Proc. Natl. Acad. Sci. USA
Vol. 77, No. 5, pp. 2445–2449, May 1980
Biochemistry

Sequences of two kinetoplast DNA minicircles of *Trypanosoma brucei*

(recombinant DNA/restriction enzymes/sequence homology/DNA sequencing)

KENNETH K. CHEN AND JOHN E. DONELSON

Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242

Communicated by William Trager, January 18, 1980

ABSTRACT **Kinetoplast DNA of *Trypanosoma brucei* is composed of a network of about 10,000 interlocked minicircle DNA molecules (1.0 kilobase) that are catenated with about 50 maxicircle DNA molecules (23 kilobases). Several different DNA·DNA hybridization techniques using individual minicircle DNA sequences cloned in *Escherichia coli* have indicated that each minicircle molecule contains about one-fourth of its sequence in common with most other minicircles and the remaining three-fourths in common with about 1 out of every 300 minicircles. We have determined the complete sequence of two cloned minicircle DNA molecules that were released from the total kinetoplast DNA network by different restriction enzymes; one minicircle is 1004 base pairs long, the other is 983 base pairs. Both are about 72% dA + dT. They share about 27% of their sequences; the largest continuous region in common is 122 base pairs of near-perfect homology. Twelve other regions of perfect homology equal to or greater than 10 base pairs are also present. Both sequences contain a large number of translation termination codons in all potential translation reading frames. The largest oligopeptide potentially specified by one minicircle sequence is 52 amino acids; the largest by the other minicircle sequence is 71 amino acids. One minicircle contains a decanucleotide sequence that is repeated in tandem five times. It is proposed that massive recombination among the interlocked minicircles in the kinetoplast DNA network may account for much of the homology observed in the two minicircle sequences.**

Trypanosomes are unicellular parasites which include the species that causes sleeping sickness in Africa and Chagas disease in South America (1). Those species that have been investigated have a single large mitochondrion per cell that is characterized by a specialized region called the kinetoplast (2, 3) located near the basal body of the cell's single flagellum. The kinetoplast contains one of the most unusual DNA structures found in nature.

Kinetoplast DNA (kDNA) consists of about 10,000 minicircle molecules and 50 maxicircle molecules, all of which are interlocked together to form a massive disk-shaped network whose width is about the diameter of a minicircle (see refs. 4–6.for recent reviews). Transmission electron micrographs of purified kDNA show a massive tangle of catenated supercoiled molecules (7, 8). Replication of the kDNA network occurs via mechanisms that are only partially understood but appear to involve the release and reattachment of individual DNA circles (6, 9, 10) until a double-sized kDNA network is obtained which separates into two as daughter cells are formed. Some of the common trypanocidal drugs, such as Berenil and ethidium compounds, strongly interact with the kDNA network (11, 12) and probably kill the trypanosome by interfering with the kDNA replication or function.

In *Trypanosoma brucei* the 50 maxicircle molecules in the kDNA network are about 23 kilobases in size and are homogeneous in sequence (13). Several RNA species hybridize to maxicircles (14–16), and it has been proposed that maxicircle DNA corresponds to normal mitochondrial DNA of other lower eukaryotes (17). The *T. brucei* minicircles are about 1 kilobase and heterogeneous in sequence, as determined by restriction enzyme analysis (18) and renaturation kinetics (19). The early renaturation analyses (19) indicated that 100 or more different minicircle sequences may occur among the 10,000 molecules in a network. The biological function of minicircle kDNA is completely unknown.

We have previously reported the isolation of individual *T. brucei* kDNA minicircle molecules in bacteria by recombinant DNA techniques (20). Several of these cloned sequences were characterized by renaturation kinetics, filter hybridizations, and heteroduplex analyses in the electron microscope. These studies indicated that each minicircle molecule has one-fourth of its sequence in common with most other minicircles and the remaining three-fourths in common with about 1 out of 300 minicircle molecules. This suggested two possible models for the sequence organization of the minicircles; one in which small homologies are scattered throughout each minicircle and one in which the homology is confined to a specific 25% of each molecule.

We present here the complete sequence of two cloned kDNA minicircles, one of 1004 base pairs and the other of 983 base pairs. A comparison of the two sequences reveals that both models for sequence organization are partially correct; the two molecules share about 12% homology within a nearly continuous stretch and another 15% that is scattered randomly throughout the rest of the two sequences for a total of 27% homology.

