Proc. Natl. Acad. Sci. USA
Vol. 78, No. 10, pp. 6116–6120, October 1981
Biochemistry

Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences

(DNA replication/3'-terminal labeling/nucleotide sequence)

JACKIE N. DODA, CATHARINE T. WRIGHT, AND DAVID A. CLAYTON

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

Communicated by Charles Yanofsky, July 6, 1981

ABSTRACT **Animal mitochondrial DNA (mtDNA) maintains a displacement loop (D loop) at the heavy strand origin of replication. These D loops represent sharply limited synthesis of heavy strands and provide a unique opportunity to examine the termination of DNA synthesis. Direct sizing at the nucleotide level indicates that the 3' ends of D-loop strands of human and mouse mtDNA are discrete and map within three to five nucleotides on the complementary template strand. In the case of human mtDNA, there is a single trinucleotide stop point 51–53 nucleotides downstream from a 15-nucleotide template sequence (3'T-A-A-C-C-C-A-A-A-A-A-T-A-C-A 5') which is repeated four times in the mouse mtDNA D-loop region 3'(T-A-A-Py-Py-A-A-A-T-T-A-C-A 5'). The stop points of the five major mouse D-loop strands are 24–63 nucleotides downstream from the four repeated template sequences. These results suggest that the arrest of D-loop strand elongation is an event determined by template sequence.**

A distinctive feature of animal mitochondrial DNA (mtDNA) is the maintenance of a primer strand at the heavy strand (H strand) origin of replication. This causes displacement of the parental H strand in this region, and the resultant novel triplex structure has been termed a displacement loop (D loop) (1–3). The D-loop H strands have been shown to consist of families of discrete lengths (4, 5). The size variability of human D-loop strands is largely due to differences at the 5' ends of the strands (4, 6), and mouse D-loop strands are more complex because of differences at both the 5' and 3' ends (4, 7) (Fig. 1).

Our knowledge of events associated with initiation and elongation of nascent DNA strands has advanced substantially in recent years, but our understanding of the mechanism of termination of DNA synthesis has remained enigmatic. The mtDNA D loop provides a unique opportunity to examine the termination of DNA synthesis. The data reported here indicate that the 3' ends of D-loop strands of human and mouse mtDNA are discrete and map within clusters of three to five adjacent nucleotides on the complementary template strand. The template strand in both systems contains distinctive sequences that can be correlated with the stop points of D-loop strand synthesis. These results suggest that the arrest of D-loop strand elongation is an event determined by template sequence.

MATERIALS AND METHODS

Growth of Cells and Isolation of mtDNA. Mouse LA9 and human KB cells were grown in suspension culture as described (8, 9). Mitochondria and the closed circular mtDNA population were isolated by the standard procedure (3, 8).

***In Vitro* 3'-End-Labeling of D-Loop Strands.** For the purpose of 3'-end nucleotide analyses, the D-loop strands were 3'-end-labeled with calf thymus deoxynucleotidyltransferase and [α-^{32}P]cordycepin as described by Tu and Cohen (10).

