THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 7, Issue of February 14, pp. 4212–4218, 1997

Printed in U.S.A.

Guide RNA-independent and Guide RNA-dependent Uridine Insertion into Cytochrome *b* mRNA in a Mitochondrial Lysate from *Leishmania tarentolae*

ROLE OF RNA SECONDARY STRUCTURE*

(Received for publication, May 28, 1996, and in revised form, October 10, 1996)

Gregory J. Connell‡§¶, Elaine M. Byrne‡, and Larry Simpson‡§‖**

From the ‡Howard Hughes Medical Institute, the §Department of Molecular, Cellular and Developmental Biology, and the ‖Department of Medical Microbiology and Immunology, University of California, Los Angeles, California 90095-1662

A primer extension assay was used for the detection of uridine insertions occurring *in vitro* in synthetic pre-edited cytochrome *b* mRNA during incubation with a *Leishmania tarentolae* mitochondrial extract. Two different activities were detected that inserted uridines within the first two editing sites: one that is dependent on the secondary structure of the mRNA but is independent of both exogenous and endogenous guide RNA, and a second that does not put the same structural constraints on the mRNA, but is dependent on the presence of a cognate guide RNA.

The mitochondrial genome of *Leishmania tarentolae*, like that of all other trypanosomatids, consists of a single network of catenated DNA maxicircles and minicircles (1–3). The maxicircle contains genes encoding proteins required for respiration and mitochondrial translation. The majority of the mRNAs transcribed from the maxicircle, however, must be edited by the addition and/or deletion of uridines (U nucleotides) to create an open reading frame (4–7).

The genetic information determining the number and location of U-insertions and deletions has been proposed to be located within the guide RNAs (gRNAs),[1] a class of small RNA molecules transcribed from both maxicircles and minicircles (8–11). gRNAs contain an anchor sequence that is complementary to mRNA immediately 3′ of the editing sites, and they also have a sequence that can function as a template for the insertions and deletions (provided G:U base pairs are allowed) (8). The 3′ end of the gRNA has a posttranscriptionally added oligo(U) tail that has been postulated to be important for editing (8).

Since *L. tarentolae* is not readily accessible to genetic analysis, the development of an *in vitro* system is crucial to understanding the mechanism of RNA editing in this organism. A mitochondrial lysate prepared from the related trypanosomatid, *Trypanosoma brucei*, has been shown to delete U nucleotides at site 1 of the mRNA encoding ATPase subunit 6 (= maxicircle unidentified reading frame 4 or MURF4) when the appropriate gRNA is added (12). The deletions can be manipulated in a predictable manner by altering the guiding nucleotides of the gRNA. Guide RNA-dependent U-insertion, which is by far the most common form of modification in kinetoplastid RNA editing, has also been recently demonstrated *in vitro* (36, 40).

An *in vitro* internal U-insertion activity has been demonstrated for two synthetic pre-edited mRNAs. The mRNAs incorporated [α-^{32}P]UTP when incubated with a mitochondrial lysate from *L. tarentolae* (13). Digestion with RNase H was used to show that U nucleotides were incorporated internally into a synthetic pre-edited cytochrome *b* (Cyb) transcript, as well as at the 3′ terminus. The 3′ end-labeling is due to a terminal uridylyl transferase activity that has been identified in mitochondrial extracts (14). Approximately 95% of the internal insertions were localized within a region containing the 15 editing sites of this RNA (13). Although the RNase H-based assay was well suited to detect internally labeled RNA, the assay does not permit the detection of U-insertions at individual editing sites. Furthermore, since the U-insertions occurred independent of added gRNA, the relationship of the *in vitro* reaction to biological editing was unclear.

We describe in this paper a primer extension assay that permits the detection of U-insertions at individual editing sites. We used this assay to characterize two different reactions. The internal U-insertion activity initially detected with the RNase H-based assay is shown to be highly dependent on the secondary structure of the mRNA and also not to be mediated by gRNA. A second U-insertion reaction is also described, which is dependent on the addition of the correct gRNA to the mitochondrial lysate.

EXPERIMENTAL PROCEDURES

Oligodeoxynucleotides—The oligonucleotides used in these studies are listed below together with a brief description of their function: FP = forward or 5′ PCR primer, RP = reverse or 3′ PCR primer. The locations of the underlined sequences in the *L. tarentolae* maxicircle sequence (GenBank entry LEIKPMAX) are indicated.

S194: TAATACGACTCACTATAGGGCTTATATGCATAAGAA (21 nt 5′ extended Cyb mRNA FP, nt 5350–5365 in LEIKPMAX). S656: TAATACGACTCACTATAGGGATATGACTTGAAGTTAAAAG (gCyb-I gRNA FP, nt 16780–16800 in LEIKPMAX). S1206: TAATACGACTCACTATA (T7 RNA polymerase promoter). S1395: TATAAAAGCGGAGAGAAAAGAAAAGGTCTTTTAACTTCAGGTTGTTTATTACGAG (Cyb mRNA edited at sites 1+2 FP, nt 5398–5450 in LEIKPMAX). S1520: TAATACGACTCACTATAGGGCCATAAATTTAATTTAAATTTTAAATAATTATAAAAGCGGAG (Cyb mRNA with 5-bp G-C stabilized helix FP, nt 5371–5410 in LEIKPMAX). S1521: GGGCCAAAAACCTAAACTAAAACCTACACC (Cyb mRNA with 5-bp G-C stabilized helix RP, nt 5456–5480 in LEIKPMAX). S1685: AAATTAAAAACCTAAAC-

* This work was supported in part by Grant AI09102 from the National Institutes of Health (to L. S.). 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.

¶ Current address: Dept. of Pharmacology, University of Minnesota, MN 55455.

** To whom correspondence should be addressed: Howard Hughes Medical Institute-UCLA, 6780 MacDonald Research Laboratories, 6775 Circle Dr. S., Los Angeles, CA 90024.

[1] The abbreviations used are: gRNA, guide RNA; PCR, polymerase chain reaction; RT, reverse transcription; nt, nucleotide(s); Cyb, cytochrome *b*; bp, base pair(s); DMS, dimthyl sulfate; FP, forward or 5′ PCR primer; RP, reverse or 3′ PCR primer.

TAAAACCTACACC (Cyb mRNA with disrupted helix RP, nt 5460–5480 in LEIKPMAX).

