Reprint Series
26 October 1990, Volume 250, pp. 512–513

SCIENCE

RNA Editing—A Novel Genetic Phenomenon?

LARRY SIMPSON

RNA Editing—A Novel Genetic Phenomenon?

LARRY SIMPSON

THE THEORY THAT GENETIC INFORMATION IS ENCODED IN nucleic acid (DNA or RNA) and flows in a unidirectional manner to determine the primary amino acid sequences of proteins has suffered few, if any, conceptual challenges since the original formulation (*1*). Reverse transcriptase, which transcribes RNA into DNA, added an interesting detour to the pathway. The discovery of RNA enzymes (*2, 3*) was a major conceptual change, but had no effect on the information flow paradigm.

The discovery of RNA editing in the mitochondria of kinetoplastid protozoa (reviewed in *4*) at first appeared to directly challenge the theory of information flow. In this process, uridine (U) residues are added or deleted at multiple, precise sites within the coding regions of mRNA's. There appeared to be no nucleic acid templates that encoded the edited information, and it was difficult to envision how such extensive and precise modifications of mRNA sequences could occur without a template. These findings raised the possibility that the information for editing could be encoded in a cryptic form in the mRNA or in proteins (*4, 5*). Soon thereafter, other types of apparently nontemplated editing of coding sequences of mRNA molecules were reported: (i) a single, precise, developmentally regulated cytosine (C)-to-U change in mammalian apolipoprotein B mRNA (*6*); (ii) multiple C-to-U or U-to-C changes (*7*) in several plant mitochondria mRNA's; (iii) nontemplated G residues in paramyxovirus P mRNA (*8*); and (iv) 54 extra C residues in mRNA for the alpha subunit of adenosine triphosphate synthetase in *Physarum* mitochondria (*9*).

The most striking examples of RNA editing are found in kinetoplastid mitochondria, which contain a large network of catenated minicircles (465 to 2500 base pairs) and maxicircles (23 to 36 kilobases) (the kinetoplastid DNA nucleoid body) (*10*). The function of minicircle DNA has been a mystery since its discovery (*11, 12*), as there was no obvious conserved coding capacity. The maxicircle molecules are homologs of informational mitochondrial DNA molecules in animal and fungal cells, and encode at least 13 genes (*10*).

Several of the maxicircle genes represent cryptogenes (*4*) in which the DNA sequences specify primary transcripts that have reading frame shifts, lack canonical translation initiation codons, and must be edited to produce translatable mRNA's. There are at least three pan-edited (*4*) cryptogenes in *Trypanosoma brucei* maxicircle DNA, which lack more than 50% of the U residues present in the mature edited transcripts (*13, 14*). Six short G-rich regions exist in both *Leishmania tarentolae* and *T. brucei*, which may also encode pan-edited RNA's, but the gene products are not yet known (*4*). RNA editing of at least four mitochondrial cryptogenes in *T. brucei* is developmentally regulated (*15, 16*) and is utilized as a translational control mechanism by regulating the abundance of translatable mRNA's.

A new class of small kinetoplastid mitochondrial RNA molecules, guide RNA's (gRNA's), was recently described (*17*). The gRNA's are short RNA molecules that can form perfect hybrids with edited mRNA sequences and possess nucleotide sequences at their 5′ ends that are complementary to the sequences of the mRNA's immediately downstream of the pre-edited regions (PER). The gRNA's do not represent classical templates for edited RNA sequences, due to the presence of abundant noncanonical G-U base pairs. The gRNA's, some of which appear to represent primary transcripts, possess unique 5′ ends and a 3′ oligoU tail added posttranscriptionally, which varies in length from 5 to 24 nucleotides (*18*). The gRNA genes are found in the maxicircle genome and within the variable regions of the minicircles (*19*), suggesting a function for these DNA molecules. The gRNA's are specific for each edited mRNA and encode the additional U residues as complementary A or G residues. In the proposed model (*17*) (Fig. 1), a hybrid is formed between the 5′ end of the gRNA and a region of the mRNA that is adjacent (3′) to the PER (3′ anchor). A stabilizing hybrid then forms between the 3′ oligoU tail of the gRNA and the GA-rich PER (5′ anchor) (*18*). Editing occurs by specific endonuclease cleavage of the mRNA within the PER at a position 3′ to the first mismatched nucleotide. Addition of U residues to, or, more rarely, deletion from the liberated 3′ hydroxyl terminus is followed by formation of a base pair between the guide A or G and the added U residues, and religation of the cleaved mRNA molecule. The putative editing enzyme complex then migrates to the next mismatch and the cycle is repeated.

