MOLECULAR AND CELLULAR BIOLOGY, Jan. 1992, p. 56–67
0270-7306/92/010056-12$02.00/0

Vol. 12, No. 1

An Intergenic G-Rich Region in *Leishmania tarentolae* Kinetoplast Maxicircle DNA Is a Pan-Edited Cryptogene Encoding Ribosomal Protein S12

DMITRI A. MASLOV,[1] NANCY R. STURM,[1] BETTINA M. NINER,[1] EILEEN S. GRUSZYNSKI,[1] MARIAN PERIS,[1] AND LARRY SIMPSON[1,2]*

Department of Biology[1] and Molecular Biology Institute,[2] University of California, Los Angeles, California 90024

Received 29 July 1991/Accepted 2 October 1991

Six short G-rich intergenic regions in the maxicircle of *Leishmania tarentolae* are conserved in location and polarity in two other kinetoplastid species. We show here that G-rich region 6 (G6) represents a pan-edited cryptogene which contains at least two domains edited independently in a 3′-to-5′ manner connected by short unedited regions. In the completely edited RNA, 117 uridines are added at 49 sites and 32 uridines are deleted at 13 sites, creating a translated 85-amino-acid polypeptide. Similar polypeptides are probably encoded by pan-edited G6 transcripts in two other species. The G6 polypeptide has significant sequence similarity to the family of S12 ribosomal proteins. A minicircle-encoded gRNA overlaps 12 editing sites in G6 mRNA, and chimeric gRNA/mRNA molecules were shown to exist, in agreement with the transesterification model for editing.

The mRNA transcripts of several mitochondrial genes in kinetoplastid protozoa are modified by a novel posttranscriptional process known as RNA editing, in which uridine residues are added to and occasionally deleted from coding regions (2, 3, 44, 48, 52). RNA editing, in effect, is a mechanism to decipher the primary transcripts of the six known cryptogenes—cytochrome *b*, cytochrome oxidase subunits II and III (COII and COIII), NADH dehydrogenase subunit 7 (ND7), and unidentified reading frames 2 (*MURF2*) and 4 (*MURF4*)—in such a way that functional translation products can be obtained by the mitochondrial translation system. However, the mitochondrial genomic information content of kinetoplastids is still somewhat limited compared with that found in other eukaryotic cells such as yeast and human cells: no F_1F_0 ATPase subunit genes and no mitochondrial tRNA genes (16, 40) have been identified in the maxicircle genome. Identification of the *MURF4* sequence as subunit 6 of the F_1 ATPase has been suggested (4), but the alignments with known ATPase subunit 6 sequences are not statistically significant (45).

Extensive or pan-editing of the ND7, COIII, and *MURF4* maxicircle cryptogenes is limited to *Trypanosoma brucei* (4, 11, 22). In *Leishmania tarentolae* and *Crithidia fasciculata* (48), the ND7 mRNA is both 5′ edited and internally edited, and the *MURF4* mRNA is pan-edited only in the 5′ portion (4). Originally it seemed that pan-editing was limited to the African trypanosomes. However, on the basis of the distinctive G+A-rich character of the preedited regions of known cryptogenes, Simpson and Shaw (48) hypothesized that six short G-rich maxicircle intergenic regions which were conserved in relative location and polarity, but not in sequence, in *L. tarentolae*, *T. brucei* and *C. fasciculata* actually represented pan-edited cryptogenes encoding additional mitochondrial genetic information. Heterogeneously sized transcripts from these regions had previously been identified in *T. brucei* (18).

We show in this report that this hypothesis is correct for G-rich region 6 (G6) in *L. tarentolae*, and we show that the putative translated polypeptide from the G6 edited mRNA is conserved in two other kinetoplastid species and is homologous to the family of ribosomal S12 proteins (RPS12). In addition, we have identified one minicircle-encoded guide RNA (gRNA) that overlaps 12 editing sites in the G6 mRNA.

MATERIALS AND METHODS

Cell culture, mitochondrial isolation, and kRNA Isolation. *L. tarentolae* (UC strain) cells were grown as described previously (46). Kinetoplast DNA networks were isolated as described previously (41). The kinetoplast mitochondrial fraction was isolated from mid-log-phase cells by flotation in Renografin density gradients as described previously (8), and kinetoplast RNA (kRNA) was isolated (49).

Oligonucleotides. Oligonucleotides for polymerase chain reaction (PCR), primer extension, and hybridization were synthesized by standard phosphoramidite methods and purified by thin-layer chromatography. The following oligonucleotides were used in this study. G6-specific primers (the nucleotide positions in GenBank entry LEIKPMAX are given in parentheses) were as follows:

S-362: (GGATCC)GCGTGAATTTTTTGAG (14677–14692)
S-363: (GAATTC)TTATTTTAATGATTA (14877–14863)
S-395: (GAATTC)T_{23}
S-440: CCaCCaCCTTCaaCTcTaaaC (14816–14798)
S-447: CGTGGaCTTaaaTaTAaaaCCaaaC (14728–14715)
S-448: CaTaTTGCATaTaaaCaaCGTCG (14761–14745)
S-452: (GGATCC)CTAATACCTATCGACC (14645–14660)
S-454: GGCTTTTACCCTAAAATAACAAAAGCAAC (14725–14696)
S-455: CCCAAAAACAAACAACCTCTTTTCTTCTTTGC (14787–14756)
S-456: CGAAAACAATTCCGTGAAATCTTTCTCC (14851–14824)

Lt154 gRNA-specific primers (the nucleotide positions in GenBank entry LEIKPMNCC are given in parentheses) were as follows:

S-315: CAAACACAACAAAAAACA (138–121)
S-397: ATAAAACACAACAAAAAA (141–123)

* Corresponding author.

Nucleotides corresponding to U residues added by editing are given in lowercase. Nonencoded restriction sites in S-362, S-363, S-395, and S-452 are indicated by parentheses.

PCR amplification and cDNA cloning. RNA was amplified from purified kRNA by reverse transcription followed by PCR. One to two micrograms of kRNA and 0.1 to 0.5 μg of 3′ primer were annealed in 50 mM KCl–10 mM Tris-HCl (pH 8.8)–5 mM $MgCl_2$–1 mM dithiothreitol at 60°C for 2 min and then at 50°C for 1 min. Single-stranded cDNA was synthesized at 37°C for 20 min with 5 U of avian myeloblastosis virus reverse transcriptase in the presence of 2.5 mM each of four deoxynucleoside triphosphates. After addition of the 5′ primer and *Taq* polymerase, 30 PCR cycles were performed in a Thermal Cycler (Perkin Elmer/Cetus), with each cycle composed of 15 s at 94°C, 2 min at 40°C, and 2 min at 72°C.

PCR products were ligated to pCR1000 vector DNA and transformed into INV1(α)F′ competent *Escherichia coli* cells, using the TA cloning kit (Invitrogen Corp.). Transformants with G6-specific inserts were selected by colony hybridization, using the labeled 5′ primer as a probe. Primers were labeled with T4 polynucleotide kinase (Bethesda Research Laboratories) and [γ-^{32}P]ATP (>7,000 Ci/mmol; ICN). Clones with amplified chimeric gRNA154/G6 RNA were sequenced at random.

DNA sequencing. Plasmid DNA was isolated from individual transformants by a boiling minilysate method. Double-stranded DNA sequencing was performed by using the Sequenase version 2.0 DNA sequencing kit (United States Biochemical) by the protocol involving T7 sequencing primer extension labeling in the presence of [α-^{32}P]dATP (800 Ci/mmol; NEN) and resolution on 6% acrylamide–8 M urea gels.

Northern (RNA) blot analysis. kRNA (5 μg) was electrophoresed in a 1.8% formaldehyde–agarose gel as described previously (6, 39). Gels were blotted onto nylon filters (Micron Separations, Inc.) and cross-linked by UV irradiation (Stratagene). Blots were hybridized with kinased oligonucleotide probes (specific activity, 10^7 to 10^8 cpm/μg). Conditions of hybridization and washes in tetramethylammonium chloride were as described previously (39, 59).

Primer extension analysis of RNA. kRNA (3 to 5 μg) and an end-labeled oligonucleotide (50 to 500 ng) were sequentially annealed at 65, 42, and 20°C for 5 min at each temperature. A set of elongation-termination reactions was performed, using 15 U of avian myeloblastosis virus reverse transcriptase per annealing reaction at 42°C for 30 min in the presence of dideoxynucleoside triphosphates (39). Extension products were analyzed on sequencing gels.

Computer analysis. Analysis of nucleic acid and protein sequences was performed on the UCLA Life Sciences VAX computer, using the University of Wisconsin Genetics Computer Group program package and the Los Alamos SEQDP program for determining statistical significance of alignments and on an IBM personal computer using the PC/GENE software package (Intelligenetics). Hydropathy analysis was performed in PC/GENE using Kyte-Doolittle (24) hydropathy values for amino acids.

Nucleotide sequence accession number. The sequence of the fully edited G6 mRNA is deposited in GenBank under accession number M74225.

