RNA editing in kinetoplastid parasites: what to do with U

Paul Sloof and Rob Benne

The order Kinetoplastida are unicellular, flagellated protozoa, most of which are parasites (for kinetoplastid physiology and taxonomy, see Refs 1, 2). They infect a wide range of plants, invertebrates and vertebrates and cause major human diseases. Millions of people, predominantly in the developing countries of Asia, Africa and South America, suffer from trypanosomiasis and leishmaniasis. Similar parasites also infect human livestock and contribute to human malnutrition. All Kinetoplastida possess a characteristic disc-shaped body in their single mitochondrion called the kinoplast, which contains mitochondrial (mt) DNA (Ref. 3). Mt gene expression has been studied extensively in three representatives of the Trypanosomatidae family: the digenetic _Leishmania tarentolae_ and _Trypanosoma brucei brucei_, which cycle between insects (sandflies and tsetse flies, respectively) and vertebrates (lizards and mammals), and the monogenetic insect parasite _Crithidia fasciculata_. Recently, investigations have been extended to a member of the Cryptobiidae family, _Trypanosoma borreli_, which is a fish parasite and only a distant relative of the trypanosomatids.

In 1986, the study of the genetic content and the expression of trypanosomatid kinetoplast DNA (kDNA) revealed the existence of an extraordinary form of transcript maturation called 'RNA editing', which involves insertions and, less frequently, deletions of uridylate (U) residues. In the ten years that have followed, many examples of kinetoplastid mt transcript editing have been described, ranging from the insertion/deletion of a limited number of Us in one or two small regions of the transcript to massive or 'pan' editing (for reviews, see Refs 9–12). In addition, other forms of RNA editing involving insertion/conversion of other nucleotides have been discovered, suggesting that processes that post- and, even co-, transcriptionally alter RNA sequences are widespread. In most cases, RNA editing plays an essential role in gene expression (see Box 1 for a glossary of terms used).

The puzzle of the origin of the genetic information for kRNA editing has been resolved by the discovery of guide RNAs (gRNAs) (see Box 2 and Fig. 1), and several models explaining how information is transferred from gRNA to pre-edited RNA have been proposed. It is only recently, however, that _in vitro_ assay systems for kRNA editing have been developed and models have been put to the test. An enzymatic cleavage/ligation ('cut-and-paste') model appears to be the most likely. In spite of this progress, it is still unclear why this extra step in the expression of kinetoplast mt genes has evolved.

From scattered genetic information to reading frames

The structure of trypanosomatid kDNAs is unique. It consists of a single network of thousands of catenated circular DNAs of two size classes: minicircles of 0.9–2.5 kb (depending on the species), which are heterogeneous in sequence and account for the major portion of the network, and 20–50 homoplasmic maxicircles with an interspecies size variation of 23–40 kb (Ref. 3). Mt gene expression appears to be a collaborative effort by both maxicircles and minicircles (see Box 3 and Fig. 2), in which the target pre-edited RNAs and a few gRNAs are maxicircle-encoded and the majority of the gRNA genes are found in minicircles. The coding region of the maxicircles in the three trypanosomatid species is similar in size and gene organization (Fig. 3); differences in maxicircle size result from the presence of a varying number of species-specific tandem repeats in the 'variable' region. The overall structure and gRNA gene content of the minicircles of the species studied

---

**Box 1. Glossary**

- **9S/12S**: Mitoribosomal RNAs of 9S and 12S.
- **ATPase**: ATP synthase.
- **CoX**: Cytochrome c oxidase.
- **cyt b**: Apocytochrome b.
- **G**: Cryptic gene producing G-rich pre-mRNA.
- **gRNA**: Guide RNA.
- **kDNA**: Kinetoplast DNA.
- **mt**: Mitochondrial.
- **MURF**: Maxicircle unassigned reading frame.
- **ND**: NADH dehydrogenase.
- **RPS12**: Ribosomal protein S12 of the small subunit.
- **TUTase**: Terminal uridylyl transferase.
Box 2. The guide RNA paradigm

