

Partially Edited mRNAs for Cytochrome b and Subunit III of Cytochrome Oxidase from *Leishmania tarentolae* Mitochondria: RNA Editing Intermediates

Nancy R. Sturm* and Larry Simpson*†

* Department of Biology

† Molecular Biology Institute

University of California

Los Angeles, California 90024

Summary

Partially edited mRNAs were selected by the polymerase chain reaction and sequenced. In the case of cytochrome b, 102 out of 106 clones displayed patterns of editing that were consistent with a strictly progressive 3' to 5' editing process, as predicted by the guide RNA model of RNA editing. In the case of cytochrome oxidase subunit III (COIII), 177 out of 304 clones displayed strictly progressive 3' to 5' patterns of editing. However, the remaining 127 COIII clones displayed unexpected patterns in which upstream editing preceded downstream editing, uridines were inserted at sites not normally edited, and purine residues were deleted. We suggest that many of these RNAs are produced by normal 3' to 5' editing of the COIII mRNA with incorrect guide RNA molecules.

Introduction

RNA editing in kinetoplastid mitochondria involves the precise insertion or, less frequently, deletion of uridine nucleotides (U's) within coding regions (and in some cases noncoding regions), thereby creating translatable mRNAs (reviewed in Simpson and Shaw, 1989; Benne, 1989). A model has been proposed for RNA editing of at least four maxicircle cryptogenes (cytochrome b [CYb], cytochrome oxidase subunit II [COII], maxicircle unidentified reading frame 2, and maxicircle unidentified reading frame 3 [MURF3]), which involves the participation of maxicircle-encoded guide RNA (gRNA) molecules in a process of progressive cycles of hybridization at the 3' side of the preedited region, endonucleolytic cleavage at the first unpaired base, addition or deletion of U's, and religation (Blum et al., 1990). A prediction of this model is that editing progresses in a strict 3' to 5' direction within each editing block, which is defined by the specific gRNA utilized.

The evidence currently available for a 3' to 5' progressive editing process is the existence of partially edited mRNAs for the cytochrome oxidase subunit III (COIII) gene of *Trypanosoma brucei*, which were interpreted as being intermediates in editing (Abraham et al., 1988). These RNAs were selected by polymerase chain reaction (PCR) amplification using a downstream primer that is edited and an upstream primer that is unedited within the preedited region. The products were analyzed by hybridization to edited and unedited probes. The fact that the reverse selection did not yield products suggested that editing progresses in the 3' to 5' direction.

A few cDNA sequences of partially edited *T. brucei* COIII mRNAs have been reported (Abraham et al., 1988). The junction sequences between the edited and unedited primers are complex and can cover up to six editing sites. Within this junction region, upstream editing in some cases precedes downstream editing and extra U's are found both at normal editing sites and, in one case, at a site not normally edited.

cDNAs truncated within the edited region have also been isolated, without selection, for the CYb and the COIII genes of *T. brucei* (Feagin et al., 1987; Abraham et al., 1988). The 5' termini of these cDNAs lie within editing sites and consist of long stretches of thymidine residues (U's in the RNA), but the cDNAs otherwise have normal editing patterns. These may represent cDNA cloning artifacts, however, since runs of thymidine residues have also been found at the 5' termini of several cDNAs of the normally unedited COI and 9S genes of *Crithidia fasciculata* (Van der Spek et al., 1990).

Whether partially edited COIII RNAs represent true intermediates in the editing process or merely some type of aberrant recombinants between edited and unedited RNAs is an important question in terms of the mechanism of editing. It is also relevant to ask whether these types of partially edited RNAs are limited to "pan-edited" cryptogenes such as the *T. brucei* COIII gene, or also occur in the cases of the "5'-edited" and "internal-edited" cryptogenes. These three types of editing differ mainly in the extent and localization of uridine addition (Simpson and Shaw, 1989); mRNAs from the three pan-edited cryptogenes in *T. brucei* exhibit uridine additions and deletions throughout the entire coding portions of the transcripts, whereas mRNAs from the 5'-edited cryptogenes have 30–40 uridine additions (and deletions in some cases) limited to the 5' regions of the transcripts, and mRNAs from the internal-edited cryptogenes have 4 or 5 uridine additions at three sites located at an internal frameshift.

