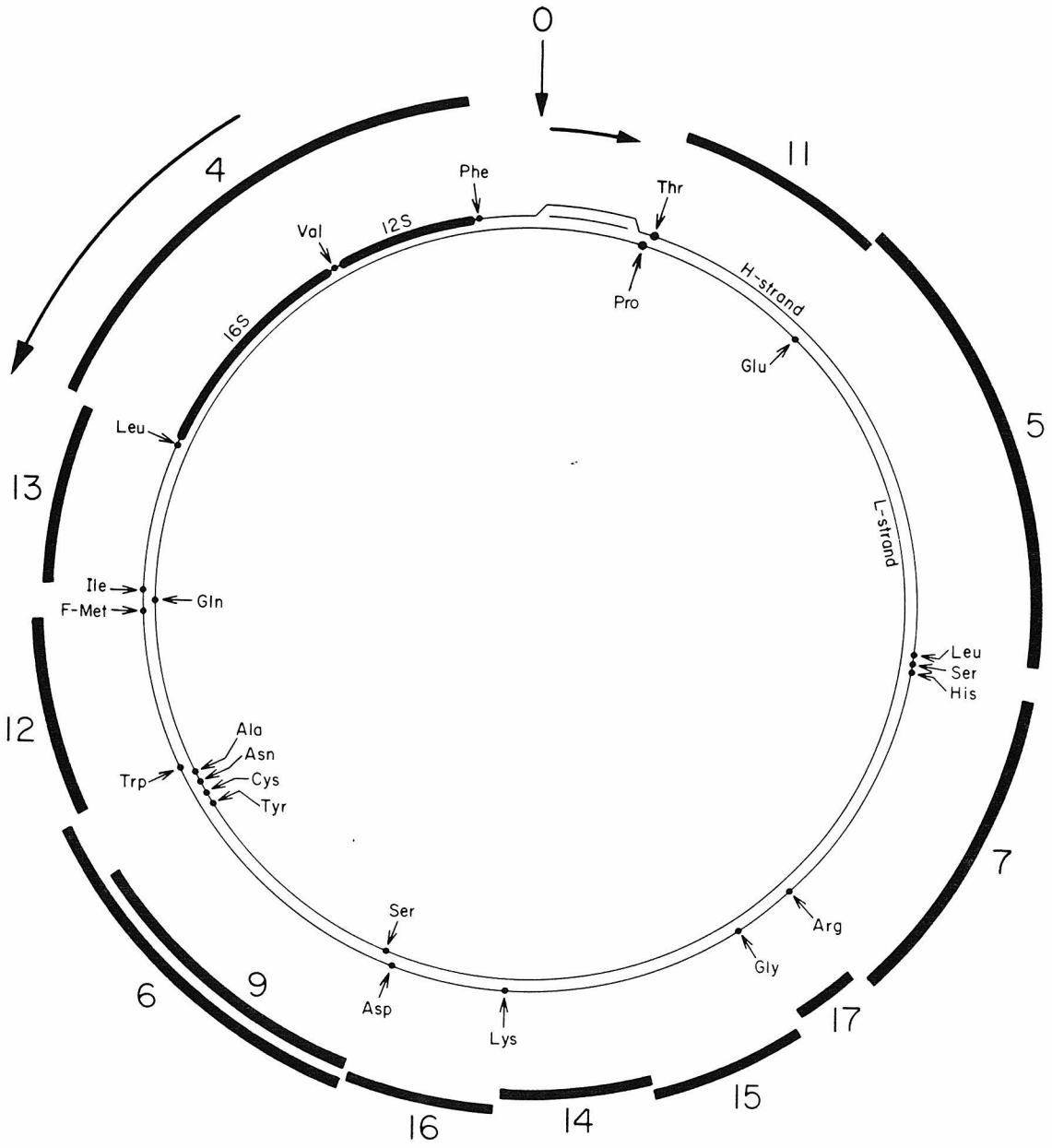

0
11
5
7
17
15
14
16
6
9
12
13
4
Phe
Thr
Val
12S
16S
Pro
H-strand
Glu
Leu
L-strand
Ile
F-Met
Gln
Leu
Ser
His
Ala
Asn
Trp
Cys
Tyr
Ser
Arg
Gly
Asp
Lys

**Figure 2**

Schematic diagram of the methodology used in the present work for the alignment of the 3'-end of polyadenylated transcripts with the DNA sequences encoding those transcripts.

Depicted in the top portion of the figure are the observed products (after electrophoretic separation in a sequencing gel) resulting from the incubation of 5'-end labeled $p(dT)_8$-dG with RNA 16 in the presence of three of the four dNTPs (-A, -T, -C and -G) or all four dNTPs ("complete"), and cDNA synthesis by reverse transcriptase. Numbered positions are those of the extended primer. A plus sign (+) indicates a positive band in the sequencing gel (as defined by criteria described in **Results**). The bottom portion of the figure shows the previously determined DNA sequence and position of the $tRNA^{Lys}$ gene (Barrell et al., 1979). Above this is the deduced sequence of RNA 16 and its alignment with the DNA sequence.

