Cell, Vol. 81, 837–840, June 16, 1995, Copyright © 1995 by Cell Press

Sense from Nonsense: RNA Editing in Mitochondria of Kinetoplastid Protozoa and Slime Molds

Minireview

Larry Simpson and Otavio H. Thiemann
Howard Hughes Medical Institute
Departments of Biology and Medical Microbiology and Immunology
University of California, Los Angeles
Los Angeles, California 90024

The reworking of the sequences of mitochondrial DNA transcripts is an intriguing genetic phenomenon that continues to create amazement even today, 9 years after the discovery of uridine (U) insertion/deletion editing in trypanosomatid mitochondria. Upon initial exposure to this phenomenon, the usual scientific questions of what and how were rapidly followed by the more philosophical question of why. In fact, some answers to both types of questions have been obtained recently, and these will be discussed in this minireview.

What?

Kinetoplastid Protozoa

The kinetoplastid protozoa contain two major taxonomic subgroups—the trypanosomatids and the bodonids/cryptobiids—as determined both through morphology and by the tools of molecular phylogeny. Phylogenetic reconstructions of trypanosomatids, using rRNA sequences, showed that the digenetic African trypanosome Trypanosoma brucei represents the deepest branch and that the digenetic lizard parasite Leishmania tarentolae and the monogenetic insect parasite Crithidia fasciculata comprise a more recently evolved monophyletic clade (Fernandes et al., 1993; Landweber and Gilbert, 1994; Maslov and Simpson, 1994). Editing has, to date, been investigated in 12 trypanosomatid species and in one cryptobiid species.

The mitochondrial genome in these cells is termed kinetoplast DNA (kDNA) and consists of two molecular spe-

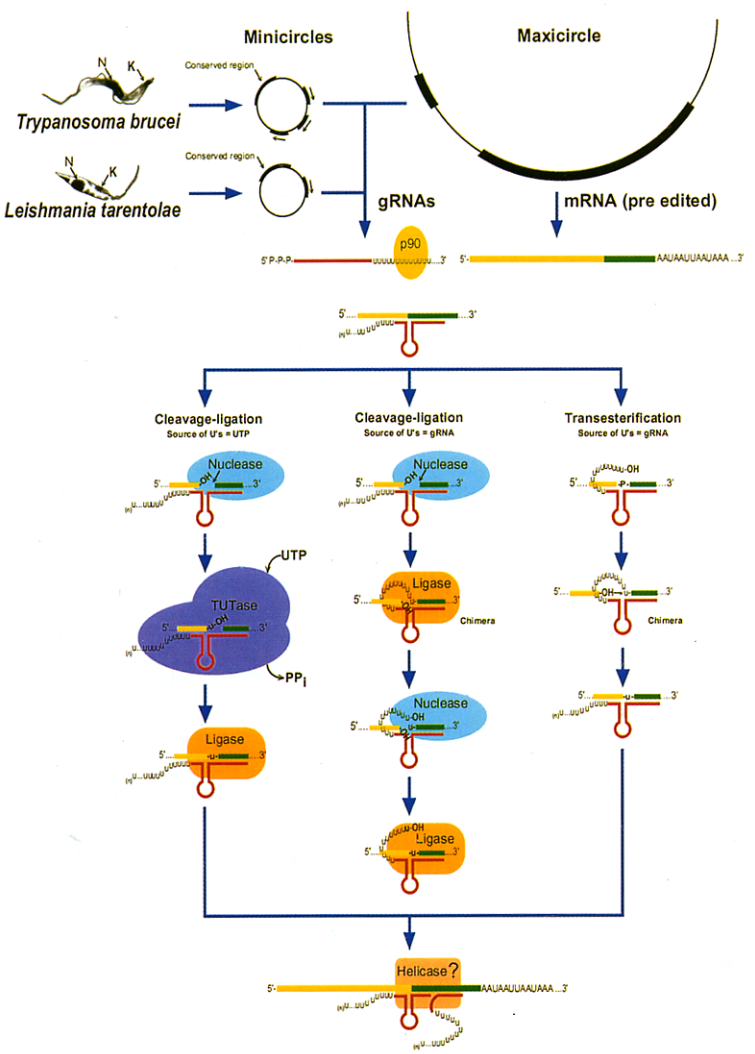

Figure 1. Diagrammatic Representation of Models of RNA Editing in Trypanosomatid Mitochondria

