Conserved Sequence Box II Directs Transcription Termination and Primer Formation in Mitochondria*

Xuan Hoi Pham, Géraldine Farge, Yonghong Shi, Martina Gaspari, Claes M. Gustafsson, and Maria Falkenberg

From the Department of Laboratory Medicine, Division of Metabolic Diseases, Karolinska Institutet, Novum, SE-141 86 Stockholm, Sweden

The human mitochondrial transcription machinery generates the RNA primers needed for initiation of heavy strand DNA synthesis. Most DNA replication events from the heavy strand origin are prematurely terminated, forming a persistent RNA-DNA hybrid, which remains annealed to the parental DNA strand. This triple-stranded structure is called the D-loop and encompasses the conserved sequence box II, a DNA element required for proper primer formation. We here use a purified recombinant mitochondrial transcription system and demonstrate that conserved sequence box II is a sequence-dependent transcription termination element in vitro. Transcription from the light strand promoter is prematurely terminated at positions 300–282 in the mitochondrial genome, which coincide with the major RNA-DNA transition points in the D-loop of human mitochondria. Based on our findings, we propose a model for primer formation at the origin of heavy strand DNA replication.

The human mitochondrial genome (mtDNA) contains two origins of replication, the origin of heavy strand replication (O_{H}) and the origin of light strand replication (O_{L}) (1). According to the strand-asymmetric model for mtDNA replication, DNA synthesis is continuous on both strands and takes place in a strand-asymmetric mode. The DNA synthesis from O_{H} is unidirectional and proceeds to displace the parental heavy strand. After leading strand synthesis has reached two-thirds of the genome, it activates O_{L}, and DNA synthesis initiates in the opposite direction (2–4). The strand-asymmetric model for mtDNA replication has been challenged by two-dimensional gel electrophoresis analyses demonstrating the presence of conventional duplex mtDNA replication intermediates, which would indicate coupled leading and lagging strand DNA synthesis (5, 6).

Transcription from the mitochondrial light (LSP) and heavy strand promoters generates transcripts of almost genomic length, which are processed to yield the individual mRNA and tRNA molecules (7, 8). LSP-dependent transcription also generates primers required for initiation of mtDNA replication at O_{H} (1, 9), and RNA attached to the newly synthesized H strand has been detected in both mouse and human cells (4, 10). The molecular mechanism governing the switch between genomic length transcription and primer formation is still not completely elucidated. According to one model, the primary LSP transcript is cleaved by an endonuclease activity at certain locations in the O_{H} region and these processed transcripts are used as primers for DNA synthesis. RNase mitochondrial RNA processing (RNase MRP) has been proposed to execute this role in primer formation because the enzyme cleaves LSP transcript model substrates in vitro close to the major initiation sites of leading strand mtDNA synthesis mapped in vivo (11, 12).

Sequence comparisons in vertebrates of mtDNA sequence downstream of the LSP have revealed three conserved sequence blocks (termed CSB I, CSB II, and CSB III). CSB II increases the stability of an RNA-DNA hybrid, and transitions from the RNA primer to the newly synthesized DNA have been mapped to sequences within or near CSB II (4, 13).

Nearly 95% of the H-strand initiation events terminate at sites placed ~600 bp downstream the initiation site (14). The arrested nascent H-strand remains stably hybridized to the circular parental molecule, forming a triple-strand structure characterized by the displaced parental H-strand. This triplex structure is known as the displacement loop (D-loop).

We have previously reconstituted a human mtDNA transcription in a pure in vitro system from promoter-containing DNA fragments, recombinant transcription factor A (TFAM), transcription factor B2 (TFB2M), and RNA polymerase (POLRMT) (15). We now report that in this defined system, a majority of LSP transcripts are prematurely terminated at CSB II, ~100 bp downstream of the promoter. Premature termination is critically dependent on the exact CSB II sequence and is also negatively affected by mutations in the nearby CSB III. Interestingly, the 3'-ends of the prematurely terminated transcripts coincide with the major points of RNA to DNA transitions in the mtDNA control region. Based on our findings we suggest a new, RNase MRP-independent mechanism for primer formation in mitochondrial DNA replication.