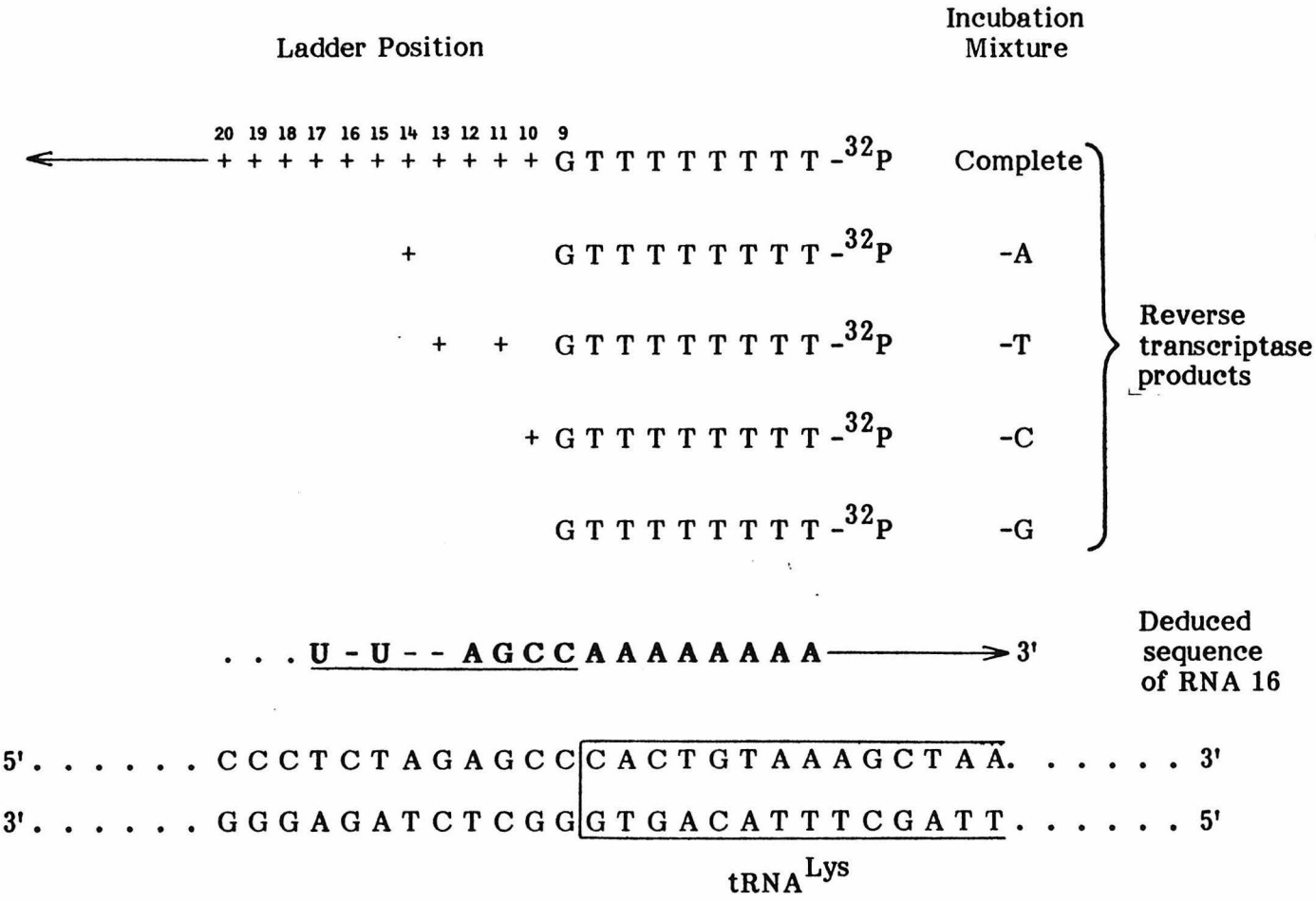

Ladder Position
Incubation Mixture
20 19 18 17 16 15 14 13 12 11 10 9
+ + + + + + + + + + + + G T T T T T T T T -32P
Complete
+ G T T T T T T T T -32P
-A
+ + G T T T T T T T T -32P
-T
+ G T T T T T T T T -32P
-C
G T T T T T T T T -32P
-G
Reverse transcriptase products
. . . U - U - - A G C C A A A A A A A A A
3'
Deduced sequence of RNA 16
5' . . . . . C C C T C T A G A G C C C A C T G T A A A G C T A A . . . . . 3'
3' . . . . . G G G A G A T C T C G G G T G A C A T T T C G A T T . . . . . 5'
tRNALys

**Figure 3**

Autoradiograms of sequencing gels showing the electrophoretic purification of 5'-end labeled $p(dT)_8$-dC primer and subsequent annealing tests using this and other similarly labeled primers with RNA 16.

Panel a: Purification, on a 15% polyacrylamide gel, of $p(dT)_8$-dC primer after 5'-end labeling with [$\gamma^{32}$P] ATP and polynucleotide kinase. The material indicated by the brackets was excised and eluted. Panel b: Electrophoretic analysis, on a 10% polyacrylamide/7 M urea sequencing gel, of the products produced by incubating $p(dT)_8$-dA (dA), $p(dT)_8$-dC (dC) or $p(dT)_8$-dG (dG) with RNA 16 in the presence of the four dNTPs and subjecting the mixture to cDNA synthesis with reverse transcriptase.