S1743: CACCATATTTCGCTTAAAAACTCTCCGCTTTTATAATTATTTAAAATTTAAATTAAATTT (Cyb reverse transcriptase (RT) normalization primer, nt 5372–5416 in LEIKPMAX). S1786: TAAAGCCTTTTCTTTTCTCTCCGCTT (RT of DMS-treated RNA, nt 5400–5426 in LEIKPMAX). S1790: GGGCCAAAAACCTAAACTAAAACCTACACCATATACTCGTAATAAATAATTGATTTATGAAGCCTTTTCTTTTCTC (Cyb mRNA with mutated anchor binding site RP, nt 5440–5480 and nt 5410–5426 in LEIKPMAX). S1795: AAAAAAAAAAAAAAATATTATTTAAAAATTTATATTATCTTTTAACTTCAAGTCATATGTGCCTATAGTGAGTCGTATTA (transcription of gCyb-I gRNA plus U-tail, nt 16758–16803 in LEIKPMAX). S1796: TATTATTTAAAAATTTATATTATCTTTTAACTTCAAGTCATATGTGCCTATAGTGAGTCGTATTAC (transcription of gCyb-I gRNA without U-tail, nt 16758–16803 in LEIKPMAX).

S1820: GTAAATTACTTTATTTAAATTTTAATATAATAATAATAATATCAC (maxicircle fragment encoding gCyb-I gRNA RP, nt 16801–16845 in LEIKPMAX). S1821: GAGTATTATAATATAATATATTTTATATCCAGTCTATCAG (maxicircle fragment encoding gCyb-I gRNA FP, nt 16705–16744 in LEIKPMAX). S1872: (CUACUACUACU)ATGACTTGAAGTTAAAAGA (3′ RACE of gCyb-I gRNA FP, nt 16780–16798 in LEIKPMAX). S1873: CAUCAUCAUCAUGGGGGGGGGGGGGGGG (3′RACE of poly(C)-tailed gCyb-I gRNA RP). S1894: TAAGTTATTATTTAAAAATTTATATTATCTTTTAACTTCAAGTCATATGTGCCTATAGTGAGTCGTATTAC (transcription of gCyb-I gRNA with 3′ extension, nt 16754–16803 in LEIKPMAX). S1899: TATTATTTAAAAATTTATATTACTTTAACTTCAAGTCATATGTGCCTATAGTGAGTCGTATTAC (transcription of gCyb-I(0,0) gRNA).

S1916: TAATACGACTCACTATAGGGATAAATTTAATTTAAATTTTAAATAATTATAAAAGCGG (Cyb mRNA with natural 5′ end FP, nt 5371–5408 in LEIKPMAX). S2003: AAATATTATTTAAAAATTTATATTATCTTTTAACTTCAAGTC ATATGTGCC (RNase H digestion of gCyb-I gRNA, nt 16758–16803 in LEIKPMAX). S2073: CACCATATACTCGTAATAAACAACCTGAAGTTAAA (Cyb RT assay primer, nt 5425–5459 in LEIKPMAX). S2074: ATAAAAATAATACAAAATATCAATTTGTACAAATAAAGCAA CTAAAAAATAATCATGCTA (Cyb RT normalization primer, nt 5514–5573 in LEIKPMAX). S2077: CACCATATACTCGTAATAAATAATTGATTTATGAA (RT of anchor-mutated Cyb mRNA). S2085: ACAAATAAAGCAACTAAAAAATAATCATGCTAAGCAAACACCAC (Cyb mRNAs RP, nt 5502–5545 in LEIKPMAX). S2093: TTCGACATGAGACACGGATCGATCCCCACAAATAAAGCAAC (3′ tagged Cyb RNAs RP, nt 5532–5545 in LEIKPMAX). S2094: TAATACGACTCACTATAGGGCGAATTGGGTACCGGGCC (5′ tagged Cyb mRNA FP). S2095: TTCGACATGAGACACGGATC (RT of 3′ tagged RNAs).

Genomic and cDNA Sequence of Cyb-I gRNA—Kinetoplast DNA was purified from stationary cells by sedimentation through a cesium chloride step gradient (15). The maxicircle sequence encoding the gCyb-I gRNA was amplified from 0.5 ng of kinetoplast DNA with 30 cycles of PCR using primers S1820 and S1821. The same two oligonucleotides were used to sequence the DNA fragment.

Kinetoplast RNA was extracted from purified mitochondria (16). A poly(C) tail was added to 5 μg of the kinetoplast RNA with poly(A) polymerase (Pharmacia Biotech Inc.). After phenol-chloroform extraction and ethanol precipitation, the RNA was annealed to S1873 and reverse-transcribed with avian myeloblastosis virus reverse transcriptase (Promega). The cDNA product was PCR-amplified using primers S1872 and S1873 and was cloned into pAMP1 (Life Technologies, Inc.) and transformed into DH5α cells for sequencing.

mRNA and gRNA Transcripts—The Cyb mRNA constructs were transcribed from PCR products that had been amplified from plasmid pNB2 (17) using the following primers: S194 and S2085 for the 21-nt 5′ extension, S1916 and S2085 for the natural 5′ end, S1520 and S1521 for the helix stabilized with 5 G-C base pairs, S1520 and S1685 for the disrupted G-C helix, S1520 and S1790 for the mutated anchor binding site, S1916 and S2093 for the natural 5′ end 3′ tag, S2094 and S2093 for the 5′ extension 3′ tag. The Cyb construct edited at both sites 1 and 2 was generated by PCR amplification with S1395 and S2085, followed by amplification of the purified product with S1916 and S2085; RT sequencing of the transcribed RNA confirmed the sequence.

The gRNAs were transcribed directly from synthesized oligonucleotide templates, as described previously (18), and the 3′ ends were sequenced by the same procedure described for the endogenous *L. tarentolae* gRNAs. All RNAs were gel-purified prior to use in the assay.

PCR—PCR was performed in 100-μl reactions containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2.5 mM $MgCl_2$, 200 μM of each dNTP, 0.5 μM of each primer, and 2.5 units of *Taq* polymerase (Perkin Elmer). The PCR profiles were 5 min at 95 °C, followed by a variable number of cycles of 95 °C for 30 s, 55 °C for 30 s or 60 s, and 72 °C for 60 s.

Dimethyl Sulfate (DMS) Modification—The Cyb mRNA construct (3 pmol) was denatured at 65 °C for 5 min in 1 mM EDTA, 50 mM HEPES (pH 7.5). For native conditions, the solution was adjusted to a final concentration of 20 mM KCl, 10 mM $MgCl_2$, 50 mM HEPES (pH 7.5), and 2 mM spermidine in a 200-μl volume. Following a 10-min 27 °C preincubation, 0.5 μl of DMS (Aldrich) was added, and the solution was incubated for another 8 min. For denaturing conditions, the RNA in 200 μl of 1 mM EDTA, 50 mM Hepes (pH 7.5) was heated with 0.5 μl of DMS at 95 °C for 45 s. Reactions were terminated with the addition of 100 μl of 1 M β-mercaptoethanol, 0.5 M HEPES (pH 7.5) and then ethanol-precipitated.

