



Published in final edited form as:

*Mol Cell*. 2016 February 4; 61(3): 364–378. doi:10.1016/j.molcel.2016.01.004.

## Antisense Transcripts Delimit Exonucleolytic Activity of the Mitochondrial 3' Processome to Generate Guide RNAs

Takuma Suematsu<sup>1, #</sup>, Liye Zhang<sup>2, #</sup>, Inna Aphasizheva<sup>1</sup>, Stefano Monti<sup>2</sup>, Lan Huang<sup>3</sup>, Qi Wang<sup>4</sup>, Catherine E. Costello<sup>4</sup>, and Ruslan Aphasizhev<sup>1, 4, \*</sup>

<sup>1</sup>Department of Molecular and Cell Biology, Boston University School of Dental Medicine, Boston, MA 02118, USA

<sup>2</sup>Section of Computational Biomedicine, Boston University School of Medicine, Boston, MA 02118, USA

<sup>3</sup>Department of Physiology & Biophysics, School of Medicine, University of California, Irvine, CA 92697, USA

<sup>4</sup>Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, USA

### SUMMARY

Small non-coding RNA biogenesis typically involves cleavage of structured precursor by RNase III-like endonucleases. However, guide RNAs that direct U-insertion/deletion mRNA editing in mitochondria of trypanosomes maintain 5' triphosphate characteristic of the transcription initiation and possess a U-tail indicative of 3' processing and uridylation. Here, we identified a protein complex composed of RET1 TUTase, DSS1 3'-5' exonuclease and three additional subunits. This complex, termed mitochondrial 3' processome (MPsome), is responsible for primary uridylation of ~800-nt gRNA precursors, their processive degradation to a mature size of 40–60 nucleotides, and secondary U-tail addition. Both strands of the gRNA gene are transcribed into sense and antisense precursors of similar lengths. Head-to-head hybridization of these transcripts blocks symmetrical 3'-5' degradation at a fixed distance from the double-stranded region. Together, our findings suggest a model in which gRNA is derived from the 5' extremity of a primary molecule by uridylation-induced, antisense transcription-controlled 3'-5' exonucleolytic degradation.

\*Correspondence: Ruslan Aphasizhev, Department of Molecular and Cell Biology, 72 E. Concord St., E424, Boston, MA 02118, USA. ruslana@bu.edu; Fax: 617 414 10-56; Phone: 617 414 1055.

<sup>#</sup>these authors contributed equally to this work

### ACCESSION NUMBERS

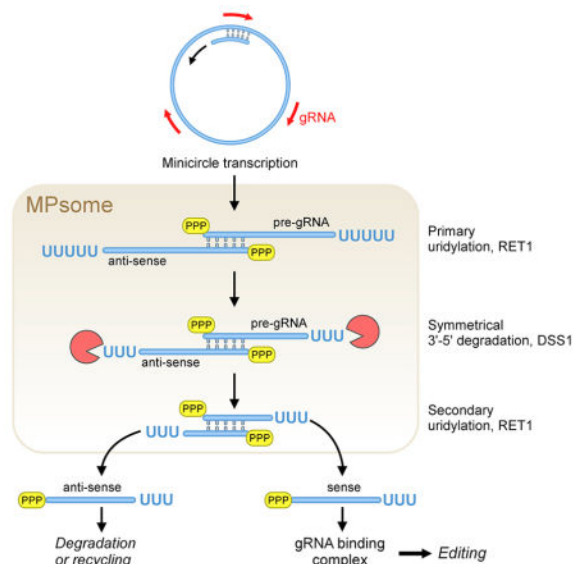
Protein-coding regions of *MPSS1*, *MPSS2* and *MPSS3* genes have been deposited in GenBank under accession codes KT282120, KT282121 and KT282122, respectively. Deep sequencing data have been deposited into Gene Expression Omnibus under accession code GSE71292.

### AUTHOR CONTRIBUTIONS

Conceptualization, RA; Methodology, TS, LZ and IA; Investigation, TS, IA, RA, LH, QW, CEC; Software: LZ, SM; Formal analysis: LZ; Writing Original Draft, RA; Writing Review and Editing, RA, TS, LZ, IA, CEC; Project Administration and Supervision: RA; Funding Acquisition, RA, CEC and LH.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Graphical Abstract



## Keywords

Trypanosoma; mitochondria; guide RNA; RNA editing; RNA decay; TUTase; exonuclease

## INTRODUCTION

*Trypanosoma brucei*, the causative agent of African sleeping sickness, is characterized by the presence of a “kinetoplast,” a disc-shaped, high-density nucleoprotein structure located in the mitochondrial lumen adjacent to the flagellar base. The kinetoplast encloses the mitochondrial genome (kDNA), which is composed of two types of catenated circles. Relatively few ~25-kb maxicircles contain genes typically found in mitochondrial genomes, such as rRNAs and subunits of respiratory complexes, while thousands of ~1-kb minicircles encode guide RNAs (gRNAs). These small non-coding RNAs direct U-insertion/deletion mRNA editing (Benne et al., 1986; Blum et al., 1990). The enzymatic cascade of mRNA cleavage, U-insertion or deletion, and ligation is catalyzed by the RNA editing core complex (RECC) (Aphasizhev et al., 2003a; Panigrahi et al., 2003). Each step of this process is directed by the secondary structure of an imperfect hybrid between pre-mRNA and gRNA (Blum et al., 1990; Kable et al., 1996; Seiwert et al., 1996). The 40–60 nucleotide (nt) gRNAs maintain 5' triphosphate typical of the transcription start site, and possess a U-tail indicative of 3' processing and uridylation. We have previously demonstrated that gRNAs are uridylated by RNA editing TUTase 1 (RET1) (Aphasizhev et al., 2002; Aphasizhev et al., 2003c).

Reminiscent of many small non-coding RNAs, gRNAs are synthesized as much longer precursor transcripts (PT). We and others have shown that the 5' end-defined, but 3' end-heterogeneous precursors span virtually the entire minicircle (Aphasizheva and Aphasizhev, 2010; Grams et al., 2000). Because each minicircle bears several gRNA genes, it follows

that the 800–1200 nt-long PT would be polycistronic and this raises the question of how such molecules are converted into mature 3' uridylated gRNAs. Although the mechanisms of small RNA processing vary among specific classes, an underlying principle invokes multiple cleavages of a structured precursor by RNase III-like enzymes (Ha and Kim, 2014; Winter et al., 2009). By the same token, an endonucleolytic model of gRNA biogenesis has been introduced along with identification of the trypanosomal mitochondrial RNase III-like endonuclease mRPN1 (Madina et al., 2011). However, the results of this study were later questioned (Carnes et al., 2015). We have previously observed that knockdown of RET1 TUTase triggers effective loss of mature gRNAs and accumulation of gRNA precursors (Aphasizheva and Aphasizhev, 2010). Conversely, the U-tail does not contribute to mature gRNA stability, which is chiefly determined by direct gRNA binding to the gRNA binding complex (Aphasizheva et al., 2014; Weng et al., 2008). These reports pose compound questions about the mechanism of nucleolytic processing and the role of uridylation in gRNA biogenesis. First, RET1 TUTase lacks activity that may account for RNA cleavage (Aphasizhev et al., 2002; Aphasizheva et al., 2004). Second, the dominant negative effect of a single amino acid mutation in RET1's active site argues against destabilization of potential protein complexes involving this TUTase (Aphasizheva and Aphasizhev, 2010). Third, RET1 is a multifunctional enzyme that also uridylates ribosomal RNAs (Adler et al., 1991) and mRNAs (Aphasizheva et al., 2011). Fourth, the endonucleolytic model leaves unclear the determinants of multicistronic PT cleavage, and particularly the instrument of precise 3' end definition.

To investigate the mechanism of guide RNA biogenesis, we focused on RET1 protein complexes and their RNA targets. We identified an ~ 900 kDa particle composed of RET1 TUTase and RNase II-like exonuclease DSS1, and three large proteins lacking annotated motifs. This complex, termed the mitochondrial 3' processome (MPsome), is responsible for initial uridylation of gRNA primary transcripts, which stimulates their processive 3'-5' degradation. Remarkably, the MPsome degrades all precursor-embedded gRNAs, but pauses before the most 5' unit. We show that the two strands of the gRNA gene are transcribed in opposite directions, giving rise to overlapping precursors of ~800 nt. It appears that hybridization between 5' regions of sense and antisense transcripts is responsible for the MPsome pausing at 10–12 nt from the double-stranded region, which defines the mature gRNA's 3' end prior to secondary uridylation. Together, our findings suggest an alternative model of small non-coding RNA biogenesis: in contrast to multiple endonucleolytic cleavages of a structured precursor by RNase III-like enzymes, guide RNA is derived from the 5' extremity of a primary molecule by uridylation-induced, antisense transcription-controlled 3'-5' exonucleolytic degradation.

## RESULTS

### TUTase, 3'-5' Exonuclease and Three Major Auxiliary Factors Constitute the Mitochondrial 3' Processome (MPsome)

RET1 silencing inhibits gRNA precursor processing even though this protein is devoid of nucleolytic activity (Aphasizheva and Aphasizhev, 2010). We hypothesized that TUTase may function as an essential component of a larger assembly and set out to investigate

RET1-associated complexes. Mouse monoclonal antibodies against recombinant RET1 were used to characterize (Figure 1A, B) and purify an ~900 kDa complex composed of five major polypeptides (Figure 1C). In addition to TUTase, DSS1 3'-5' exonuclease (Tb927.9.7210) was identified by mass spectrometry. DSS1, a member of the RNR superfamily (Figure S1A), was previously annotated by homology to a catalytic subunit of the yeast degradosome (Dziembowski et al., 1998; Mattiaccio and Read, 2008). Three large proteins lacking recognizable motifs, Tb927.11.9150, Tb927.10.9000 and Tb927.3.2770, were also detected with high confidence and named mitochondrial processome subunits MPSS1, 2 and 3, respectively (Table S1). Remarkably, the majority of RET1, DSS1 and non-enzymatic subunits MPSS1, 2 and 3 are sequestered into the ~900 kDa complex (Figures 1A and S1).

To confirm association with RET1, DSS1 and MPSS1, 2 and 3 were overexpressed as TAP-tagged proteins, purified from RNase-treated mitochondrial extracts and analyzed by mass spectrometry (Figure 1D, Table S1). A label-free quantitative strategy was applied to calculate the relative abundance of a given protein in each purification and to quantify interactions based on distributed normalized spectral abundance factor (dNSAF, (Fang et al., 2012; Zhang et al., 2010)). An interaction model predicts extensive contacts between RET1 and other complex components, and less pronounced binary interactions between DSS1, MPSS1 and MPSS2; MPSS3 appears to bind exclusively to RET1 (Figure 1E). To corroborate this model, affinity purified fractions (Figure 1D) were adjusted for bait proteins and probed with RET1 and DSS1 antibodies. Quantitative immunoblotting demonstrated that the relative amount of RET1 in MPSS3 TAP-purified fraction exceeds that of MPSS1 and MPSS2 by ~5-fold, and that of DSS1 by ~10-fold (Figure 1F). To verify TUTase's predicted protein contacts, RET1 was co-synthesized with MPsome subunits in a coupled transcription-translation reticulocyte system in the presence of [<sup>35</sup>S]-methionine and immunoprecipitated (Figure 1G). In agreement with mass spectrometry data, stable RET1-DSS1 and RET1-MPSS1 interactions were readily detected, while incorporation of MPSS2 required RET1-MPSS1 pre-assembly. The RET1- MPSS3 binding was also detected, but the close migration of two proteins in SDS PAGE precluded an unambiguous validation. To assess pairwise interactions further, TAP-tagged MPSS2 and MPSS3 were co-synthesized with enzymatic subunits (Figure 1H). It appears that MPSS3 indeed binds to RET1 only while MPSS2 engages in direct contact with DSS1. Collectively, these results demonstrate that the enzymes of seemingly competing activities, TUTase and potential 3'-5' exonuclease, are confined within an ~ 900 kDa protein complex, the mitochondrial 3' processome (MPsome).