MATERIALS AND METHODS

Recombinant plasmid DNAs were isolated from *Escherichia coli* HB101 cells as described (21). Purified DNAs were digested with various restriction enzymes according to protocols suggested by the commercial suppliers (New England BioLabs, Beverly, MA or Bethesda Research, Rockville, MD). The restriction fragments were resolved on 5% polyacrylamide gels (22) and eluted from the gels.

All of the restriction fragments whose sequences were determined were generated by restriction enzymes that cleave DNA to produce staggered ends with recessed 3′ termini. These 3′ termini were radiolabeled by incubation of the restriction fragment with *E. coli* DNA polymerase I (Boehringer Mannheim) and deoxynucleoside [α-^{32}P]triphosphates (Amersham; specific activity 2000 mCi/μmol; 1 Ci = 3.7×10^{10} becquerels) and the radioactive ends were separated as described (23). The Maxam and Gilbert sequencing method (24)

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.

Abbreviation: kDNA, kinetoplast DNA.

was used, and the chemical cleavage products were resolved by electrophoresis through 0.38-mm-thick gels of 6% (85 cm long) or 20% (40 cm long) acrylamide (25).

Sequences were analyzed by the computer program of Korn *et al.* (26) with the IBM 370/168 computer available at the University of Iowa.

Experiments involving recombinant plasmids containing kDNA inserts were conducted under P2 + EK1 containment as specified by the National Institutes of Health Recombinant DNA Guidelines of December, 1978.

RESULTS AND DISCUSSION

DNA Sequencing Strategy. Preliminary restriction enzyme analyses were performed on six different minicircle kDNA inserts cloned at the *Pst* I site of pBR322 (20) to identify two cloned minicircle kDNA sequences that possessed a distribution of restriction enzyme cleavage sites suitable for total sequence determination by the Maxam and Gilbert procedure (24). *Taq* I was the only enzyme tested that cleaved all six kDNA sequences at least once, although several other enzymes cleaved more than one kDNA. The kDNA inserts of pkT51 and pkT201 were chosen for further characterization. Neither kDNA insert contains a *Hpa* II site. However, *Hpa* II cleaves the vector, pBR322, sequence about 50 base pairs on either side of the *Pst* I site (27, 28) at which the kDNAs were cloned. Therefore, the largest *Hpa* II fragment in both pkT51 and pkT201 contains the complete kDNA insert. About 400 μg of these two *Hpa* II fragments was eluted from preparative 5% acrylamide gels to serve as the starting material for the sequence determinations.

Fig. 1 shows the strategy used to determine the two sequences. The minicircle kDNA sequence in pkT51 was initially released from the kDNA network of *T. brucei* clone 18E2 by digestion with a mixture of *Hha* I, *Hpa* II, and *Alu* I (20). It was cloned at the *Pst* I site of pBR322 via the poly(dG)·poly(dC) tailing method (20), which regenerates *Pst* I sites at the two boundaries between the kDNA insert and the pBR322 sequence. Sequences of 250–350 nucleotides were obtained from the several different labeled termini, as shown in Figs. 1 and 2, to provide the complete continuous sequence shown in Fig. 3*B*. This minicircle is 983 base pairs long and is 71.2% dA + dT. The base composition of the strand shown is 37.4% A, 33.8% T, 20.1% G, and 8.6% C. In recombinant plasmid pKT51 this minicircle sequence is flanked by a 26-base-pair dG·dC boundary on one side and 29 dG·dC base pairs on the other side, as indicated in Fig. 1 but not shown in Fig. 3*B*. Because of these