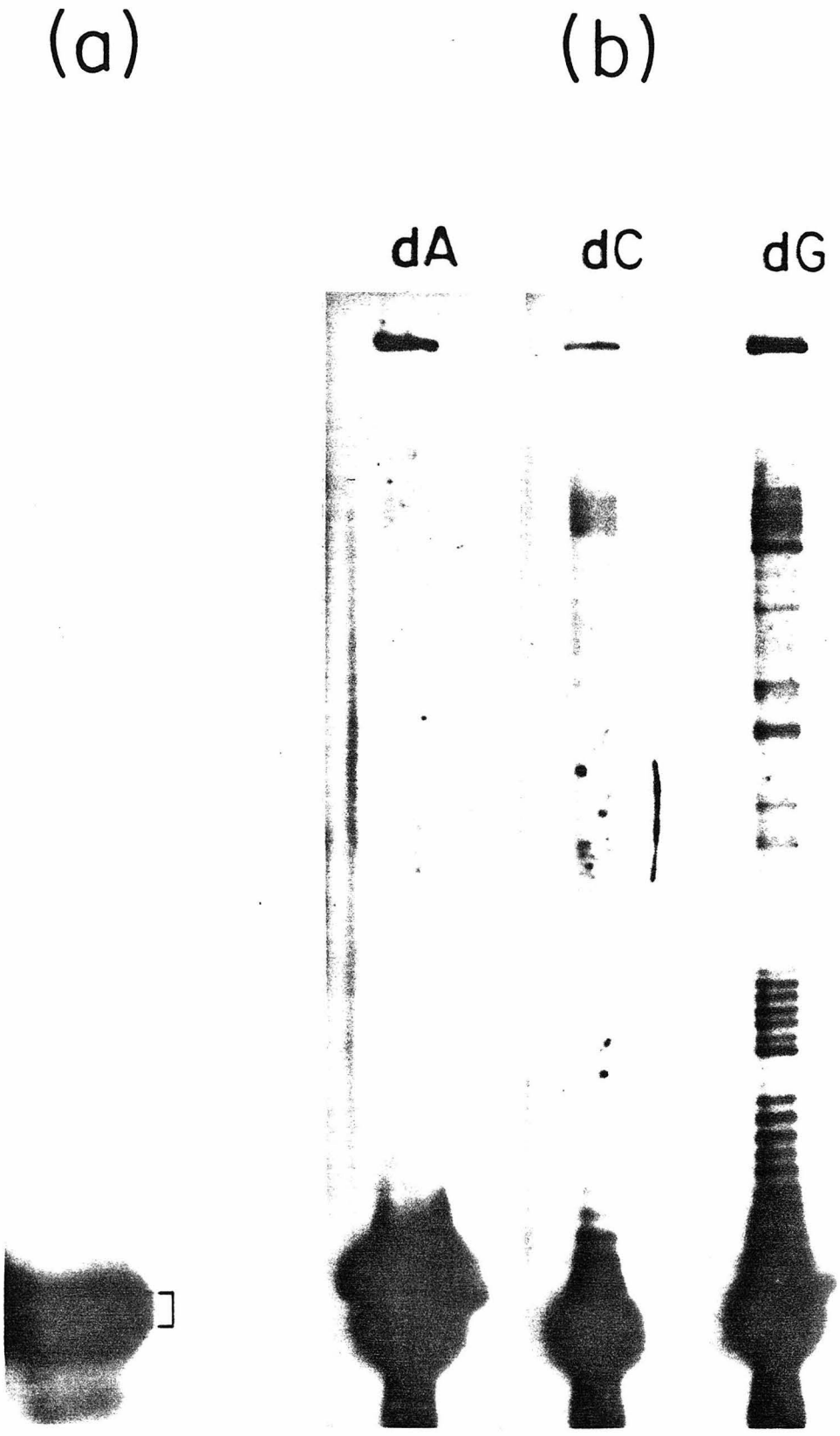

(a)
(b)
dA
dC
dG

**Figure 4**

Sequencing gels of mitochondrial poly(A)-containing RNA species 13, 12, 9 and 16.

Autoradiograms of 10% polyacrylamide/7 M urea gels showing the results obtained when each RNA species was incubated with its complementary primer and subjected to cDNA synthesis in the presence of three of the four deoxyribonucleoside triphosphates. Letters above each lane indicate the nucleotide omitted from the reaction mixture; L indicates a control ladder reaction in which all four nucleotides are present. Ticks and numbers to the right of each panel indicate the various ladder positions. A horizontal arrow refers to a band interpreted as a positive result (as defined by criteria described in the **Results** section).

