

Small ncRNA transcriptome analysis from kinetoplast mitochondria of *Leishmania tarentolae*

Monika J. Madej¹, Juan D. Alfonzo² and Alexander Hüttenhofer^{1,*}

¹Innsbruck Biocenter, Division of Genomics and RNomics, Innsbruck Medical University, Fritz-Pregl-Strasse. 3, 6020 Innsbruck, Austria and ²Department of Microbiology, Ohio State Biochemistry Program and The RNA Group, The Ohio State University, 484 West 12th Ave., Columbus, OH 43210, USA

Received October 17, 2006; Revised December 21, 2006; Accepted December 22, 2006

ABSTRACT

Gene expression in mitochondria of kinetoplastid protozoa requires RNA editing, a post-transcriptional process which involves insertion or deletion of uridine residues at specific sites within mitochondrial pre-mRNAs. Sequence specificity of the RNA editing process is mediated by oligo-uridylated small, non-coding RNAs, designated as guide RNAs (gRNAs). In this study, we have analyzed the small ncRNA transcriptome from kinetoplast mitochondria of *Leishmania tarentolae* by generating specialized cDNA libraries encoding size-selected RNA species. Through this screen, a significant number of novel oligo-uridylated RNA species, which we have termed oU-RNAs, has been identified. Most novel oU-RNAs are present as stable RNA species in mitochondria as assessed by northern blot analysis. Thereby, novel oU-RNAs show similar expression levels and sizes as previously reported for canonical gRNAs. Several oU-RNAs are transcribed from both strands of the maxicircle and minicircles components of the mitochondrial genome, from regions where up till now no transcription has been reported. Two stable oU-RNAs exhibit an anchor sequence in antisense orientation to known gRNAs and thus might regulate editing of respective pre-mRNAs. A number of oU-RNAs map in antisense orientation to non-edited protein-coding genes suggesting that they might function by a different mechanism. In addition, our screen shows that all kinetoplast-derived RNAs are prone to some degree of uridylation.

INTRODUCTION

Leishmania tarentolae belongs to the order Kinetoplastida which comprises flagellated protozoan parasites that cause a wide spectrum of diseases (1–4). All Kinetoplastida species possess a single mitochondrion, the kinetoplast, containing its own genome (kinetoplast DNA, kDNA) (5). kDNA consists of thousands of circular DNA molecules exhibiting two different sizes, designated as maxicircles and minicircles, which are catenated in a network (5).

Maxicircles, with a size of 23–40 kb, depending on the species, are homogenous in sequence and are reminiscent of typical eukaryal mitochondrial genomes in that they encode small and large subunit ribosomal RNAs (rRNAs) and a number of respiratory chain proteins. In contrast to most eukaryal mitochondria: (i) kinetoplast mitochondrial genomes completely lack tRNA genes, (ii) hence, all tRNAs are encoded by the nuclear genome and imported from cytosol and (iii) some of the mitochondrial mRNA genes are encoded as so-called cryptogenes, which have to be edited by the insertion or deletion of uridine (U) residues at specific sites in the pre-mRNAs to generate translatable open reading frames (ORFs) (6,7).

Editing is catalyzed by a ribonucleoprotein complex, termed the editosome (8,9). Precise sequence information required for mRNA editing is provided by small oligo (U)-tailed RNAs, termed guide RNAs (gRNAs), which base-pair to pre-edited mRNA downstream to an editing site (10,11). gRNAs possess three functional domains: an anchor domain, which anneals to the pre-edited mRNA, a guiding domain, which directs U insertion or deletion and a 3'-oligo(U) tail, which is added post-transcriptionally. The 3'-oligo(U) tail is thought to tether the purine-rich 5'-cleavage fragment of the pre-mRNA intermediate (see below) to the editosome complex (10,12,13). Annealing of the gRNA to the

*To whom correspondence should be addressed. Tel: +43(0) 512 9003 70250; Fax: +43(0) 512 9003 73100; Email: alexander.huettnerhofer@i-med.ac.at
Correspondence may also be addressed to Juan D. Alfonzo. Tel: +1 614 292 0004; Fax: +1 614 292 0141; Email: Alfonzo.1@osu.edu

pre-mRNA, targets endonuclease cleavage of the pre-mRNA immediately upstream of the anchor duplex, thus generating a 5'-cleavage fragment that is the substrate for either a 3' terminal uridylyl transferase (TUTase) in case of U insertion or a 3' to 5' exonuclease for U deletion (8,9).

Several transcripts containing more than one mRNA have been identified, indicating that maxicircles are transcribed polycistronically and subsequently processed by a yet unknown mechanism (14–16). Thereby, maturation activities for maxicircle transcripts include cleavage to generate monocistrons, polyadenylation of mRNAs and addition of oligo(U) tails to rRNAs and gRNAs.

Most gRNAs are encoded on heterogeneous minicircles of 0.8–2.5 kb in size, depending on the species, while a much smaller number is encoded on the maxicircle genome (11,17,18). Both maxicircle- and minicircle-encoded guide RNAs are thought to originate from independent promoters since transcripts can be labeled with [γ - 32 P] GTP by guanylyltransferase which specifically recognizes RNAs with a 5' di- or triphosphate, characteristic for unprocessed primary transcripts (12,19). However, no consensus-promoter elements for gRNAs have been described to date. The number of gRNAs encoded per minicircle also varies depending on the species. In *L. tarentolae*, minicircles are thought to encode a single gRNA, while in *T. brucei* three to four gRNA genes per minicircle are observed (20,21). Transcription from minicircles of *T. brucei* proceeds polycistronically (22). While the processing of the 3'-ends of gRNA precursors has been demonstrated, no evidence for processing of 5'-ends has been obtained (22). Therefore, a model for minicircle transcription within *T. brucei* has been proposed, in which transcription of polycistronic guide RNAs initiates at a point that corresponds precisely to that of the mature 5'-end of each gRNA. Following transcription, the 3'-end of the first gRNA (in a polycistron) is processed to produce a monocistron, while the downstream sequence is degraded (22). According to this model, in *L. tarentolae* where minicircles encode a single gRNA, initiation of transcription is determined by the 5'-end of the gRNA gene itself. Therefore each minicircle of *L. tarentolae* is predictably transcribed as a single transcription unit.

In this study, we have analyzed the small transcriptome of kinetoplast mitochondria by generating cDNA libraries encoding small non-coding RNA (ncRNA) species sized from 15–600 nt. Thereby, we identified novel, stable, oligo(U)-tailed maxicircle- as well as minicircle-encoded ncRNAs bearing resemblance to canonical gRNAs. In addition, we show that minicircles of *L. tarentolae* represent multiple transcription units, where transcription occurs from both strands of the minicircles. Our results demonstrate that a majority of transcripts expressed from the kinetoplast genome are prone to some degree of oligo-uridylation, an observation which might shed light on kinetoplast genome evolution.

MATERIALS AND METHODS

Cell culture, isolation of mitochondria and RNA extraction

Leishmania tarentolae-UC strain cells were grown in a New Brunswick Bioflow IV fermenter at 27°C in brain–heart infusion (BHI) medium (Difco) supplemented with 10 µg/ml hemin. LEM125 strain mitochondria were isolated as previously described (23). Kinetoplast RNA was isolated using the TRIzol reagent (Gibco BRL), as described previously (24).

Oligonucleotides

All oligonucleotides used in this study are listed in the Supplementary Data.

