MOLECULAR AND CELLULAR BIOLOGY, Mar. 1989, p. 1365–1367
0270-7306/89/031365-03$02.00/0

Vol. 9, No. 3

Conserved Sequence Blocks in Kinetoplast Minicircles from Diverse Species of Trypanosomes

DAN S. RAY

Department of Biology and Molecular Biology Institute, University of California, Los Angeles, California 90024

Received 1 September 1988/Accepted 12 December 1988

Kinetoplast DNA minicircles from various species of trypanosomes are heterogeneous in nucleotide sequence to various degrees but in all instances contain a conserved sequence region of 100 to 200 base pairs present in one, two, or four copies per minicircle. Comparison of the conserved sequence regions of minicircles from eight species of trypanosomes revealed a common sequence motif consisting of three conserved sequence blocks (CSBs) present in the same order and with similar spacing in all species. In addition to the invariant 12-base-pair universal minicircle sequence (CSB-3), a 10-base-pair sequence (CSB-1) and an 8-base-pair sequence (CSB-2) are highly conserved in all minicircles. The overlap of CSB-1 and CSB-3 with previously identified 5′ termini of newly synthesized minicircle H and L strands, respectively, and the presence of this conserved sequence motif in minicircles from diverse species suggest that these CSBs may determine a common mechanism of minicircle replication.

The mitochondrial DNA of trypanosomes consists of two types of circular DNA species, minicircles and maxicircles, catenated into an enormous network (for recent reviews, see references 20, 23, 25, and 26). Individual networks contain thousands of minicircles but only a few maxicircles. Maxicircles have been found to contain mitochondrial genes analogous to those of other eucaryotic cells, whereas the minicircles have no known function. Despite the lack of an established role for minicircles, the abundance of these circular molecules and their replication free of the DNA network (8) has stimulated interest in the study of minicircle replication mechanisms (20, 23).

Although minicircles of most species are heterogeneous in sequence, a common feature of minicircle sequence organization is the presence of a conserved region of about 100 to 200 base pairs in all minicircles of a given species. Thus, although minicircles from *Trypanosoma brucei* and *Leishmania tarentolae* have extensive sequence heterogeneity, in each instance there is a single conserved region and a variable region (10, 11). In *Crithidia fasciculata* and *Trypanosoma lewisi* there are two copies of the conserved sequence present as direct repeats located 180° apart on the minicircle (19, 27). In *Trypanosoma cruzi* there are four copies of the conserved sequence present as direct repeats located 90° apart (7).

The longest sequence conserved among species is a 12-nucleotide sequence, designated the universal minicircle sequence (UMS) (16). In both *Trypanosoma equiperdum* (16, 17) and *C. fasciculata* (3) a small gap of 4 to 8 nucleotides containing ribonucleotides at the 5′ terminus has been found in the light (L) strand within the UMS in newly replicated minicircles. It has therefore been proposed that the UMS serves as the origin for the continuous synthesis of the L strand of the minicircle.

Synthesis of the minicircle heavy (H) strand is discontinuous (2, 12, 13). Minicircles having a nascent H strand are highly gapped and release short H-strand fragments upon denaturation. Gaps in the nascent H strands are filled and covalently joined after separation of the daughter molecules. Specific discontinuities remain in the minicircle H strand of *C. fasciculata* even after all other gaps have been repaired. These remaining H-strand discontinuities are nicks and are located at specific sites about 100 base pairs away from each of the two UMS sequences in the *C. fasciculata* minicircle (2, 3). Unlike the single discontinuity found in the newly synthesized L strand, the specific nicks in the nascent H strand are frequently present in both of the conserved regions of the *Crithidia* minicircle in individual molecules. The 5′ termini at these specific nick sites in the H strand appear to have an unusual structure which might account for their persistence (3). We have proposed a unidirectional D-loop replication model in which these specific nick sites represent specific origins of H-strand synthesis distinct from the initiations of the Okazaki-like fragments (3, 20).