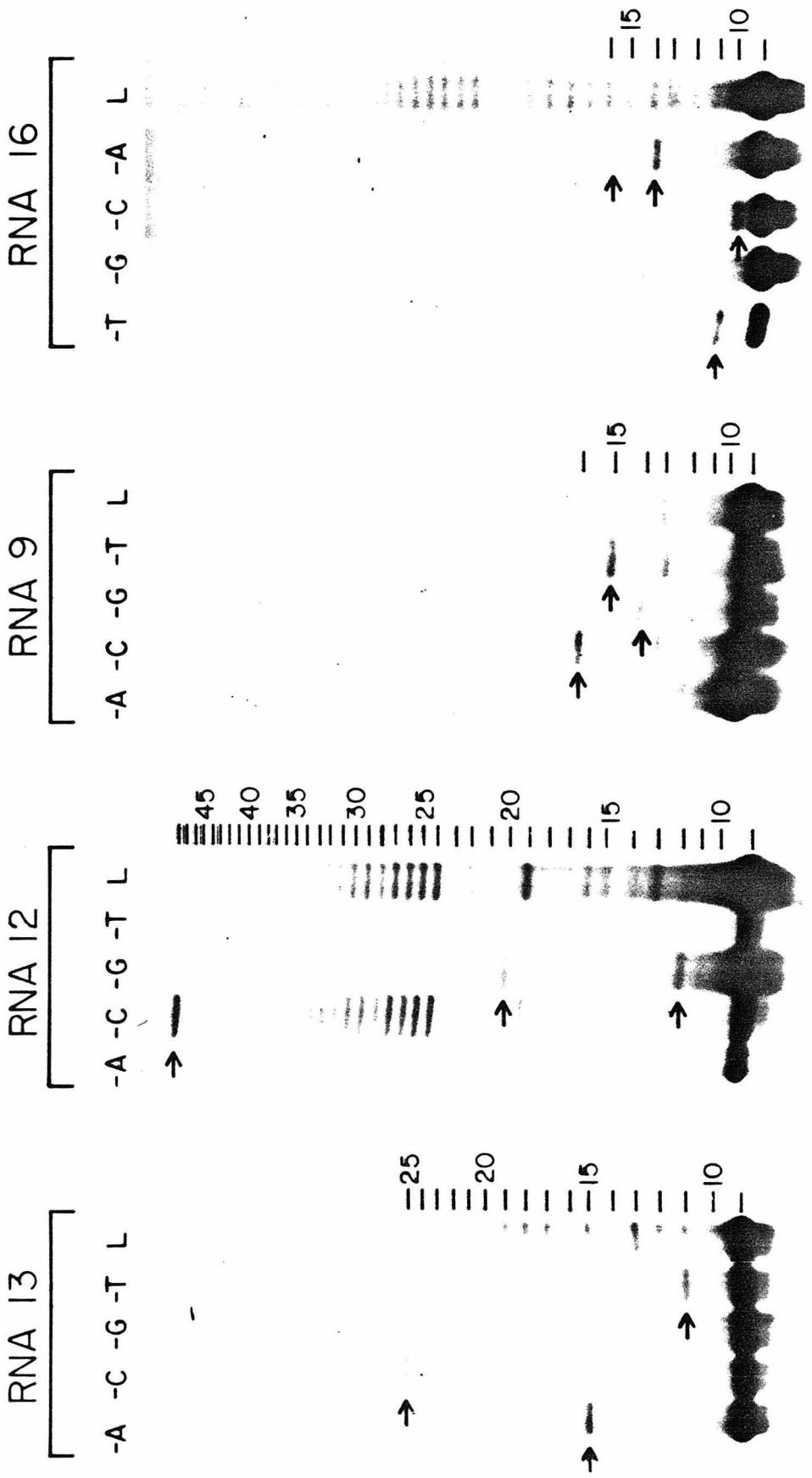

RNA 16
-T -G -C -A L
RNA 9
-A -C -G -T L
RNA 12
-A -C -G -T L
RNA 13
-A -C -G -T L

**Figure 5**

Sequencing gels of mitochondrial poly(A)-containing RNA species 14, 15, 7 and 11.

Gel conditions and numerical and letter designations are the same as those described in the legend of Figure 4.

