

Generation of Unexpected Editing Patterns in *Leishmania tarentolae* Mitochondrial mRNAs: Misediting Produced by Misguiding

Nancy R. Sturm,^{*†} Dmitri A. Maslov,^{*}
Beat Blum,^{*} and Larry Simpson^{*‡§}

^{*}Department of Biology

[‡]Howard Hughes Medical Institute

[§]Molecular Biology Institute

University of California, Los Angeles

Los Angeles, California 90024

Summary

We have analyzed the generation of unexpected patterns of RNA editing, i.e., those not following a strict 3' to 5' progression, which occur in junction regions between fully edited and preedited sequences. Evidence is presented that these patterns are generated by misediting due to specific events of misguiding. Misediting can occur through the interaction of inappropriate gRNAs with mRNAs or appropriate gRNAs in an incorrect fashion. Four possible mechanisms for the generation of misedited sequences are presented. Chimeric molecules have been detected in steady-state mitochondrial RNAs that are composed of misguiding gRNAs covalently linked to mRNAs at misediting sites by the 3' oligo(U) tail. We propose that misediting within junction regions can be corrected by appropriately acting gRNAs.

Introduction

RNA editing in mitochondria of kinetoplastid protozoa involves the addition and deletion of uridine residues (U's) within coding regions of maxicircle DNA transcripts (Benne, 1989; Simpson and Shaw, 1989; Simpson, 1990). The information for the exact sites and the precise number of U's is provided by specific guide RNAs (gRNAs) that can form perfect hybrids (allowing G–U and occasional A–C pairing) with the mature edited mRNAs (Blum et al., 1990; Van der Spek et al., 1991). The 5' portions of specific gRNAs also have complementary sequences of variable length to the mRNA sequence just 3' of the preedited region (PER) of the specific target mRNA. These anchor duplex regions are thought to provide the specificity for initiation of the editing process at the 3' end of the PER, and the process then continues in a 5' direction. The gRNAs are encoded both in the maxicircle (Blum et al., 1990; Van der Spek et al., 1991) and the minicircle (Sturm and Simpson, 1990a, 1991; Pollard et al., 1990) components of the kinetoplast DNA.

Two specific hypotheses—the enzyme cascade (Blum

et al., 1990) and the transesterification (Blum et al., 1991; Cech, 1991) models—have been proposed for gRNA-mediated RNA editing, and the available evidence is not sufficient to decide between them. In both models, however, base pairing of the edited mRNA with the gRNA provides the driving force for the 3' to 5' editing that occurs within a single editing block, which is defined as the edited sequence determined by a single gRNA. The overall 3' to 5' polarity of editing observed within editing domains (Abraham et al., 1988; Sturm and Simpson, 1990b; Decker and Sollner-Webb, 1990), which are defined as edited sequences determined by multiple overlapping gRNAs, is due to the fact that the anchor duplexes of gRNAs other than the first require edited mRNA sequences (Maslov and Simpson, 1992 [this issue of *Cell*]).

An analysis of a large collection of partially edited mRNAs for the cytochrome b (*CYb*) gene in *Leishmania tarentolae* agreed well with the mandatory processivity within a block and a domain as suggested by these models (Sturm and Simpson, 1990b). However, 42% of cloned partially edited mRNAs for the cytochrome oxidase subunit III (*COIII*) gene of *L. tarentolae* (Sturm and Simpson, 1990b) and the majority of a library of partially edited mRNAs for *CYb* and *COIII* genes of *Trypanosoma brucei* (Decker and Sollner-Webb, 1990) exhibited unexpected editing patterns at junction regions between fully edited and unedited sequences. We have suggested that the generation of these unexpected editing patterns is due to 3' to 5' misediting by inappropriate gRNAs (Sturm and Simpson, 1990a, 1990b; Maslov et al., 1992). We have also shown an example of misediting in a *COIII* mRNA caused by the formation of a secondary anchor by the *gCOIII-1* gRNA just upstream of a single nucleotide loopout (Blum et al., 1991).