Cells are stained with Giemsa. Abbreviations: K, kinetoplast; N, nucleus. Anchor sequence in the mRNA is in green, the preedited region in yellow. The p90 protein from T. brucei (Köller et al., 1994) is shown as an example of an oligo U–binding protein. The gRNA stem–loop structure shown has not yet been tested experimentally. Indication of the enzymatic activities by colored areas is not meant to imply any structural information. The question mark next to the helicase (Missel and Goringer, 1994) indicates that this is a hypothetical role for this activity.

cies: ~50 catenated maxicircles (23–36 kb in size, depending on the species) that contain the rRNA genes, protein-coding genes, a few guide RNA (gRNA) genes, and ~5–10,000 catenated minicircles that encode gRNA genes. The transcripts of 11 or 12 out of the 20 identified maxicircle structural genes are edited to varying extents (for reviews see Simpson et al., 1993; Hajduk et al., 1993; Stuart, 1993; Benne, 1994).

gRNAs are antisense to portions of edited mRNAs (allowing GU base pairs), and can form short "anchor" duplexes with preedited mRNA just downstream of the sequence that is to be edited (Figure 1). The observed overall 3′ to 5′ polarity of editing within a multiple gRNA-mediated editing domain is due to the creation of upstream anchor sequences by downstream editing. In L. tarentolae, complete sets of gRNAs overlapping by the anchor sequences have been described for the *RPS12* and *MURF4* cryptogenes.

Owing to the GU "wobble" base pairing, it is possible for gRNAs of different guiding sequences to encode the same editing information, but only a single example of such a "redundant" gRNA has been found in L. tarentolae (Thiemann et al., 1994). In contrast, there is a much higher abundance of redundant gRNAs in T. brucei (Corell et al., 1993).

After a gRNA-mediated "block" has been edited, the mRNA–gRNA duplex must be melted for the adjacent upstream gRNA to form an anchor duplex for the initiation of the next round of editing. This can occur by means of an RNA helicase activity (Missel and Goringer, 1994), by means of a lower stability of the edited mRNA–gRNA duplex due to an abundance of GU base pairs (Simpson et al., 1993), or by both means.

The number of different gRNAs, which are transcribed mainly from the minicircle molecules, varies from species to species. In T. brucei, there may be over 300 different minicircle classes, and each minicircle usually encodes three gRNAs between sets of 18-mer inverted repeats. To date, over 63 T. brucei gRNAs have been identified (Corell et al., 1993). In L. tarentolae, 13 maxicircle-encoded and 43 minicircle-encoded gRNAs have been cloned and identified so far (Thiemann et al., 1994), of a predicted total of ~80 gRNAs. In this species, there is a single gRNA gene per minicircle.

The extent of the gRNA complexity is correlated with the mode of replication of the kDNA. The single kDNA network, which appears after cell lysis as a cup-shaped sheet of catenated DNA circles, is organized in situ as a disk-shaped structure termed the nucleoid body, which has approximately the width of a single minicircle. In T. brucei, the kDNA is stationary, and newly replicated minicircles are recatenated at two sites termed replisomes located at either side of the nucleoid (Ferguson et al., 1994). In C. fasciculata (and also in L. tarentolae), the nucleoid body apparently rotates, and newly replicated minicircles are recatenated along the entire periphery of the network (Pérez-Moraga and Englund, 1993). The large gRNA complexity and redundancy in T. brucei apparently limit the loss of minicircle classes by missegregation at mitochondrial division, which would more readily occur with a stationary kinetoplast nucleoid; in L. tarentolae and C. fasciculata, the molecules are more randomly distributed in the replicated network, and this is correlated with a decreased minicircle complexity and redundancy.

The gRNA complexity of an old laboratory strain of L. tarentolae was found to be significantly less than that of a recently isolated strain (Thiemann et al., 1994). The latter contained at least 32 additional minicircle-encoded gRNAs that were not present in the former. These gRNAs encoded information for the editing of transcripts of cryptogenes that encode several components of complex I of the respiratory chain: ND3, ND8, and ND9. It was hypothesized that specific minicircle sequence classes were lost during the culture history of the old laboratory strain and that the mutant cells survived the subsequent disruption of editing owing to the lack of a requirement for these protein products during the culture stage of the life cycle. It was concluded that, as a consequence of the localization of gRNA genes on minicircles, editing is a labile genetic trait that is easily lost in the absence of a selective pressure to maintain the edited product.