RESULTS

Sequences of partially edited G6 RNAs. To detect and analyze the putative edited RNAs from the G6 region, an approach based on the known 3′-to-5′ progression of editing was used (1, 9, 52). Partially edited RNAs were obtained by PCR amplification of total mitochondrial RNA, using a 3′ oligo(dT) primer (S-395) which would hybridize to the poly(A) [or poly(AU)] tail and a 5′ genomic sequence primer (S-362). The deduced edited consensus sequences were used for synthesis of primers for direct sequence analysis of mature edited mRNA. As shown in Fig. 1, the 5′ primer sequence was located at the 5′ side of the G6 region within the putative preedited region, and therefore PCR amplification should select for 3′-edited, 5′-unedited transcripts.

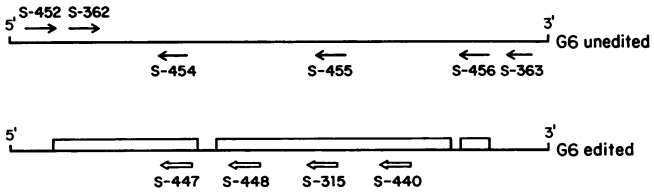

FIG. 1. Localization of oligonucleotides on G6 RNAs. Thin lines represent unedited sequence; open boxes represent the major editing domains. To facilitate the juxtaposition of the corresponding regions from both sequences, they are shown as being of equal length. Thin and thick arrows represent unedited and edited sequence-specific oligonucleotides, respectively. The directions of arrows correspond to the 5′-3′ polarity.

The PCR products were cloned, and colony hybridization was performed with the S-362 oligonucleotide. Positive clones were then size selected for the presence of inserts larger than 200 bp, and 26 clones were obtained and sequenced. The sequences are shown in three groups in Fig. 2A and B and diagrammatically in Fig. 2C.

One clone (Fig. 2A, clone 1) corresponded to unedited G6 sequence. Twenty clones in this group (clones 2 to 21) had a partial distribution of editing events with an overall 3′-to-5′ progression. To better visualize the 3′-to-5′ polarity of editing, the correctly edited regions are indicated by shadowing in Fig. 2A and B and by open boxes in Fig. 2C. Edited sequences at the 3′ ends of these clones showed perfect consistency, thereby allowing the deduction of a reliable consensus for the fully edited sequence in this region. In eight clones (clones 2, 11, 13, 14, 16, 19, 20, and 21), a portion of edited sequence was found which either differed from the consensus or contained unexpected (54) editing events which deviated from the strictly unidirectional 3′-to-5′ process. These deviations occurred in the junction regions between the correctly edited and the unedited portions of the sequences; they are underlined in Fig. 2A and B and diagrammed as thick black lines in Fig. 2C.

Three clones (clones 22 to 24) contained short, almost identical blocks of the correctly edited sequence in their 5′ portions (Fig. 2C, open boxes). Clones 23 and 24 also contained extended edited regions in their 3′ portions encompassing more than half the length of the molecule. Such an alternation of edited and unedited sequences is consistent with the existence of two separate domains of editing, as was shown previously to occur in the pan-edited ND7 gene of *T. brucei* (22). Within each domain, the pattern of editing corresponded to that observed in the first group of clones. Junction regions with unexpected edited sequences were seen at the 5′ ends of these clones. Further evidence for the presence of two separate domains of editing will be presented below.

Two clones, 25 and 26, showed extended areas of incorrectly edited sequences at the junction regions between

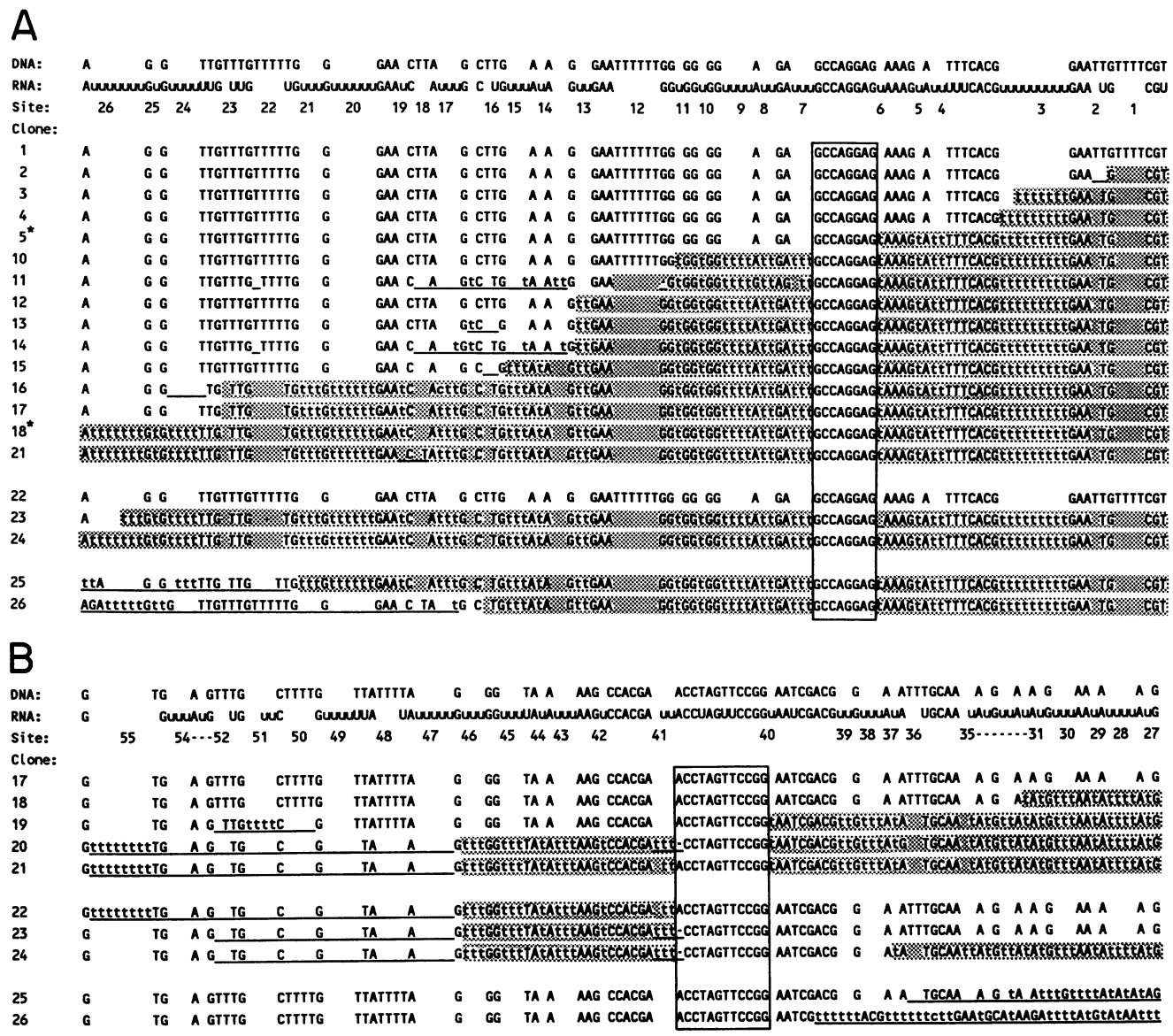

FIG. 2. Partially edited G6 cDNA sequences. kRNA was amplified by using a 3′ oligo(T) primer (S-395) and a 5′ genomic primer (S-362). The PCR products were cloned, and the clones were sequenced. Sequences of clones were aligned so as to show the 3′-to-5′ progression of expected editing events. (A) 3′-terminal portions of the cDNA sequences. Editing sites are numbered in a 3′-to-5′ direction. Correctly edited sequences are shadowed. Unexpected edited sequences are underlined. Each original clone was renumbered according to its position in the alignment. Clones with identical patterns are indicated with asterisks; clone pattern 5 was seen in five clones; clone pattern 18 was seen in three clones. The unedited domain connection sequence between the major 3′ domain and the putative 3′-terminal domain is boxed. (B) 5′-terminal portion of the cDNA sequences. Clones 1 to 16 are not presented since their sequences, like that of clone 17, are unedited in this portion. The 5′-terminal G corresponds to the last nucleotide of the primer used to amplify these intermediates. The unedited domain connection sequence between the major 5′ and 3′ domains is boxed. (C) Schematic representation of the sequences shown in panels A and B. Dashed boxes show the portions of sequences presented in panels A and B. Thin lines, unedited sequence; open boxes, correctly edited sequence; thick lines, unexpected partially edited sequence.

edited and unedited sequences. In both clones, however, correct editing has occurred at the 3′ ends.