Alternative explanations for misediting within junction regions have also been proposed. Decker and Sollner-Webb (1990) have suggested that editing occurs randomly at multiple sites within a defined region, with hybridization of gRNA to correctly edited mRNA protecting the mRNA from further random editing. The unexpected patterns that occur within junction regions represent sequences that are not yet complementary to the gRNA guide sequences. Koslowsky et al. (1991) have proposed that regions of lower thermodynamic stability in the initially imperfect gRNA–mRNA hybrids are targets for editing. Cycles of progressive realignments and editing are continued until perfect correspondence between mRNA and gRNA is achieved. Like random editing, this model suggests that misedited junction sequences represent natural intermediates of the editing process.

In this paper, we present additional experimental and analytical evidence for the hypothesis in which misediting is mediated by misguiding gRNAs that mediate editing in a strict 3' to 5' direction. These incorrectly edited sequences require reediting with the correct gRNA in the correct context and guiding frame to produce a mature mRNA.

[†]Present address: Boyce Thompson Institute for Plant Research, Cornell University, Tower Road, Ithaca, New York 14853.

Results

Several Misediting Patterns in Partially Edited *COIII* mRNAs Correlate with Specific *RPS12* and *MURF4* gRNAs

The gRNA transcribed from minicircle Lt154 was first identified as being complementary to unexpected editing pattern 4E-2 of a partially edited *COIII* mRNA (Sturm and Simpson, 1990b). This gRNA was then shown to mediate correct editing of sites 24–32 in the *RPS12* mRNA (*gRPS12-IV*) (Maslov et al., 1992). The *gRPS12-IV* gRNA was the first example of a gRNA that has the potential for specifying both a mature editing pattern for the corresponding mRNA and a misedited pattern for a heterologous mRNA.

The misediting patterns in junction regions of partially edited *L. tarentolae COIII* mRNAs could be separated into subgroups according to the extent of the 3' to 5' progression of editing (see Figures 4 and 5 in Sturm and Simpson, 1990b). As shown in Figures 1A and 1B, several of these misediting patterns correspond well with specific *RPS12* and *MURF4* gRNAs. This evidence suggests that these *RPS12* and *MURF4* gRNAs may be responsible for the generation of at least some of the observed *COIII* misediting patterns by the formation of false anchors and subsequent misguiding.

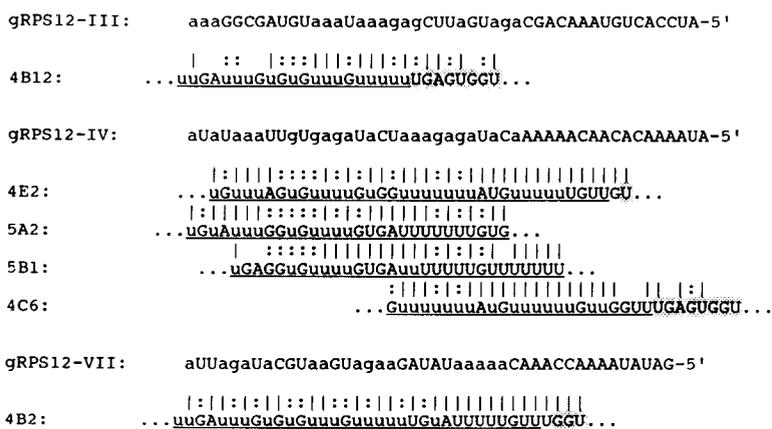
Additional correlations of misedited sequences and specific gRNAs were observed for several partially edited se-

quences from G-rich region 5 (Simpson and Shaw, 1989) and the gRNAs for *ND7-I*, *MURF2-II*, and *MURF4-IV* (D. A. M., E. S. Gruszynski, N. R. S., and L. S., unpublished data).

