

Kinetoplast DNA Minicircles Encode Guide RNAs for Editing of Cytochrome Oxidase Subunit III mRNA

Nancy R. Sturm* and Larry Simpson*†

* Department of Biology

† Molecular Biology Institute

University of California

Los Angeles, California 90024

Summary

Guide RNAs (gRNAs) for the editing of sites 1–8 of COIII mRNA and an “unexpected” partially edited COIII mRNA are encoded in the variable regions of specific kinetoplast DNA minicircles. The gRNAs can form 37 and 44 nucleotide perfect hybrids (allowing for G-U base pairs) with edited mRNAs. The gRNAs were detected on Northern blots and shown to have unique 5' ends situated close to the beginning of the potential base pairing with the edited mRNAs. We suggest that kinetoplast DNA minicircle molecules in general may encode gRNAs for editing of cryptogene mRNAs by a mechanism similar to that previously proposed for editing by maxicircle-encoded gRNAs.

Introduction

The mitochondrial DNA of the kinetoplast protozoa is known as kinetoplast DNA (kDNA). Each kDNA network contains approximately 10^4 minicircles and 50 maxicircles, all catenated together in a highly organized fashion. The minicircle component of the kDNA has long been a mystery in terms of a possible genetic function (reviewed in Simpson, 1986). The minicircles contain one to four conserved regions, depending on the species, and a variable region (Kidane et al., 1984; Ray, 1989; Simpson, 1986). The multiple copies within the network comprise a species-specific number of different sequence classes as defined by the variable region sequences.

A 240 nucleotide minicircle transcript has been reported from *Trypanosoma brucei* (Rohrer et al., 1987) and *Criethidia fasciculata* (Fouts and Wolstenholme, 1979), and a family of up to approximately 24 low abundance transcripts each differing by 1 nucleotide and migrating ahead of tRNAs in acrylamide gels has been detected in *Leishmania tarentolae* mitochondrial RNA (Simpson et al., 1989). Maxicircle transcripts of similar electrophoretic mobility have also been detected in mitochondrial RNA from *L. tarentolae* (Simpson et al., 1989). These maxicircle transcripts have been identified as guide RNAs (gRNAs; Blum et al., 1990). The function of the minicircle transcripts is unknown; they do not appear to serve a standard coding function since the known minicircle sequences do not contain conserved open reading frames.

A model for the editing of maxicircle cryptogene mRNAs involving the participation of maxicircle-encoded gRNA molecules has been presented previously (Blum et al., 1990). This model postulates a 3' to 5' progressive process with multiple cycles of specific cleavages at sites of

mRNA–gRNA mismatch, addition or deletion of U's at the site to create a perfect mRNA–gRNA match (allowing for G-U base pairs), and finally religation of the edited mRNA site. The gRNAs provide the information for the edited sequence in the form of complementary A or G residues at the sites of addition of U residues in the mRNA and the absence of complementary residues at the sites of encoded U residues that are to be deleted.

Appropriate maxicircle-encoded gRNAs for four of the five known maxicircle cryptogenes in *L. tarentolae* have been identified. However, no appropriate gRNAs were identified for the 5'-edited cytochrome oxidase subunit III (COIII) mRNA. Editing of COIII mRNA appears to have a more relaxed level of fidelity than editing of cytochrome b (CYb) mRNA, for which over 95% of partially edited transcripts follow a strict 3' to 5' process, since more than 40% of partially edited COIII mRNAs had patterns that did not agree with a strict 3' to 5' editing process interpreted in terms of the mature COIII RNA sequence (Sturm and Simpson, 1990).

In this paper we present the characterization of two minicircle-encoded gRNA molecules that could provide the information for editing of sites 1–8 of COIII mRNA and editing of one of the “unexpected” partially edited COIII mRNAs. We suggest that encoding of gRNAs represents one of the major genetic functions of the minicircles.