We have addressed these questions and have analyzed the kinetics of the editing process in detail for two 5'-edited cryptogenes of *L. tarentolae*—CYb and COIII—by sequencing multiple PCR-derived clones of partially edited mRNAs and attempting to arrange these sequences into a series consistent with the gRNA editing model.

Results

Amplification of Partially Edited Transcripts

Two edited and two unedited PCR primers each overlapping three to four editing sites at the 3' and 5' portions of the edited regions of the CYb and COIII mRNAs (Figure 1) were used for PCR amplification from either mitochondrial RNA or total cell RNA. The amplified products were cloned and sequenced.

Two basic types of sequences of partially edited mRNAs were obtained by use of the 3'-edited:5'-unedited primer sets. The first type showed a strict 3' to 5' editing direction-

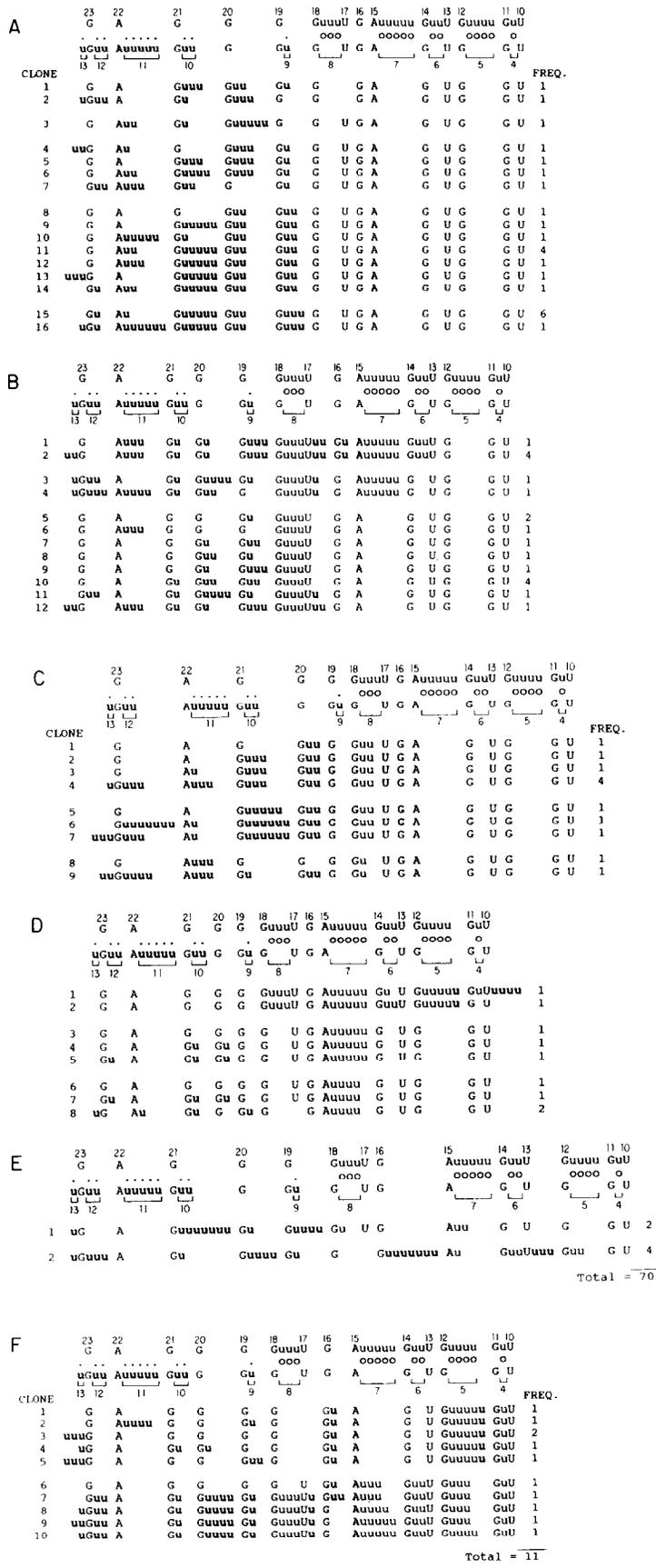


Figure 4. Alignments of Unexpected Partially Edited COIII Sequences without Purine Deletions

