# Kinetoplast DNA in Trypanosomid Flagellates

LARRY SIMPSON

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

I. Introduction ... 120
II. Minicircle DNA: Structure and Complexity ... 120
A. *Trypanosoma brucei* ... 121
B. *Trypanosoma gambiense* ... 127
C. *Trypanosoma equiperdum* ... 127
D. *Trypanosoma evansi* ... 129
E. *Trypanosoma mega* ... 129
F. *Trypanosoma lewisi* ... 129
G. *Trypanosoma cruzi* ... 130
H. *Crithidia* Species ... 130
I. *Leishmania tarentolae* ... 133
J. *Phytomonas davidi* ... 141
III. Evolution of Minicircle Sequence Heterogeneity ... 141
IV. Transcription of Minicircle DNA ... 148
V. Replication of Minicircle Kinetoplast DNA ... 149
VI. Maxicircle DNA: Isolation, Cloning, and Restriction Mapping ... 150
A. *Trypanosoma brucei* ... 151
B. *Trypanosoma evansi* ... 154
C. *Trypanosoma equiperdum* ... 154
D. *Trypanosoma mega* ... 155
E. *Trypanosoma cruzi* ... 155
F. *Leishmania tarentolae* ... 155
G. *Crithidia* Species ... 157
H. *Herpetomonas samuelpessoa* ... 160
VII. Replication of Maxicircle DNA ... 160
A. Replication by Rolling-Circle Model ... 160
B. Autonomous Replication Sequences in Maxicircle DNA ... 160
VIII. Transcription of Maxicircle DNA ... 161
A. *Trypanosoma brucei* ... 161
B. *Leishmania tarentolae* ... 163
IX. Kinetoplast Ribosomes ... 166
A. Isolation of Ribosomes ... 166
B. DNA Sequences of 9 and 12 S RNA Genes ... 166
C. Secondary Structures of 9 and 12 S RNAs ... 167
X. Genomic Organization of Maxicircle DNA ... 167
A. Maxicircle Structural Genes ... 167
B. Evolution of Maxicircle DNA: Structural and Ribosomal Genes ... 172
C. Evolution of Maxicircle DNA: Divergent Region ... 172



Copyright © 1986 by Academic Press, Inc
All rights of reproduction in any form reserved

XI. Unusual Kinetoplast DNAs .............................. 173
XII. Conclusions .............................. 174
References .............................. 176

## I. Introduction

This review will emphasize recent developments in the application of recombinant DNA technology to problems of the structure, replication, and transcription of the unusual mitochondrial DNA in the kinetoplastid protozoa known as kinetoplast DNA. Kinetoplast DNA will be discussed in terms of its two molecular components, minicircles and maxicircles, and species-dependent variations will be emphasized.

## II. Minicircle DNA: Structure and Complexity

The minicircle component of the kinetoplast DNA is the most unusual aspect of this mitochondrial genetic system. Circular DNA molecules are not uncommon in nature nor in other mitochondrial genetic systems, but nowhere in nature does one find thousands of small circles catenated together into a single giant network of DNA such as found in the kinetoplastid protozoa. The function of minicircle DNA is still a mystery, but much has been learned recently about the structure and replication of this DNA in several kinetoplastid species: *Trypanosoma brucei, Trypanosoma equiperdum, Leishmania tarentolae, Crithidia fasciculata,* and *Crithidia luciliae.* In general minicircle DNA from all these species has the following general characteristics: (1) circles exist catenated with each other, although there is a small percentage of unattached circles that may have a functional significance; (2) within any one species the circles are of fairly uniform size although the size varies from species to species; (3) there is a species-dependent variable amount of sequence heterogeneity among the minicircles from any one clonal population; (4) the sequence heterogeneity is expressed on the level of the individual minicircle as a variable region and a constant region and sequence changes among the minicircle population in a kinetoplastid species occur rapidly in nature; and (5) there is no apparent sequence homology between the minicircles of a given species and the maxicircle DNA. In addition to these general common properties of minicircle DNA, there are many differences between the various kinetoplastid species, and these differences may prove

to have phylogenetic significance. Therefore, I will discuss results from each species in succession.

