Footprints of a trypanosomatid RNA world: pre-small subunit rRNA processing by spliced leader addition *trans*-splicing

Mario Gustavo Mayer[1], Marcos Gonzaga dos Santos[2],
Maria Fernanda Laranjeira da Silva[2], Lucile Maria Floeter-Winter[2]/+

[1]Laboratório de Genética, Instituto Butantan, São Paulo, SP, Brasil
[2]Departamento de Fisiologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, Brasil

The addition of a capped mini-exon [spliced leader (SL)] through trans-*splicing is essential for the maturation of RNA polymerase (pol) II-transcribed polycistronic pre-mRNAs in all members of the Trypanosomatidae family. This process is an inter-molecular splicing reaction that follows the same basic rules of* cis-*splicing reactions. In this study, we demonstrated that mini-exons were added to precursor ribosomal RNA (pre-rRNA) are transcribed by RNA pol I, including the 5' external transcribed spacer (ETS) region. Additionally, we detected the SL-5'ETS molecule using three distinct methods and located the acceptor site between two known 5'ETS rRNA processing sites (A' and A1) in four different trypanosomatids. Moreover, we detected a polyadenylated 5'ETS upstream of the* trans-*splicing acceptor site, which also occurs in pre-mRNA* trans-*splicing. After treatment with an indirect* trans-*splicing inhibitor (sinefungin), we observed SL-5'ETS decay. However, treatment with 5-fluorouracil (a precursor of RNA synthesis that inhibits the degradation of pre-rRNA) led to the accumulation of SL-5'ETS, suggesting that the molecule may play a role in rRNA degradation. The detection of* trans-*splicing in these molecules may indicate broad RNA-joining properties, regardless of the polymerase used for transcription.*

Key words: external transcribed spacer - spliced leader RNA - RNA polymerase I -
spliceosome - RNA polymerase II carboxy-terminal domain

Trypanosomatids are a remarkable group of early branching protozoa that have developed unusual strategies to become efficient eukaryotic parasites (Simpson et al. 2006). Although data on the regulation of gene expression by the modification of chromatin at sites of transcription initiation exist (Siegel et al. 2009, Thomas et al. 2009), trypanosomatids primarily regulate gene expression at the post-transcriptional level via unique RNA metabolism (Clayton 2002). Unlike most eukaryotes, protein-coding genes in trypanosomatids are transcribed from a few constitutive sites that do not present conserved sequences and are not organised with conventional RNA polymerase (pol) II promoter regions (Martinez-Calvillo et al. 2003, Worthey et al. 2003). Post-transcription, long polycistronic pre-mRNAs undergo two successive cleavage events at intergenic regions to convert the polycistronic pre-mRNA to monocistronic mRNA. The first cleavage is a *trans*-splicing of a pre-capped 39-nt exon (mini-exon) from the highly abundant spliced leader (SL) RNA to the 5' UTR boundary of a gene in the pre-mRNA substrate (Liang et al. 2003, Mayer & Floeter-Winter 2005). This inter-molecular splicing reaction obeys many of the same basic rules as *cis*-splicing reactions, including the GU-AG donor-acceptor site rule

and the requirement of a polypyrimidine tract upstream of the acceptor site. A second cleavage event occurs after the *trans*-splicing event at sites located from 150 nt (*Trypanosoma brucei*) to 500 nt (*Leishmania*) upstream of the *trans*-splice acceptor site. Additionally, this cleavage event marks the site to be polyadenylated during the maturation of the preceding mRNA (LeBowitz et al. 1993, Hug et al. 1994). The conventional polyadenylation signals characteristic of higher eukaryotes appear to be absent in trypanosomatid pre-mRNAs. The coupled *trans*-splicing and polyadenylation reactions lead to the production of a large number of mature mRNAs from a single transcription initiation event.

Although first observed in trypanosomatids, the SL addition *trans*-splicing during mRNA maturation has been detected in many other organisms, such as nematodes, flatworms, dinoflagelates, cnidarians, nonvertebrate chordates, rotifers, ctenophores, sponges and arthropods (Davis 1996, Douris et al. 2010). In trypanosomatids, all mRNAs are targets of SL addition, but there is no evident functional reason to explain why only some of the transcripts are *trans*-spliced in other organisms. Furthermore, it is not clear if this processing evolved earlier in the metazoan lineage or appeared repeatedly among diverse metazoan lineages (Nilsen 2001).

Another unusual feature of the trypanosomatid gene expression repertoire is the use of the powerful RNA pol I machinery to produce large quantities of *T. brucei* stage-specific variant surface glycoprotein and procyclin precursor (pre-mRNA) (Rudenko et al. 1991, Zomerdijk et al. 1991). These pre-mRNAs are also processed through *trans*-splicing (Huang & Van der Ploeg 1991). The RNA pol I transcription of natu-

ral pre-mRNAs has been reported only for *T. brucei*; however, with the use of artificial constructs, the RNA pol I-controlled expression of protein-coding genes has been reported for the majority of trypanosomatids (Lee & Van der Ploeg 1997). We have shown that the *T. cruzi* and *Leishmania (Leishmania) amazonensis* rRNA pol I promoters can drive the expression of the chloramphenicol acetyltransferase reporter gene, even when no heterologous *trans*-splice acceptor site is added to the reporter construct (Tyler-Cross et al. 1995, Uliana et al. 1996, Floeter-Winter et al. 1997). These experiments suggest that the 5' external transcribed spacer (ETS) rDNA region can encompass a native 3' acceptor site and that pre-rRNA can be *trans*-spliced.

It is interesting to note that trypanosomatid RNA pol II does not present the canonical carboxy-terminal domain (CTD) responsible for transcription modulation. Instead, it contains a non-canonical CTD that lacks the heptapeptide repeats usually found in Metazoa. In *T. brucei,* this pseudo-CTD (approximately 285 amino acids) is important in gene expression regulation. Moreover, the 95 amino acids at the carboxy-terminus of the largest subunit of RNA pol II are essential for parasite survival (Das & Bellofatto 2009). Thus, we speculate that the joining of RNA molecules could be an ancient process that is evolutionarily conserved in organisms with partially specialised RNA polymerases.

In this paper, using multiple methods, we show that the pre-rRNA of four divergent trypanosomatids can be processed through *trans*-splicing using their 5'ETS sequences as acceptor substrates. Polyadenylated 5'ETS molecules were also detected, indicating that these processes may be coupled similarly to pre-mRNA processing. A modulation of SL-5'ETS molecules was achieved using *trans*-splicing and exosome inhibitors, such as sinefungin or 5'fluorouracil. Although the functional significance of these findings is unclear, we hypothesise that mini-exon addition and/or polyadenylation could serve as tags to drive the destruction of waste molecules in the nuclear compartment.

MATERIALS AND METHODS

Organisms and culture - Promastigotes of *L. (L.) amazonensis* (MHOM/BR/73/M2269) were cultivated at 25°C in 199 medium (Invitrogen, Grand Island, NY, USA) supplemented with 10% foetal calf serum (FCS) (Invitrogen), 10 µg/mL haemin, 100 U/mL penicillin and 100 µg/mL streptomycin as previously described (Kapler et al. 1990). Epimastigotes of *Trypanosoma cruzi* (Y strain) were grown at 28°C in liver infusion tryptose (LIT) medium supplemented with 10% FCS as previously described (Camargo 1964). *Crithidia fasciculata* (ATCC 30267) was grown at 25°C in LIT medium. Procyclic *T. brucei* (strain 427) was grown at 28°C in SDM-79 medium supplemented with 10% FCS as previously described (Brun & Schonenberger 1979).