(A-E) The sequences generated using the 3'-edited:5'-unedited primer set have been subgrouped according to common 3' patterns and extent of editing. For each subgroup, the preedited and fully edited junction region sequence for COIII are indicated above the alignments. See legends of Figures 1 and 2 for details of labeling.

(F) Unexpected partially edited COIII mRNA sequences obtained by PCR amplification using the 3'-unedited:5'-unedited primer set. Out of 30 clones sequenced, 19 were completely unedited and 11 clones (shown here) were partially edited.

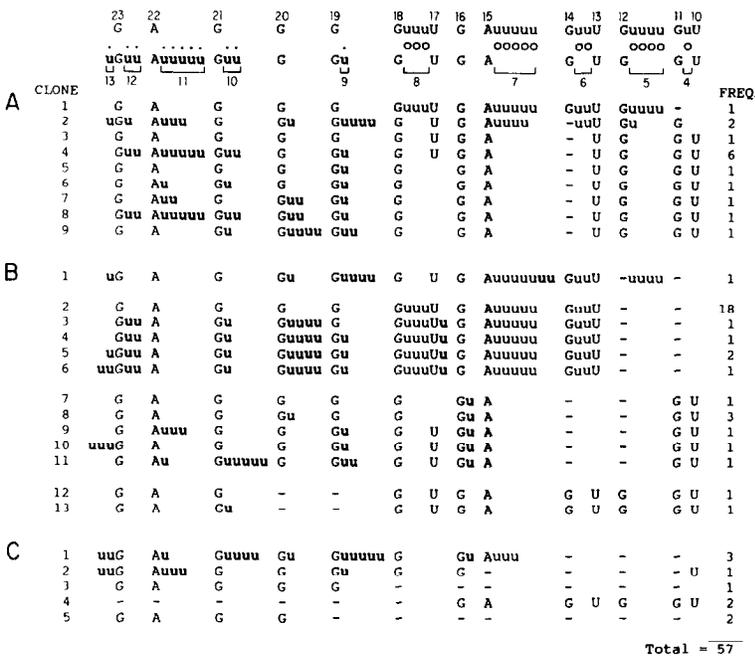


Figure 5. Alignments of Unexpected Partially Edited COIII Sequences with Purine Deletions. Deletions of purines are indicated by dashes. Deletions of encoded U's are indicated by empty spaces, because of the ambiguity involved in determining whether a particular U was encoded or deleted and then inserted. The precise placement of the deletions is ambiguous in some cases, and deletions have been located in a manner to minimize the apparent editing complexity.

or additions at downstream sites and insertion of U's at sites not normally edited. Two additional types of unexpected sequence patterns were observed in the COIII case: the addition of extra U's at normal sites (Figure 4) and the deletion of purines (Figure 5). Certain 3' deletions occurred more frequently than others. The rare massive deletions (Figure 5C, clones 3-5) might be PCR artifacts, but the specific deletions of G(14), and the two groups of U(10), G(11), and G(12) and G(12), U(13), and G(14) occur too frequently to represent PCR artifacts, although this possibility cannot be eliminated.

PCR products obtained by amplification using the COIII 3'-unedited:5'-unedited primer set represented a mixture of completely unedited clones (19/30) and unexpected partially edited clones (11/30), as shown in Figure 4F. This evidence suggests that the relative abundance of partially edited COIII mRNAs is approximately equivalent to that of unedited COIII mRNA.