### A. *Trypanosoma brucei*

From the initial optical reassociation studies of Steinert *et al.* (1976) it was realized that the kDNA of *T. brucei* was perhaps the most complex in terms of sequence heterogeneity and the most rapidly changing in terms of evolution. The DNA reassociated with a single copy complexity of 300 times the size of the minicircle, and no hybridization between different *T. brucei* strains was observed. The high complexity of kDNA from *T. brucei* was confirmed by the $S_1$ $C_0t$ analysis of Donelson *et al.* (1979) in which [$^3$H]kDNA renatured with a single copy $C_0t_{1/2}$ 316 times the $C_0t_{1/2}$ when a cloned 1.1-kb minicircle (pkT51) was the excess driver (Fig. 1). However, there was a minor component (15%) of the kDNA which renatured at a rate suggestive of repetitive sequences. Renaturation of a cloned minicircle, pkT51, in the presence of excess DNA from two other cloned minicircles, pkT3 and pkT204, showed that approximately 25% of the sequence of the pkT51 insert is homologous to sequences in both the pkT3 and pkT204 inserts (Fig. 1). Stuart (1979) confirmed that there are approximately 300 different minicircle sequence classes in *T. brucei* 164 kDNA and showed that the minicircle classes differ in relative abundance (Stuart and Gelvin, 1982). For example, a *Hin*dIII cut minicircle cloned into the *Hin*dIII site of pBR322 renatured at a rate implying the existence of 500 copies per network, whereas *Bam*HI cleaved uncloned minicircles renatured at a rate implying the existence of 60 copies per network (Fig. 2).

However, assuming a *T. brucei* network mass of $4.2 \times 10^9$ Da that is composed of 15% maxicircle DNA of 22-kb molecules and 85% minicircle DNA of 1.1-kb molecules, one can calculate that there are approximately 45 maxicircles and 5500 minicircles (Stuart and Gelvin, 1980). Furthermore, assuming that there are about 300 different minicircle sequence classes, this would imply about 18 copies of each sequence class if the classes are of equal size. The finding (Stuart and Gelvin, 1980) that one cloned minicircle is present in 500 copies implies that there is a wide divergence in frequency of the different sequence classes in *T. brucei*.

The sequence heterogeneity of *T. brucei* minicircle DNA shown by $C_0t$ analysis is evidenced also by restriction enzyme analysis on gels (Fig. 3). Digestion of *T. brucei* kDNA with most restriction enzymes yields low release of unit-length linearized minicircles with no smaller bands (Borst *et al.*, 1980a, 1981a; Brunel *et al.*, 1980a; Fairlamb *et al.*, 1978; Simpson and Simpson, 1980). Most of the DNA remains catenated in the network

that does not enter the gel. This low release of mainly once-cut minicircles by most restriction enzymes is due both to the presence of many minicircle sequence classes and to the high percentage of AT in the minicircle DNA of *T. brucei*. However, digestion with *Taq*I, *Alu*I, or *Mbo*I does release most of the minicircles from the network and produce fragments of less than unit length, implying that sites for these enzymes are present more than once in conserved regions of the minicircles. The apparently single bands of unit-length minicircle DNA released by digestion with enzymes such as *Eco*RI, *Pst*I, *Hpa*II, *Hae*III, *Hin*dIII, and *Hha*I were