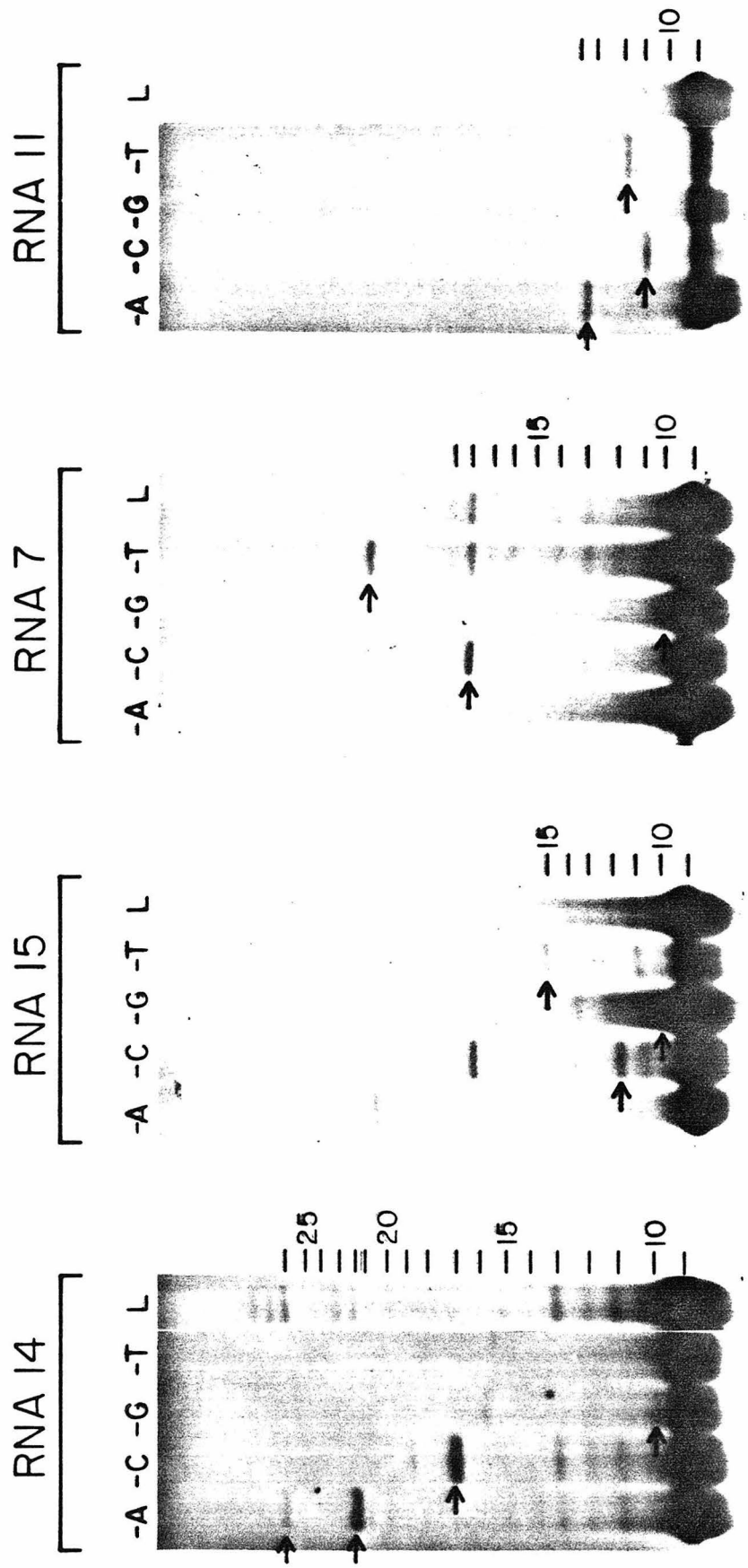

RNA 14
-A -C -G -T L
25
20
15
10
RNA 15
-A -C -G -T L
15
10
RNA 7
-A -C -G -T L
15
10
RNA 11
-A -C -G -T L
10

**Figure 6**

Alignment of the 3'-end proximal regions of mitochondrial poly(A)-containing RNA species 13, 12, 9 and 16 with their corresponding DNA sequences.

The identities of specific nucleotides in the 3' proximal region of each RNA species, as determined in the present work, are shown aligned with previously published DNA sequences. A 5' proximal segment of the sequence of the tRNA genes lying on the 3'-side of the RNA coding stretch is also shown. A dash (-) indicates an unidentified base.