Purine Deletions Occur during Mitochondrial Isolation

The percentage of COIII clones with purine deletions varied significantly with the RNA preparation used for PCR amplification. As shown in Table 1, 47% of the clones from one kRNA preparation had purine deletions, whereas only 7% of the clones from another kRNA preparation and 14% of the clones from a total cell RNA preparation had purine deletions. The percentage of unexpected clones without purine deletions was approximately identical for all RNA preparations. We interpret these results as indicating an increase in the steady-state abundance of the purine deletions in partially edited COIII mRNAs as a result of some uncontrolled parameter during the process of organelle isolation for kRNA preparation number 2. It should be noted that no purine deletions were observed in the 106

partially edited CYb RNA clones sequenced, suggesting that the generation of purine deletions could be related to the specific editing process of the COIII mRNAs, in which there is a large U deletion block.

Kinetic Implications from Clone Frequencies of Partially Edited mRNAs
CYb

Assuming unbiased processes of cDNA synthesis, PCR amplification, and cloning, the appearance of a higher frequency of clones for a particular partial editing pattern could be interpreted as indicating a greater relative abundance of these RNAs in the original steady-state RNA population. In the case of the expected CYb editing patterns, statistically significant accumulations of clones were found for sequence patterns A-1, A-3, and A-8, as shown in the histogram in Figure 2A. We attribute the apparent increased relative abundance of specific partially edited sequences to pauses in the editing process at

Table 1. COIII Intermediates Amplified from Different RNA Preparations

Type	Total Cell RNA (%) ^a	Kinetoplast RNA Preparation 1 (%) ^b	Kinetoplast RNA Preparation 2 (%) ^c
Expected	66	69	28
Unexpected, without purine deletions	19	24	24
Unexpected, with purine deletions	14	7	47

^a 77 clones were sequenced.

^b 153 clones were sequenced.

^c 74 clones were sequenced.

those particular sites. The pause after the addition of three U's at site 5 (clone A-3) is consistent with the predicted transition from gRNA-I-mediated editing to gRNA-II-mediated editing in the RNA editing model for the CYb mRNA (Blum et al., 1990). The pause after two U additions to site 8 (clone A-8) does not correspond to any predicted transition. However, it should be noted that all of the four unexpected CYb sequence patterns show only partial editing at site 8 (Figure 2B), suggesting a difficulty with the editing of this site, perhaps due to the relatively high number of U's to be added.

COIII

Owing to the greater number of COIII clones examined, the spectrum of expected partially edited RNAs is more complete than that obtained for CYb (Figure 3). Only two of the possible expected partially edited intermediates are lacking in the U addition region in this data set (between clones 20 and 21 and clones 22 and 23). Statistically significant accumulations of clones with sequence patterns 1, 6, 13, 16, 24, and 25 suggest a stall before the editing of site 4 and after the complete editing of sites 5, 7, 8, 12, and 13. The accumulation of clones containing pattern 16 is particularly striking, perhaps indicating a shift in the putative gRNA guiding editing through the deletion region to another guiding the upstream additions. The relatively high number of clones containing partial editing of the deletion block (sites 4–8) suggests that deletion of U's may occur more slowly than addition of U's in the editing process.

Nonrandom Occurrence of Certain Partially Edited Patterns

The unexpected sequence patterns generated with the 3'-edited:5'-unedited primer set could be aligned into subgroups (Figures 4A–4E and 5). The subgroupings were derived primarily from analysis of the 3'-terminal editing patterns. The Figure 4A patterns exhibit complete deletions in sites 4 through 8 and varying amounts of upstream additions. The patterns in Figure 4B exhibit complete or partial deletions at sites 4–7 and no deletions and some excess additions at site 8, with variable numbers of additions upstream. The patterns in Figure 4C all exhibit complete deletions at sites 4 through 7 and partial deletions at site 8. The Figure 4D patterns include a progression of various additions and deletions at sites 4 through 8 and varying upstream additions. Figure 4E contains two unique patterns.

