Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 351–355, January 1985
Biochemistry

Priming of human mitochondrial DNA replication occurs at the light-strand promoter

(displacement loop/regulatory sequences/template signals/transcription)

DAVID D. CHANG AND DAVID A. CLAYTON

Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305

Communicated by I. Robert Lehman, September 17, 1984

ABSTRACT **Individual promoters for transcription of each strand of human mtDNA are located near the origin of heavy-strand DNA replication in the displacement-loop region. Initiation of heavy-strand synthesis represents the first event in mtDNA replication. Analyses of the 5′ and 3′ map positions of displacement-loop nucleic acids from mitochondria of cultured human cells reveal a close correspondence between the 3′ ends of RNA, whose 5′ ends map at a unique site, and the 5′ ends of DNA strands. The 5′ ends of the RNA species all map at nucleotide position 407 in the genomic sequence, which corresponds exactly to the major 5′ transcriptional start site, determined previously *in vitro*, that is contained within the light-strand promoter sequence. Displacement-loop heavy-strand DNAs map immediately adjacent to the 3′ termini of these RNAs, and these transition points between RNA and DNA lie within short conserved sequence blocks in the template sequence. The simplest interpretation of these data is that replication is initiated at the major transcriptional promoter with subsequent precise cleavage of primary transcripts to provide the appropriate primer species.**

Mammalian mtDNAs have two separate and distinct origins of replication (1). The origin of heavy-strand (H-strand) synthesis is located within the displacement-loop (D-loop) region of the genome, and the origin of light-strand (L-strand) synthesis is within a cluster of five tRNA genes two-thirds of the way around the circular genome. Replication begins by the initiation of H-strand synthesis on the D-loop template sequence. The initiation of L-strand synthesis occurs only after H-strand synthesis is at least two-thirds complete. Thus the primary event in mammalian mtDNA replication is initiation of H-strand synthesis.

The D-loop region of human mtDNA contains, in addition to the origin of H-strand replication, two transcriptional promoters (HSP and LSP), one for each strand of mtDNA (2). These promoters direct synthesis of polycistronic transcripts for the expression of most, if not all, genes encoded in each strand. The proximity and upstream location of one of these promoters (LSP) to the origin of H-strand replication raise the possibility that this promoter also functions to generate transcripts involved in priming of DNA synthesis. We present evidence that this single promoter (LSP) is responsible for generating primer RNA in addition to expressing structural genes. Furthermore, we have identified potential control sequences that may function to ensure a proper transition from RNA to DNA synthesis.

MATERIALS AND METHODS

Isolation and Nuclease Limit-Digestion of Mitochondrial Nucleic Acids. Nucleic acids from human KB cell mitochondria isolated by discontinuous sucrose gradient centrifugation were purified by hot-phenol extraction as described (3). DNase I digestion was done as described elsewhere (2). Ten micrograms of mitochondrial nucleic acids was treated with 2 μg of RNase-free DNase I (Miles) in 50 μl of 50 mM Tris Cl (pH 8.0)/10 mM $MgCl_2$/1 mM EDTA at 30°C for 30 min. The reaction was terminated by phenol/chloroform (1:1, vol/vol) extraction and DNase I-resistant nucleic acids were recovered by ethanol precipitation. RNase T_1 digestion was carried out similarly with 100 units of RNase T_1 (Bethesda Research Laboratories), except that the $MgCl_2$ was omitted from the reaction buffer. Alkali treatment was performed with 0.3 M KOH/1 mM EDTA as described (4).