The absence of certain minicircle-encoded gRNAs in the old laboratory strain of L. tarentolae has led to an illustrative situation of misediting caused by misguiding in transcripts of the *G5* cryptogene (Thiemann et al., 1994). A set of partially edited *G5* transcripts exhibit correct editing of block I, since this is mediated by a maxicircle-encoded gRNA, but have upstream misediting, which is apparently mediated by two noncognate gRNAs.

Editing also exists in the cryptobiid, Trypanoplasma borreli (Lukes et al., 1994; Maslov and Simpson, 1994). The mitochondrial gene order in T. borreli is different from that in the maxicircle genome of the trypanosomatids, but editing and pan-editing appear to proceed by a similar mechanism. However, there are no minicircles in this species. Instead, there is a 1 kb repetitive component organized in tandem repeats in 200 kb circular molecules. This component encodes small transcripts with similarities to gRNAs from trypanosomatids (Maslov and Simpson, 1994).

Physarum polycephalum

Mitochondrial DNA transcripts are modified by the insertion of single nonencoded C residues at multiple precise sites. At a lower frequency, insertions of G and U residues, insertions of certain dinucleotides, and C to U substitutions have also been observed (Mahendran et al., 1994; Gott et al., 1993). There is a striking nonrandom pattern to the editing process, in which Cs are inserted at intervals of 26–10 nt throughout the coding regions of the transcripts from all of the structural genes yet identified and at intervals of ~40 nt in the rRNAs. Of the five mitochondrial tRNAs examined, four also are edited by C (and U) insertions that create canonical base pairs in stem regions. This represents a striking example of a mixed nucleotide insertional and substitutional editing system.

How?

Kinetoplastid Protozoa

The edited sequence information resides in the gRNA molecules, but the precise mechanism for the insertion and

deletion of U's is not yet clear. Two basic mechanisms have been proposed, with several variations of each (see Figure 1) (Simpson et al., 1993). Both models propose that a mismatch between a gRNA and a preedited mRNA identifies an editing site, but disagree on the mechanistic aspects of this process. The transesterification model has the advantage of an appealing simplicity and a mechanistic relationship to RNA splicing. However, the only evidence for this model so far is the presence of the predicted intermediate mRNA–gRNA chimeric molecules in steady-state kRNA. These could very well be artifacts or even true intermediates of an enzymatic cleavage–ligation process. In fact, formation of chimeric molecules has been shown to occur in vitro using synthetic preedited mRNA and a cognate gRNA upon incubation with mung bean nuclease and T4 RNA ligase (Piller et al., 1995).

The enzyme cascade model invokes an endonuclease cleavage at a base mismatch, the transfer of a U either directly from UTP or from the 3′ end of the gRNA, and a ligation. Several of the predicted enzymatic activities cosediment in ribonucleoprotein (RNP) complexes from mitochondrial extracts.

RNP Complexes Possibly Involved in Editing

Two gRNA-containing RNP complexes in mitochondrial extracts from T. brucei can be detected by sedimentation in glycerol gradients: a 19S complex I that contains gRNA, TUTase, RNA ligase, and mRNA–gRNA chimera-forming activity, and a 35S complex II that has in addition preedited RNA but lacks tightly bound TUTase (Pollard et al., 1992). A preedited domain-specific endonuclease (Simpson et al., 1993) cosediments with complex I from T. brucei extracts (Piller et al., 1995), and it was concluded that chimera formation occurs through a cleavage–ligation mechanism rather than by transesterification (Rusche et al., 1995). This conclusion is strengthened by the finding of two adenylated proteins that show properties characteristic of RNA ligase that cosediment with both complex I and complex II (Sabatini and Hajduk, 1995).

However, it is not yet established that the in vitro chimera-forming activity represents the in vivo mechanism, especially since the in vitro chimeras show differences from the in vivo chimeras (Simpson et al., 1993). In addition, there is no evidence that chimeras actually represent true intermediates in the editing process.

Several groups have identified by gel retardation analysis mitochondrial RNP complexes from T. brucei that interact with exogenous synthetic gRNAs (Goringer et al., 1994; Köller et al., 1994). In addition, eight gRNA-interacting proteins were detected ranging from 9–124 kDa (Köller et al., 1994; Leegwater et al., 1995). The 90 kDa protein interacted specifically with the 3′ oligo U tail of gRNAs. In C. fasciculata, a 65 kDa mitochondrial protein was found to have a high affinity for the 3′ oligo U tails of gRNAs (Leegwater et al., 1995). Proteins of 88 kDa and 30 kDa also bind to gRNA 3′ oligo U tails, but with a lower affinity and specificity. In L. tarentolae mitochondrial extract, an oligo U–binding 60 kDa protein and a gRNA-binding 94 kDa protein were detected (Byrne et al., 1995).