The unexpected sequences generated with the 3'-unedited:5'-unedited primer set shown in Figure 4F also can be placed into two subgroups of related patterns.

As is the case with the expected partially edited sequence patterns, the unexpected sequence patterns frequently could be aligned in terms of a progressive 3' to 5' model (e.g., Figure 4A-8, -9, -11, and -14, Figure 4C-1 through -4, and Figure 4D-3 through -5).

Addition of U's was found to occur between every nucleotide pair, including at sites not edited in the mature RNA, specifically, G(19)/G(20) and A(15)/G(16). Deletion of U's that are present in the preedited mRNA and normally present in the correctly edited mature mRNA occurred in

several cases. For example, U(17) was deleted in eight clones (Figures 4A, 4D, and 4E). In line with the gRNA editing model, we propose that incorrect U additions and U deletions are correctable by U additions in subsequent cycles of editing with the appropriate gRNA.

There are several patterns that could represent reinitiation of the editing process at the 3' editing terminus with the correct gRNA sequence information, after a complete cycle of incorrect editing has already occurred. For example, patterns B-2 and B-12 in Figure 4 possess identical upstream unexpected editing profiles in the addition region, but the downstream deletion regions vary, as if another cycle of editing had initiated 3' and had progressed, in the case of pattern B-12, partway through the deletion block. Other possible examples of this phenomenon are provided by patterns 4B-3 and 4B-11, patterns 4D-5 and 4D-7, and patterns 4B-3 and 5B-5.

The partially edited COIII patterns with purine deletions shown in Figure 5 fell into several distinct subgroups. Specific deletion of G(11) and G(12) (and U[10]) occurred in 24 clones, deletion of G(14) in 14 clones, and deletion of G(12) and G(14) (with deletion of U[13]) in seven clones. The purine deletions occur almost exclusively in the 3' region where multiple U deletions also occur. We assume that deletions of purines are not correctable by additional cycles of gRNA editing and that these molecules represent dead ends in the process.

The various unexpected partially edited sequence patterns in the COIII mRNAs occurred in a nonrandom distribution in our data set. The nonrandom distribution of certain sequence patterns was confirmed statistically by performing a chi-square test on the observed frequencies of the unexpected COIII sequences: the probability of randomly generating these collections of sequences was less than 0.001%. We propose that the majority of the unexpected COIII partially edited patterns are generated not by misediting with the correct gRNA template but rather by accurate editing with incorrect gRNAs. The discovery of a minicircle-encoded gRNA sequence specifying the precise editing information for one of the unexpected COIII clones strengthens this argument (Sturm and Simpson, 1990).

Discussion

We have examined the sequences of the unedited–edited junction regions of a large number of independent PCR-amplified cDNA clones of partially edited mRNAs for the CYb and the COIII genes. The fact that a series of partially edited RNA sequences was obtained for both genes that could be arranged in a strictly progressive 3' to 5' fashion is consistent with the interpretation that these RNAs are intermediates in editing, and suggests that editing is occurring in a 3' to 5' direction a single site at a time, and, within the site, a single nucleotide at a time. The evidence is particularly strong in the case of the CYb gene, for which 96% of the clones obtained were of the expected type. The observed differences in relative steady-state abundances of several of the partially edited transcripts can be interpreted as pauses in the editing process, such

as that created by the putative shift from CYb gRNA-I- to CYb gRNA-II-mediated editing. This interpretation must be a tentative one since it is necessarily dependent on the assumption that the PCR amplification and cloning processes are unbiased. It is possible that the PCR technique preferentially amplifies certain RNAs owing to secondary structures, although this is unlikely in view of the 50°C incubation temperature for the reverse transcriptase reaction and the 55°C annealing and 72°C elongation temperatures in the PCR reaction.