Results

Identification of a Minicircle Sequence Complementary to Sites 1–8 of Edited COIII mRNA

Three complete (Kidane et al., 1984) and five partial minicircle sequences (Muhich and Simpson, 1986) are known from *L. tarentolae* kDNA. A computer search for potential minicircle transcripts that could form perfect hybrids (assuming G-U base pairs) with edited COIII mRNA identified a 37-mer region within the D-12 partial minicircle sequence, as shown in Figure 1. The COIII editing sites covered by this putative transcript include the 3' U addition sites 1–3 and all the U deletion sites 4–8. There is a 10 bp “3' anchor” sequence that presumably allows the initial hybridization between the unedited mRNA and the gRNA. The gRNA-edited mRNA hybrid contains a higher percentage of G-U base pairs upstream of the 3' anchor region than present in the known maxicircle-encoded gRNA-edited mRNA hybrids. No G-U base pairs are found in the 3' anchor sequence. In line with the nomenclature suggested for maxicircle-encoded gRNAs, we designate this putative gRNA transcript COIII gRNA-I.

The polarity of a *L. tarentolae* minicircle sequence (or any kinetoplast minicircle sequence) can be defined with reference to the polarity of the conserved 12-mer sequence motif GGGGTGGTGTA (Kidane et al., 1984; designated CSB-3 by Ray, 1989) within the conserved region. The D-12 minicircle partial sequence (Muhich and Simpson, 1986) does not contain this motif but does have

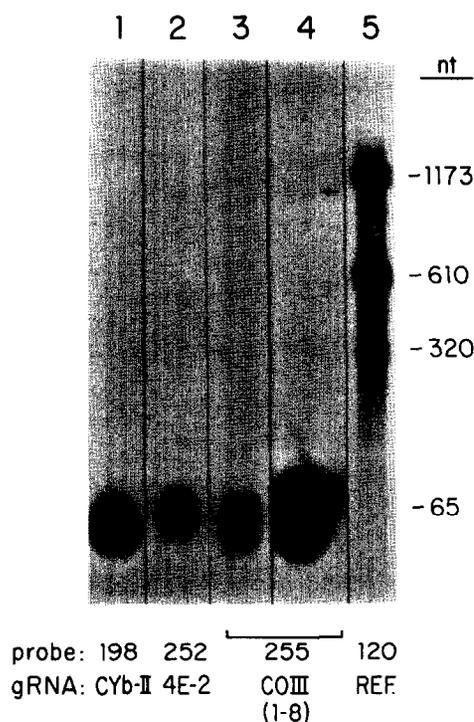


Figure 3. Northern Blot Analysis of Minicircle-Encoded gRNAs
Northern blots of total kRNA separated in formaldehyde-agarose were hybridized with labeled oligonucleotide probes for gRNAs. The probe and gRNA are indicated below each lane. Lanes 1-3 were exposed for the same time. Lane 4 is a reexposure of lane 3 to reveal the minor high molecular weight band (arrow). Lane 5 is a rehybridization of a stripped blot with pLt120 insert DNA (Simpson et al., 1985), to provide molecular weight markers.

a short region of similarity (indicated in Figure 1C) with the variable region of the pKSR1/Lt19 minicircle, thereby tentatively defining both the localization of this sequence within the minicircle and the polarity of the putative gRNA transcript.

Identification of a Minicircle Sequence Complementary to an Unexpected Partially Edited COIII mRNA

Polymerase chain reaction selection of partially edited COIII mRNAs using downstream edited and upstream unedited primers yielded a high percentage (42%) of clones containing sequence patterns that were unexpected in terms of the strict 3' to 5' gRNA model for editing of the COIII mRNA (Sturm and Simpson, 1990). We hypothesized that these unexpected patterns were due to mechanistically correct editing with incorrect gRNAs (Sturm and Simpson, 1990). A computer search of both maxicircle and minicircle DNA was performed in an attempt to locate the genomic source of these putative gRNAs. A 44-mer region within the variable region of the Lt154 minicircle sequence was identified that could produce a transcript capable of forming a perfect hybrid with the unexpected partially edited sequence 4E-2 (Figure 4 in Sturm and Simpson, 1990), as shown in Figure 2. The putative 3' anchor sequence for this hybrid could be at least 7 bp,

assuming that editing with this gRNA is occurring on an mRNA that has already been partially edited with the COIII gRNA-I. As is the case for the COIII gRNA-I-edited mRNA hybrid, there is a high percentage of G-U base pairs in this hybrid.