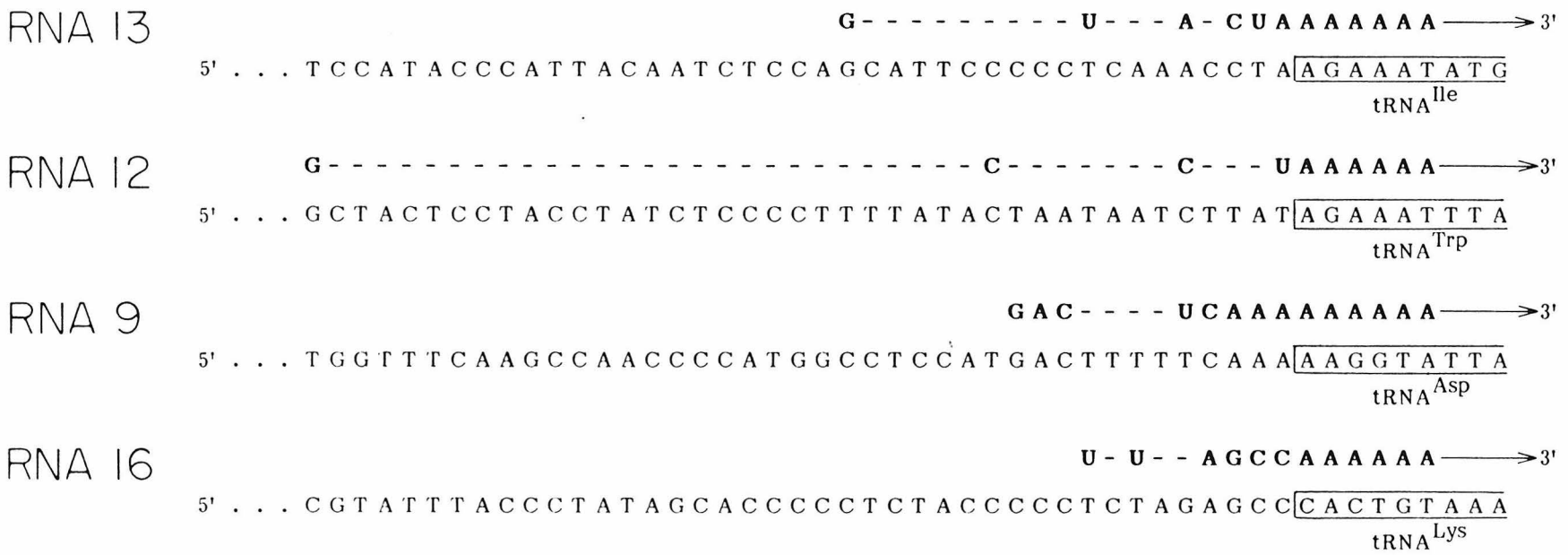

RNA 13
G - - - - - - - - - U - - - A - C U A A A A A A A → 3'
5' . . . T C C A T A C C C A T T A C A A T C T C C A G C A T T C C C C C T C A A A C C T A A G A A A T A T G
tRNA Ile
RNA 12
G - - - - - - - - - - - - - - - - - - - - - - - - - - - C - - - - - - - C - - - U A A A A A A → 3'
5' . . . G C T A C T C C T A C C T A T C T C C C C T T T T A T A C T A A T A A T C T T A T A G A A A T T T A
tRNA Trp
RNA 9
G A C - - - - U C A A A A A A A A A A → 3'
5' . . . T G G T T T C A A G C C A A C C C C A T G G C C T C C A T G A C T T T T T C A A A A A G G T A T T A
tRNA Asp
RNA 16
U - U - - A G C C A A A A A A → 3'
5' . . . C G T A T T T A C C C T A T A G C A C C C C C T C T A C C C C C T C T A G A G C C C A C T G T A A A
tRNA Lys

**Figure 7**

Diagram illustrating the juxtaposition of the 12S rRNA or poly(A)-containing RNA coding sequences and the flanking tRNA genes on their 3'-side.

The sequence of the 3'-end proximal region of each RNA was aligned with the DNA sequence (Barrell et al., 1979; Sanger et al., 1980; Dubin and Baer, 1980; F. Sanger and B. Barrell, personal communication). The arrows indicate the nucleotide immediately adjacent to the 5'-terminal residue of the flanking tRNA gene, or in the case of RNA 14, to the 5'-terminal residue of RNA 15.