Generation of a cDNA library encoding small RNA species and 5'/3' RACE assay

About 70 µg of kinetoplast RNA was treated with DNase RQ1 (Promega) and was subsequently size-fractionated by denaturing 8% (w/v) PAGE (7 M urea, ×1 TBE buffer). RNAs in the size range between ~15 and 600 nt were excised from the gel in three fractions (a) from ~600 to 90 nt, (b) from ~90 to 70 nt (tRNA fraction) and (c) from ~70 to 15 nt, passively eluted, and ethanol-precipitated. Eluted RNA fractions were poly(C)-tailed. The C-tailing reaction was carried out in a volume of 50 µl containing 1× C-tailing buffer (50 mM Tris/Cl pH 8.0, 200 mM NaCl, 10 mM MgCl₂, 2 mM MnCl₂, 0.4 mM EDTA, 1 mM DTT), 2 mM CTP and 2 U of poly (A) polymerase (Invitrogen). Poly (C)-tailed RNAs were treated with tobacco acid pyrophosphatase (Epicentre Technologies) and ligated to a 5'-oligonucleotide linker as described (25). Poly(C)-tailed and 5'-adaptor-ligated RNAs were reverse transcribed using SuperScript II reverse transcriptase (Invitrogen) and amplified by PCR, followed by cloning into a pGEM-T vector (Promega). Additionally, in order to enrich our screen for uridylated transcripts, a modified uridine-track anchor PCR primer (3'libPCR-U-anchor) was used. For 5' and 3' RACE (Rapid amplification of cDNA ends) assay, internal PCR-primer specific for selected oU-RNAs and 5'libPCR-c primer or 3'libPCR-c primer, respectively, were used. Amplified products were cloned into a pGEM-T vector (Promega) and sequenced on an ABI Prism 3100 capillary sequencer (PE Applied Biosystems).

Sequence analysis of the cDNA library

cDNA clones were sequenced using the M13 reverse primer and the BigDye terminator cycle sequencing reaction kit (PE Applied Biosystems). Sequencing reactions were run on an ABI Prism 3100 (Perkin Elmer) capillary sequencer. Subsequently, sequences were analyzed with the LASERGENE sequence analysis program package (DNASTAR, Madison, USA), followed by a BlastN search against the GenBank database (NCBI).

Northern blot analysis

A 5 µg of total kinetoplast RNA was denatured for 1 min at 95°C, separated on a 8% (w/v) denaturing

polyacrylamide gel (7M urea, 1× TBE buffer) and transferred onto a Hybond-N⁺ membrane (Amersham Biosciences) using the semi-dry blotting apparatus (BioRad). For electrophoresis on formaldehyde-agarose gels, 5 μg of total kinetoplast RNA was denatured for 15 min at 60°C and run on a formaldehyde-1.2% agarose gel (2% (w/v) formaldehyde, 1× RP buffer). RNA was transferred onto a Hybond-N nylon membrane by capillary blotting and UV-cross-linked with 120 mJ in a UV-Stratalinker (Stratagene). Oligonucleotides were 5'-end-labeled with [γ -³²P] ATP and T4 polynucleotide kinase (Promega). Hybridization was carried out at 45°C in 1 M sodium phosphate buffer (pH 6.2), 7% (w/v) SDS for 12 h. Blots were first washed twice at room temperature in 2× SSC buffer (20 mM sodium phosphate, pH 7.4; 0.3 M NaCl; 2 mM EDTA), 0.1% (w/v) SDS for 15 min and then at 58°C in ×0.1 SSC, 0.5% (w/v) SDS for 1 min. Membranes were exposed on a phosphorimager cassette (Fujifilm) and signal was quantified with Image Quant software (molecular dynamics, version 5.2) and normalized to the hybridization signal obtained for 9S rRNA.

Primer extension analysis

A 2 μg of kinetoplast RNA was hybridized to 0.3 pmol of DNA primer labeled with [γ -³²P] ATP and T4 polynucleotide kinase (Promega). In a total volume of 10 μl containing 25 mM Tris/Cl pH 8.4 and 30 mM KCl, RNA was first denatured at 94°C for 2 min and then hybridized with primer by slow cooling to 42°C. Subsequently, extension mixture was added to a total volume of 20 μl containing: 10 mM MgCl₂, 100 mM Tris/Cl pH 8.4, 10 mM DTT, 0.5 mM dNTP, 0.5 U of AMV Reverse Transcriptase (Promega) and reverse transcription was performed at 42°C for 45 min. Extension product was separated on 8% (w/v) denaturing polyacrylamide gel. Gel was exposed to Kodak MS-1 film from 30 min to 3 h.

RESULTS AND DISCUSSION

Library construction and sequence analysis of cDNA clones

The goal of this study was to characterize the small ncRNA transcriptome of *L. tarentolae* mitochondria. To that end we generated cDNA libraries from size-selected mitochondrial RNAs in the range from ~15 to 600 nt (materials and methods section). Thereby, a total of 600 cDNA clones were sequenced and further analyzed by first grouping identical cDNA clones, followed by bioinformatic analysis (BLASTN) on their genomic location (materials and methods section).

From the 600 cDNAs analyzed, 422 were assigned to randomly distributed fragments of kinetoplast or nuclear-encoded rRNAs, likely to represent degradation products, or nuclear-encoded mitochondrial tRNAs which are imported into kinetoplasts. We therefore focused on the remaining 178 clones from which 22% represented annotated gRNAs, while the remaining 78% mapped to distinct locations within the maxicircle or minicircle genome that had not been previously annotated. Remarkably, all cDNA clones regardless of their

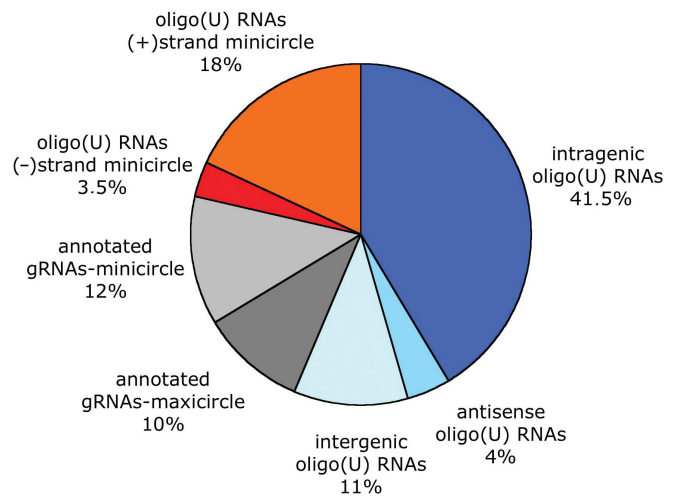


Figure 1. Sequence analysis of 178 cDNA clones representing different RNA species derived from the kinetoplast genome of the *L. tarentolae* US strain. cDNA clones were grouped into different classes, depending on their genomic location and are shown as percent of total number of clones. Oligo(U) RNAs: oligo-uridylated RNAs.

localization within the kinetoplast genome, exhibited oligo (U) tails at their 3'-ends (Figure 1). Thus, we refer to these novel unassigned gRNA candidates as oU-RNAs (oligo(U) RNAs). In contrast, oligo(U) tails were not observed for mitochondrial tRNAs (encoded by the nuclear genome), which are imported into kinetoplasts or other nuclear-encoded RNAs, which represent most likely contaminations of the mitochondrial RNA preparation.