We have compared the organization of the conserved sequence regions from various species of trypanosomes and report here the existence of a common sequence motif in the conserved regions of minicircles from eight species. In addition to the 12-base-pair UMS, two additional conserved sequence blocks (CSB-1 and CSB-2) have been identified, including one (CSB-1) which overlaps the 5′ termini of newly synthesized minicircle H strands in *C. fasciculata* (2).

Conserved sequence motif shared by minicircles from diverse species of trypanosomes. We examined the nucleotide sequences of the conserved regions of minicircles from eight species of trypanosomes. In addition to the conservation of the 12-base-pair UMS in all instances, two additional CSBs are present in the conserved regions from these species (Table 1). CSB-1 is a 10-base-pair sequence in which a core hexanucleotide sequence and residues 1 and 10 are highly conserved in all of the minicircle sequences. The conservation of this particular sequence block is significant in view of the previous identification of a putative H-strand replication origin within this CSB in *C. fasciculata* (2). Specific discontinuities in the newly replicated H strand of *C. fasciculata* minicircles were observed within each of the two copies of CSB-1 (one in each of the two conserved regions). CSB-1 also lies within a 29-base-pair sequence in which 26 of 29 residues are identical between *C. fasciculata* and *L. tarentolae* minicircles (27).

CSB-2 is an octanucleotide sequence which is almost perfectly conserved at six of the eight residues. CSB-2 is located from 27 to 30 base pairs on the 3′ side of CSB-1 in six of the eight trypanosome species. The CSBs of *T. brucei* and

TABLE 1. Minicircle CSBs

Organism or consensus	CSB-1	x[a]	CSB-2	y[b]	CSB-3	CSRs[c] per minicircle	Reference
C. fasciculata Cf-C1[d]	AGGGGCGTTC	28	TCCCGTTC	45	GGGGTTGGTGTA	2	27
L. mexicana amazonensis[e]	. . . . .	30	C	48	. . . . .	1	21
L. tarentolae[f]	. . . . .	27	C	47	. . . . .	1	11
T. brucei[g]	. T . . . . . G .	20	. . . . . G .	41	. . . . .	1	10
Trypanosoma congolense[h]	. A	29	t c . . . . A .	47	. . . . .	1	15
T. cruzi[i]	. . . . .	29	C . . . . . A .	49	. . . . .	4	7
T. equiperdum[j]	. T . . . . . G .	21	. . A . . . G .	38	. . . . .	1	1
T. lewisi[k]	g a . . . . . t c	29	C . . . . . AT	47	. . . . .	2	19
Consensus[l]	AgGGGCGTTC		cCCCGTNC		GGGGTTGGTGTA		

[a] Average distance (in base pairs) between CSB-1 and CSB-2.
[b] Average distance (in base pairs) between CSB-2 and CSB-3.
[c] CSRs, conserved sequence regions.
[d] These sequences are identical in both of the conserved sequence regions on opposite sides of minicircles from the major sequence class.
[e] CSB-1 and CSB-3 are identical in at least five minicircle clones. One T residue is deleted from CSB-2 in one of five minicircle clones.
[f] Consensus sequence based on three minicircle clones.
[g] Consensus sequence based on six minicircle clones.
[h] CSB-1 and CSB-3 are identical in two minicircle clones. The first two residues of CSB-2 are TC in one clone and CG in the other.
[i] Consensus sequence based on five minicircle clones (total of 20 conserved regions).
[j] From direct sequence analysis of the single sequence class of minicircles.
[k] Consensus sequence based on two minicircle clones (total of four conserved regions). The first residue is CSB-1 is G in three of four copies of CSB-1; the second residue is A in two of four copies. The last two residues of CSB-1 are TC in both copies of CSB-1 in one minicircle and CA in both copies in the other.
[l] Capital letters indicate identical residues at the indicated position for at least five of eight species; lowercase letters indicate identity for at least four of eight species.

T. equiperdum have a slightly different spacing between the CSB elements. In these species, CSB-1 and CSB-2 are separated by only 20 or 21 base pairs. The spacing between CSB-2 and CSB-3 is also highly conserved, ranging from 45 to 48 base pairs in six of the species. Again, *T. equiperdum* and *T. brucei* have smaller spacings of 38 and 41 base pairs, respectively, between these CSBs.