Heterogeneity of polyadenylation sites in partially edited G6 transcripts. As shown in Fig. 3, the partially edited mRNAs had heterogeneous sites of 3′ poly(A) tail attachment. Nine clones had poly(A) tails beginning at a site 40 nucleotides (nt) downstream of the first editing site. This position is located 8 nt upstream of the 5′ end of the mRNA for the adjacent ND5 gene. Other clones showed sites of polyadenylation located more upstream within the adjacent AU-rich sequence (5, 57). The first nucleotide of the tail in many clones coincided with the position of a genomically encoded A residue. One clone contained a single uridine within the poly(A) tail and another a run of five U's between the poly(A) tail and the genomically encoded sequence, but most did not have an oligo(U) sequence preceding the poly(A) sequence as was observed by Decker and Sollner-Webb (9) with *T. brucei*. Therefore, the proposed rapid acquisition of

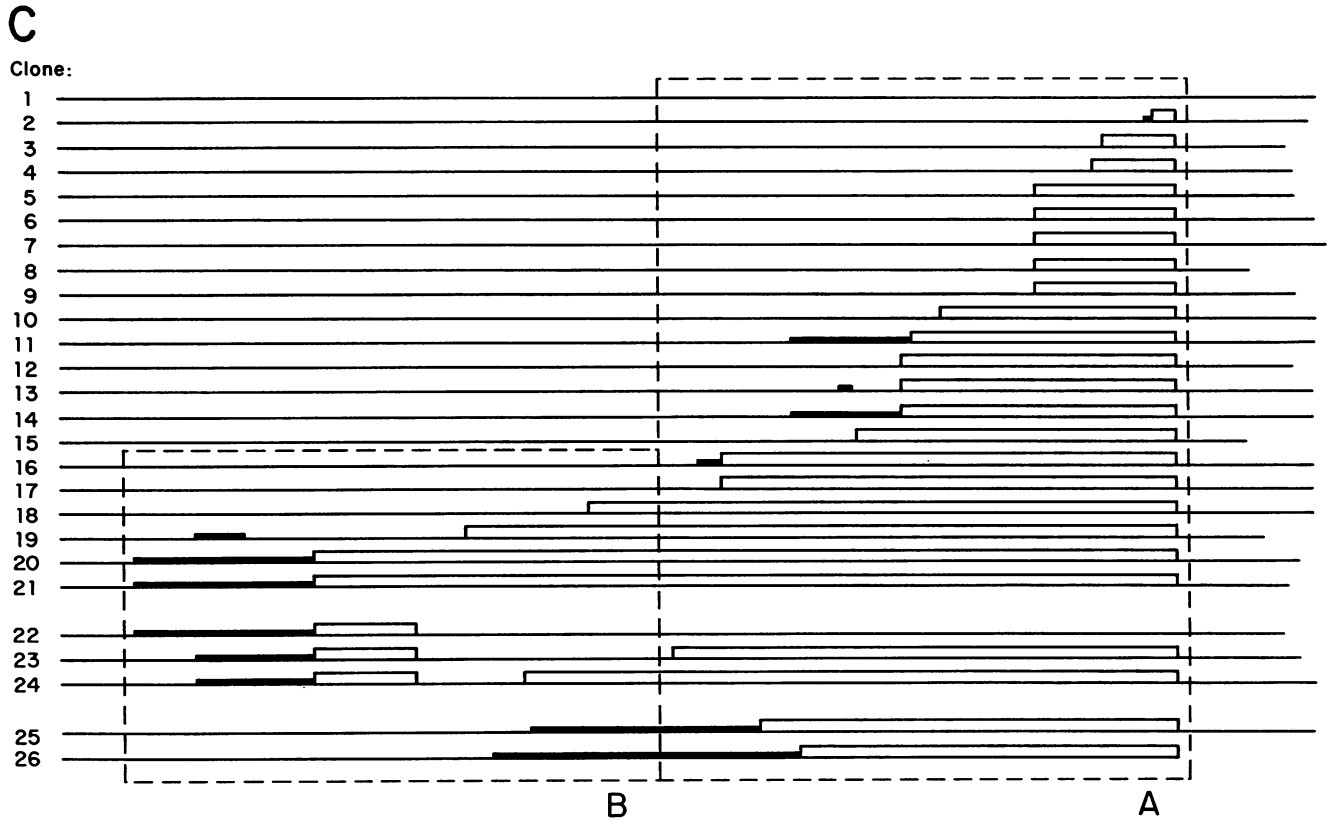

FIG. 2—*Continued.*

uridines on the 3′ end of a messenger prior to the polyadenylation might not be a universal feature.

G6 is a pan-edited cryptogene. To determine the mature edited mRNA sequence, G6 RNA was sequenced directly.

```
                                                        ND5
          2    1  Ter                              *     |
DNA:  ..GAATTGTTTTCGTAAGCAATAAGTAATCATTAAAATAATTTTATTTAGATAGTATATTGTA..
RNA:  ..GAA UG    CGUAAGCAAUAAGUAAUCAUUAAAAUAAUUUUAUUUAGAU(a)n

9     ..........................................................U(a)n
1     ..........................................................Uaaaaau(a)n
1     ..........................................................Guuuuu(a)n
1     ..........................................................Uaaag(a)n
1     ..........................................................U(a)n
1     ..........................................................Uc(a)n
4     ..........................................................U(a)n
1     ..........................................................U(a)n
1     ..........................................................U(a)n
1     ..........................................................Ugg(a)n
1     ..........................................................U(a)n
2     ..........................................................U(a)n
```

FIG. 3. Polyadenylation sites in G6 partially edited RNAs. The 3′-terminal G6 genomic sequence (nt 14840 to 14900 in LEIKPMAX) (DNA) and the edited G6 RNA sequence (RNA) are shown. 1 and 2, editing sites 1 and 2; Ter, termination codon for the 85-amino-acid polypeptide; *, position of the most frequently occurring polyadenylation site; ND5, 5′ end of the ND5 mRNA. Only the last genomically encoded nucleotide is shown for the sequences of the cDNA clones presented below. The nonencoded nucleotides are shown in lowercase. The number of clones with each pattern is given on the left.

Edited sequence at the 3′ end was confirmed using the S-363 primer, which anneals downstream of the first editing site (Fig. 1). This primer should hybridize with all G6 transcripts, unedited, edited, or partially edited. The mature edited G6 sequence was observed, indicating that the mature edited RNA is the most abundant G6 species. Edited sequence-specific primers (S-440, -315, -448, and -447) which were used for sequencing more-upstream regions gave homogeneous sequence ladders (Fig. 4) which confirmed and extended the consensus sequences of the partially edited transcripts. The position of the 5′ end, localized by primer extension (data not shown), corresponds to a C residue (nt 14645 in LEIKPMAX), which is separated by only a few nucleotides from the upstream adjacent G5 region and by 29 nt from the last G6 editing event. Some limited 5′-end heterogeneity was observed. A limited 5′-end heterogeneity was also observed for unedited G6 transcripts, using the S-455 unedited primer (data not shown). In addition, most unedited transcripts seemed to lack 50 nt from the 5′ end. The 5′ ends determined by primer extension analysis should, of course, be confirmed by S1 protection studies.

The complete sequence of mature edited G6 RNA is presented in Fig. 5A together with the genomic sequence. The putative 5′ ends of the mature transcripts are indicated by asterisks. In the fully edited transcript, 117 uridines are added and 32 uridines are deleted, thereby extending the size of the RNA from 245 to 330 nt [excluding the poly(A) tail].

There are several editing domains in the G6 cryptogene. The existence of a 5′ editing domain (defined by sites 41 to 62) is indicated by clones 22 to 24 in Fig. 2. This 5′ domain

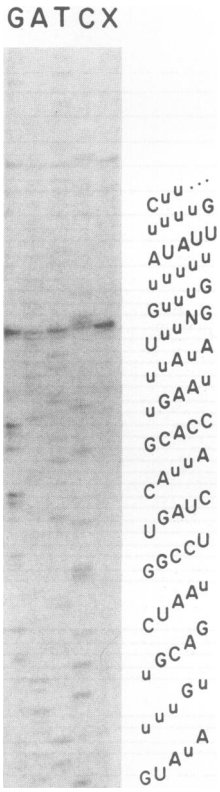

FIG. 4. Portion of G6 edited RNA sequence determined with oligonucleotide S-440. Primers were extended in the presence (lanes G, A, T, and C) and absence (lane X) of dideoxynucleotides. The sequence shown to the right of the panel represents the sequence of the corresponding RNA templates. Uridines added by editing are shown in lowercase.

is separated from the 3′ editing domain (defined by sites 1 to 40) by a 12-mer unedited sequence.

Additional evidence for the existence of separate editing domains was obtained by another selective PCR amplification of mitochondrial RNA and screening of clones with specific oligomeric probes. Two unedited primers (S-456 and S-452) (Fig. 1) were used to obtain a collection of cDNA clones. These clones were screened for editing in the 5′ domain by using the edited oligomer, S-447. The presence or absence of editing in the 3′ domain was tested by using the unedited oligomer, S-455.

The sequences of the 5′ domains of representative clones from both hybrid selections are presented in Fig. 6. Clones 1 to 9 represent sequences from 13 of the 16 clones which contained unedited 3′-domain sequences. A 3′-to-5′ progression of expected editing events (shadowed) can be seen in addition to unexpected editing patterns (underlined) at the junction regions between edited and unedited sequences. Clones 14 to 25 represent sequences from 15 of the 25 clones which contained edited 3′-domain sequences. In this case also, a 3′-to-5′ progression of expected editing events can be seen in addition to a variety of unexpected editing events at the junction regions. This evidence clearly indicates that the 5′ domain exists as a separate entity and is separated from the 3′ editing domain by the 12-mer domain connection sequence (boxed).

The existence of clones 14 to 25 which contain edited 3′-domain internal sequences but an unedited 3′-terminal sequence (due to the amplification with primer S-456) suggests that an additional 3′ editing domain covering sites 1 to 6 (Fig. 2A) may exist. This must be confirmed by additional study.