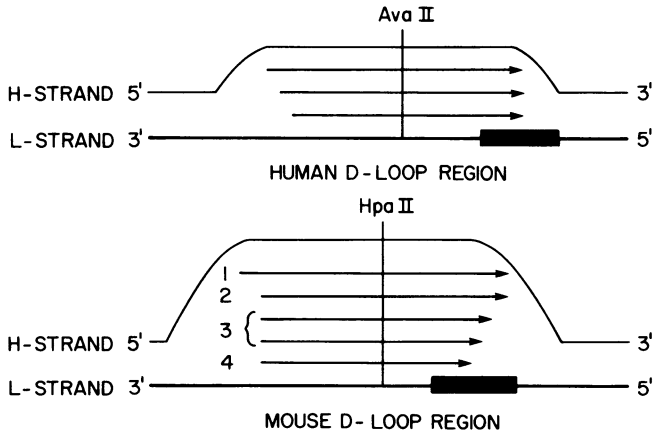

FIG. 1. D-loop region of human and mouse mtDNA. Human mtDNA maintains three major D-loop strands whose 5' ends map at different locations (6) and whose 3' ends map at one location (4). Mouse mtDNA maintains five D-loop strands of which four have the same 5'-end distribution (7). A fifth species is longer at the 5' end by approximately 80 nucleotides (7). Four of the five mouse D-loop strands are different in overall length by an amount consistent with length differences at the 3' end (5). In both human and mouse mtDNA, D-loop strands are of H-strand sequence. The *Ava* II and *Hpa* II restriction sites utilized for structural mapping of the 3' ends are shown. In the case of mouse mtDNA, four fractions from a preparative gel (see Fig. 2) were analyzed and they correspond to the assignment, 1–4 shown here. The solid blocks denote the light-strand (L-strand) template sequence shown in Fig. 6.

Hybridization and Restriction of D-Loop Strands. Isolated D-loop strands were hybridized to template light strands (L strands) in 70% formamide/0.5 M NaCl/40 mM 1,4-piperazinediethanesulfonic acid, pH 7.0/2.5 mM EDTA for 1 hr at 30°C. Human D-loop strands were hybridized to genome length L strands (9) and subsequently cleaved with *Ava* II (Bethesda Research Laboratories, Rockville, MD) in 20 mM Tris·HCl, pH 7.4/30 mM NaCl/10 mM $MgCl_2$ for 1 hr at 37°C. Mouse D-loop strands were hybridized to genome length L strands of mouse mtDNA (7) and subsequently cleaved with *Hpa* II (Bethesda Research Laboratories) in 20 mM Tris·HCl (pH 7.4)/7 mM $MgCl_2$/1 mM dithiothreitol for 1 hr at 37°C.

Nucleotide Sequence Analysis of D-Loop Strands. The chemical procedure of Maxam and Gilbert (11) was utilized for sequence determinations of 3'-end-labeled D-loop strands and 5'-end-labeled restriction fragments used as sizing ladders. 5'-End-labeling procedures were as reported (12).

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.

3′-End-Terminal Nucleotide Analysis. Human D-loop strands were individually digested with micrococcal nuclease (Worthington) in 10 mM NaCl/10 mM Tris·HCl, pH 8.0/10 mM $CaCl_2$ for 1.5 hr at 37°C. The digestion products were spotted on cellulose thin-layer chromatography sheets and chromatographed in 0.1 M sodium phosphate, pH 6.8/$(NH_4)_2SO_4$/*n*-propanol, 100:60:2 (vol/wt/vol). Nonradioactive 3′-monophosphate markers were visualized under a short-wave UV lamp, and the ^{32}P-labeled products were located by autoradiography.

RESULTS

***In Vitro* 3′-End-Labeling and Isolation of D-Loop Strands.** The strategy for 3′-end-labeling of D-loop strands was designed to minimize exposure of D-loop strands to unnecessary preliminary purification steps (5, 7). The total closed circular human or mouse mtDNA population was denatured and the resultant free D-loop strands were 3′-end-labeled with [α-^{32}P]cordycepin triphosphate. Because D-loop strands provide the only discrete 3′ ends in such samples, they can be directly and individually isolated from gels after electrophoresis (Fig. 2).

Structural Mapping of 3′ Ends of Human D-Loop Strands. The three major human D-loop strands shown in Fig. 2 were individually excised from the gel, electroeluted, and hybridized to excess L-strand template. The resultant DNA hybrids were digested with *Ava* II and subjected to electrophoretic sizing in parallel with the DNA sequence of the *Ava* II–*Bam*HI human mtDNA restriction fragment labeled at the same *Ava* II site near the 3′ end of human mtDNA D-loop strands (Fig. 1). This approach permits an exact nucleotide assignment of the 3′ ends of human D-loop strands after correction for the fact that the strands have been lengthened by one nucleotide due to the addition of the [α-^{32}P]cordycepin 3′-end-label. The 3′-end positions of each of the three human D-loop strands were identical and limited to a trinucleotide 3′ G-T-C 5′ in the L-strand template sequence (Fig. 3). Approximately two-thirds of all D-loop strands map at the template G as judged from the relative intensities of the three species present. The fact that the 3′-end-labeling protocol does not introduce 3′-end heterogeneity was demonstrated by 3′-end-labeling a mtDNA restriction fragment of known sequence. This fragment was 3′-end-labeled at a *Pst* I site and yielded the predicted unique DNA sequence (data not shown).