RNA and DNA extractions - Trypanosomatid RNA and DNA were isolated from late log-phase cultures. RNA was isolated with TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The RNA sample was then digested with RNase-free DNase (Fermentas Life

Sciences, Burlington, Ontario, Canada) and extracted with TRIzol LS reagent. DNA was extracted with a previously described "mini-prep" procedure (Medina-Acosta & Cross 1993); however, alkaline hydrolysis (50 mM NaOH and 2 mM EDTA, pH 8.0) was performed at 65°C for 60 min in place of RNase digestion.

Oligonucleotides - RT - 5'CGGTTGGTTCATGTATTAGCTTGGCGTTTC3'; Rev1 - 5'TCTGATGTAATGAGCCATGCGCAGATTC3'; Rev2 - 5'ATGCATGGCTAATGTCCTTG3'; SL from *L. (L.) amazonensis* and *C. fasciculata* - 5'GCTATATAAGTATCAGTTTC3'; SL from *T. brucei* - 5'AACGCTATTATTAGAACAGTTTC3'; SL from *T. cruzi* - 5'AACGCTATTATTGATACAGTTTC3'; SSU S4 - 5'GATCCAGCTGCAGGTTCACC3'; SSU S12 - 5'GGTTGATTCCGTCAACGGAC3'; αets - 5'GGCAGAATCAACCAG3'; SRTA - 5'CATACACGACTCACGTTGTCACACA3'; PA - 5'GAGACGCGTGTGGCGCAC3'; Pext 2 - 5'GGCAGAATCAACCAG'; ETSense 1 - 5'CTGCGCGTGCTGTGTAC3'; ETSense 2 - 5'AGTTTCCAGCCTTACTGCTTG3'; Oligo dT (16) adaptor - 5'GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT3'; Adaptor - GACTCGAGTCGACATCGA3'.

Reverse transcription (RT) - To minimise RTstops due to a highly structured template, we used a hot-start protocol. According to the reverse transcriptase manufacturer's instructions (Invitrogen), 5 µg RNA and 2 pmol RT primer were denatured at 65°C for 5 min in a mixture containing 10 nmol dNTPs. After denaturation, the temperature was decreased to 50°C. After 2 minu at 50°C, a mixture preheated to the same temperature was added to adjust the reaction conditions to 50 mM Tris-Cl, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 5 mM DTT and 40 U RNaseOUT RNase inhibitor (Invitrogen). After 2 min, 200 U SuperScript III reverse transcriptase (Invitrogen) was added and the reaction was incubated for 60 min. The reaction was then inactivated through incubation at 70°C for 15 min. A mock reaction without reverse transcriptase was performed as a negative control for DNA contamination.

cDNA amplification through semi-nested polymerase chain reaction (PCR) - During the first round of PCR, 1/10 of the reverse-transcribed cDNA from each species was PCR-amplified using a common Rev1 primer. A species-specific forward SL primer was also used for *L. (L.) amazonensis, C. fasciculata, T. brucei* and *T. cruzi*. The PCR was performed in a final volume of 50 µL containing 10 mM Tris-Cl, pH 8.8, 50 mM KCl, 0.08% Nonidet-P40, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 10 µM of each primer and 1.25 U *Taq* DNA polymerase (Fermentas Life Sciences, Burlington, Ontario, Canada). A 2-min pre-incubation step at 94°C was followed by 35 cycles of 94°C for 1 min, 57°C for 1 min and 72°C for 1.5 min, with a final step at 72°C for 10 min. One microliter of a 10^4 dilution of the PCR product was then used as the template in a second round of PCR using the same SL primers and Rev2, a nested reverse primer. The reaction and cycling conditions were identical to the first PCR, with the exception that the annealing temperature was 58°C. The PCR products were then analysed on a 2% agarose gel, with a 100 bp DNA ladder as a marker (Fermentas). For the DNA amplifications, DNA (50 µg) was either PCR-

amplified using the same semi-nested strategy and conditions used for cDNA amplification (as described above) or directly amplified as described for the second nested PCR reaction. To determine the DNA quality, a control PCR was performed in parallel using the ribosomal RNA small subunit (SSU) universal primers S4 and S12. In this reaction, 0.5 ng DNA was amplified with 10 mM Tris-Cl, pH 8.8, 50 mM KCl, 0.08% Nonidet-P40, 1.5 mM MgCl$_2$, 0.2 mM each dNTP, 10 µM of each primer and 1.25 U *Taq* DNA polymerase (Fermentas). The cycling conditions were the same as described above; however, PCR was only performed for 30 cycles (Uliana et al. 1994). The amplified products were analysed through electrophoresis on a 2% agarose gel. To determine the *T. cruzi* Y strain 5'ETS sequences around the acceptor sites, 50 ng of genomic DNA was amplified with the p66.2 5'ETS sense (Dietrich et al. 1993) and Rev2 primers under the same conditions as the S4-S12 amplifications; however, the reaction was performed under the following cycling parameters: an initial denaturation step at 94ºC for 2 min was followed by 30 cycles of 94ºC for 1 min, 55ºC for 1 min and 72ºC for 1.5 min, with a final extension step at 72ºC for 10 min. The PCR products were cloned into the PGEM®-T Easy (Promega, Madison, WI, USA) vector according to the manufacturer's instructions and clones were selected based on their length after digestion with the *Not*I restriction enzyme (Fermentas Life Sciences, Burlington, Ontario, Canada). Nucleotide sequences were obtained using the BigDye® Terminator v3.1 Cycle Sequence Kit (Applied Biosystems, Carlsbad, CA, USA) and an ABI PRISM® 3100 GeneticAnalyser.

Primer extension (PE) - Total *L. (L.) amazonensis* RNA (25 µg) was used as a template in PE analysis. PE was performed with 5 x 10^6 cpm of T4 polynucleotide kinase 5' end-labelled ETS specific Pext 2 primer. After annealing at 37ºC for 15 min, 200 U of M-MLV reverse transcriptase (Invitrogen) was added and the extension was performed at the same temperature for 50 min according to the manufacturer's suggested conditions. Subsequently, the reaction was inactivated at 70ºC for 15 min, precipitated with ammonium acetate and ethanol and analysed on an 8% polyacrylamide denaturing gel alongside a sequencing reaction as a length marker.

