MOLECULAR AND CELLULAR BIOLOGY, Feb. 1987, p. 925–931
0270-7306/87/020925-07$02.00/0

Vol. 7, No. 2

The Excised Leader of Human Cytochrome *c* Oxidase Subunit I mRNA Which Contains the Origin of Mitochondrial DNA Light-Strand Synthesis Accumulates in Mitochondria and Is Polyadenylated

GEORGE GAINES,† CESARE ROSSI, AND GIUSEPPE ATTARDI*

Division of Biology, California Institute of Technology, Pasadena, California 91125

Received 8 July 1986/Accepted 6 November 1986

We identified a polyadenylated RNA species which contains the origin of human mitochondrial DNA light-strand synthesis and the surrounding complementary sequences of the four light-strand-encoded tRNAs. This RNA (RNA 9L) is probably derived from the leader portion of RNA 6 which is excised during the formation of the mature cytochrome *c* oxidase subunit mRNA (RNA 9). The high degree of secondary structure of this RNA is presumably responsible for its anomalous electrophoretic behavior in denaturing polyacrylamide gels.

The tRNA punctuation model for RNA processing in mammalian mitochondria proposes that the tRNA sequences interspersed between the rRNA- and protein-coding sequences in the polycistronic transcripts of the mitochondrial DNA (mtDNA) heavy (H) strand act as processing signals for the formation of the rRNAs, mRNAs, and tRNAs (15, 16). The formation of cytochrome *c* oxidase subunit I (COI) mRNA (RNA 9) represents an unusual situation in the maturation of mRNAs in human mitochondria. This RNA is not derived directly from processing of the primary transcript but rather is derived via an intermediate (RNA 6) (9, 15) by the removal of a 5′ segment which is complementary to the four tRNAs encoded in the light (L) strand and to the template sequence for the initiation of L-strand synthesis (2) (Fig. 1). Using a transcription system involving isolated mitochondria, we identified an RNA species (designated RNA 9L) which corresponds to the leader portion of RNA 6 and which also occurs in vivo. The in vivo accumulation and properties of RNA 9L have given new insights into the processing of RNA in human mitochondria, at the same time hinting at a possible physiological role of this RNA species.

Identification of the excised leader of COI mRNA (RNA 9L). Transcription of mtDNA and RNA processing in isolated HeLa cell mitochondria have been shown to faithfully reproduce the in vivo events (8). Figure 2 shows the autoradiographic patterns of mRNAs which had been labeled in vivo for 4 h with $^{32}P_i$ in the presence of 0.1 μg of actinomycin D per ml (12) (Fig. 2b) or in isolated organelles for 30 min with [α-^{32}P]UTP as a precursor (7, 8) (Fig. 2a), fractionated on oligo(dT)-cellulose (15), and electrophoresed through 5% polyacrylamide–7 M urea gels. Although the relative migration of the RNA species in this gel system was different from that observed in a CH_3HgOH-agarose gel system (1), the components indicated have been identified by S1 nuclease mapping experiments (J. Montoya, G. Gaines, M. King, C. Rossi, and G. Attardi, manuscript in preparation). Of special interest are the three bands migrating in the region of RNA 17 (mRNA of subunit ND3 of NADH dehydrogenase [5]) in the in vitro-labeled RNA. These were also present in the in vivo-labeled RNA, although in Fig. 2 they appeared much less intense, owing to the shorter exposure of the autoradiogram (see below). The slowest-migrating band was retained on oligo(dT)-cellulose and was, therefore, presumably polyadenylated, while the two faster-migrating bands were not retained. S1 nuclease protection analysis with in vitro-labeled RNA species eluted from the gel and an M13-cloned mtDNA H-strand fragment containing a portion of the RNA 17-coding sequence showed that the two slowest-migrating RNA species mapped in the region corresponding to RNA 17 and were, therefore, the polyadenylated and nonpolyadenylated forms of RNA 17 (designated 17b and 17u, respectively). In contrast, the fastest-migrating RNA species appeared to be unrelated to RNA 17 (data not shown). The size of this RNA species (~320 nucleotides [nt]) suggested that it might be the leader of COI mRNA excised from RNA 6. Hence, a second S1 nuclease protection experiment was performed with the M13-cloned mtDNA H-strand fragment from the *Eco*RI site at position 5279 to the *Hin*dIII site at position 6208 (mp19.HE2), which corresponds to the 3′-end-proximal 233 nt of RNA 12, $tRNA^{Trp}$, and the 5′-end-proximal 628 nt of RNA 6 (2, 15) (Fig. 1). The results of this experiment clearly indicated that the RNA in question is an H-strand transcript from the above-mentioned region. Further S1 nuclease protection experiments with the subfragments *Eco*RI-*Hpa*I and *Hpa*I-*Hin*dIII of the *Eco*RI-*Hin*dIII fragment (derived from the M13 clone mp8.HE2) (Fig. 1) yielded 5′-end- and 3′-end-protected RNA fragments of 100 and 204 nt, respectively, showing that RNA 9L corresponds precisely to the H-strand segment complementary to the four tRNA genes surrounding the origin of L-strand synthesis (Fig. 1 and data not shown). This RNA was therefore designated 9L (for RNA 9 leader).