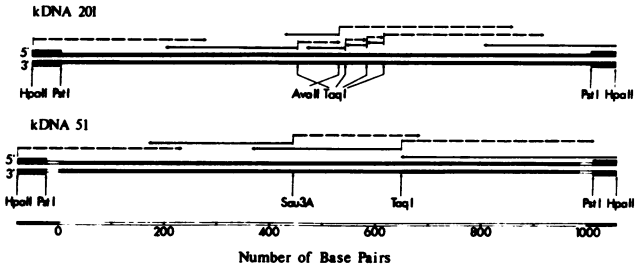

FIG. 1. Diagrams showing the strategy for determining the sequences of kDNA 201 (*Upper*) and kDNA 51 (*Lower*). Only those restriction enzyme sites used in the actual sequence analyses are shown. The indicated *Hpa* II sites are in the pBR322 sequence and the *Pst* I sites are at the boundaries of the pBR322 sequence and the cloned kDNA sequences. The poly(dG)·poly(dC) boundaries flanking the kDNA 51 sequence are indicated by open spaces within the thick lines representing the two strands of kDNA 51. The solid horizontal arrows indicate sequences obtained from the top strands of the two kDNAs; the dashed horizontal arrows indicate sequences obtained from the bottom strands.

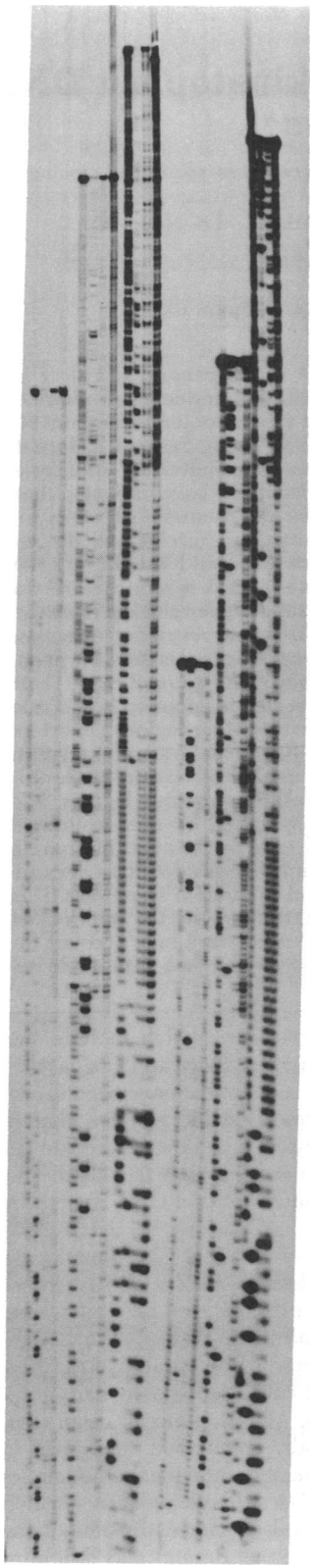

FIG. 2. Example of an autoradiogram of a DNA sequencing gel. The two *Hpa* II/*Taq* I restriction fragments of kDNA 51 labeled at the external *Hpa* II sites (Fig. 1) are shown. The four cleavage reactions of each of the two fragments were applied three different times to the gel at 10-hr intervals. From left to right the order of the four cleavage reactions are A, G, C, and C + T. Approximately the first 75 nucleotides can be determined from the last application, the next 125 from the second application, and the next 125 from the first application. Prominent features of these two sequences are the 26-nucleotide poly(dC) boundary of the left fragment (third column from the left) and the 29-nucleotide poly(dC) boundary of the right fragment (the extreme right column).

poly(dG)·poly(dC) boundaries, it was not possible to unambiguously determine the exact nucleotides at which the kDNA sequence begins and ends in the recombinant plasmid. The nucleotides at positions 1 and 983 in Fig. 3*B* represent a best estimate of these two nucleotides based on the recognition sequences of *Hha* I, *Hpa* II, and *Alu* I. There is also a possibility that the mixture of three enzymes initially used to digest the