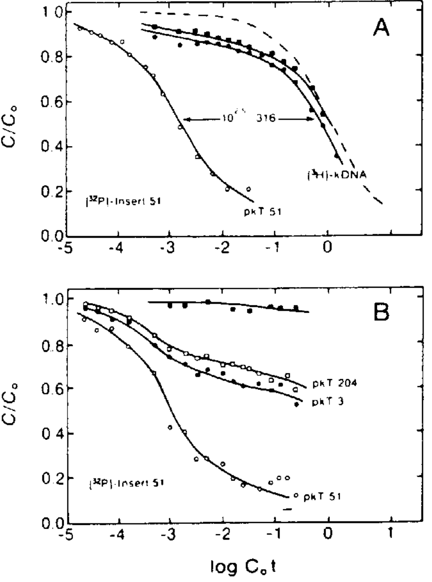


FIG. 1. Renaturation kinetics of the *Trypanosoma brucei* kDNA insert of pkT51 which had been labeled *in vitro* with $^{32}P$. (A) The renaturation of $^{32}P$-labeled pkT51 insert DNA in the presence of an excess of pkT51 DNA (open circles) or an excess of total network kDNA (solid circles). The network kDNA was labeled with $^{3}H$ *in vitro* prior to shearing so that its renaturation could also be monitored (solid squares). The broken line is a model curve generated by the standard renaturation equation when it is assumed that a single-copy DNA sequence has a log $C_0t_{1/2}$ = 0.15. (B) The renaturation of $^{32}P$-labeled pkT51 insert DNA (open circles), pkT3 DNA (solid circles), pkT204 DNA (open squares), or pBR322 DNA (solid squares). The initial concentration, $C_0$, of the excess unlabeled driver DNAs refers to the concentration of only the inserted kDNA sequence when a pkT recombinant plasmid is the driver DNA. In all of the experiments $C_0$ was chosen so that the final time point was taken 11 hours after the start of the reaction. About 4000 $^{32}P$ cpm were sampled at each time point. The pBR322 sequence does not drive the renaturation of the kDNA insert 51 within the time span of the renaturation experiments (solid squares). Reprinted from Donelson *et al.* (1979) with permission.

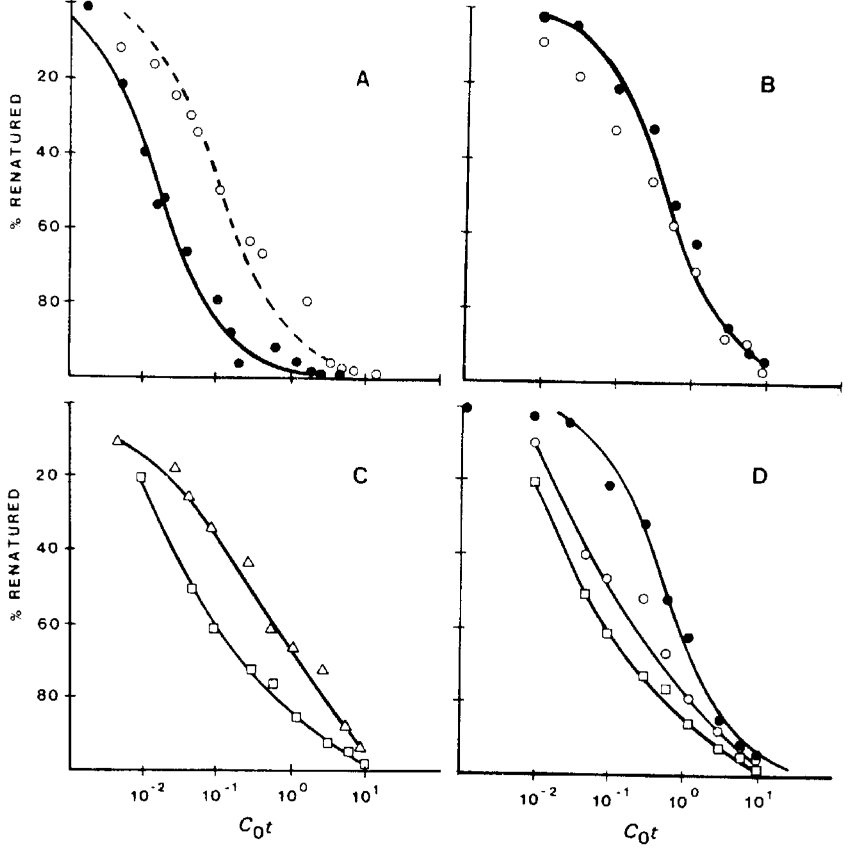


FIG. 2. Renaturation kinetics of nick-translated kDNA sequences of *T. brucei*. (A) kDNA and homologous driver (open circles) compared to lambda DNA standard (solid circle). Both lines are theoretical single component curves. (B) Maxicircle (open circles) and cloned maxicircle fragment (solid circles) driven by kDNA. (C) Cloned *Hin*dIII minicircle (open squares) and *Bam*HI minicircle (open triangles) sequences driven by kDNA. (D) Cloned *Hin*dIII minicircle (open squares), kDNA (open circles), and cloned maxicircle segment (solid circles) driven by kDNA. The curve through the cloned maxicircle segment points is a theoretical single component curve. Reprinted from Stuart and Gelvin (1980) with permission.

shown to be composed of 5–10 closely spaced bands by acrylamide gel electrophoresis (Simpson and Simpson, 1980) (Fig. 4), implying either a substantial minor length heterogeneity or the existence of gel mobility anomalies, as shown to exist in *L. tarentolae* minicircle DNA (Challberg and Englund, 1980; Simpson, 1979).