The evidence for this model is circumstantial but convincing: (i) gRNA's for five cryptogenes in *L. tarentolae* exist that can form perfect hybrids with edited mRNA's, with the unique 5′ ends of the gRNA's localized close to the beginning of the hybrid regions; (ii) multiple partially edited molecules have been detected in steady-state RNA, which show the expected 3′ to 5′ polarity of partial editing (*20, 21*); (iii) enzymes for several of the predicted activities exist in purified mitochondria of *L. tarentolae* (*22*)—a terminal uridylyl transferase (TUTase), an RNA ligase, and a cryptic, site-specific endonuclease (*23*); and (iv) imprecise editing of synthetic pre-edited

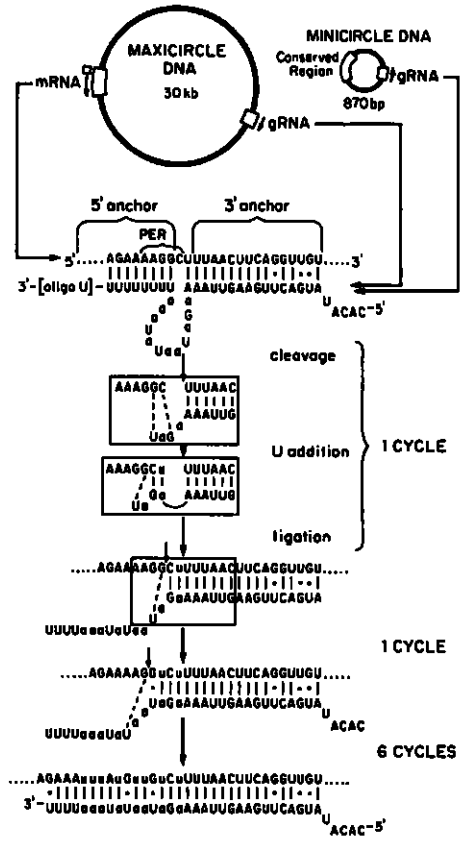

Fig. 1. The gRNA model for RNA editing. One cycle is shown for the 5′ editing of the CYb mRNA with gRNA-1. The guide nucleotides in the gRNA are shown as a, and the added U residues in the mRNA as u. The cleavage sites are indicated by arrows. G-U base pairs are indicated by *.

Department of Biology and Molecular Biology Institute, University of California, Los Angeles, CA 90024.

mRNA's has been detected with crude mitochondrial lysates from *L. tarentolae* (*23*).

A variant of the gRNA editing model has been proposed to account for the presence, in *T. brucei* mitochondria, of a high frequency of partially edited cytochrome oxidase subunit III (COIII) and cytochrome b (Cyb) RNA's, which exhibit unexpected (that is, not precisely 3′ to 5′) patterns of editing (*24*). In this model, random insertion and deletion of U residues occur between every nucleotide within editing domains, which are defined by specific gRNA's. Editing occurs as a result of selection of the correctly edited mRNA sequence by the formation of a perfect hybrid with the gRNA rather than by a directed unidirectional mismatch repair. A fairly high percentage (42%) of unexpected, partially edited mRNA's was also observed in *L. tarentolae* for the COIII gene (*21*). However, the majority of these patterns may arise from correct editing with incorrect gRNA's that may guide the editing of yet unidentified cryptogenes (*19*). It may be that functional modulation of amino acid sequences of mitochondrial proteins occurs through the use of different gRNA's for editing.

More compelling support for the gRNA model remains to be obtained. For example, isolation of an editing complex that contains gRNA, riboendonuclease, TUTase, exonuclease, and RNA ligase is crucial. In addition, the following questions are of interest: (i) Are gRNA's complexed with proteins, thus representing the functional equivalent of small nuclear ribonucleoproteins in the splicing paradigm? (ii) What is the mechanism for the sequential interaction of the multiple gRNA's, proposed to be required for editing of the CYb or MURF2 transcripts in *L. tarentolae* or for editing of the pan-edited mRNA's (COIII, ND7, and ATP6) in *T. bruccei*? (iii) How are the secondary structures of mRNA and gRNA involved in the specific binding of an editing complex? (iv) Are ribozymes involved in the editing process?