### MPsome Subunits Are Essential for Mitochondrial RNA Processing and Parasite Viability

The potential roles of DSS1, RET1 and MPSS1-3 in mitochondrial RNA processing were examined by RNAi knockdowns and protein overexpression in procyclic (insect) actively respiring developmental form of *T. brucei* (Figures 2A and S2A, B). DSS1 expression caused a moderate phenotype while the D202A mutation (dominant negative, DN) in the metal binding site triggered a severe cell growth inhibition. Likewise, overexpression of active and inactive RET1 variants elicits moderate and severe cell growth inhibition, respectively, (Aphasizheva and Aphasizhev, 2010) while the impacts of MPSS1, 2 and 3

overexpression were minimal (Figure S2B). RNAi silencing of MPsome subunits produced cell growth defects ranging from severe (DSS1) to moderate (MPSS3). To assess whether RET1 downregulation compromises its binding partners and vice versa, we analyzed the relative abundance of RET1 and DSS1 in inducible RNAi cell lines for MPsome components (Figures 2B and S2C). Although efficient knockdowns of targeted nuclear mRNAs have been achieved (Figure 2C), it appears that the depletion of any single subunit does not elicit noticeable decline in abundances of enzymatic components. We next used quantitative RT-PCR to evaluate the effects of MPsome knockdowns on mitochondrial mRNAs and rRNAs (Figure 2C). The magnitude of observed changes in relative abundance varied, with DSS1 DN overexpression being the strongest and MPSS3 RNAi being the weakest, but the overall outcomes were similar to those obtained in RET1 silencing experiments (Aphasizheva and Aphasizhev, 2010). Specifically, we observed a significant downregulation of edited transcripts and moderate loss of 9S and 12S ribosomal RNAs, while the pre-edited and unedited mRNAs were either upregulated or remained steady. We conclude that the MPsome components function as a stable complex in a mitochondrial RNA processing pathway that is essential for the parasite's viability.

### MPsome Is Responsible for Guide RNA Processing

RET1 TUTase has been implicated in several 3' modification reactions, including uridylation of rRNAs (Aphasizheva and Aphasizhev, 2010), mature gRNAs (Aphasizhev et al., 2003c) and mRNAs (Aphasizheva et al., 2011; Ryan and Read, 2005). We noticed, however, that DSS1 knockdown most prominently affected edited mRNAs while exerting no appreciable impact on the RNA editing core (RECC) and gRNA binding (GRBC) complexes (Figure S2D). In addition, mass spectrometric analyses of editing complexes (Aphasizheva et al., 2014) and the MPsome (Table S1) indicate a lack of stable interaction between these pathways. Therefore, we hypothesized that the MPsome is essential for biogenesis of gRNAs required for RNA editing reactions and investigated gRNA processing defects in respective knockdowns. Selective labeling of 5' triphosphorylated RNAs (Blum and Simpson, 1990) was used to demonstrate gRNA loss and accumulation of heterogeneous 0.8–1.2 kb precursors upon repression of either RET1 or DSS1 (Figure 3A). This finding was corroborated by Northern blotting of individual gRNAs encoded, like most gRNAs, in the minicircle genome (gCO3[147]), or a single maxicircle-encoded gMurf2[II] (Figure 3B). Overexpression of active DSS1 and RET1 increased the yield of mature gRNAs, while inactivation of respective enzymes by a point mutation caused gRNA decline and precursor accumulation (Figure 3C, D and Table S2). To establish the global nature of compromised gRNA processing, total cellular RNA was subjected to RNA-Seq (pipeline in Figure 4A). In agreement with Northern blotting results, PTs that could be mapped to known minicircle sequences (Table S5) were upregulated in RET1 and DSS1 RNAi cells (Figure 3E). Repression of MPSS1 and MPSS2 produced results virtually indistinguishable from those of enzymatic components, while MPSS3 RNAi led to minor upregulation of gRNAs and their precursors (Figure 3F and Table S2). Finally, we inquired whether additional proteins abundantly present in DSS1 and MPSS1-3 affinity purified fractions are essential for gRNA processing and analyzed outcomes of RNAi knockdowns for potential candidates termed MPSS4-8 (Figure S3). Although repression of MPSS4-7 triggered severe cell growth inhibition, none of the knockdowns caused significant gRNA processing defects.

Collectively, these findings demonstrate that both expected enzymatic activities of the MPsome, 3' uridylation and 3'-5' RNA degradation, are essential for the gRNA precursor processing.

### Uridylation Stimulates Processive 3'-5' Degradation of Guide RNA Precursors

Inhibition of gRNA processing upon enzyme inactivation implicates DSS1 as the MPsome's hydrolytic subunit. The autonomous DSS1, however, is inactive ((Mattiaccio and Read, 2008) and this study), which suggests essential contributions of other MPsome components to the exonucleolytic activity. Although gRNA processing defects may be attributed to putative DSS1 3'-5' exonuclease, the precursor build up in RET1 DN cells ((Aphasizheva and Aphasizhev, 2010) and Figure 3D) could not be explained by perturbation of the complex. We reasoned that accumulation of gRNA precursors in RET1 DN cells may reflect the contribution of uridylation to effective substrate recognition by DSS1. To test this hypothesis, we performed a paired-end RNA-Seq on 600–1200 nt fraction of the total RNA and added a single-end sequencing reads from the same sample fragmented to 200–300 nt (Figure 4A and Tables S3, S4). In agreement with the overall increase in precursor abundance, the number of minicircles covered by PT reads increased 4- to 5-fold in DSS1 and RET1 knockdowns (Tables S3 and S4). Importantly, in the parental cell line, more than 30% of precursors were modified by addition of ~12 Us, while either RET1 or DSS1 depletion led to shortening of U-tails to 4–5 nucleotides (Table S4). A minor PT fraction contained random A/U tails and short A-stretches at the 3' end (Figure 4B). These results indicate that gRNA precursors are indeed uridylated and, combined with inhibition of gRNA processing by RET1 DN mutation, implicate this 3' modification in facilitating precursor recognition. Furthermore, the impairment of uridylation by DSS1 knockdown demonstrates that RET1 and DSS1 activities are interdependent within the MPsome. Thus, initial uridylation of long gRNA precursors by the MPsome-embedded RET1 appears to be essential for their efficient recognition and degradation by DSS1 within the same complex.

A dramatic difference between the lengths of gRNA precursor and mature gRNA, and the lack of apparent intermediates indicates that the MPsome catalyzes a processive 3'-5' degradation. To test this hypothesis *in vivo*, we performed UV-crosslinking immunoprecipitation with RET1 antibody followed by deep sequencing of partially digested RNAs (HITS-CLIP). As a positive control for mature gRNA binding, the GRBC1 subunit of the gRNA binding complex was affinity purified from UV-irradiated and mock-treated cells (Figure 4C). In addition, small mitochondrial RNAs from the parental cell line have been sequenced. Mapping RET1- and GRBC1-crosslinked RNA fragments along with small RNAs and their precursors accumulating in DSS1 DN and RET1 RNAi cells to a representative minicircle illustrated that both DNA strands are transcribed (Figure 4D). Importantly, CLIP-derived MPsome positioning along RNA precursors mirrored that of mature small RNAs locations in the minicircle. We conclude that the MPsome processively degrades precursor transcripts and pauses in the vicinity of mature small RNAs.

The steady-state level of small RNA is likely to be determined by several factors, including minicircle copy number, transcription rate, processing efficiency, function and decay. We next inquired whether the relative abundance of sense and antisense strands varies between



precursor transcripts and small RNAs, and how post-transcriptional events may contribute to these differences. To complement RET1 and GRBC CLIP and PT sequencing data, we performed rapid affinity pulldowns (RAP) with TAP-tagged GRBC1 and RET2, which represent core subunits of the gRNA binding (GRBC) and RNA editing enzymatic core (RECC) complexes, respectively. Associated RNAs were fragmented or size-selected to isolate small RNAs (30–75 nt) and sequenced (Table S6). Although this purification strategy does not depend on UV-crosslinking efficiency and is designed to detect unstable interactions (Aphasizheva et al., 2011), we were unable to identify precursors in either gRNA binding or core editing complexes while small RNAs from the same fractions produced consistent libraries. For each duplex, the log<sub>2</sub> ratios of read counts from sense and antisense strands were calculated to represent the relative abundance of two strands in precursors, small mitochondrial RNAs, and species identified in CLIP and RAP assays. The heat-maps ordered by hierarchical clustering at the assay and duplex levels highlight the two predominant outcomes of sense and antisense processing: in most cases, the relative abundance increases from PT to mature small RNAs (N=948), while in fewer examples (N=222) this trend is reversed (Figure 4E). In both cases RET1-CLIP reads clustered with precursor sequences, which supports RET1 association with precursors. However, the close distance between sequences derived from small RNA fraction and GRBC RAP suggests that the enrichment of one strand from the small RNA duplex is determined by its binding to the GRBC complex. Thus, it appears that the post-processing sequestration of small RNAs by the gRNA binding complex is largely responsible for their accumulation in the steady-state population.

### Antisense Transcription of gRNA Gene Defines Structure of the Processing Intermediate

In *T. brucei*, minicircles typically encode several gRNAs, which suggests a multicistronic nature of long precursors. Given that the vast majority of small mitochondrial RNAs bear 5' triphosphates (Aphasizheva et al., 2014), it is conceivable that only the 5' extremity of each PT matures into gRNA. These considerations raise a question of how the PT degradation is terminated at a fixed distance of 40–60 nt from the 5' end prior to U-tail addition. Based on our earlier discovery of a single antisense RNA for the gRNA involved in editing of ATPase subunit 6 mRNA (Aphasizheva and Aphasizhev, 2010) and the precise juxtaposition of small RNA reads mapped to both minicircle strands (Figure 4D), we hypothesized that a double-stranded region formed by the overlapping sense and antisense transcription units originating from the gRNA gene may impede the MPsome activity. We next analyzed the impacts of inhibited RET1 and DSS1 activities on gRNA that participates in editing of RPS12 mRNA, gRPS12[100], and its potential antisense counterpart. Remarkably, a synchronous loss of similarly-sized gRNA and anti-gRNA, and accumulation of their respective precursors were observed in both genetic backgrounds (Figure 5A, B). Accumulation of antisense precursors was also detected for gCO3[147] and gMurf2[II] guide RNAs investigated earlier (Figure 5C). Provided that radiolabeled probes were of similar specific activity and hybridization free energy, we note an approximately 10-fold lower relative abundance of antisense transcripts. We also note that precursor abundance typically exceeds that of corresponding gRNA (Figure 5A, B), indicating that some precursors may be degraded entirely without yielding mature molecules.