Analysis of cloned minicircle DNA has greatly assisted our understand-

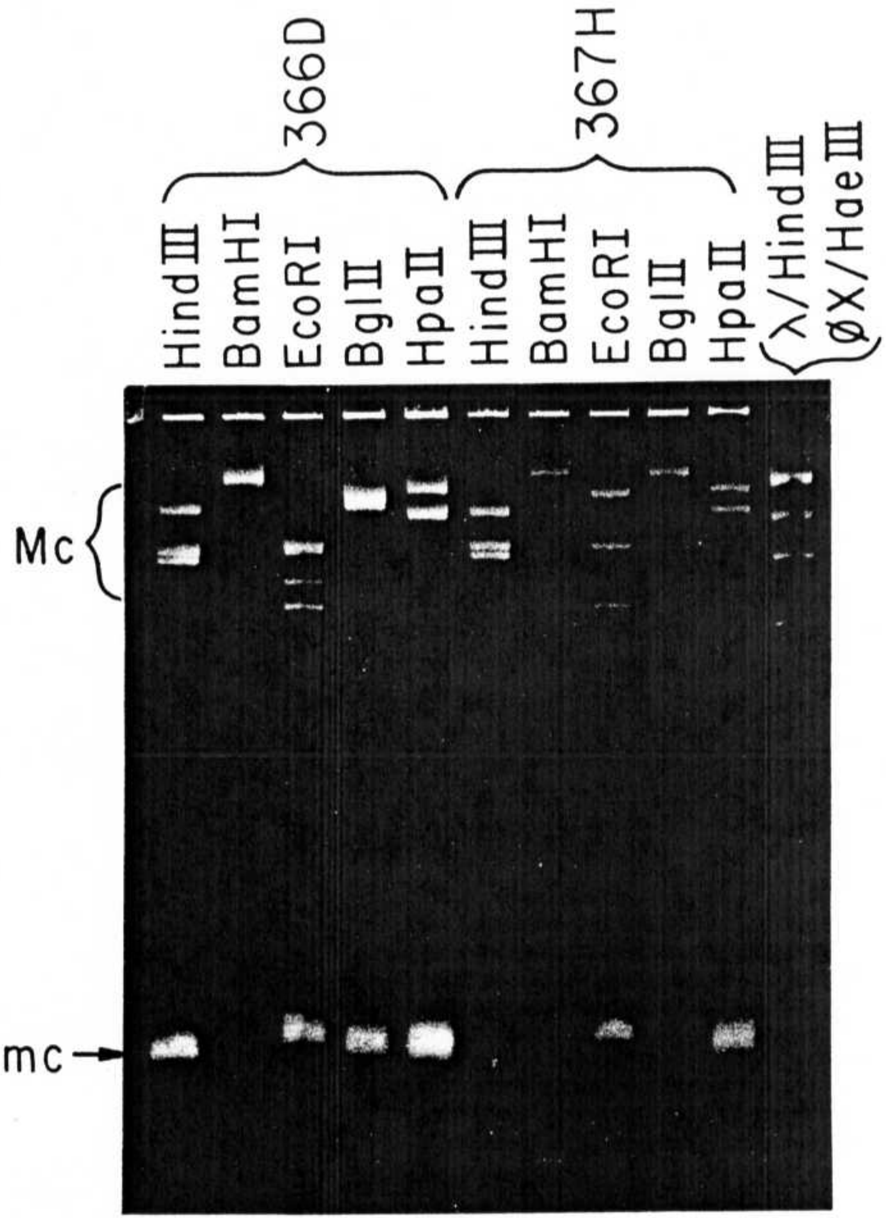


FIG. 3. Comparison of digestions of kDNA from two clonal strains (366D and 367H) of *T. brucei* procyclics in 0.8% agarose. The minicircle (mc) and maxicircle (Mc) regions are indicated. The molecular weight references are λ/*Hin*dIII and ϕXRF/*Hae*III fragments. Reprinted from Simpson and Simpson (1980) with permission.