The higher molecular weight species in Fig. 3, lanes A–C, represent uncleaved D-loop strands that presumably were not

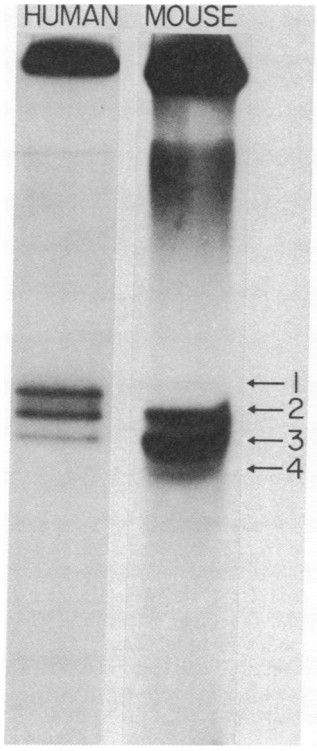

FIG. 2. Preparative 4% polyacrylamide 7 M urea gel of human and mouse D-loop strands 3′-end-labeled with [$\alpha^{32}P$]cordycepin. In the case of human D-loop strands, each of the three major species was individually excised from the gel and isolated by electroelution. Mouse mtDNA D-loop strands are not as well resolved as human and were excised as four individual fractions, as noted in the figure, and isolated by electroelution. Fraction 1 was the least abundant species and did not reproduce well on the print. The radioactive species at the top of each lane are form II circular molecules. The higher molecular weight radioactive species in the mouse sample are elongated daughter strands.

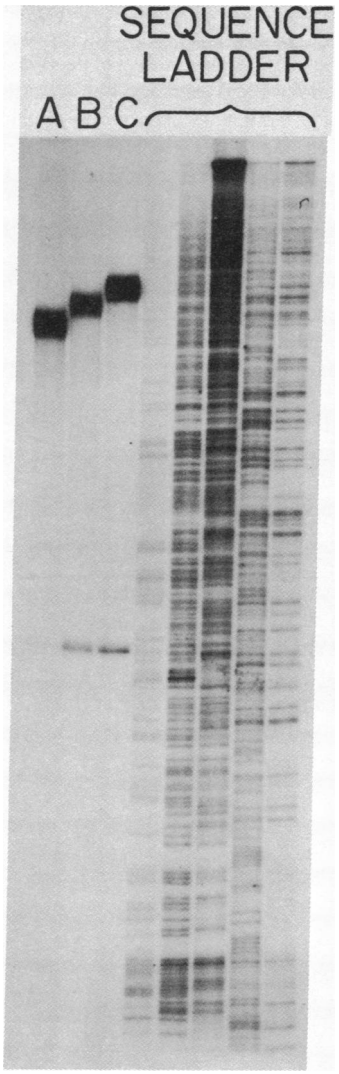

FIG. 3. Structural mapping of the 3′ ends of human D-loop strands. Each of the three human D-loop strands isolated as in Fig. 2 was hybridized to L-strand template and digested with *Ava* II (Fig. 1). The three separate digests were electrophoresed in parallel with a DNA sequence ladder generated by chemical cleavage (11) of a human mtDNA restriction fragment labeled at the 5′-end at the same *Ava* II site. This analysis permitted a direct nucleotide assignment of the 3′-end map positions of D-loop strands by comparison with the corresponding nucleotide in the DNA sequence lanes. The 3′-end map position of D-loop strands is offset by two nucleotides in relation to a correspondingly mobile species in a DNA sequence lane. This is because the 3′ ends have one additional nucleotide due to the 3′-terminal [α-^{32}P]cordycepin, and the actual length of a fragment in a DNA sequence lane is one nucleotide less than a fragment containing the base that was cleaved to generate that particular base-specific signal (11). Lanes A, B, and C: *Ava* II digestion products of each of the three major human D-loop strands in increasing order of overall D-loop strand size. The low-abundance species approximately 12 nucleotides longer than the major digestion products has not been further characterized. The right five lanes show the parallel DNA sequence ladder: G, G+A, A>C, C+T, and C, shown left to right.