Novel maxicircle-encoded oU-RNAs

Up till now, 15 gRNA genes have been described which are encoded on the maxicircle component of the *L. tarentolae* mitochondrial genome. These gRNAs are predominantly located within intergenic regions with three exceptions: (i) guide RNA gCO2, which is part of the 3'-end of the COII pre-mRNA and edits this pre-mRNA in *cis*; (ii) guide RNA gCYb-II, which overlaps the 3'-end of 9S rRNA in an antisense orientation and (iii) guide RNA gMURF2-I, which is located within the COII gene in an antisense orientation (10). Through our screen, we identified 24 novel oU-RNA species which map to different locations within the maxicircle genome including intergenic as well as intragenic regions, both in sense and antisense orientations (Table 1).

For the 13 novel oU-RNAs, we verified their expression by northern blot analysis and compared their expression to previously described maxicircle encoded gRNAs (i.e. gND7-I, gND7-II and gCyb-II, Figure 2) (18). To that end, equal amounts of total kinetoplast RNA was loaded onto each lane and hybridized to radio-labeled oligonucleotides complementary to respective gRNAs or oU-RNAs (materials and methods section). As a loading control, in addition, expression of 9S rRNA was analyzed. Hybridization signals of novel oU-RNA candidates were normalized to the expression level of 9S rRNA and subsequently quantified on a phosphorimager with

Table 1. Novel oligo(U) RNAs encoded on the maxicircle component of the mitochondrial genome of *L. tarentolae* categorized in four groups depending on their genomic location

Name	Sequence	cDNA clones	Northern blot	Exact location/Remarks
oU-RNAs derived from coding region: sense orientation, overlapping 5'termini of cryptogene				
Lmax1a	<u>ATTTATTTAGCCGACTACAGTATAAATAAAACACACTTGTATA(T)₆₋₁₀</u>	6	4.5	ND7 pre-edited domain II
Lmax1b	<u>ATTTTTTACCGACTACAGTAAATATATAAAAGACACTTGTAT(T)₆</u>	2	4.5	ND7 pre-edited domain II
Lmax2	<u>AAATAATATAAAGCGGAGAAAAGAAAAGGGCTTAACTTCAAGGT(T)₇</u>	3	3	Cytb pre-edited domain
Lmax3a	<u>TTATATAAAATTTATCAGATTAAGATAAATAAGGGTTGGAAAGAAATTG CGAATAGGAAGGGAT CCAGAAAGGAAATTAAGAACCCGAGAAAGGGAG(T)₁₀</u>	3	-	ATP6 pre-edited region, 3 cDNAclones exhibiting heterogeneous 3'end
Lmax3b	<u>TTATATAAAATAATAATCAGATTAAGATAAATAAGGGTTGGAAAGAAATTG CGAATAGGAAGGGAT CCAGAAAGGAAATTAAGAACCCGAGAAAGGGAG GAGAG(T)₁₄</u>			
Lmax3c	<u>TTATATAAAATAATAATCAGATTAAGATAAATAAGGGTTGGAAAGAAATTG CGAATAGGAAGGGAT CCAGAAAGGAAATTAAGAACCCGAGAAAGGGAG GAGAG(T)₁₃</u>			
Lmax4	<u>CTAATACCTATCGACCTATATAAATTATAATGGGTTGGAAAGGGAG TTTGCTTTTGTAAAT TAGGGTAAAGCCAC(T)₁₃</u>	1	-	RPS12 pre-edited domain II
oU-RNAs derived from coding region: sense orientation, of intragenic location				
Lmax5	<u>AGTAGACAGTGTACCCGATGAAGACCTGGATTAAGTAAAGTGTAGTT(T)₁₀</u>	23	3.5	ND7 pre-edited domain I
Lmax6	<u>CATGTCATGCACTTAGTAGCAGTTATCGGAAATGTGTAGT(T)₈</u>	4	-	ND7 not edited region
Lmax7	<u>AAGAAATTAGCACTGGTTCAAAATAAGAAATGGCTGCTCTTTGTTTTT(T)₂</u>	5	3	ND1 not edited gene
Lmax8	<u>ATATTTTAGGGIGATTTAATTTCCAAATAGGTTTCAATAATTAT(T)₁₁</u>	7	-	ND1 not edited gene
Lmax9	<u>AGTAAATTTAGAAAGTGTATTATATGATAGGAGACTTACGTTATT(T)₈</u>	2	1.3	CO2 in the vicinity of gMURF2-I, not edited domain
Lmax10	<u>GAAGAGTCGTAGTAAATGTGACACATTTGAAACATCTGATAAAAAATTT(T)₇</u>	7	0.8	Cytb not edited domain
Lmax11	<u>AGTTTAAATTTAGTGTGAAATTTGCATTTGTTTTGTTATACCAG(T)₅₋₁₂</u>	6	-	CO3 not edited domain
Lmax12	<u>ATTAACATAATTTAAAGTGTCCATAGAAAATTTAAAATTTATAACAATC(T)₈₋₁₃</u>	4	1.5	12S rRNA
oU-RNAs derived from coding region: antisense orientation				
Lmax13	<u>ATATAATCACAGTAAAATAATATAAAAAATAAACATACACACAAAAA(T)₁₀</u>	1	2.8	ND5
Lmax14	<u>ATCTATTAATAATAAATAAATAATAT(T)₁₁</u>	1	0.7	ND5
Lmax15	<u>TAGACCAATGCATAAAAAATACATGCTATCATAATACTAAATTA(T)₈</u>	1	3.2	MURF1
Lmax16	<u>ACATAAAAATATGTAGTTTTT(T)₈</u>	1	n.d.	MURF1
Lmax17	<u>TAATCATATC ATG AAGGGTCA TAT TGGATACTTAATAGAAAACATTATA (T)₁₁</u>	1	3	ND4
Lmax18	<u>TAATATACCTCAGGATGACCACAAAAATCAAAAAATATGTGAAATAGTA(T)₉</u>	1	-	CO1
Lmax19	<u>ACACAGGATAGTCAGATAATCTACGAGGAATGCATACATACCTAAA(T)₁₃</u>	1	-	CO1
oU-RNAs derived from divergent region (DR)				
Lmax20	<u>ATATGATTTGTCGACAGTTTAAAGTACGTTTGTCTAAATTTATAGA(T)₆₋₁₃</u>	7	0.7	strand minus
Lmax21a	<u>GAAATCGT AGTTAAATTAATAAAGCACAAAACATTTT(T)₃₋₁₂</u>	6	-	plus
Lmax21b	<u>GAAATCGTAGTAAATTAATAAATAATAAACAAA(T)₁₂</u>	1	-	plus
Lmax21c	<u>GAAATCGTAGTAAATTAATAAATAAAC(T)₄</u>	1	-	plus
Lmax22	<u>AAATTTGAATCAAGTAAATTTCCGAAAATTTTGACAGAAAAGTTCCA(T)₇</u>	3	0.8	plus
Lmax23	<u>CTTATAATTTTGTGACCATACAAAAATTCGTTCAATAATTTT</u>	1	-	minus
Lmax24	<u>GATAGAATTGAAATAAATAATTGACTTATTACCTTTT</u>	1	-	plus

Name: designation of oligo(U) RNA species; cDNA clones: number of independent cDNA clones identified from each RNA species in the cDNA library; Northern blot: number indicates the relative expression level of the respective oligo(U) RNA with respect to gND7-II, whose expression is assessed as 1; pre-edited domains within oligo(U) RNAs are indicated in bold; (T)_x: number of uridine residues not encoded by the mitochondrial genome; 5'-extensions of oU-RNAs, as assessed by 5'-RACE are indicated in italics; sequence polymorphisms between Lmax21a, b and c are underlined.

