ELSEVIER

Molecular and Biochemical Parasitology 68 (1994) 155–159

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

Short communication

Editing and misediting of transcripts of the kinetoplast maxicircle G5 (ND3) cryptogene in an old laboratory strain of *Leishmania tarentolae*

Dmitri A. Maslov [a], Otavio Thiemann [a], Larry Simpson [a,b,c,*]

[a] *Department of Biology, University of California, Los Angeles, CA 90024 USA*
[b] *Howard Hughes Medical Institute, Los Angeles, CA 90024 USA*
[c] *Department of Medical Microbiology and Immunology, University of California, Los Angeles, CA 90024 USA*

Received 20 July 1994; accepted 8 August 1994

Keywords: RNA editing; Mis-editing; *Leishmania tarentolae*; Kinetoplast; Kinetoplastid protozoa; G-rich region; ND3

RNA editing in the mitochondrion (kinetoplast) of trypanosomatid protozoa involves a post-transcriptional addition and deletion of uridilate residues [1–4]. The sites of editing are specified by small guide RNA transcripts (gRNAs), which are encoded in both kinetoplast DNA minicircles and maxicircles. The process initiates with a short region of base-pairing (anchor formation) between a specific gRNA and a pre-edited mRNA, and generally proceeds in a 5′ direction with the insertion and deletion of uridilate (U) residues. The precise sites and number of Us inserted and/or deleted is determined by base-pairing with guide adenine or guanine nucleotides in the gRNA. Two basic models have been proposed for the mechanism of this process, one involving two successive transesterifications [5,6] and the other involving an enzymatic cleavage-ligation [7,8], with the source of the U residues being either UTP or the 3′ non-encoded oligo(U)-tail of the gRNA. Several ribonucleoprotein complexes containing gRNAs and thought to be involved in editing have been detected and partially characterized [9–12].

The 3′ to 5′ polarity of editing within an editing domain results from the formation of anchors for upstream gRNAs by downstream editing [13–15]. In partially edited RNAs, the junction regions between fully edited and unedited sequences are frequently incorrectly edited [16–18]. These may represent normal intermediates of editing [17,19], or may be 'misedited' sequences produced during an abortive editing process [18,20] which are subsequently re-edited correctly. Some evidence for misediting by non-cognate gRNAs forming spurious anchor hybrids and for correct re-editing of misedited sequences has been previously presented [20]. An alternate hypothesis ('match protection') is that editing is essentially random within a block and that the correct insertions and deletions are frozen in place by

Abbreviations: gRNA, guide RNA; ND3, NADH dehydrogenase subunit 3

* Corresponding author. Tel.: 310-825-4215; Fax: 310-206-8967; E-mail: simpson@biovx1.dnet.nasa.gov.

SSDI 0166-6851(94)00160-X

base-pairing with the cognate gRNA [17]. Another hypothesis ('progressive realignment') is that editing proceeds in regions of lowest thermodynamic stability and not necessarily 3′ to 5′ within an editing block [19].

In *Leishmania tarentolae*, each minicircle sequence class encodes a single gRNA situated in the variable region adjacent to the 'bend' and the conserved region [21], and the number of different gRNAs is determined essentially by the number of different minicircle sequence classes [13]. We reported recently that, in an old laboratory strain of *L. tarentolae* (UC strain) [22], many kinetoplast DNA minicircles sequence classes encoding gRNAs for the editing of several pan-edited cryptogenes (G1–G5) are absent. On the other hand, in a recently isolated strain of *L. tarentolae* (LEM125) [23] additional minicircle sequence classes are present which encode gRNAs responsible for the editing of these cryptogenes, and the pre-edited transcripts of these genes are productively edited in this strain [24].

When we attempted to clone partially edited mRNAs from the G5 cryptogene (may represent ND3) [25] from the UC strain by the standard RT-PCR method, a variety of edited sequences were obtained. In this report we show that some of these represent misedited sequences derived from correct editing with gRNAs for other genes, as a result of the absence of the cognate gRNAs in this strain.