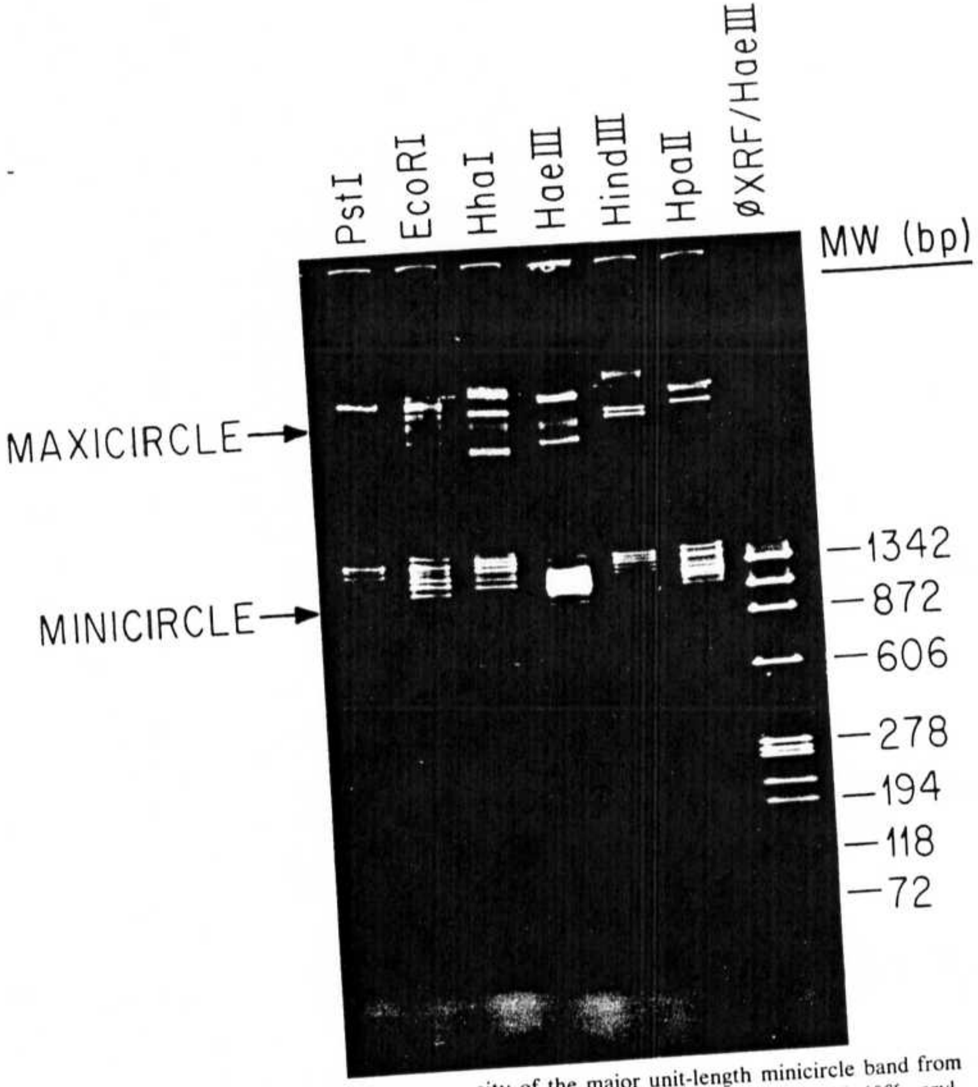


FIG. 4. Demonstration of heterogeneity of the major unit-length minicircle band from 366D procyclic *T. brucei* kDNA in an acrylamide gradient gel. Conditions: 3.5–10% acrylamide with a 3.0% stacking region. The molecular weight references are $\phi$XRF/*Hae*III fragments. Reprinted from Simpson and Simpson (1980) with permission.