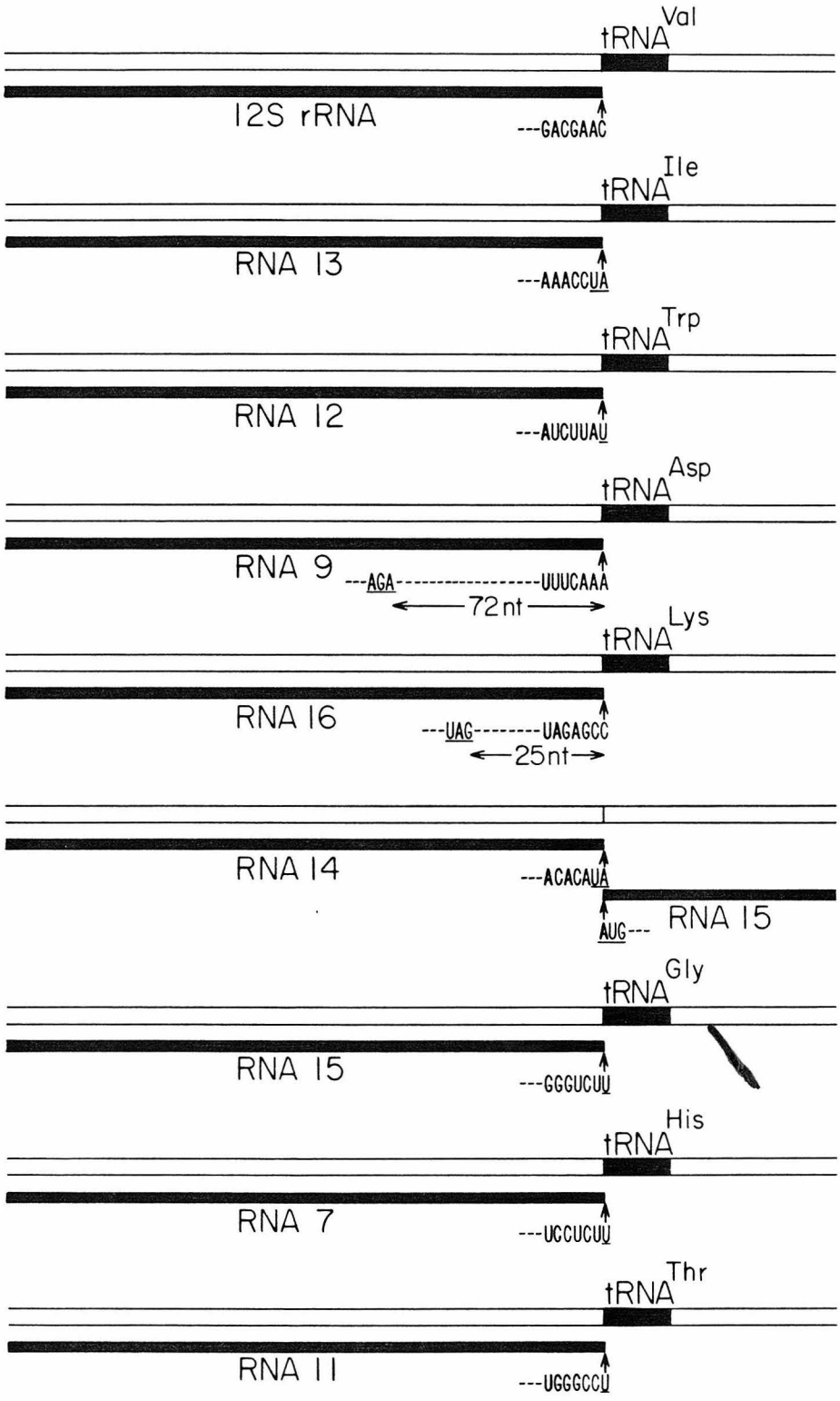

tRNA Val
12S rRNA
---GACGAAC
tRNA Ile
RNA 13
---AAACCUA
tRNA Trp
RNA 12
---AUCUUAU
tRNA Asp
RNA 9
---AGA----------UUUCAAA
72nt
tRNA Lys
RNA 16
---UAG--------UAGAGCC
25nt
RNA 14
---ACACAUA
AUG--- RNA 15
tRNA Gly
RNA 15
---GGGUCUU
tRNA His
RNA 7
---UCCUCUU
tRNA Thr
RNA 11
---UGGGCCU

**Figure 8**

Diagram illustrating the juxtaposition of the 12S and 16S rRNA or poly(A)-containing RNA coding sequences and the flanking tRNA genes on their 5'-side.

The sequence of the 5'-end proximal region of each RNA was aligned with the DNA sequence (Crews and Attardi, 1980; Eperon et al., 1980; Barrell et al., 1979; Ojala et al., 1980; Montoya, Ojala and Attardi, manuscript in preparation. The arrows indicate the nucleotide immediately adjacent to the 3'-terminal residue of the flanking tRNA gene, or in the case of RNAs 9, 15 and 11, the 5'-terminal nucleotide.

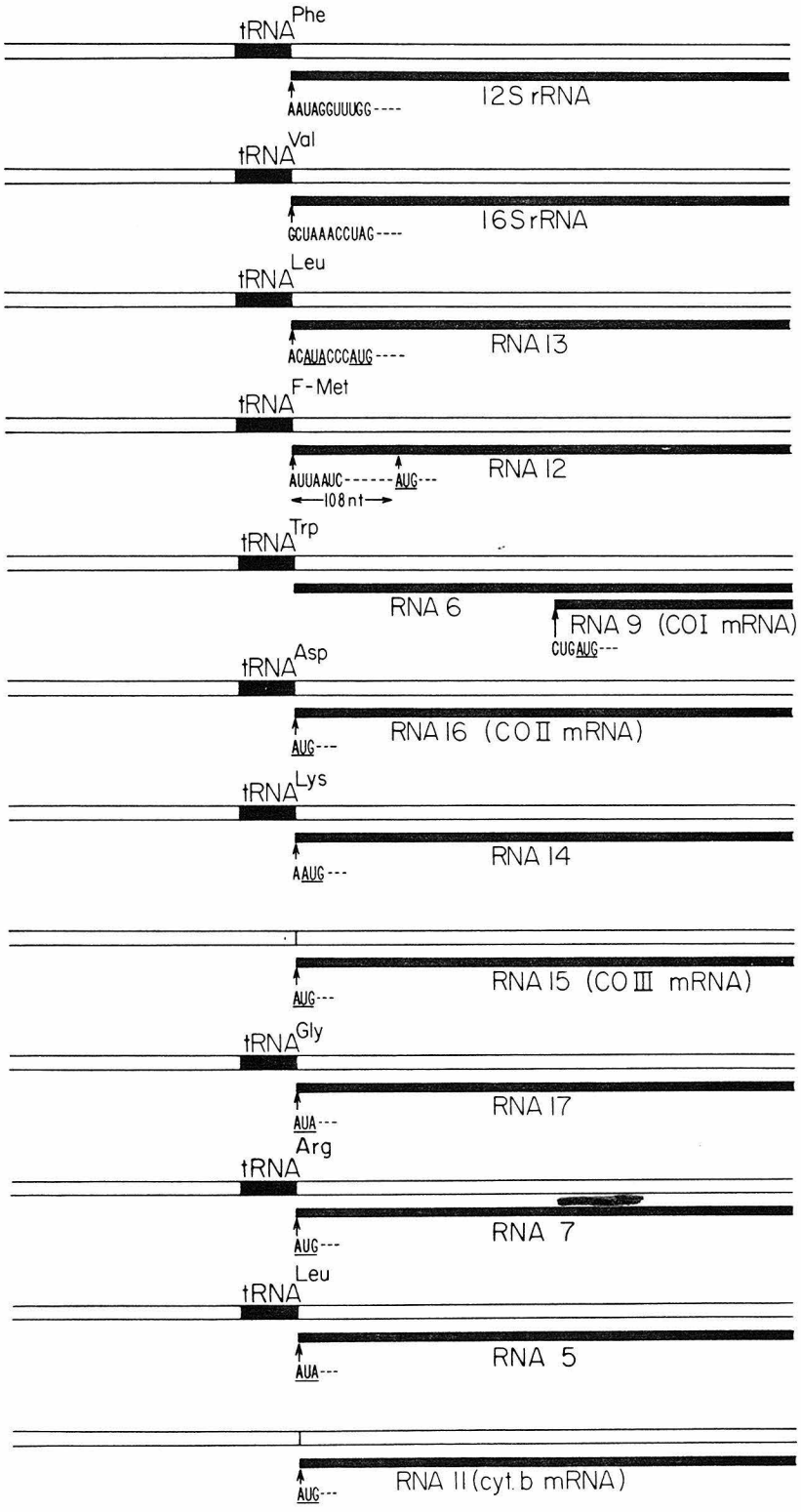
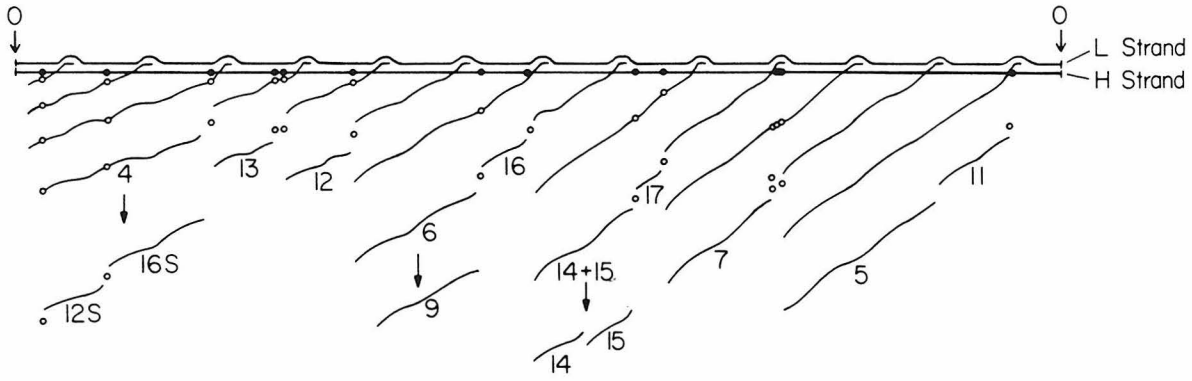