Confirmation of an Interaction between the Misguiding *gRPS12-IV* gRNA and the *COIII* mRNA by Detection of Chimeric Molecules

Direct evidence for the involvement of *gRPS12-IV* gRNA in misediting of *COIII* mRNA was obtained by polymerase chain reaction (PCR) amplification of the corresponding gRNA-mRNA chimeric molecules, which are predicted by the transesterification model of editing. We selected chimeric molecules composed of misguiding gRNAs covalently linked to mRNAs, using a *COIII* mRNA-specific 3' primer and a *gRPS12-IV*-specific 5' primer. Chimeric molecules were obtained in which *gRPS12-IV* was attached to *COIII* mRNA at several sites. The chimeric molecules in Figure 2A are shown in foldback configurations indicating the most likely local base pairing as determined by the local alignment GCG BESTFIT program (see Experimental Procedures). In each case, the base pairing between the gRNA and mRNA included a potential duplex anchor, which could have resulted in an editing event leading to the attachment of the misguiding gRNA. In some cases, prior editing by *gCOIII-I* (molecules 1 and 4) or another putative misguiding gRNA (molecule 5) would be required to form a stable anchor. In molecules 2, 3, and 6, the

A



B



Figure 1. Correlation of *COIII* Misedited Sequence Patterns with *RPS12* and *MURF4* gRNA Sequences

The *COIII* sequences represent unexpected editing patterns of partially edited *COIII* mRNAs obtained by Sturm and Simpson (1990b). The clones are designated by the original figure and clone numbers in Sturm and Simpson (1990b). The patterns were grouped according to shared sequence motifs. Only single representatives of such groups are shown here matching with the putative misguiding gRNAs. gRNA sequences are shown 3' to 5'. G-C and A-U base pairs are indicated by a vertical line, G-U base pairs by a colon. Uridines added by editing are shown as u's. Edited mRNA sequences are indicated by stippling. Misedited sequences are underlined. (A) *gRPS12-III*, *gRPS12-IV*, and *gRPS12-VII* gRNAs. (B) *gMURF4-II* gRNA.

mRNA. Figure 6B shows that the sequence of *gRPS12-VI* corresponds to that of edited mRNA up to editing site 46. The exact location of the 3' end of this gRNA is unknown, but if the 3' end extends beyond this site, the last several bases would result in generation of a short misedited sequence covering sites 47–50, a situation that was actually observed in clones 4 and 14 (see also Figure 6 of Maslov et al., 1992). A 7 bp anchor then could be formed between this sequence and *gRPS12-I* gRNA, which normally edits sites 1–6. The subsequent misediting could create the unexpected patterns seen in a number of *RPS12* clones.

Discussion

We have proposed four basic mechanisms by which misedited sequences could be generated by misguiding in a strict 3' to 5' fashion, as shown schematically in Figure 7. Possible examples of each type of misguiding are presented. The correlations of misedited patterns and gRNAs provide strong, though indirect, evidence in favor of the misguiding hypothesis. If unexpected editing patterns were derived from random editing, as suggested by Decker and Sollner-Webb (1990), or from the reediting–realignment cycles of the dynamic interaction model of Koslowsky et al. (1991), the probability of finding any pattern correlations at all between unexpected patterns and gRNA sequences would be extremely low.