Approximately 0.2 pmol of the modified RNA was extended with primers S1685, S1786, S2073, and S2085 as described by Inoue and Cech (19). The extension products were resolved on a 12% polyacrylamide, 7 M urea gel and scanned on a PhosphorImager (Molecular Dynamics). A ratio of the intensity of the RT termination signal quantitated under native conditions to that measured under denaturing conditions was determined for the modified bases (A and C nucleotides). This ratio in each reaction was normalized to the base of maximal activity (highest native/denatured value). A mean normalized ratio for each site was obtained from at least two independent modification reactions. Positions complicated by modification-independent RT terminations were not quantitated.

Isolation of Mitochondria—L. tarentolae (UC strain) cells were grown to a late log phase density of 10^8 cells/ml as described previously (20). After washing twice in 0.25 M sorbitol, 20 mM HEPES, 2 mM EDTA (pH 7.5) (SHE), the cells were resuspended at a density of 10^9 cells/ml and disrupted in a Stansted Power Fluid apparatus (Energy Service Co., Washington, DC) at 1200 p.s.i. The crude mitochondrial fraction was pelleted, resuspended in 0.25 M sorbitol, 20 mM Hepes (pH 7.5), 2 mM $MgCl_2$, digested with DNase I to remove contaminating nuclear DNA, and washed by centrifugation in SHE. The crude fraction was taken up in 76% Renografin and a purified mitochondrial fraction isolated by flotation in Renografin density gradients as described previously for hypotonically broken cells (15), and resuspended in 20 mM HEPES (pH 7.5), 10% glycerol, and 100 mM KCl for storage at −80 °C. After thawing on ice, the protease inhibitors leupeptin and Pefabloc SC (Boehringer Mannheim) were added, for the Cyb assay, to final concentrations of 10 μg/ml and 1 mg/ml, respectively, and the mitochondria were lysed by addition of Triton X-100 to a final concentration of 0.3%.

Lysate-mediated Reactions—In a 20-μl volume containing 12 mM HEPES (pH 7.5) and 0.1 mM EDTA, 3 pmol of the mRNA construct were added to 30 pmol of the appropriate gRNA and denatured at 65 °C for 3 min. The RNA solution was adjusted in a 50-μl volume to a final concentration of 5 mM HEPES (pH 7.5), 0.04 mM EDTA, 20 mM KCl, 3 mM potassium phosphate (pH 7.5), 10 mM $MgCl_2$, 20 mM dithiothreitol, 2 mM spermidine, 10 μg/ml leupeptin, 1 mg/ml Pefabloc SC, 1 mM ATP, 1 mM GTP, and 1 mM UTP. Reactions were initiated by the addition of approximately 10^9 cell equivalents of the mitochondrial lysate containing about 30 μg of protein. After incubation at 27 °C for 50 min (Cyb), reactions were terminated by phenol-chloroform extraction followed by ethanol precipitation.

RT Primer Extension Assay—The lysate-treated Cyb mRNA constructs were purified in a 6% polyacrylamide, 7 M urea gel. A gel slice containing the full-length RNA plus molecules up to 40% longer was excised to ensure that those RNAs with U-additions would also be included. RNA was eluted from the gel at 37 °C for 1 h in 1 mM EDTA, 0.1% SDS, 0.5 M NH_4 acetate, 10 mM $MgCl_2$, and 0.3 M sodium acetate and then ethanol-precipitated using 10 μg of glycogen (Boehringer Mannheim) as a carrier. Approximately 0.7 pmol of each treated RNA sample was assayed for U-insertions by RT using a primer annealed immediately 3′ of the editing sites. The RNA was heated at 65 °C for 3 min in a 7.5-μl volume containing 10 pmol of primer S2073, 10 pmol of a normalization primer, 60 mM NaCl, 10 mM dithiothreitol, and 50 mM Tris-HCl (pH 8.3). The solution was made 5 mM in $MgCl_2$ and annealed 10 min at 42 °C. The extensions were performed in a 5-μl volume containing 2 μl of the primer/RNA solution and 200 μM ddCTP, 375 μM dGTP, 1.3 μM [α-^{32}P]dATP (800 μCi/mmol), 5 mM $MgCl_2$, 60 mM NaCl, 10 mM dithiothreitol, 50 mM Tris-HCl (pH 8.3), and 1 unit of avian myeloblastosis virus reverse transcriptase (Promega). Reactions were allowed to proceed at 42 °C for 30 min prior to electrophoresis in a 12% polyacrylamide, 7 M urea gel. All reactions were quantitated using a PhosphorImager (Molecular Dynamics).

Two different oligonucleotides, which annealed immediately 3′ of an encoded U, were used to normalize the extension signal; S1743 anneals 5′ of the editing sites and S2074 at the 3′ end of the RNA. The extension