The recent discovery that kinetoplast minicircle DNA encodes gRNA's that function in the editing of maxicircle transcripts remains to be investigated, especially in view of the extensive minicircle sequence divergence observed between species of kinetoplastids. Putative gRNA genes have been identified in *T. brucei* (*14, 25*), and gRNA-like transcripts have been identified in *T. equiperdum* that are precisely located between pairs of 18-bp inverted repeats (*26*) in the variable regions of the minicircles. These inverted repeats could represent remnants of transposition events that involved the migration of mobile gRNA genes between maxicircle and minicircle DNA.

One function of RNA editing in kinetoplastids is to provide translational regulation of mitochondrial gene expression. However, the evolutionary origin of this type of editing and the effect of such split genes on the evolution of the mitochondrial genome remain to be explored. It is possible that this process originally represented a general mechanism for the modification or repair of RNA sequences prior to the origin of polymerase enzymes. Perhaps modern mitochondrial RNA editing in trypanosomes in an atavistic remnant of a primitive RNA sequence modification process. It remains to be seen if the trypanosome type of RNA editing is present in higher eukaryotes.

REFERENCES

1. F. H. C. Crick, *Symp. Soc. Exp. Biol.* **12**, 138 (1958).
2. T. Cech, *Science* **236**, 1532 (1987).
3. S. Altman *et al.*, *Gene* **82**, 63 (1989).
4. L. Simpson and J. Shaw, *Cell* **57**, 355 (1989).
5. R. Benne, *Biochim. Biophys. Acta* **1007**, 131 (1989).
6. S.-H. Chen, X. Li, W. S. L. Liao, J. H. Wu, L. Chan, *J. Biol. Chem.* **265**, 6811 (1990).
7. P. S. Covello and M. W. Gray, *Nature* **341**, 662 (1989).
8. S. M. Thomas, R. A. Lamb, R. G. Paterson, *Cell* **54**, 891 (1988).
9. R. Mahendran, M. R. Spottswood, D. L. Miller, *ibid.*, in press.
10. L. Simpson, *Annu. Rev. Microbiol.* **41**, 363 (1987).
11. G. Riou and E. Delain, *Proc. Natl. Acad. Sci. U.S.A.* **62**, 210 (1969).
12. L. Simpson and A. M. da Silva, *J. Mol. Biol.* **56**, 443 (1971).
13. J. E. Feagin, J. Abraham, K. Stuart, *Cell* **53**, 413 (1988).
14. G. J. Bhat, D. J. Koslowsky, J. E. Feagin, B. L. Smiley, K. Stuart, *ibid.* **61**, 885 (1990).
15. J. E. Feagin, D. Jasmer, K. Stuart, *ibid.* **49**, 337 (1987).
16. J. Feagin and K. Stuart, *Mol. Cell. Biol.* **8**, 1259 (1988).
17. B. Blum, N. Bakalara, L. Simpson, *Cell* **60**, 189 (1990).
18. B. Blum and L. Simpson, *ibid.* **62**, 391 (1990).
19. N. R. Sturm and L. Simpson, *ibid.* **61**, 879 (1990).
20. J. Abraham, J. Feagin, K. Stuart, *ibid.* **55**, 267 (1988).
21. N. R. Sturm and L. Simpson, *ibid.* **61**, 871 (1990).
22. N. Bakalara, A. M. Simpson, L. Simpson, *J. Biol. Chem.* **264**, 18679 (1989).
23. A. M. Simpson, N. Bakalara, E. Gruszynski, L. Simpson, unpublished data.
24. C. J. Decker and B. Sollner-Webb, *Cell* **61**, 1001 (1990).
25. V. W. Pollard, S. P. Rohrer, E. F. Michelotti, K. Hancock, S. Hadjuk, *ibid.*, in press.
26. D. Jasmer and K. Stuart, *Mol. Biochem. Parasitol.* **18**, 321 (1986).