RNase protection - The SL5'ETSSSU and ETSSSU DNA fragments cloned in the p-GEM®-T Easy vector were used as the templates for the synthesis of anti-sense riboprobes. The SL5'ETSSSU insert was obtained with real-time (RT)-PCR with Rev2 and the *L. (L.) amazonensis* SL oligonucleotides. The ETSSSU fragment was obtained with PCR using Rev2 and PA oligonucleotides. The correct insert orientation was verified through sequencing. The plasmid DNA containing the probe sequence was digested with *Nco*I and used as the template to synthesise high specific-activity riboprobes with a MAXIscript® In Vitro Transcription Kit (Ambion/Applied Biosystems, USA) using 50 µCi Uridine- 5'- α^{32}P (Amersham Redivue SP6/T7 grade - 800 Ci/mmol, 20 mCi/mL, GE Healthcare, UK). The probes were purified on a denaturing 5% polyacrylamide gel. RNase and SP6 RNA polymerase of RAPIII – ribonuclease protection assay (Ambion, Applied Biosystems/Life Technologies, Carlsbad, CA, USA). Each probe was purified through Sephadex G20 filtration and then hybridised to different amounts of *L. (L.) amazonensis* total RNA using the RPAIII™ kit (Ambion) according to the manufacturer's protocol. Briefly, 100,000 cpm of each probe was separately mixed with total RNA and 10 µL hybridisation buffer. The mixtures were heated at 94ºC for 3 min and hybridised overnight at 42ºC. Subsequently, each mixture was precipitated and the pellets were dissolved and treated with RNase according to the instructions for the RPAIII kit. The protected products were fractionated with a denaturing 5% polyacrylamide/8 M urea gel. A non-labelled RNA marker (Decade™ Markers, Ambion) was used to estimate molecular sizes. After drying, the gel was exposed in a Phosphoimager (Molecular Dynamics, USA).

Detection of polyadenylation in the ETS - RTwas performed as described above for cDNA preparation, except that the extension temperature was 50ºC and an oligo-dT (16) adaptor primer was used. The cDNA products were amplified through a three-round PCR strategy (Fig. 1). First, 1/20 of the RT reaction was amplified for 35 cycles (94ºC for 20 s, 52ºC for 30 s and 72ºC for 45 s) using an ETS sense 1 and an AS adaptor primer. The PCR products were then diluted 100x and amplified for another round of 35 cycles under identical conditions with ETS-2 and AS adaptor primers. Finally, after another 100X dilution of the PCR products, a third amplification was performed under the same reaction conditions and the products were analysed through agarose gel electrophoresis, cloned and sequenced as described above.

Sinefungin or 5-fluorouracil (5-FU) treatments - Promastigotes of *L. (L.) amazonensis* (1-2 x 10^7) were treated with 1 µg/mL sinefungin (Sigma-Aldrich, MO, USA) as described previously (Clayton et al. 2008). The treatment was stopped by lysing the parasites with TRIzol reagent, as described for the RNA extraction.

The same number of promastigotes was treated with 5-FU (Sigma-Aldrich, Saint Louis, MO, USA) at the indicated concentrations for 90 min. The treatment was stopped by the addition of TRIzol Reagent, as described above.

Quantitative RT-PCR (qRT-PCR) - The obtained cDNA was diluted to the equivalent of 250 ng of RNA in water and used for qRT-PCR with the addition of 0.2 µM of the indicated primers. We used the SYBR Detection System (Applied Biosystems, USA) according to the manufacturer's instructions in a 50 µL final reaction volume with a 7300 System (Applied Biosystems, USA). The reactions were initially maintained at 50ºC for 2 min and then held at 94ºC for 10 min. A total of 50 cycles were then performed, consisting of a denaturation step at 94ºC for 30 s, an annealing step at 50ºC for 30 s and an extension step at 72ºC for 30 s. Known amounts of the cloned DNA products were diluted to produce a standard curve for each pair of primers. The primers used to amplify arginase and SSU rRNA are described elsewhere (Uliana et al. 1994, dos Santos et al. 2011).

Data analysis - Statistical significance was determined using a Student's *t* test (p < 0.05).

RESULTS

Pre-rRNA is a substrate for trans-splicing in trypanosomatids - To verify the addition of an SL by *trans*-splicing in the 5'ETS regions of pre-rRNAs of four trypanosomatid species (*T. brucei*, *T. cruzi*, *L. (L.) amazonensis* and *C. fasciculata*), we developed a semi-nested RT-PCR strategy (Fig. 2A, B). The total RNA preparations of these trypanosomatids were reverse transcribed using an oligonucleotide complementary to a conserved SSU rRNA region (a ribosomal RNA small subunit-coding region) as a primer (Fig. 2B - RT). The cDNAs were initially amplified with an SL species-specific forward primer and a nested SSU rRNA primer complementary to a second conserved region (Fig. 2B - Rev1). The PCR products were then used in a second PCR with the same SL primer and a second nested SSU rRNA conserved primer (Fig. 2B - Rev2). The agarose gel analyses of the RT-PCR products revealed different patterns of possible *trans*-spliced products for each organism and no products in the mock reactions (Fig. 2C). A single product was detected for *L. (L.) amazonensis*, whereas two bands were observed for all of the other trypanosomatids. Negative results from similar reactions using DNA as a template indicated that these RT-PCR products did not arise from recombination at the DNA level because a 540 bp product was detected only when genomic DNA was used as the template (Fig. 2D). Furthermore, GeneDB sequence analyses of the *T. brucei* strain indicated that the SL and rDNA sequences were located on different chromosomes, excluding the possibility of a *cis*-splicing event, at least in that species.

Pre-rRNA trans-splicing adheres to canonical splicing rules - The RT-PCR products were cloned and sequenced to verify that they represented *bona fide* pre-SSU rRNA *trans*-spliced molecules. The 5' ends of all of the products were composed of the species-specific mini-exon forward primer (represented in red lowercase in Fig. 3A), followed by the downstream mini-exon sequence (represented in black uppercase in Fig. 3A) joined to the corresponding species-specific 5'ETS sequence. The downstream mini-exon sequence was initially omitted from the design of the primer to confirm that the complete mini-exon was added to the 5'ETS. The mini-exon addition sites were located through the alignment of the 5'ETS of the RT-PCR product with genomic rDNA sequences. The joining occurred immediately upstream of the canonical GU donor splice site of the SL RNA (Fig. 3B). The double-banded patterns that were observed for *C. fasciculata*, *T. brucei* and *T. cruzi* were explained by the detection of two 5'ETS joining sites, whereas *L. (L.) amazonensis* had a single joining site. The canonical AG dinucleotide splice acceptor site and its associated polypyrimidine tract were found immediately preceding the 5'ETS joining sites in all of the genomic sequences (Fig. 3B). Moreover, the polypyrimidine tracts contained more than twice as many T residues as C residues. Thus, hybrid SL-5'ETS generation obeys the same basic rules as pre-mRNA *trans*-splicing; it generally occurs at the first AG dinucleotide 68 nt or less downstream of a T-rich polypyrimidine tract of 8-25 nt (Benz et al. 2005). An exception was observed for the downstream acceptor site in *T. cruzi*, which was located at a greater distance downstream (Fig. 3B).