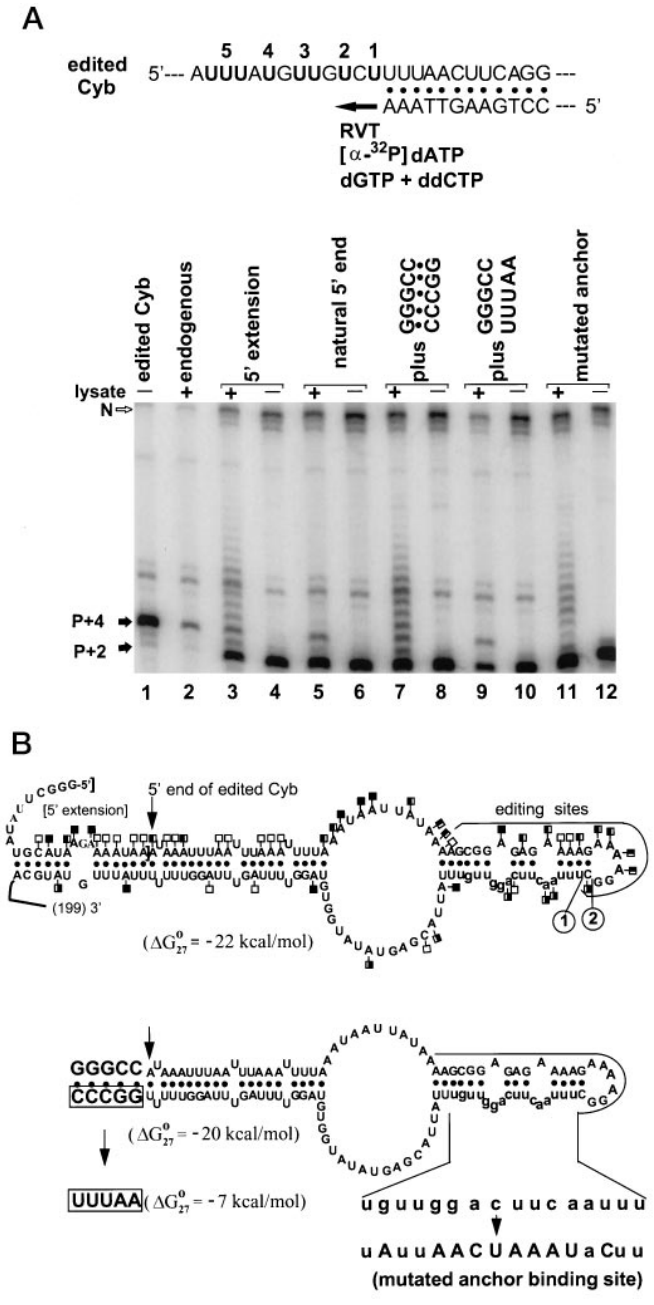

FIG. 1. **Guide RNA-independent U-insertions in Cyb mRNA constructs.** *A*, *upper panel*, diagram of primer extension assay of fully edited mRNA. U nucleotides added by editing are in *boldface*. *Lower panel*, primer extension assay of Cyb mRNA constructs after incubation of RNAs in mitochondrial lysate. No exogenous gRNA was added. The substrate RNAs are indicated above each lane: *edited Cyb*, a Cyb transcript edited at sites 1 and 2; *endogenous*, no RNA added to lysate; *5′ extension*, Cyb mRNA construct with the 5′ extension shown in *B*; *natural 5′ end*, Cyb mRNA construct with the mature 5′ end; *plus GGGCC/CCCGG*, Cyb mRNA construct with substitution of 5′ upstream sequence and terminal 3′ 65 nt with an artificial 5-bp GC-helix; *plus GGGCC/UUUAA* = Cyb mRNA construct lacking a stable helix as shown in *B*; *mutated anchor*, Cyb mRNA construct with mutated gRNA anchor binding site and the GGGCC/CCCGG terminal helix as shown in *B*, *lower panel*. Presence or absence of lysate is indicated by + or −. The extension product used for normalization of RNA quantity is indicated by *N* (see "Experimental Procedures"). The primer + 2 (P+2) and primer + 4 (P+4) extension products are indicated on the *left*. The additional bands shown in *lanes 1* and *2* migrating between the P+4 band and the N band represent artifacts of the normalization extension. These artifact bands are also present in experimental *lanes 3–12*. *B*, *upper panel*, a predicted secondary structure for the Cyb mRNA construct containing the 21-nt 5′-extended sequence plus 3 G nucleotides added by T7 transcription. The natural 5′ end is indicated by an *arrow*, and editing sites are indicated by a *bracket*; the first two sites are *numbered*. The anchor-binding sequence is in *lowercase*. DMS modification of this RNA was performed under both native and denaturing conditions (see "Experimental Procedures"). A ratio of the intensity of the RT termination signal measured under native conditions with that under denaturing conditions was determined for each modified position. A base was defined as protected from modification (□) if this ratio is less than 50% of maximum, partially protected (◧) if the ratio is between 50% and 75%, and fully reactive (■) if greater than 75%. The structure is 3 kcal/mol removed from the most stable fold, but it is in better agreement with the DMS modification results; the helix containing the 5′-extended sequence is also present in the optimal fold. *Lower panel*, the sequences and stabilities of additional constructs used in this study.

of S1743 generates some artifact bands that are not present with S2074, but it was used for Fig. 1*A* because the annealing site of S2074 is not present on some of the assayed constructs. S2074 was used for all other reactions. Since the lysate-mediated U-insertions occur infrequently, the signal intensity from the normalization primer is much greater than that from the assay primer, and it can interfere with quantitation. For this reason, the intensity of the normalization extensions was decreased with the addition of ddT to the 3′ end of the primers by treatment with terminal transferase (Boehringer Mannheim). This prevented 80% of primer S2074 and 99% of primer S1743 from being extended. The extension signal from all of the primers was linear with the RNA concentration provided that at least a 7-fold molar excess of the oligos were used in the reactions.

Indirect Primer Extension Assay—RT-PCR of the 3′ tagged Cyb mRNA constructs eliminated complications from endogenous edited mRNA and also chimeras and other potential ligation products. After gel purification, the lysate-treated 3′ tagged constructs were reverse-transcribed using primer S2095. The cDNA derived from the Cyb mRNA construct containing the mature 5′ end was treated with RNase A and amplified with four cycles of PCR using S1916 and S2095. The 3′ tagged construct containing the cytochrome oxidase III extension also has a 67-base 5′ tag taken from the PNB2 vector (17). Amplification of the cDNA synthesized from this construct was with primers S2094 and S2095. RNA was produced for use in the RT assay by T7 transcription of the PCR products.

RESULTS

Direct Primer Extension Assay for Cyb mRNA—The assay described in Fig. 1*A* was used to detect U-insertions that occur within the first two editing sites of the Cyb pre-edited transcript. The mature *in vivo* edited mRNA has a single U at each of these sites. A DNA oligonucleotide annealed to the RNA immediately downstream of the first editing site was extended with RT in the presence of [α-^{32}P]dATP and the appropriate deoxy- and dideoxynucleoside triphosphates. Incorporation of label into the cDNA extension products is indicative of U-insertions in the mRNA. The extension of a second primer outside of the editing sites (indicated by *N* in Fig. 1*A*) was used to normalize the U-insertion signal.

As shown in Fig. 1*A*, the major cDNA extension product of a synthetic Cyb transcript edited at sites 1 and 2 was of the expected size: the length of the primer + 4 bases (P+4, *lane 1*). Assay of the Cyb mRNA endogenous to the lysate resulted in a product of the same size (*lane 2*). This is expected, since approximately 95% of the Cyb mRNA from late log phase *L. tarentolae* cells is fully edited (21). The bands migrating between the P+4 band and the major normalization N band, which are present in all lanes (see "Experimental Procedures"), represent artifacts of this normalization extension and are probably the result of the priming of the assay primer at an ectopic site on the normalization primer (Fig. 1*A*).

A major advantage of the primer extension assay is that insertions occurring in less than 1 in 20,000 molecules can be detected. At this level of sensitivity, however, artifacts resulting from misincorporation by the RT become significant. An example of such an artifact is the primer + 2 (P+2) band that appeared in the assays of all of the pre-edited Cyb mRNA

constructs tested, independent of exposure to the mitochondrial lysate (*lanes 3–12*). The artifact band was not seen with the edited Cyb transcript in *lanes 1* and *2* since extensions containing the misincorporation would still be extended up to the dideoxy termination site (primer + 4). The appearance of this band is consistent with the misincorporation of a labeled adenosine (A) opposite the cytidine (C) immediately 5′ of the first editing site. In any case, the bands of interest are only those that are dependent on incubation of pre-edited mRNAs with the mitochondrial lysate.