The polarity of the sense strand of this 44-mer within the Lt154 minicircle was identical to that of the 37-mer in the D-12 minicircle in Figure 1. Both polarities agreed with that determined previously for the small minicircle transcripts identified by Simpson et al. (1989) in mitochondrial tRNA preparations.

Detection of Minicircle-Encoded gRNAs by Northern Blot Hybridization

Oligonucleotide probes for both identified minicircle-encoded gRNA sequences were hybridized to Northern blots of formaldehyde-agarose gels of total kinoplast RNA (kRNA). As shown in Figure 3, each probe hybridized to gRNA bands that approximately comigrated with maxicircle-encoded CYb gRNA-II, indicating an approximate size of 65 nucleotides and an approximate relative abundance less than that of the CYb gRNA-II. A minor high molecular weight band, which may represent a precursor RNA, could be visualized with the COIII gRNA-I probe after a longer exposure time (arrow, lane 4).

A sense COIII gRNA-I oligonucleotide probe (S-260) did not detect any transcripts from the opposite strand (data not shown).

Localization of Minicircle gRNA 5' Ends

Primer extension sequence analysis of the 5' ends of the two identified gRNAs was performed as shown in Figure 4, using the primers indicated in Figures 1 and 2. The 5' ends of the gRNAs, shown by the strong stops in the primer extension ladders, are indicated by arrows in the sequences in Figures 1 and 2.

3' End Heterogeneity of the gRNAs

Acrylamide gel analysis (Figure 5) of the two identified minicircle-encoded gRNAs showed extensive length heterogeneity, which can be attributed to the 3' end since the 5' ends were shown to be unique. Both gRNAs migrated ahead of tRNA and were not abundant enough to be visible in the ethidium bromide-stained gel (Figure 5A). The COIII gRNA-I (Figure 5B) showed a larger family of bands differing by single nucleotides than the 4E-2 gRNA. The sizes of the most abundant species of the two gRNAs were approximately 56 and 61 nucleotides, respectively, as compared with the 65 nucleotide size of the most abundant maxicircle-encoded gRNA, CYb gRNA-II, in lane 1. By analogy with the maxicircle-encoded gRNAs, the 3' end length heterogeneity of the minicircle-encoded gRNAs may also be due to nontranscribed oligo(U) tails of variable length (Blum and Simpson, unpublished data), but this must be demonstrated by direct sequence analysis. Identical minicircle gRNA gel profiles were seen using total cell RNA (data not shown), suggesting that the 3' heterogeneity was not produced by the addition of U's during mitochondrial isolation.

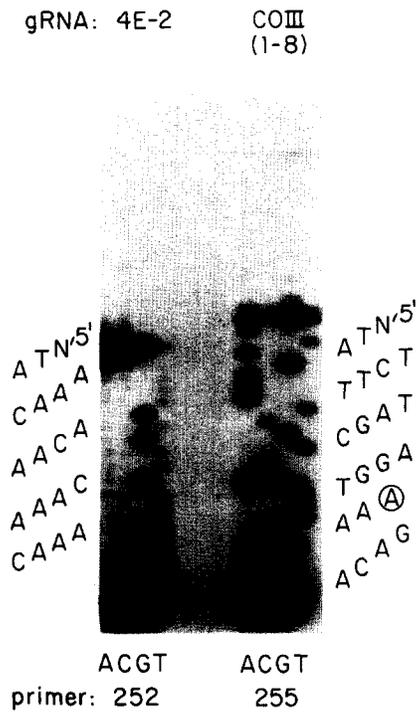


Figure 4. Primer Extension Sequencing of the 5' Ends of the Minicircle-Encoded gRNAs

The S-255 primer (COIII gRNA-I) and the S-252 primer (4E-2 gRNA) were used for sequence analysis of total kRNA. The RNA sequences to the 5' ends (strong extension products) are shown next to the ladders. The circled A residue is present in the RNA sequence but not in the published D-12 sequence (Muhich and Simpson, 1986). It is present in the D-12-1 DNA sequence of uncloned kDNA (Figure 1B).