Comparison of the conserved sequence regions of minicircles from eight species of trypanosomes shows the presence of two additional CSBs in addition to the 12-base-pair UMS (CSB-3). RNA-primed initiation of the minicircle L strand occurs within CSB-3 in both *C. fasciculata* (3) and *T. equiperdum* (16). In addition, specific nicks in the H strand within CSB-1 in *C. fasciculata* have unusual 5′ termini and possibly represent specific origins of H-strand synthesis. The significance of CSB-2 remains to be determined, but its conservation and positioning between CSB-1 and CSB-3 could reflect a role in minicircle replication as well.

CSBs are common features of viral (6, 9, 14, 18, 22, 24) and mitochondrial (4, 5, 28) origins of replication. The adenovirus replication origin consists of three functionally distinct domains (A, B, and C) which are essential for initiation of replication. Specific sequence blocks within these domains are conserved among all adenovirus serotypes and have been shown to represent binding sites for sequence-specific DNA-binding proteins. The cellular proteins NF-I and NF-III bind to specific sites within domains B and C, respectively, and are both required for efficient initiation at the adenovirus origin (18, 22). In addition to their role in initiation of replication, both proteins have been shown to function as activators of eucaryotic transcription.

Evolutionarily conserved sequence blocks (CSBs I, II, and III) have also been identified in mammalian mitochondrial genomes in the region of the H-strand origin (4, 5). Primer RNA for H-strand synthesis begins at a promoter adjacent to CSB-III and involves endonucleolytic cleavage of the RNA at a unique position between CSB-II and CSB-III. While the precise roles of these CSBs in the mitochondrial initiation mechanism remain to be determined, the fixed distance between the RNA cleavage site and CSB-II in both mouse and human mitochondrial DNAs implicates CSB-II in the recognition process (4).

The apparent utilization of common mechanisms of minicircle DNA replication by various species of trypanosomes, the conservation of the order, spacing, and sequences of the CSBs, and the location of specific 5′ termini of each of the strands of newly synthesized minicircle DNA within specific CSBs strongly suggest that this conserved sequence motif is a determinant of a universal minicircle replication mechanism. The identification of proteins that bind to specific sites within the conserved sequence motif could provide new opportunities for testing this hypothesis.

This work was supported by Public Health Service research grant AI20080 from the National Institutes of Health.

LITERATURE CITED

1. **Barrois, M., G. Riou, and F. Galibert.** 1981. Complete nucleotide sequence of minicircle kinetoplast DNA from *Trypanosoma equiperdum*. Proc. Natl. Acad. Sci. USA **78**:3323–3327.
2. **Birkenmeyer, L., and D. S. Ray.** 1986. Replication of kinetoplast DNA in isolated kinetoplasts from *Crithidia fasciculata*. J. Biol. Chem. **261**:2362–2368.
3. **Birkenmeyer, L., H. Sugisaki, and D. S. Ray.** 1987. Structural characterization of site-specific discontinuities associated with replication origins of minicircle DNA from *Crithidia fasciculata*. J. Biol. Chem. **262**:2384–2392.
4. **Chang, D. D., and D. A. Clayton.** 1987. A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. EMBO J. **6**:409–417.
5. **Chang, D. D., W. H. Hauswirth, and D. A. Clayton.** 1985. Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. EMBO J. **4**:1559–1567.
6. **Deb, S., A. L. DeLucia, C. P. Baur, A. Koff, and P. Tegtmeyer.** 1986. Domain structure of the simian virus 40 core origin of replication. Mol. Cell. Biol. **6**:1663–1670.
7. **Degrave, W., S. P. Fragoso, C. Britto, H. van Heuverswyn, G. Z. Kidane, M. A. B. Cardoso, R. U. Mueller, L. Simpson, and C. M. Morel.** 1988. Peculiar sequence organization of kinetoplast DNA minicircles from *Trypanosoma cruzi*. Mol. Biochem. Parasitol. **27**:63–70.
8. **Englund, P. T.** 1979. Free minicircles of kinetoplast DNA

networks in *Crithidia fasciculata*. J. Biol. Chem. **254**:4895–4900.