Detection of gRNA-independent U-insertions—The internal U-additions initially detected by an RNase H-based assay (13) were within a Cyb mRNA construct (pNB2 RNA) that contained at the 5′ end 73 nt of vector sequence in addition to 24 nt of upstream sequence that included part of the adjacent cytochrome oxidase III gene (22). As shown in Fig. 1*A* (*lanes 3* and *4*), the RT assay of an RNA containing 21 nt of this 5′ upstream sequence resulted in a ladder of extension products that is dependent on prior treatment with the mitochondrial lysate. The ladder is consistent with the RNA having been modified in the lysate with a heterogeneous number of U-insertions. Each successive band, progressing up the ladder, resulted from the primer extension of a subpopulation of RNAs containing an increasing number of inserted U nucleotides. For quantitation, these bands must be normalized for the increase in the amount of incorporated label. Approximately 1 in every 100 molecules contained at least one U-insertion. The insertions appear to be restricted to sites within the RNA that are normally edited *in vivo*; the extension of several DNA primers through non-edited regions of the lysate-treated RNA failed to detect inserted U nucleotides (data not shown).

Removal of the upstream sequence from the 5′ end of the Cyb mRNA construct resulted in the loss of the extension ladder (Fig. 1*A*, *lanes 5* and *6*); the remaining primer + 4 signal results from the detection of the Cyb mRNA endogenous to the lysate (*lane 2*). The secondary structure predicted by MFOLD (39), suggested that the role of the upstream sequence could be to stabilize a helix. As shown in Fig. 1*B*, part of this sequence is protected from DMS modification, which is consistent with the presence of a duplex.

Substitution of both the 5′ upstream sequence and the terminal 3′ 65 nucleotides of the construct with an artificial 5-bp G-C helix (Fig. 1*B*) preserved the U-insertion activity (Fig. 1*A*, *lanes 7* and *8*). Disruption of the artificial G-C helix by substitution of one strand with an AU sequence resulted in loss of the insertion activity, thus confirming the importance of the stabilized helix (Fig. 1*A*, *lanes 9* and *10*).

The U-insertions occurred in the absence of added gRNA. It was previously proposed that endogenous gRNAs could be mediating the U-insertions into the Cyb pre-edited region as assayed by the RNase H method (13). However, a change in the anchor binding site of the Cyb mRNA construct that should completely inhibit the binding of endogenous gRNA ($\Delta G^{o} > 0$) only resulted in a 40% decrease in the level of U-insertion into sites 1 and 2, as assayed by primer extension (Fig. 1*A*, *lanes 11* and *12*). In addition, a second independent Cyb mRNA construct also having a completely disrupted anchor binding site (S1 construct in Ref. 13) showed the same level of U-insertions as the wild type RNA (data not shown). This evidence suggests that the *in vitro* U-insertion activity occurring in sites 1 and 2 is independent of endogenous gRNA.

Indirect Primer Extension Assay for Cyb mRNA—A potential complication of the direct primer extension assay is the difficulty in differentiating between U nucleotides inserted within an editing site of an mRNA and U nucleotides present at the assayed sites as part of a gRNA-mRNA chimeric molecule.

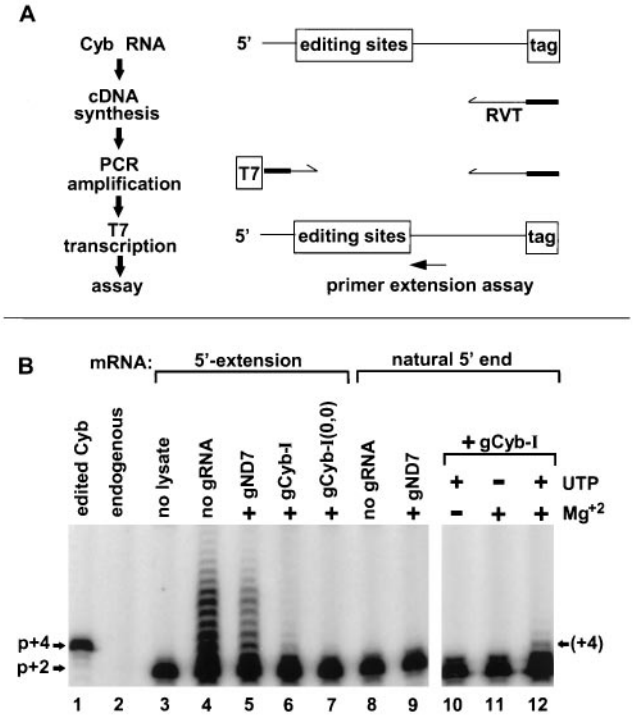

FIG. 2. **Indirect primer extension assay using RNA transcribed from RT-PCR amplified cDNA.** *A*, a cDNA copy of lysate-treated RNA was synthesized using a primer complementary to a tag sequence on the 3′ end of the substrate mRNA, thereby preventing reverse transcription of endogenous RNAs. The cDNA was amplified using a 5′ PCR primer sense to the 5′ end of the Cyb mRNA; molecules such as chimeras that do not contain both ends of the Cyb mRNA construct would not be amplified. RNA for the RT primer extension assay was transcribed from a T7 promoter incorporated into the 5′ PCR primer. *B*, use of the indirect assay to detect U-insertions. The substrate RNA is indicated above each lane. *Lanes 10–13* represent a separate experiment using the Cyb mRNA construct with the natural 5′ end in which the effect of addition of 1 mM UTP and 10 mM Mg^{2+} were analyzed. The primer+4 (P+4) bands are indicated.

Chimeras result from the covalent linkage of the 3′ oligo(U) tail of the gRNA to the mRNA at an editing site and are known to exist in steady state kinetoplast RNA and to be created *in vitro* by incubation of synthetic gRNA with mitochondrial extract (23–26). It is unclear whether chimeras represent true intermediates of the editing process or are a side product of the reaction.

Even though gRNA was not added to the reactions of Fig. 1, it is possible that chimeric molecules could be formed from the small pool of gRNAs that are endogenous to the mitochondrial lysate. The majority of the chimeras formed *in vitro* copurify with the Cyb mRNA construct containing the natural 5′ end; the distance between the 5′ end of this RNA and the first editing site is approximately the same size as the gRNA. The maximal possible contribution of these chimeric molecules to the primer extension signal, therefore, would not exceed that in *lane 5* of Fig. 1*A*.