Trans-spliced pre-SSU rRNA analyses revealed open reading frames (ORFs) that contained no more than 45 amino acid residues; thus, it is unlikely that they represent functional coding molecules. A BLASTP search of
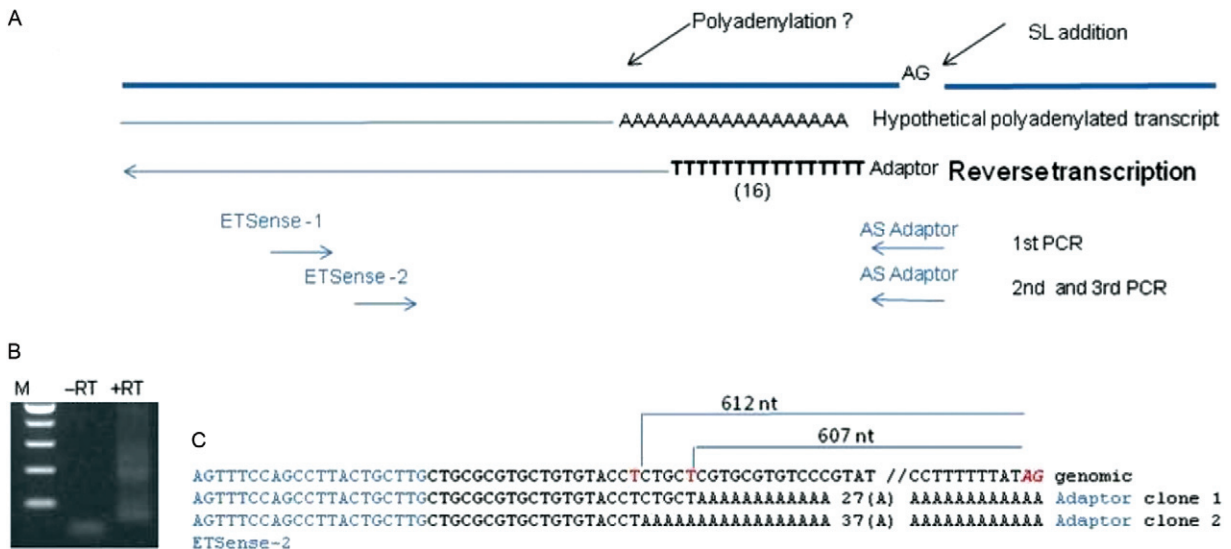

Fig. 1: detection of a polyadenylated external transcribed spacer (ETS) molecule. A: semi-nested reverse transcription-polymerase chain reaction (RT-PCR) strategy used to detect polyadenylated molecules (RT with oligodT (16)/adaptor primer, a subsequent round of PCR amplification with adaptor anti-sense primer and ETSense-1 and 2 rounds with adapter ETSense-2 primers); B: PCR products from the final three rounds were analyzed by agarose gel electrophoresis. RT- and RT+ correspond to products obtained from RT reactions with and without reverse transcriptase and M corresponds to the fragment length marker; C: nucleotide sequence analysis of cloned products. The ETSense primer is shown in blue letters and red Ts indicate polyadenylation points. The red AG indicates the *trans*-splicing point. SL: spliced leader.

the *T. brucei* genome for each candidate peptide sequence was performed in the homologous data bank (genedb.org/genedb/tryp/blast.jsp) using default parameters and no sequence identities were found. Furthermore, none of the putative RNAs were found in the *T. brucei* EST database, suggesting that a functional peptide is unlikely to be produced from these sequences (not shown).

The SL-5'ETS can be detected through alternative methods - To verify that the SL-5'ETS molecule was not a product of spurious RT-PCR, we identified this molecule in the *L. (L.) amazonensis* total RNA preparation using PE and RNase protection analyses (Fig. 4). For PE assays, a labelled oligonucleotide complementary to the 5'ETS sequence was used as a primer in the RT of *L. (L.) amazonensis* total RNA (Fig. 4A - αets), producing fragments of 73-77 nt. Because the CAP structure causes reverse transcriptase slippage, these fragments were expected (Fig. 4B).

For the RNase protection analysis, we constructed one plasmid containing the *trans*-spliced 5'ETS sequence followed by a portion of the SSU rRNA sequence (Fig. 4A - αSLETSSSU) and another plasmid containing the 5'ETS sequence and a portion of the SSU rRNA sequence (Fig. 4A - αETSSSU). In both plasmids, the designed sequence was cloned adjacent to an SP6 promoter region. The in vitro transcription of the plasmid DNA using SP6 RNA polymerase produced labelled probes encompassing sequences complementary to the target RNA. These probes were used in a hybridisation reaction with a large amount of total RNA, followed by RNase digestion. The protected products, fractionated with denaturing polyacrylamide gels, presented a strongly labelled 107 nt fragment for both probes, which was the expected 5'ETS-SSU portion of the target. However, the probe for SL-5'ETS (αSLETSSSU) also produced a small amount of 140 nt protected fragments and a fragment corresponding to SL alone (33 nt) (Fig. 4C). The amount of the 107 nt fragment detected was congruent with the large amount of ribosomal transcription in the cell. The small amount of SL-5'ETS RNA detected corroborated the RT-PCR results because the molecule was detected through a two-round nested reaction. It is important to note that the RNase protection experiments were performed using a large amount of input total RNA to allow for the visualisation of the small amount of protected SL-5'ETS RNA. However, it is difficult to make a direct quantitative comparison between the results obtained with the two methods. It seems clear that the non-*trans*-spliced rRNA fragments are saturated in the RNase protection assay and therefore, it is not easy to correlate the 10^4-fold difference in the amount of these molecules and the number of spliced molecules as determined through qRT-PCR. The other fragments that were protected with both probes (Fig. 4C) are most likely pre-RNA processing intermediates.

rRNA 5'ETS can be polyadenylated - Natural pre-mRNA *trans*-splicing processes are coupled to the polyadenylation of the upstream mRNA sequence in the precursor molecule. Thus, we searched for 5'ETS polyadenylated sequences in a total RNA preparation from *L. (L.) amazonensis* with a nested PCR strategy with
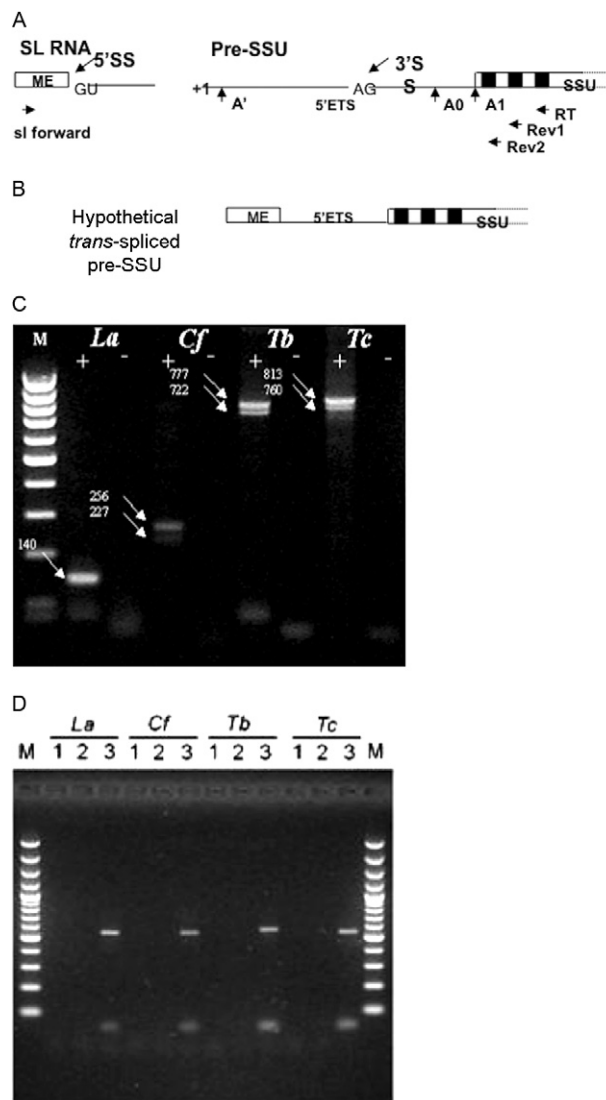