Labeling and Isolation of DNA Restriction Fragments. The nuclease S1-protection probe used in the 5′-end analyses of D-loop H-strand (DH)-DNA and DH-primer RNA was generated from a clone containing a 741-base-pair *Sau*3A1 fragment of human mtDNA cloned in pBR322 (pKB741). After *Sau*3A1 digestion, DNA was end-labeled with T4 polynucleotide kinase (Bethesda Research Laboratories) and [γ-^{32}P]ATP (>5000 Ci/mmol, 1 Ci = 37 GBq; Amersham) and cleaved with *Ava* II, and the 658-base-pair restriction fragment was purified by gel electrophoresis. The 3′-end-labeled S1 probes were generated from a BAL-31 deletion clone (contains human mtDNA nucleotide positions 1–407). After *Bam*HI digestion, the restriction fragments were treated with 5 units of T4 DNA polymerase (P-L Biochemicals), [α-^{32}P]dTTP (>400 Ci/mmol, Amersham), and 1 mM unlabeled dATP in 15 μl of 60 mM Tris Cl (pH 8.0)/6 mM $MgCl_2$/6 mM 2-mercaptoethanol at 11°C for 20 min to uniquely label at nucleotides 406 and 407. The labeled DNA was subsequently cleaved with *Eco*RI and the L strand was isolated by electrophoresis in a 5% polyacrylamide (50:1 acrylamide/*N*,*N*′-methylenebisacrylamide, wt/wt) gel. The second 3′-end-labeled S1 probe was prepared by labeling pKB741 at the *Bal* I site (mtDNA nucleotide position 321) with *Escherichia coli* DNA polymerase I (large fragment) (New England Biolabs) and [α-^{32}P]dGTP (>400 Ci/mmol, Amersham). The DNA was then cleaved with *Eco*RI and labeled L strand was isolated as described above.

Nuclease S1-Protection Analyses. The labeled DNA probe was hybridized to mitochondrial nucleic acids as described in the figure legends. The nuclease S1 treatment was carried out in a 300-μl reaction mixture containing 2000 units of nuclease S1 (Boehringer Mannheim)/0.28 M NaCl/50 mM sodium acetate (pH 4.6)/4.5 mM $ZnSO_4$ at 40°C for 30 min. The reaction was terminated by phenol/chloroform (1:1) extraction and nuclease S1-resistant DNA was recovered by ethanol precipitation and analyzed by electrophoresis under denaturing conditions in 6% or 8% polyacrylamide gels.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: H strand, heavy strand; D loop, displacement loop; L strand, light strand; HSP, heavy-strand promoter; LSP, light-strand promoter; DH-DNA, D-loop heavy-strand DNA; CSB, conserved sequence block(s).

RESULTS

Precise Identification of 5′ Termini of Nucleic Acids Around the Origin of H-Strand Replication. Mammalian mtDNA exhibits a complex, species-specific set of D-loop strands (DH-DNA) (5, 6). For human mtDNA, at least four major DH-DNA species with different 5′ termini have been identified by direct end-labeling analyses of isolated DH-DNAs; these 5′ termini are the same as those of nascent H strands (7, 8). We extended this analysis to include all nucleic acids, including DH-DNAs, around the origin of H-strand replication by using nuclease S1 protection to determine accurately their 5′ end positions (Fig. 1). Total mitochondrial nucleic acid isolates were hybridized to a *Sau*3A1/*Ava* II restriction fragment (nucleotides 1–658) whose L strand was 5′-end-labeled at the *Sau*3A1 site (Fig. 2). The position of label in the probe is at least 110 nucleotides downstream from any known DH-DNA 5′ terminus. Therefore, this probe will hybridize to all DH-DNAs and transcripts initiating from an upstream promoter. Upon nuclease S1 treatment of such hybrids, a family of discrete fragments will appear whose sizes accurately reflect the 5′ termini of mitochondrial nucleic acids.

The resulting nuclease S1-resistant fragments were analyzed on a denaturing 6% polyacrylamide gel (Fig. 2). The nature of the hybridized mitochondrial nucleic acid was determined by treating the nucleic acid with DNase I, RNase T_1, or alkali prior to hybridization. With untreated nucleic acid (lane 2), nuclease S1-resistant species are seen whose 5′ ends map to nucleotide positions around 110, 147, 168, and 190. These nucleotide positions are identical to the 5′ ends of previously reported DH-DNAs (5, 7). They are DNase I-sensitive (lane 3), but remain unchanged after either RNase T_1 or alkali pretreatment (lanes 4 and 5), indicating the absence of any 5′ ribonucleotides. In addition, three minor species that are sensitive to DNase I but resistant to RNase T_1 or alkali pretreatment are visible in Fig. 2. The 5′ termini of these minor DH-DNAs map around nucleotides 220, 310, and 440, upstream of previously reported DH-DNA 5′ ends.