RNA editing requires a high level of accuracy because insertion/deletion of the wrong number of uridylates (Us) could produce an untranslatable, frame-shifted mRNA. In 1990, a major breakthrough was achieved with the discovery of short (40–70 nt) RNAs, all containing a sequence that is complementary to an edited mRNA segment.\(^{14}\) In addition to Watson and Crick base pairing, the complementarity frequently involves G–U pairs, which precludes a conventional template function for these RNAs. Therefore, they were called guide RNAs (gRNAs). Comparison of different gRNAs revealed the presence of a few common structural elements (Fig. 1), which provided the clues to gRNA function on which early models were based.\(^{9–14}\) In the 5' region they all contain a so-called anchor sequence that can basepair with a pre-mRNA immediately 3' of an editing region. It was envisaged that the annealing of the anchor to its complement represents the first step of the editing process. Recently, the anchor-sequence-dependent involvement of gRNAs in editing has been directly demonstrated in in vitro systems using mitochondrial (mt) extracts of Trypanosoma brucei and synthetic pre-mRNA and gRNA variants.\(^{15–19}\) Immediately 3' of the anchor, a sequence is located that is complementary to an edited region of a mRNA. This region specifies the pattern of deletions and insertions; in other words, it contains the 'information' for editing.\(^{15–18}\) The third domain is an oligo(U) extension of 5–24 residues at the 3' terminus of the gRNA, which is probably added post-transcriptionally by mt terminal uridylyl transferase (TUTase).\(^{9–12,18–20}\) The U tail may provide extra stability to the gRNA-pre-edited RNA duplex during editing.

Site selection with the aid of gRNAs is not unique to trypanosome editing because in other RNA-based transactions the same principle appears to be used. In some vertebrate and invertebrate mRNAs, adenosine residues that are converted to inosines by deamination are selected via intramolecular basepairing with other sections of the pre-mRNA (reviewed in Ref. 20). In addition, mispairing in tRNA stems is corrected by editing events that result in the incorporation of the correct base (nucleotide).\(^{21,22}\) Last, but not least, the sites of 2'-O-ribose methylation in ribosomal RNAs are determined by basepairing with small nucleolar gRNAs (Ref. 22).

---

![Fig. 1](image-url). The guide RNA (gRNA) paradigm. Schematic representation of a guide RNA in which the anchor region is annealed to a section of pre-edited RNA immediately downstream of a region containing four editing sites (marked by arrowheads). Editing is guided by the informational part of the gRNA. Edited uridylate residues in mRNA and guiding nucleotides in gRNA are lower case. Editing results in an extension of the mRNA–gRNA duplex, in which Watson and Crick basepairs are indicated by bars and G–U pairs by asterisks. gRNAs contain an oligo(U) tail at the 3’ terminus.
The search for extensively edited RNAs in L. tarentolae was hampered by the fact that the strains that were studied initially had lost the ability to produce all but one of the panedited RNAs during the ~50 years that they had been cultivated in the lab. Only G6-derived edited RNA was found, encoding ribosomal protein S12 (RPS12). In a strain recently isolated from the wild, however, the six panedited RNAs mentioned above were present. During cultivation, many minicircle classes were lost, resulting in a much smaller complement of minicircle-encoded gRNAs and loss of editing ability. This not only means that editing is a labile genetic trait that is easily lost but also that the products of five of the six panedited RNAs are not essential for L. tarentolae under the conditions of cultivation. Owing to a marginal degree of sequence similarity, the identification of the proteins in question is not yet conclusive, but it has been suggested that some of them may be subunits of a complex I-type NADH dehydrogenase (ND3, ND8, ND9) in L. tarentolae. In line with the hypothesis that ND subunits are not required in cultivated trypanosomatids, we have recently shown that a typical complex I-type NADH dehydrogenase is absent in cultured C. fasciculata. Interestingly, other RNAs encoding ND subunits are also untranslatable in cultured trypanosomatids. ND7 RNA lacks an in-frame translational initiation codon in both L. tarentolae and C. fasciculata, in which it is edited in two small regions by maxicircle-encoded gRNAs, whereas (unedited) ND1 RNA suffers from the same defect in C. fasciculata and T. brucei.