Fig. 2: polymerase chain reaction (PCR) detection of *trans*-spliced precursor ribosomal RNA (pre-rRNA). A: schematic diagram of spliced leader (SL) RNA and pre-rRNA molecules involved in putative pre-rRNA *trans*-splicing. For the SL RNA, the boxed ME represents the mini-exon, the black line shows the intron and the 5'SS is the GU donor splice site. For the pre-rRNA, the +1 represents the transcription start site, the black line shows the 5'ETS, the 3'SS is the hypothetical 3'splice site, the boxed small subunit (SSU) is the rRNA, the black boxes represent the three SSU-conserved regions for trypanosomatids and A', A0, and A1 are known trypanosomatid pre-rRNA processing sites; B: schematic diagram of hypothetical *trans*-spliced pre-rRNA. Arrows indicate the SL forward primer for PCR and the complementary primers used for reverse transcription (RT) and two subsequent nested-PCRs (Rev1 and Rev2); C: white arrows indicate RT-PCR products obtained by semi-nested RT-PCR in the presence (+) or (-) absence of reverse transcriptase from total RNA of *Trypanosoma brucei* (*Tb*), *Trypanosoma cruzi* (*Tc*), *Leishmania (Leishmania) amazonensis* (*La*), and *Crithidia fasciculata* (*Cf*). Fermentas Ladder (100 bp) was used as fragment lengths marker (M) but, after cloning, the fragment lengths were determined by sequencing and are shown by the number (in bp) tagged by white arrows; D: trypanosomatid DNA amplification. Line 1: PCR products from trypanosomatid genomic DNA obtained by the same strategy as for RT-nested PCR from total RNA of *Tb*, *Tc*, *La* and *Cf*; 2: PCR using SL-forward and Rev2 primers; 3: PCR using SSU-derived S4 and S12 primers (see Materials and Methods).

an oligo(dT) adaptor primer that paired with forward 5'ETS-based primers (Fig. 1A). The electrophoresis of the second round of PCR produced a smear, but not a marked product (Fig. 1B). However, the sequencing of two cloned products from the third round of PCR identified polyadenylated 5'ETS molecules. The poly-A tail was added either 612 or 607 nt upstream of the mapped AG *trans*-splicing acceptor site (Fig. 1C).

SL-5'ETS molecule quantities can be modulated through trans-splicing or exosome inhibitors - Sinefungin is an S-adenosylmethionine analogue that inhibits SL RNA methylation and indirectly inhibits the *trans*-splicing mRNA maturation process in trypanosomatids (McNally & Agabian 1992). We prepared the total RNA of *L. (L.) amazonensis* promastigotes after a 30-min or 60-min treatment with the antibiotic and then performed a quantitative RT-PCR to quantify the number of SL-5'ETS molecules (Fig. 5A, B). For the qRT-PCR procedure, we used a reverse primer based on the SSU rRNA sequence

to produce cDNA (Fig. 5A). The PCR was performed with a nested primer in the SSU rRNA region (rev) paired to an SL-based primer (sl) (Fig. 5A). The amount of 5'ETS-SSU rRNA was also measured using the rev primer paired to an ETS-based forward primer (srta) (Fig. 5A). Although the precursor molecule 5'ETS-SSU was not affected by sinefungin, the amount of SL-5'ETS was reduced following antibiotic treatment (Fig. 5B), indicating that 5'ETS *trans*-splicing responds to sinefungin in a manner similar to mRNA. Interestingly, there was an increase of four orders of magnitude in the number of 5'ETS-SSU molecules compared to SL-5'ETS molecules, which explains the low levels of ETS-SSU observed with RT-PCR, PE and RNase protection experiments.

We also treated promastigotes of *L. (L.) amazonensis* with different concentrations of 5-FU. This drug is intracellularly converted into 5-fluoro-UMP, a uridine analogue that is not metabolised to pseudouridine, which is a modification required for effective rRNA processing,
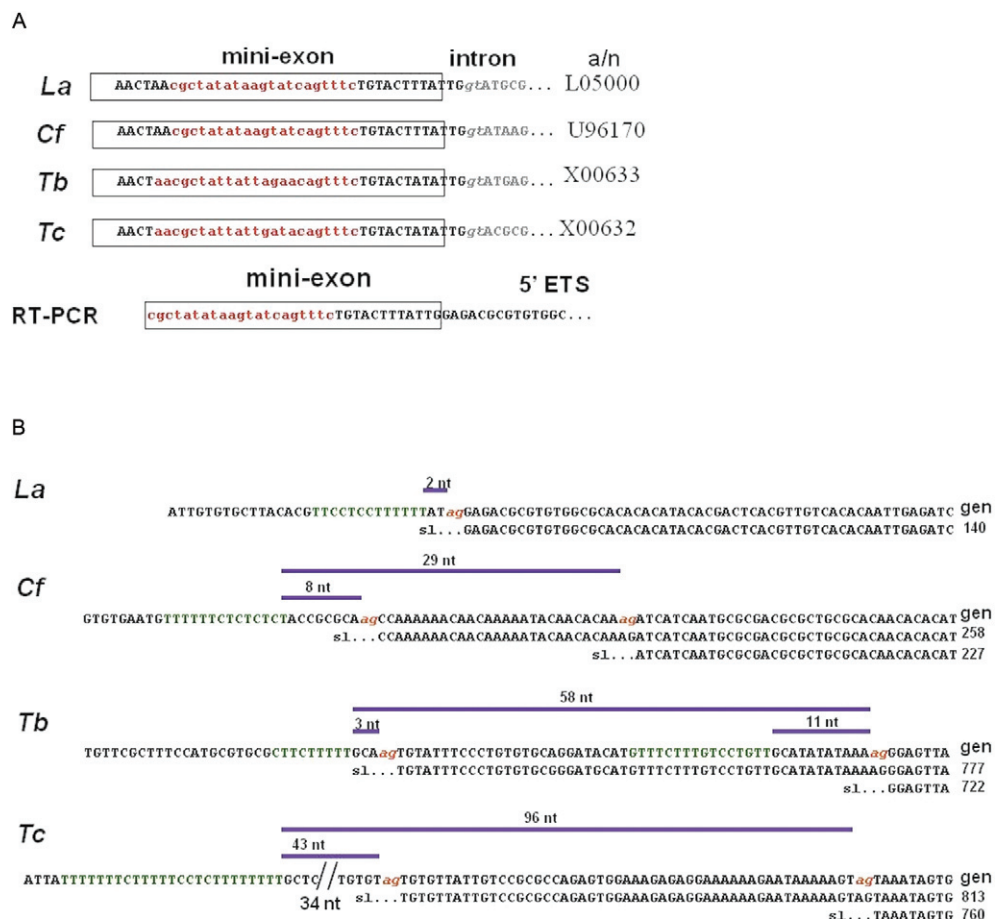