The 5′ terminus of the only major RNA detected in this experiment maps at nucleotide 407 (Fig. 2, lanes 2 and 3) where L-strand transcription initiates (2, 10). Since the probe DNA for S1 protection is labeled at the *Sau*3A1 site, the 3′ end of this RNA must extend to at least this position, beyond the 5′ termini of all DH-DNA (Fig. 1). The exact 3′ extent of this species was not determined in this study. However, its 5′-end map position and minimum 3′ extension beyond the origin of H-strand replication indicate that it represents unprocessed primary transcripts from the LSP. Also seen in Fig. 2 are minor RNA species whose 5′ termini are located between nucleotides 200 and 300 (lanes 2 and 3). Since no transcriptional initiation site exists in this region, it is likely they are processed intermediates of primary transcripts.

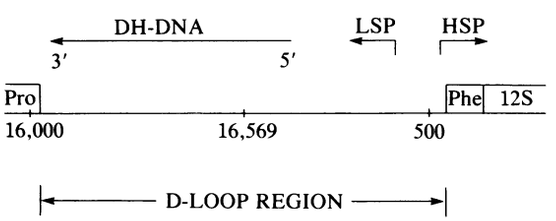

FIG. 1. Organization of the human mtDNA D-loop region. The genetic organization of the D loop, including the surrounding structural genes for $tRNA^{Pro}$ (Pro), $tRNA^{Phe}$ (Phe), and 12S rRNA (12S), is shown. Also indicated are the 5′- and 3′-terminal positions of the major DH-DNA species and the location and transcriptional polarities of the LSP and the HSP. The nucleotide numbers refer to the genomic sequence (9).