Genomic Origin of the gRNAs

The gRNA oligonucleotide probes were used on Southern blots of total kDNA digested with several restriction enzymes. As shown in Figure 6, the COIII gRNA-I probe hybridized with a single unit-length MspI minicircle fragment and a single HaeIII minicircle fragment, and the unexpected 4E-2 gRNA probe hybridized with single minicircle fragments of different sizes liberated by MspI, HaeIII, and HindIII, respectively. As a control, the maxicircle-encoded CYb gRNA-II probe hybridized with the expected maxicircle fragments in each case. These results confirm that the COIII gRNA-I and the 4E-2 gRNA represent transcripts of distinct minicircle sequence classes and do not hybridize with maxicircle DNA.

Direct sequence analysis of uncloned kDNA using the same oligonucleotide primers used for the Northern blot analysis showed a perfect correspondence of the genomic DNA sequences and the gRNA sequences (data not shown). A few single nucleotide differences, both within and flanking the gRNA regions, were noted with respect to the published cloned minicircle sequences originally used for identification of the gRNA sequences by computer analysis. These minor differences, noted in Figures 1 and 2, could be due to sequencing errors, changes introduced during cloning, or actual polymorphisms in the minicircle DNA population. It is clear from the direct se-

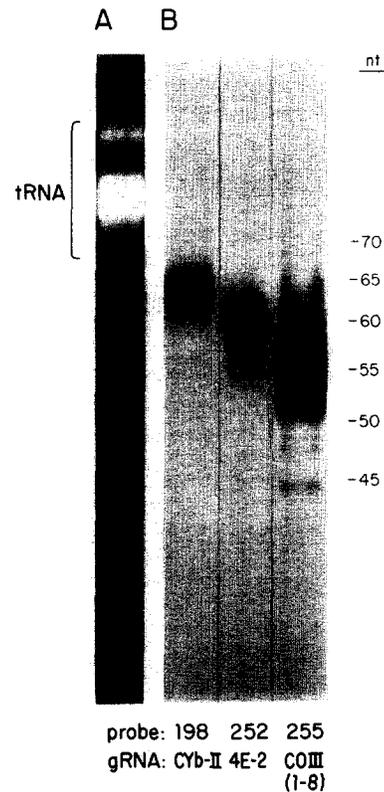


Figure 5. High Resolution Acrylamide Gel Analysis of Minicircle-Encoded gRNAs

Total kRNA was separated in 8% acrylamide-8 M urea and the gel electroblotted onto Nytran filters for hybridization with specific labeled gRNA oligonucleotide probes. The ethidium bromide-stained gel is shown in (A), with the tRNA region indicated. In (B), the probes and the specific gRNAs are indicated beneath each lane. Lane 1, control hybridization of the maxicircle-encoded CYb gRNA-II with the S-198 probe. Lane 2, hybridization with the 4E-2 gRNA probe. Lane 3, hybridization with the COIII gRNA-I probe.

quence analysis of total kinetoplast DNA, however, that the identified gRNAs represent transcripts from the variable regions of the major minicircle sequence classes within the two minicircle families involved.

Discussion

We have identified two minicircle-encoded gRNAs that may be involved in RNA editing. The COIII gRNA-I contains the sequence information for editing of sites 1-8 of the 5'-edited COIII mRNA, for which no appropriate maxicircle-encoded gRNAs had been identified (Blum et al., 1990). This gRNA has several characteristics in common with known maxicircle-encoded gRNAs: a 10 bp "3' anchor" sequence, a unique 5' end 7 nucleotides downstream (with respect to the mRNA polarity) of the beginning of the duplex mRNA-gRNA region, and a heterogeneous 3' end. The length of the duplex region that could be formed with edited COIII mRNA is 37 bp.

In addition, another minicircle-encoded gRNA was identified that contains the sequence information for the editing of the 4E-2 unexpected partially edited COIII mRNA.