To investigate potential duplex formation on a transcriptome-wide scale, we performed RNA-Seq of 30–75 nt-long 5′ mitochondrial RNAs that were phosphorylated and treated sequentially with Terminator™ 5′-3′ exonuclease, and tobacco acid pyrophosphatase (Figure 4A). Resultant sequences were analyzed for their capacity for direct RNA editing and to form intermolecular duplexes. Guide RNAs, as defined by mapping to fully-edited mRNAs, amounted to 60,441 species among 133,801 unique small RNA assemblies (Table S7). Small RNAs that could not be mapped to edited mRNAs have been previously termed non-gRNAs, or ngRNA (Aphasizheva et al., 2014), and we will continue to use this definition. To predict potential duplexes, we performed pair-wise NCBI BLAST (megablast) and identified reverse complementary candidates with an E-value cutoff of  $1\text{E}^{-5}$  and maximum of 10 alignment outputs. Under these parameters, 1523 pairs have been identified in a small RNA population with sense and antisense lengths distributed as two broad peaks of 40–50 and 60–70 nucleotides (Figure 5D). Accordingly, the overall length of predicted hybrids fell into the 60–100 nt range (Figure 5E). Although conventional cloning and sequencing led to an assumption that mature gRNAs terminate with 10–15 Us (Thiemann and Simpson, 1996), we find that non- and monouridylated molecules constitute a significant fraction of small mitochondrial RNAs (Figure 5F). Accounting for gRNA representation, the hybrids can be classified into three classes: gRNA-gRNA (n=263), gRNA-ngRNA (n=966), and ngRNA-ngRNA (n=294). Between these duplex combinations, the lengths of double stranded regions were remarkably similar, although gRNAs in general tend to have longer median U-tails than ngRNAs (Figure 5G).

Mapping small RNAs to representative minicircles illuminated pervasive transcription of both DNA strands. Although the number of transcription units per minicircle varied dramatically, the consistent and precise juxtaposition of sense and antisense small RNAs, as defined by the polarity of the conserved sequence block 3 (CSB-3, minicircle origin of replication, (Ray, 1989)), provided further indication of their interdependent processing (Figure 5H). In agreement with Northern blotting, sense transcripts, which include some previously annotated gRNAs, appear to be present at higher levels than antisense RNAs. Because gRNAs of the same sequence or base-pairing capacity can be encoded in different minicircles (Kosłowsky et al., 1992; Savill and Higgs, 2000), we next inquired whether transcripts produced from diverse minicircles may form duplex structures. The putative processing intermediates were classified according to length of 5′ and 3′ overhangs into three classes with either short or long single-stranded regions on either strand (Figure 5I). The bi-directional transcription of gRNA gene would produce an intermediate depicted by Class I in which the 5′ ends are base-paired; the 1–2 nt “overhangs” are most likely caused by errors in the beginning of a sequencing run. The predicted structures of Class II and Class III hybrids suggest that regions of complementarity between small RNAs transcribed from distinct minicircles do exist. However, testing potential pairs by Northern blotting did not detect size differences between sense and antisense RNAs that would correspond to Class III (data not shown). Class II intermediates are also uncertain because the exonuclease would be expected to hydrolyze single-stranded RNA up to the double-stranded region, thereby generating a substrate that cannot be uridylated by RET1 TUTase (Aphasizhev et al., 2002; Aphasizheva et al., 2004). In conclusion, we identified numerous precisely initiated sense and antisense transcripts that may overlap to create a double-stranded region and



demonstrated that similarly-sized sense and antisense precursors undergo a symmetrical degradation to produce small RNAs.

### Antisense RNA-based Mechanism of the 3' End Definition

Outcomes of reverse genetic, deep sequencing and *in vivo* crosslinking studies indicate that the MPsome uridylylates and then processively degrades precursor RNAs before pausing at a fixed distance from the duplex formed by overlapping 5' extremities of sense and antisense transcripts. To directly test these mechanistic predictions, we purified TAP-tagged DSS1 along with DSS1 D202A and verified the composition of associated complexes by mass spectrometry (Table S1). For enzymatic assays, DSS1 variants were normalized by quantitative Western blotting (Figure S4A) and tested for activity with 5' radiolabeled synthetic gRNA (Figure 6A) and uniformly labeled 150 nt RNA (Figure S4B). Purified MPsome processively degraded single-stranded RNA to 4–5 nucleotides as the final product (FP) while the mutation in the catalytic metal binding site abolished the 3'-5' exonuclease activity. RNAs that terminate with 3' phosphate or 2' deoxyribose were degraded with similar efficiency (Figure S4C, D), as expected for the RNase II-mediated cleavage (Frazao et al., 2006). RNA degradation was observed at broad sub-millimolar concentrations of magnesium and manganese ions with an optimum at ~50  $\mu$ M, but zinc ions effectively inhibited the reaction at concentrations above 10  $\mu$ M (Figure 6B). The MPsome's TUTase activity resembled that of the recombinant RET1 in terms of optimal magnesium concentration, but showed much lower processivity: ~25 nucleotides at the maximum (Figure 6C) vs. hundreds of uridines, respectively (Figure S4E). Remarkably, at physiological concentration of  $Mg^{2+}$  ions (2–3 mM), the U-tailing activity of the MPsome-embedded RET1 TUTase resembles the *in vivo* gRNA uridylation pattern (Figure 5G).

To determine whether initial uridylation facilitates precursor recognition by the MPsome, we tested 3'-5' exonuclease activity on a synthetic 800-nt gCO3[147] precursor that terminated with either encoded sequence or additional 15 uridines (Figure 6D). Although both substrates were degraded with equally high processivity, the 3' U-tail provided for more efficient substrate recognition, as reflected by accumulation of final product and decline of the input RNA. The MPsome's remarkable processivity on a structured gRNA precursor suggested that the chemical energy of RNA hydrolysis may be used to unwind secondary structures (Lee et al., 2012). We next assembled 152-nt pre-gCO3[147] and 123-nt antisense molecule to create a 25-bp double-stranded region with 127-nt overhang on the labeled sense strand, and analyzed the exonucleolytic and unwinding activities of MPsome-embedded DSS1. As shown in Figure 6E, the MPsome indeed possesses dsRNA unwinding activity, which strictly depends on RNA hydrolysis and can be likewise modulated by divalent ions (Figure S4F). It follows that RNA hydrolysis must precede the unwinding; hence, we reasoned that the MPsome would require a single-stranded RNA of a certain length to initiate the reaction and to accumulate the energy required for strand separation. Experiments with three RNA hybrids containing identical double-stranded regions and various 3' overhangs demonstrated that substrates bearing less than a 10-nt single-stranded region are refractory to digestion and unwinding (Figure 6F). Apparently, the full activity requires 10–20 nt overhang; incidentally a median overhang of ~12 nt (Figure 5G) is consistent with an optimal length of RNA substrate for RET1 TUTase, which would be

expected to uridylate the trimmed 3' end. The lack of catalytic activity and discernible RNA binding motifs in DSS1 protein (Figure S1A) indicate that substrate recognition is delegated to other subunits. Therefore, the hydrolytic, unwinding and constrained uridylation activities ought to be considered as properties of the assembled MPsome.

Although a preference for uridylated substrates and the capacity to degrade long precursors explain phenotypic manifestations of DSS1 and RET1 knockdowns, generation of a defined 3' end would require the MPsome to pause or abort at a certain point. In light of its prominent strand separation activity, we next inquired whether properties of the double-stranded region formed by sense-antisense transcripts can modulate the degradation activity. To design model substrates, we selected a previously characterized gCO3[147] guide RNA and truncated the minicircle sequence-deduced precursor to represent the first 150 nt, and added a 15-nt U-tail. Two antisense molecules, 35 and 55 nt, were synthesized to resemble three and five turns of RNA helix, respectively. In addition, we applied data-driven unbiased selection of abundant RNA duplexes that map to the same region (positions 560–600) of the edited CO3 mRNA and then queried the overlapping sequences for the most abundant sense/antisense pair. The selected 92 nt-long gRNA precursor, designated gCO3[147\*], forms a 31-bp double-stranded region with respective antisense molecule, which falls within the median for the predicted duplex regions (Figure 5G). As expected, a single-stranded RNA was effectively degraded to 4–5 nt. However, in the presence of antisense RNAs, we observed MPsome pausing within 10–12 nt from the double-stranded region, as demonstrated by electrophoresis under denaturing conditions (Figure 6G). To verify that the duplex with trimmed 3' overhang was indeed produced, the reaction products were separated on native gel under conditions that retain trimmed duplexes and degraded RNA (Figure 6H). Interestingly, there appears to be no direct correlation between the calculated free energy of duplex formation and the efficiency of pausing: the more stable 35-bp duplex involving gCO3[147] ( $\Delta G = -58$  kcal/mol) was marginally less efficient than the data-supported gCO3[147\*] pair characterized by a free energy of  $-48$  kcal/mol. Nonetheless, the pausing threshold does exist as the shorter hybrid with  $\Delta G$  of  $-28$  kcal/mol (Figure 6F) was fully degraded. Conversely, extending the double-stranded region beyond the threshold, i.e., 55-bp duplex, provided no additional gain. Indeed, distribution of  $\Delta G$  frequency shows that 90% of predicted sense-antisense duplexes fall into the  $-(30-60)$  kcal/mol range ((Figure S4H, (Freier et al., 1986)). Finally, we inquired whether a chemical moiety at the 5' end of the antisense molecule may influence the pausing efficiency. However, including the same RNA with 5' hydroxyl, monophosphate or triphosphate produced no appreciable differences (Figure 6I).

Under all conditions tested, a significant fraction of input RNA was degraded to a few nucleotides indicating a stochastic nature of MPsome pausing, which also explains why the steady-state levels of mature small RNAs are typically lower than those of precursors (Figures 4 and 5). Although participation of other factors in more efficient MPsome pausing cannot be excluded, we have not observed such effects by including potentially relevant proteins. Examples include complexes implicated in gRNA binding, such as GRBC1/2 (Figure S5) and gRNA-mRNA annealing, such as MRP1/2 ((Aphasizhev et al., 2003b; Muller et al., 2001), Figure S6).