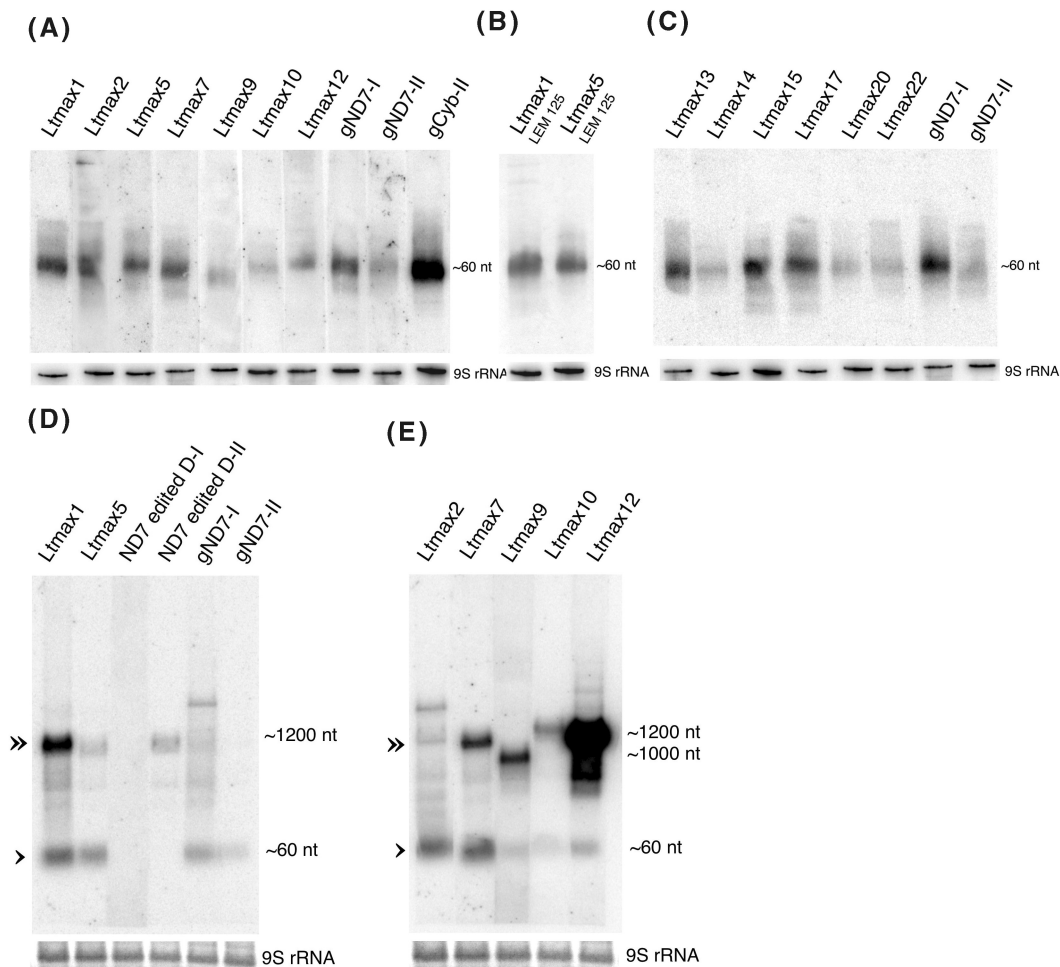


Figure 2. Northern blot analysis of selected maxicircle-derived oligo(U) RNAs and reported guide RNAs; sizes of RNAs, indicated on the right, are estimated by comparison with an internal RNA marker and reported gRNAs. (A) Intragenic oligo(U) RNAs (derived from the UC strain). (B) Intragenic oligo(U) RNAs (derived from the LEM 125 strain). (C) antisense oligo(U) RNAs and oligo(U) RNAs from the divergent region. (D–E) intragenic maxicircle-derived oligo(U) RNAs; oligo(U) RNAs as well as corresponding mRNAs from which they are derived from are simultaneously detected by probes directed against the oligo(U) RNA portion; oligo(U) RNAs are indicated by a single arrow (\blacktriangleright), corresponding mRNAs are indicated by a double arrow ($\blacktriangleright\blacktriangleright$); 9S rRNA was used as an internal control to assess for equal loading.

respect to the reported guide RNA gND7-II, whose expression level was arbitrarily assigned as 1 (Table 1). Accordingly, for the other two reported guide RNAs, gND7-I and gCyb-II, expression levels of 4-fold and 16-fold (compared to gND7-II), respectively, were determined (Figure 2, Table 1). By this method, we show that expression levels of novel gRNAs candidates are in the range of those determined for gND7-I and gND7-II, i.e. between 1-fold and 4-fold, respectively (Table 1). Hybridization signals of oU-RNAs appear as diffuse bands on autoradiograms of northern blots. This can be explained by the variable length of oligo(U) tails of oU-RNAs consistent with DNA sequences of the corresponding cDNA clones showing heterogeneous lengths of 3'-oligo(T) tails.

By northern blot analysis, sizes of oU-RNAs as well as canonical gRNAs range around 60 nt (Figure 2). By chemical and enzymatic probing it has previously been demonstrated that gRNAs, despite their

heterogeneous sequences, fold into a conserved secondary structure comprising two thermodynamically unstable stem-loops, essential for recognition by editing factors (26–28). Therefore, folding of gRNAs into this canonical secondary structure is restricted by their size.

Thus, in order to determine the exact size, including precise location of 5' and 3' termini, we performed 5' and 3' RACE assays for selected oU-RNAs and for the reported gND7-I guide RNA. The sequence of gND7-I has been previously published as a considerably shorter gRNA compared to other canonical guide RNAs (10); however, this finding is in contrast to the determined size of about 60 nt, as assessed by northern blot analysis. By 5' and 3' RACE assays, we confirmed the reported 5' terminus of gND7-I, while the 3' terminus showed an extension of additional 17 nt (Table S1). To assess, whether novel oU-RNAs are present in the same size range, as reported for canonical gRNAs, we selected some of oU-RNAs for 5' and 3' RACE analysis,

from which several independent cDNA clones (i.e. Ltmax1 and Ltmax5) and also from which single cDNA clone (i.e. Ltmax13-15, Ltmax17) were obtained. From these, Ltmax14 and Ltmax17 were cloned as RNA species significantly shorter than canonical gRNAs (Table 1, Figure 2). The sizes of obtained cDNA clone sequences for Ltmax1 and Ltmax5 are in the range of canonical gRNAs (i.e. about 60 nt); however, since they are derived from protein coding regions we wanted to analyze whether they would be part of longer, heterogeneous transcripts.