Apparently, in trypanosomatid mitochondria, gene expression is rapidly shut down in the absence of selective pressure not only via the loss of minicircles but also by mutations that accumulate in maxicircles.

**The mechanism of RNA editing**

The fact that trypanosomatid mtRNAs are edited at internal sites implies that at some stage the RNA ribose-phosphate backbone has to be cleaved, followed by the addition or removal of Us and resealing of the RNA strand. The first model for editing that was developed after the discovery of gRNAs assumed that these steps were carried out by enzymes, whose involvement in editing was inferred from suggestive properties and their (sometimes assumed) presence in mitochondria. This 'enzyme cascade' model, outlined in Fig. 4, proposed the following consecutive steps: (1) editing site selection by the formation of the anchor duplex (see Box 1 and Fig. 1), (2) pre-edited RNA cleavage at the first mismatch immediately 5' of the gRNA-pre-edited RNA duplex by endonuclease(s) and (3) U addition to the newly generated 3' end of the 5' cleavage product by 3' terminal uridylyl transferase (TUTase) and UTP (insertion), or removal of U(MP) by a 3' exonuclease (deletion) and rejoicing of the two RNA moieties by RNA ligase.

A second model, proposed shortly afterwards, was based on the observations that gRNAs possess a 3' oligo(U) tail and that chimeric molecules, in which a RNA is covalently linked via this oligo(U) tail to an editing site of pre-edited RNA (Ref. 31), exist in mtRNA. This and similar models envisaged that chimeric molecules could be intermediates of an editing process in which it is edited in two small regions by maxicircle-encoded gRNAs, whereas (unedited) ND1 RNA suffers from the same defect in C. fasciculata and T. brucei.
which the gRNA oligo(U) tail serves as a U donor (insertion) or acceptor (deletion) in two successive transesterification reactions (for discussion, see Refs 9-12, 31, 32). The attraction of this model was the possible analogy and common evolutionary origin of editing and splicing reactions. Recent experiments, however, have proved it to be wrong. First, the possible link between editing and the 'RNA world' has been weakened by the observation that chimeric molecules are not made by transesterification in vitro but by the combined action of endonuclease and RNA ligase in a pathway reminiscent of the enzyme cascade model (for a review, see Ref. 9). The next and more serious blow to this model came recently from in vitro assay systems for U deletion and U insertion. In these studies, a detailed kinetic analysis of products and possible intermediates of editing reactions was made, which fully supported the enzyme cascade model. Three observations are crucial. First, 3' pre-edited RNA cleavage products are not predicted by the transesterification pathway, yet they are generated at an early stage of editing reactions and reach a steady-state level characteristic of reaction intermediates. Second, chimeric RNAs appear subsequent to edited products and continue to accumulate. It seems likely, therefore, that the bulk of the chimeric RNAs produced in these systems represents
aberrant end products and not intermediates of the editing process.\(^{12-19}\) Finally, inserted Us appear to be derived from free UTP, as U insertion is gRNA- and UTP-dependent.\(^{12}\) Efficient in vitro systems in which the editing intermediates and products can be directly visualized have not yet been developed for \(L. tartarotne\), but recent reports describe both gRNA-dependent and gRNA-independent U-insertion activities, the characteristics of which support the model outlined in Fig. 4 (Refs 33–35).