Partially edited G5 transcripts were PCR-amplified from UC strain kRNA using a 3′ oligo-dT primer or a 3′ downstream genomic primer (S-290: TTAGAATTCAATAAGAGAATACAC), and 5′ upstream genomic primers (S-291: TTTGGATCCAAGGTTAAAAATAGG and S-292: AGGGGATCC TTTGGGGAGGAGGGA) (Fig. 1). The products were cloned and sequenced and some of the sequences are shown in Fig. 2. Several different classes of partially edited sequences could be detected based on sequence alignments. Seven poly(A)-containing clones shown in Fig. 2A and an additional 22 clones that were obtained with the downstream genomic primer, S-290, exhibited an editing pattern at their 3′ ends which is consistent with correct editing by the initial cognate gRNA (gND3-I), which mediates the editing of block I of this transcript in the LEM125 strain (data not shown). Thiemann et al. [24] showed that the gND3-I RNA is encoded in the maxicircle

Fig. 1. Localization of the RT-PCR oligonucleotide primers. A partial map of the maxicircle that includes the NADH dehydrogenase subunit 4 (ND4), G-rich region 5 (G5 or ND3) and ribosomal protein S12 (RPS12) genes is shown. The position of the genes above or below the line refers to their polarity. Black boxes represent pan-edited cryptogenes, and the open box, a non-edited gene.

DNA and is present in both the LEM125 and UC strains. This explains the presence of a subset of partially edited mRNAs in the UC strain which are correctly edited in block I.

These clones also contained incorrectly edited junction regions. As shown in Fig. 2A, the sequence of one of these clones (clone 1.7) is complementary to the sequence of another maxicircle-encoded gRNA (gG4-IV) for the G4 cryptogene, which is not edited productively in the UC strain. In another clone (not shown), the misedited pattern produced by gG4-IV created an anchor sequence for another non-cognate gRNA (gMURF4-IV), which could extend this sequence in the 5′ direction. This is a clear example of misediting involving two successive misguiding gRNAs.

Additional evidence was obtained with another group of seventeen clones of partially edited G5 transcripts, some of which are shown in Fig. 2B–D. They were distinguished from the first group by the location of polyadenylation sites approximately 25 nucleotides upstream from the sites in the clones in Fig. 2A. Some 3′-end heterogeneity was also observed in this group of clones. None of these shorter clones contained the correct block I editing pattern exhibited by the clones in Fig. 2A. The editing patterns could be classified into four subgroups, three of which are shown in Figs. 2B–D. One long poly(A)-containing clone (clone 1.4, Fig. 2B) and some clones obtained with S-290 as a 3′ primer (clones 3.15, 3.15, Fig. 2C and 3.35, Fig. 2D) also fell within these subgroups. The sequences of the clones shown in Fig. 2B were complementary to the sequence of another non-cognate gRNA (gND7-I)