In L. tarentolae, several classes of mitochondrial RNP complexes were identified that are possibly involved in editing (Peris et al., 1994). One class of complexes that were operationally termed T complexes and found to sediment in glycerol gradients at 10S–13S contains RNAs that are metabolically labeled in vitro with [α-^{32}P]UTP, giving rise to a pattern of approximately 6 bands in native gels. gRNAs are exclusively localized in the T-IV complexes, which also exhibit TUTase activity in an in-gel assay. T complexes were also detected by gel retardation experiments using labeled synthetic gRNAs (Byrne et al., 1995). The T complexes, or at least T-IV, may be involved in the maturation of gRNAs.

In Vitro Editing Systems

Deletions of U's

A breakthrough in the analysis of the mechanism of editing was achieved by the demonstration of gRNA-mediated in vitro U deletion activity by a dideoxy-terminated primer extension assay through the use of a tagged synthetic ATPase subunit 6 preedited mRNA, a synthetic cognate gRNA for the first editing site, and a mitochondrial extract from T. brucei (Seiwert and Stuart, 1994). Decreasing the number of guiding nucleotides in the gRNA produced a corresponding increase in the number of U's deleted from the mRNA. The results show clearly that base pairing interactions with the gRNA determine the number of U's deleted in vitro. Interestingly, no additional upstream editing events were detected, including predicted U insertions, suggesting either that essential components are missing in the in vitro system that restrict the editing to one site or that U deletion may have a different mechanism than U insertion.

Additions of U's

There is currently no in vitro system in which a gRNA-mediated insertion of U residues has been demonstrated. However, a mitochondrial extract from L. tarentolae was found to direct the incorporation of U's into the preedited regions of synthetic *CYb* mRNA and *ND7* mRNA substrates (Frech et al., 1995). This internal U incorporation does not, however, represent precise complete editing. The activity sediments in glycerol gradients as an RNP complex of 20S–25S, clearly separated from the TUTase-containing T complexes (Peris et al., 1994). The relationship of this internal U incorporation activity to in vivo editing remains to be investigated.

Physarum

The mechanism for this mixed insertion–substitution type of RNA editing is still a complete mystery. The editing process is highly efficient and accurate. Analysis of amplified partially edited RNAs showed no evidence for a polarity of editing as occurs in trypanosomes (Miller et al., 1993). The presence of multiple types of editing in Physarum mitochondria may be indicative of multiple mechanisms.

Why?

This is the most difficult but also the most interesting question. The rhetorical answer—why RNA splicing?—is clearly not a true response to this important question, but does emphasize the point that biological phenomena are in part frozen remnants of complex evolutionary histories and cannot be completely explained teleologically merely in terms of energy efficiency and the most parsimonious

mechanisms. The demonstration that pan-editing is a primitive character within the trypanosomatids and even within the kinetoplastids and that the pan-edited cryptogenes are apparently substituted by 5′-edited and even unedited genes during the evolution of these species perhaps by a retroposition mechanism involving partially edited RNAs (Simpson and Maslov, 1994) strongly suggests that editing requires a selective pressure to be maintained. This selective pressure is evidenced by the regulation of editing in the biphasic life cycle of the African trypanosomes (Stuart, 1993), but the possibility of regulation has not yet been examined in Leishmania, and there is as yet no explanation for the retention of editing in monogenetic species such as Crithidia.

The presence of pan-editing in T. brucei and also in the cryptobiid T. borreli suggests that complex editing involving multiple gRNAs was present in an ancestor of the entire kinetoplastid lineage (Lukes et al., 1994; Maslov and Simpson, 1994). It is tempting to speculate that gRNA genes were initially organized in tandem repeats, which later became circularized but remained attached by catenation to form the kDNA network characteristic of the trypanosomatids.

The relationship of the complex mixed insertional–substitutional editing in Physarum to the trypanosomatid editing is unclear.

The evidence for the evolution of kinetoplastid RNA editing leads to several alternative scenarios. One is that such editing is primitive, but not truly ancient, and first evolved in the mitochondrion of the kinetoplastid protozoa (Covello and Gray, 1993). Another scenario is that editing is truly ancient, was present in the eubacterial ancestor of the protomitochondrion, and perhaps even existed in the RNA world as a mechanism for creation of enzymatic RNA sequences. The apparent absence of the U insertion/deletion type of editing in modern eubacteria and in higher organisms together with the many common features of mitochondria clearly makes the former hypothesis more plausible, but the possibility of a polyphyletic origin of mitochondria in eukaryotes should not be ignored.