hybridized. In each case, only one D-loop strand was present, indicating that the preparative separation procedure was effective (Fig. 2). The size heterogeneity of each is due to the trinucleotide variability at the 3′ end and heterogeneity at the 5′ end (4–6). Thus, the overall size distribution of each full-length D-loop strand is self-consistent.

3′-End Nucleotide Analysis of Human D-Loop Strands. Based on the 3′-end map position determined by the direct sizing shown in Fig. 3, one would predict that the major 3′-end nucleotides of each human D-loop strand would be C > A,G. Each of the three major human D-loop strands 3′-end-labeled with [α-^{32}P]cordycepin and isolated as in Fig. 2 were subjected to digestion with micrococcal nuclease, which generates 3′-phosphate mononucleotides. Thus, D-loop strands 3′-end-labeled with [α-^{32}P]cordycepin will have ^{32}P in the *in vivo* 3′-terminal nucleotide upon digestion with micrococcal nuclease. The digestion products were chromatographed, and the identified major nucleotides were dC and dA with dC > dA (Fig. 4). There were minor amounts of dG and T, particularly in the shortest human D-loop strand. This indicates that the major 3′-end nucleotide is dC, a result consistent with the mapping data (Fig. 3). These data suggest that the preferred stop points for the two largest D-loop strands are at the template dinucleotide 3′ G-T 5′ whereas the smallest D-loop strand exhibits a relatively more frequent stop at C in the template trinucleotide 3′G-T-C 5′. The significance of a minor amount of T at the 3′ end of the smallest D-loop strand is not clear. However, we note that the next 5′ nucleotide in the template is A.

The argument that the 3′ termini of human D-loop strands are mainly limited to the two assigned nucleotides dC and dA is strengthened by data obtained from chemical sequence analysis of 3′-end-labeled D-loop strands. The three individual D-loop strands 3′-end-labeled and isolated as shown in Fig. 2 were subjected to direct chemical sequence analysis of each strand as well as of a mixture of all three (data not shown). Although

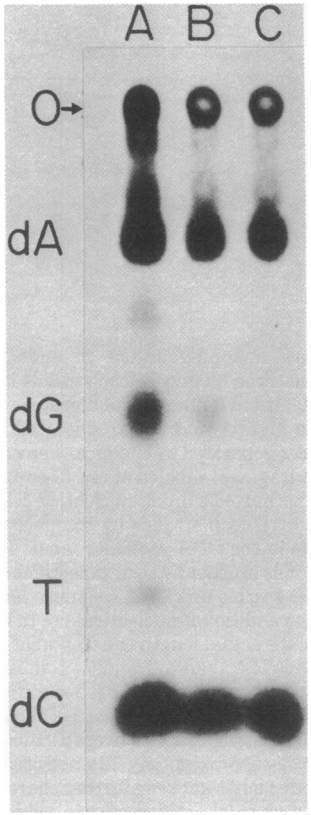

FIG. 4. Chromatograph of human D-loop strand 3′-end nucleotides. Each of the three human D-loop strands 3′-end-labeled with [α-^{32}P]cordycepin was digested to completion with micrococcal nuclease and the digestion products were subjected to chromatography. The nucleotide assignments are shown. Lanes A, B, and C are 3′-end nucleotides of each of the three human D-loop strands in increasing order of overall D-loop strand size. The identity of the minor radioactive species migrating between dA and dG in lane C is unknown, although its mobility is similar to that predicted for rA (7). O, origin.