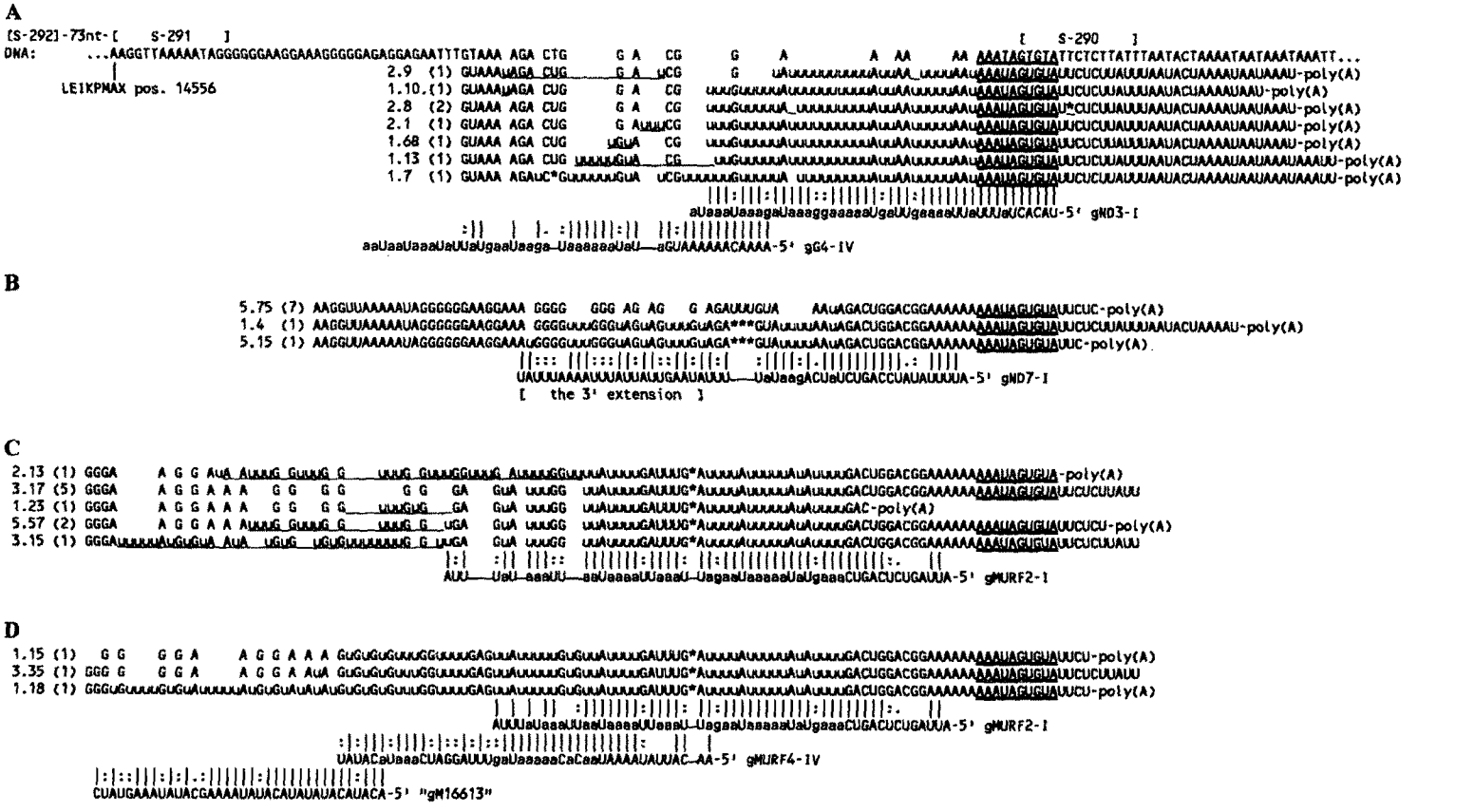

Fig. 2. Partially edited and misedited G5 mRNA clones from the UC strain of *L. tarentolae*. The clones are grouped into four subgroups (A-D). The sequences of the subgroups are aligned according to the anchor sequence for gND3-I (double underlined), so as to make the positions of the 3′ ends comparable. Inserted U's are indicated by lowercase, deleted Us by asterisks. Editing patterns for which gRNAs were not found are underlined. PCR amplifications were performed using as primers, oligo-dT and S-291 (clones 1.10, 1.6B, 1.13, 1.7, 1.4, 1.23, 1.15 and 1.18), oligo-dT and S-292 (clones 2.9, 2.8, 2.1, 2.13), S-290 and S-291 (3.17, 3.35), S-290 and S-292 (5.75, 5.57). The numbers of clones obtained with identical sequences are given in parentheses next to the clone designations.

which normally mediates the editing of block I of the ND7 cryptogene. Interestingly, the 5′ extension of this misedited pattern could be produced by a 3′-end extension of this gRNA. Blum et al [26] showed that the majority of the gND7-I RNAs in the UC strain had 3′ ends as indicated in Fig. 2B. This suggests that a subclass of gND7-I RNAs with extended 3′ transcribed sequences were responsible for this misediting pattern.