***L. tarentolae* G6 encodes an 85-amino-acid polypeptide which shows evolutionary conservation.** The mature edited G6 RNA contains an open reading frame which encodes a polypeptide of 85 amino acids (Fig. 5A). The initiation and termination codons for the G6 polypeptide are genomically encoded. Editing terminates just upstream of the initiator AUG, and the 3′ untranslated region is not edited. A hydropathy plot (see Fig. 8A) of the G6 polypeptide reveals two hydrophobic domains connected by a relatively hydrophilic amino acid sequence, which coincides with the dodecamer domain connection sequence between the 5′ and 3′ editing domains.

An open reading frame encoded by the genomic 3′-domain sequence is in frame with the edited 5′-domain-encoded open reading frame. G6 RNA which is edited only in the 5′ domain would encode a continuous open reading frame of 70 amino acids (Fig. 5B), which shares the N-terminal hydrophobic portion with the 85-amino-acid G6 polypeptide, while the rest of the sequence which is derived from the genomic sequence is more hydrophilic; the amino acid sequence of the C-terminal region showed no similarity with anything in the data bases. The corresponding transcripts have been observed after selective PCR amplification as shown above, but the abundance was too low in steady-state RNA to be detected by primer extension (data not shown).

It was shown previously that no sequence similarity could be detected by dot matrix analysis between the equivalent genomic G6 regions in *L. tarentolae* and *T. brucei* (47). However, it can be seen in Fig. 5C for three kinetoplastid species that if T residues are deleted, the genomically encoded purine and cytosine residues can be aligned. If U residues are then added to the *T. brucei* and *Crithidia oncopelti* purine backbone sequences according to their occurrence in the G6 edited RNA in *L. tarentolae*, continuous open reading frames can be obtained in both species. Wherever the purine sequences differed from the *L. tarentolae* sequence by either base substitutions or additions/deletions, the relative positions of the purines were adjusted to provide a conservative amino acid substitution where possible. The derived *T. brucei* and *C. oncopelti* polypeptides are easily aligned with the 85-amino-acid G6 polypeptide from *L. tarentolae*, and the alignments are statistically significant by the Monte Carlo shuffling test in PC/GENE (20 standard deviation units for *L. tarentolae* versus *C. oncopelti*, 25 standard deviation units for *L. tarentolae* versus *T. brucei*). Despite the somewhat arbitrary nature of the purine nucleotide alignment procedure, the similarity of the derived putative G6 amino acid sequences is highly significant. Direct sequences of the edited G6 mRNAs of the other two species are required to confirm this conclusion completely. This tentative evolutionary conservation evidence, although indirect, indicates that the 85-amino-acid G6 polypeptide is a real translation product. The evidence for the identification of this polypeptide as a homolog of RPS12 will be presented below.

A similar analysis of the putative 70-amino-acid G6 polypeptide encoded by the 5′-domain-edited, 3′-domain-unedited transcript indicates that there is no evolutionary conservation of the genomically encoded portion of this polypeptide between these three species (data not shown). This suggests that this polypeptide may not be a translation product in vivo, despite the existence of a low abundance of this partially

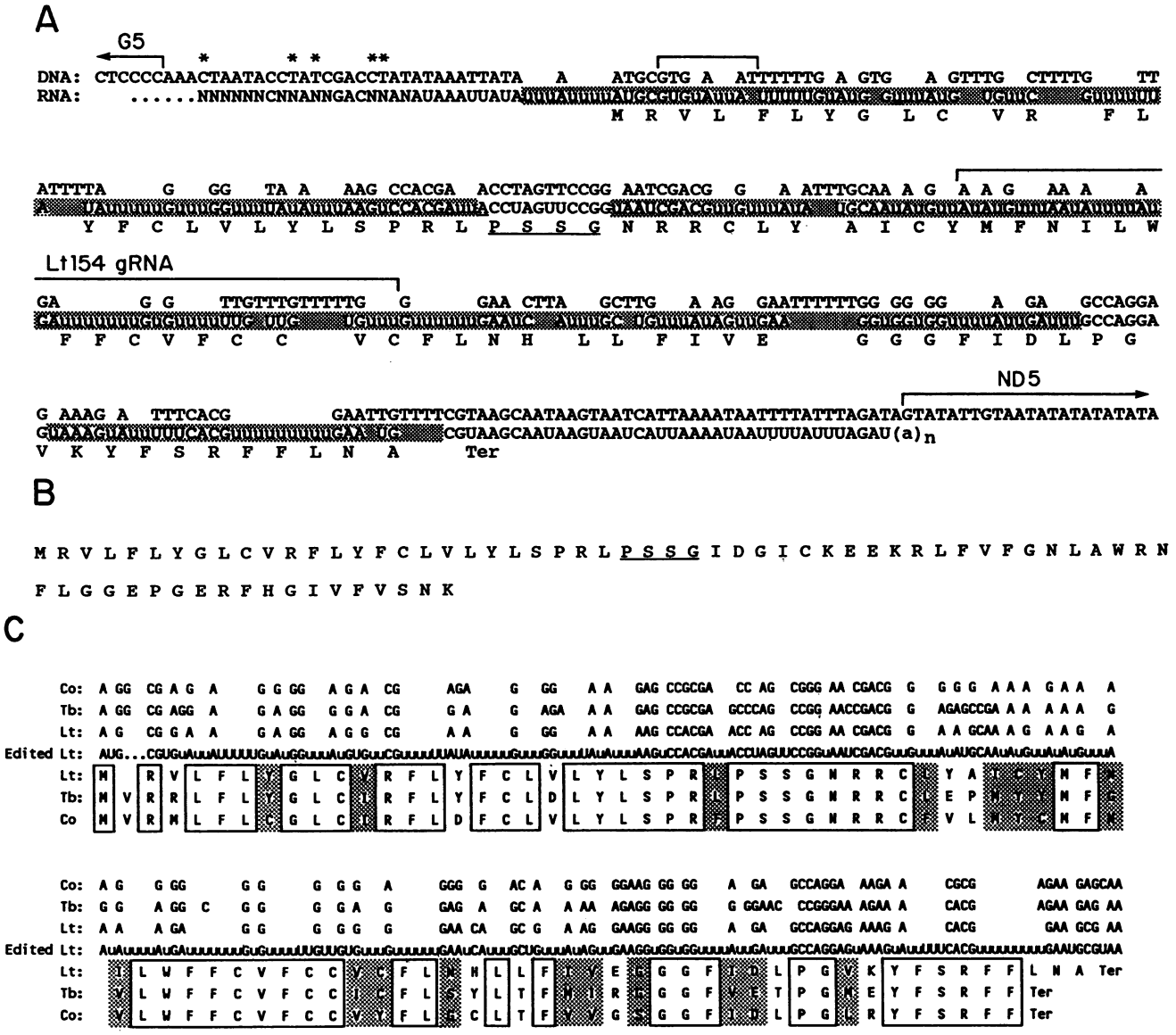

FIG. 5. The pan-edited G6 cryptogene. (A) DNA, G6 genomic sequence (nt 14636 to 14913 in LEIKPMAX). Adjacent portions and polarity of G5 region and ND5 gene are shown by arrows above the sequence. Asterisks identify the positions of the 5′ ends of edited RNA. The bracket shows the sites where most unedited transcripts are truncated. RNA, sequence of the fully edited G6 mRNA. The two major editing domains are indicated by shadowing. Unresolved nucleotides at the 5′ end are shown as N. Uridines added by editing are shown as u. An ambiguous position presumptively identified as uridine is indicated as u. A possible A-U polymorphism is indicated by A. The most frequently occurring 3′-terminal sequence is presented. The amino acid sequence of the 85-amino-acid G6 polypeptide is given beneath the mRNA sequence. The tetrapeptide PSSG encoded by the unedited domain connection sequence between two editing domains is underlined. The localization of the region edited by the Lt154 gRNA is indicated. (B) Amino acid sequence of the alternative 70-amino-acid G6 polypeptide. The N-terminal portion is common with the 85-amino-acid polypeptide. (C) Possible evolutionary conservation of G6 85-amino-acid polypeptide in three kinetoplastid species. Genomic sequences of the G6 regions from *L. tarentolae* (Lt), *C. oncopelti* (Co) (17, 27–29) (GenBank accession number X56015), and *T. brucei* (Tb) with encoded U residues deleted were aligned, using the known occurrence of U residues in the edited *L. tarentolae* sequence as explained in the text. Beneath the nucleotide sequences are shown the putative translated amino acid sequences for the edited RNAs from each species. Matches in all three sequences are boxed, and conservative amino acid substitutions are shadowed. The pairwise similarity scores (including conservative replacements) are 82% for Lt versus Co and 84% for Lt versus Tb.

edited mRNA species or that only the N-terminal portion is functional and conserved in sequence and the C-terminal portion is diverged between species.

The G6 protein has sequence similarity with the family of RPS12. PROFILE analysis (13), using the G6 amino acid sequence from *L. tarentolae* aligned with the two putative G6 amino acid sequences from *C. oncopelti* and *T. brucei*, gave several possible significant sequence similarities. RPS12 from several organisms represented 18 of the top 100 hits.