**EXPERIMENTAL PROCEDURES**

Transcription in Vitro—Recombinant proteins POLRMT, TFAM, and TFB2M were expressed and purified as described previously (15). We cloned DNA fragments corresponding to bp 1–477 (LSP) and 277–477 of human mtDNA into pUC18.
Using the 1–477 plasmid as template, we employed overlap extension PCR to mutate or delete the CSB I, CSB II, and CSB III elements. All mutant constructs were sequenced before use.

After linearization, we used the plasmid constructs in transcription assays as previously described (15). For the time course experiment in Fig. 1B, 25-μl fractions were taken at the indicated times from a 200-μl reaction. The pulse-chase experiment in Fig. 1C was performed in the same way as the time course experiment, but 4 μl of 100 mM cold UTP was added 5 min after the transcription reaction was initiated and fractions (25 μl) were taken at 0, 5, 10, 20, 30, and 60 min after addition of the cold UTP. For T7 RNA polymerase transcription we used 1× T7 transcription buffer supplied by the manufacturer (New England Biolabs). To introduce a T7 promoter upstream of mtDNA sequences, we subcloned wild-type and mutant mtDNA fragments (bp 1–393) into pBluescript SKII.

To introduce a T7 promoter upstream of mtDNA sequences, we subcloned wild-type and mutant mtDNA fragments (bp 1–393) into pBluescript SKII. Using the 1–477 plasmid as template, we employed overlap primer extension PCR to mutate or delete the CSB I, CSB II, and CSB III elements. All mutant constructs were sequenced before use.

Polymerase reaction buffer supplied by the manufacturer. The mtDNA fragments (bp 1–393) into pBluescript SKII.

Polymerase reaction buffer supplied by the manufacturer. The mtDNA sequences, we subcloned wild-type and mutant mtDNA fragments (bp 1–393) into pBluescript SKII. England Biolabs). To introduce a T7 promoter upstream of

**RESULTS**

POLRMT, TFB2M, and TFAM can initiate transcription in vitro from linear DNA fragments containing mitochondrial promoter sequences (Fig. 1A). Unexpectedly, transcription from LSP generated not only the expected runoff transcript (RO), but also a shorter, prematurely terminated transcript (PT). We used phosphoimaging to calculate the relative molar amounts of the RO and PT transcripts. More than 50% of all transcription events initiated from LSP resulted in the shorter PT product (data not shown). Two lines of evidence suggested that the PT transcripts were derived from true termination of transcription. First, in a time course experiment, the relative ratios of RO and PT transcripts were stable over time, arguing against the possibility that the PT transcripts were due to POLRMT pausing (Fig. 1B). Second, we performed a pulse-chase experiment. After an initial 5-min pulse of radiolabeled UTP, we added an 8000-fold excess of cold UTP and followed the LSP transcription reaction for 60 min. If the PT transcripts were caused by POLRMT pausing, the ratio of PT to RO transcripts would decrease progressively with time, due to resumption of elongation by the stalled enzyme. In fact, the shorter transcript persisted for up to 60 min of incubation without any obvious change of the ratio between terminated and runoff transcripts (Fig. 1C). We conclude that PT transcripts are formed by true transcription termination and not by POLRMT pausing.

The PT products terminated ~100 bp downstream of LSP in a region encompassing the conserved sequence elements CSBI, CSB II, and CSB III (Fig. 2A). To examine whether these elements could influence PT formation, we replaced the individual CSB sequences with the complement sequence and monitored effects on transcription termination in vitro (Fig. 2B). The CSB III mutation caused a mild decrease in transcription termination, whereas the mutant form of CSB II nearly abolished PT formation (Fig. 2B, lanes 3 and 4). Replacement of CSB I did not affect transcriptional termination and did not display any synergistic effects in combination with the mutated forms of CSB II or CSB III (Fig. 2B, lanes 1, 5, and 6). We conclude that PT formation is critically dependent on the CSB II element.
TFAM influences mitochondrial transcription activity, and a high affinity TFAM binding site is situated close to the 3' of our PT transcript (16). We investigated whether TFAM concentrations could influence PT formation in vitro (Fig. 2C). The relative ratio of PT to RO transcript remained largely unchanged over different TFAM concentrations. Replacement of the CSB II sequence with its complement nearly abolished PT formation at all TFAM concentrations. The absence of premature transcription events in turn generated higher levels of the RO transcript.