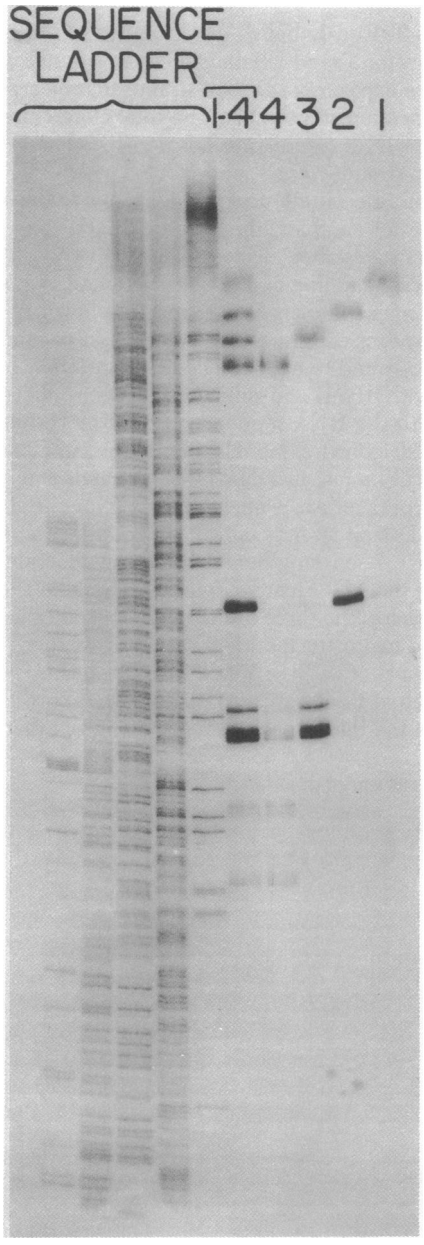

FIG. 5. Structural mapping of the 3′ ends of mouse D-loop strands. Four separate fractions of the mouse D-loop strand population were isolated as in Fig. 2 and hybridized to L-strand template. The four hybrid populations were digested with *Hpa* II (Fig. 1) and electrophoresed in parallel with a DNA sequence ladder generated by chemical cleavage (11) of a mouse mtDNA restriction fragment labeled at the 5′ end at the same *Hpa* II site. The assignment of 3′-end map positions is as noted in the legend to Fig. 3. The left five lanes show the parallel DNA sequence ladder with the individual chemistries in the same left-to-right order as in Fig. 3.