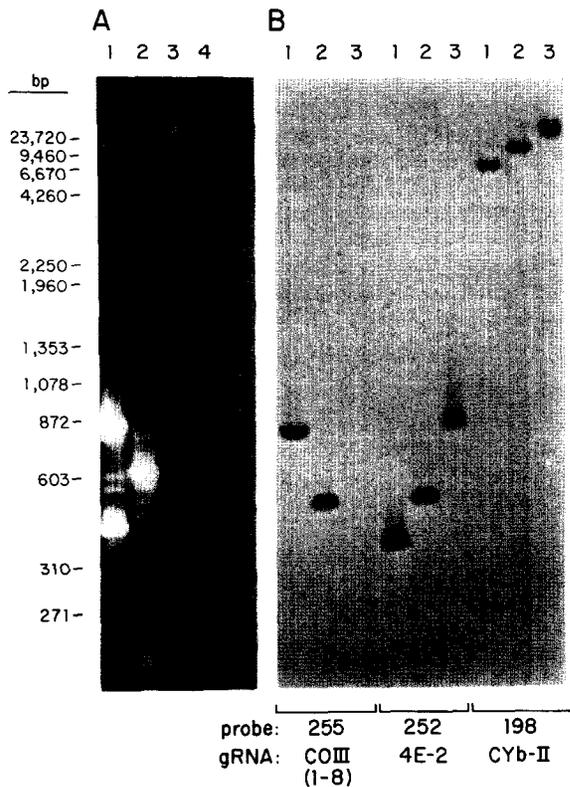


Figure 6. Southern Blot Analysis of gRNAs
Total kinetoplast DNA, digested with MspI (lane 1), HaeIII (lane 2), or HindIII (lane 3), was electrophoresed in agarose and blotted onto Nytran for hybridization with specific gRNA probes, as indicated beneath each autoradiograph. (A) is the ethidium bromide-stained gel, showing the lower minicircle fragments and the upper maxicircle fragments. Lane 4, λ /HindIII and ϕ X174/HaeIII size markers.

We suggest that this gRNA is normally involved with the editing of an unidentified cryptogene mRNA. In addition, this same gRNA can form a 38-mer hybrid (with one mismatch) with the 5A-2 unexpected partially edited COIII mRNA (Sturm and Simpson, 1990). We hypothesize that additional maxicircle- or minicircle-encoded gRNAs exist for other unidentified cryptogene mRNAs which could account for most if not all of the unexpected partially edited COIII patterns. A limitation on the search for additional minicircle-encoded gRNAs is the limited number of minicircle sequences known. In the case of *L. tarentolae*, three complete minicircle sequences and five partial sequences are known, out of an estimated 10–20 total sequence classes present in the network (Kidane et al., 1984; Muhich and Simpson, 1986).

The identification of minicircle-encoded gRNAs defines a genetic function of kinetoplast minicircle DNA. We speculate that multiple minicircle-encoded gRNAs may also contain the information for the pan-editing (Simpson and Shaw, 1989) of COIII mRNA as well as the maxicircle unidentified reading frame 3 and maxicircle unidentified reading frame 4 mRNAs in *T. brucei*. This hypothesis is consistent with several previous observations: First, the sequence complexity of minicircle DNA in *T. brucei*, which is 300 times the complexity of a single minicircle (Stuart,

1983; Donelson et al., 1979; Simpson and Simpson, 1978), is correlated with the occurrence of pan-editing of three maxicircle mRNAs (Feagin et al., 1988; Stuart, personal communication) in this species. Second, a homogeneous minicircle DNA population in *T. equiperdum* (Barois et al., 1982) is correlated with a maxicircle DNA that has undergone massive deletions (Frasch et al., 1980). Third, homogeneous minicircle DNA populations in several strains of *T. evansi* are correlated with loss of maxicircle DNA (Borst et al., 1987). Fourth, small minicircle transcripts have been described that have the same electrophoretic migration properties in acrylamide as maxicircle-encoded gRNAs, and have the same polarity as the minicircle gRNAs (Simpson et al., 1989).