9. **Fairman, M. P., and B. Stillman.** 1988. Cellular factors required for multiple stages of SV40 DNA replication *in vitro*. EMBO J. **7**:1211–1218.
10. **Jasmer, D. P., and K. Stuart.** 1986. Conservation of kinetoplastid minicircle characteristics without nucleotide sequence conservation. Mol. Biochem. Parasitol. **18**:257–269.
11. **Kidane, G. Z., D. Hughes, and L. Simpson.** 1984. Sequence heterogeneity and anomalous electrophoretic mobility of kinetoplast minicircle DNA from *Leishmania tarentolae*. Gene **27**: 265–277.
12. **Kitchin, P. A., V. A. Klein, and P. T. Englund.** 1985. Intermediates in the replication of kinetoplast DNA minicircles. J. Biol. Chem. **260**:3844–3851.
13. **Kitchin, P. A., V. A. Klein, B. I. Fein, and P. T. Englund.** 1984. Gapped minicircles. A novel replication intermediate of kinetoplast DNA. J. Biol. Chem. **259**:15532–15539.
14. **Li, J. J., K. W. C. Peden, R. A. F. Dixon, and T. Kelly.** 1986. Functional organization of the simian virus 40 origin of DNA replication. Mol. Cell. Biol. **6**:1117–1128.
15. **Nasir, A., G. A. Cook, and J. E. Donelson.** 1987. Sequences of two kinetoplast minicircle DNA's of *Trypanosoma (Nannomonas) congolense*. Mol. Biochem. Parasitol. **24**:295–300.
16. **Ntambi, J. M., and P. T. Englund.** 1985. A gap at a unique location in newly replicated kinetoplast DNA minicircles from *Trypanosoma equiperdum*. J. Biol. Chem. **260**:5574–5579.
17. **Ntambi, J. M., T. A. Shapiro, K. A. Ryan, and P. T. Englund.** 1986. Ribonucleotides associated with a gap in newly replicated kinetoplast DNA minicircles from *Trypanosoma equiperdum*. J. Biol. Chem. **261**:11890–11895.
18. **O'Neill, E. A., and T. J. Kelly.** 1988. Purification and characterization of nuclear factor III (origin recognition protein C), a sequence-specific DNA binding protein required for efficient initiation of adenovirus DNA replication. J. Biol. Chem. **263**: 931–937.
19. **Ponzi, M., C. Birago, and P. A. Battaglia.** 1984. Two identical symmetrical regions in the minicircle structure of *Trypanosoma lewisi* kinetoplast DNA. Mol. Biochem. Parasitol. **13**:111–119.
20. **Ray, D. S.** 1987. Kinetoplast DNA minicircles: high-copy-number mitochondrial plasmids. Plasmid **17**:177–190.
21. **Rogers, W. O., and D. F. Wirth.** 1988. Generation of sequence diversity in the kinetoplast DNA minicircles of *Leishmania mexicana amasonensis*. Mol. Biochem. Parasitol. **30**:1–8.
22. **Rosenfeld, P. J., E. A. O'Neill, R. J. Wides, and T. J. Kelly.** 1987. Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. Mol. Cell. Biol. **7**:875–886.
23. **Ryan, K. A., T. A. Shapiro, C. A. Rauch, and P. T. Englund.** 1988. The replication of kinetoplast DNA in trypanosomes. Annu. Rev. Microbiol. **42**:339–358.
24. **Sekimizu, K., D. Bramhill, and A. Kornberg.** 1988. Sequential early stages in the *in vitro* initiation of replication at the origin of the *Escherichia coli* chromosome. J. Biol. Chem. **263**:7124–7130.
25. **Simpson, L.** 1986. Kinetoplast DNA in trypanosomatid flagellates. Int. Rev. Cytol. **99**:119–179.
26. **Simpson, L.** 1987. The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. Annu. Rev. Microbiol. **41**:363–382.
27. **Sugisaki, H., and D. S. Ray.** 1987. DNA sequence of *Crithidia fasciculata* kinetoplast minicircles. Mol. Biochem. Parasitol. **23**:253–263.
28. **Walberg, M. W., and D. A. Clayton.** 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. **9**:5411–5421.