Editing of the COIII mRNA appears to represent a less specific process than editing of the CYb mRNA. Approximately 42% of the 304 partially edited cDNA clones examined showed patterns that varied significantly from a strictly progressive 3' to 5' fashion. Most of these unexpected patterns, however, fell into distinct subsets. We have proposed an explanation for the appearance of a majority of the unexpected COIII mRNAs: these are specific partially edited products resulting from interaction of incorrect gRNA molecules with the COIII mRNA. Some suggestive evidence was also obtained from our analysis for the reediting of incorrectly edited COIII mRNAs with the proper gRNAs. The generation of partially edited mRNAs with purine deletions is thought to represent the only incorrigible error in this process, and this may be accentuated at some uncontrolled step in the mitochondrial isolation process, at least in the case of the COIII mRNAs.

One possible explanation for the hypothesized decrease in specificity of gRNA-mRNA hybridization in the case of the COIII mRNA is a consequence of G-U base pairing. Owing to the ability of G residues to pair with either C or U, and U residues with either A or G, an ambiguity is introduced into the hybridization potential of preedited, partially edited, or fully edited mRNAs. A and C residues in the mRNA can form base pairs only with U or G residues, respectively. However, G or U residues in the mRNA represent positions that could accept nucleotide transitions in the gRNA. An examination of the analyzed junction region of COIII partially edited mRNA shows only two A residues, whereas the CYb mRNA junction region shows nine A residues (neither contains C residues).

We show in the accompanying paper (Sturm and Simpson, 1990) that a major difference between the editing of the CYb mRNA and the COIII mRNA is that the COIII editing involves a minicircle-encoded gRNA for sites 1-8. We also show that there is another minicircle-encoded gRNA specifying the editing sequence information for the unexpected COIII partially edited RNA in Figure 4E-2, as predicted by our hypothesis for the generation of unexpected editing patterns (Sturm and Simpson, 1990).

The recent *in vivo* pulse-chase results of Volloch et al. (1990) on the synthesis of unedited and pan-edited COIII mRNA in *T. brucei* appear to indicate that there is no precursor-product relationship between unedited mRNA and edited mRNA, which would render our interpretation of the partially edited mRNAs untenable. However, these results are based on the conclusion that actinomycin D *in vivo* selectively inhibits maxicircle transcription of un-

edited RNAs and not the synthesis of edited RNAs, and this was not clearly demonstrated. We feel that the most parsimonious interpretation of our steady-state RNA data together with the additional evidence presented previously (Blum et al., 1990) is that the partially edited CYb and COIII molecules represent intermediates in a post-transcriptional editing process, as predicted by the gRNA model.

Experimental Procedures

Kinetoplast Isolation and RNA Extraction

A kinetoplast mitochondrial fraction was isolated from *L. tarentolae* (UC strain) cells grown in Difco Brain Heart Infusion medium (Simpson and Braly, 1970) to a density of $70-150 \times 10^6/\text{ml}$ by isopycnic Renografin density gradients as previously described (Simpson and Simpson, 1978). kRNA was isolated as described (Simpson and Simpson, 1978) and extensively treated with RNAase-free DNAase I to remove kinetoplast DNA. Total cell RNA was isolated as described (Simpson and Simpson, 1978).

PCR Amplification of RNA and Cloning

The primers were synthesized by standard phosphoramidite methods and contain EcoRI and BamHI sites for forced cloning into pGEM-7Zf, as indicated in Figure 1. Amplification by PCR was performed by creating single-stranded DNA templates from approximately 1 μg of kRNA by extending the appropriate 3' primer (0.5-1.0 μg) with AMV reverse transcriptase at 50°C for 2 min. Heat inactivation of the enzyme was followed by addition of Taq polymerase and the appropriate 5'-unedited primer. The reaction mix consisted of 50 mM KCl, 10 mM Tris-HCl (pH 8.4-8.8), 6-10 mM MgCl₂, 1 mM DTT, and the four dNTPs at 1-1.5 mM (Sturm et al., 1989). Twenty-five to 30 PCR cycles were routinely performed in a Thermal Cycler (Perkin Elmer/Cetus) with a step cycle profile of 15 s at 94°C, 15 s at 55°C, and 1 min at 72°C. The PCR products were then digested with EcoRI and BamHI, purified in low melting agarose, and cloned into the pGEM-7Zf plasmid vector. Multiple PCR amplifications and multiple clonings were performed to obtain the sequences analyzed.