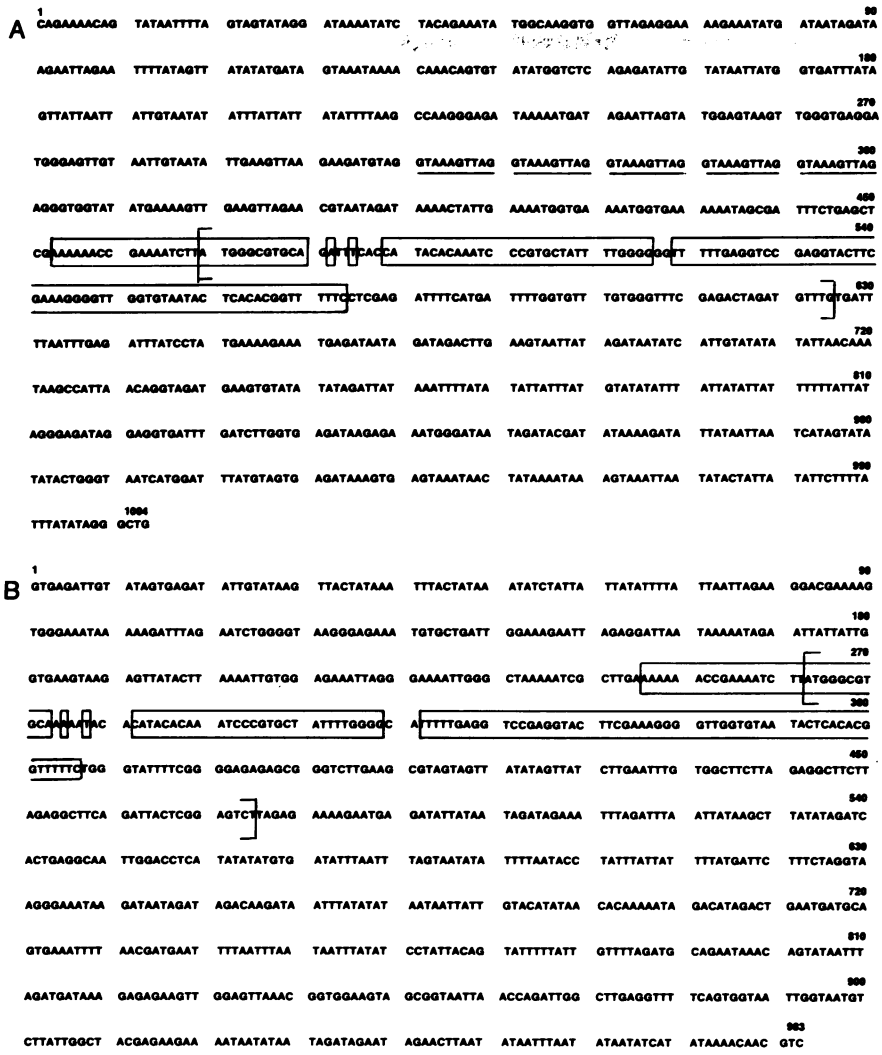

FIG. 3. Nucleotide sequence of (*A*) the upper strand of kDNA 201 and (*B*) the lower strand of kDNA 51 shown in Fig. 1. The boxes indicate a 122-base-pair region of near-perfect homology between the two kDNAs as discussed in the text. The brackets show the largest series of potential translation codons in each kDNA that begins with a methionine codon and ends with a termination codon. This region is 52 codons in kDNA 201 and 71 codons in kDNA 51. The underlining in the kDNA 201 sequence indicates five tandem repeats of a decanucleotide sequence. The numbers indicate the number of nucleotides away from one of the boundaries with pBR322 DNA in the recombinant plasmid.

network kDNA may have excised a small fragment from the original sequence of minicircle kDNA 51 before it was cloned. At present, there is no way to check this possibility.

Minicircle kDNA 201 was initially released from the kDNA network of *T. brucei* clone 18E2 by *Pst* I digestion (20). *Pst* I liberates very few minicircles from the network (20) so that its six-nucleotide recognition sequence, d(C-T-G-C-A-G), occurs only infrequently in the total kDNA. The kDNA 201 sequence was cloned at the *Pst* I site of pBR322 by ligation with T4 DNA ligase (20). Fig. 1 shows how this kDNA sequence was determined and Fig. 3*A* shows its sequence of 1004 base pairs. The base composition of the strand shown is 37.8% A, 34.8% T, 21.2% G, and 6.2% C for an overall dA + dT content of 72.6%. Note that the two strands of both minicircles have an unequal distribution of dG and dC residues.

During the sequencing experiments, it was observed that the actual length of some (but not all) kDNA subfragments turned out to be as much as 30% smaller than the length predicted from their migration on acrylamide gels. This anomaly has also been observed for restriction fragments of *Leishmania* minicircle kDNA (6, 29). The reason for this is not understood but may be related to the high dA + dT content or to an unusual secondary structure of the DNA fragments during electrophoresis as suggested by Englund (6).

