ELSEVIER

Molecular and Biochemical Parasitology 79 (1996) 229–234

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

Short communication

Analysis of the 3′ uridylylation sites of guide RNAs from *Leishmania tarentolae*

Otavio H. Thiemann[a], Larry Simpson[a,b,c,*]

[a]*Department of Molecular, Cellular and Developmental Biology, UCLA, Los Angeles, CA 90095-1662, USA*
[b]*Howard Hughes Medical Institute, 6780 MacDonald Building, 675 Circle Drive S., UCLA, Los Angeles, CA 90095-1662, USA*
[c]*Department of Medical Microbiology and Immunology, UCLA, Los Angeles, CA 90095-1662, USA*

Received 22 December 1995; revised 1 May 1996; accepted 6 May 1996

Keywords: RNA editing; *Leishmania*; gRNA

The templates for RNA editing in kinetoplastid mitochondria are small 3′ oligo-uridylylated RNA molecules termed guide RNA (gRNAs), which show complementarity to the mature edited RNAs, provided G-U base pairs are allowed. The gRNAs are encoded mainly in the minicircle component of the kinetoplast DNA, but a few are also encoded in the maxicircle DNA. Each gRNA mediates the editing of a single 'block' of mRNA sequence [1–3].

Little is known about processing of gRNA transcripts. The 5′ ends of minicircle- and maxicircle-encoded gRNAs from *L. tarentolae* are fairly homogeneous, as determined by primer extension sequencing [4–6]. Guide RNAs are an excellent substrate for GTP-capping by vaccinia virus guanylyl transferase and therefore possess 5′ di- or tri-phosphates which may represent primary 5′ ends [7].

* Corresponding author. Howard Hughes Medical Institute, 6780 MacDonald Building, 675 Circle Drive S., UCLA, Los Angeles, CA 90095-1662, USA.

The 3′ termini of several maxicircle-encoded gRNAs from *L. tarentolae* and *Trypanosoma brucei* have been directly examined by S1 protection and RNA sequencing [4,5,7], and by sequencing PCR-amplified gRNA/mRNA chimeric molecules [8–10]. Several maxicircle gRNA genes in *L. tarentolae* have short oligo[T] sequences which may represent termination signals [8,11]. Some evidence for premature termination (or processing) of the *L. tarentolae* maxicircle-encoded MURF4-II gRNA was provided by the observation of several minor 5′-GTP-cappable species which appeared to terminate at stretches of encoded uridine (U) residues [7].

However, nothing is known about termination of transcription and/or 3′ end processing of minicircle-encoded gRNAs. By a combination of sequencing chimeric molecules and direct analysis of 3′ ends of gRNAs cloned by 3′ RACE, extensive heterogeneous 3′ truncations were observed for several maxicircle-encoded gRNAs from *Crithidia fasciculata* [12], and the authors suggested that the apparent 3′ to 5′ progression of editing within an

PII S166-6851(96)02653-9

editing block observed in partially edited mRNAs from *L. tarentolae* might be an artifact of such 3′ truncations. However, no evidence for such 3′ truncations was observed in a 3′ RACE library of minicircle-encoded gRNAs from another strain of *C. fasciculata* [13].

Since the extent of 3′ end heterogeneity of gRNAs can affect the interpretation of experimental data on the polarity of editing, we have directly addressed this question in the case of the *L. tarentolae* LEM125 strain by sequencing a collection of clones from a gRNA library constructed by 3′/5′ RACE [14]. We showed previously that the LEM125 strain has a larger repertoire of minicircle-encoded gRNAs than the UC strain, probably as a result of the loss of multiple minicircle sequences classes in the UC strain due to its extended culture history.

The library construction method maintained the genomically encoded 3′ termini as well as portions of the 3′ oligo[U] tails of the gRNAs, thus allowing an analysis of the extent of 3′ end heterogeneity of the transcripts. The 5′ ends of the gRNAs were also preserved in the cDNA clones, provided there was no premature termination of reverse transcription during the first strand synthesis. In order to eliminate PCR-amplified identical cDNA clones from the dataset, unique clones were selected for analysis by the length of the cloned oligo[U] tail and the extent of genomically encoded 3′ end sequence.