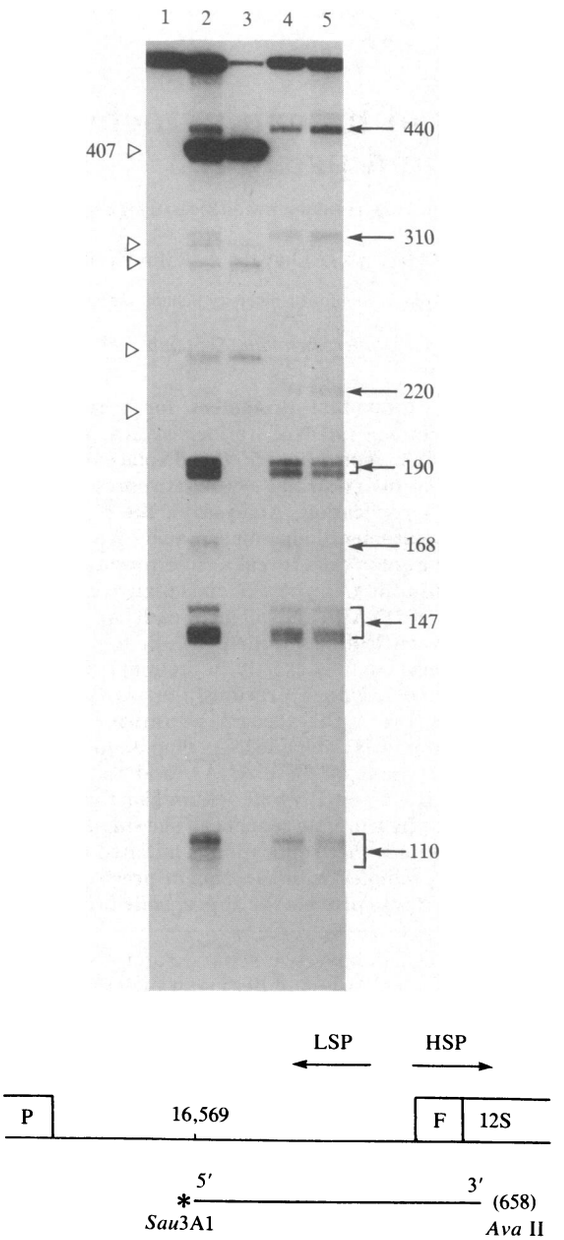

FIG. 2. Location of the 5′ termini of DH-DNA and H-strand transcripts by nuclease S1-protection analyses. Ten micrograms of nucleic acid from sucrose step-gradient purified mitochondria was nuclease limit-digested and hybridized to a *Sau*3A1/*Ava* II fragment 5′-end-labeled at the *Sau*3A1 site. The hybridization was carried out in 80% (vol/vol) formamide/0.4 M NaCl/50 mM Hepes (pH 6.4)/1 mM EDTA at 40°C for 4 hr. The hybridized nucleic acids were treated with nuclease S1 and analyzed by denaturing 6% polyacrylamide gel electrophoresis. The nucleotide positions of the 5′ termini of DH-DNA (arrows) are shown on the right. The 5′ map position of the major H-strand transcript (nucleotide 407) from the LSP and minor H-strand RNA species are indicated on the left (open arrowheads). Lane 1, probe DNA only; lane 2, untreated mitochondrial nucleic acids; lane 3, DNase I-treated sample; lane 4, RNase T_1-treated sample; lane 5, alkali-treated sample. The location of a *Sau* 3A1/*Ava* II restriction fragment used as the S1-protection probe relative to the D loop and the position of the label on the probe (asterisk) are shown below (compare with Fig. 1). Single-letter amino acid symbols are used to denote the genes for $tRNA^{Pro}$ (P) and $tRNA^{Phe}$ (F).

Characterization of H-Strand Transcripts Mapping Immediately Upstream of the Origin of H-Strand Replication. The above experiment identified only the 5′ termini of DNA or

RNA that extend at least to the *Sau*3A1 site and, therefore, failed to provide any information regarding transcripts that are processed or terminated upstream of this site. To characterize such transcripts, their 3′ ends were positioned by nuclease S1-protection analyses. The probe DNA was generated from a BAL-31 deletion clone that contains mtDNA sequence positions 1–407 at the *Eco*RI/*Bam*HI sites of pBR322. After *Bam*HI digestion, the DNA was 3′-end-labeled at the LSP initiation site (nucleotides 406 and 407) with T4 DNA polymerase and subsequently cleaved with *Eco*RI. Thus, this probe will hybridize to all transcripts initiating from the LSP. Since the probe is 3′-end-labeled, the size of a resulting nuclease S1-resistant species will correspond to the 3′-end map position of a given nucleic acid.