To exactly define the 3'-ends of the PT transcripts, we separated transcription reactions on an 8% denaturing urea polyacrylamide gel (Fig. 3). Premature transcription termination took place at nucleotide positions 300–282 (Fig. 3, lane 2). Termination therefore occurs immediately downstream of CSB II, which spans positions 315–299 (Fig. 2A). The observed transcription termination was DNA sequence dependent, and a deletion of the 304–290 region abolished transcription termination (Fig. 3, lane 3).

Our data did not rule out the possibility that the PT transcripts were produced from a second initiation site, downstream of LSP. To investigate this possibility, we first monitored in vitro transcription on a series of LSP promoter constructs in which 2 bp at a time were mutated from positions 1 to 20 with the transcription initiation site corresponding to position +1 (17). Mutations had identical effects on both RO and PT products, supporting the notion that both transcripts are initiated at the core LSP promoter (Fig. 4A). Second, we monitored premature transcription termination on a truncated template that had been linearized by restriction enzyme cleavage at position 277, just downstream of the premature transcription termination sites at nucleotide positions 300–282 (Fig. 4B). The truncated template produced a RO transcript of only 131 nt but did not affect the length of the PT products, further supporting the
notion that these transcripts are produced by termination and are not due to the existence of a second transcription initiation site downstream of LSP.

To better define the sequence requirements for transcription termination we made a series of mutant constructs (Fig. 5, A and B). PT formation was completely abolished in the 319–289 mutant, in which CSB II and an additional 12 base pairs downstream of this sequence element had been replaced with the complement sequence (Fig. 5B, lane 4). The 294–290 and 299–295 mutants, overlapping with the sites of transcription termination but positioned outside the conserved CSB II motif, had no distinct effect on PT formation (Fig. 5B, lanes 6 and 7). In contrast, the 304–300 mutant, corresponding to the last part of CSB II, caused a drastic decrease in transcription termination (Fig. 5B, lane 3). Therefore, sequences within the 3’-end of the CSB II element are required for PT formation, whereas sequences immediately downstream of CSB II do not affect this process.

We next investigated whether CSB II-dependent termination was specific to POLRMT or whether the same sequence element could affect the structurally related bacteriophage T7 RNA poly-
We monitored the ability of the T7 enzyme to transcribe the D-loop region and observed a prominent prematurely terminated transcript of the expected size (Fig. 5C, lane 2). On a template lacking the CSB II element (the 319–289 mutant), the prematurely terminated transcript was completely abolished. We also used the 304–300 mutant template and obtained similar results for POLRMT and T7 RNA polymerase (data not shown). We conclude that the primary DNA sequence and probably also structural features of CSB II may stimulate transcription termination by both POLRMT and the T7 RNA polymerases.

The CSB II region has previously been implicated in RNA primer formation for D-loop synthesis (4, 13). To directly compare transcription products formed by termination at CSB II with the start sites for DNA synthesis in D-loop formation, we isolated mitochondria from HeLa cells and thrombocytes. After purification, mtDNA was linearized with SacII treatment. We annealed a radioactively labeled oligonucleotide to nt 75–95 in the D-loop region and used primer elongation to map 5′-ends in the DNA, presumably corresponding to the start sites of DNA synthesis. Thermophilic DNA polymerases can copy RNA templates when they get a jump start on DNA covalently linked to RNA (18). To overcome this problem we performed primer elongation before and after RNase H treatment. E. coli RNase H will specifically degrade the RNA part of the D-loop and thereby allowed us to exactly map the RNA to DNA transition sites. Without RNase H treatment, no significant signal was detected at the CSB II region, but we observed free 5′-ends at positions 151C and 191A, which have been reported previously as sites of initiation for mtDNA synthesis (Fig. 6A, lanes 6 and 8). After RNase H treatment we observed the signal from positions Cys-151 and Ala-191 remained unchanged (Fig. 6A, compare lanes 5 and 6 or lanes 7 and 8), suggesting that no RNA remains associated with mtDNA at these positions. In contrast, RNase H treatment generated a strong signal from the CSB II region (Fig. 6A, compare lanes 5 and 6 or lanes 7 and 8), revealing RNA to DNA transitions at these locations. RNA to DNA transitions took place within a region immediately downstream of CSB II, with a maximum at positions 301–299 and 292–289 of the human mtDNA (Fig. 6B). These transitions therefore coincide with the 300–282 region in which we had observed premature termination of transcription (Fig. 6C).