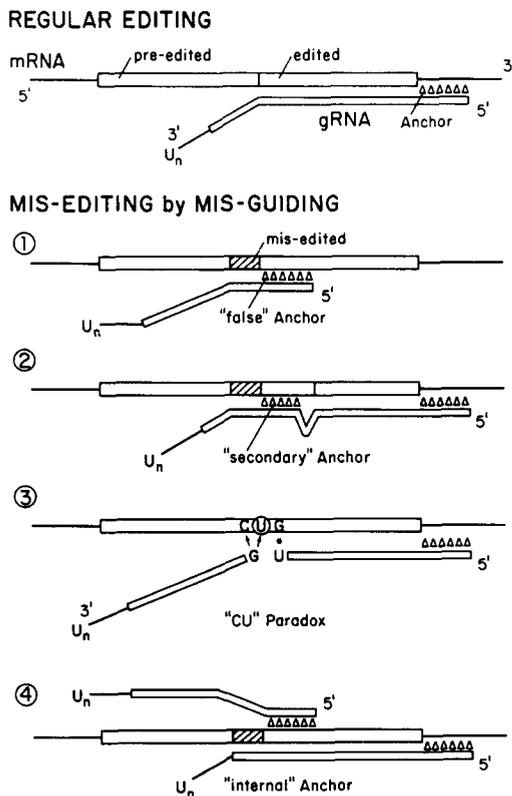


Figure 7. Schematic Diagrams of Normal Editing and Misediting Produced by Misguiding
The anchors are indicated by carets, and misedited sequences are indicated by cross-hatching. See text for explanation of models.

The formation of anchor duplexes (i.e., false anchors) by inappropriate gRNAs represents one type of initiating event in misediting (Figure 7, misguiding type 1). In addition to the indirect evidence for this mechanism by correlating misedited sequences with known gRNA sequences, some direct evidence has been obtained in the form of the occurrence of chimeric molecules with inappropriate gRNA–mRNA pairs.

Wobble G–U base pairing is the major factor responsible for the appearance of spurious anchor formations. The presence of more G and U residues in the *L. tarentolae COIII* preedited region than in the *CYb* preedited region and the subsequent greater potential for wobble base pairing was suggested to be responsible for the observed higher frequency of misediting in the former (Sturm and Simpson, 1990b). In this regard, Maslov and Simpson (1992) have shown that a characteristic feature of many proper anchors is the virtual absence of G–U base pairs.

The frequency of occurrence of a particular misediting pattern should also depend on the relative abundance of a corresponding misguiding gRNA. In the case of *RPS12* (Figures 2 and 6 in Maslov et al., 1992), many misedited patterns were associated with the 5' domain of the molecule, where a false anchor for *gRPS12-I* gRNA can be created by the putative 3' extended end of *gRPS12-VI* gRNA. *gRPS12-I* is one of the most abundant gRNAs in steady-state kinetoplast RNA (Maslov and Simpson, 1992).

The misediting produced by the formation of a secondary anchor that is created independently or as a result of a loopout or bulge in either the gRNA or the mRNA (Figure 7, misguiding type 2) is due to a shift in the gRNA–mRNA guiding frame. This mechanism, which was first observed in a gRNA–mRNA chimeric molecule by Blum et al. (1991), has been expanded in the dynamic interaction model of Koslowsky et al. (1991) to include multiple misalignments and misediting occurring as normal intermediates of the editing process.

The CU paradox (Feagin, 1990), misediting produced by mishybridization of a gRNA guide G residue to a U in the mRNA 3' of a C (Figure 7, misguiding type 3), is another example of frameshift misediting. We have proposed that deletions of U's are caused either by a 3' exonuclease trimming of the unpaired terminal U in the cleaved mRNA (Blum et al., 1990) or by an initial transesterification 5' of the unpaired U residue rather than 3' and a subsequent incorporation of the U into the gRNA tail after the second transesterification (Blum et al., 1991; Cech, 1991). In either case, base pairing of the guide G to the U would prevent the deletion and initiate misediting. We speculate that there is an equilibrium between base pairing of the G to the U or to the C, which is perhaps driven toward the latter by the greater free energy of the G–C base pair. This would allow U deletion and correct editing or reediting of the misedited sequence.

Finally, some other misediting events tend to occur at the locations of the 3' ends of gRNAs, possibly reflecting gRNA 3' heterogeneity or misediting mediated by the oligo(U) tail (Figure 7, misguiding type 4).

Misedited patterns produced by misguiding events can

not be extended unless a false anchor sequence is inadvertently formed for a misguiding gRNA. A possible example of this has been presented in the case of *RPS12* editing.