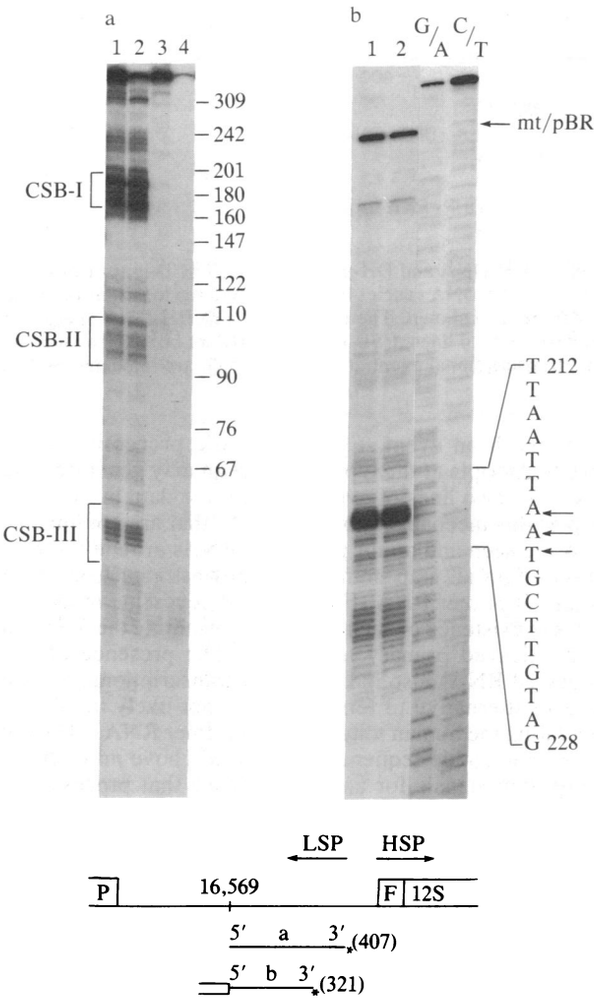

FIG. 3. Location of 3′ termini of DH-primer RNAs. The 3′ map position of primer RNAs are identified by S1 protection. The probe DNAs (shown schematically at the bottom) were generated by 3′-end-labeling a BAL-31 deletion clone containing human mtDNA sequence positions 1–407 with T4 DNA polymerase and [α-^{32}P]dTTP at nucleotides 406 and 407 (probe a) and by 3′-end-labeling pKB741 at the *Bal* I site at nucleotide 321 with *E. coli* DNA polymerase I (large fragment) and [α-^{32}P]dGTP (probe b). Labeled DNA probes were strand-separated and L strands were purified by gel electrophoresis and hybridized at 65°C for 12 hr in 0.3 M NaCl/40 mM Tris Cl (pH 7.8)/2 mM EDTA to 10 μg of untreated mitochondrial nucleic acids (lanes 1), 10 μg of DNase I-treated mtRNA (lanes 2), 10 μg of RNase T_1-treated mitochondrial nucleic acids (lane 3), or 5 μg of *E. coli* tRNA (lane 4). (*a*) Hybridization with probe a. The size markers based on *Hpa* II digests of pBR322 are given at the right and the positions of conserved sequence blocks (CSB)-I, -II, and -III are shown at the left. (*b*) Hybridization with probe b (lanes 1 and 2). The 3′ map positions of major DH-primer RNAs at CSB-I are identified to the nucleotide position by comparing the size of nuclease S1-resistant fragments against the sequence ladders generated from partial chemical cleavage of the labeled probe b DNA (lanes G/A and C/T). The L-strand nucleotide sequence (positions 212–228) around the 3′ termini of DH-primer RNAs (arrows) is shown on the right. mt/pBR refers to the junction between cloned mtDNA and vector pBR322 on the probe DNA.

The results of the experiment described above are shown in Fig. 3*a*. The pattern and intensity of protected fragments in the untreated sample (lane 1) are identical to those in the DNase I-pretreated sample (lane 2), suggesting the absence of any significant internal deoxyribonucleotide moieties. Upon RNase T_1 pretreatment, none of the protected bands are seen except the full-length species protected by mtDNA (lane 3). Furthermore, the 5′-end analyses (Fig. 2) demonstrate that most RNAs in this region have a common terminus mapping at the LSP initiation site. Therefore, all protected species visible in Fig. 3*a* must represent transcripts initiating from the LSP. The 3′ termini of these RNAs are heterogeneous; however, they can be grouped into three families mapping approximately at nucleotides 220, 310, and 360 (they correspond to the nuclease S1-resistant DNA fragments of 190, 100 and 50 nucleotides, respectively, in Fig. 3*a*).

The exact 3′-end location of the most abundant family (around nucleotide 220) was mapped to the nucleotide level by sizing the protected fragments against sequencing ladders generated by partial degradation of the labeled DNA used as the S1-protection probe. The probe was 3′-end-labeled at the *Bal* I site (nucleotide 321) with *E. coli* DNA polymerase (large fragment) and subsequently digested with *Hha* I, which cleaves within the pBR322 vector. The results, shown in Fig. 3*b*, indicate two major 3′-end positions. One is a heterogeneous zone centered at nucleotide 217. The distribution of species in this group is similar to that of the most abundant family seen in Fig. 3*a*. The second species maps at nucleotide 1, the junction between the cloned mtDNA and pBR322 sequence. This RNA represents the primary H-strand transcript whose 3′ end extends beyond the origin of H-strand replication.