Fig. 3: species-specific spliced leader (SL) sequences and mapping *trans*-splicing signalling points. A: mini-exon sequences are boxed. SL forward primers for *Trypanosoma brucei* (*Tb*), *Trypanosoma cruzi* (*Tc*), *Leishmania (Leishmania) amazonensis* (*La*), and *Crithidia fasciculata* (*Cf*) are in red lowercase. The SL intron sequence is unboxed, with the 5' donor splice site in gray lowercase italics. GenBank accessions (a/n) are given; B: alignments of genomic (gen) and *trans*-spliced 5'ETS sequences of *La* (U21687, DQ907229), *Cf* (Y00055, DQ907227, DQ907228), *Tb* (AF416290, DQ907226, DQ907225, DQ907226) and *Tc* (DQ907222, DQ907223, DQ907224). The AG acceptor site is in red lowercase italics and the green sequence is the polypyrimidine tract. Real-time polymerase chain reaction product lengths in bp and positions where the mini-exon sequences, corresponding to organisms specific as in A (sl) were joined are indicated. Purple bars represent distances from polypyrimidine tracts to acceptor sites; SSU: small subunit.

resulting in exosome inhibition (Lenz et al. 1994, Fang et al. 2004). After these treatments, the amount of SL-ETS in the total RNA preparation was measured with qRT-PCR as described in Fig. 5A. Notably, treatment with 5-FU caused a dose-dependent accumulation of the ETS *trans*-spliced molecules (Fig. 5C); this effect was not observed for arginase mRNA, a control (Fig. 5C).

DISCUSSION

In this study, we detected a ribosomal RNA precursor containing a mini-exon sequence at the 5' end of the ETS region in four divergent species within the Trypanosomatidae family (Simpson et al. 2006), three heteroxenic pathogens and one monoxenic, suggesting that this processing is a common feature of all trypanosomatids. The SL (mini-exon) and 5'ETS joining obeyed the canonical

rules of SL addition to pre-mRNA in *trans*-splicing reactions. Moreover, the sequence analysis of pre-SSU rRNA *trans*-spliced products revealed that the 3' acceptor site was consistently located between the two most distantly known 5'ETS rRNA processing sites (A' and A1) (Fig. 2A) (Hartshorne & Toyofuku 1999, Schnare et al. 2000). Because the RT-PCR products encompassed a 5'ETS region linked to the 5' end of SSU, it is clear that *trans*-splicing occurred before A1 cleavage. This result does not exclude the possibility that SL addition can also occur after A1 processing, but the minimum acceptable substrate of the *trans*-splicing reaction is pre-SSU rRNA.

Recent studies using RNA-seq technology and a parallel sequencing SL trapping approach mapped a considerable number of splicing and polyadenylation sites in *T. brucei* pre-mRNAs (Kolev et al. 2010, Nilsson et al. 2010, Siegel et al. 2010). It is interesting to note that these studies revealed heterogeneity in pre-mRNA processing, with one to three alternative *trans*-splicing sites being
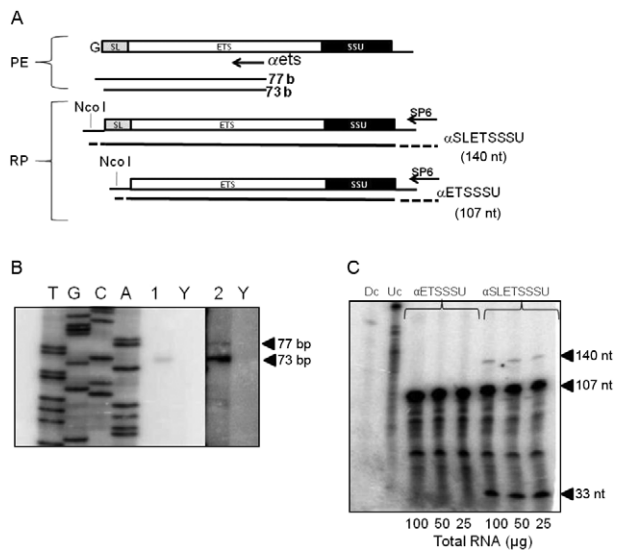

Fig. 4: primer extension or RNase protection detection of *trans*-spliced precursor ribosomal RNA (pre-rRNA). A: schematic representation of the method for the detection of spliced leader (SL)-external transcribed spacer (ETS) molecule by primer extension (PE). The relative position of the αETS primer (white box), products and the expected length (black line with nucleotides number) considering the stops caused by methylated G at CAP in the mini-exon sequence (gray box) are indicated below the ETS box. Schematic representation of antisense RNA probes produced by SP6 in vitro transcription corresponding to the complementary SL-5'ETS molecule (in C αSLETSSU) and to the complementary 5'ETS-SSU (αETSSU) precursor molecules for RNase protection (RP). Continuous line, with predicted length inside parenthesis, represent the complementary probe, broken lines indicate the plasmid sequences that also encompasses the probe, but are digested by RNase; B: denaturing PAGE analysis of the primer extension product (1: less exposed; 2: more exposed). Y is the lane containing yeast RNA as control. TGCA are the lanes of the sequencing reaction used as a fragment length marker. Arrows indicate the length of the detected products; C: different amounts of total *Leishmania (Leishmania) amazonensis* RNA (Lanes 100, 50 and 25 μg) were submitted to RNase protection analysis as described in Materials and Methods. The probes are as represent in A. Dc and Uc corresponds to Lanes containing the probe hybridized to control RNA (yeast) and the undigested probe, respectively. The length of digested products was estimated by the non-labelling RNA marker migration and are indicated in nt. SSU: small subunit.
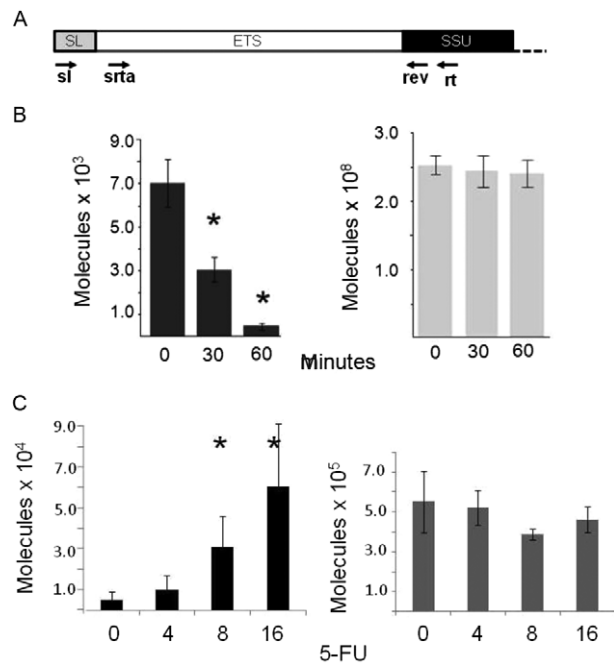

Fig. 5: sinefungin inhibits *trans*-splicing of precursor ribosomal RNA (pre-rRNA) and 5-fluorouracil (5-FU) treatment results in of *trans*-spliced pre-rRNAs accumulation. A: schematic representation of reverse transcription-polymerase chain reaction (RT-PCR) amplicons. Primer positions are indicated; B: RNA molecules by real-time PCR quantification of SL-5'ETSSSU using sl/rev primers (black) or 5'ETSSSU using srta/rev primers (light gray) after treatment of *Leishmania (Leishmania) amazonensis* with 1 μg/mL sinefungin for indicated times. Values represent the means [+/- standard deviation (SD)] of two independent experiments in duplicate. The difference compared to the control is statistically significant (asterisks mean $p < 0.05$); C: RNA molecules by RT-PCR quantification of SL-5'ETSSU (black) and arginase mRNA (dark gray) after treatment of *L. (L.) amazonensis* with the indicated concentration of 5-FU, in μM, for 90 min. The values represent the means (+/- SD) of two independent experiments in duplicate. The difference compared to the control is statistically significant (asterisks mean $p < 0.05$). ETS: external transcribed spacer; SL: spliced leader; SSU: small subunit.