Contrary to this assumption, we confirmed the 5' and 3' termini of Ltmax1, Ltmax5 and Ltmax13 (as assessed by sequence analysis of the cDNA library), while from Ltmax14 no RACE data could be obtained, most likely due to the suboptimal low T_m of the oU-RNA specific PCR primer. The 5' terminus of Ltmax15 exhibited a 4 nt extension at its 5' end, while the determined 3' terminus was in agreement with previous cDNA sequencing data. Also, the 3' terminus of Ltmax17 oU-RNA was in agreement with the previous cDNA sequencing results, while the 5' terminus amplified by RACE assays exhibited a 24nt extension, in agreement with the northern blot analysis (Table 1). Thus, sizes of novel oU-RNAs determined by 5' and 3' RACE experiments appear to be in the same range of about 60 nt, as assessed for canonical gRNAs.

oU-RNAs derived from coding regions: sense orientation. We obtained 12 distinct oU-RNAs encoded within various protein coding genes and—in one case—within the 12S rRNA gene. The intragenic location of these RNAs is reminiscent of the previously identified independently transcribed intragenic gRNA gMURF2-II, which is entirely localized within the ND4 gene of *T. brucei* (19). We divided the novel maxicircle-encoded intragenic oU-RNAs into two groups: the first group contains oU-RNAs which are most likely processed from mRNAs; the second group comprises oU-RNAs, which are presumably independently transcribed and thus might resemble canonical gRNAs (see below).

The first group is represented by Ltmax1 to Ltmax4 oU-RNAs (Table 1). The 5'-ends of these RNAs precisely coincide with the mature 5'-ends of the corresponding mRNAs encoding ND7, Cytb, ATP6 and RPS12, respectively. The 3'-ends of Ltmax1 and Ltmax2 oU-RNAs map downstream to the pre-edited domain, at a location where the anchor domain of the cognate gRNA anneals to the mRNA (Figure 3A). Therefore, Ltmax1 and Ltmax2 oU-RNAs might be processed from respective mRNAs by endonuclease cleavage at the mismatched position within the gRNA/mRNA anchor duplex (Figure 3A). Since editing requires endonucleolytic cleavage of the pre-edited mRNA at the mismatch position upstream to the gRNA/mRNA anchor duplex, these results show that endonuclease cleavage also occurs within the anchor sequence.

The second group comprises intragenic oU-RNAs which in contrast to Ltmax1, 2, 3 and 4, described above, do not initiate at the same 5'-termini as the genes they are located in (Table 1, Figure 3B). Seven RNA species map to different positions within various mRNA

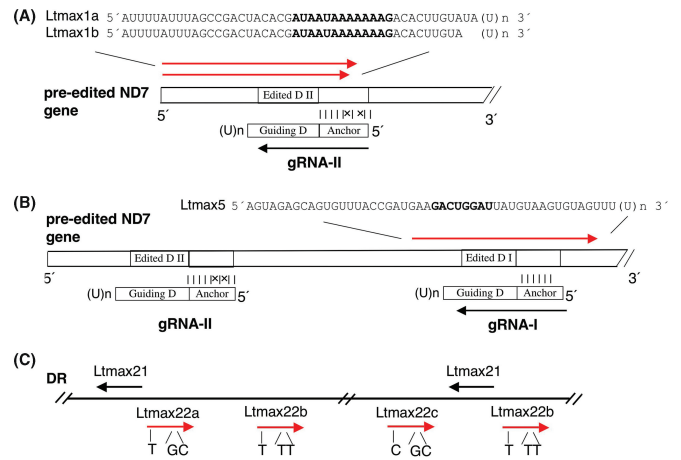


Figure 3. Location of oligo(U) RNAs within the maxicircle kinetoplast genome, (not drawn to scale). (A) Ltmax1 is indicated by a red arrow; the 5'-end of Ltmax1 coincides with the reported 5'-end of the ND7 gene while its 3'-end localizes downstream from the pre-edited domain at a mismatched position within the gRNA anchor domain. (B) Ltmax5 is indicated by a red arrow, flanking the pre-edited domain I of ND7 gene. (C) Ltmax21 loci indicated by black arrows and Ltmax21a/b/c loci varying in three nucleotides are indicated by red arrows.

genes (including edited and not edited genes) and one candidate is encoded within the 12S rRNA gene. All identified RNAs from this group were cloned as multiple cDNA clones with homogenous 5'-and 3'-termini. We suggest that these oU-RNA species might be independently transcribed from their respective protein coding genes they are located in. All, except for Ltmax11 and Ltmax12 oU-RNAs, possess a short, genome-encoded oligo(U)-track at their 3'-termini which serves as a termination signal for independently transcribed gRNAs (10). At this point it is difficult to assign any promoter elements to these RNAs, mostly for the lack of known regulatory sequence motifs within the kinetoplast genome. Sequence comparison of regions upstream to these ncRNAs did not reveal obvious consensus promoter elements.

In order to compare the expression levels of oU-RNAs with those of the overlapping mRNAs we performed northern blot analysis by agarose gel electrophoresis (in contrast to denaturing PAGE used previously, see materials and methods section) using oU-RNA-specific probes. Thereby, expression of both mRNAs and oU-RNAs can be monitored within the same experiment. As shown in Figures 2D and E, distinct hybridization signals of about 60 nt in size, corresponding to oU-RNAs or control gRNAs, are observed. Higher running bands (about 1200 nt in size) indicate expression levels of corresponding mRNAs or 12S rRNA (for Ltmax12), respectively. By phosphoimaging, we quantified hybridization signals of oU-RNAs versus corresponding mRNAs (data not shown). Thereby, we determined a higher expression level of Ltmax2 and Ltmax5, both derived from the pre-edited domains, compared to their corresponding mRNAs. This implicates a significant level of expression and/or stability of these RNAs (Figure 2D and E).

Interestingly, employing oligonucleotides directed against Ltmax1 and Ltmax5 oU-RNAs, which overlap two separate pre-edited domains of the ND7 gene (Figure 3A and B), we observed two significantly differing hybridization signals for ND7 mRNA (Figure 2D): the pre-edited domain I probe (corresponding to Ltmax1) exhibits an about 18-fold stronger hybridization signal than the oligonucleotide probe specific for pre-edited domain II (corresponding to Ltmax5), while expression of Ltmax1 and Ltmax5 oU-RNAs appear at about the similar levels (Figure 2D, Table 1). This might imply that ND7 mRNA is more efficiently edited at the domain II since it hybridizes less strongly to the—unedited—Ltmax5 oligonucleotide probe. This model is corroborated by northern blot analysis employing oligonucleotide probes directed against the fully edited domains I and II of the ND7 mRNA: as shown in Figure 2D, the hybridization signal specific for the edited domain I of the ND7 mRNA is detectable only after exposition of the blot for several days (data not shown), while the hybridization signal corresponding to edited domain II is readily visible after several hours of exposure.

Ltmax1 and Ltmax5 exhibit antisense elements to the anchor domain of two previously described guide RNAs gND7-I and gND7-II which mediate editing at two separate domains of ND7 gene: domain II and domain I, respectively (Figure 3A and B). In case of other edited mitochondrial genes, sequentially overlapping gRNAs give rise to fully edited mRNAs in 3' to 5' direction (29). Therefore, Ltmax1 and Ltmax5 might play a role in regulation of—sequential—editing of the two edited ND7 domains.