**DISCUSSION**

RNase MRP can cleave in vitro transcribed L-strand RNA at many of the major initiation sites of DNA replication mapped in vivo (11, 12). The relevance of these findings has been questioned, because the majority of RNase MRP is localized to the nucleolus, where the enzyme plays an important role in rRNA processing, and the amounts of detectable full-length MRP RNA in mitochondria have also been judged too small to attribute a mitochondrial function to RNase MRP (19–21). Finally, there is a lack of genetic data to support a role for RNase MRP in primer formation, because to the best of our knowledge there have been no reports of robust mutants in RNase MRP, which distinguishes effects on rRNA processing in the nucleolus from mtDNA replication.

The data presented here do not allow us to exclude the RNase MRP-dependent model for primer formation but suggest an alternative mechanism. In our experiments, we find that CSB II is a powerful transcription termination element. Mutations in CSB II abolish transcription termination, demonstrating the requirement of specific DNA sequences. We mapped the 3′-ends of the prematurely terminated transcripts and found that transcription terminates just downstream of CSB II between positions 300–282 in the mitochondrial genome. These sites coincide with the major RNA-DNA transition points in the D-loop of human mitochondria. The excellent correspondence between transcription termination and initiation of DNA replication leads us to propose that primer forma-
tion may be directed by sequence-specific DNA elements in human mitochondria, without involvement of RNase MRP.

The human mitochondrial transcription and DNA replication machineries are structurally and functionally related to bacteriophage T7 (22). The mitochondrial RNA polymerase interacts sequence specifically with promoters and displays significant sequence similarity to the monomeric bacteriophage T7 RNA polymerase (23, 24). Sequence similarities also exist between the T7 DNA polymerase and POLγ, both of which are classified as family A DNA polymerases (4, 25). During the initiation of bacteriophage DNA replication, the T7 DNA polymerase displaces the RNA polymerase at a specific region and uses the newly synthesized transcript as a primer for DNA synthesis (26, 27). Therefore, no primer processing is required to initiate T7 bacteriophage DNA replication. It is tempting to speculate that this mechanism has been conserved in mammalian mitochondria. CSB I, located just downstream of CSB II, may function as a loading site for the DNA replication machinery, because this sequence element has been shown to adopt a single-stranded conformation in vivo (28).

We have here focused our attention on Ox4 and the major RNA-DNA transition sites mapped in this region by others (4) and us. There have also been reports of other initiation sites of mtDNA synthesis (9, 29–31). These sites were all mapped by primer extension or two-dimensional gel electrophoresis, and at these mtDNA initiation sites there have been no reports of RNA to DNA transitions. Until RNA to DNA transitions have been demonstrated, we cannot exclude the possibility that these other sites are formed by processing, e.g., removal of the RNA primer together with a section of newly synthesized mtDNA (32–34). An interesting feature of the maturation of T4 lagging strand fragments is that T4 RNase H acts predominantly as an exonuclease, removing the primer and ~30 nucleotides of adjacent DNA before synthesis of the upstream fragment is completed by DNA polymerase (33). It remains to be shown whether hydrolysis of DNA adjacent to RNA primers also is a feature of mtDNA synthesis, but RNase H may play an important role in mammalian mtDNA replication because Rnase1−/− mice display a significant decrease in mitochondrial DNA content, leading to apoptotic cell death (35).