The significance of pairwise alignments of the G6 protein with known RPS12 was measured by performing Monte

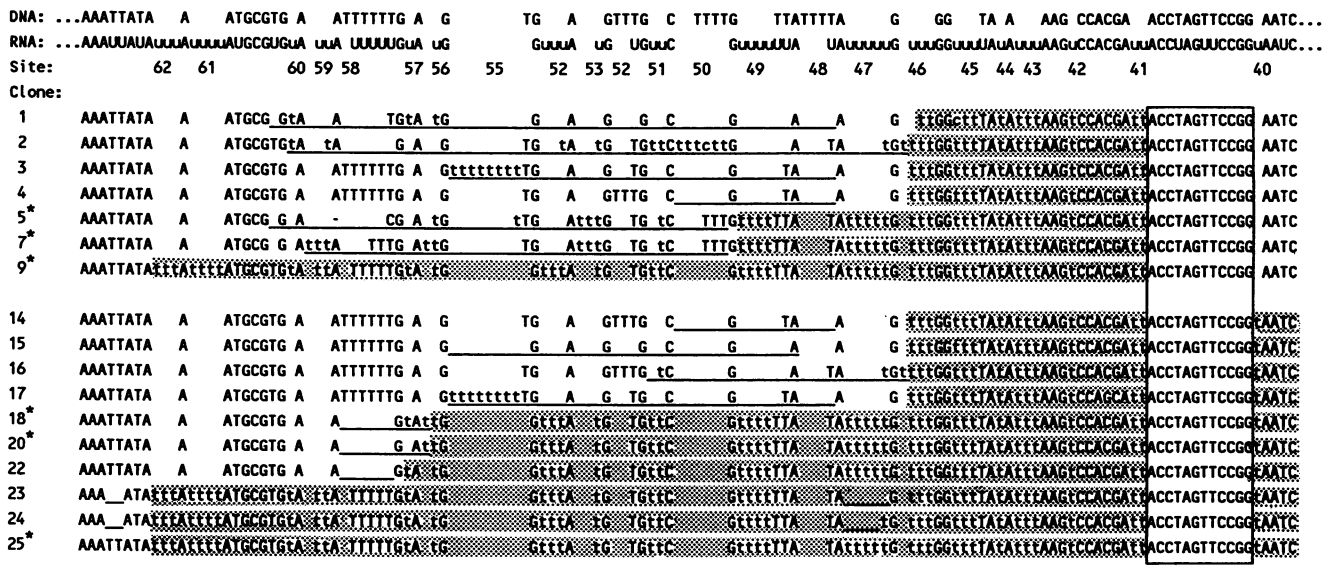

FIG. 6. Evidence that editing of the 5′ domain can occur independently of editing of the 3′ domain. Editing intermediates were PCR amplified by using the S-456 unedited 3′ primer and the unedited S-452 5′ primer. Only the 5′-domain portions of the sequences are shown. Clones 1 to 9 are unedited in the 3′ domain (sites 1 to 40) but contain partially or completely edited sequence at sites 41 to 62. Clones 14 to 25 are unedited at sites 1 to 6, completely edited at sites 7 to 40, and contain variable extents of editing at sites 41 to 62. The asterisks indicate that these patterns were found in additional clones. Clone patterns 5, 7, 18, and 20 were found in two clones each. The pattern of clone 9 was found in five clones, and the pattern of clone 25 was found in four clones. Additional symbols are the same as in Fig. 2.

Carlo shuffling analysis using the program SEQDP (19). As shown in Table 1, Z values for alignments of known RPS12 with each other were 19 to 58 standard deviation units. Alignments of the G6 protein with several chloroplast RPS12 and the *Escherichia coli* RSP12 showed Z values of 6 to 9 standard deviation units, which indicates that these alignments are probably not due to chance. However, alignments of the G6 protein with mitochondrial RPS12 sequences from *Zea mays* and *Paramecium tetraurelia* were not statistically significant. We interpret the low Z values for the G6/RPS12 chloroplast and eubacterial alignments as being due to sequence divergence, as has been found for some other kinetoplastid mitochondrial proteins (43).

A multiple alignment of seven RPS12 sequences from different organisms together with the G6 sequence is shown in Fig. 7. Regions of absolute or significant conservative matches of the known RPS12 sequences with the G6 sequence are boxed, and two regions of known functional significance in RPS12 are marked I and II. Mutations in these latter regions confer streptomycin resistance (region I) or streptomycin dependence (region II). The G6 protein is 40 to 60 amino acids shorter than the RPS12, and the gaps are in the amino-terminal and carboxy-terminal regions, which represent the least conserved regions of RPS12 in terms of absolute matches. The PSSG peptide in the G6 protein, which is encoded by the domain connection sequence, lies within conserved region I.

The hydropathy profile of the G6 protein is compared with representative profiles of RPS12 in Fig. 8. The conserved regions shown in the multiple alignment are boxed, and regions I and II are indicated by underlining. The RPS12 patterns show a clear similarity to each other. The G6 pattern is also quite similar to the RPS12 patterns, with the exception of the substitution of a hydrophobic peak for a hydrophilic dip between regions I and II.

The Z value results together with the multiple alignment and the similar hydropathy patterns suggest that the G6

TABLE 1. Statistical significance of sequence similarities between G6 polypeptide and ribosomal protein S12 from eight species[a]

cCr		cEg	cMp	cCp	Ec	Ml	mPt	mZm
kLt	8.9 ± 0.9	8.0 ± 0.4	7.3 ± 0.6	6.5 ± 0.7	7.7 ± 0.5	2.4 ± 0.2	0.02 ± 0.6	0.5 ± 0.1
cCr		50.2 ± 2.1	52.9 ± 3.8	56.1 ± 5.1	53.1 ± 3.2	46.6 ± 4.0	22.8 ± 3.2	36.8 ± 3.2
cEg			50.5 ± 5.3	49.0 ± 2.4	46.9 ± 4.7	43.7 ± 5.2	22.4 ± 1.1	36.1 ± 2.6
cMp				57.5 ± 5.1	47.5 ± 1.8	46.7 ± 2.8	22.4 ± 2.7	37.1 ± 3.4
cCp					53.1 ± 3.2	46.6 ± 4.0	22.8 ± 1.9	36.8 ± 3.2
Ec						45.6 ± 5.3	23.0 ± 1.8	34.7 ± 2.3
Ml							18.9 ± 2.0	31.4 ± 3.7
mPt								23.8 ± 2.4

[a] SEQDP program of Kanehisa (19). Values represent the average number of standard deviation units for 10 SEQDP trials (10 × 100 random sequence comparisons), indicating the significance that the calculated distance between two sequences is not due to chance. Values greater than 3 standard deviation units are considered significant. kLt, *L. tarentolae* kinetoplast; cCr, *Chlamydomonas reinhardtii* chloroplast; cEg, *Euglena gracilis* chloroplast; cMp, *Marchantia polymorpha* chloroplast; cCp, *Cyanophora paradoxa* cyanelle; Ec, *E. coli*; Ml, *Micrococcus luteus*; mPt; *P. tetraurelia* mitochondria; mZm, *Z. mays* mitochondria.

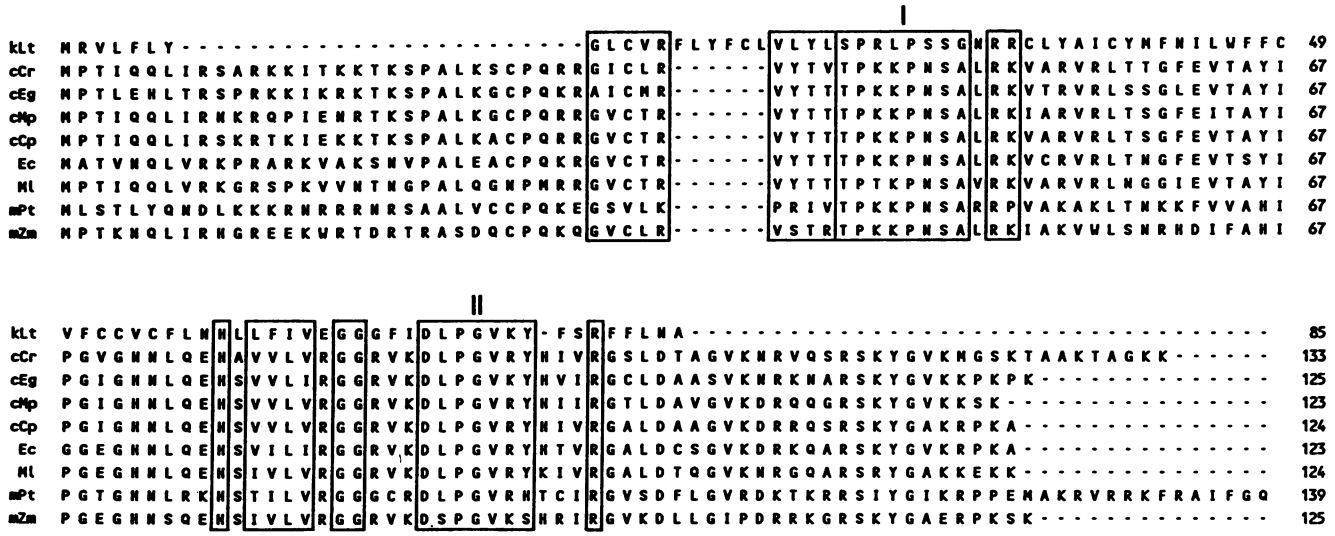

FIG. 7. Multiple sequence alignment of RPS12 from several species with the G6 protein from *L. tarentolae*. The alignments were performed by using CLUSTAL and BESTFIT (University of Wisconsin Genetics Computer Group). Gaps are indicated by dashes. Regions conserved relative to the G6 sequence in terms of matches or conservative changes are boxed. I and II refer to the regions which are involved in determining streptomycin resistance and dependence in several known RPS12. The amino acid numbers on the right refer to individual proteins. Abbreviations and references: kLt, kinetoplast *L. tarentolae*; cCr, chloroplast *Chlamydomonas reinhardtii* (26); cEg, chloroplast *Euglena gracilis* (31); cMp, chloroplast *Marchantia polymorpha* (34); cCp, cyanelle (chloroplast) *Cyanophora paradoxa* (23); Ec, *E. coli* (37); Ml, *Micrococcus luteus* (33); mPt, mitochondrial *P. tetraurelia* (38); mZm, mitochondrial *Z. mays* (15).

protein represents a diverged homolog of the RPS12 family which lacks the N-terminal and C-terminal regions.