The ND7 gene codes for one of the subunits of complex I (designated as NADH:Ubiquinone oxidoreductase). It has been previously suggested that complex I is not required for growth in culture (30). Since all of the oU-RNAs from this study are isolated from the UC strain of *L. tarentolae*, which has been maintained in culture for an extended period of time, we verified expression of Ltmax5 and Ltmax1 in the recently isolated LEM125 strain. We could demonstrate similar expression levels for Ltmax5 and Ltmax1 in LEM strain125 to those observed in the UC strain (Figure 2B), indicating that these RNAs are stably expressed in both strains.

oU-RNAs derived from coding regions: antisense orientation. Seven oU-RNAs, Ltmax13 to Ltmax19, are transcribed in antisense orientation to protein-coding genes (Table 1). The abundant expression and/or stability of four ncRNA species was confirmed by northern blot analysis (Figure 2C). Transcripts in antisense orientation to protein-coding genes might be expected to be rapidly degraded. Stable expression of these candidates argues against the assumption that they merely represent processing or degradation products. The characteristic feature of these oU-RNAs is their antisense orientation to genes not known to be edited. Thus, they might influence expression of their complementary mRNAs by an antisense-like mechanism, rather than being involved in RNA editing.

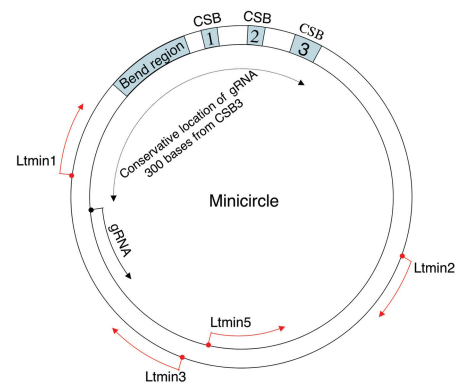


Figure 4. Organization of a canonical minicircle genome. Conserved sequence boxes (CSB) 1–3, bend region and conserved location of gRNA are indicated. Locations of identified oligo(U) RNAs Ltmin1–3 and Ltmin5 are indicated by red arrows, (not drawn to scale).

oU-RNAs derived from the divergent region. Five of the novel oU-RNAs, Ltmax20 to Ltmax24, map to the non-coding region of the maxicircle genome, containing variable numbers of species-specific tandem repeats, also termed ‘Divergent region’ (DR; Table 1). Ltmax20 and Ltmax21 map to the locus which is repeated while Ltmax22 to 24 are expressed from different, non-repeated regions within the DR. The repeated locus of Ltmax20 is homogenous in sequence, while the repeated locus of Ltmax21 differs in sequence by three nucleotides (Figure 3C). Interestingly, we obtained different isoforms from Ltmax21, designated as Ltmax21a/b/c indicating their expression from distinct loci. Although several independent cDNA clones from Ltmax20 to Ltmax24 could be identified in our screen, only Ltmax20 and 22 could be confirmed by northern blot analysis. Despite these oU-RNAs being expressed from intergenic loci, where most of the canonical annotated gRNAs have been mapped, we have been unable to identify their potential pre-mRNA targets. Previously, a single gRNA without a known cryptogene target, designated as gM150, has been reported (31). In our library, we identified two independent cDNA clones encoding gM150, with one cDNA clone corresponding to the annotated sequence, while a second cDNA clone was exhibiting a 22 nt long 3'-extension (Supplementary Data, Table S1). It has been suggested that gM150 might presumably be involved in misediting of RPS12 pre-mRNA (31). Thus, those intergenic oU-RNAs described in this study could be involved in a similar misediting process.

Novel minicircles-encoded oU-RNAs

Minicircle genomes of *L. tarentolae* have so far only been reported to encode one gRNA species each (21). All minicircles contain a conserved region and a variable region (20). The conserved region harbors two to three conserved sequence box motifs (CSB)—conserved in all trypanosomatids—in which CSB3 comprises the origin of replication for both strands (Figure 4) (11). The variable region encodes a single gRNA gene, located on the minus strand of the minicircle genome, which is localized

approximately 300 nt downstream from the CSB3 motif. This conserved location of gRNA genes was proposed to play an important role in expression of minicircle-encoded gRNAs (11).

In contrast, in our screen, we have identified stable minicircle-derived oU-RNAs, in the size range of canonical gRNAs, but mapping to locations in the minicircle genome that differ from those reported for canonical gRNA genes (Table 2). These include Ltmin1 to Ltmin4 which are encoded on the plus strand, while Ltmin5 is encoded on the minus strand of minicircles (Figure 4). Thereby, Ltmin1 is represented by four identical cDNA clones, while for other oU-RNAs from this class only single cDNA clones were obtained. The expression levels of Ltmin1 and Ltmin2 oU-RNAs were quantified by northern blot analysis and compared to the levels of known minicircle encoding gRNAs. Ltmin1, which is encoded on minicircle RPS12-Va and localized between the reported canonical gRNA-gRPS12-Va and CSB1 has about 50% lower expression level than the canonical gRPS12-Va (Figure 5). Ltmin2 gene, which maps about 300 nt upstream from CSB3 on minicircle gATP6-III, exhibits about 50% higher expression level compared to the canonical gA6-III RNA (Figure 5). Our data supports the view that the identified oU-RNAs represent novel RNA species independently transcribed from the reported canonical gRNAs.

Heterogeneous minicircle-derived oU-RNAs. In order to enrich our cDNA library for oligo-uridylylated ncRNAs, we additionally used a modified 3'-PCR-primer containing an oligo(U)-anchor for PCR amplification (see materials and methods section) and sequenced about 100 cDNA clones. By this approach we identified several of the oU-RNAs identified in this study as well as annotated gRNAs (Supplementary Data, Table S1). We also obtained several oU-RNA species exhibiting heterogeneous 5'- and 3'-ends mapping predominantly to the plus strand of minicircles within the variable and conserved regions (Table 3, Supplementary Data Table S2). Thereby, identified cDNA clones map to five different minicircles, encoding the following reported gRNAs: gND3-IX, gND9-VIIb, gRPS12-IIIa, gRPS12-IV and gATP6-III. Among the plus-strand encoded oU-RNAs, we also identified single cDNA clones located in antisense orientation to previously reported guide RNAs gRPS12-IIIa and gA6-IIIa, in analogy to Ltmax1 and Ltmax5 oU-RNAs, described above.

We also observed three annotated minicircle-derived gRNAs exhibiting heterogeneous sizes (for sequence information see: Supplementary Data, Table S1). Thereby, gG4-IIIa is represented by two cDNA clones, one of 50 nt in size while the second clone contains an additional 10 nt extension at its 3' end. While this gRNA has been described to be encoded as a longer transcript in the UC strain (30), our results indicate that instead it might exhibit a heterogeneous 3'-end due to imperfect 3'-end processing or synthesis. This gRNA is not thought to be functional since all of the other minicircle-encoded gRNAs in those editing cascades are missing from the UC strain. Two other gRNAs, gRPS12-Va and gA6-IIIa,

Table 2. Novel oligo(U) RNAs encoded on the minicircle component of the mitochondrial genome of *L. tarentolae*

Name	Sequence	cDNA clones	Northern blot	Location-minicircle
Plus strand encoded				
Ltmin1	AATGTCTGCCCA	4	+	RPS12-Va
	AGTGGCGTTTA			
	GGGTGCATACGT			
	GACAGAATTTAGA(T) ₁₂			
Ltmin2	CGTTGAGCATAG	1	+	A6-IIIa
	GCATTGGTAAT			
	AGGAGGCTATAAT			
Ltmin3	GAGCAATTATAA(T) ₁₄	1	+	RPS12-VIII
	ATAGGTAGAGGT			
	CCGAGAGGAT			
	TTGATCAGACATA			
Ltmin4	GTGGCGGTAA(T) ₁₀	1	-	A6-Ia
	AAAGTGTCCAGT			
	GGGGTTAAGA			
	GTGATAAAAATTGT			
Minus strand encoded				
Ltmin5	GTCATGAAAATT(T) ₅	1	+	RPS12-IIIa
	AACACTATGCTCTC			
	GAACAATTGT			
	AATATAAGTTATAA			
	TTAGTATT(T) ₄			

Name: designation of oligo(U) RNA species; cDNA clones: number of independent cDNA clones identified from each RNA species; Northern blot: the presence or absence of hybridizations signals in northern blot analysis is indicated by (+) or (-), respectively. Location minicircle: designation of minicircle to which the oligo(U) RNA is mapping; (T)_N: number of uridine residues not encoded by the mitochondrial genome.