The notable aspect of the 3′-end nucleotide positions of these H-strand transcripts is that they align remarkably well with the 5′ ends of DH-DNA species at nucleotides 220 and 310. Although we could not demonstrate covalent linkage between these transcripts and DH-DNAs, such a precise alignment of their termini strongly implicates these transcripts in priming of DNA synthesis.

DISCUSSION

The precise locations of DH-DNAs and upstream H-strand transcripts are summarized in Fig. 4. All *in vivo* RNAs upstream of the origin of H-strand replication have common 5′ ends at nucleotide 407. This position is identical to the initiation site from the LSP and, therefore, these H-strand RNAs must all be products of transcriptional events from the LSP (2, 10). Furthermore, the 3′ termini of two of these transcripts abut the DH-DNA 5′ ends, suggesting that these RNAs are involved in priming of DNA synthesis. The LSP also directs synthesis of transcripts extending beyond the replication origin and, presumably, into the coding region for the expression of L-strand genes (11, 12). Since there is no other promoter downstream from the LSP, we conclude that a single promoter generates transcripts used in replication priming and for expressing structural genes.

The precise alignment of putative primer RNA 3′ ends to DH-DNA 5′ ends implies that the transition from RNA synthesis to DNA synthesis occurs at these junctions. The exact processes that lead to the specific transition are not known. However, the 3′ termini of primer RNAs are located within

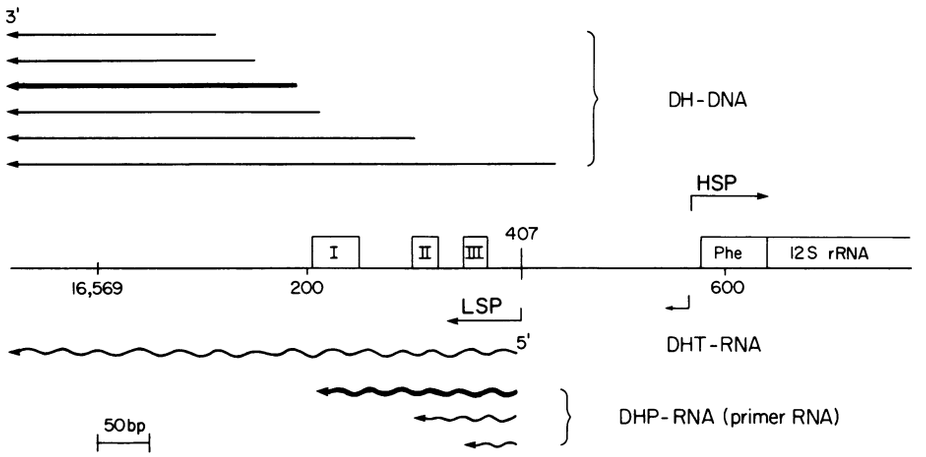

FIG. 4. Schematic diagram of the origin of H-strand replication and locations of DH-DNA and DH-primer RNAs. The D-loop region near the structural genes for $tRNA^{Phe}$ and 12S rRNA is shown. The extent and 5' termini of DH-DNA species (solid lines) and the locations of primer RNA (DHP-RNA) and unprocessed transcript (DHT-RNA) species (wavy lines) are also shown. The most prominent DH-DNA species (5' termini at nucleotide 190) and DHP-RNA (3' termini at nucleotide 217) are indicated by bold lines. CSB-I, -II, and -III are boxed. The location and transcriptional direction from LSP and HSP, as well as the minor H-strand transcriptional event at nucleotide 559, are indicated by bent arrows. The nucleotide numbers refer to the genomic sequence. bp, Base pairs.

three conserved sequence blocks (CSB-I, CSB-II, and CSB-III) that have been recognized in vertebrate mtDNAs (13, 14). Furthermore, corresponding DH-DNA 5' ends that abut the primer RNA 3' ends are found in CSB-I and CSB-II (Fig. 5). Although the LSP for mouse mtDNA has not yet been identified, a correlation between these conserved sequences and the transition to DNA synthesis has been observed (unpublished data). Such an association between conserved nucleotide sequences and switch regions for elongation of primer RNAs strongly implicates CSB elements in ensuring proper transition from RNA to DNA synthesis.