To test *in vivo* whether the secondary uridylation of a properly trimmed gRNA is also accomplished by the MPsome, we sequenced small RNAs and analyzed their uridylation patterns in RET1 RNAi and DSS1 RNAi backgrounds. Controls included GRBC1/2 dual RNAi, which causes downregulation of mature gRNAs, and knockdown of the gRNA binding complex subunit 4 (GRBC4), which triggers gRNA accumulation. In addition, we analyzed small RNAs that accumulate in dual knockdown of structural subunits MP18/24 of the RNA editing core complex (Aphasizheva et al., 2014). In a remarkable parallel to gRNA precursors (Table S4), the small RNA sequencing demonstrated that the U-tails are shortened in both RET1 TUTase RNAi and in DSS1 RNAi, but remain unaffected in other backgrounds (Table S8). The dependence of the RET1 activity on the presence of DSS1 protein demonstrates the MPsome's participation in primary uridylation of gRNA precursors and in secondary uridylation of trimmed gRNAs. Overall, our data provide a strong support for the model of gRNA maturation that invokes antisense RNA-mediated pausing of the uridylation-stimulated 3'-5' exonucleolytic degradation as the means of defining the gRNA 3' end. To our knowledge, this pathway represents a fundamentally different mechanism of small non-coding RNA biogenesis (Figure 7).

## DISCUSSION

The discovery of guide RNA led to the cascade model of U-insertion/deletion mRNA editing, which designated small RNAs as instruments of targeting enzymatic complexes to specific RNA sequences (Blum et al., 1990). Although superficially similar to the RNAi and CRISPR-Cas systems that evolved to protect genomes by destroying nucleic acid targets, RNA editing appears to have molded ancient catalytic modules into unique protein assemblies capable of changing RNA sequence. Intrinsically mutagenic, trypanosomal editing also represents a unique mechanism of small RNA-mediated information transfer in which local secondary structure of gRNA-mRNA hybrid dictates the positioning and extent of editing events. Whereas biogenesis of siRNA, miRNA and most other small RNAs invariably involves RNase-III mediated cleavage of a structured precursor, the gRNA's unique structural features (triphosphorylated 5' and uridylated 3' ends) suggested that guide RNAs may be produced by a distinctive mechanism. In this work we show that RET1, initially implicated in gRNA uridylation (Aphasizhev et al., 2002; Aphasizhev et al., 2003c; Ernst et al., 2003), functions as a subunit of the mitochondrial 3' processome (MPsome). This stable protein complex is composed of RET1, DSS1 3'-5' exonuclease and three large subunits, MPSS1-3. Sequestering the enzymes with seemingly opposing transferase and hydrolase activities into the MPsome represents an ultimate example of a functional coupling between TUTase and RNase-II-like exonuclease, an evolutionary conserved RNA decay pathway (Chang et al., 2013; Lim et al., 2014; Malecki et al., 2013), which apparently also functions in small RNA processing.

We provide evidence that the MPsome's TUTase activity initially uridylates gRNA precursors to facilitate their recognition by DSS1 and to induce processive 3'-5' degradation. The hydrolytic reaction provides chemical energy for propelling the MPsome and unwinding relatively stable secondary structures along the 800–1200 nt substrates. The MPsome pausing in the vicinity of gRNA's 3' terminus prior to the secondary uridylation event defines the length of the mature molecule and is the key feature of the proposed model

(Figure 7). To elucidate the mechanism of pausing, we investigated transcriptional profile of the minicircle genome and identified gRNA-sized transcripts that map as antisense to gRNA genes. Remarkably, antisense transcripts are also transcribed as precursors, similar to those of gRNAs in length and behavior upon MPsome inactivation. These findings led to an understanding that both strands of gRNA gene are transcribed, giving rise to long symmetric precursors that overlap with their 5' regions. Such precursors may be considered multicistronic because they cover the entire minicircle and most minicircles encode more than one gRNA. However, our data indicate that only the 5' extremity of each precursor is processed into gRNA or gRNA-like molecule. Further, the predicted duplex formed by overlapping 5' ends of sense and antisense molecules emerged as the main factor that delimits precursor degradation by the MPsome. To that extent, most of the inferences derived from proteomics, reverse genetics, deep sequencing and *in vivo* binding analyses are consistent with properties of the purified MPsome acting on model RNA substrates. Specifically, the MPsome possesses both uridylation and exonuclease activities, with the former resembling gRNA uridylation and the latter showing near absolute processivity. Most importantly, we demonstrate a stochastic pausing of the purified MPsome at 10–12 nucleotides downstream from the double-stranded region formed by overlapping transcripts. The MPsome's RNA hydrolysis-driven unwinding activity appears to be finely-tuned to travel through a long precursor, but pause at a fixed distance from sufficiently stable duplex. The length of the remaining 3' overhang (10–12 nt) likely reflects the protein complex footprint and constitutes a single-stranded RNA of necessary length to be recognized by RET1 TUTase. It seems plausible that MPsome pausing enables RET1 to compete for RNA substrate and to add the U-tail, thereby disengaging the degradation machinery.

Following this scenario, the duplex in the processing intermediate covers the gRNA's "anchor" region required for hybridization with mRNA target during editing; hence, strand separation would be required to liberate functional single-stranded gRNA. Although this step needs further investigation, the gRNA release may be accomplished by an asymmetric degradation of antisense RNA as we consistently observed lower levels of the latter. The gRNA binding complex is emerging as downstream vehicle for binding, stabilizing and delivering the single-stranded gRNA into the editing cascade ((Aphasizheva et al., 2014; Weng et al., 2008).

In *T. brucei*, all but one of the gRNAs are encoded in minicircles, while in related Trypanosomatids, such as *L. tarentolae*, many more are borne by the maxicircle genome (Sturm and Simpson, 1991). However, gRNA transcription and maturation patterns appear to be similar irrespective of the genomic context, and distinct from those of pre-mRNAs and pre-rRNAs (Aphasizheva and Aphasizhev, 2010; Clement et al., 2004). Here we show that transcription generates precursors to functional gRNAs and *trans*-acting antisense RNAs that delineate the gRNA precursor processing. Although transcription start sites can be readily inferred from mapping triphosphorylated RNAs to known minicircles, potential promoters with a reasonable degree of conservation are yet to be identified. Overall, our findings illuminate the guide RNA processing pathway in mitochondria of trypanosomes and introduce a distinctive mechanism of small non-coding RNA biogenesis.

## EXPERIMENTAL PROCEDURES

### Trypanosome Culture, RNAi and Protein Expression

The RNAi plasmids were generated by cloning ~500-bp gene fragments into p2T7-177 vector for tetracycline-inducible expression (Wickstead et al., 2002). The constructs were transfected into procyclic 29–13 *T. brucei* strain (Wirtz et al., 1999). RNAi was performed as described (Weng et al., 2008). For inducible protein expression in PF *T. brucei*, full-length genes were cloned into pLEW-MHTAP vector (Jensen et al., 2007).

### Mitochondrial Isolation, Glycerol Gradients and Affinity Purification

Mitochondrial fraction was isolated as described except omitting the Percoll density gradient (Pelletier et al., 2007). Mitochondrial pellets were extracted into 25 mM HEPES (pH 7.6), 125 mM KCl, 12 mM MgCl<sub>2</sub> and 1.2 % Nonidet P-40 (NP-40) and fractionated on 10%–30% glycerol gradient. The conventional TAP procedure was performed as described (Aphasizhev and Aphasizheva, 2007). For the rapid pulldown, rabbit IgG was coupled to Dynabeads M-270 Epoxy (Invitrogen) according to (Oeffinger et al., 2007) and used in total cell extract.

### RNA Analysis

The change in relative abundance was calculated based on qRT-PCR or Northern blotting data assuming the ratio between analyzed transcripts and control RNAs in mock-induced cells as 1 or 100%, respectively. Membranes and gels were exposed to phosphor storage screens and analyzed with ImageQuant software (GE Healthcare).

### Enzymatic Assays

MPsome activity assays were carried out at 27°C in 40 µl reaction containing 50 mM Tris-HCl (pH 8.0), 0.05% NP-40, 1 mM DTT, 2 units of RNaseOut ribonuclease inhibitor, 1 mM of MgCl<sub>2</sub>, 2 nM of DSS1 from TAP-purified fractions, and 10 nM of labeled ssRNA or dsRNA.

See Supplemental Experimental Procedures for detailed protocols.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

We thank members of our laboratories, Dmitri Maslov and Larry Simpson for discussions, advice and technical assistance. We also thank Maria Schumacher for providing MRP1/2 co-expression plasmid. This work was supported by NIH grants AI091914 and AI101057 to RA, P41 GM104603 to CEC and GM074830 to LH.

### References

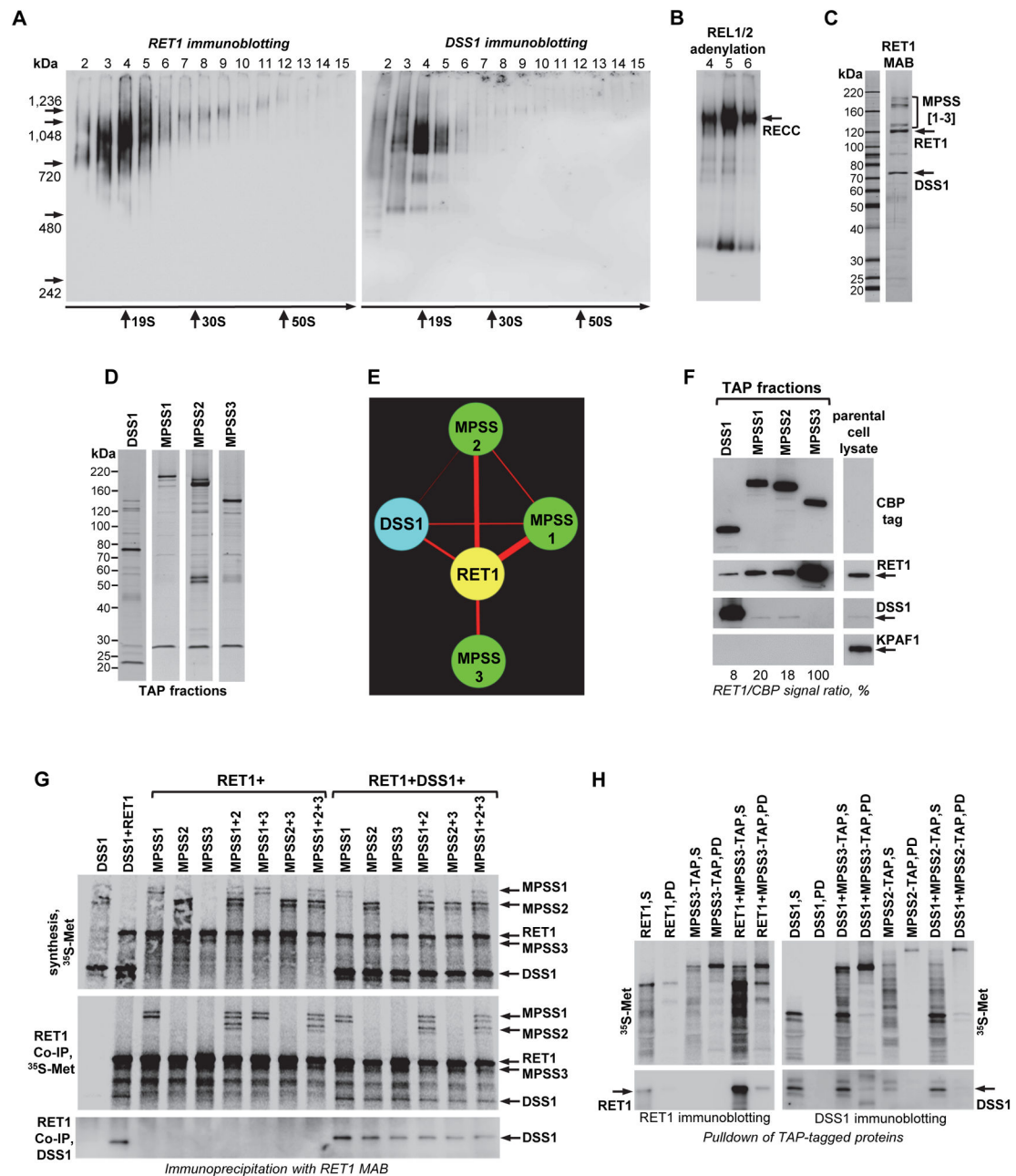
Adler BK, Harris ME, Bertrand KI, Hajduk SL. Modification of *Trypanosoma brucei* mitochondrial rRNA by posttranscriptional 3' polyuridine tail formation. *Mol Cell Biol.* 1991; 11:5878–5884. [PubMed: 1719373]