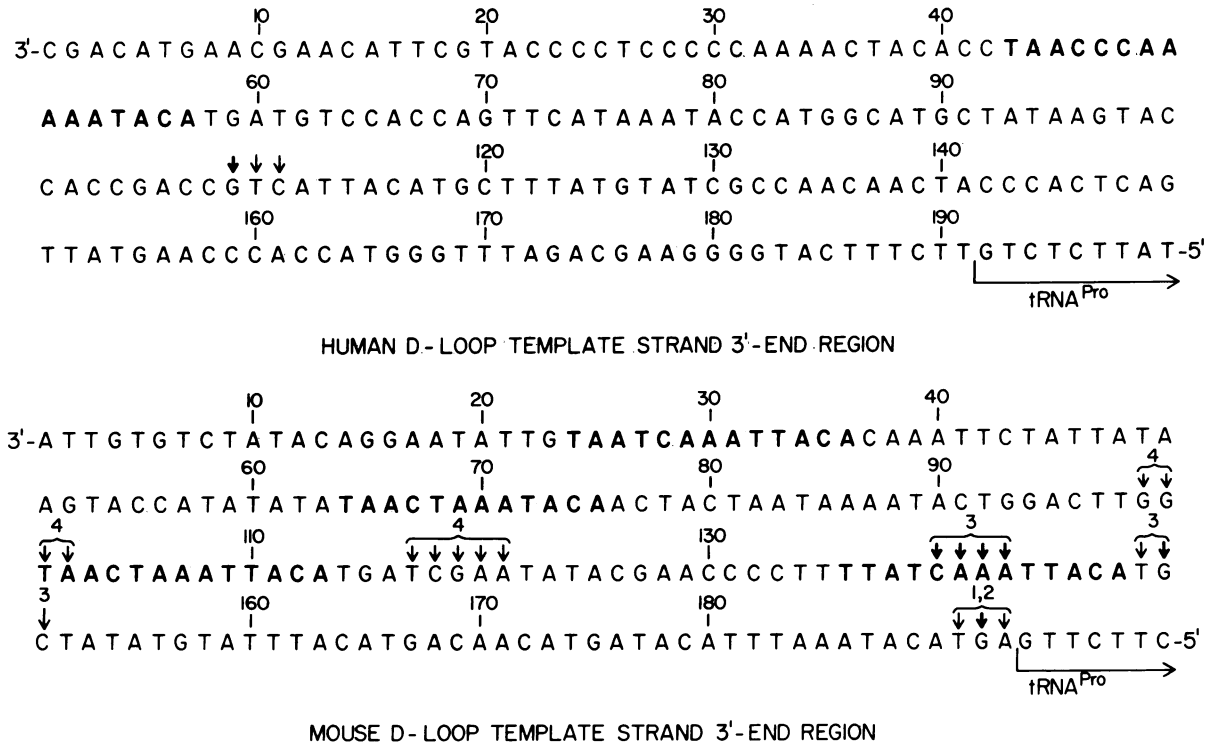

FIG. 6. Nucleotide assignment of the 3' ends of human and mouse D-loop strands. (*Upper*) The 200 nucleotides of the human mtDNA D-loop template region (Fig. 1). This sequence is L-strand template of human KB cell mtDNA and differs at nucleotide 91 (C) from the corresponding nucleotide 16124 (T) in the sequence obtained by Anderson *et al.* (13). The L strand codes for $tRNA^{Pro}$ starting at nucleotide 192. The predominant 3'-end-nucleotide of human D-loop strands is dC which maps at nucleotide 109 (bold arrow). The other 3'-end-nucleotides of human D-loop strands are dA and dG which map at nucleotides 110 and 111 (arrows). The putative template stop signal is shown in bold (nucleotides 43–57). (*Lower*) The 200 nucleotides of the mouse mtDNA D-loop template region (Fig. 1). This sequence is L-strand template and codes for $tRNA^{Pro}$ starting at nucleotide 194. The 3'-end-nucleotide map positions of mouse D-loop strands are denoted by arrows. Bold arrows denote clearly preferred stop points as judged from the original autoradiograms. The number assigned to the arrows represents the particular D-loop strand(s) shown in Fig. 1. The putative template stop signals are shown in bold (nucleotides 24–36, 64–75, 101–113, and 136–148, respectively).

difficult to interpret in their own right, the resultant autoradiograms could be read and were reconcilable with the known template sequence *if* one assumed an approximately equal distribution of two 3' ends mapping at the template 3' G-T 5' position. We therefore conclude that the predominant species of each of the three major human D-loop strands has a 3'-terminal dC which maps at a single G in the template. A less-abundant species has a 3'-terminal dA which maps at a single T in the template immediately 5' to the G. A minor portion of the D-loop strands, particularly in the case of the smallest D-loop strand, has a 3'-terminal dG which maps at a single C in the template immediately 5' to the GT.

Structural Mapping of 3' Ends of Mouse D-Loop Strands. Mouse mtDNA maintains at least five distinguishable D-loop strands (7). These strands differ in length by a variable amount (Fig. 1) and were separated into four fractions by preparative gel electrophoresis (Fig. 2). These fractions were excised from the gel, electroeluted, and hybridized to excess L-strand template. The resultant DNA hybrids were digested with *Hpa* II and subjected to electrophoretic sizing in parallel with the DNA sequence of the *Hpa* II–*Pst* I mouse mtDNA restriction fragment labeled at the same *Hpa* II site present near the 3' ends of mouse D-loop strands. As in the case of human D-loop strands, this approach permits a precise map position assignment of the 3' ends.