Another group of clones shown in Fig. 2C had incorrectly edited sequences which were complementary to the sequences of the gMURF2-I RNA, which normally mediates the editing of block I of transcripts of the MURF2 cryptogene. This is another clear example of misediting by a non-cognate gRNA. However, in this case, the incorrectly edited sequences upstream of this misedited region could not be correlated with any known gRNA sequences in the UC strain. It is possible that there are still some low abundance gRNAs in the UC strain that have not yet been detected. It is also possible that incorrect editing patterns can be produced in some cases by an alternate mechanism.

A third group of clones shown in Fig. 2D had incorrectly edited sequences which were complementary to three different non-cognate gRNAs, each of which created an anchor sequence for the adjacent upstream gRNA as in the case of normal productive editing [13]. The initial gRNA (gMURF2-I), formed a short false anchor with unedited G5 mRNA sequence and mediated misediting for at least 33 nucleotides, thereby creating a false anchor for the gMURF4-IV RNA, which mediates misediting for another 32 nucleotides. A third putative gRNA, gM16613, which was deduced by computer analysis of the known maxicircle sequence (GenBank entry LEIKPMAX), could form a short false anchor and continue the observed misediting for another 30 nucleotides.

It is of some interest that none of the 3′ truncated clones shown in Fig. 2B-D exhibited the correct block I editing pattern in spite of the fact that the anchor sequence is present in all but one clone (Fig. 2C). In other cryptogenes in *L. tarentolae*, distances between the polyadenylation sites and the cognate block I anchors vary from as little as 7 nt in G4, to as much as 45 nt in G3 [24]. In the correctly edited G5 transcripts in the LEM125 strain, the distance between the block I anchor and the 3′ end of the mRNA is 34 nucleotides [24], which is approximately identical to the distance in the 'long' clones in Fig. 2A. However, in the 'short' clones in Fig. 2B-D, which do not exhibit correct block I editing, the distance is only 0–6 nucleotides. This suggests the possibility of some minimal space requirements for the assembly of an editing on the 3′ end of a pre-edited mRNA.

The probability that the extensive matches between the observed incorrectly edited sequences and the non-cognate gRNAs (which are present in the UC strain of *L. tarentolae*) is due to chance is extremely low. Furthermore, neither the 'match protection' model [17] nor the 'progressive realignment' model [19] can provide an adequate explanation for these incorrectly edited sequences. These data demonstrate that in the absence of cognate gRNAs for a particular editing cascade, extensive misediting by non-cognate gRNAs can occur. The same phenomenon appears to occur in recently isolated *L. tarentolae* cells containing a full complement of gRNAs, but to a lesser degree. The existence in *T. brucei* of redundant gRNAs which differ in sequence but contain the identical editing information is most likely responsible for the high frequency of misedited sequences observed in partially edited RNAs in that species [17].

Our results do not eliminate the possibility that alternate mechanisms or a sloppiness in the editing machinery are responsible for some incorrect editing, since non-cognate gRNAs have not been identified in every case. They do, however, provide additional evidence for the validity of the misediting hypothesis.

Acknowledgements

The authors acknowledge the participation of B. Niner, E. Gruzsynski and Dr. N. Sturm in obtaining the partially edited G5 mRNA clones. We also thank members of the Simpson lab for discussions. The work was supported in part by a research grant AI-09102 to L.S. from the National Institutes of Health. D.A.M. is on leave of absence from Moscow State University. O.T. was supported by a predoctoral fellowship from CNPq, Brasil.

References

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

[2] Stuart, K. (1993) RNA editing in mitochondria of African trypanosomes, In: RNA Editing: The Alteration of Protein Coding Sequences of RNA (Benne, R., ed.), pp. 25–52. Ellis Horwood, New York.

[3] Hajduk, S.L., Harris, M.E., and Pollard, V.W. (1993) RNA editing in kinetoplastid mitochondria. FASEB J. 7, 54–63.

[4] Benne, R. (1994) RNA editing in trypanosomes. Eur. J. Biochem. 221, 9–23.