Other features of gRNA functioning in this enzyme cascade model were experimentally verified in the in vitro systems. First, gRNAs were rendered ineffective by deletion of the anchor sequence, illustrating its importance in selecting the editing sites. The primary function of the gRNA-pre-edited RNA interaction could be to direct the endonucleolytic cleavage of the pre-edited RNA phosphodiester bond 5' of the duplex. Subsequently, the number of Us inserted or deleted in a given site is specified by the sequence of the informational domain of the gRNAs. It can be changed in a predictable fashion by alterations in the gRNA and/or pre-edited RNA sequence. However, the gRNA sequence does not specify the number of Us to be added to or deleted from the 5' cleavage product, as 3' RNAs with a varying number of Us are found. The exact number of Us is probably determined at the ligation step when a 5' RNA with the correct number of Us is selected by basepairing with the gRNA, resulting in the correct juxtaposition with the 5' end of the 3' cleavage product for ligation to occur. During these transactions, the pre-edited RNA cleavage products may be kept together by basepairing of the gRNA U tail to a purine-rich sequence that is present upstream of most editing sites, as a pre-edited RNA containing a small editing region from ND7 RNA from \(C. fasciculata\), which lacks such a purine-rich region, could not be edited in vitro. This emphasizes that not all of the details of the editing process have been fully elucidated, as \(C. fasciculata\) ND7 RNA is efficiently edited in vivo.\(^{25}\) Chimeric RNAs may be generated as side products at the ligation step when, instead of the 3' hydroxyl of the 5' cleavage product, the 3' hydroxyl of the gRNA U tail is ligated to the 3' cleavage product.

Fig. 4 shows the editing of a hypothetical single-site RNA but, in reality, editing regions contain many more sites, requiring multiple cycles of the same sequence of events: each cycle is initiated by cleavage of the partially edited RNA at the next editing site, directly 5' of the extending gRNA-pre-mRNA duplex. However, it should be realized that the in vitro systems are not very efficient because only one editing cycle is completed in a few percent of the added substrate RNAs (in \(L. tartarotne\) the efficiency is lower). This means that the exact order in which the sites of a region are edited cannot be determined yet. Nevertheless, the location of the in vitro cleavage sites predicts a strict 3' to 5' order of editing site selection during multiple cycles.
REVIEWS

Questions for future research

- What does the editosome look like and what is the mechanism of editing? The contours of the editing mechanism and the machinery that is operative are beginning to emerge now that in vitro systems for editing have been developed. It is clear, nevertheless, that many aspects of editing in vivo have yet to be elucidated.

- What is the evolutionary origin of editing? The discovery of editing in T. borrei and other bodonids suggests that editing has been present in the kinetoplast lineage for more than 500 million years. However, as it has not (yet) been found outside the kinetoplastid lineage, it is not known how 'old' RNA editing really is. The search for editing in other species should continue therefore.

- What (if any) is the function of RNA editing? It has been speculated that RNA editing is an RNA-repair mechanism to compensate for DNA mutations. Alternatively, it may increase genetic flexibility by allowing the production of multiple proteins from one gene or by providing an extra level of regulation of gene expression.

- How is RNA editing regulated? Very little is known of the signals that regulate the transcription/stability of maxi- and minicircle RNAs. The available evidence suggests that gRNAs are constitutively produced, even in life cycle stages in which its target RNA is not edited. This could mean that the editing of a particular transcript is not regulated by the concentration of the corresponding gRNAs, but the monumental task of measuring the concentration of all of the gRNAs involved has not yet been completed.

Proteins involved in editing

Although formal proof is still lacking, an impressive amount of evidence supports the inferred role in editing of site-specific endonuclease(s), TUTase and RNA ligase. First, following glycerol gradient analysis of mt extracts from T. brucei, the relevant enzymatic activities were found to (co-)sediment in a broad peak at 20-40S in fractions that also contained gRNA (Refs 37-41) and that accurately catalysed in vitro editing reactions, although the different activity peaks did not completely coincide. Moreover, the requirement of the in vitro insertion reaction for UTP (Ref. 18) suggests the involvement of TUTase, whereas the in vitro deletion studies indicate that the removal of UMP is not the result of a reverse reaction of TUTase but rather of 3' U exonuclease activity. The involvement of RNA ligase in the editing reactions is possible because of the requirement of the deletion reaction for hydrolysis of the α-β-phosphate bond of ATP (Ref. 15) and the observation that the pre-edited RNA cleavage intermediates have 3' OH and 5' P-termini. Last, but not least, a seven-polypeptide complex has been purified that catalyses deletion editing and contains all of the activities mentioned so far in this section, but no gRNAs (Ref. 42). These results indicate that an 'editosomal' complex containing the proteins directly involved in the U insertion/deletion process with only a few others appears to be sufficient to achieve editing in vitro. As the efficiency of the in vitro systems is low, this may not be the whole story, and efficient editing in vivo may use more proteins. These could be involved in correct folding of the participating RNAs, the assembly of gRNA–pre-mRNA duplexes, duplex disassembly to make edited mRNAs available for translation (indeed a helicase was recently described in T. brucei), the ordered 'coming and going' of gRNAs and the overall regulation of editing. For this reason, several groups are studying the characteristics of proteins that have affinity for gRNA and pre-edited RNA (Refs 44-48) in UV cross-linking and/or gel retardation analysis. Some of these proteins have been purified and their cloning is underway, but a specific role in editing has not yet been assigned to any of them.