used for each pre-mRNA and a higher number of poly-adenylation sites (Siegel et al. 2010). In another study, Helm et al. (2008) detected alternative *trans*-splicing in the pre-mRNA transcribed from the coding gene sequence of luciferase integrated in a reverse orientation in the rDNA *locus* of *T. brucei* (Helm et al. 2008). The deep sequencing of poly (A)-enriched SL-containing cDNA libraries of *T. brucei* found SL-tRNAsec hybrid molecules, which are polycistronically transcribed by RNA pol II in these organisms (Aeby et al. 2010). Taken together, these observations indicate that *trans*-splicing in those organisms may be a highly permissive process, most likely assuring the expression of key products for parasite survival (Kolev et al. 2010). Keeping this *trans*-splicing permissivity in mind, our findings are not surprising, although this is the first time that mini-exon sequences have been detected in an rRNA precursor molecule. However, it is surprising that *trans*-splicing signalling is conserved in all of the hundreds of rDNA copies in each of the four trypanosomatids studied. This sequence conservation limits the hypothesis of the existence of an RNA pol II transcription leak in the first copy of the rRNA gene tandem array, which could explain the low-copy-number SL-5'ETS molecule. Therefore, it is a biological challenge to elucidate a physiological meaning for this pre-rRNA processing that would explain the selective pressure necessary to preserve the conservation of *trans*-splicing signalling in rDNA.

It is unlikely that the *trans*-spliced pre-SSU rRNA molecule codes for functional peptides; however, the generated RNA molecules could have a regulatory function, as has been suggested for the highly abundant *L. (L.) major* ORF-less *trans*-spliced RNAs (Iribar & Cruz 2002). Alternatively, the highly structured *trans*-spliced 5'ETS transcripts with prematurely terminated ORFs could be processed through an RNA surveillance mechanism (Houseley et al. 2006). Thus, these organisms could use a powerful default RNA *trans*-splicing pathway to process the most abundant structural RNAs in the cell. Nuclear-localised orthologues of the *Saccharomyces cerevisiae* 5'--> 3'exoribonucleases XRNA and XRND have been identified in *T. brucei* and could participate in this processing (Li et al. 2006).

We used sinefungin and 5-FU to modulate the concentration of SL-5'ETS molecules in *L. (L.) amazonensis* promastigotes, which confirmed that these molecules were present as naturally occurring intermediates. A low copy number of *trans*-spliced 5'ETS molecules may be the result of a highly efficient degradation pathway for molecules tagged for degradation. Blocking SL addition with sinefungin treatment led to a rapid decay of *trans*-spliced 5'ETS. Regarding treatment with 5-FU, it is interesting to note that in addition to exosome inhibition, this drug can generate structurally aberrant RNAs by inhibiting the conversion of uridine to pseudouridine and preventing the degradation of highly structured RNA molecules (Allmang et al. 2000). This mechanism could also explain the observed accumulation of SL-5'ETS molecules when parasites were treated with 5-FU. Because 5-FU did not alter the expression of arginase mRNA, we can say that SL-5'ETS is not a by-product of *trans*-splicing.

Previously, polyadenylated SLRNA molecules were detected in *T. brucei* and *Leishmania (Leishmania) donovani* and implicated in the control of the levels of mature SLRNA (Pelle & Murphy 1993), (Lamontagne & Papadopoulou 1999). Furthermore, RT-PCR experiments showed that more than half of the rRNA fragment precursors (the furthest 3' region of the fragmented large subunit rRNA) were polyadenylated in both *L. (V.) braziliensis* and *L. (L.) donovani* (Decuypere et al. 2005). Recently, a *T. brucei* orthologue of the yeast putative RNA helicase Mtr4p and a non-canonical poly-(A) polymerase were implicated in the polyadenylation-dependent quality control of nuclear rRNA transcripts (Cristodero & Clayton 2007, Etheridge et al. 2009). The polyadenylation of nuclear rRNA transcripts has been described in fungus (Fleischmann & Liu 2001, Fleischmann et al. 2004) and yeast (Kuai et al. 2004) and human cells (Slomovic et al. 2006). A poly (A)-stimulated degradation pathway in yeast was also described (LaCava et al. 2005). We also detected polyadenylated 5'ETS molecules in *Leishmania (Leishmania) amazonensis*, indicating that this event could participate in the generation of molecules tagged at their 3'ends for nuclease attack, as suggested for the other polyadenylated rRNA fragment precursors. However, we still cannot determine whether 5'ETS polyadenylation is dependent on 5'ETS *trans*-splicing.

In *Caenorhabditis elegans,* an organism that also presents SL *trans*-splicing (Blumenthal 2005), we could not detect homologous *C. elegans* SL-5'ETS molecules (not shown). *C. elegans* uses a complex RNA pol II containing a CTD domain (Garrido-Lecca & Blumenthal 2010). In other eukaryotes, this domain has been implicated in coupling transcription to capping, splicing and polyadenylation (Maniatis & Reed 2002). Although there are descriptions of non-canonical CTDs (Schumacker et al. 2003, Das & Bellofatto 2009), trypanosomatid RNA pol II lacks the classical CTD domain (Evers et al. 1989). Thus, the presence of *trans*-splicing signals in both protein-coding and rRNA gene transcription products in trypanosomatids could play a role in RNA processing in a manner independent of the RNA pol used. This hypothesis explains our initial observations of the successful expression of a reporter gene driven by the RNA pol I promoter in constructs in which no acceptor site was added (Tyler-Cross et al. 1995, Uliana et al. 1996, Floeter-Winter et al. 1997).

This study provides solid evidence that natural trypanosomatid RNA pol I-transcribed pre-rRNA is processed by SL addition *trans*-splicing. This RNA processing event could be the footprint of a gene expression regulation strategy that allowed trypanosomatids to survive without specialised RNA polymerases. Alternatively, the relaxed rules for the conservation of *trans*-splicing sites allowed for the evolution of trypanosomatids that use ancient spliceosome-catalysed RNA reactions to produce a broad spectrum of RNAs.

REFERENCES

Aeby E, Ullu E, Yepiskoposyan H, Schimanski B, Roditi I, Muhlemann O, Schneider A 2010. tRNASec is transcribed by RNA polymerase II in *Trypanosoma brucei* but not in humans. *Nucleic Acids Res 38*: 5833-5843.

Allmang C, Mitchell P, Petfalski E, Tollervey D 2000. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res 28*: 1684-1691.

Benz C, Nilsson D, Andersson B, Clayton C, Guilbride DL 2005. Messenger RNA processing sites in *Trypanosoma brucei. Mol Biochem Parasitol 143*: 125-134.

Blumenthal T 2005. *Trans*-splicing and operons. *WormBook*: 1-9.

Brun R, Schonenberger M 1979. Cultivation and *in vitro* cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. *Acta Trop 36*: 289-292.