Sequencing

Single transformants were grown and the plasmid DNA was isolated by a boiling miniprep procedure. Individual clones were sequenced using the Sequenase kit (United States Biochemical). A and G ladders were initially obtained for CYb and G and T ladders for COIII. Complete sequencing was performed in ambiguous cases.

The clones were separated into "expected" and "unexpected" subgroups, where the former sequences followed a strict 3' to 5' pattern progressively resembling the known mature sequence (Shaw et al., 1988). Within each subgroup the sequences were ordered based primarily on the number of sites affected, and secondarily on the number of U's affected.

Statistical Analysis

Error bars were determined for the expected frequencies at $\pm 1.5 \times$ (standard error). For the unexpected clones, a standard chi-square test was performed, comparing any single site relative to the other possible editing sites. The editing of each site was scored as plus or minus, and the quality of editing was not taken into account for this initial test. The observed frequencies were tested against those expected by random chance.

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 Mary Leonora Schulte Memorial Scholarship. We would like to thank Dr. N. Bakalara, Dr. B. Blum, Dr. A. Simpson, Dr. J. Shaw, and Eileen Gruszynski for helpful discussions and technical advice, and Niles Lehman for statistical advice.

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

- Abraham, J. M., Feagin, J. E., and Stuart, K. (1988). Characterization of cytochrome c oxidase III transcripts that are edited only in the 3' region. *Cell* 55, 267-272.
- Benne, R. (1989). RNA editing in trypanosome mitochondria. *Biochim. Biophys. Acta* 1007, 131-139.
- 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.
- Feagin, J. E., Jasmer, D. P., and Stuart, K. (1987). Developmentally regulated addition of nucleotides within apocytochrome b transcripts in *Trypanosoma brucei*. *Cell* 49, 337-345.
- Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1988). Creation of AUG initiation codons by addition of uridines within cytochrome b transcripts of kinetoplastids. *Proc. Natl. Acad. Sci. USA* 85, 539-543.
- Shaw, J. M., Feagin, J. E., Stuart, K., and Simpson, L. (1988). Editing of kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons. *Cell* 53, 401-411.
- Shaw, J. M., Campbell, D., and Simpson, L. (1989). Internal frameshifts within the mitochondrial genes for cytochrome oxidase subunit II and maxicircle unidentified reading frame 3 of *Leishmania tarentolae* are corrected by RNA editing: evidence for translation of the edited cytochrome oxidase subunit II mRNA. *Proc. Natl. Acad. Sci. USA* 86, 6220-6224.
- Simpson, L., and Braly, P. (1970). Synchronization of *Leishmania tarentolae* by hydroxyurea. *J. Protozool.* 17, 511-517.
- Simpson, L., and Shaw, J. M. (1989). RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. *Cell* 57, 355-366.
- Simpson, L., and Simpson, A. M. (1978). Kinetoplast RNA from *Leishmania tarentolae*. *Cell* 14, 169-178.
- Sturm, N., and Simpson, L. (1990). Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. *Cell* 61, this issue.
- Sturm, N., Degrave, W., Morel, C., and Simpson, L. (1989). Sensitive detection and schizodeme classification of *Trypanosoma cruzi* cells by amplification of kinetoplast minicircle DNA sequences: use in diagnosis of Chagas disease. *Mol. Biochem. Parasitol.* 33, 205-214.
- Van der Spek, H., Speijer, D., Arts, G.-J., Van den Burg, J., Van Steeg, H., Sloof, P., and Benne, R. (1990). RNA editing in transcripts of the mitochondrial genes of the insect trypanosome *Crithidia fasciculata*. *EMBO J.* 9, 257-262.
- Volloch, V., Schweitzer, B., and Rits, S. (1990). Uncoupling of the synthesis of edited and unedited COIII RNA in *Trypanosoma brucei*. *Nature* 343, 482-484.