Comparison of the Two kDNA Minicircle Sequences. The two complete kDNA sequences were analyzed by a computer program (28) for sequence symmetries, homologies, and potential translation reading frames. The most striking feature to emerge from this sequence analysis is the extent and distribution of the homologous regions between the two molecules (Fig. 4). The two circular molecules are represented by linear lines that are aligned to facilitate a comparison of the locations of the shared sequences. The largest homology is a sequence of 122 base pairs that is identical in the two molecules except for

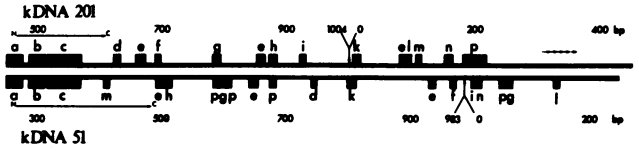

FIG. 4. Diagram showing the relative locations of perfect homologies between kDNAs 51 and 201 that are equal to or greater than 10 base pairs (bp). The two circular kDNA sequences are represented by two lines aligned so that each begins with the first position of the 122-base-pair near-perfect homology discussed in the text. This region is indicated by the lower case letters "a," "b," and "c." The other letters indicate the other smaller corresponding regions of homology. The numbers indicate nucleotide positions and correspond to the nucleotide numbers shown in Fig. 3. The horizontal arrows with N at the left end and C at the right show the largest regions in the two kDNAs that could potentially code for a polypeptide as discussed in the text. The five tandem arrows above the lines on the right indicate five repeats of a decanucleotide sequence in kDNA 201.

two small regions. The first is a region of eight base pairs in which six are different; the second consists of two adjacent base pairs that differ in the two molecules. In total, 114 out of the 122 positions are identical. This 122-base-pair region is also distinguished by the fact that it has a dG + dC content of nearly 50% whereas the remaining sequences of the two molecules are only about 23% dG + dC. The only *Taq* I site of kDNA 51 is contained within this 122-base-pair sequence. One of the four *Taq* I sites in kDNA 201 is in the same position within the homology as it is in kDNA 51. The other three *Taq* I sites are located just beyond the extremities of the 122-base-pair homology. Given that these *Taq* I sites are either within or just outside the homologous region, it is of interest that Borst and Hoeijmakers (5) have found that *Taq* I cleaves all, or nearly all, of the minicircles in the *T. brucei* kDNA network. In addition, *Taq* I cleaves all six cloned minicircles kDNAs characterized at the beginning of this study. Furthermore, previous renaturation kinetics indicates that kDNA minicircles share about 25% of their sequences (20). All of these findings are consistent with the view that most, if not all, *T. brucei* minicircles contain a region similar in sequence to this 122-base-pair region.

In addition to this region of striking homology, kDNAs 51 and 201 also contain other common sequences that would not be expected to occur on a strictly random basis. Fig. 4 shows those regions of 10 base pairs or longer that are identical in sequence. There are 12 such regions, the longest of which is 15 base pairs. Together they account for approximately 150 base pairs of additional homology. It is difficult to decide on the exact number of base pairs that fall into this group because some of the homologous regions overlap, or nearly overlap, one another. There is no obvious pattern to the locations of these small homologous regions; e.g., homology "d" occurs 54 nucleotides after the 122-base-pair region in kDNA 201 and 372 nucleotides after it in kDNA 51. One such homologous region, homology "e," is repeated three times in both kDNA 51 and kDNA 201. Another sequence, homology "p," is repeated at four different locations in kDNA 51 but occurs only once in kDNA 201 as a series of related overlapping sequences.

The small homologous regions shown in Fig. 4 occur in the same strand as the 122-base-pair homology. If the kDNA 201 strand shown in Figs. 3*A* and 4 is compared with the strand of kDNA 51 that is complementary to the one shown in Figs. 3*B* and 4, only four perfect homologies 10 base pairs or longer occur. When all of the perfect homologies 10 base pairs or longer are summed, they account for about 27% of the two sequences, in agreement with the 25% homology that we predicted earlier from renaturation kinetics (C_0t analysis) of other cloned kDNA minicircles (20). If the constraints on the computer program that searched for homologies are relaxed somewhat so that it detects nonperfect homologies greater than 15 base pairs in which only 85% of the positions are a perfect match, then over 30 such homologies are found. Although these homologies have a very low probability of occurring on a random basis, they were difficult to analyze further because there are so many. At least some of these nonperfect homologies probably arise because of the high dA + dT content of the kDNA sequences.