Anomalous electrophoretic behavior of RNA 9L. During the mapping of RNA 9L, it was observed that RNA 9L migrated much more slowly in a 10% polyacrylamide–7 M urea gel than expected for its size. This was most clearly seen when the mRNA was fractionated by bidimensional electrophoresis in denaturing gels, the first and second runs being done in 5 and 10% polyacrylamide–7 M urea gels, respectively. All the components of the total in vitro-labeled RNA, including

* Corresponding author.
† Present address: Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637.

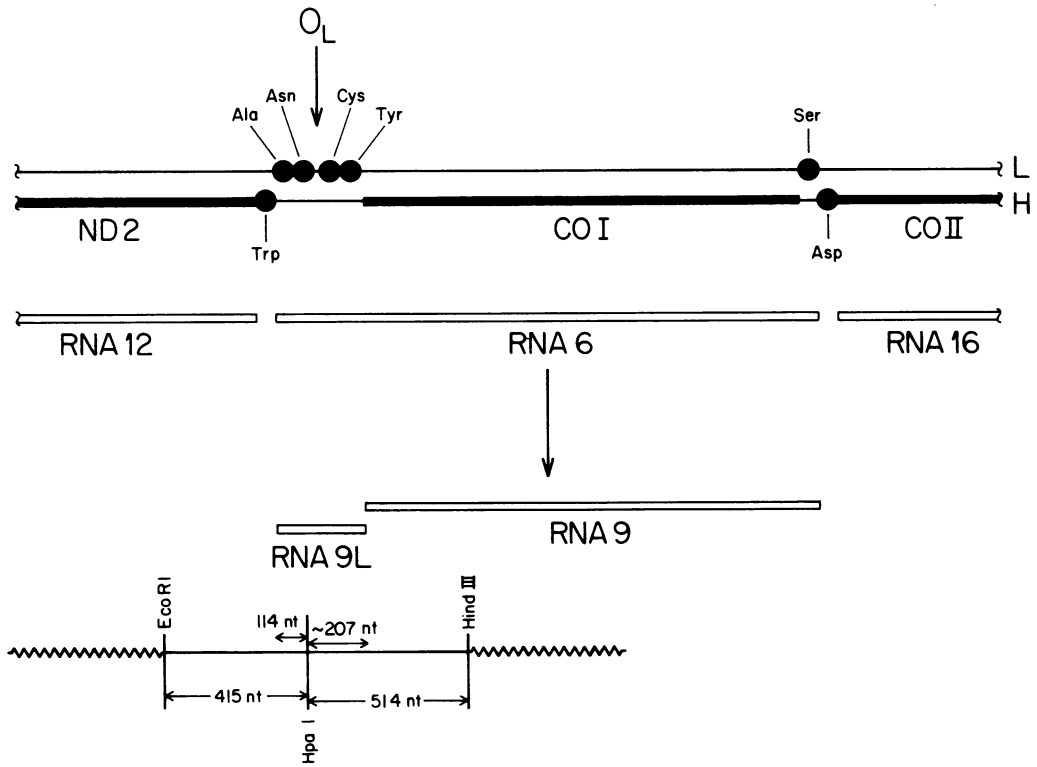

FIG. 1. Portion of the HeLa cell mitochondrial DNA genetic and transcription maps illustrating the COI gene region and the adjacent regions (2). The mapping positions of the COI mRNA precursor (RNA 6) and of the mature COI mRNA (RNA 9) are shown, along with that of RNA 9L, as determined in the present work by S1 nuclease mapping and primer extension experiments. In the lower part of the diagram, the positions of the *Eco*RI-*Hpa*I and *Hpa*I-*Hin*dIII fragments (derived from the M13 clone mp8.HE2; see the text) used in the S1 nuclease protection analysis of RNA 9L and the sizes of the protected segments expected for the leader of RNA 6 are shown. The maps also show the 3'-proximal portion of the gene for a subunit of NADH dehydrogenase (ND2; formerly URF2 [5]), the 5'-proximal portion of the gene for cytochrome *c* oxidase subunit II (COII), and the corresponding transcripts. O_L, Origin of L-strand DNA synthesis.