G6 edited RNA exists in two size classes. The expected size of the unedited transcripts, most of which are 5′ truncated, is approximately 200 nt. An RNA of this size was detected on a Northern blot using two different unedited oligonucleotide probes (Fig. 9). The S-454-specific band was consistently broader and stronger, probably reflecting the contribution of a heterogeneous population of intermediates partially edited in the 3′ domain.

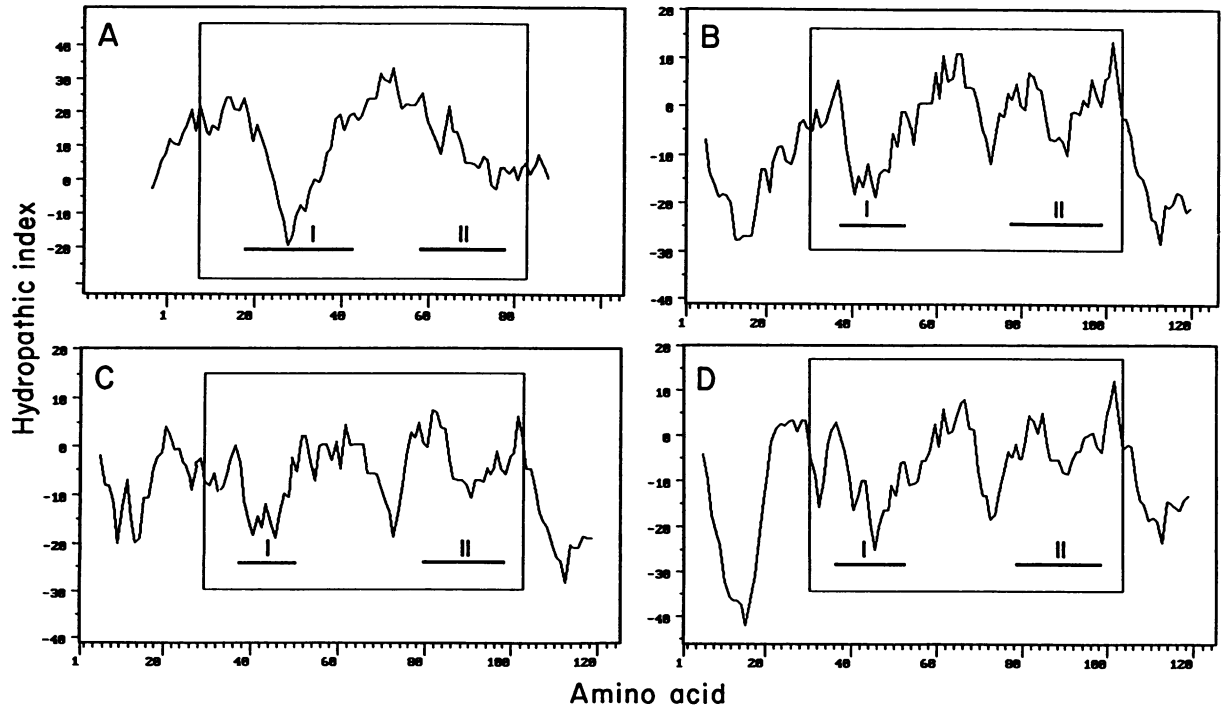

FIG. 8. Comparison of hydropathy profiles of the *L. tarentolae* G6 protein and several RSP12. A window of 11 amino acids was used. The RPS12 regions that are conserved in the G6 protein are boxed, and regions I and II are underlined. (A) *L. tarentolae* G6 protein; (B) *Euglena gracilis* chloroplast RPS12 (31); (C) *E. coli* RPS12 (37); (D) *P. tetraurelia* mitochondrial RPS12 (38).

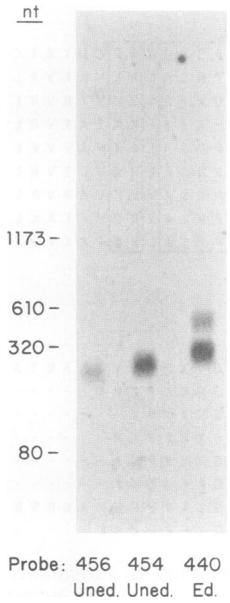

FIG. 9. Northern analysis of G6 RNAs. The oligonucleotides (5′ end labeled with ^{32}P) indicated below each lane were used to probe Northern blots of kRNA. The reference nucleotide sizes shown on the left were derived from the hybridization of duplicate filters with labeled pLt120 maxicircle DNA (55) and from the position of tRNA in the stained gel.

An edited sequence-specific probe detected two RNA bands of 300 and 450 nt. The lower-molecular-weight band showed the expected size increase of 90 nt due to the additional U residues added by editing. The broadness of this band is probably due to the presence of partially edited RNAs in addition to the mature species. The higher-molecular-weight band probably consists solely of mature edited RNA, since the S-454 probe, which could detect partially edited RNAs, did not hybridize with this band. As shown by Bhat et al. (5) for the *MURF4* gene of *L. tarentolae*, the major cause of the size difference between the two classes of transcripts is probably the length of the poly(A) or poly(AU) tail, but this was not examined in the case of the G6 RNAs.

Involvement of a minicircle-encoded gRNA in editing of the G6 transcript. A portion of the G6 editing information could be provided by the Lt154 minicircle-encoded gRNA (55) (Fig. 10A). Assuming that edited sites 21 to 23 serve as a 3′ anchor, this gRNA could mediate the addition of 23 U residues in nine sites beginning at site 24 (Fig. 2). This gRNA was shown previously to encode the editing information for an unexpected partially edited COIII RNA (55).

The chimeric Lt154 gRNA/G6 mRNA molecules predicted by the transesterification model for RNA editing (6) were amplified from kRNA, using S-363 as a G6-specific 3′ primer and S-397 as a gRNA 154-specific 5′ primer. All three chimerics obtained were joined at editing site 24 with oligo(U) tails of one, six, and nine U's in length (Fig. 10B). Sequence 1 had a deletion of the gRNA similar to that reported previously in ND7 chimeric amplifications (7). All mRNA editing sites 3′ of the site of chimeric junction were mature.

DISCUSSION

The demonstration that the G6 transcript is pan-edited to produce a mature mRNA that encodes a homolog of RPS12 suggests that the other five G-rich regions (48) may also be

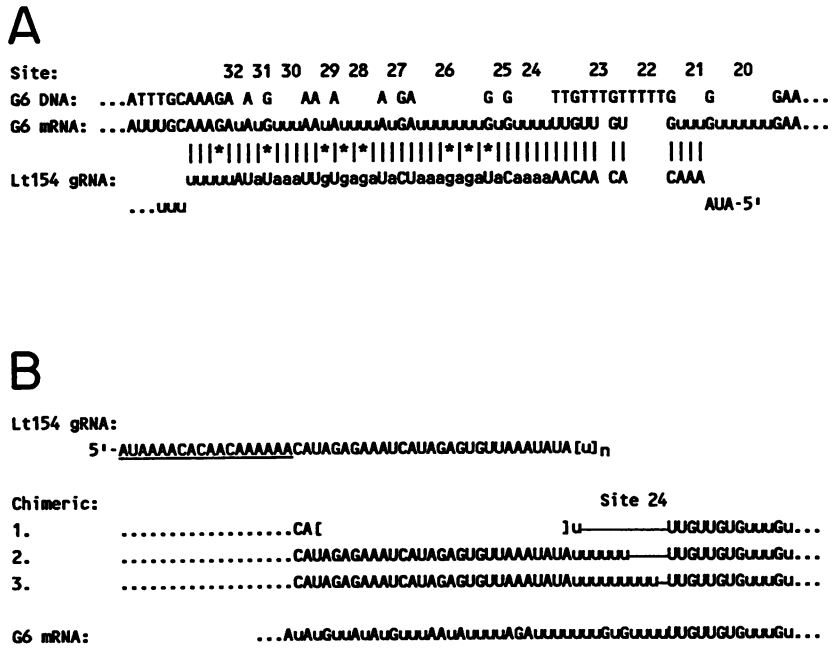

FIG. 10. Evidence that minicircle-encoded Lt154 gRNA encodes information for nine editing sites. (A) Edited G6 RNA aligned with Lt154 gRNA. U residues added to the mRNA and guide nucleotides in the gRNA are shown in lowercase. Editing site numbers are shown above the mRNA sequence. (B) Chimeric Lt154 gRNA/G6 mRNA molecules. The sequence of the gRNA is shown on the top; the oligo(U) tail is indicated by $[u]_n$, and the 5′ PCR primer (S-397) sequence is underlined. The sequence of the mature G6 mRNA is given at the bottom. S-363 was used as the 3′ PCR primer. Only the gRNA/mRNA junction regions of the chimeric molecules are shown. The gRNA deletion in sequence 1 is indicated by brackets.

pan-edited cryptogenes, although this must be verified by direct sequence analysis of the edited mRNAs. Preliminary evidence for pan-editing of G5 transcripts has been obtained (14), although a fully edited G5 RNA sequence has not yet been detected. The presence of six additional cryptogenes would represent a significant increase in the known genetic information contained in the mitochondrial genome of kinetoplastids.