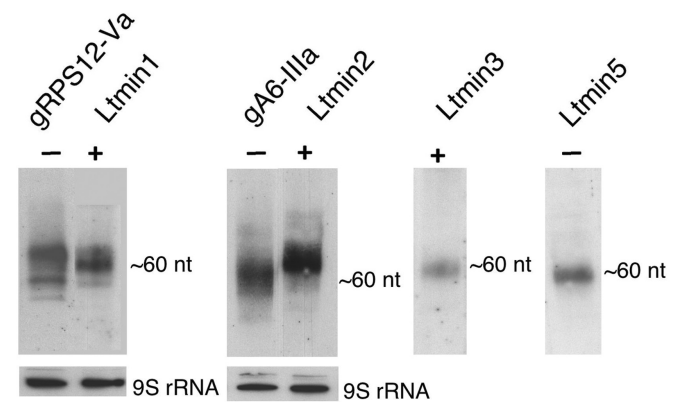


Figure 5. Northern blot analysis of novel minicircle-derived oligo(U) RNAs as well as known canonical guide RNAs, (+): plus strand encoded RNA (-): minus strand encoded RNA. 9S rRNA was used as an internal control to assess for equal loading; the size of each RNA species, indicated on the right, is estimated by comparison with an internal RNA marker.

were obtained with 5'-end extensions of 7 and 25 nt, respectively. This might suggest that transcription can initiate upstream of the proposed canonical transcription initiation start of a gRNA gene. This assumption was further corroborated by primer extension analysis for two annotated gRNAs and two novel oU-RNAs identified in this study, indicating transcription starts at sites upstream

Table 3. Heterogeneous oligo(U) RNAs mapping to various minicircles from *L. tarentolae*

Minicircle	Minicircle location	cDNA clones	Remarks
minicircle ND3-IX	195-616	15	variable region
minicircle RPS12-IIIa	669-729	2	variable region
	256-417	3	variable region
minicircle A6-IIIa	541-565	1	antisense to gRPS12-IIIa
	615-679	3	variable region
	86-35	2	variable region
minicircle ND9-VIIIb	541-614	1	antisense to gA6-IIIa
	848-786	4	conserved region
several minicircles	overlapping CSB1	2	conserved region

Minicircle location indicates the region to which map various heterogeneous oligo(U) RNAs, for sequences see Supplementary-Data Table S2; the location within minicircle region is indicated with respect to the location of conserved sequence box 3 (1-13) as a polarity marker; cDNA clones, number of cDNA clones representing independent RNA species.

from the reported mature 5'-ends of these RNA species (Figure 6).

CONCLUSIONS

In this study, we have identified a total of 29 novel oligo-uridylated RNA species (oU-RNAs) from the *L. tarentolae* kinetoplast genome, where 24 oU-RNAs are encoded on the maxicircle genome. We have also identified five novel minicircle-encoded gRNA-like RNAs, mapping either to the plus or minus strand of minicircles. So far, no transcription products have been reported from the plus strand of the minicircle genome. Thus, these RNA species provide the first evidence for transcription from both strands of the minicircle genome in *L. tarentolae*.

We note that several minicircle-encoded oU-RNA transcripts exhibit heterogeneous sizes according to their corresponding cDNA sequences. This is consistent with a model that transcription is not confined to a single transcription initiation site of a single gRNA transcription unit, as previously reported (21). The presence of these heterogeneous oU-RNA transcripts might be due to transcription from different promoters or processing from larger minicircle transcripts. Indeed, primer extension experiments indicate that transcription of gRNAs can initiate upstream from the annotated 5'-ends of a gRNA (Figure 6).

We were so far unable to identify any cryptogene mRNA targets for oU-RNAs from either maxi- or minicircles. Therefore, the biological role of these novel oligo-uridylated gRNA-like transcripts remains elusive at this point. It thus might be argued, that at least some of these oU-RNAs represent spurious transcription products rather than functional gRNA species. Although we cannot exclude this possibility, for thirteen maxicircle-derived

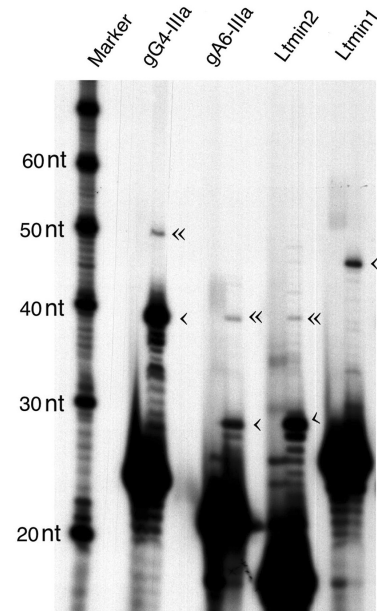


Figure 6. Primer extension analysis, assessing the 5'-ends of minicircle-encoded gRNAs: Major 5'-ends of the reported gG4-IIIa, gA6-IIIa as well as Ltmin1-2, identified in this screen are marked by an arrow (◁); 5'-extensions of reported gRNAs and Ltmin2 are indicated by a double arrow (◀). As a control, an equal amount of labeled primer, as used in primer extension reaction, was loaded next to extended product. Left: RNA size marker.

and four minicircle-derived oU-RNAs we could verify their expression by northern blot analysis. This indicates (a) a high expression level or (b) a significant stability of these RNA species. Thereby, expression levels and/or stability as well as the size (i.e. around 60 nt) of most novel oU-RNAs was found to be in the same range of reported, canonical gRNAs (see above).

Recently, it was demonstrated that in *T. brucei* gRNA-L07 can direct alternative editing of the COIII gene, leading to the generation of an alternatively edited mRNA which is translated into a novel protein (32). Therefore, alternative editing of a single gene can increase protein diversity in *T. brucei*. In *L. tarentolae*, mRNAs exhibiting alternative editing patterns were reported for several genes (e.g. COIII, RPS12 and ND3) (31,33). The alternative editing pattern of the COIII or ND3 (G5) gene was shown to be directed due to incorrect binding of certain gRNAs which normally direct proper editing of other cryptogenes. (31,33). Alternative editing of the RPS12 gene was proposed to be promoted by guide RNA gM150 which does not have any other known mRNA target (31). Thereby, misediting of RPS12 mRNA is proposed to be initiated by gM150, generating an anchor sequence for a putative second misguiding gRNA (31). Hence, novel gRNA candidates, identified by our screen, might be similarly involved in the generation of alternative editing patterns corresponding to alternative open reading frames in *L. tarentolae*.