RNAs 17b and 17u, migrated in the second dimension along the diagonal, except RNA 9L, which migrated much more slowly than expected (Fig. 3a). After fractionation of the in vitro-labeled RNA by oligo(dT)-cellulose chromatography, it was found that the majority of the slower-migrating component was in the unbound fraction (Fig. 3b, panels 2 and 4). When RNA that had been labeled in vivo as in the experiment in Fig. 2 and fractionated into oligo(dT)-cellulose-bound and -unbound fractions was analyzed, the slower-migrating component was also observed, with the majority, however, in the bound fraction (Fig. 3b, panels 1 and 3). The labeling of polyadenylated RNA 9L after a 4-h in vivo exposure of the cells to $^{32}P_i$ was about one-half of that of RNA 17b. This value was estimated from the intensities of the bands in the 10% polyacrylamide–7 M urea gel in another experiment in which a thick slice was cut from the RNA 17b-RNA 9L region of a 5% gel and the RNA was further electrophoresed in a 10% gel (data not shown).

The anomalous behavior of RNA 9L in the 10% polyacrylamide–7 M urea gel persisted even after it was boiled for 5 min in 50 mM CH_3HgOH, indicating that the unusual structure of RNA 9L withstood denaturation or reformed easily upon removal of the CH_3HgOH, despite the presence of 7 M urea.

Structure of RNA 9L. Anomalous electrophoretic migration in denaturing polyacrylamide gels has been associated with branched or lariat RNA structures (6, 17, 18). On the other hand, the high degree of secondary structure expected for RNA 9L could conceivably have been responsible for its unusual electrophoretic properties. Several experiments were done to examine the structure of RNA 9L. An oligodeoxynucleotide (number 1) which is complementary to a 22-nt segment of RNA 9L between nt 120 and 141 from the 5' end of RNA 6 was synthesized (15) (Fig. 4a). The oligodeoxynucleotide was hybridized to in vitro-labeled RNA 9L, and the hybrid was subjected to treatment with RNase H, an endonuclease which digests the RNA portion of RNA-DNA hybrids. Two major bands corresponding to RNA segments of about 183 and 120 nt were observed (Fig. 5a); a fainter band corresponding to an RNA segment of about 177 nt was also seen. These results excluded a lariat structure for RNA 9L in which the oligonucleotide is complementary to the circular portion of the lariat (model 4 in Fig. 4b) but could not distinguish among models 1, 2, and 3.

Primer extension experiments were then carried out with the same oligodeoxynucleotide (number 1) labeled at its 5' end with [γ-^{32}P]ATP (10), and unlabeled, electrophoretically purified RNA 9L as a template. After electrophoresis of the extended primer through a denaturing 15% polyacrylamide gel, a major band representing a product whose 3' end corresponds to the T or A at position 5580 or 5581, respectively, was observed (Fig. 5b). This result ruled out model 2 in Fig. 4b for RNA 9L but did not exclude model 3. To rule out model 3, we synthesized another oligodeoxynucleotide (number 2), which is complementary to a 20-nt segment of RNA 6 between positions 5884 and 5903 (2); this segment

was expected to include the putative 3′-end-proximal 17 nt of RNA 9L (if this RNA extended up to the 5′ end of COI mRNA) and the first 3 nt of COI mRNA (4, 13). This oligonucleotide, labeled at its 5′ end, was used as a primer in a reverse trancriptase reaction with unlabeled RNA 9L as a template. Figure 4b (models 5 and 6) shows the results expected for a linear structure or a lariat structure, and Fig. 5c shows the results obtained. After electrophoresis of the extended primer through a 5% polyacrylamide–7 M urea gel, the longest extended primer observed had a size of ~335 nt. There appeared to be several major bands faster moving than the 335-nt band. These presumably resulted from stops in the reverse transcriptase progression caused by the high degree of secondary structure of RNA 9L. These results appeared to exclude conclusively a lariat structure for RNA 9L. They also implied that RNA 9L extends on the 3′ end to a position which is probably less than 10 to 12 nt from the 5′ end of COI mRNA, if one assumes a minimum overlap of 5 to 7 nt between RNA 9L and the oligodeoxynucleotide as being necessary for priming.

In further experiments, a segment of human mtDNA including the RNA 9L-coding sequence, plus ~300 base pairs upstream and ~300 base pairs downstream, was cloned in the pGEM-1 vector (Promega Biotec). After appropriate cuts of the plasmid, transcripts of the insert were made in vitro by using the SP6 or T7 promoter and the corresponding RNA polymerase and were electrophoresed through a 5 or 10% polyacrylamide–7 M urea gel. It was found that any transcript which corresponded to all or a major portion of the RNA 9L region, independently of the coding strands, migrated anomalously in the 10% denaturing gel (data not shown). The RNA 9L sequence or its complementary sequence was, therefore, capable of conferring abnormal electrophoretic properties on transcripts having extra sequences at the 5′ end or 3′ end or both.