The method diagrammed in Fig. 2*A* was used to completely eliminate any possible contribution of chimeric molecules and endogenous edited Cyb mRNA to the U-insertion signal. The mRNA construct used in this assay was modified at the 3′ end with a tag sequence. After incubation with the lysate, a cDNA copy of this RNA was synthesized using a primer complementary to the tag sequence, thereby eliminating any reverse transcription of endogenous Cyb mRNA. The cDNA was then PCR-amplified using the tagged 3′ primer together with a 5′ primer specific to the 5′ end of the Cyb mRNA; since chimeras and other potentially complicating ligation products do not contain

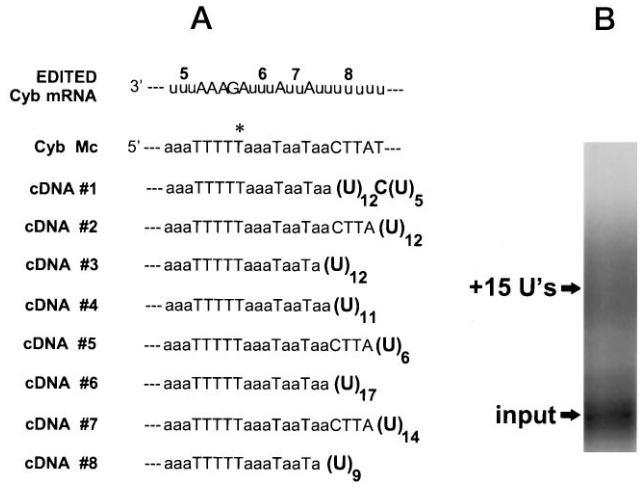

FIG. 3. **gCyb-I gRNA can potentially template U-insertions for the first seven editing sites of the Cyb mRNA.** *A*, editing sites 5–8 of fully edited Cyb mRNA are shown, with U nucleotides added by editing indicated by *u*. The sequence at the 3′ end of gCyb-I gRNA was obtained by directly sequencing maxicircle (*Mc*) DNA and several cDNA clones; the guiding nucleotides are in *lowercase*, and the T residue not present in the originally reported gRNA sequence is indicated by an *asterisk*. The cDNA clones showed a limited heterogeneity at the 3′ end of the genomically encoded sequence and a heterogeneity in the number of posttranscriptionally added U nucleotides (*bold*). *B*, a ^{32}P-labeled gCyb-I gRNA lacking an oligo(U) tail was incubated in mitochondrial lysate under the conditions used to detect the internal U-insertions, and the RNA was isolated and analyzed by acrylamide gel electrophoresis. Sequencing of the 3′ end of the RNA by 3′ RACE as described under "Experimental Procedures" confirmed that the elongation is caused by U-addition.

the 5′ end of the mRNA, they would not be PCR-amplified. The RT primer extension assay using [α-^{32}P]dATP was performed on RNA produced by T7 transcription of the RT-PCR DNA product.

When the pre-edited Cyb transcript containing the 5′ upstream sequence and the 3′ tag was incubated with mitochondrial lysate in the absence of added gRNA, the indirect assay produced a ladder of primer extension products similar to that obtained from the direct assay of the non-tagged Cyb mRNA (compare *lane 4* in Fig. 2*A* and *lane 3* in Fig. 1*A*). This confirms that chimeras did not significantly contribute to the extension signal obtained with the direct primer extension assay in Fig. 1*A*. The absence of a band in *lane 2* of Fig. 2*B* also indicates that this indirect assay succeeded in eliminating the primer + 4 signal arising from the edited Cyb mRNA that is endogenous to the lysate.

gCyb-I gRNA—The information required to correctly edit Cyb mRNA is potentially contained within two overlapping gRNAs. A single nucleotide error in the *L. tarentolae* maxicircle sequence (22) resulted in the original published sequence of the maxicircle-encoded gCyb-I gRNA being truncated by 20% (8). The corrected sequence of this gRNA potentially is able to template the U-insertions for the first seven editing sites and part of site 8, instead of the first five editing sites as previously suggested (Fig. 3*A*). This would create an anchor for the overlapping gCyb-II gRNA that could then be used to edit the remaining sites, in agreement with the scheme proposed for *Crithidia fasciculata* (27).

Cloning and sequencing of the endogenous gCyb-I gRNAs from the mitochondrial fraction revealed a limited heterogeneity of the 3′ uridylylation site, as well as heterogeneity in the number of posttranscriptionally added 3′ terminal U nucleotides (Fig. 3*A*) as is common with other gRNAs.

The gCyb-I gRNA used in the assay for *in vitro* U-insertions was transcribed from a DNA template that did not encode an oligo(U) tail, since *in vitro* transcription of gRNA by T7 RNA polymerase directly from a template encoding an oligo(U) 3′ tail resulted in extensive misincorporation (data not shown). Cloning and sequencing of the 3′ ends of individual gRNA molecules lacking the oligo(U) tail indicated that 90% of these RNAs were correctly transcribed and terminated. Fig. 3*B* shows that incubation of these transcripts with the mitochondrial extract under the conditions employed to detect U-insertions resulted in the 3′ addition of 15 ± 5 U nucleotides to approximately 20% of the transcript population by the endogenous mitochondrial terminal uridylyl transferase.

gRNA-dependent U-insertion Activity—Addition of the cognate gCyb-I gRNA to the reaction with the Cyb mRNA construct containing the 5′-extended sequence dramatically decreased the overall level of U-insertions, leaving a primer + 4 band and a background ladder (Fig. 2*B*, *lane 6*). Little insertion was seen on incubation with the gCyb-I(0,0) gRNA lacking guiding nucleotides for editing sites 1 and 2 (Fig. 2*B*, *lane 7*). Incubation with a non-cognate gRNA for editing of NADH dehydrogenase subunit 7 mRNA had a small inhibitory effect on the gRNA-independent U-insertion ladder.

Addition of the cognate Cyb gRNA to the Cyb mRNA containing the natural 5′ end brought about the appearance of a primer + 4 U-insertion band and a faint background ladder (Fig. 2*B*, *lane 12*). This gRNA-dependent U-insertion activity requires both Mg^{2+} and added UTP (Fig. 2*B*, *lanes 10–12*). Although the insertions are dependent on the presence of guiding nucleotides for sites 1 and 2, increasing the number of guiding nucleotides failed to result in a corresponding increase in the number of U nucleotides inserted into these sites (data not shown).

As described in the previous section, gRNA was synthesized without a 3′ oligo(U) tail, but it was added during the incubation by the endogenous terminal uridylyl transferase activity (Fig. 3). In a control experiment, the addition of an oligo(U) tail (15 ± 5 U nucleotides) with poly(A) polymerase to the 3′ end of the synthetic gCyb-I gRNA prior to incubation with lysate did not increase the level of insertion (data not shown).

The presence of the primer + 2 artifact band partially obscures the region of interest in these experiments and makes quantitation difficult, but it is clear that gRNA-dependent U-insertion does occur, and this is particularly evident with the mRNA substrate containing the natural 5′ end in *lane 12* of Fig. 2*B*.