Finally, our study suggests that, in addition to gRNAs and rRNAs, all RNA transcripts encoded by the

kinetoplast genome are prone to some degree of oligouridylation. Previously, it has been demonstrated, that terminal uridylyl transferase 1 (TUT1) exhibits no sequence specificity towards its targets (34,35). Thus, in *T. brucei* oligo(U) tails have also been observed at the 3'-ends of numerous unedited and partially edited pre-mRNAs (36). Interestingly, none of the tRNA species identified in our cDNA library exhibited oligo(U) tails at their 3'-ends. Conversely, oligo(U) tails of gRNAs are essential for the efficiency of the editing process (12,37). Since addition of uridine residues to 3'-ends of tRNAs, prior to aminoacylation, would probably abolish their function, mitochondrial tRNAs from kinetoplasts escape uridylation by a yet unknown mechanism. Since all tRNAs in *L. tarentolae*, as well as in other kinetoplastid protozoa, are encoded by the nuclear genome and subsequently imported into kinetoplast mitochondria the promiscuous TUT1 activity in kinetoplasts might have exerted evolutionary pressure on the kinetoplast genome to transfer all tRNA genes to the nucleus.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

This work was supported by an Austrian grant FWF 171370 and a German DFG grant 457-1/2 to A.H. and by an American Heart association grant and an MCB-0620707 grant from the National Science Foundation to J. D. A. Funding to pay the Open Access publication charge was provided by Austrian grant FWF 171370.

Conflict of interest statement. None declared.

REFERENCES

- Vickerman, K. (1976) *Biology of Kinetoplastida*. Academic press, London, England, Vol.1.
- Lipoldova, M. and Demant, P. (2006) Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. *Nat. Rev. Genet.*, **7**, 294–305.
- Kioy, D., Jannin, J. and Mattock, N. (2004) Human african trypanosomiasis. *Nat. Rev. Microbiol.*, **2**, 186–187.
- Dutra, W.O., Rocha, M.O. and Teixeira, M.M. (2005) The clinical immunology of human chagas disease. *Trends Parasitol.*, **21**, 581–587.
- Simpson, L. (1987) The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. *Annu. Rev. Microbiol.*, **41**, 363–382.
- Simpson, A.M., Suyama, Y., Dewes, H., Campbell, D.A. and Simpson, L. (1989) Kinetoplastid mitochondria contain functional tRNAs which are encoded in nuclear DNA and also contain small minicircle and maxicircle transcripts of unknown function. *Nucleic Acids Res.*, **17**, 5427–5445.
- Simpson, L., Sbicego, S. and Aphasizhev, R. (2003) Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA*, **9**, 265–276.
- Peris, M., Simpson, A.M., Grunstein, J., Liliental, J.E., Frech, G.C. and Simpson, L. (1997) Native gel analysis of ribonucleoprotein complexes from a *Leishmania tarentolae* mitochondrial extract. *Mol. Biochem. Parasitol.*, **85**, 9–24.
- Rusche, L.N., Cruz-Reyes, J., Piller, K.J. and Sollner-Webb, B. (1997) Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *EMBO J.*, **16**, 4069–4081.
- 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.
- 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.
- Blum, B. and Simpson, L. (1990) Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo(U) tail involved in recognition of the preedited region. *Cell*, **62**, 391–397.
- 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.
- Feagin, J.E., Jasmer, D.P. and Stuart, K. (1985) Apocytochrome b and other mitochondrial DNA sequences are differentially expressed during the life cycle of *Trypanosoma brucei*. *Nucleic Acids Res.*, **13**, 4577–4596.
- Read, L.K., Myler, P.J. and Stuart, K. (1992) Extensive editing of both processed and preprocessed maxicircle CR6 transcripts in *Trypanosoma brucei*. *J. Biol. Chem.*, **267**, 1123–1128.
- Koslowsky, D.J. and Yahampath, G. (1997) Mitochondrial mRNA 3' cleavage/polyadenylation and RNA editing in *Trypanosoma brucei* are independent events. *Mol. Biochem. Parasitol.*, **90**, 81–94.
- Gao, G., Kapushoc, S.T., Simpson, A.M., Thiemann, O.H. and Simpson, L. (2001) Guide RNAs of the recently isolated LEM125 strain of *Leishmania tarentolae*: an unexpected complexity. *RNA*, **7**, 1335–1347.
- Simpson, L., Wang, S.H., Thiemann, O.H., Alfonso, J.D., Maslov, D.A. and Avila, H.A. (1998) U-insertion/deletion edited sequence database. *Nucleic Acids Res.*, **26**, 170–176.
- Clement, S.L., Mingler, M.K. and Koslowsky, D.J. (2004) An intragenic guide RNA location suggests a complex mechanism for mitochondrial gene expression in *Trypanosoma brucei*. *Eukaryot. Cell*, **3**, 862–869.
- Simpson, L. (1997) The genomic organization of guide RNA genes in kinetoplastid protozoa: several conundrums and their solutions. *Mol. Biochem. Parasitol.*, **86**, 133–141.
- 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.
- Grams, J., McManus, M.T. and Hajduk, S.L. (2000) Processing of polycistronic guide RNAs is associated with RNA editing complexes in *Trypanosoma brucei*. *EMBO J.*, **19**, 5525–5532.
- Bralley, S.A. (1974) Biomaterials in clinical use. *Biomater. Med. Devices Artif. Organs*, **2**, 225–234.
- Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**, 156–159.
- Huttenhofer, A. and Vogel, J. (2006) Experimental approaches to identify non-coding RNAs. *Nucleic Acids Res.*, **34**, 635–646.
- Schmid, B., Riley, G.R., Stuart, K. and Goring, H.U. (1995) The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res.*, **23**, 3093–3102.
- Hermann, T., Schmid, B., Heumann, H. and Goring, H.U. (1997) A three-dimensional working model for a guide RNA from *Trypanosoma brucei*. *Nucleic Acids Res.*, **25**, 2311–2318.
- Schumacher, M.A., Karamooz, E., Zikova, A., Trantirek, L. and Lukes, J. (2006) Crystal structures of *T. brucei* MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell*, **126**, 701–711.
- Simpson, L., Thiemann, O.H., Savill, N.J., Alfonso, J.D. and Maslov, D.A. (2000) Evolution of RNA editing in trypanosome mitochondria. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 6986–6993.
- 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.

31. 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.
32. Ochsenreiter,T. and Hajduk,S.L. (2006) Alternative editing of cytochrome c oxidase III mRNA in trypanosome mitochondria generates protein diversity. *EMBO Rep.*, **7**, 1128–1133.
33. 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.
34. Aphasizheva,I., Aphasizhev,R. and Simpson,L. (2004) RNA-editing terminal uridylyl transferase 1: identification of functional domains by mutational analysis. *J. Biol. Chem.*, **279**, 24123–24130.
35. Aphasizhev,R. (2005) RNA uridylyltransferases. *Cell. Mol. Life. Sci.*, **62**, 2194–2203.
36. Decker,C.J. and Sollner-Webb,B. (1990) RNA editing involves indiscriminate U changes throughout precisely defined editing domains. *Cell*, **61**, 1001–1011.
37. Aphasizhev,R., Sbicego,S., Peris,M., Jang,S.H., Aphasizheva,I., Simpson,A.M., Rivlin,A. and Simpson,L. (2002) Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell*, **108**, 637–648.