Conclusions. RNA 9L identified in this work most probably represents the excised leader of RNA 6. Previous evidence has indicated that RNA 6 is a short-lived intermediate (9). S1 nuclease protection and primer extension experiments localized the 5′ end of RNA 9L to the T or A at position 5580 or 5581, respectively, i.e., to positions immediately adjacent to $tRNA^{Trp}$ (2).The position of the 3′ end of RNA 9L was, on the contrary, not determined precisely. The size expected for an RNA derived from endonucleolytic cleavage of RNA 6 immediately upstream of the 5′ end of COI mRNA is 321 nt (2) (Fig. 4a). The size estimated for in vitro-labeled RNA 9L, which was mostly nonpolyadenylated (~320 nt; Fig. 5a), and the results of the of the S1 nuclease protection, RNase H digestion, and primer extension experiments are compatible with RNA 9L being the complete leader of COI mRNA. Further work, however, is needed to determine the precise 3′ end of RNA 9L.

The finding of an RNA species with ends mapping at the positions indicated above provides some insights into the mechanism of RNA processing in mammalian mitochondria. The accumulation of RNA 9L itself indicates that the processing machinery does not recognize efficiently the anti-tRNA sequences as processing signals. The 5′ end of RNA 9L molecules appears to derive from cleavage of the polycistronic H-strand transcript at the 3′ end of H-strand-encoded $tRNA^{Trp}$, and there is no evidence of cleavage at the 5′ end of anti-$tRNA^{Ala}$, situated 7 nt downstream. Previously, it was shown that the mature polyadenylated RNA 9 (COI mRNA) contains the anti-$tRNA^{Ser}$ sequence in the 3′-untranslated region (2, 16), again indicating that, in general, the sequences complementary to tRNAs are not

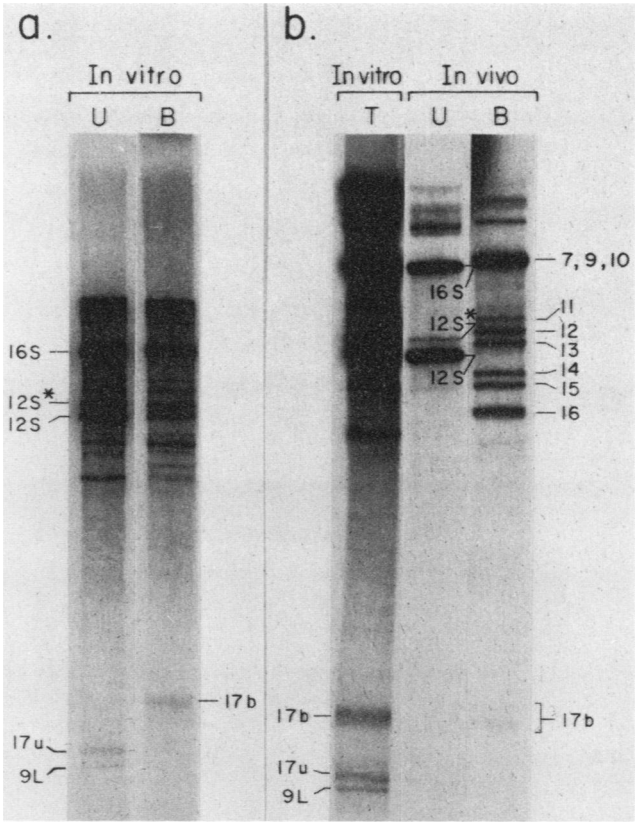

FIG. 2. Electrophoretic patterns in 5% polyacrylamide–7 M urea gels of RNA labeled in isolated mitochondria or in vivo in the presence of 0.1 μg of actinomycin D per ml and fractionated on oligo(dT)-cellulose columns. T, Total RNA; U, oligo(dT)-cellulose-unbound RNA; B, oligo(dT)-cellulose-bound RNA. In panel a, the RNA in lane B was derived from 2.5 times the amount of mitochondria as the RNA in lane U; in panel b, the RNA in lane B was derived from 10 times the amount of cells as the RNA in lane U. 12S* is a 12S rRNA precursor (7).

recognized by the processing enzyme(s). However, it is clear that at least one of the anti-tRNA sequences can act as a processing signal: in fact, cleavage at the 5′ end and, possibly, also at the 3′ end of anti-$tRNA^{Glu}$ is required for the formation of RNA 11 and RNA 5 (3, 13, 15). It seems likely that the structures of the individual anti-tRNAs and the surrounding sequences determine the efficiency of the anti-tRNAs as signals for the mitochondrial processing enzymes.