A characteristic feature of misedited patterns is a preferential localization to a fairly well-defined junction region between completely edited and preedited sequences within an editing domain in partially edited molecules. This is functional, for misediting outside of an editing domain could not be repaired by reediting. Limitation to a junction region is a natural consequence of the misguiding mechanisms illustrated as types 2, 3, and 4 in Figure 7. In the case of false anchor formation (misguiding type 1), the limitation to sequences within an editing domain could be a result of the same mechanism that limits normal editing to a junction region. Bakker et al. (A. Bakker, M. Peris, and L. S., unpublished data) have shown recently that gRNAs are bound to a 200–700 kd complex containing terminal uridylyltransferase (Bakalara et al., 1989) and several other proteins and have suggested that this complex is involved with the specific presentation of gRNAs to the editing site. In addition, Harris and Hajduk (1992), Koslowsky et al. (1992), and B. B. and L. S. (submitted) have shown that in vitro formation of gRNA–mRNA chimeric molecules requires the addition of mitochondrial extract, presumably owing to a requirement for similar if not identical protein factors. Limitation of misediting to the junction region that would result from the presentation by such a complex of an inappropriate gRNA could form a false anchor in that region and initiate misediting in a 5' direction.

We have shown that multiple examples of misediting can be explained by a local 3' to 5' processive model involving a misguiding mechanism, and we suggest that this could provide a general explanation for the occurrence of unexpected editing patterns in partially edited RNAs. Misguiding represents mechanically correct RNA editing, but misedited mRNAs must always be corrected by in-frame editing mediated by the appropriate gRNA to obtain the mature mRNA.

Experimental Procedures

Cell Culture and Kinetoplast RNA Isolation

L. tarentolae cells (UC strain) were grown as described previously (Simpson and Braly, 1970). Cells were harvested at mid-log phase and used for mitochondrial isolation for the preparation of kinetoplast RNA as described (Braly et al., 1974; Simpson and Simpson, 1978), except that kinetoplast RNA was isolated from the enriched kinetoplast fraction without centrifugation through Renografin.

Chimeric Amplification, Cloning, and Sequencing

The chimeric molecules were amplified by RNA PCR using a *COIII* mRNA-specific 3' primer, S-71, which is downstream of the preedited region, and a *Lt154* gRNA (*gRPS12-IV*)-specific 5' primer, S-397, as described (Blum et al., 1991). The products of the amplification were ligated directly from the PCR reaction mix using the TA Cloning Kit (Invitrogen). Colonies were screened by colony hybridization with an internal *gRPS12-IV*-specific primer (S-252) or selected on the basis of color and screened for inserts by sizing on agarose gels. Sequencing was performed on plasmid DNA extracted by the boiling method miniprep procedure with the Sequenase Kit (US Biochemical Company).

Oligonucleotides

Oligonucleotides were synthesized by standard phosphoramidite methods on an ABS 341A DNA Synthesizer. The following is a list of oligonucleotides used in this study.

S-71 (*COIII* mRNA 3' PCR primer, nucleotides 4708–4692 in LEIKP-MAX): GTCTACAAATAAGTCAC.

S-252 (*gRPS12-IV* probe): AACTCTATGATTTCTCTAT.

S-397 (*gRPS12-IV* 5' PCR primer): ATAGAATTCATAAAACACAA-CAAAAAA.

S-466 (*gM150* probe, nucleotides 102–120 in LEIKP-MAX): AACAAAT-TGTTTCATATA.

Computer Analysis

To search for the most stable gRNA–mRNA anchor sequences, the UWGCG BESTFIT program was used with a modified data file, SWGAPDNA.CMP, in which mismatches were given a value of –0.9; A–C base pairs, 0.01; G–C base pairs, 1.0; A–U base pairs, 0.5; and G–U base pairs, 0.25. The gap weight was 100, and the gap length weight was 2.

Acknowledgments

We thank all the members of the Simpson laboratory for stimulating discussion and technical advice. This research was supported in part by research grant AI-09102 to L. S. from the National Institutes of Health. N. R. S. was a predoctoral trainee supported in part by United States Public Health Service National Research Service Award GM-07104.

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 USC Section 1734 solely to indicate this fact.

Received April 30, 1992.