As expected, the 3' ends of mouse mtDNA terminated at more template sites than do human D-loop strands (Fig. 5). The two largest D-loop strands (fractions 1 and 2, Fig. 2) terminated at a trinucleotide 3' T-G-A 5' in the template L strand. The third and fourth largest D-loop strands (fraction 3, Fig. 2) terminated at two positions, 3' C-A-A-A 5' and 3' T-G-C 5', on the template. These two sites are separated by five nucleotides in the template strand. Because fraction 3 is a mixture of two D-loop strands, we can only assign these termini to both strands. The smallest D-loop strand (fraction 4, Fig. 2) terminated at two positions, 3' G-G-T-A 5' and 3' T-C-G-A-A 5', separated by 14 nucleotides in the template strand. As before, the higher molecular weight species in Fig. 5 represent uncleaved D-loop strands which demonstrate the effectiveness of the isolation procedure and the additional 5'-end heterogeneity previously reported (5, 7).

DISCUSSION

Stop Points of D-Loop Strand Synthesis. Human mtDNA presents the simplest D-loop anatomy with respect to arrest of strand synthesis. The results here demonstrate that the major stop point of human D-loop synthesis is limited to a single nucleotide position on the template strand (Fig. 6). Each of the three major human D-loop strands exhibits a 3'-end distribution with the predominant species mapping at a single G in the template strand. The remaining population maps at the immediately adjacent 5'-dinucleotide TC in the template trinucleotide 3' G-T-C 5'.

The distribution of 3' ends of mouse mtDNA D-loop strands is more complex. As shown in Fig. 6, there are five stop points for D-loop strand synthesis which vary in breadth from three to five nucleotides on the template strand. The two largest of the five major mouse D-loop strands have a unique trinucleo-

HUMAN D-LOOP TEMPLATE

3'- T A A C C C A A A A A T A C A -5'

MOUSE D-LOOP TEMPLATE

3'- { T A A T C A A A T T A C A
T A A C T A A A T A C A
T A A C T A A A T T A C A
T T A T C A A A T T A C A } -5'

MOUSE CONSENSUS SEQUENCE

3'- T A A PyPyA A A T T A C A -5'

FIG. 7. Human and mouse template stop signals. The template sequences implicated in the termination of D-loop strand synthesis are shown.

tide stop point, one nucleotide of which is immediately adjacent to the first coding nucleotide for $tRNA^{Pro}$. The preferred stop point for these two strands is the 3'-adjacent template nucleotide which is the penultimate noncoding nucleotide prior to the $tRNA^{Pro}$ gene. It is not clear whether the three smaller D-loop strands have one or two stop points each because it is difficult to resolve these D-loop strands in a preparative isolation (Fig. 2). However, the smallest major D-loop strand appears to consist of two 3'-end populations which map at two sites separated by 14 nucleotides.

D-Loop Strands Terminate Near Specific Template Sequences. Previous work has suggested that the arrest of DNA synthesis in mtDNA D loops is not due to thermodynamic considerations (3) or to the lack of a proper available 3'-OH for extension (14). The data here demonstrate that the 3'-end deoxyribonucleotides of human D-loop strands are those predicted by complementarity to their precise template map positions.

The simplest explanation for the arrest of mtDNA D-loop synthesis at discrete points is that elongation stops due to a template-sequence-directed event. The most obvious possibilities are properties of either primary sequence or potential secondary structure at or near the 3' ends of D-loop strands. There is no striking correlation between potential template secondary structure and the 3' ends of D-loop strands common to these two systems (unpublished data).

In support of the hypothesis that primary sequence may effect termination of DNA synthesis is the fact that a highly homologous sequence of 12–15 nucleotides is repeated four times in mouse mtDNA and once in human mtDNA (Fig. 6). In the case of human mtDNA, the 5' end of the 15-nucleotide template sequence 3' T-A-A-C-C-C-A-A-A-A-A-T-A-C-A 5' lies 51–53 nucleotides upstream from the trinucleotide stop point of DNA synthesis. The mouse mtDNA template contains three 13-nucleotide sequences and one 12-nucleotide sequence which are >75% homologous to each other. The 3'-end map positions of mouse D-loop strands are 24–63 nucleotides downstream from the 5' end of the putative corresponding member of this family of sequences. The mouse template consensus sequence is 3' T-A-A-Py-Py-A-A-A-T-T-A-C-A 5' which is 80% homologous to the predicted single stop signal in human mtDNA.