The identification of the G6 85-amino-acid polypeptide as a homolog of the RPS12 family raises several interesting questions. Protection experiments have defined the binding of RPS12 to specific regions of the *E. coli* 16S rRNA secondary structure (51). Several of these regions are missing in the *L. tarentolae* 9S small rRNA secondary structure (10), suggesting a rationale for the missing amino-terminal and carboxy-terminal portions of the G6 protein. These regions of the RPS12 proteins in other organisms have been proposed to represent rRNA-binding domains (25).

The significance of the greater relative similarity of the G6 sequence with chloroplast RPS12 sequences than with mitochondrial RPS12 sequences from *P. tetraurelia* or *Z. mays* is uncertain. One possible explanation is a more rapid divergence of the other mitochondrial genes than the kinetoplastid gene. On the other hand, we speculate that this could be due to a horizontal gene transfer from a chloroplast genome which may have been present in the ancestral cell to the kinetoplastid and euglenoid lineages and that the chloroplast organelle was lost in the former but retained in the latter lineage. The evolutionary relatedness of kinetoplastid and euglenoid cells is supported by morphological considerations (21), by nuclear rRNA alignments (12, 50), and by the presence of *trans*-splicing of nuclear-encoded mRNAs in both lineages (56).

Editing of the G6 transcripts occurs in a 3′-to-5′ manner in two or three editing domains. The overall regulation of editing in the different domains remains to be explained. The first example of a pan-edited gene with separately edited domains was the ND7 gene of *T. brucei* (22). In this case, the editing of each domain is developmentally regulated: RNA with complete editing in both domains was present only in bloodstream trypanosomes, while in procyclic cells, the 3′ domain was edited incompletely or unedited. Although a detectable steady-state level of the 5′-domain-edited, 3′-domain-unedited G6 mRNA in *L. tarentolae* promastigotes was not found by primer extension analysis, the possibility exists that the alternative 70-amino-acid polypeptide is synthesized during another stage of the parasite life cycle.

Unexpected editing patterns are characterized by deviations from the unidirectional 3′-5′ polarity of events and/or by additions and deletions in wrong positions (1, 10, 54). Usually these patterns are found in the junction region between the fully edited sequence on the 3′ side and the unedited sequence on the 5′ side. Two models have been proposed to explain the existence of these editing patterns. According to the model of Sturm and Simpson (54), unexpected patterns could be generated by accurate editing with an incorrect gRNA (misguiding). The model of Decker and Sollner-Webb (9) assumes that multiple indiscriminate additions and deletions of uridines occur in and around editing sites, and the correct combinations of changes are then fixed by hybridization with gRNA. The finding in this study that the Lt154 minicircle gRNA, which was previously shown to act as a misguiding gRNA in producing an unexpected partially edited COIII mRNA pattern (54), acts as a functional gRNA for the G6 mRNA is strong evidence for the generation of at least some unexpected junction region patterns by the misguiding model.

In addition, some indirect evidence in favor of the existence of misguiding gRNAs can be obtained by an analysis of the junction regions in the G6 partially edited clones 20 to 24 in Fig. 2. In all of these cases, deletions in sites 48, 50, and 52 were carried out correctly, while no additions occurred in sites 47, 49, and 51. The existence of the same complex pattern in five different clones is not consistent with a random process but could be explained by participation of a deletion-specific gRNA. This putative gRNA may actually guide an initial step in the editing of this region and be followed by U-addition editing of the same region, or it may represent a misguiding gRNA.

The existence of a delay in completion of editing of the 5′ domain in the *L. tarentolae* G6 mRNA until the 3′ domain is edited is indicated by the sequences of clones 22, 23, and 24 in Fig. 2. The edited and unexpected partially edited sequences in the 5′ domains of these clones show a striking similarity, whereas the extent of editing in the 3′ domains differs significantly. A similar situation was described for three editing intermediates of the ND7 gene in *T. brucei* (22). The temporal delay could be related to the action of misguiding gRNAs and the subsequent 3′ correction of the resulting pattern with the correct gRNA.

The total genomic gRNA content in *L. tarentolae* is still unknown. To date, eight maxicircle-encoded gRNAs have been identified (6, 53) and 17 different minicircle sequences classes have been cloned (20, 30), yielding a subtotal of 25 different gRNAs. We estimate that eight gRNAs would be required for complete editing of the G6 transcript and an additional 40 to 45 gRNAs for transcripts from the remaining five G-rich regions. A single minicircle-encoded gRNA that overlaps 12 editing sites in the G6 mRNA has been identified. To date, all identified gRNAs for pan-edited genes in *T. brucei* and *L. tarentolae* are minicircle DNA encoded (4, 22, 35, 36, 55). This number of gRNA genes could easily be within the coding capacity of the complex minicircle DNA in *T. brucei* (43, 44) but would appear to exceed the minicircle coding capacity in *L. tarentolae* (20, 32, 58). Possible explanations include the presence of additional minicircle minor sequence classes, the possible existence of microsequence heterogeneity of minicircle DNA in *L. tarentolae* perhaps limited to the gRNA region, or the possibility that G-rich regions 1 to 5 in *L. tarentolae* may no longer be functional due to the extended period of time that this organism has been maintained in culture and that gRNAs for these regions may no longer be required. Another theoretical possibility is an importation of gRNAs which may be encoded in the nucleus. In regard to the mis editing-misguiding hypothesis for the generation of unexpected editing patterns, the number of gRNAs required does not correspond to the number of patterns, since one gRNA could mediate multiple unexpected editing patterns by forming false anchors in different regions. In addition, by looping out just upstream of the correct anchor, the gRNA could form a secondary anchor and create an unexpected pattern by editing the wrong positions. It remains to be seen, however, whether there are sufficient gRNAs to account for all the observed expected and unexpected editing patterns or whether an imprecision in the editing process (10) is responsible for some of these patterns. This is still an open question until the total genomic gRNA content is well defined.

The G6-RPS12 cryptogene system represents a useful experimental model for a detailed analysis of the precise

scenario of RNA maturation events involved in pan-editing in general.

ACKNOWLEDGMENTS

This research was supported in part by grant AI09102 to L.S. from the National Institutes of Health. N.R.S. was a predoctoral trainee on NIH training grant GM07104.

We thank Beat Blum for discussions and G. Jayarama Bhat for kindly providing a preprint of his paper.