The question why will not be answered until we obtain a more complete knowledge of the variety and distribution of this or related types of RNA editing in other eukaryotic cells or in prokaryotic cells, as well as an understanding of the physiological role of editing. One thing that is certain is that investigation of this and related RNA editing phenomena will continue to illuminate the wonderful diversity of genetic mechanisms in living systems.

References

Benne, R. (1994). Eur. J. Biochem. *221*, 9–23.

Byrne, E., Bringaud, F., and Simpson, L. (1995). Mol. Biochem. Parasitol., in press.

Corell, R. A., Feagin, J. E., Riley, G. R., Strickland, T., Guderian, J. A., Myler, P. J., and Stuart, K. (1993). Nucl. Acids Res. *21*, 4313–4320.

Covello, P. S., and Gray, M. W. (1993). Trends Genet. *9*, 265–268.

Ferguson, M. L., Torri, A. F., Pérez-Morga, D., Ward, D. C., and Englund, P. T. (1994). J. Cell Biol. *126*, 631–639.

Fernandes, A. P., Nelson, K., and Beverley, S. M. (1993). Proc. Natl. Acad. Sci. USA *90*, 11608–11612.

Frech, G. C., Bakalara, N., Simpson, L., and Simpson, A. M. (1995). EMBO J. *14*, 178–187.

Goringer, H. U., Koslowsky, D. J., Morales, T. H., and Stuart, K. (1994). Proc. Natl. Acad. Sci. USA *91*, 1776–1780.

Gott, J. M., Visomirski, L. M., and Hunter, J. L. (1993). J. Biol. Chem. *268*, 25483–25486.

Köller, J., Nörskau, G., Paul, A. S., Stuart, K., and Goringer, H. U. (1994). Nucl. Acids Res. *22*, 1988–1995.

Landweber, L. F., and Gilbert, W. (1994). Proc. Natl. Acad. Sci. USA *91*, 918–921.

Leegwater, P., Speijer, D., and Benne, R. (1995). Eur. J. Biochem. *227*, 780–786.

Lukes, J., Arts, G. J., van den Burg, J., de Haan, A., Opperdoes, F., Sloof, P., and Benne, R. (1994). EMBO J. *13*, 5086–5098.

Mahendran, R., Spottswood, M. S., Ghate, A., Ling, M.-l., Jeng, K., and Miller, D. L. (1994). EMBO J. *13*, 232–240.

Maslov, D. A., and Simpson, L. (1994). Mol. Cell. Biol. *14*, 8174–8182.

Miller, D., Mahendran, R., Spottswood, M., Costandy, H., Wang, S., Ling, M. L., and Yang, N. (1993). Semin. Cell Biol. *4*, 261–266.

Missel, A., and Goringer, H. U. (1994). Nucl. Acids Res. *22*, 4050–4056.

Pérez-Moraga, D. L., and Englund, P. T. (1993). Cell *74*, 703–711.

Peris, M., Frech, G. C., Simpson, A. M., Bringaud, F., Byrne, E., Bakker, A., and Simpson, L. (1994). EMBO J. *13*, 1664–1672.

Piller, K. J., Decker, C. J., Rusche, L. N., and Sollner-Webb, B. (1995). Mol. Cell. Biol., in press.

Pollard, V. W., Harris, M. E., and Hajduk, S. L. (1992). EMBO J. *11*, 4429–4438.

Rusche, L., Piller, K. J., and Sollner-Webb, B. (1995). Mol. Cell. Biol., in press.

Sabatini, R., and Hajduk, S. L. (1995). J. Biol. Chem. *270*, 1–8.

Seiwert, S. D., and Stuart, K. (1994). Science *266*, 114–117.

Simpson, L., and Maslov, D. A. (1994). Science *264*, 1870–1871.

Simpson, L., Maslov, D. A., and Blum, B. (1993). In RNA Editing: The Alteration of Protein Coding Sequences of RNA, R. Benne, ed. (New York: Ellis Horwood), pp. 53–85.

Stuart, K. (1993). In RNA Editing: The Alteration of Protein Coding Sequences of RNA, R. Benne, ed. (New York: Ellis Horwood), pp. 25–52.

Thiemann, O. H., Maslov, D. A., and Simpson, L. (1994). EMBO J. *13*, 5689–5700.