REFERENCES
1. Shadel, G. S., and Clayton, D. A. (1997) Annu. Rev. Biochem. 66, 409–435
2. Robberson, D. L., Kasamatsu, H., and Vinograd, J. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 737–741
3. Tapper, D. P., and Clayton, D. A. (1981) J. Biol. Chem. 256, 5109–5115
4. Kang, D., Miyako, K., Kai, Y., Irie, T., and Takeshige, K. (1997) J. Biol. Chem. 272, 15275–15279
5. Holt, I. J., Lorimer, H. E., and Jacobs, H. T. (2000) Cell 100, 515–524
6. Yang, M. Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E., Jacobs, H. T., and Holt, I. J. (2002) Cell 111, 495–505
7. Clayton, D. A. (1991) Annu. Rev. Cell Biol. 7, 453–478
8. Ojala, D., Montoya, J., and Attardi, G. (1981) Nature 290, 470–474
9. Chang, D. D., and Clayton, D. A. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 351–355
10. Chang, D. D., Hauswirth, W. W., and Clayton, D. A. (1985) EMBO J. 4, 1559–1567
11. Bennett, J. L., and Clayton, D. A. (1990) Mol. Cell. Biol. 10, 2191–2201
12. Lee, D. Y., and Clayton, D. A. (1997) Genes Dev. 11, 582–592
13. Xu, B., and Clayton, D. A. (1995) Mol. Cell. Biol. 15, 580–589
14. Brown, W. M., Shine, J., and Goodman, H. M. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 735–739
15. Falkenberg, M., Gaspari, M., Rantanen, A., Trišunovic, A., Larsson, N. G., and Gustafsson, C. M. (2002) Nat. Genet. 31, 289–294
16. Fisher, R. P., Topper, J. N., and Clayton, D. A. (1987) Cell 50, 247–258
17. Gaspari, M., Falkenberg, M., Larsson, N. G., and Gustafsson, C. M. (2004) EMBO J. 23, 4606–4614
18. Mosig, G., Colowick, N., Gruidl, M. E., Chang, A., and Harvey, A. J. (1995) FEMS Microbiol. Rev. 17, 83–98
19. Kiss, T., Marshallsay, C., and Filipowicz, W. (1992) EMBO J. 11, 3737–3746
20. Kiss, T., and Filipowicz, W. (1992) Cell 70, 11–16
21. Lygerou, Z., Mitchell, P., Peltelski, E., Seraphin, B., and Tollervey, D. (1994) Genes Dev. 8, 1423–1433
22. Korhonen, J. A., Pham, X. H., Pellegrini, M., and Falkenberg, M. (2004) EMBO J. 23, 2423–2429
23. Masters, B. S., Stohl, L. L., and Clayton, D. A. (1987) Cell 51, 89–99
24. Gaspari, M., Larsson, N. G., and Gustafsson, C. M. (2004) Biochim. Biophys. Acta 1659, 148–152
25. Ye, F., Carrodegua, J. A., and Bogenhagen, D. F. (1996) Nucleic Acids Res. 24, 1481–1488
26. Fuller, C. W., and Richardson, C. C. (1985) J. Biol. Chem. 260, 3185–3196
27. Fuller, C. W., and Richardson, C. C. (1985) J. Biol. Chem. 260, 3197–3206
28. Cantatore, P., Daddabbo, L., Fracasso, F., and Gadaleta, M. N. (1995) J. Biol. Chem. 270, 25020–25027
29. Crews, S., Ojala, D., Posakony, J., Nishiguchi, J., and Attardi, G. (1979) Nature 277, 192–198
30. Fish, J., Raule, N., and Attardi, G. (2004) Science 306, 2098–2101
31. Yasukawa, T., Yang, M. Y., Jacobs, H. T., and Holt, I. J. (2005) Mol. Cell 18, 651–662
32. Hollingsworth, H. C., and Nossal, N. G. (1991) J. Biol. Chem. 266, 1888–1897
33. Bhagwat, M., and Nossal, N. G. (2001) J. Biol. Chem. 276, 28516–28524
34. Chapados, B. R., Chai, Q., Hosfield, D. J., Qiu, J., Shen, B., and Tainer, J. A. (2001) J. Mol. Biol. 307, 541–556
35. Cerretelli, S. M., Frolova, E. G., Feng, C., Grinberg, A., Love, P. E., and Crouch, R. J. (2003) Mol. Cell 11, 807–815