Another striking feature of the two kDNAs sequences is the high frequency of the three (DNA) termination codons, TAA, TGA, and TAG. One strand of the 1004-base-pair kDNA 201 contains 101 termination codons whereas the other strand has 58. The two strands of the 983-base-pair kDNA 51 contain 90 and 63 termination codons. Thus, the first nucleotide of a termination codon occurs, on the average, once every 12 nucleotides. These termination codons are fairly uniformly distributed among the six translation reading frames of each molecule. Nevertheless, each molecule does contain one stretch of amino acid codons flanked by a methionine codon at the 5′ end and a termination codon at the 3′ end that could potentially code for a small polypeptide. In kDNA 201 this polypeptide is 52 amino acids long; in kDNA 51 it is 71 amino acids. As shown in Figs. 3 and 4, the codons for the initial methionine and the next 35 amino acids of both potential polypeptides occur within the 122-base-pair homology. Therefore, the NH_2-terminal half of the two potential polypeptides would be the same except for those amino acids specified by the two small differences in the 122-base homology. It has recently been proposed that one of the three (RNA) termination codons, UGA, serves as a tryptophan codon in yeast mitochondria (30). This raises the possibility that one or two of the three termination codons could also specify amino acids in trypanosome mitochondria and extend the length of these potential polypeptides. In addition, split genes containing intervening sequences have been found in yeast mitochondrial DNA (31). If kDNA likewise contains split genes, the intervening sequences could not be detected from the DNA sequence alone. Nevertheless, there is no evidence that kDNA minicircle sequences code for polypeptides *in vivo*. Efforts to detect RNA species that hybridize to minicircle kDNA of *T. brucei* have failed (5, 6) although the experiments may not have been sensitive enough to detect small amounts of unstable RNAs complementary to only some minicircle sequences. In contrast, Fouts and Wolstenholme (32) have detected RNA transcripts that are homologous to the minicircle kDNA of *Crithidia acanthocephali*. These transcripts are 240 nucleotides long (about 10% of the minicircle kDNA length) and hybridize to only one of the two minicircle strands, the H strand. Because it is unlikely that minicircle sequences of one organism are transcribed whereas those of another organism are not, the possibility remains that RNA transcripts and translation products are specified by the sequences shown in Fig. 3.

Individual kDNA minicircles have been detected which appear to be undergoing replication (33, 34). These molecules have a single replication "bubble" that can be observed in the electron microscope (35, 36). In general, replication origins in other DNA molecules whose sequences have been determined have been found to possess at least one distinctive sequence symmetry which is usually a 2-fold rotational symmetry (dyad symmetry). For example, the replication origin of simian virus 40 DNA (37) contains both a large palindrome and several large nonperfect dyad symmetries, whereas the replication origin of λ DNA (38) contains a large dyad symmetry. The sequences of kDNAs 51 and 201 were searched for similar dyad symmetries which might indicate their origins of replication. The largest such region in the two sequences occurs in kDNA 51 between positions 16 and 56. This region can form a potential hairpin loop with an 11-base-pair stem and a 17-nucleotide loop. No analogous symmetry is present in kDNA 201. Both kDNA sequences also contain several smaller nonidentical and nonperfect dyad symmetries that are not large enough to be considered statistically significant in a 1000-nucleotide sequence of 72% dA + dT content. Thus, if kDNA minicircle replication begins at a dyad symmetry, it does so at a symmetry that is not distinguished by its size or its occurrence in both kDNAs.

The kDNA 201 sequence contains one unusual tandem repetitive sequence. The decanucleotide sequence, d(G-T-A-A-A-G-T-T-A-G), is repeated in tandem five times between positions 310 and 360 for a total repeat length of 50 nucleotides. This region is shown in Fig. 3*B* by underlining and in Fig. 4 by short horizontal arrows. Neither this sequence nor a similar large tandem repetitive region occurs in kDNA 51. The kDNA 51 sequence does, however, contain two unrelated 11-nucleotide tandem repeats (positions 31–52 and 438–459).