The evidence discussed above has clearly indicated a linear structure for RNA 9L, implying that the anomalous electrophoretic behavior of RNA 9L is due to its capacity to form an unusually stable secondary structure, which impedes its progress through a 10% polyacrylamide–7 M urea gel. The most stable potential secondary structure of the molecule, as determined by the Nussinov-Jacobson algorithm (14), has a long, central, base-paired segment with a terminal loop and some interstitial loops, as well as several other stem and loop structures; one of the latter is the evolutionarily conserved stem and loop structure which contains the origin of L-strand synthesis (19). With due caution concerning a possible extrapolation to the in vivo situation, it is interesting that the above configuration does not exhibit any cloverleaf structure. To our knowledge, this

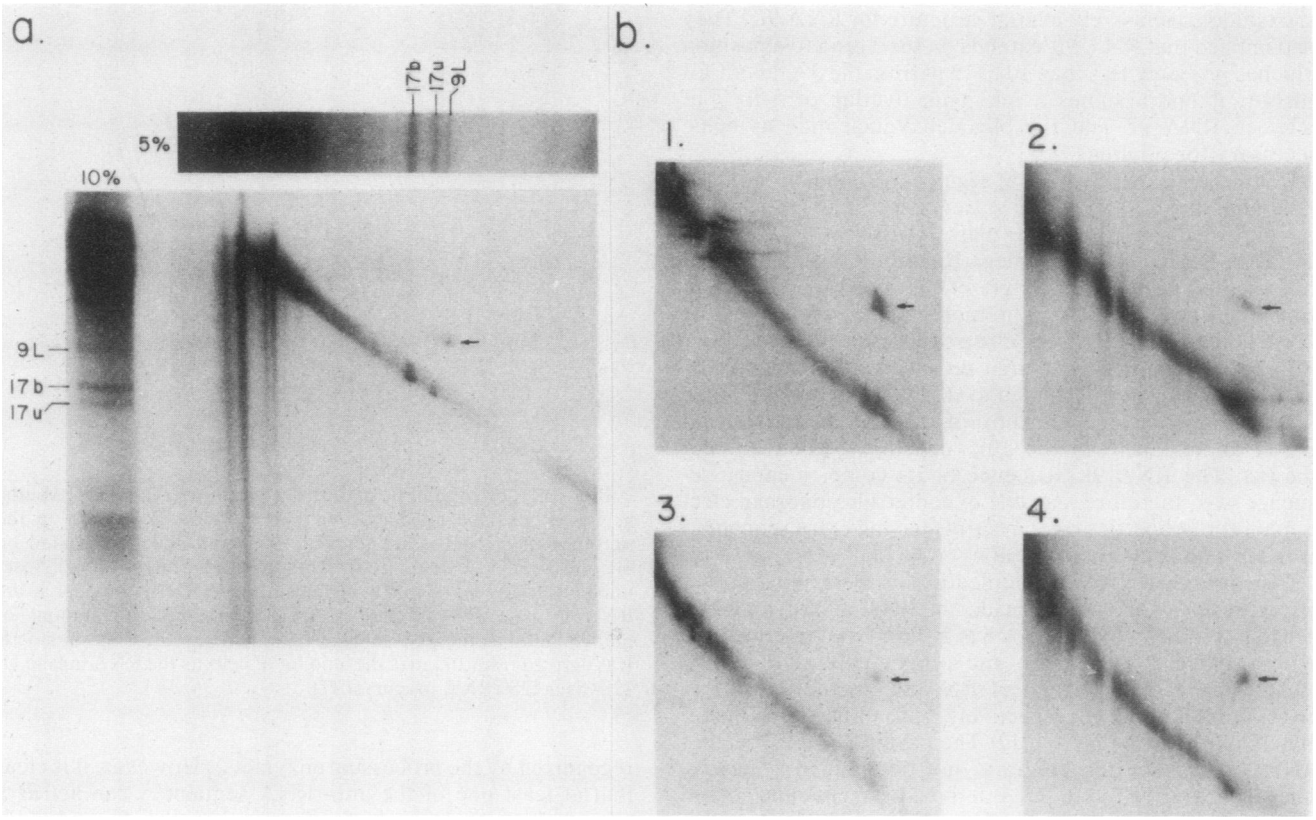

FIG. 3. Bidimensional polyacrylamide gel electrophoresis of mitochondrial RNA labeled in isolated organelles or in vivo. (a) Two samples of total RNA labeled in isolated mitochondria were run in a 5% polyacrylamide–7 M urea gel. A strip containing one of the lanes was placed over a 10% polyacrylamide–7 M urea gel and electrophoresed in the second dimension; a sample of the same RNA preparation was run in parallel. In the autoradiogram, the second lane of the first electrophoresis is shown at the top. (b) Samples of oligo(dT)-cellulose-bound (panels 1 and 2) or -unbound (panels 3 and 4) RNA labeled in vivo (panels 1 and 3) or in isolated organelles (panels 2 and 4) were subjected to bidimensional electrophoresis as described for panel a. The sample in panel 1 contained RNA derived from 10 times the amount of cells as the sample in panel 3, while the samples in panels 2 and 4 contained equivalent amounts of RNA. The arrows point to RNA 9L.