LEM125 gRNA clones representing known UC strain gRNAs were analyzed first: three clones of gRPS12-II, two clones of gRPS12-VIII, two clones of gCOIII-I, five clones of gA6-I. The gRNA sequences were aligned with the edited mRNA sequences previously determined for the UC strain (Fig. 1A). Two general features were observed from these alignments: a remarkable homogeneity of the oligo[U] insertion sites within a given gRNA sequence class and a consistency with the 3′ end positions previously determined or predicted for the UC strain gRNAs [4,7,15].

The alignments of the LEM125-specific gRNA clones with edited mRNA sequences shown in Fig. 1B also illustrate a remarkable homogeneity of the 3′ ends of the gRNA transcripts.

The alignments of the gRNA clones also showed a high degree of homogeneity of the genomically encoded 5′ ends, as was shown previously for several maxicircle-encoded and minicircle-encoded gRNAs by direct primer extension sequencing [4,5,10]. The 5′ deletions observed with clones of gND9-V and gG4-II are probably due to errors in the reverse transcription step (Fig. 1B).

Two clones, g194 (gND8-VII) and g207 (gND8-IX) showed multiple nucleotide substitutions throughout the guiding portions of the gRNAs (Fig. 1B). These gRNAs edit blocks VII and IX of the ND8 mRNA. The mismatches between g194 and g183, and between g207 and g159, are mostly A to G transitions, which would not affect the hybridization of the gRNA to the mRNA or the guiding of U insertions [10,14]. The 3′ end of g207 is truncated by three nucleotides, which could potentially produce mis-editing of site 84 in ND8 mRNA by the addition of three instead of five U's. This gRNA has the same 5′ end as the redundant g159, although the anchor sequence has two transitions which will lower its stability compared to g207 (Fig. 1B). The limited heterogeneity at the 3′ ends of these gRNAs is consistent with the above observations for other gRNAs from the library. These examples (g194 and g207) represent 'redundant' gRNAs, which can encode the same editing information but have different sequences. Redundant gRNAs exist in relatively high abundance in *T. brucei* [10] and *T. cruzi* [16], but only a single pair of redundant gRNAs has been previously described in *L. tarentolae* LEM125: gND3-IIIa and b [14], and no redundant gRNAs were found in the UC strain of *L. tarentolae*.

We showed previously that the Lt19 gRNA in the UC strain of *L. tarentolae* corresponded to the G4-III gRNA in the LEM125 strain, but lacked an editing function in the UC strain due to the absence of the upstream gRNAs in the G4 editing cascade. We also showed that the Lt19 gRNA was 18 nucleotides longer at the 3′ end than the G4-III gRNA [14]. No evidence for the presence of a 3′-extended gG4-III gRNA in LEM125 kinetoplast RNA was obtained by primer extension sequencing and PCR amplification using a specific 3′ primer (data not shown).