It is possible that the cleavage sites recognized in the variable regions of minicircle DNA by mung bean nuclease in the presence of formamide (Muhich and Simpson, 1986) flank gRNA transcriptional units. McCutchan et al. (1984) reported that mung bean nuclease under these conditions produced duplex cleavages at sites before and after genes in DNA from a variety of lower eukaryotic cells.

It remains to be explained why minicircle variable-region sequences have evolved so rapidly in nature. They can be used as markers for strains or schizodemes (Morel et al., 1980; reviewed in Simpson, 1986), and show no cross-hybridization between species (Simpson, 1986). It is possible that there is little constraint on sequence changes within the variable region outside the short gRNA sequence, and within the gRNA sequence, nucleotide transitions are often acceptable without loss of function. If we assume a single gRNA sequence per minicircle sequence class, then this limited sequence similarity may have been overlooked in previous analyses.

In conclusion, we have extended the original concept of gRNA editing of maxicircle transcripts to include gRNAs derived from minicircles as well as from maxicircles. The evolutionary origin of a dichotomy of gRNA genetic information between two classes of DNA molecules in the single kinetoplastid mitochondrion is obscure but fascinating. Minicircles may have evolved by duplication of the original maxicircle-encoded gRNA genes in the form of closed circular self-replicating plasmids, which remained catenated with maxicircle DNA and with each other and aided in the partitioning of mitochondrial DNA to the two daughter mitochondria at mitosis, in addition to providing gRNA transcripts. Alternatively, minicircle-encoded gRNA genes may represent the more primitive state, and some may have integrated into maxicircle DNA in the course of evolution. It is clear that many questions remain to be answered regarding the function and evolution of this unusual organelle genetic system.

Experimental Procedures

Cell Culture, Mitochondrial Isolation, and kRNA Isolation

L. tarentolae cells (UC strain) were grown as described (Simpson and Braly, 1970). Cells at mid log phase were used for mitochondrial isolations. The kinetoplast mitochondrial fraction and total kRNA were isolated as described (Simpson and Simpson, 1978).

Oligonucleotide Probes and Primers

Oligonucleotides were synthesized and purified as described previ-

ously (Blum et al., 1990). Oligonucleotides were 5' end-labeled using T4 polynucleotide kinase (Bethesda research Laboratories) and [γ - 32 P] ATP (>7000 Ci/mmol; ICN) as described (Blum et al., 1990). The following oligonucleotides were used for this study (the gRNA, the nucleotide localization, and the GenBank sequence entry are provided in parentheses):

S-198 (CYb gRNA-II): 5'-CTCATGTAAACCTTTGTTAC (LEIKP-MAX = L.t. maxicircle, nucleotides 2250–2270)

S-252 (4E-2 gRNA): 5'-ACACTCTATGATTCTCTAT (LEIKPMNCC = Lt154 minicircle, nucleotides 102–121)

S-255 (COIII gRNA-I): 5'-GAATAGTGTTCATCTCTC (LEIKPMCC = D-12 minicircle, nucleotides 125–144)

S-260 (antisense COIII gRNA-I): 5'-GATGAAAACACTATTCGTAG (LEIKPMCC, nucleotides 121–140)

Guide nucleotides are underlined.

Northern and Southern Blot Analysis and Primer Extension Sequence Analysis

All procedures were performed as described previously (Blum et al., 1990). The high resolution acrylamide gel consisted of 8% acrylamide–8 M urea run for 9 hr at 50 mA; the stained gel was electroblotted onto a Nytran filter as described previously (Simpson et al., 1989).

Computer Analysis

This was performed as described previously (Blum et al., 1990). L. tarentolae minicircle sequences were down-loaded from GenBank for analysis.

Acknowledgments

This research was supported in part by a grant to L. S. from the National Institutes of Health (AI09102). N. R. S. was a predoctoral trainee on NIH training grant GM07104 and was supported in part by the Dr. Ursula Mandel Scholarship. We would like to thank Dr. B. Blum, Dr. A. Simpson, and Dr. N. Bakalara for helpful discussions and technical advice. We also thank Dr. A. Simpson for allowing the use of her electron micrograph of kinetoplast DNA for the cover of this issue.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 14, 1990; revised April 5, 1990.