In summary, the most distinguishing features to arise from the comparison of these two minicircle sequences are the extent and distribution of the homologies, the occurrence of some repetitive sequences, and the partially homologous segments coding for small polypeptides. The ubiquity of small homologous elements throughout the two sequences suggests that massive recombination may occur among different minicircle sequences to generate a randomized version of a unique ancestoral minicircle sequence. Two previous reports support this conclusion. First, when *C. acanthocephali* is initially grown in ^2H_2O and then grown for two generations in H_2O, all of the minicircles possess the intermediate densities expected if the minicircles exchange sequences by recombination (34). Second, electron microscopy of kDNA minicircles often reveals the figure-eight dimer structures (39) that are associated with genetic recombination of circular DNA molecules in other systems (40, 41).

An analysis of the two kDNA sequences shown in Fig. 3 does not provide direct evidence for a genetic function of the minicircles although it does not formally rule out the possibility. In addition to the potential small polypeptides discussed above, the 122-base-pair homology may provide other unique functions for the organism [for example, regulator RNA molecules to control gene expression (42)] and the remaining (dA + dT)-rich sequence may simply be spacer DNA similar to the (dA + dT)-rich spacer DNA found between genes in yeast mitochondrial DNA (43, 44). This spacer region in minicircle DNA might simply enlarge the size of the individual molecules to facilitate their replication or to permit a specific number of supertwists per molecule. Still other possibilities for minicircle function, suggested by other investigators (5), range from a repository for insertion sequences, to a structural role in cell division or a scaffolding that aids in the segregation of maxicircles after replication. Although the two sequences reported here do not distinguish among these possibilities, the availability of cloned kDNA minicircles of known sequence should facilitate the design of further experiments to determine their biological function. In particular, it should be possible to use kDNAs 51 and 201 as radioactive probes to conduct a more sensitive search for the occurrence of kDNA minicircle transcripts *in vivo* via the reverse Southern blot technique of Alwine *et al.* (45). In addition, it will be of interest to compare these two sequences with minicircle sequences of other *T. brucei* clones and other trypanosomatids.

We thank James L. Hartley and Dennis L. Miller for advice on the chemical cleavage reactions and sequencing gels, William Murphy for help with some of the early experiments on pkT201, and Dr. Paul T. Englund for providing a preprint of his review article on kinetoplast DNA prior to publication. This research was supported by National Science Foundation Grant PCM 76-13461 and National Institutes of Health Grant GM-21696.