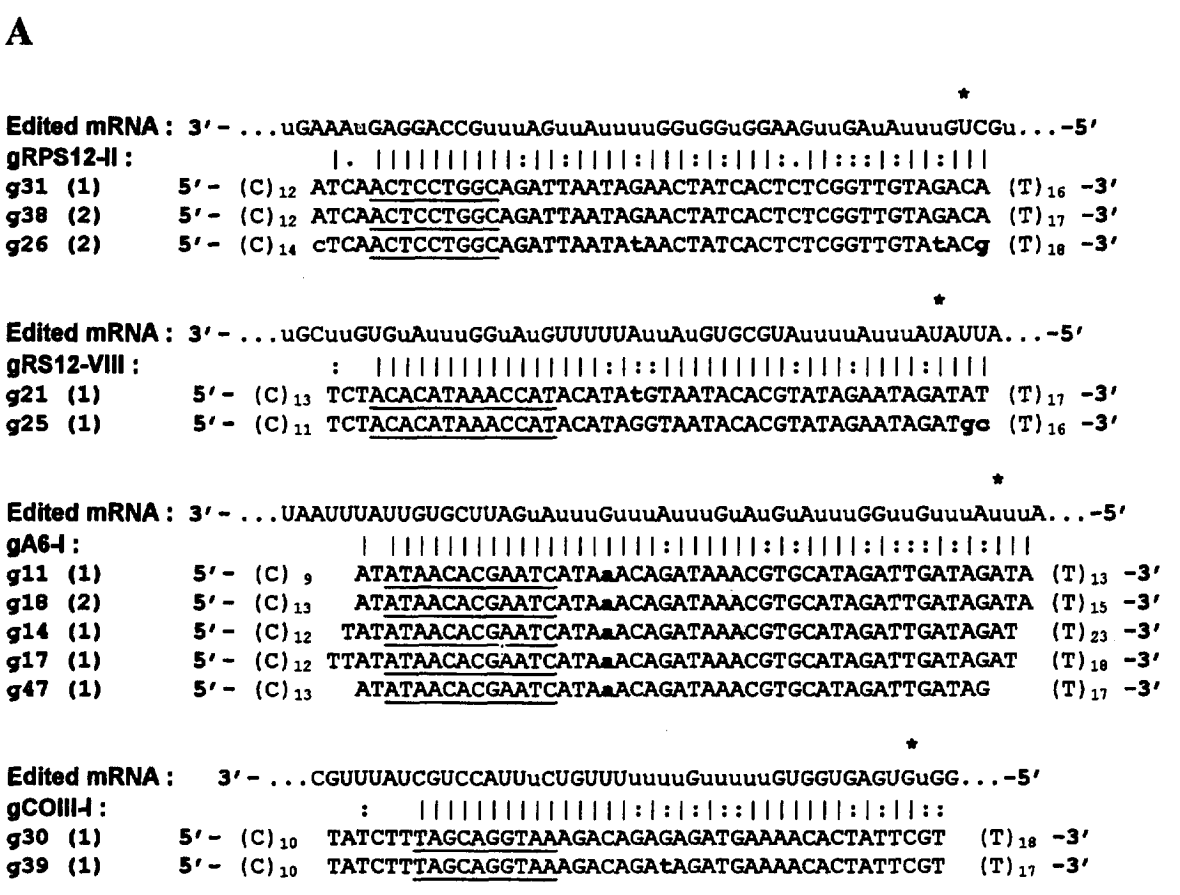

Fig. 1(a).

In order to investigate the extent of 3′ end heterogeneity in the G4-III gRNA population, three unique clones (**a**, **b** and **c** in Fig. 1A) derived by RT-PCR using a 5′ gG4-III-specific primer, and one clone (g232) from the original 3′/5′ RACE library were analyzed (Fig. 1B). Alignment of these clones with the edited mRNA sequence showed the same low level of 3′ end heterogeneity as evidenced above for other LEM125 gRNAs. The deletions in clone **c** may be due to premature termination of the reverse transcriptase in the first strand synthesis. All other clones showed identical 5′ ends as the Lt19 gRNA [5].

It is of some interest that the entire sequence of the gG4-III-encoding minicircle from LEM125 was found to be identical to the previously pub lished sequence of the Lt19 minicircle (data not shown), suggesting that these represent homologous genetic elements. The presence of an 18 nucleotide extended transcribed 3′ sequence for the presumed non-functional Lt19 gRNA in the UC strain of *L. tarentolae*, in spite of the fact that the Lt19 minicircle sequence is identical to the G4-III homologue in the LEM125 strain, suggests that 3′ end processing occurs to produce the shorter transcript in LEM125, but this must be established directly.

In summary, the majority of the LEM125 gRNAs analyzed had a homogeneous 3′ end which, in most cases overlapped the last predicted editing site. The limited amount of 3′ end heterogeneity could in some cases, lead to misediting by producing shifts in the 'guiding frame', as previously suggested [9]. Possible examples of this type of