DISCUSSION

Sensitive and specific primer extension assays were used to detect U-insertions occurring *in vitro* at sites 1 and 2 of pre-edited Cyb mRNA. The insertions were dependent on incubation of the RNAs in a detergent lysate of a purified mitochondrial fraction from *L. tarentolae*. With a 5′-extended Cyb mRNA transcript, a ladder of extension products representing insertions of up to 15 U nucleotides was observed in the absence of added gRNA. The extension ladder was not observed with a transcript having the 5′ end characteristic of fully edited Cyb mRNA ("natural 5′ end"). The possibility that contaminating gRNA-mRNA chimeric molecules were the source of the extension ladders was eliminated by performing the extension assay on RNAs transcribed from 3′-tagged RT-PCR-amplified cDNAs. Evidence was also presented that the U-insertions with the 5′-extended mRNA substrate can occur in the absence of a cognate anchor sequence, which indicates that the activity is independent of endogenous as well as exogenous gRNA.

These *in vitro* U-insertions are unlikely to have been mediated by misguiding by the endogenous gRNA pool (26, 28), as is known to occur at a very low level during *in vivo* editing of the Cyb mRNA (29), for several reasons. First, U-insertions were

detected in the Cyb mRNA constructs having the natural 5′ end only after exogenous gRNA was added to the lysate; the endogenous gRNA pool was not sufficient to mediate the reaction. When exogenous gRNA was added to the lysate, the presence of the 5′ extension on the Cyb mRNA construct did not significantly increase the level of U-insertion, suggesting that it does not strengthen the interaction with gRNA. Thus, if endogenous gRNAs were mediating insertions into the Cyb mRNAs with the 5′ upstream sequence, they would also have been expected to mediate insertions into the RNAs with the natural 5′ end.

Second, most, if not all, of the gRNAs in the UC laboratory strain of *L. tarentolae* used in this study have been identified (30, 31), and none of these is capable of mediating the insertions detected in the Cyb RNAs with a mutated anchor binding site (Fig. 1*B*, *lanes 11* and *12*). For example, a search of the UC strain gRNA collection for those complementary to the mutated anchor binding sites did not find any interactions with a ΔG_{27} better than −6.7 kcal/mol. The ΔG_{27} for the binding of gCyb-I gRNA to its anchor binding site is −19.1 kcal/mol. As a result, the K_d for the fortuitous binding of gRNAs to the mutated anchor binding sites would be approximately 10^9 times greater than that of the cognate gRNA interacting with its mRNA. Since the most abundant gRNA in the UC strain is only 500 times more abundant than gCyb-I gRNA,[2] the disfavorable dissociation constant for potential fortuitous interactions would not be compensated by a greater gRNA abundance. In the absence of other factors, insertions mediated by endogenous gRNAs interacting with the mutated anchor binding sites should be relatively rare events. Although it is possible that other components could be stabilizing the gRNA-mRNA interaction, the binding energy contributed by the anchor sequence and the evolutionary pressure required to maintain it cannot be ignored.

In this regard, it has been proposed that additional factors must be required to stabilize the gRNA-mRNA interaction, since some of the reported gRNA anchor sequences are too short to form a stable duplex on their own (32). The 8-base anchor binding site previously proposed for gCyb-II gRNA of *L. tarentolae* was one of the shortest in the literature (8). However, editing of the first seven sites by the corrected gCyb-I gRNA sequence potentially extends this anchor to 14 bases, which is closer to the mean anchor length of the other *L. tarentolae* gRNAs (33). Likewise, a 4-base anchor has been proposed for the same gRNA in *C. fasciculata* (32), but this would also be increased by the editing mediated by gCyb-I gRNA (27). There are probably other examples of putative gRNAs in the literature that are either a result of sequencing errors or fortuitous sequence similarity. Short anchor sequences might be considered an indicator of these complications, and this will eventually be tested with the *in vitro* system.

There is no evidence that the 5′-extended pre-edited mRNAs required for the gRNA-independent insertions *in vitro* exist in *vivo*. Primer runoff analysis of *L. tarentolae* endogenous edited Cyb RNA localized the 5′ end to the position shown in Fig. 1*B* (data not shown), which is in agreement with the previous mapping of the major Cyb transcript (34). Mapping of the 5′ end of pre-edited Cyb RNA from *L. tarentolae* gave ambiguous results, as the primer used to extend the pre-edited transcript cross-hybridized with additional RNAs (data not shown). The 5′ ends of both the major edited and pre-edited transcripts from *T. brucei*, however, were reported to be located at approximately the same position as that of the *L. tarentolae* edited transcript (35). Although it is possible that the Cyb transcript

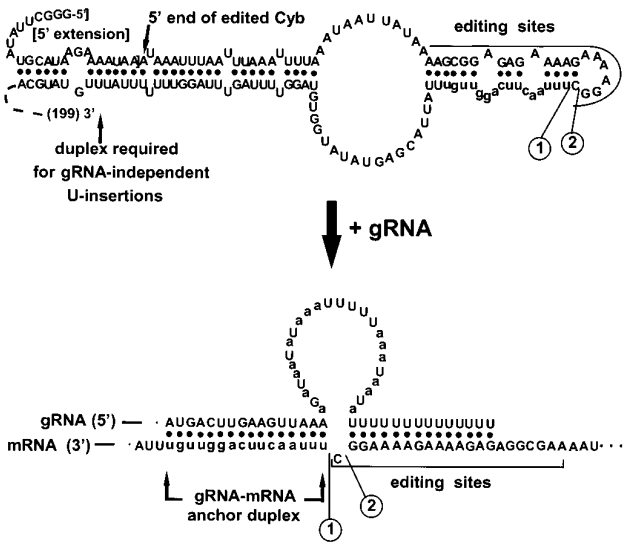

FIG. 4. **Possible relationship of secondary structure of the 5′-extended Cyb mRNA to the secondary structure of the postulated gRNA-mRNA hybrid.** *Upper panel*, secondary structure of the Cyb mRNA construct. Editing sites 1 and 2 are indicated, and the entire pre-edited region is *underlined*. The anchor-binding sequence is in *lowercase*. The helix required for gRNA-independent U-insertion into this RNA is indicated. *Lower panel*, secondary structure of the postulated gCyb-I gRNA/Cyb mRNA hybrid. Editing sites 1 and 2 are indicated, and the entire pre-edited region is *underlined*. The anchor helix is indicated, which we speculate fulfills a similar role in gRNA-dependent U-insertion activity as the indicated helix in the mRNA sequence shown above does in the gRNA-independent activity. The guiding A and G nucleotides in the single-stranded loop are in *lowercase*.

in *L. tarentolae* is made as part of a rapidly processed polycistronic RNA, there is not yet any evidence to support this hypothesis. It is interesting, however, that similar *in vitro* results have been obtained with ND7 pre-edited mRNA, in which a 5′-extended construct showed a higher level of gRNA-independent U-insertions than the non-extended construct (36).

Leaving aside the question of the biological relevance of the gRNA-independent reaction, it is possible that the structure of the RNAs supporting this *in vitro* reaction could be mimicking the RNA structures occurring during gRNA-mediated editing. For example, the duplex formed with the Cyb mRNA 5′ extension, which is required for the gRNA-independent reaction, may serve the same function as the anchor duplex created by the gRNA binding to the mRNA, as diagrammed in Fig. 4. Annealing of the cognate gRNA containing a complementary anchor sequence with the 5′-extended mRNA construct would be predicted to disrupt the mRNA secondary structure, thereby accounting for the observed inhibitory effect of the addition of exogenous gRNA on the gRNA-independent U-insertion activity.