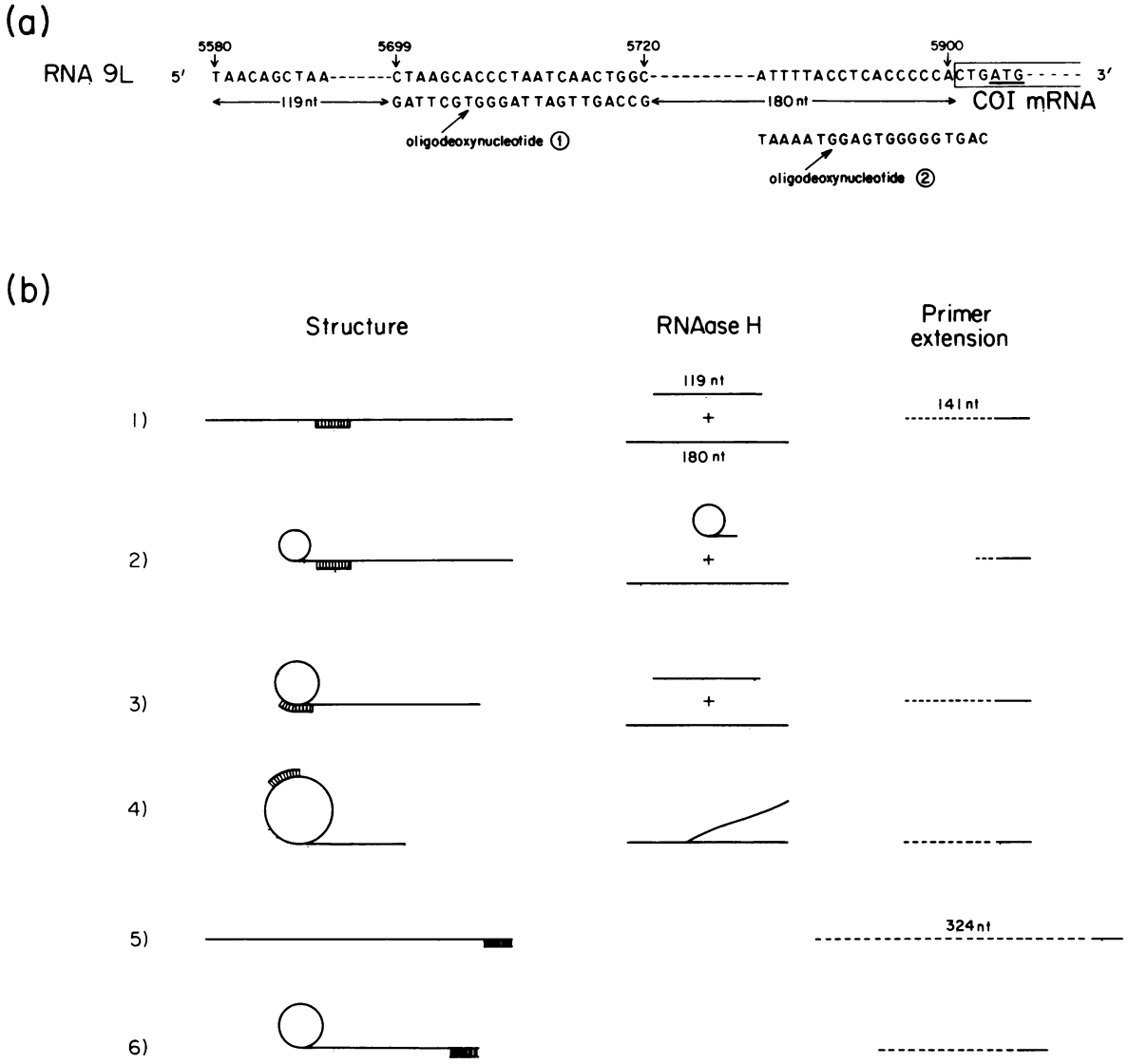

FIG. 4. Scheme of the rationale of the experiments carried out to determine the structure of RNA 9L. (a) Partial sequence of the coding DNA sequence (sense strand) for RNA 9L and sequences of synthetic oligodeoxynucleotides number 1 and number 2 used in the RNase H and primer extension experiments. (b) Diagrammatic representation of the possible structures of RNA 9L and of the corresponding predicted results in the RNase H and primer extension experiments. See the text for details.