[5] Cech, T.R. (1991) RNA editing: World's smallest introns. Cell 64, 667–669.

[6] Blum, B. and Simpson, L. (1992) Formation of gRNA/mRNA chimeric molecules in vitro, the initial step of RNA editing, is dependent on an anchor sequence. Proc. Natl. Acad. Sci. USA 89, 11944–11948.

[7] 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.

[8] Harris, M., Decker, C., Sollner-Webb, B., and Hajduk, S. (1992) Specific cleavage of pre-edited mRNAs in trypanosome mitochondrial extracts. Mol. Cell. Biol. 12, 2591–2598.

[9] Pollard, V.W., Harris, M.E., and Hajduk, S.L. (1992) Native mRNA editing complexes from Trypanosoma brucei mitochondria. EMBO J. 11, 4429–4438.

[10] Göringer, H.U., Koslowsky, D.J., Morales, T.H., and Stuart, K. (1994) The formation of mitochondrial ribonucleoprotein complexes involving guide RNA molecules in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. USA 91, 1776–1780.

[11] Read, L.K., Göringer, H.U., and Stuart, K. (1994) Assembly of mitochondrial ribonucleoprotein complexes involves specific guide RNA (gRNA)-binding proteins and gRNA domains but does not require preedited mRNA. Mol. Cell. Biol. 14, 2629–2639.

[12] Peris, M., Frech, G.C., Simpson, A.M., Bringaud, F., Byrne, E., Bakker, A., and Simpson, L. (1994) Characterization of two classes of ribonucleoprotein complexes possibly involved in RNA editing from Leishmania tarentolae mitochondria. EMBO J. 13, 1664–1672.

[13] 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, 459–467.

[14] Souza, A.E., Myler, P.J., and Stuart, K. (1992) Maxicircle CR1 transcripts of Trypanosoma brucei are edited, developmentally regulated, and encode a putative iron-sulfur protein homologous to an NADH dehydrogenase subunit. Mol. Cell. Biol. 12, 2100–2107.

[15] Sugisaki, H. and Takanami, M. (1993) The 5'-terminal region of the apocytochrome *b* transcript in *Crithidia fasciculata* is successively edited by two guide RNAs in the 3' to 5' direction. J. Biol. Chem. 268, 887–891.

[16] 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.

[17] Decker, C.J. and Sollner-Webb, B. (1990) RNA editing involves indiscriminate U changes throughout precisely defined editing domains. Cell 61, 1001–1011.

[18] Sturm, N.R. and Simpson, L. (1990) Partially edited mRNAs for cytochrome b and subunit III of cytochrome oxidase from Leishmania tarentolae mitochondria: RNA editing intermediates. Cell 61, 871–878.

[19] 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.

[20] Sturm, N.R., Maslov, D.A., Blum, B., and Simpson, L. (1992) Generation of unexpected editing patterns in Leishmania tarentolae mitochondrial mRNAs: misediting produced by misguiding. Cell 70, 469–476.

[21] 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. Nucleic Acids Res. 19, 6277–6281.

[22] Parrot, L. and Foley, H. (1939) Sur la fréquence de la leishmaniose du gecko dans le Sud oranais. Arch. Inst. Pasteur d'Algerie 17, 231–232.

[23] Wallbanks, K., Maazoun, R., Canning, E., and Rioux, J. (1985) The identity of *Leishmania tarentolae* Wenyon 1921. Parasitology 90, 67–78.

[24] Thiemann, O.T., Maslov, D.A. and Simpson. L. (1994) Disruption of RNA editing in *Leishmania tarentolae* due to the loss of minicircle-encoded guide RNA genes. Submitted for publication.

[25] Read, L.K., Wilson, K.D., Myler, P.J., and Stuart, K. (1994) Editing of *Trypanosoma brucei* maxicircle CR5 mRNA generates variable carboxy terminal predicted protein sequences. Nucleic Acids Res. 22, 1489–1495.

[26] Blum, B. and Simpson, L. (1990) Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo-(U) tail involved in recognition of the pre-edited region. Cell 62, 391–397.