References

- Barrois, M., Riou, G., and Galibert, F. (1982). Complete nucleotide sequence of minicircle kinetoplast DNA from *Trypanosoma equiperdum*. Proc. Natl. Acad. Sci. USA 78, 3323–3327.
- Blum, B., Bakalara, N., and Simpson, L. (1990). A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. Cell 60, 189–198.
- Borst, P., Fase-Fowler, F., and Gibson, W. (1987). Kinetoplast DNA of *Trypanosoma evansi*. Mol. Biochem. Parasitol. 23, 31–38.
- Donelson, J., Majiwa, F., and Williams, R. (1979). Kinetoplast DNA minicircles of *Trypanosoma brucei* share regions of sequence homology. Plasmid 2, 572–588.
- Feagin, J. E., Abraham, J. M., and Stuart, K. (1988). Extensive editing of the cytochrome c oxidase III transcript in *Trypanosoma brucei*. Cell 53, 413–422.
- Fouts, D., and Wolstenholme, D. (1979). Evidence for a partial transcript of the small circular component of kinetoplast DNA of *Crithidia acanthocephali*. Nucl. Acids Res. 6, 3785–3804.
- Frasch, A., Hajduk, S., Hoeijmakers, J., Borst, P., Brunel, F., and Davison, J. (1980). The kDNA of *Trypanosoma equiperdum*. Biochim. Biophys. Acta. 607, 397–410.
- Kidane, G., Hughes, D., and Simpson, L. (1984). Sequence heterogeneity and anomalous electrophoretic mobility of kinetoplast minicircle DNA in *Leishmania tarentolae*. Gene 27, 265–277.
- McCutchan, T., Hansen, J., Dame, J., and Mullins, J. (1984). Mung bean nuclease cleaves *Plasmodium* genomic DNA at sites before and after genes. Science 225, 625–628.
- Morel, C., Chiari, E., Camargo, E., Mattei, D., Romanha, A., and Simpson, L. (1980). Strains and clones of *Trypanosoma cruzi* can be characterized by restriction endonuclease fingerprinting of kinetoplast DNA minicircles. Proc. Natl. Acad. Sci. USA 77, 6810–6814.
- Muhich, M., and Simpson, L. (1986). Specific cleavage of kinetoplast minicircle DNA from *Leishmania tarentolae* and identification of several additional minicircle sequence classes. Nucl. Acids Res. 14, 5531–5556.
- Ray, D. (1989). Conserved sequence blocks in kinetoplast DNA minicircles from diverse species of trypanosomes. Mol. Cell. Biol. 9, 1365–1367.
- Rohrer, S. P., Michelotti, E. F., Torri, A. F., and Hajduk, S. L. (1987). Transcription of kinetoplast DNA minicircles. Cell 49, 625–632.
- Simpson, L. (1986). Kinetoplast DNA in trypanosomid flagellates. Int. Rev. Cytol. 99, 119–179.
- Simpson, L., and Braly, P. (1970). Synchronization of *Leishmania tarentolae* by hydroxyurea. J. Protozool. 17, 511–517.
- Simpson, L., and Shaw, J. (1989). RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. Cell 57, 355–366.
- Simpson, L., and Simpson, A. (1978). Kinetoplast RNA from *Leishmania tarentolae*. Cell 14, 169–178.
- Simpson, A., Neckelmann, N., de la Cruz, V., Muhich, M., and Simpson, L. (1985). Mapping and 5' end determination of kinetoplast maxicircle gene transcripts from *Leishmania tarentolae*. Nucl. Acids Res. 13, 5977–5993.
- Simpson, A., Suyama, Y., Dewes, H., Campbell, D., and Simpson, L. (1989). Kinetoplastid mitochondria contain functional tRNAs which are encoded in nuclear DNA and also contain small minicircle and maxicircle transcripts of unknown function. Nucl. Acids Res. 17, 5427–5445.
- Stuart, K. (1983). Kinetoplast DNA, a mitochondrial DNA with a difference. Mol. Biochem. Parasitol. 9, 93–104.
- Sturm, N., and Simpson, L. (1990). Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. Cell 61, this issue.