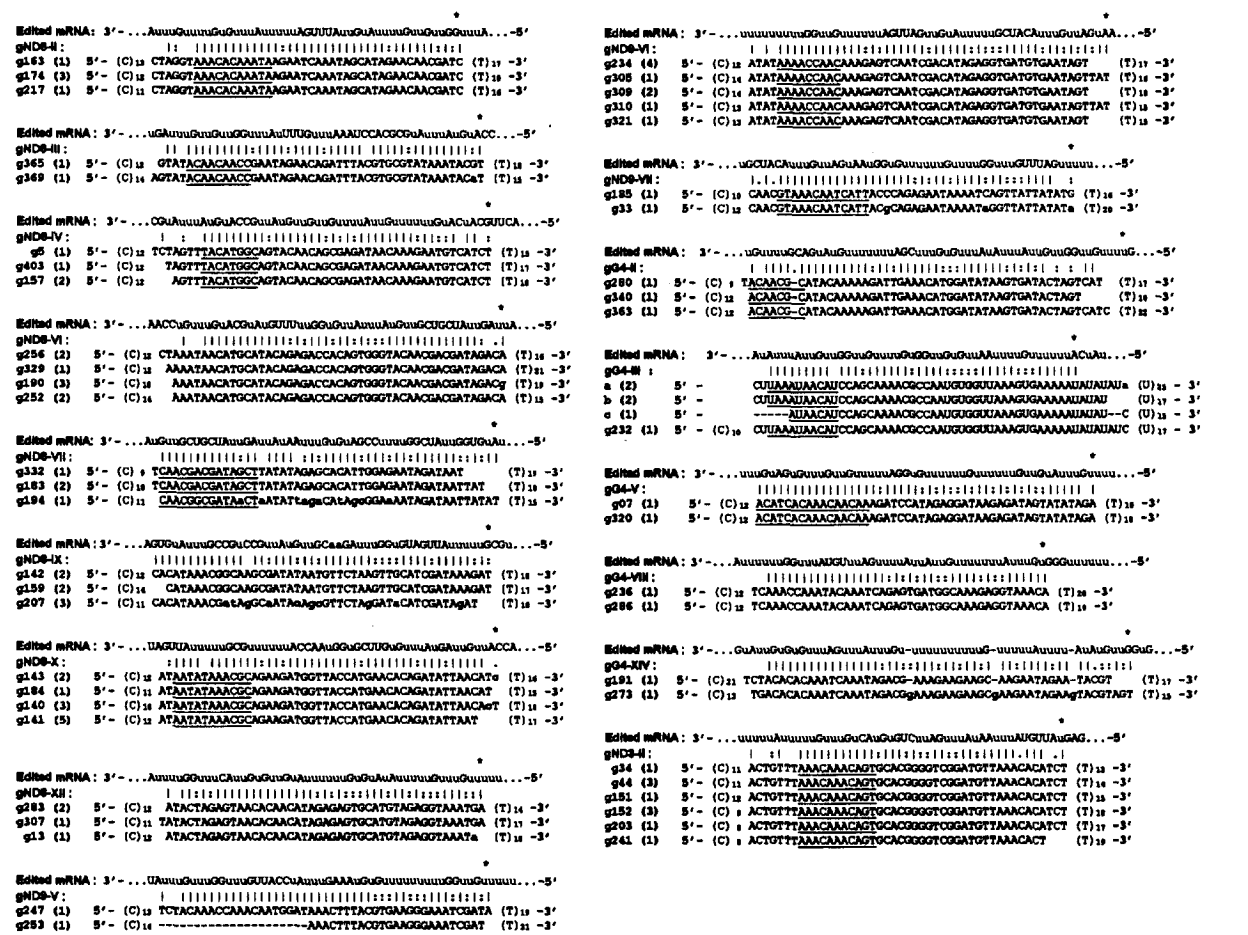

Fig. 1. Alignment of gRNAs from *L. tarentolae* LEM125. (A) gRNA clones representing known gRNAs from the UC strain: gRPS12-II, gRPS12-VIII, gA6-I and gCOIII-I. The gRNA clones obtained from the 3′/5′ RACE library by random selection and sequencing are shown aligned with the cognate edited mRNA block in each case (3′ to 5′). The U's added by editing in the mRNA are in lower case. U-deletions are not shown. Canonical base pairs between the gRNA and mRNA are indicated by vertical lines (|) and G-U base pairs by colons (:). The gRNA anchor sequences are underlined. The clone numbers are shown at the left, and the numbers of clones which shared the same sequence and had the same number of T residues at the 3′ end are indicated in parentheses. Only a single example of each was used for this analysis. The number of T residues in the 3′ oligo[U] tail of each clone is indicated in parenthesis at the right, and the number of C residues added during the cDNA cloning to the 5′ end is indicated in parentheses on the left. Nucleotide mismatches between the LEM125 and the known UC gRNAs are indicated in bold lower case. The consensus genomically-encoded 3′ end is indicated by an asterisk above the edited mRNA. Deletions in the gRNA clones are indicated by dashes. (B) Alignment of gRNAs from *L. tarentolae* LEM125 and their cognate mRNA edited blocks. For details see legend to Fig. 1A. The anchor sequences when known are indicated by underlines. Lower case bold characters indicate mismatches between the gRNA clones, and can indicate the presence of redundant gRNAs (g194/gND8-VII and g207/gND8-IX) or errors occurring in the reverse transcriptase or PCR steps. gG4-III **a**, **b** and **c** represent three unique clones obtained by RT-PCR using a 5′ primer specific to this anchor sequence.

misediting include gND8-IV, gND8-VI, gND8-X and gND9-VII. These misediting events due to 3′-extended gRNAs would be corrected by the following gRNA in the editing cascade. Similar results were observed with chimeric molecules from *T. brucei* [10].

It is not known if primary transcripts terminate at the 3′ ends observed in these 3′/5′ RACE

library or are endonucleolytically processed prior to addition of uridylylate residues by the mitochondrial terminal uridylyl transferase.