At least three possibilities exist for making certain that a change from RNA to DNA synthesis occurs. First, interaction between the template sequence and mitochondrial polymerases may cause this event. This mechanism demands enzymatic machinery having both RNA and DNA synthesizing capacity. No free primer RNA, such as seen in our study, would be detectable unless primer RNA is precisely removed from DNA at the transition junction. Primer RNA would be expected to exist covalently linked to DNA. Second, transcription may terminate at or near the replication origin. DNA polymerase could then utilize the free 3'-hydroxyl end of RNA for DNA synthesis. The presence of transcripts extending beyond the replication origin argues against such an event. Finally, specific processing of primary transcripts at the replication origin may generate primer RNAs. Two lines of evidence from the data in this study support this mechanism; primer RNAs that are not linked to DNA are seen and minor processed RNAs around the replication origin all map immediately downstream from the 3' termini of primer RNA (Fig. 2). This processing model predicts the existence of such RNAs, representing the 3' portion of the cleaved primary transcripts. The presence of such processed RNAs also indicates that transcriptional attenuation, an alternative to termination, is not likely involved in generating the human mitochondrial primer RNAs. The conserved nucleotide sequences mentioned above may serve as recognition signals for an endonuclease that processes the primary transcript at the replication origin.

The processing model requires only a single transcriptional event to generate both a primer RNA and the precursor RNA for structural genes. The cleaved 5' portion of a primary transcript is used as a primer and the remainder then serves as a precursor for further processing to distinct mRNAs and tRNAs. This raises a question about how mtDNA replication is regulated. Since the rate of transcription from the LSP (and hence the rate of primer RNA formation) exceeds the rate of overall mtDNA replication (1),

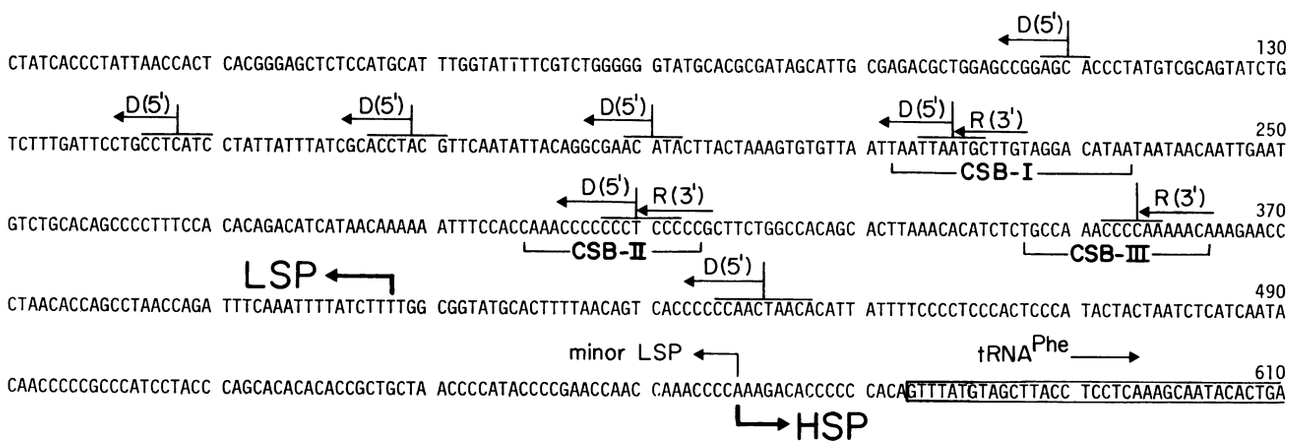

FIG. 5. The nucleotide sequence around the H-strand replication origin. The L-strand sequence of human mtDNA between nucleotides 11 and 610 is shown. D(5') and R(3') denote the 5'-termini of DH-DNA and 3' termini of DHP-RNA, respectively. CSB-I, -II, and -III and the transcriptional initiation sites and directions from HSP and both major and minor LSPs are also shown.