References

- Abraham, J. M., Feagin, J. E., and Stuart, K. (1988). Characterization of cytochrome c oxidase III transcripts that are edited only in the 3' region. *Cell* 55, 267–272.
- Bakalara, N., Simpson, A. M., and Simpson, L. (1989). The *Leishmania* kinetoplast-mitochondrion contains terminal uridylyltransferase and RNA ligase activities. *J. Biol. Chem.* 264, 18679–18686.
- Benne, R. (1989). RNA editing in trypanosome mitochondria. *Biochim. Biophys. Acta* 1007, 131–139.
- Blum, B., Bakalara, N., and Simpson, L. (1990). A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60, 189–198.
- Blum, B., Sturm, N. R., Simpson, A. M., and Simpson, L. (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.
- Braly, P., Simpson, L., and Kretzer, F. (1974). Isolation of kinetoplast-mitochondrial complexes from *Leishmania tarentolae*. *J. Protozool.* 21, 782–790.
- Cech, T. R. (1991). RNA editing: world's smallest introns? *Cell* 64, 667–669.
- Decker, C. J., and Sollner-Webb, B. (1990). RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell* 61, 1001–1011.
- Feagin, J. E. (1990). RNA editing in kinetoplastid mitochondria. *J. Biol. Chem.* 265, 1–4.
- Harris, M. E., and Hajduk, S. L. (1992). Kinetoplastid RNA editing: in vitro formation of cytochrome b gRNA–mRNA chimeras from synthetic substrate RNAs. *Cell* 68, 1091–1099.
- Koslowsky, D. J., Bhat, G. J., Read, L. K., and Stuart, K. (1991). Cycles of progressive realignment of gRNA with mRNA in RNA editing. *Cell* 67, 537–546.
- Koslowsky, D. J., Gooringer, H. U., Morales, T. H., and Stuart, K. (1992). In vitro guide RNA/mRNA chimaera formation in *Trypanosoma brucei* RNA editing. *Nature* 356, 807–809.
- Maslov, D. A., and Simpson, L. (1992). The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell* 70, this issue.

- Maslov, D. A., Sturm, N. R., Niner, B. M., Gruszynski, E. S., Peris, M., and Simpson, L. (1992). An intergenic G-rich region in *Leishmania tarentolae* kinetoplast maxicircle DNA is a pan-edited cryptogene encoding ribosomal protein S12. *Mol. Cell. Biol.* *12*, 56–67.
- Pollard, V. W., Rohrer, S. P., Michelotti, E. F., Hancock, K., and Hajduk, S. L. (1990). Organization of minicircle genes for guide RNAs in *Trypanosoma brucei*. *Cell* *63*, 783–790.
- Simpson, L. (1990). RNA editing—a novel genetic phenomenon. *Science* *250*, 512–513.
- Simpson, L., and Braly, P. (1970). Synchronization of *Leishmania tarentolae* by hydroxyurea. *J. Protozool.* *17*, 511–517.
- Simpson, L., and Shaw, J. (1989). RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa. *Cell* *57*, 355–366.
- Simpson, L., and Simpson, A. M. (1978). Kinetoplast RNA of *Leishmania tarentolae*. *Cell* *14*, 169–178.
- Sturm, N. R., and Simpson, L. (1990a). Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. *Cell* *61*, 879–884.
- Sturm, N. R., and Simpson, L. (1990b). Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from *Leishmania tarentolae* mitochondria: RNA editing intermediates. *Cell* *61*, 871–878.
- Sturm, N. R., and Simpson, L. (1991). *Leishmania tarentolae* minicircles of different sequence classes encode single guide RNAs located in the variable region approximately 150 bp from the conserved region. *Nucl. Acids Res.* *19*, 6277–6281.
- Van der Spek, H., Arts, G.-J., Zwaal, R. R., Van den Burg, J., Sloof, P., and Benne, R. (1991). Conserved genes encode guide RNAs in mitochondria of *Crithidia fasciculata*. *EMBO J.* *10*, 1217–1224.