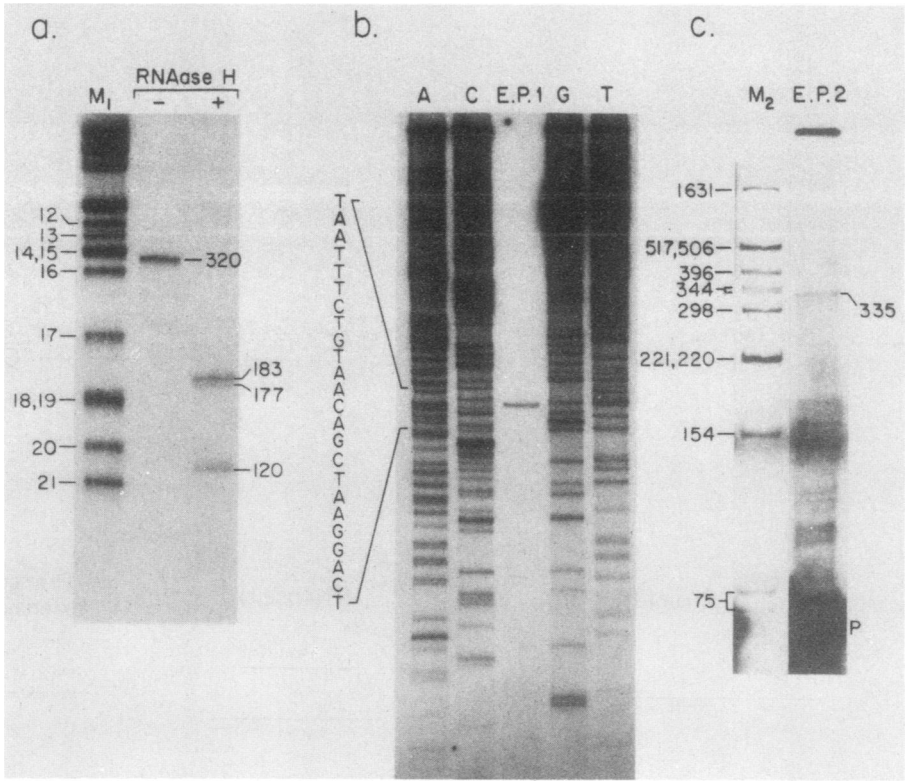

FIG. 5. Results of the RNase H and primer extension experiments. (a) RNA 9L labeled in isolated organelles and purified by electrophoresis in a 5% and then in a 10% polyacrylamide–7 M urea gel was hybridized with synthetic oligodeoxynucleotide number 1 (Fig. 4a), digested (+) or not digested (−) with RNase H, and electrophoresed through a 5% polyacrylamide–7 M urea gel. M_1, 3′-end-labeled *Hpa*II digest of HeLa cell mtDNA. (b) Unlabeled RNA 9L purified as described above and hybridized with 5′-end-labeled oligodeoxynucleotide number 1 (Fig. 4a) was used as a template for primer extension with avian myeloblastosis virus reverse transcriptase and unlabeled deoxynucleoside triphosphates. The size of the product obtained (E.P.1 [extended primer 1]) is compared with the sequence ladder produced by primer extension reactions carried out with the same unlabeled oligodeoxynucleotide as a primer, a single-stranded M13 clone (mp8.HE2) containing the L strand of the mtDNA segment between positions 5274 and 6203 as a template, the Klenow fragment of *Escherichia coli* DNA polymerase I, [α-^{32}P]dCTP, and the three other unlabeled deoxynucleoside triphosphates in the presence of one of the four dideoxynucleotides. A, C, G, and T, Reactions carried out in the presence of the dideoxy derivatives of T, G, C, and A, respectively. The bases indicated at the tops of the lanes and the nucleotide sequence pertain to the sense strand. (c) Unlabeled RNA 9L hybridized with 5′-end-labeled oligodeoxynucleotide number 2 (Fig. 4a) was used as a template for primer extension with avian myeloblastosis virus reverse transcriptase and unlabeled deoxynucleoside triphosphates. E.P.2, Extended primer 2; M_2, 3′-end-labeled *Hin*fI digest of pBR322 DNA; P, primer.

is the first time that an unusual electrophoretic behavior in denaturing gels of a linear RNA molecule due to secondary structure has been reported.