These data are consistent with the precise 3′ to 5′ polarity of editing within an editing block suggested by the data of Sturm and Simpson [17] obtained from a library of partially edited cytochrome *b* mRNAs. The extensive junction region misediting observed in cytochrome oxidase subunit III and RPS12 mRNAs in *L. tarentolae* and in multiple mRNAs in *T. brucei* can be explained, at least in part, by the proposed misguiding mechanisms [9]. The existence of a large number of redundant minicircle-encoded gRNAs in *T. brucei*, which overlap to variable extents [10,18], is consistent with the higher level of misediting in the former species, since the heterogeneous length of these redundant gRNA could possibly account for the misincorporation of U's at the boundaries of two edited blocks. These events would be corrected by the subsequent gRNA in the editing cascade, producing a correctly edited mRNA.

Acknowledgements

We would like to thank Dr. Dmitri Maslov for helpful suggestions and assistance, Dr. Georges Frech for a detailed review and discussion of this manuscript and Anthea Ramos for technical assistance. O.H.T. was supported by a predoctoral fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil). This work was supported in part by grant AI-09102 from the National Institutes of Health to L.S.

References

[1] 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.

[2] 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.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] Riley, G.R., Corell, R.A. and Stuart, K. (1994) Multiple guide RNAs for identical editing of *Trypanosoma brucei*apocytochrome *b* mRNA have an unusual minicircle location and are developmentally regulated. J. Biol. Chem. 269, 6101–6108.

[11] 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.

[12] Arts, G.J., Van der Spek, H., Speijer, D., Van den Burg, J., Van Steeg, H., Sloof, P. and Benne, R. (1993) Implications of novel guide RNA features for the mechanism of RNA editing in *Crithidia fasciculata*. EMBO J. 12, 1523–1532.

[13] Yasuhira, S. and Simpson, L. (1995) Minicircle-encoded guide RNAs from *Crithidia fasciculata*. RNA 1, 634–643.

[14] Thiemann, O.H., Maslov, D.A. and Simpson, L. (1994) Disruption of RNA editing in *Leishmania tarentolae* by the loss of minicircle-encoded guide RNA genes. EMBO J. 13, 5689–5700.

[15] Sturm, N.R. and Simpson, L. (1990) Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. Cell 61, 879–884.

[16] Avila, H. and Simpson, L. (1995) Organization and complexity of minicircle-encoded guide RNAs from *Trypanosoma cruzi*. RNA 1, 939–947.

[17] 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.

[18] Corell, R.A., Feagin, J.E., Riley, G.R., Strickland, T., Guderian, J.A., Myler, P.J. and Stuart, K. (1993) *Trypanosoma brucei* minicircles encode multiple guide RNAs which can direct editing of extensively overlapping sequences. Nucleic Acids Res. 21, 4313–4320.