tRNA Phe
AAUAGGUUUGG----
12S rRNA
tRNA Val
GCUAAACCUAG----
16S rRNA
tRNA Leu
ACAUAACCCAUG----
RNA 13
tRNA F-Met
AUUAAUC------AUG---
108nt
RNA 12
tRNA Trp
RNA 6
RNA 9 (COI mRNA)
CUGAUG---
tRNA Asp
AUG---
RNA 16 (COII mRNA)
tRNA Lys
AAUG---
RNA 14
AUG---
RNA 15 (COIII mRNA)
tRNA Gly
AUA---
RNA 17
tRNA Arg
AUG---
RNA 7
tRNA Leu
AUA---
RNA 5
AUG---
RNA 11 (cyt. b mRNA)

**Figure 9**

Proposed model for mtDNA H-strand transcription.

Illustrated here is a diagrammatic representation of the processing of nascent mitochondrial RNA chains. ● refers to a tRNA gene; o, to a mature tRNA; O, to the origin of H-strand replication. Direction of transcription proceeds from left to right.


0
0
L Strand
H Strand
4
13
12
16
17
11
6
16S
12S
9
14+15
7
5
14
15

To the future . . .

*I should like to rise and go where the golden apples grow;*
*Where below another sky Parrot Islands anchored lie,*
*And, watched by cockatoos and goats, lonely Crusoes building boats;*
*Where in sunshine reaching out Eastern cities, miles about,*
*Are with mosque and minaret among sandy gardens set,*
*Where the Great Wall round China goes, and on one side the desert blows,*
*And with bell and voice and drum, cities on the other hum;*
*Where are forests, hot as fire, wide as England, tall as a spire,*
*Full of apes and coconuts and the Negro hunters' huts;*
*Where the knotty crocodile lies and blinks in the Nile,*
*And the red flamingo flies hunting fish before his eyes;*
*Where in jungles, near and far, man-devouring tigers are,*
*Lying close and giving ear lest the hunt be drawing near,*
*Or a comer-by be seen swinging in a palanquin;*
*Where among the desert sands some deserted city stands,*
*All its children, sweep and prince, grown to manhood ages since,*
*Not a foot in street or house, not a stir of child or mouse,*
*And when kindly falls the night, in all the town no spark of light.*
*There we'll come, my man and I, and as the hours pass swiftly by,*
*We'll light a fire in the gloom of some dusty dining room.*
*See the pictures on the walls, heroes, fights, and festivals;*
*And in a corner find the toys of the old Egyptian boys.*

Modified, with thanks, from the original by Robert Louis Stevenson.