Conclusions and prospects

Despite the progress that has been made during ten years of RNA editing research, the question of why kinetoplastid RNA editing (or other forms of RNA editing) exists is hard to answer. It may have developed as an RNA repair mechanism to compensate for genomic mutations or it could serve to increase genetic flexibility; for example, as an extra control point for gene expression or as a means to produce multiple proteins from a single cryptogene. There is no evidence to support any of these possibilities. In fact, even formal proof that edited RNAs are used by the translational machinery is lacking because kinetoplastid mt proteins have not yet been sequenced. The study of the phylogeny of RNA editing in kinetoplastids suggests that RNA editing is an ancient trait and that panediting is the ancestral form, retrotransposition giving rise to 'edited' genes, whose transcripts require little or no editing, in more recently evolved species (see Ref. 50 for extensive discussion). It remains possible, therefore, that RNA editing no longer has a function but represents a molecular fossil that kinetoplastids, with the exception of species such as Trypanosoma equiperdum, which have a mutated maxi-circle and no longer edit mt RNAs (Ref. 51), cannot easily get rid of in vivo. However, T. equiperdum and other species that have undergone (extensive) mutation of the mt DNA do not have a functional mitochondrion and have lost the ability to propagate in the insect host.

An ongoing search for editing in other organisms, such as Euglena, archaeabacteria or α-proteobacteria (the closest bacterial relatives of mitochondria), and the unravelling of the molecular details of the mechanism may prove essential to obtain further clues to the evolutionary significance of this process. In all studies, the in vitro editing systems will continue to play a crucial role, the primary aim being to increase efficiency so that more than one cycle of editing and panediting with multiple gRNAs can be studied. Undoubtedly, this will involve the painstaking development of purification procedures for labile and/or low abundant proteins and protein–RNA complexes. Now that efficient transfection of trypanosomatids is routine, the final test for the model in Fig. 4, or for any other model, has to be performed in vivo with cloned genes of candidate editing factors. There is no reason to believe that the next ten years of RNA editing will be less exciting than the first ten.

Acknowledgements

We thank the members of the trypanosome group for stimulating discussions and D. Speijer for allowing us to include pre-publication results. Our research is supported by the Netherlands Foundation for Chemical Research (SON) and the Foundation for Medical and Health Research (MW), which are subidized by the Netherlands Foundation for Scientific Research (NWO). We apologize to all our colleagues whose work could not be properly cited because of space limitations.
TIM comes of age

This month we are celebrating 50 issues of Trends in Microbiology. To mark this occasion, we would like to hear what you think about the journal. Which sections do you find particularly relevant, interesting or informative? Are there any topics that you’re dying to read about? Do you have any suggestions for Web pages to be included in the Microweb section or literature to be included in Horizons? Or are there areas where you would like to see changes? Help us improve on the success of the first 50 issues of TIM by sending your comments to:

Caroline Ash (Editor)
Trends in Microbiology,
68 Hills Road,
Cambridge,
UK CB2 1LA.
tel: +44 1223 315961,
fax: +44 1223 464430,
e-mail: tim@elsevier.co.uk