ing of the observed sequence heterogeneity. Donelson *et al.* (1979) cloned minicircles from *T. brucei* clone 18E2 either by C-tailing minicircles released by digestion with *Hae*III, or with a mixture of *Hpa*II, *Hha*I, and *Alu*I and cloning into the *Pst*I site of pBR322, or by direct ligation of *Hin*dIII- or *Pst*I-released minicircles into the appropriate sites in pBR322. One minicircle, pkT3, was labeled *in vitro* and used as a probe in a Southern blot hybridization of kDNA digested with several enzymes. Hybridization was observed in all minicircle bands released by digestion with six enzymes. Similar results were obtained by colony hybridization, using three minicircle clones as probes. In general the probes hybridized strongly to themselves and weakly to all the other minicircle clones. Brunel *et al.* (1980a,b) cloned *Hin*dIII minicircles from *T. brucei* 427–60 in pBR322 and also *Eco*RI minicircles in λgtWES/λB (together with a maxicircle fragment, or with λB) and tested the cloned inserts for digestion with nine enzymes. The only sites found were single sites for *Eco*RI, *Hha*I, *Hae*III, and two to three sites for *Taq*I; all four of the clones studied contained different patterns of restriction sites. However, when the λ clones were spotted onto filters and probed with five of the pBR322 clones, some hybridization was observed with all clones, implying common sequences.

Simpson and Simpson (1980) cloned *Bam*HI-released minicircles from *T. brucei* 366D into pBR322 and tested 25 clones by colony hybridization using one, pTb7, as a probe. Four of the 25 clones showed strong hybridization, but all showed some hybridization, implying the existence of more than one semihomologous minicircle sequence class with a single *Bam*HI site.

Some heteroduplex studies of *T. brucei* minicircles have been reported, but the small size of these molecules precludes obtaining much information from this method (Brunel *et al.*, 1980a; Donelson *et al.*, 1979).

A direct confirmation and extension of the hybridization results was reported by Chen and Donelson (1980) in a sequence analysis of two cloned minicircles from *T. brucei* (Fig. 5). pk51 was derived from a digest of kDNA with a mixture of *Hha*I, *Hpa*II, and *Alu*I, tailed with G and inserted into the *Pst*I site of pBR322. pk201 was released from kDNA by digestion with *Pst*I and ligated directly into the *Pst*I site of pBR322. Both molecules contained approximately 72% AT, as expected from the low buoyant density of the kDNA ($\rho$ = 1.690 g/ml). The most striking aspect was a nearly perfect homology of 122 bp. There were also 12 other common regions containing 50% AT. pk201 contained a decanucleotide sequence repeated 5 times in tandem. The only region of dyad symmetry was an 11-bp sequence in pk51. There was a high frequency of termination codons, TAA, TGA, and TAG, in all reading frames in both mole-

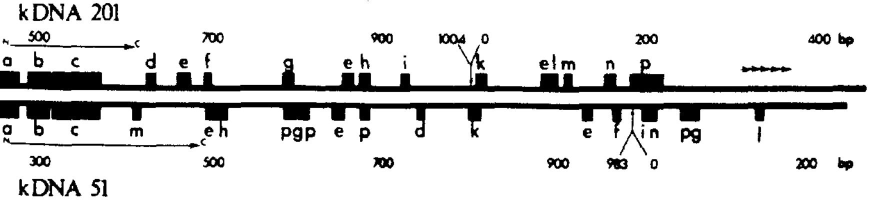


FIG. 5. Diagram showing the relative locations of perfect homologies between *T. brucei* cloned minicircle kDNAs 201 and 51 that are equal to or greater than 10 bp. The two circular kDNA sequences are represented by two lines aligned so that each begins with the first position of the 122-bp near-perfect homology. The other letters indicate the other smaller corresponding regions of homology. The numbers indicate nucleotide positions. 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. The five tandem arrows above the lines on the right indicate five repeats of a decanucleotide sequence in kDNA 201. Reprinted from Chen and Donelson (1980) with permission.

cules; the longest open reading frames were 52 amino acids in pk201 and 71 amino acids in pk51. Interestingly, these small open reading frames encompassed the 122-bp common region. The sequence data nicely confirmed the evidence from the $C_0t$ analysis that the different minicircle classes had approximately 25% of their sequences in common, since summation of all the perfect homologies of 10 bp or greater gave a value of 27%.