- Aphasizhev R, Aphasizheva I. RNA Editing Uridylyltransferases of Trypanosomatids. *Methods Enzymol.* 2007; 424:51–67.
- Aphasizhev R, Aphasizheva I, Nelson RE, Gao G, Simpson AM, Kang X, Falick AM, Sbicego S, Simpson L. Isolation of a U-insertion/deletion editing complex from *Leishmania tarentolae* mitochondria. *EMBO J.* 2003a; 22:913–924. [PubMed: 12574127]
- Aphasizhev R, Aphasizheva I, Nelson RE, Simpson L. A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA.* 2003b; 9:62–76. [PubMed: 12554877]
- Aphasizhev R, Aphasizheva I, Simpson L. A tale of two TUTases. *Proc Natl Acad Sci U S A.* 2003c; 100:10617–10622. [PubMed: 12954983]
- Aphasizhev R, Sbicego S, Peris M, Jang SH, Aphasizheva I, Simpson AM, Rivlin A, Simpson L. Trypanosome Mitochondrial 3' Terminal Uridylyl Transferase (TUTase): The Key Enzyme in U-insertion/deletion RNA Editing. *Cell.* 2002; 108:637–648. [PubMed: 11893335]
- Aphasizheva I, Aphasizhev R. RET1-catalyzed Uridylylation Shapes the Mitochondrial Transcriptome in *Trypanosoma brucei*. *Mol Cell Biol.* 2010; 30:1555–1567. [PubMed: 20086102]
- Aphasizheva I, Aphasizhev R, Simpson L. RNA-editing terminal uridylyl transferase 1: identification of functional domains by mutational analysis. *J Biol Chem.* 2004; 279:24123–24130. [PubMed: 15060068]
- Aphasizheva I, Maslov D, Wang X, Huang L, Aphasizhev R. Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes. *Mol Cell.* 2011; 42:106–117. [PubMed: 21474072]
- Aphasizheva I, Zhang L, Wang X, Kaake RM, Huang L, Monti S, Aphasizhev R. RNA binding and core complexes constitute the u-insertion/deletion editosome. *Mol Cell Biol.* 2014; 34:4329–4342. [PubMed: 25225332]
- Benne R, Van den Burg J, Brakenhoff J, Sloof P, Van Boom J, Tromp M. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell.* 1986; 46:819–826. [PubMed: 3019552]
- Blum B, Bakalara N, Simpson L. A model for RNA editing in kinetoplastid mitochondria: “Guide” RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell.* 1990; 60:189–198. [PubMed: 1688737]
- Blum B, Simpson L. Guide RNAs in kinetoplastid mitochondria have a nonencoded 3' oligo-(U) tail involved in recognition of the pre-edited region. *Cell.* 1990; 62:391–397. [PubMed: 1695552]
- Carnes J, Lerch M, Kurtz I, Stuart K. Bloodstream form *Trypanosoma brucei* do not require mRPN1 for gRNA processing. *RNA.* 2015; 21:28–35. [PubMed: 25404564]
- Chang HM, Triboulet R, Thornton JE, Gregory RI. A role for the Perlman syndrome exonuclease Dis3L2 in the Lin28-let-7 pathway. *Nature.* 2013; 497:244–248. [PubMed: 23594738]
- Clement SL, Mingler MK, Koslowsky DJ. An intragenic guide RNA location suggests a complex mechanism for mitochondrial gene expression in *Trypanosoma brucei*. *Eukaryot Cell.* 2004; 3:862–869. [PubMed: 15302819]
- Dziembowski A, Malewicz M, Minczuk M, Golik P, Dmochowska A, Stepień PP. The yeast nuclear gene DSS1, which codes for a putative RNase II, is necessary for the function of the mitochondrial degradosome in processing and turnover of RNA. *Mol Gen Genet.* 1998; 260:108–114. [PubMed: 9829834]
- Ernst NL, Panicucci B, Igo RP Jr, Panigrahi AK, Salavati R, Stuart K. TbMP57 is a 3' terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol Cell.* 2003; 11:1525–1536. [PubMed: 12820966]
- Fang L, Kaake RM, Patel VR, Yang Y, Baldi P, Huang L. Mapping the protein interaction network of the human COP9 signalosome complex using a label-free QTAX strategy. *Mol Cell Proteomics.* 2012; 11:138–147. [PubMed: 22474085]
- Frazao C, McVey CE, Amblar M, Barbas A, Vonnrhein C, Arraiano CM, Carrondo MA. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature.* 2006; 443:110–114. [PubMed: 16957732]



- Freier S, Kierzek R, Jaeger J, Sugimoto N, Caruthers M, Neilson T, Turner D. Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA*. 1986; 83:9373–9377. [PubMed: 2432595]
- Grams J, McManus MT, Hajduk SL. Processing of polycistronic guide RNAs is associated with RNA editing complexes in *Trypanosoma brucei*. *EMBO J*. 2000; 19:5525–5532. [PubMed: 11032819]
- Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol*. 2014; 15:509–524. [PubMed: 25027649]
- Jensen BC, Kifer CT, Brekken DL, Randall AC, Wang Q, Drees BL, Parsons M. Characterization of protein kinase CK2 from *Trypanosoma brucei*. *Mol Biochem Parasitol*. 2007; 151:28–40. [PubMed: 17097160]
- Kable ML, Seiwert SD, Heidmann S, Stuart K. RNA editing: a mechanism for gRNA-specified uridylylation insertion into precursor mRNA [see comments]. *Science*. 1996; 273:1189–1195. [PubMed: 8703045]
- Koslowsky DJ, Riley GR, Feagin JE, Stuart K. Guide RNAs for transcripts with developmentally regulated RNA editing are present in both life cycle stages of *Trypanosoma brucei*. *Mol Cell Biol*. 1992; 12:2043–2049. [PubMed: 1373804]
- Lee G, Bratkowski MA, Ding F, Ke A, Ha T. Elastic coupling between RNA degradation and unwinding by an exoribonuclease. *Science*. 2012; 336:1726–1729. [PubMed: 22745434]
- Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, Kim VN. Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. *Cell*. 2014; 159:1365–1376. [PubMed: 25480299]
- Madina BR, Kuppan G, Vashisht AA, Liang YH, Downey KM, Wohlschlegel JA, Ji X, Sze SH, Sacchettini JC, Read LK, Cruz-Reyes J. Guide RNA biogenesis involves a novel RNase III family endoribonuclease in *Trypanosoma brucei*. *RNA*. 2011; 17:1821–1830. [PubMed: 21810935]
- Malecki M, Viegas SC, Carneiro T, Golik P, Dressaire C, Ferreira MG, Arraiano CM. The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J*. 2013; 32:1842–1854. [PubMed: 23503588]
- Mattiacio JL, Read LK. Roles for TbDSS-1 in RNA surveillance and decay of maturation by-products from the 12S rRNA locus. *Nucleic Acids Res*. 2008; 36:319–329. [PubMed: 18032430]
- Muller UF, Lambert L, Goring HU. Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *EMBO J*. 2001; 20:1394–1404. [PubMed: 11250905]
- Oeffinger M, Wei KE, Rogers R, Degrasse JA, Chait BT, Aitchison JD, Rout MP. Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat Methods*. 2007; 4:951–956. [PubMed: 17922018]
- Panigrahi AK, Schnauffer A, Ernst NL, Wang B, Carmean N, Salavati R, Stuart K. Identification of novel components of *Trypanosoma brucei* editosomes. *RNA*. 2003; 9:484–492. [PubMed: 12649499]
- Pelletier M, Read LK, Aphasizhev R. Isolation of RNA Binding Proteins Involved in Insertion/deletion Editing. *Methods Enzymol*. 2007; 424:69–96.
- Ray D. Conserved sequence blocks in kinetoplast DNA minicircles from diverse species of trypanosomes. *Mol Cell Biol*. 1989; 9:1365–1367. [PubMed: 2542768]
- Ryan CM, Read LK. UTP-dependent turnover of *Trypanosoma brucei* mitochondrial mRNA requires UTP polymerization and involves the RET1 TUTase. *RNA*. 2005; 11:763–773. [PubMed: 15811918]
- Savill NJ, Higgs PG. Redundant and non-functional guide RNA genes in *Trypanosoma brucei* are a consequence of multiple genes per minicircle. *Gene*. 2000; 256:245–252. [PubMed: 11054554]
- Seiwert SD, Heidmann S, Stuart K. Direct visualization of uridylylation in vitro suggests a mechanism for kinetoplastid RNA editing. *Cell*. 1996; 84:831–841. [PubMed: 8601307]
- Sturm NR, Simpson L. *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*. 1991; 19:6277–6281. [PubMed: 1720240]
- Thiemann OH, Simpson L. Analysis of the 3' uridylylation sites of guide RNAs from *Leishmania tarentolae*. *Mol Biochem Parasitol*. 1996; 79:229–234. [PubMed: 8855560]

- Weng J, Aphasizheva I, Etheridge RD, Huang L, Wang X, Falick AM, Aphasizhev R. Guide RNA-Binding Complex from Mitochondria of Trypanosomatids. *Mol Cell*. 2008; 32:198–209. [PubMed: 18951088]
- Wickstead B, Ersfeld K, Gull K. Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. *Mol Biochem Parasitol*. 2002; 125:211–216. [PubMed: 12467990]
- Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol*. 2009; 11:228–234. [PubMed: 19255566]
- Wirtz E, Leal S, Ochatt C, Cross GA. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol*. 1999; 99:89–101. [PubMed: 10215027]
- Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem*. 2010; 82:2272–2281. [PubMed: 20166708]



**Figure 1. Identification and Characterization of the Mitochondrial 3' Processome (MPsome)**  
 (A) Mitochondrial fraction from the insect stage of *T. brucei* was extracted with detergent and soluble contents were separated for 5 hours at 178,000 g in 10%–30% glycerol gradient. Each fraction was further resolved on 3%–12% Bis-Tris native gel. Positions of native protein standards are indicated by arrows. RET1- and DSS1-containing complexes were visualized by immunoblotting. Thyroglobulin (19S) and bacterial ribosomal subunits were used as apparent S-value standards. See also Figure S1B.