On the basis of (i) the extent of in vivo labeling of the polyadenylated form of this RNA relative to polyadenylated RNA 17 after a 4-h exposure of the cells to $^{32}P_i$, (ii) the steady-state amount of RNA 17 (3), and (iii) the similarity of the in vivo rates of synthesis of RNA 6 and RNA 17 (3), one can estimate a steady-state amount of polyadenylated RNA 9L of 100 to 200 molecules per cell. The majority of RNA 9L is in the polyadenylated form in vivo, an observation which may reflect a need for stabilization of this RNA against degradation. However, one cannot exclude the fact that polyadenylation is a required step associated with the processing of the total H-strand polycistronic transcripts (12). On the basis of the rate of synthesis and the estimated steady-state amount of RNA 9L, one can calculate a half-life for this RNA species of 15 to 30 min, shorter than that observed for the mRNAs (averaging about 60 min) but considerably longer than that of the large L-strand transcripts (3, 9). Further work is required to determine whether this RNA species has any role in regulating the initiation of L-strand synthesis, a possibility suggested by its mapping position and its in vivo accumulation in mitochondria.

These investigations were supported by grants GM-11726 and T32GM-07616 from the National Institutes of Health Public Health Service. One of us (C.R.) was a recipient of a Gosney Fellowship and a Fellowship of the Associazione Italiana per la Ricerca sul Cancro.

We thank Michael King for his generous gift of several M13 clones of human mtDNA. Anne Chomyn for the computer analysis, and Jan Hoh for help in the sequencing experiments. The technical assistance of Arger Drew is gratefully acknowledged.

LITERATURE CITED

1. **Amalric, F., C. Merkel, R. Gelfand, and G. Attardi.** 1978. Fractionation of mitochondrial RNA from HeLa cells by high-resolution electrophoresis under strongly denaturing conditions. J. Mol. Biol. **118**:1–25.
2. **Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden, and I. G. Young.** 1981. Sequence and organization of the human mitochondrial genome. Nature (London) **290**:457–465.
3. **Attardi, G.** 1984. RNA synthesis and processing in mito-

chondria, p. 227–290. *In* D. Apirion (ed.), Processing of RNA. CRC Press, Inc., Boca Raton, Fla.
4. **Chomyn, A., M. W. Hunkapiller, and G. Attardi.** 1981. Alignment of the aminoterminal amino acid sequence of human cytochrome c oxidase subunits I and II with the sequence of their putative mRNAs. Nucleic Acids Res. **9**:867–877.
5. **Chomyn, A., P. Mariottini, M. W. J. Cleeter, C. I. Ragan, A. Matsuno-Yagi, Y. Hatefi, R. F. Doolittle, and G. Attardi.** 1985. Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature (London) **314**:592–597.
6. **Domdey, H., B. Apostol, R.-J. Lin, A. Newman, E. Brody, and J. Abelson.** 1984. Lariat structures are *in vivo* intermediates in yeast pre-mRNA splicing. Cell **39**:611–621.
7. **Gaines, G., and G. Attardi.** 1984. Intercalating drugs and low temperatures inhibit synthesis and processing of ribosomal RNA in isolated human mitochondria. J. Mol. Biol. **172**:451–466.
8. **Gaines, G., and G. Attardi.** 1984. Highly efficient RNA-synthesizing system that uses isolated human mitochondria: new initiation events and in vivo-like processing patterns. Mol. Cell. Biol. **4**:1605–1617.
9. **Gelfand, R., and G. Attardi.** 1981. Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable. Mol. Cell. Biol. **1**:497–511.
10. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 122–123. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Montoya, J., T. Christianson, D. Levens, M. Rabinowitz, and G. Attardi.** 1982. Identification of initiation sites for heavy strand and light strand transcription in human mitochondrial DNA. Proc. Natl. Acad. Sci. USA **79**:7195–7199.
12. **Montoya, J., G. Gaines, and G. Attardi.** 1983. The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell **34**:151–159.
13. **Montoya, J., D. Ojala, and G. Attardi.** 1981. Distinctive features of the 5′-terminal sequences of the human mitochondrial mRNAs. Nature (London) **290**:465–470.
14. **Nussinov, R., and A. B. Jacobson.** 1980. Fast algorithm for predicting the secondary structure of single-stranded RNA. Proc. Natl. Acad. Sci. USA **77**:6309–6313.
15. **Ojala, D., C. Merkel, R. Gelfand, and G. Attardi.** 1980. The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell **22**:393–403.
16. **Ojala, D., J. Montoya, and G. Attardi.** 1981. The tRNA punctuation model of RNA processing in human mitochondria. Nature (London) **290**:470–474.
17. **Padgett, R. A., M. M. Konarska, P. J. Grabowski, S. F. Hardy, and P. A. Sharp.** 1984. Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science **205**:898–903.
18. **Ruskin, B., A. R. Krainer, T. Maniatis, and M. R. Green.** 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing *in vitro*. Cell **38**:317–331.
19. **Wong, T. W., and D. A. Clayton.** 1985. In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light strand synthesis. Cell **42**:951–958.