1. Jones, A. W. (1967) *Introduction to Parasitology* (Addison-Wesley, Reading, MA), pp. 49–70.
2. Steinert, M. (1960) *J. Biophys. Biochem. Cytol.* **8**, 542–546.
3. Pitelka, D. (1961) *Exp. Cell Res.* **25**, 87–93.
4. Borst, P. & Hoeijmakers, J. H. J. (1979) *Plasmid* **2**, 20–40.
5. Borst, P. & Hoeijmakers, J. H. J. (1979) in *Extrachromosomal DNA* ICN-UCLA Symposium on Molecular and Cellular Biology, eds. Cummings, D., Borst, P., David, I., Weissman, S. & Fox, C. F. (Academic, New York), Vol. 15, pp. 515–532.
6. Englund, P. T. (1980) in *Biochemistry and Physiology of Protozoa*, eds. Levandowsky, M. & Hutner, S. H. (Academic, New York), 2nd Ed., Vol. 4, in press.
7. Simpson, A. M. & Simpson, L. (1974) *J. Protozool.* **21**, 774–781.
8. Fairlamb, A. H., Weislogel, P. O., Hoeijmakers, J. H. J. & Borst, P. (1978) *J. Cell Biol.* **76**, 293–309.
9. Englund, P. T., DiMaio, D. C. & Price, S. S. (1977) *J. Biol. Chem.* **252**, 6208–6216.
10. Englund, P. T. (1978) *Cell* **14**, 157–168.
11. Newton, B. A. (1974) in *Trypanosomiasis and Leishmaniasis*, Ciba Foundation Symposium 20, New Series (Elsevier, Amsterdam), pp. 285–301.
12. Riou, G. (1976) in *Biochemistry of Parasites and Host-Parasite Relationships*, ed. Van Den Bossche, H. (Elsevier, Amsterdam), pp. 237–244.
13. Fairlamb, A. H., Weislogel, P. O., Hoeijmakers, J. H. J. & Borst, P. (1978) *J. Cell Biol.* **76**, 293–309.
14. Simpson, L. & Simpson, A. M. (1978) *Cell* **14**, 169–178.
15. Hoeijmakers, J. H. J. & Borst, P. (1978) *Biochim. Biophys. Acta* **521**, 407–411.
16. Masuda, H., Simpson, L., Rosenblatt, H. & Simpson, A. M. (1979) *Gene* **6**, 51–73.
17. Steinert, M. & Van Assel, S. (1975) *Exp. Cell Res.* **96**, 406–409.
18. Kleisen, C. M. & Borst, P. (1975) *Biochim. Biophys. Acta* **407**, 474–478.
19. Steinert, M., Van Assel, S., Borst, P. & Newton, B. A. (1976) in *The Genetic Function of Mitochondria*, eds. Saccone, C. & Kroon, A. M. (Elsevier/North Holland, Amsterdam), pp. 71–81.
20. Donelson, J. E., Majiwa, P. A. O. & Williams, R. O. (1979) *Plasmid* **2**, 572–588.
21. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. (1974) *Cell* **3**, 315–325.
22. Maniatis, T., Jeffrey, A. & Van deSande, H. (1975) *Biochemistry* **14**, 3787–3794.
23. Nichols, B. & Donelson, J. E. (1978) *J. Virol.* **25**, 429–434.
24. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
25. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110.
26. Korn, L. J., Queen, C. L. & Wegman, M. N. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4401–4405.
27. Sutcliffe, J. G. (1978) *Proc. Natl. Acad. Sci. USA* **25**, 3737–3741.
28. Sutcliffe, J. G. (1978) *Nucleic Acids Res.* **5**, 2721–2728.
29. Simpson, L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1585–1588.
30. Macino, G., Coruzzi, G., Nobrega, F. G., Li, M. & Tzagoloff, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3784–3785.
31. Bos, J. L., Heyting, C., Borst, P., Arnberg, A. C. & Van Bruggen, E. F. J. (1978) *Nature (London)* **275**, 336–338.
32. Fouts, D. L. & Wolstenholme, D. R. (1979) *Nucleic Acids Res.* **6**, 3785–3804.
33. Simpson, L., Simpson, A. M. & Westley, R. D. (1974) *Biochim. Biophys. Acta* **349**, 161–172.
34. Manning, J. E. & Wolstenholme, D. R. (1976) *J. Cell Biol.* **70**, 406–418.
35. Brack, Ch., Delain, E., Riou, G. & Festy, B. (1972) *J. Ultrastruct. Res.* **39**, 568–579.
36. Englund, P. T. (1979) *J. Biol. Chem.* **254**, 4895–4900.
37. Subramanian, K. N., Dhar, R. & Weissman, S. (1977) *J. Biol. Chem.* **252**, 355–367.
38. Denniston-Thompson, K., Moore, D. D., Kruger, K. E., Furth, M. E. & Blattner, F. R. (1977) *Science* **198**, 1051–1056.
39. Brack, Ch. & Delain, E. (1975) *J. Cell Sci.* **17**, 287–306.
40. Holiday, R. (1964) *Genet. Res.* **5**, 282–304.
41. Potter, H. & Dressler, D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3698–3702.
42. Davidson, E. H. & Britten, R. J. (1979) *Science* **204**, 1052–1058.
43. Miller, D. M., Martin, N. C., Hung, D. P. & Donelson, J. E. (1979) *J. Biol. Chem.* **254**, 11735–11740.
44. Martin, N. C., Miller, D. L., Donelson, J. E., Sigurdson, C., Hartley, J. L., Moynihan, P. S. & Hung, D. P. (1979) in *Extrachromosomal DNA*, ICN-UCLA Symposium on Molecular and Cellular Biology, eds. Cummings, D., Borst, P., David, I., Weissman, S. & Fox, C. F. (Academic, New York), Vol. 15, pp. 357–376.
45. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.