(B) The well-characterized RNA editing core complex (RECC, ~1.2 MDa) served as separation marker. RECC was exposed by self-adenylation of RNA ligases REL1 and REL2 in the presence of [ $\alpha$ - $^{32}$ P]ATP (peak fractions are shown).

(C) Immunoaffinity purification of RET1-associated proteins. Purification was performed from RNase-treated mitochondrial extract. The final fraction was separated on 8%–16% SDS gel and stained with Sypro Ruby. The migration area of MPSS1-3 proteins is shown by a bracket.

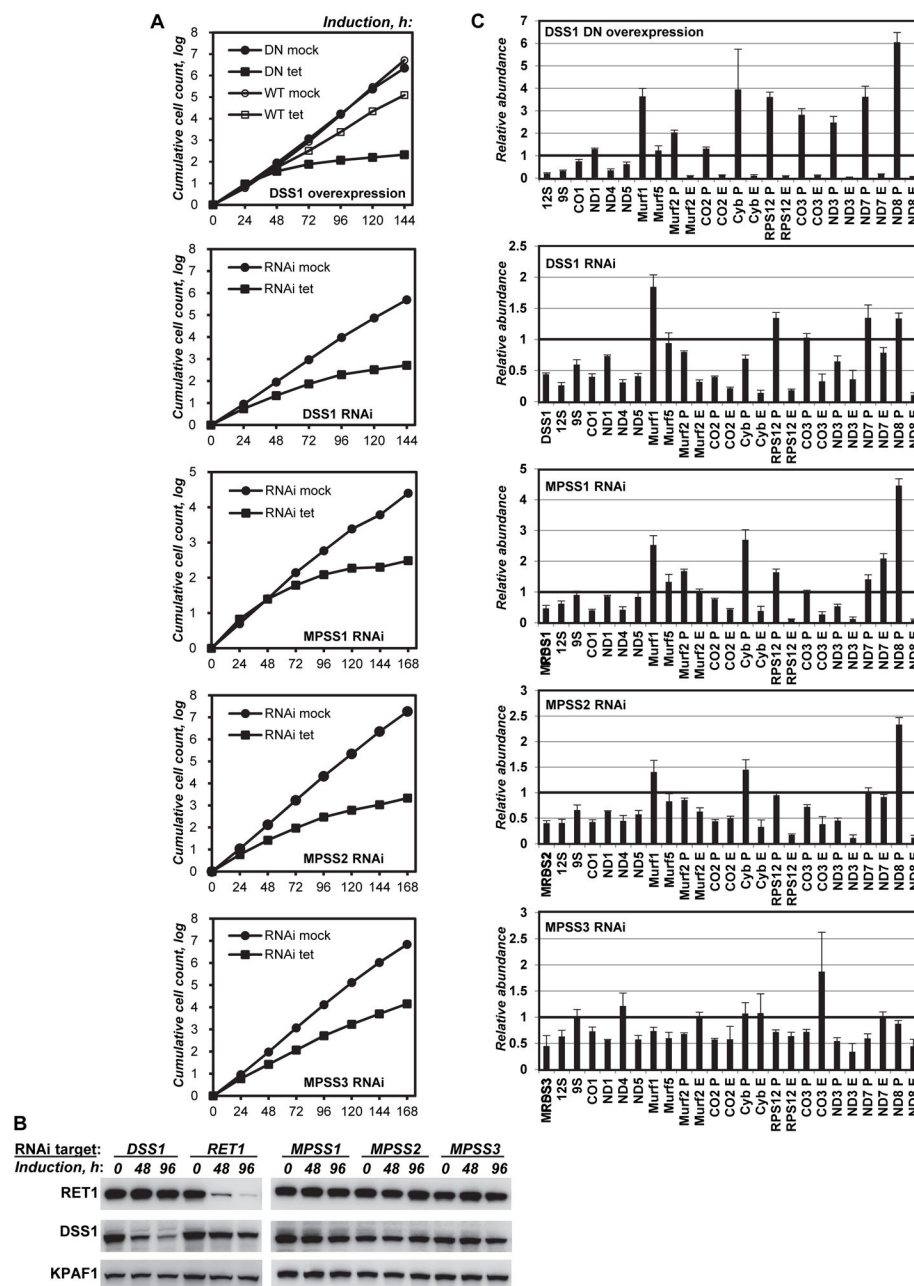
(D) Tandem affinity purifications of MPsome components. Final fractions were separated on 8%–16% SDS gel and stained with Sypro Ruby.

(E) MPsome protein interactions model. The network of RNase-resistant interactions was generated in Cytoscape software from bait-prey pairs in which the prey protein was identified with at least 20 unique peptides (Table S1). The edge thickness correlates with dNSAF values.

(F) Tandem affinity-purified complexes were normalized by immunoblotting with antibody against calmodulin binding peptide (bait), and probed for RET1 and DSS1. Antibody against kinetoplast polyadenylation factor KPAF1 served as negative control. See also Figure S1C.

(G) *In vitro* analysis of RET1 interactions within the MPsome. Synthesis: coupled transcription-translation reticulocyte system was used for protein synthesis in the presence of [ $^{35}$ S] methionine. Co-IP: RET1 and co-precipitated proteins were eluted from antibody-coated magnetic beads, separated on 8%–16% SDS PAGE and exposed to phosphor storage screen. DSS1 was also visualized by immunoblotting.

(H) *In vitro* analysis of selected pairwise protein interactions within the MPsome. C-terminally TAP-tagged proteins MPSS2 and MPSS3 were co-synthesized with RET1 or DSS1, and isolated with IgG-coated magnetic beads. Individual proteins were identified by radioactive signals or by immunoblotting. S, synthesis; PD, pulldown.



**Figure 2. Effects of DSS1 and MPSS1-3 Repression on Cell Viability and Mitochondrial RNAs**  
 (A) Growth kinetics of procyclic parasite suspension cultures after mock induction, protein and RNAi expression. See also Figure S2A, B.  
 (B) Immunoblotting of cell lysates with antibodies against RET1 and DSS1. KPAF1, loading control. See also figure S2C.  
 (C) Quantitative real-time RT-PCR analysis of RNAi-targeted nuclear mRNAs, and mitochondrial rRNAs and mRNAs. The assay distinguishes edited, corresponding pre-edited transcripts, and unedited mRNAs. RNA levels were normalized to  $\beta$ -tubulin mRNA. RNAi was induced for 72 hours. Error bars represent the standard deviation from at least three

replicates. The thick line at “1” reflects no change in relative abundance; bars above or below represent an increase or decrease, respectively. P, pre-edited mRNA; E, edited mRNA. See also Figure S2D.

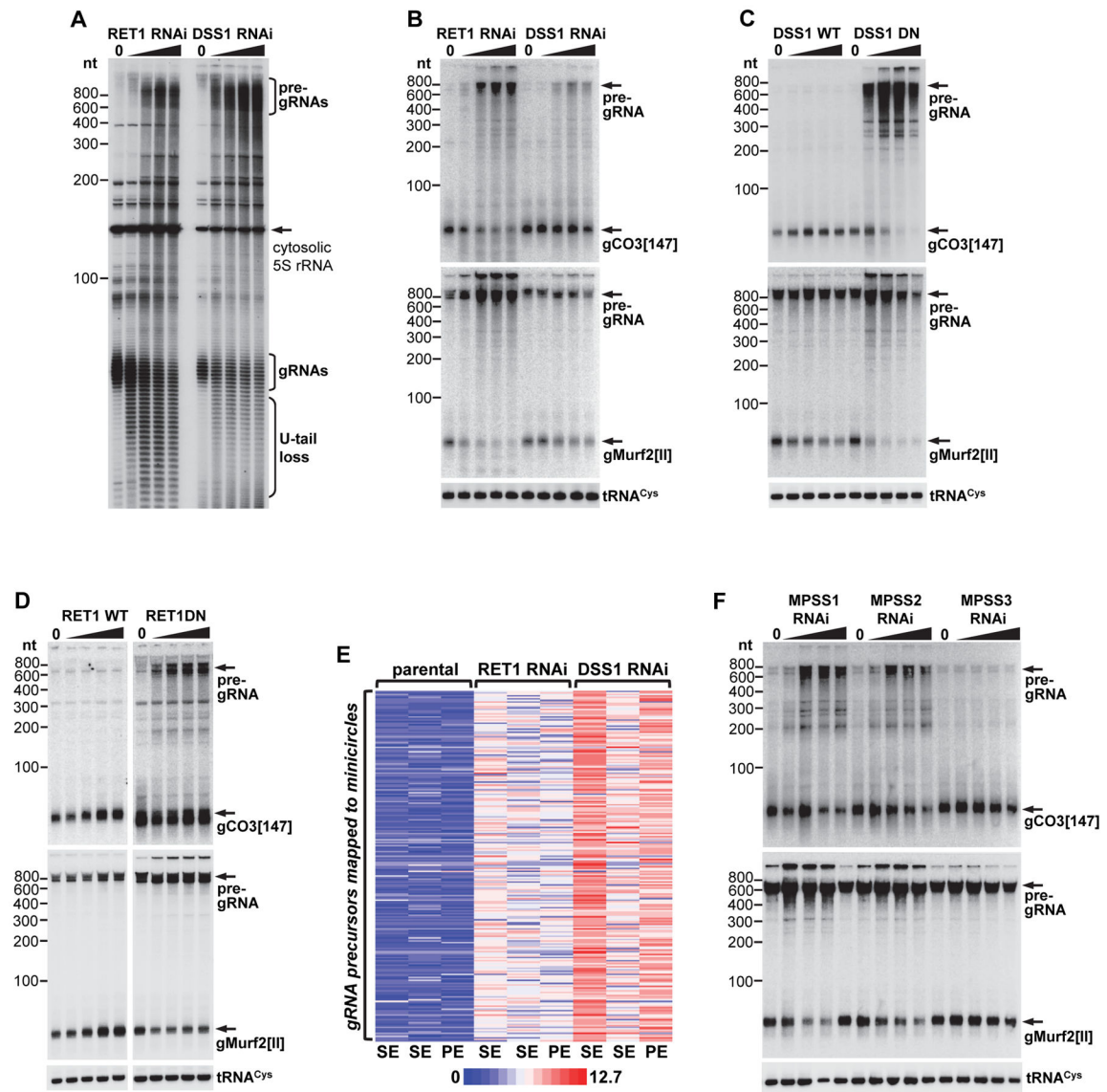
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript





**Figure 3. Inhibition of gRNA Processing upon Repression of MPsome Components**

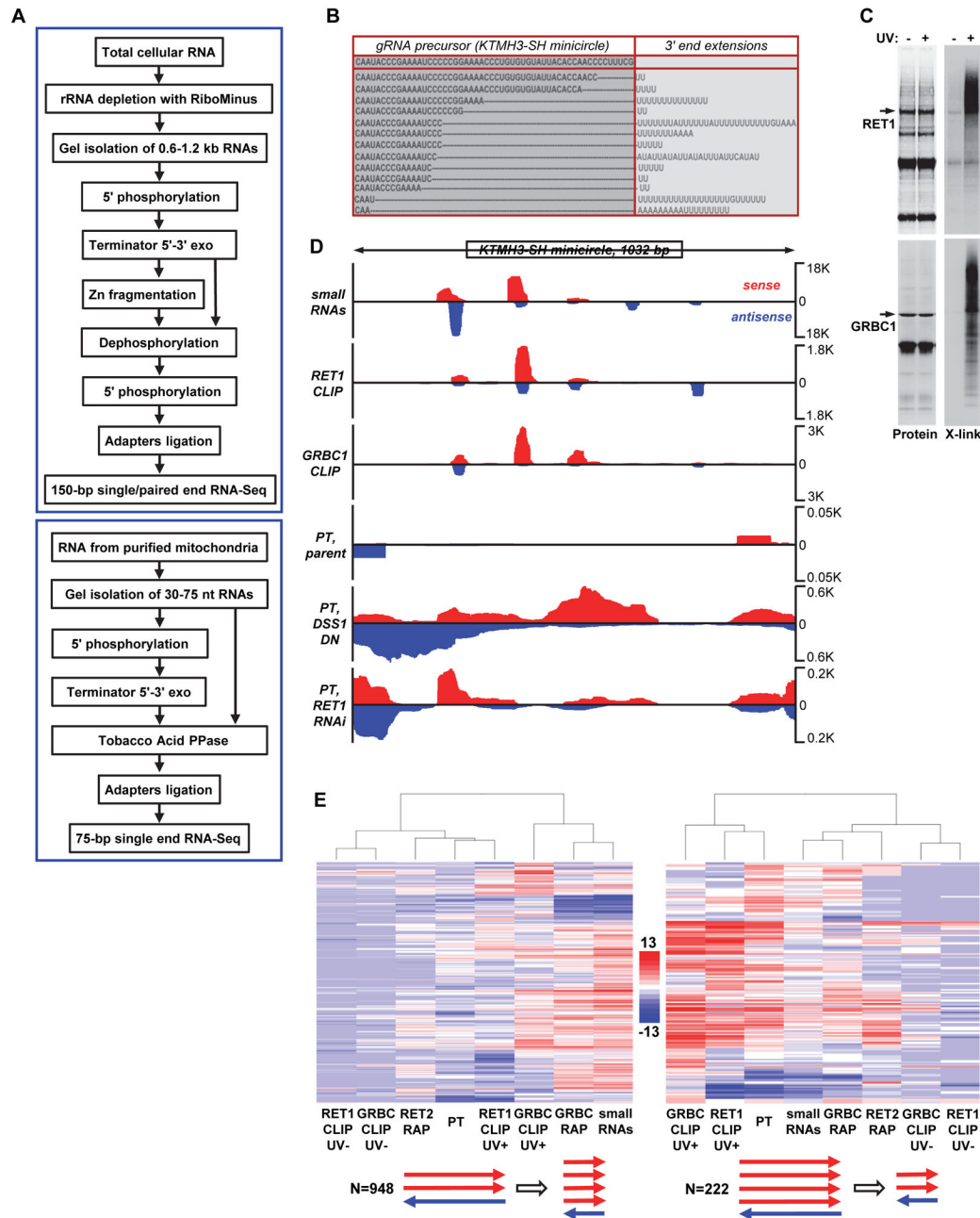
(A) Selective labeling of primary transcripts. RNAi expression was induced for 24, 48, 72 or 96 hours. Total RNA samples were 5'-labeled with [ $\alpha$ -<sup>32</sup>P]GTP in the presence of vaccinia virus guanylyltransferase and separated on 10% polyacrylamide/8M urea gel.

(B–D) Northern blotting of gCO3[147] minicircle- and gMurf2[II] maxicircle-encoded gRNAs in RET1 and DSS1 RNAi and protein overexpression cell lines. Total RNA was isolated from mock and tetracycline-induced cells at 24-hour time intervals, separated on 10% polyacrylamide/8M urea gel, and electroblotted onto membrane.

(E) RNA-Seq analysis of gRNA precursors. Total RNA samples from parental and RNAi cell lines induced for 72 hours were partially depleted of cytoplasmic rRNAs and separated on 5% polyacrylamide/8M urea gel. RNA was eluted from the 0.6–1.2 kb region, fragmented with Zn<sup>2+</sup> ions to ~200 nt fragments, and subjected to single-end (SE) 150 bp sequencing on Illumina HiSeq 2500 platform. Paired-end (PE) 150 bp sequencing was also

performed on unfragmented precursors. RNA reads were mapped to minicircle database (Table S5). To generate a heat-map, the read counts that mapped to individual minicircles were summed up for each replicate (SE, SE and PE runs) and grouped by RNA source. Parental, RNA isolated from 29–13 procyclic cell line.

(F) Northern blotting of gCO3[147] minicircle- and gMurf2[II] maxicircle-encoded gRNAs in MPSS1, MPSS2 and MPSS3 RNAi cell lines was performed as in panels (B–D). See also Figure S3.

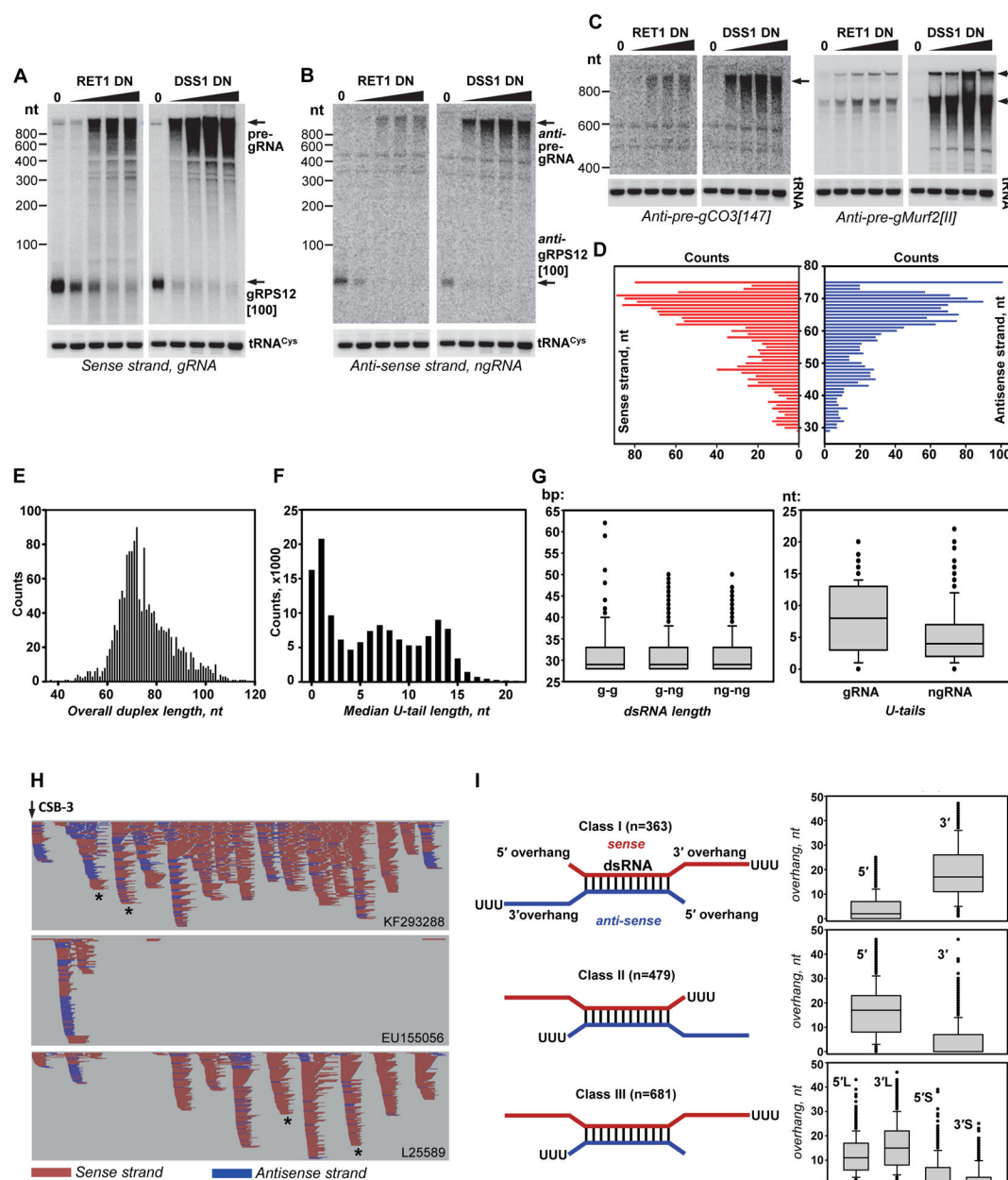


**Figure 4. *In vivo* Positioning and Substrate Specificity of the Mitochondrial Processome**  
(A) Workflow for sequencing of small mitochondrial RNAs and their precursors. Terminator<sup>TM</sup> 5'-3' exonuclease degrades 5'-monophosphorylated RNAs. Tobacco acid pyrophosphatase hydrolyzes phosphoric ester bond, thereby converting the 5'-triphosphorylated terminus into monophosphorylated RNA, which can be ligated to an adapter.  
(B) Multiple sequences alignment of gRNA precursor 3' tails to a representative minicircle. The 3' end reads were obtained from paired-end sequencing of gel-isolated unfragmented precursors from DSS1 RNAi cell line. DNA reference is provided in the upper row.

(C) Isolation of *in vivo* RNA-protein crosslinks of the MPsome (RET1) and gRNA binding complex (GRBC1) subunits. Immunoaffinity purified fractions were subjected to partial RNase digestion and RNA fragments bound to the protein were radiolabeled. Upon separation on SDS gel, RNA-protein crosslinks were transferred onto nitrocellulose membrane. Protein patterns were visualized by Sypro Ruby staining (left panels) and radioactive signals by exposure to phosphor storage screen. RNA was eluted from areas indicated by brackets and sequenced. UV, live parasites were subjected to UV-irradiation (+) or mock-treated (–).

(D) *In vivo* positioning of the MPsome and guide RNA binding complex. Strand-specific reads obtained from sequencing of small mitochondrial RNAs, RNA fragments crosslinked to RET1 TUTase and GRBC1, and precursors that accumulate in DSS1 DN and RET1 RNAi cells were mapped to the same minicircle as in Figure 4B. K, raw read counts in thousands.

(E) Relative abundance of two strands in a sense-antisense pair varies between the precursor transcript and mature small RNA. For each duplex pair, the log2 ratios of read counts from sense and antisense strands were calculated to represent the relative abundance of two strands in HITS-CLIP data, RNA isolated from rapid affinity pulldowns (RAP), and precursor transcripts (PT) and small RNA sequencing. In the hierarchically clustered heat-map, every row is a sense-antisense pair and every column is an experiment.