These potential stop signals for DNA replication are aligned in Fig. 7. We have no information as to whether these sequences could interact directly with the enzymatic assembly which synthesizes mtDNA and thereby cause termination by release or arrest of the polymerizing activity. Alternatively, these sequences could be recognized by other molecules, such as RNA or protein, which might block elongation. This seems unlikely in view of the fact that elongation proceeds past these sequences, albeit for short distances. Whatever the exact role of these putative stop signals, the fact remains that, based on Poisson statistics, the probability of this specific 15-nucleotide human mtDNA sequence appearing by chance within the 200 nucleotides shown in Fig. 6 is less than 3×10^{-6}. The probability of four mouse consensus sequences appearing within this region of mouse mtDNA (Fig. 6) is less than 10^{-15}.

The assignment of these template sequences is based on the assumption that the number and position of termination sites should be reflected by the frequency and location of putative stop signals. Other blocks of sequence in these regions show homology to each other. However, in these cases there is no correlation between the frequency of template sequences and the number of termination sites. Comparison of sequences related to those in Fig. 7 shows significantly less homology to the most highly conserved nucleotides within this set. We note that the mouse sequence shown in Fig. 6 is >74% A+T; however, the mouse mtDNA genome is >63% A+T (15). Other regions of the mouse mtDNA genome exhibit increased A+T content, and in the case of the human sequence shown in Fig. 6 the overall base composition is the same as that of the entire genome (13). It therefore is unlikely that base composition in these regions plays a critical role in the termination of D-loop strand synthesis.

Because there is only one stop point for human D-loop strands, the fact that human mtDNA is predicted to, and does, contain only one such sequence argues strongly for a functional role for these sequences. Because this sequence does not appear anywhere else in the human mtDNA genome (13) or outside the D-loop region shown here for the mouse mtDNA genome (15), we believe that it is unlikely that they do not play a significant role in terminating DNA synthesis. If so, this represents an assignment of a short, defined template sequence to the discrete termination of DNA synthesis. Further studies with altered mtDNA template sequences are necessary in order to define precisely the exact nucleotides in these or possibly other sequences that are required for termination of DNA synthesis.

We thank M. W. Walberg for useful discussions of DNA sequence homologies. This research was supported by Grants CA-12312 from the National Cancer Institute and NP-9 from the American Cancer Society, Inc. D.A.C. is a Faculty Research Awardee of the American Cancer Society (FRA-136).

1. Kasamatsu, H., Robberson, D. L. & Vinograd, J. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2252–2257.
2. Berk, A. J. & Clayton, D. A. (1974) *J. Mol. Biol.* **86**, 801–824.
3. Bogenhagen, D. & Clayton, D. A. (1978) *J. Mol. Biol.* **119**, 49–68.
4. Brown, W. M., Shine, J. & Goodman, H. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 735–739.
5. Gillum, A. M. & Clayton, D. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 677–681.
6. Tapper, D. P. & Clayton, D. A. (1981) *J. Biol. Chem.* **256**, 5109–5115.
7. Gillum, A. M. & Clayton, D. A. (1979) *J. Mol. Biol.* **135**, 353–368.
8. Bogenhagen, D. & Clayton, D. A. (1974) *J. Biol. Chem.* **249**, 7991–7995.
9. Battey, J. & Clayton, D. A. (1980) *J. Biol. Chem.* **255**, 11599–11606.
10. Tu, C.-P. D. & Cohen, S. N. (1980) *Gene* **10**, 177–183.
11. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 449–560.
12. Van Etten, R. A., Walberg, M. W. & Clayton, D. A. (1980) *Cell* **22**, 157–170.
13. 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.
14. Eichler, D. C., Wang, T. S.-F., Clayton, D. A. & Korn, D. (1977) *J. Biol. Chem.* **252**, 7888–7893.
15. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. (1981) *Cell* **26**, in press.