REFERENCES

1. **Abraham, J., J. Feagin, and K. Stuart.** 1988. Characterization of cytochrome c oxidase III transcripts that are edited only in the 3′ region. Cell **55**:267–272.
2. **Benne, R.** 1985. Mitochondrial genes in trypanosomes. Trends Genet. **6**:117–121.
3. **Benne, R.** 1989. RNA editing in trypanosome mitochondria. Biochim. Biophys. Acta **1007**:131–139.
4. **Bhat, G. J., D. J. Koslowsky, J. E. Feagin, B. L. Smiley, and K. Stuart.** 1990. An extensively edited mitochondrial transcript in kinetoplastids encodes a protein homologous to ATPase subunit 6. Cell **61**:885–894.
5. **Bhat, G. J., P. J. Myler, and K. Stuart.** 1991. The two ATPase 6 mRNAs of Leishmania tarentolae differ at their 3′ ends. Mol. Biochem. Parasitol. **48**:139–150.
6. **Blum, B., N. Bakalara, and L. Simpson.** 1990. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. Cell **60**:189–198.
7. **Blum, B., N. R. Sturm, A. M. Simpson, and L. Simpson.** 1991. Chimeric gRNA-mRNA molecules with oligo(U) tails covalently linked at sites of RNA editing suggest that U addition occurs by transesterification. Cell **65**:543–550.
8. **Braly, P., L. Simpson, and F. Kretzer.** 1974. Isolation of kinetoplast-mitochondrial complexes from Leishmania tarentolae. J. Protozool. **21**:782–790.
9. **Decker, C. J., and B. Sollner-Webb.** 1990. RNA editing involves indiscriminate U changes throughout precisely defined editing domains. Cell **61**:1001–1011.
10. **de la Cruz, V., J. A. Lake, A. M. Simpson, and L. Simpson.** 1985. A minimal ribosomal RNA: sequence and secondary structure of the 9S kinetoplast ribosomal RNA from Leishmania tarentolae. Proc. Natl. Acad. Sci. USA **82**:1401–1405.
11. **Feagin, J. E., J. Abraham, and K. Stuart.** 1988. Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei. Cell **53**:413–422.
12. **Gray, M. N.** 1989. The evolutionary origins of organelles. Trends Genet. **5**:294–299.
13. **Gribskov, M., R. Luthy, and D. Eisenberg.** 1990. Profile analysis. Methods Enzymol. **183**:146–159.
14. **Gruszynski, E., B. Niner, N. Sturm, D. Maslov, and L. Simpson.** Unpublished data.
15. **Gualberto, J. M., H. Wintz, J. H. Weil, and J. M. Grienenberger.** 1988. The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are co-transcribed. Mol. Gen. Genet. **215**:118–127.
16. **Hancock, K., and S. L. Hajduk.** 1990. The mitochondrial tRNAs of *Trypanosoma brucei* are nuclear encoded. J. Biol. Chem. **265**:19208–19215.
17. **Horvath, A., D. A. Maslov, and A. A. Kolesnikov.** 1990. The nucleotide sequence of the 12S ribosomal RNA gene from kinetoplast DNA of the protozoan *Crithidia oncopelti*. Nucleic Acids Res. **18**:2811.
18. **Jasmer, D., J. Feagin, M. Payne, and K. Stuart.** 1987. Variation of G-rich mitochondrial transcripts among stocks of African trypanosomes. Mol. Biochem. Parasitol. **22**:259–272.
19. **Kanehisa, M. I.** 1982. Los Alamos sequence analysis package for nucleic acids and proteins. Nucleic Acids Res. **10**:183–196.
20. **Kidane, G., D. Hughes, and L. Simpson.** 1984. Sequence heterogeneity and anomalous electrophoretic mobility of kinetoplast minicircle DNA in *Leishmania tarentolae*. Gene **27**:265–277.
21. **Kivic, P. A., and P. L. Walne.** 1984. An evaluation of a possible phylogenetic relationship between the Euglenophyta and the Kinetoplastida. Origins Life **13**:269–288.
22. **Koslowsky, D. J., G. J. Bhat, A. L. Perrollaz, J. E. Feagin, and K. Stuart.** 1990. The MURF3 gene of *Trypanosoma brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. Cell **62**:901–911.
23. **Kraus, M., M. Gotz, and W. Loffelhardt.** 1990. The cyanelle str operon from *Cyanophora paradoxa*: sequence analysis and phylogenetic implications. Plant Mol. Biol. **15**:561–573.
24. **Kyte, J., and R. Doolittle.** 1982. Hydropathy index of proteins. J. Mol. Biol. **157**:105–132.
25. **Liljas, A.** 1991. Comparative biochemistry and biophysics of ribosomal proteins. Int. Rev. Cytol. **124**:103–136.
26. **Liu, X. Q., N. W. Gillham, and J. E. Boynton.** 1989. Chloroplast ribosomal protein gene rps12 of *Chlamydomonas reinhardtii*. Wild-type sequence, mutation to streptomycin resistance and dependence, and function in *Escherichia coli*. J. Biol. Chem. **264**:16100–16108.
27. **Maslov, D. A., N. S. Entelis, A. A. Kolesnikov, and G. N. Zaitseva.** 1982. Kinetoplast DNA from *Crithidia oncopelti*. Restriction mapping of maxicircle DNA. Bioorg. Chem. **8**:676–685. (In Russian.)
28. **Maslov, D. A., N. S. Entelis, A. A. Kolesnikov, G. N. Zaitseva, N. P. Kouzmin, and I. I. Fodor.** 1981. Cloning and restriction mapping of the maxicircle fragment of *Crithidia oncopelti* kinetoplast DNA. Dokl. Akad. Nauk. SSSR **261**:1271–1273.
29. **Maslov, D. A., K. Gwang, and A. A. Kolesnikov.** Unpublished data.
30. **Maslov, D. A., and L. Simpson.** Unpublished data.
31. **Montandon, P., and E. Stutz.** 1984. The genes for the ribosomal proteins S12 and S7 are clustered with the gene for the EF-Tu protein on the chloroplast genome of *Euglena gracilis*. Nucleic Acids Res. **12**:2851–2859.
32. **Muhich, M., and L. Simpson.** 1986. Specific cleavage of kinetoplast minicircle DNA from *Leishmania tarentolae* and identification of several additional minicircle sequence classes. Nucleic Acids Res. **14**:5531–5556.
33. **Ohama, T., F. Yamao, A. Muto, and S. Osawa.** 1987. Organization and codon usage of the streptomycin operon in *Micrococcus luteus*, a bacterium with a high genomic G+C content. J. Bacteriol. **169**:4770–4777.
34. **Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, S. Sano, K. Umesono, Y. Shiki, M. Takeuchi, Z. Chang, S. Aota, H. Inokuchi, and H. Ozeki.** 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymopha* chloroplast DNA. Nature (London) **322**:572–574.
35. **Pollard, V. W., and S. L. Hajduk.** 1991. *Trypanosoma equiperdum* minicircles encode three distinct primary transcripts which exhibit guide RNA characteristics. Mol. Cell. Biol. **11**:1668–1675.
36. **Pollard, V. W., S. P. Rohrer, E. F. Michelotti, K. Hancock, and S. L. Hajduk.** 1990. Organization of minicircle genes for guide RNAs in *Trypanosoma brucei*. Cell **63**:783–790.
37. **Post, L. E., and M. Nomura.** 1980. DNA sequences from the str operon of *Escherichia coli*. J. Biol. Chem. **255**:4660–4666.
38. **Pritchard, A. E., J. J. Seilhamer, R. Mahalingam, C. L. Sable, S. E. Venuti, and D. J. Cummings.** 1990. Nucleotide sequence of the mitochondrial genome of *Paramecium*. Nucleic Acids Res. **18**:173–180.
39. **Shaw, J., J. E. Feagin, K. Stuart, and L. Simpson.** 1988. Editing of mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons. Cell **53**:401–411.
40. **Simpson, A. M., Y. Suyama, H. Dewes, D. Campbell, and L. Simpson.** 1989. Kinetoplastid mitochondria contain functional tRNAs which are encoded in nuclear DNA and also small minicircle and maxicircle transcripts of unknown function. Nucleic Acids Res. **17**:5427–5445.
41. **Simpson, L.** 1979. Isolation of maxicircle component of kinetoplast DNA from hemoflagellate protozoa. Proc. Natl. Acad. Sci. USA **76**:1585–1588.
42. **Simpson, L.** 1986. Kinetoplast DNA in trypanosomid flagellates.

Int. Rev. Cytol. **99**:119–179.
43. **Simpson, L.** 1987. The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication and evolution. Annu. Rev. Microbiol. **41**:363–382.
44. **Simpson, L.** 1990. RNA editing—a novel genetic phenomenon? Science **250**:512–513.
45. **Simpson, L.** Unpublished data.
46. **Simpson, L., and P. Braly.** 1970. Synchronization of *Leishmania tarentolae* by hydroxyurea. J. Protozool. **17**:511–517.
47. **Simpson, L., N. Neckelmann, V. de la Cruz, A. Simpson, J. Feagin, D. Jasmer, and K. Stuart.** 1987. Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. J. Biol. Chem. **262**:6182–6196.
48. **Simpson, L., and J. Shaw.** 1989. RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. Cell **57**:355–366.
49. **Simpson, L., and A. Simpson.** 1978. Kinetoplast RNA from *Leishmania tarentolae*. Cell **14**:169–178.
50. **Sogin, M., J. Gunderson, H. Elwood, R. Alonso, and D. Peattie.** 1989. Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Gardia lamblia*. Science **243**:75–77.
51. **Stern, S., T. Powers, L. M. Changchien, and H. F. Noller.** 1988. Interaction of ribosomal proteins S5, S6, S11, S12, S18 and S21 with 16 S rRNA. J. Mol. Biol. **201**:683–695.
52. **Stuart, K., J. Feagin, and J. Abraham.** 1989. RNA editing: the creation of nucleotide sequences in mRNA—a minireview. Gene **82**:155–160.
53. **Sturm, N. R., D. A. Maslov, B. Blum, and L. Simpson.** Unpublished data.
54. **Sturm, N. R., and L. Simpson.** 1990. Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. Cell **61**:871–878.
55. **Sturm, N. R., and L. Simpson.** 1990. Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. Cell **61**:879–884.
56. **Tessier, L. H., M. Keller, R. L. Chan, R. Fournier, J. H. Weil, and P. Imbault.** 1991. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-splicing in *Euglena*. EMBO J. **10**:2621–2626.
57. **Van der Spek, H., D. Speijer, G.-J. Arts, J. Van den Burg, H. Van Steeg, P. Sloof, and R. Benne.** 1990. RNA editing in transcripts of the mitochondrial genes of the insect trypanosome *Crithidia fasciculata*. EMBO J. **9**:257–262.
58. **Wesley, R. D., and L. Simpson.** 1973. Studies on kinetoplast DNA. III. Kinetic complexity of kinetoplast and nuclear DNA from *Leishmania tarentolae*. Biochim. Biophys. Acta **319**:267–276.
59. **Wood, W., J. Gitschier, L. Lasky, and R. Lawn.** 1985. Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. Proc. Natl. Acad. Sci. USA **82**:1585–1588.