### B. *Trypanosoma gambiense*

Riou and Barrois (1981) found that the minicircles from the kDNA networks of *T. gambiense* were similar to those from *T. brucei* in terms of cleavage by restriction enzymes. The only enzymes that cleaved the majority of the network minicircles and produced fragments of less than unit length were *Taq*I, *Hin*fI, and *Xba*I. The patterns with these enzymes indicated extensive minicircle sequence heterogeneity. They also found that the minicircles hybridized somewhat with minicircle DNA from the Pasteur Institute strain of *T. equiperdum*.

### C. *Trypanosoma equiperdum*

Unlike most other kinetoplastid species, the kDNA minicircles from *T. equiperdum* are homogeneous in base sequence. This was first shown by Riou and Barrois (1979) for a strain of *T. equiperdum* from the Pasteur Institute and by Frasch *et al.* (1980) for a strain from the ATCC (30019). However, these two strains differ signficantly in the maxicircle compo-

nents, indicating a possible molecular pleomorphism among strains of this species.

A unique circular restriction enzyme map for the uncloned minicircle from the Pasteur *T. equiperdum* strain was derived by Riou and Barrois (1979) and the complete nucleotide sequence was reported by Barrois *et al.* (1982). The molecule contained 73% AT, had a 6-fold tandem repeat of a 12-bp sequence, and had three small dyad symmetries of 9–11 bp distributed equidistantly from each other. In terms of possible codogenic function, the longest open reading frames were 18 and 22 amino acids (Fig. 6). A sequence of 130 bp was strikingly homologous to the 122-bp common sequence to *T. brucei* minicircles pkT51 and pkT201; it differed from pkT201 by 12 bp and from pkT51 by 15 bp. The remainder of the *T. equiperdum* molecule had no significant homology with the *T. brucei* sequences. This cross-species partial sequence conservation is intriguing in terms of the functional significance of the minicircle. Both Chen and Donelson (1980) and Barrois *et al.* (1982) speculated that the conserved

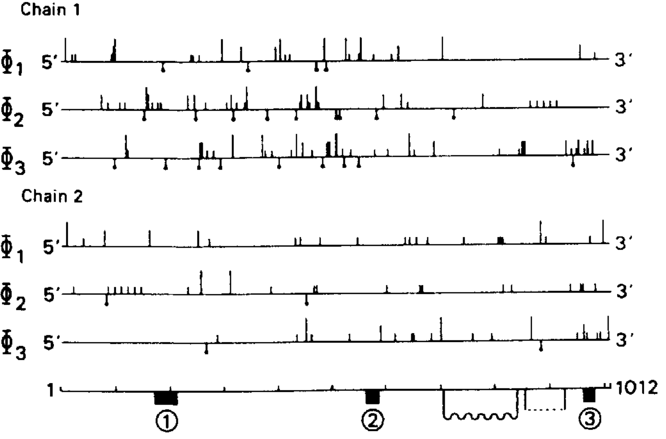


FIG. 6. Diagram showing the distribution of nonsense (vertical bars) and initiation (inverted bars with solid circles) codons in the six reading frames of *T. equiperdum* minicircle DNA. Three reading frames called $\phi_1$, $\phi_2$, and $\phi_3$ were initiated from the 5′ end of each DNA strand. On chain 1, $\phi_1$ is initiated by its first triplet AAT, $\phi_2$ by ATC, and $\phi_3$ by TCA. On chain 2, $\phi_1$ is initiated by its first triplet ATT, $\phi_2$ by TTC, and $\phi_3$ by TCT. The length of the kDNA minicircle does not correspond to a multiple of three nucleotides. Therefore, going through *Hin*fI site, phase $\phi_1$ is changed into phase $\phi_2$. Vertical bars have different lengths, depending on the stop codon (small-sized vertical bar, TAA; medium-sized vertical bar, TAG; large-sized vertical bar, TGA). (Bottom) Schematic representation of the minicircle, with the locations of the three dyad symmetries (circled numbers 1, 2, and 3), the DNA region homologous to *T. brucei* minicircles (wavy line), and the six repeats of 12 bp each (dashed line). Reprinted from Barrois *et al.* (1982) with permission.