**Figure 5. Sense and Antisense Transcription of the Minicircle Genome**

(A) Northern blotting of gRPS12[100] minicircle-encoded gRNAs in RET1 and DSS1 DN lines. Total RNA was isolated from mock- and tetracycline-induced cells at 24-hour time intervals and separated on 10% polyacrylamide/8M urea gel.

(B) The same membrane as in (A) was stripped, re-exposed to phosphor storage screen to confirm signal elimination, and hybridized with a reverse complement probe for gRPS12[100].

(C) Reverse complement hybridization probes were designed for gCO3[147] minicircle- and gMurf2[II] maxicircle-encoded gRNAs in RET1 and DSS1 DN cell lines. Total RNA was

separated on 5% polyacrylamide/8M urea gel; under these conditions small RNAs are no longer retained in the gel.

(D) Length distribution of sense and antisense small RNAs that form potential hybrid intermediates. RNA was purified from highly-enriched mitochondrial fraction and 30–75 nt molecules were gel-isolated and treated with Terminator exonuclease prior to RNA-Seq. The U-tails were excluded during duplex search, but added back in the final assembly (Table S7).

(E) Overall length distribution of putative duplex intermediates.

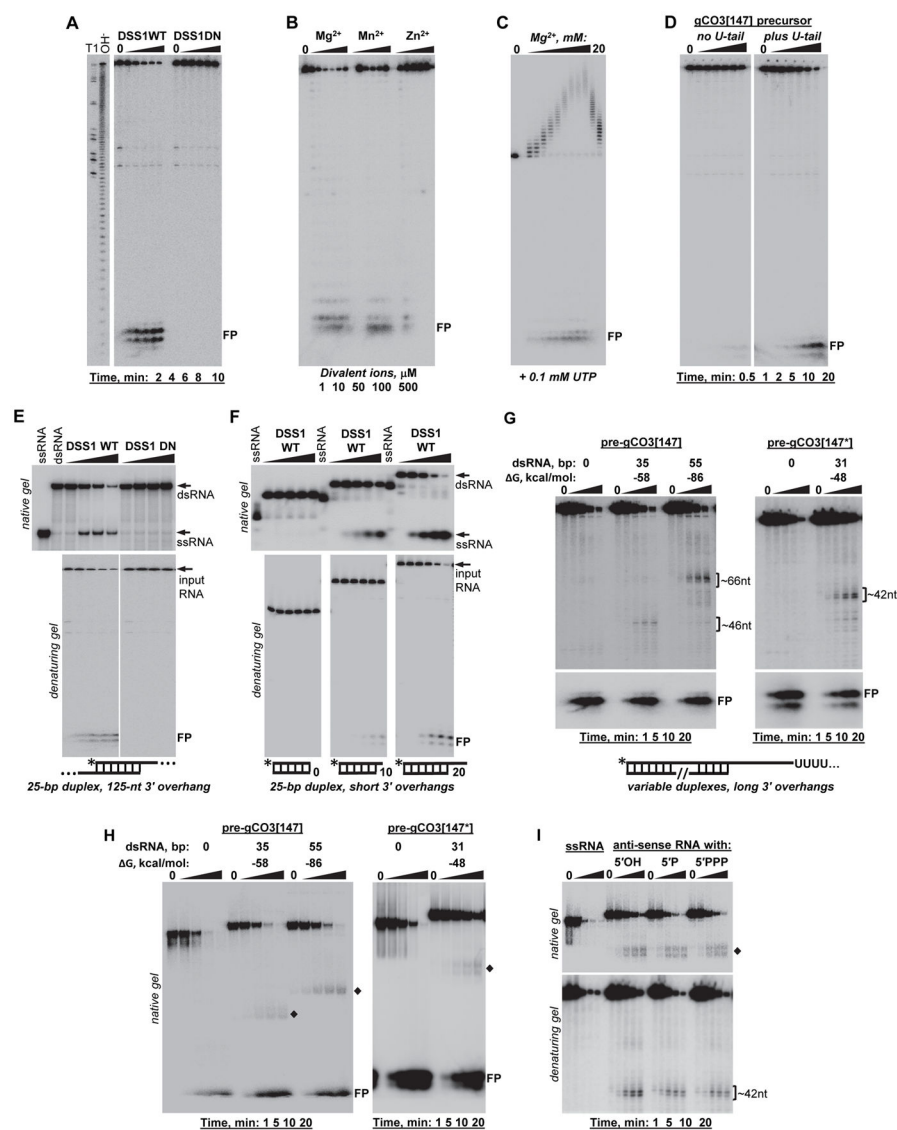
(F) U-tail length distribution among small mitochondrial RNAs.

(G) The duplexes were classified as gRNA-gRNA (g-g), gRNA-ngRNA (g-ng) and ngRNA-ngRNA (ng-ng) based on the gRNA annotation for sense and antisense transcripts. The boxplots were generated for the length of double-stranded regions in identified hybrids and for U-tail length in all gRNA and ngRNAs (Table S7).

(H) Small RNAs mapping to representative minicircles. Guide RNAs annotated in original GenBank submissions are indicated by asterisks.

(I) The duplexes were classified into three groups based on the overhang lengths. Class I duplexes have a longer 3' overhang vs. 5' overhang. Class II duplexes have longer 5' overhang vs. 3' overhang. In Class III, the transcripts from one strand have both longer 5' and 3' overhang sequences (L) while the opposite strand overhangs are short (S). The boxplots represent the differences in overhang length.





**Figure 6. Enzymatic and RNA Unwinding Activities of the Purified MPsome**

(A) The MPsome displays 3'-5' exonucleolytic activity. Tandem affinity-purified DSS1 and DSS1 D202A (DN) complexes were incubated with 5' radiolabeled gCO3[147] guide RNA. Products were resolved on 15% polyacrylamide/8M urea gel. T1, partial RNA digestion by guanine-specific RNase T1; OH<sup>-</sup>, partial RNA hydrolysis by sodium hydroxide. See also Figure S4.

(B) Metal ions modulate the MPsome's exonuclease activity. Synthetic 24-mer RNA was incubated with DSS1 complex for 10 min at varying concentrations of divalent ions.

(C) Magnesium ion dependence of the MPsome-embedded TUTase activity. Reactions were performed with DSS1 complex and 24-mer RNA in the presence of 0, 0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 2.5, 5, 10 and 20 mM of MgCl<sub>2</sub> and 0.1 mM of UTP for 20 min. Extension and degradation products were separated on 15% polyacrylamide/8M urea gel.

(D) Uridylation stimulates RNA recognition by the MPsome. Synthetic 5' radiolabeled gRNA precursors with or without 3' U-tail were incubated with DSS1 complex for indicated time periods and separated on 10% polyacrylamide/8M urea gel.

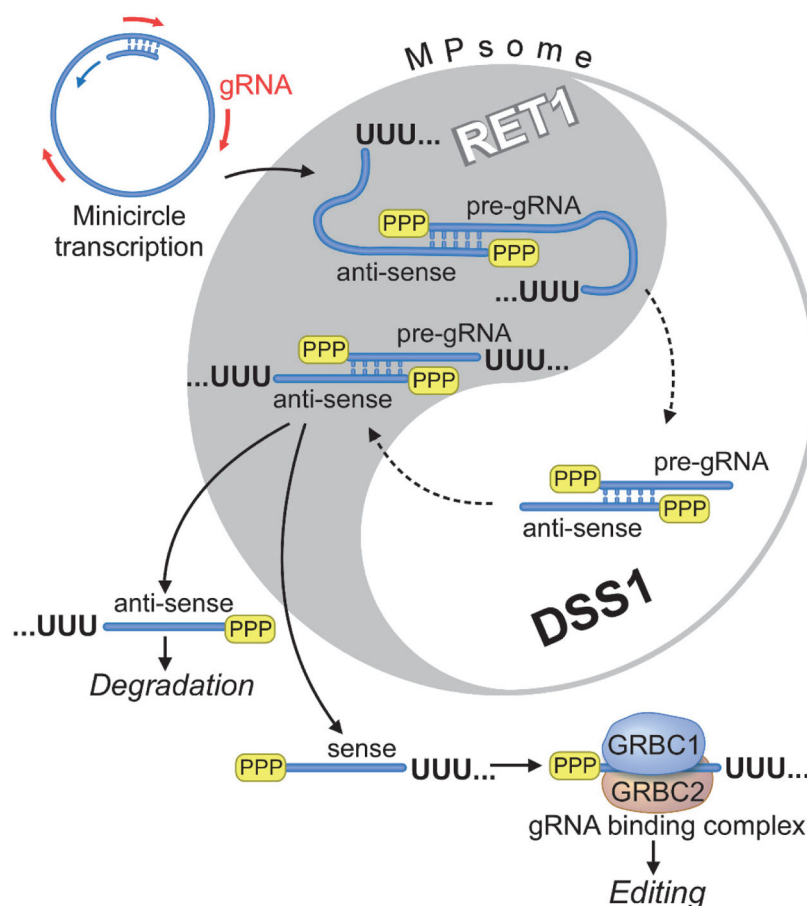
(E) RNA hydrolysis-dependent unwinding activity of the MPsome. Reaction with pre-assembled duplex RNAs was terminated by adding EDTA to 10 mM and SDS to 0.5%, and resolved on 7% native Tris-borate gel (upper panels) or 15% polyacrylamide/8M urea denaturing gel.

(F) The MPsome requires a single-stranded RNA to initiate RNA hydrolysis and duplex unwinding.

(G) The MPsome pauses at 10–12 nt before sufficiently stable double-stranded region. The duplex lengths and two-state hybridization free energies (UNAFold, <http://mfold.rna.albany.edu>) are positioned above gel panels. Abortive products are shown by brackets. Reaction products were separated on 10% polyacrylamide/8M urea denaturing gels. See also Figure S4G, H.

(H) MPsome pausing generates RNA duplex with shortened 3' overhang. Products from the same reactions as in (G) were separated on 7% native Tris-borate gel. Positions of duplexes with trimmed 3' overhangs are shown by diamonds.

(I) Chemical nature of the 5' end in antisense RNA does not affect MPsome pausing. Antisense RNA was produced by *in vitro* transcription with T7 RNA polymerase and dephosphorylated (5'OH), re-phosphorylated (5'P) or used directly (5'PPP) to assemble duplex substrate with pre-gCO3[147\*].



**Figure 7. Guide RNA Biogenesis in Mitochondria of *Trypanosoma brucei***

Bidirectional transcription of gRNA gene generates sense and antisense precursor RNAs with overlapping 5' regions. Mitochondrial 3' processome (MPsome) composed of RET1 TUTase, DSS1 3'-5' exonuclease and additional subunits catalyzes three sequential processing reactions: primary precursor uridylation, processive precursor degradation and secondary uridylation of the mature gRNA. Primary uridylation stimulates hydrolytic activity of DSS1, which provides energy for unwinding the secondary structures along gRNA precursor. Nonetheless, the MPsome stochastically pauses at 10–12 nt from sufficiently stable duplex regions. The MPsome pausing allows RET1 TUTase to perform secondary uridylation. This step may disengage the MPsome from the duplex intermediate. Double-stranded RNA likely undergoes active unwinding before mature gRNA can be sequestered by the gRNA binding complex and delivered into the editing pathway.