mechanisms must exist that coordinate these two processes. One product of transcription from the LSP is a small polyadenylylated "7S RNA" ≈230 nucleotides long (ref. 15 and unpublished data). The 5′ terminus of 7S RNA is identical to the LSP initiation site. Although the exact 3′ terminus of 7S RNA is not known, the relative size and genomic location of 7S RNA indicate that its 3′ end may coincide with the 3′ end of the most abundant primer RNA at nucleotide 217. The size of this primer RNA, deduced from its 5′ and 3′ termini, is 191 nucleotides. The apparent size difference of 40 nucleotides can be explained by the extent of polyadenylylation of 7S RNA. If the primer RNA is polyadenylylated posttranscriptionally, this process, which reduces the number of potentially functional primer RNAs, may regulate the rate of replication. We note that the formation of DH-DNA itself, which can be regarded as an abortive or attenuated replication event, may also control the rate of overall mtDNA replication.

The above priming model does not account for DH-DNAs whose 5′ termini map between nucleotide 110 and 190 and around 440. We were unable to detect any specific candidate primer RNAs for these species. As for the species between nucleotides 110 and 190, it is possible that they are derived from one or more longer DH-DNAs, either by 5′-exonucleolytic trimming or specific endonucleolytic cleavage. The 5′ end of DH-DNA at nucleotide 440 is located upstream of the LSP initiation site and, therefore, cannot have been primed by the transcripts from the LSP. Although there is a minor LSP around nucleotide 559 observed in *in vitro* transcription assays (2), no specific *in vivo* transcripts map upstream of the DH-DNA at nucleotide 440. Whether priming of this DH-DNA involves a transcriptional event from the minor LSP or a second, as yet unknown, mechanism remains to be determined.

A major conclusion is that human mtDNA H-strand synthesis is primed by transcripts ≈50–190 nucleotides long. The primer RNAs likely are generated by specific cleavage, directed by CSB elements, of primary H-strand transcripts from the LSP. The fact that no covalently linked RNA-DNA species are detected suggests that the cleavage process is extremely efficient. It may thus be feasible to characterize the postulated processing endonuclease from human mitochondria.

This investigation was supported by Grant GM-33088-14 from the National Institutes of Health. D.D.C. is a Medical Scientist Training Program Trainee of the National Institutes of Health (GM-07365-08).

1. Clayton, D. A. (1982) *Cell* **28,** 693–705.
2. Chang, D. D. & Clayton, D. A. (1984) *Cell* **36,** 635–643.
3. Tapper, D. P., Van Etten, R. A. & Clayton, D. A. (1983) *Methods Enzymol.* **97,** 426–434.
4. Gillum, A. M. & Clayton, D. A. (1979) *J. Mol. Biol.* **135,** 353–368.
5. Gillum, A. M. & Clayton, D. A. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 677–681.
6. Brown, W. M., Shine, J. & Goodman, H. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 735–739.
7. Tapper, D. P. & Clayton, D. A. (1981) *J. Biol. Chem.* **256,** 5109–5115.
8. Tapper, D. P. & Clayton, D. A. (1982) *J. Mol. Biol.* **162,** 1–16.
9. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., 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. & Young, I. G. (1981) *Nature (London)* **290,** 457–465.
10. Walberg, M. W. & Clayton, D. A. (1983) *J. Biol. Chem.* **258,** 1268–1275.
11. Montoya, J., Christianson, T., Levens, D., Rabinowitz, M. & Attardi, G. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7195–7199.
12. Clayton, D. A. (1984) *Annu. Rev. Biochem.* **53,** 573–594.
13. Walberg, M. W. & Clayton, D. A. (1981) *Nucleic Acids Res.* **9,** 5411–5421.
14. Wong, J. F. H., Ma, D. P., Wilson, R. K. & Roe, B. A. (1983) *Nucleic Acids Res.* **11,** 4977–4995.
15. Ojala, D., Crews, S., Gelfand, R. & Attardi, G. (1981) *J. Mol. Biol.* **150,** 303–314.