Addition of exogenous synthetic gCyb-I gRNA was found to stimulate the insertion of U nucleotides into Cyb mRNA constructs with the natural 5′ end, and this stimulation was dependent on the presence of guiding nucleotides in the cognate gRNA (data not shown). The major primer + 4 band observed by the indirect assay in Fig. 2*B*, *lane 12*, could be interpreted as representing a single U insertion into each of the two editing sites as *in vivo*, or alternatively, as two insertions occurring within a single site. A clear demonstration of *in vitro* U-insertions into pre-edited mRNA being determined by base pairing with guiding nucleotides has been obtained recently using a pre-edited mRNA for NADH dehydrogenase subunit 7 and the cognate gRNA (36).

The creation of a recognition element for the assembly of the

[2] O. Thiemann and L. Simpson, unpublished results.

U-insertion machinery may represent an important role of the gRNA in the gRNA-mediated reaction. The recognition element could be the double-stranded RNA formed by the gRNA-mRNA interaction that may be mimicked by the intramolecular helix necessary for the gRNA-independent reaction. Alternatively, the recognition element could be an indirect consequence of the formation of the double-stranded RNA. However, DMS modification of the editing sites within the Cyb mRNA was not significantly altered by the addition of the 5′ extension, suggesting that the stabilized duplex supporting the gRNA-independent reaction does not result in major global changes to the structure (data not shown). Other parts of the mRNA in addition to the stabilized duplex, however, may still be involved in editing site recognition. For example, it has previously been proposed that the presence of the editing sites within a single-stranded loop is important for an endonucleolytic cleavage hypothesized by the enzyme cascade model to be part of the editing reaction (17, 37), but only limited evidence has been provided in support of this hypothesis (38).

Acknowledgments—We thank J. Alfonzo and O. Thiemann for helpful discussions. We also thank G. Frech for supplying the S1 construct.

REFERENCES

1. Simpson, L. (1972) *Int. Rev. Cytol.* **32,** 139–207
2. Simpson, L. (1987) *Annu. Rev. Microbiol.* **41,** 363–382
3. Simpson, L. (1986) *Int. Rev. Cytol.* **99,** 119–179
4. Simpson, L., and Shaw, J. (1989) *Cell* **57,** 355–366
5. Stuart, K. (1993) in *RNA Editing: The Alteration of Protein Coding Sequences of RNA* (Benne, R., ed) pp. 25–52, Ellis Horwood, New York
6. Hajduk, S. L., Harris, M. E., and Pollard, V. W. (1993) *FASEB J.* **7,** 54–63
7. Simpson, L., and Thiemann, O. H. (1995) *Cell* **81,** 837–840
8. Blum, B., Bakalara, N., and Simpson, L. (1990) *Cell* **60,** 189–198
9. Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1990) *Cell* **63,** 783–790
10. Sturm, N. R., and Simpson, L. (1990) *Cell* **61,** 879–884
11. Corell, R. A., Feagin, J. E., Riley, G. R., Strickland, T., Guderian, J. A., Myler, P. J., and Stuart, K. (1993) *Nucleic Acids Res.* **21,** 4313–4320
12. Seiwert, S. D., and Stuart, K. (1994) *Science* **266,** 114–117
13. Frech, G. C., Bakalara, N., Simpson, L., and Simpson, A. M. (1995) *EMBO J.* **14,** 178–187
14. Bakalara, N., Simpson, A. M., and Simpson, L. (1989) *J. Biol. Chem.* **264,** 18679–18686
15. Braly, P., Simpson, L., and Kretzer, F. (1974) *J. Protozool.* **21,** 782–790
16. Simpson, L., and Simpson, A. (1978) *Cell* **14,** 169–178
17. Simpson, A. M., Bakalara, N., and Simpson, L. (1992) *J. Biol. Chem.* **267,** 6782–6788
18. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol.* **180,** 51–62
19. Inoue, T., and Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 648–652
20. Simpson, L., and Braly, P. (1970) *J. Protozool.* **17,** 511–517
21. Feagin, J. E., Shaw, J. M., Simpson, L., and Stuart, K. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 539–543
22. de la Cruz, V., Neckelmann, N., and Simpson, L. (1984) *J. Biol. Chem.* **259,** 15136–15147
23. Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (1991) *Cell* **65,** 543–550
24. Blum, B., and Simpson, L. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 11944–11948
25. Koslowsky, D. J., Goringer, H. U., Morales, T. H., and Stuart, K. (1992) *Nature* **356,** 807–809
26. Harris, M. E., and Hajduk, S. L. (1992) *Cell* **68,** 1091–1099
27. Sugisaki, H., and Takanami, M. (1993) *J. Biol. Chem.* **268,** 887–891
28. Sturm, N. R., Maslov, D. A., Blum, B., and Simpson, L. (1992) *Cell* **70,** 469–476
29. Sturm, N. R., and Simpson, L. (1990) *Cell* **61,** 871–878
30. Maslov, D. A., and Simpson, L. (1992) *Cell* **70,** 459–467
31. Thiemann, O. H., Maslov, D. A., and Simpson, L. (1994) *EMBO J.* **13,** 5689–5700
32. Van der Spek, H., Arts, G.-J., Zwaal, R. R., Van den Burg, J., Sloof, P., and Benne, R. (1991) *EMBO J.* **10,** 1217–1224
33. Simpson, L., Maslov, D. A., and Blum, B. (1993) in *RNA Editing: The Alteration of Protein Coding Sequences of RNA* (Benne, R. ed) pp. 53–85, Ellis Horwood, New York
34. Simpson, A., Neckelmann, N., de la Cruz, V., Muhich, M., and Simpson, L. (1985) *Nucleic Acids Res.* **13,** 5977–5993
35. Feagin, J., Jasmer, D., and Stuart, K. (1987) *Cell* **49,** 337–345
36. Byrne, E., Connell, G. J., and Simpson L. (1996) *EMBO J.* **15,** 6758–6765
37. Harris, M., Decker, C., Sollner-Webb, B., and Hajduk, S. (1992) *Mol. Cell. Biol.* **12,** 2591–2598
38. Piller, K. J., Decker, C. J., Rusche, L. N., Harris, M. E., Hajduk, S. L., and Sollner-Webb, B. (1995) *Mol. Cell. Biol.* **15,** 2916–2924
39. Zuker, M. (1989) *Science* **244,** 48–52
40. Kable, M. L., Seiwert, S. D., Heidmann, C., and Stuart, K. (1996) *Science* **273,** 1189–1195