Camargo EP 1964. Growth and differentiation in *Trypanosoma cruzi*. I. Origin of metacyclic trypanosomes in liquid media. *Rev Inst Med Trop Sao Paulo 6*: 93-100.

Clayton CE 2002. Life without transcriptional control? From fly to man and back again. *Embo J 21*: 1881-1888.

Clayton CE, Schwede A, Stewart M, Robles A, Benz C, Po J, Wurst M, Queiroz R, Archer S 2008. Control of mRNA degradation in trypanosomes. *Biochem Soc Trans 36*: 520-521.

Cristodero M, Clayton CE 2007. Trypanosome MTR4 is involved in rRNA processing. *Nucleic Acids Res 35*: 7023-7030.

Das A, Bellofatto V 2009. The non-canonical CTD of RNAP-II is essential for productive RNA synthesis in *Trypanosoma brucei. PLoS ONE 4*: e6959.

Davis RE 1996. Spliced leader RNA *trans*-splicing in metazoa. *Parasitol Today 12*: 33-40.

Decuypere S, Vandesompele J, Yardley V, De Donckeri S, Laurent T, Rijal S, Llanos-Cuentas A, Chappuis F, Arevalo J, Dujardin JC 2005. Differential polyadenylation of ribosomal RNA during post-transcriptional processing in *Leishmania. Parasitology 131*: 321-329.

Dietrich P, Soares MB, Affonso MH, Floeter-Winter LM 1993. The *Trypanosoma cruzi* ribosomal RNA-encoding gene: analysis of promoter and upstream intergenic spacer sequences. *Gene 125*: 103-107.

dos Santos MG, da Silva MFL, Zampieri RA, Lafraia RM, Floeter-Winter LM 2011. Correlation of *meta 1* expression with culture stage, cell morphology and infectivity in *Leishmania (Leishmania) amazonensis* promastigotes. *Mem Inst Oswaldo Cruz 106*: 190-193.

Douris V, Telford MJ, Averof M 2010. Evidence for multiple independent origins of *trans*-splicing in metazoa. *Mol Biol Evol 27*: 684-693.

Etheridge RD, Clemens DM, Gershon PD, Aphasizhev R 2009. Identification and characterization of nuclear non-canonical poly(A) polymerases from *Trypanosoma brucei. Mol Biochem Parasitol 164*: 66-73.

Evers R, Hammer A, Kock J, Jess W, Borst P, Memet S, Cornelissen AW 1989. *Trypanosoma brucei* contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. *Cell 56*: 585-597.

Fang F, Hoskins J, Butler JS 2004. 5-fluorouracil enhances exosome-dependent accumulation of polyadenylated rRNAs. *Mol Cell Biol 24*: 10766-10776.

Fleischmann J, Liu H 2001. Polyadenylation of ribosomal RNA by *Candida albicans. Gene 265*: 71-76.

Fleischmann J, Liu H, Wu CP 2004. Polyadenylation of ribosomal RNA by *Candida albicans* also involves the small subunit. *BMC Mol Biol 5*: 17.

Floeter-Winter LM, Souto RP, Stolf BS, Zingales B, Buck GA 1997. *Trypanosoma cruzi*: can activity of the rRNA gene promoter be used as a marker for speciation? *Exp Parasitol 86*: 232-234.

Garrido-Lecca A, Blumenthal T 2010. RNA polymerase II C-terminal domain phosphorylation patterns in *Caenorhabditis elegans* operons, polycistronic gene clusters with only one promoter. *Mol Cell Biol 30*: 3887-3893.

Hartshorne T, Toyofuku W 1999. Two 5'-ETS regions implicated in interactions with U3 snoRNA are required for small subunit rRNA maturation in *Trypanosoma brucei. Nucleic Acids Res 27*: 3300-3309.

Helm JR, Wilson ME, Donelson JE 2008. Different trans RNA splicing events in bloodstream and procyclic *Trypanosoma brucei. Mol Biochem Parasitol 159*: 134-137.

Houseley J, Lacava J, Tollervey D 2006. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol 7*: 529-539.

Huang J, Van der Ploeg LH 1991. Requirement of a polypyrimidine tract for *trans*-splicing in trypanosomes: discriminating the PARP promoter from the immediately adjacent 3' splice acceptor site. *Embo J 10*: 3877-3885.

Hug M, Hotz HR, Hartmann C, Clayton C 1994. Hierarchies of RNA-processing signals in a trypanosome surface antigen mRNA precursor. *Mol Cell Biol 14*: 7428-7435.

Iribar MP, Cruz AK 2002. Base compositional bias in *trans*-spliced sequences of unknown function in *Leishmania major. Exp Parasitol 100*: 1-5.

Kapler GM, Coburn CM, Beverley SM 1990. Stable transfection of the human parasite *Leishmania major* delineates a 30-kilobase region sufficient for extrachromosomal replication and expression. *Mol Cell Biol 10*: 1084-1094.

Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, Tschudi C 2010. The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathog 6*: e1001090.

Kuai L, Fang F, Butler JS, Sherman F 2004. Polyadenylation of rRNA in *Saccharomyces cerevisiae. Proc Natl Acad Sci USA 101*: 8581-8586.

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell 121*: 713-724.

Lamontagne J, Papadopoulou B 1999. Developmental regulation of spliced leader RNA gene in *Leishmania donovani* amastigotes is mediated by specific polyadenylation. *J Biol Chem 274*: 6602-6609.

LeBowitz JH, Smith HQ, Rusche L, Beverley SM 1993. Coupling of poly(A) site selection and *trans*-splicing in *Leishmania. Genes Dev 7*: 996-1007.

Lee MG, Van der Ploeg LH 1997. Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu Rev Microbiol 51*: 463-489.

Lenz HJ, Manno DJ, Danenberg KD, Danenberg PV 1994. Incorporation of 5-fluorouracil into U2 and U6 snRNA inhibits mRNA precursor splicing. *J Biol Chem 269*: 31962-31968.

Li CH, Irmer H, Gudjonsdottir-Planck D, Freese S, Salm H, Haile S, Estevez AM, Clayton C 2006. Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation. *Rna 12*: 2171-2186.

Liang XH, Haritan A, Uliel S, Michaeli S 2003. *trans* and *cis* splicing in trypanosomatids: mechanism, factors and regulation. *Eukaryot Cell 2*: 830-840.

Maniatis T, Reed R 2002. An extensive network of coupling among gene expression machines. *Nature 416*: 499-506.

Martinez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K, Myler PJ 2003. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell 11*: 1291-1299.

Mayer MG, Floeter-Winter LM 2005. Pre-mRNA *trans*-splicing: from kinetoplastids to mammals, an easy language for life diversity. *Mem Inst Oswaldo Cruz 100*: 501-513.

McNally Kp, Agabian N 1992. *Trypanosoma brucei* spliced-leader RNA methylations are required for *trans*-splicing *in vivo*. *Mol Cell Biol 12*: 4844-4851.

Medina-Acosta E, Cross GA 1993. Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. *Mol Biochem Parasitol 59*: 327-329.

Nilsen TW 2001. Evolutionary origin of SL-addition *trans*-splicing: still an enigma. *Trends Genet 17*: 678-680.

Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi I, Ochsenreiter T 2010. Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. *PLoS Pathog 6*: e1001037.

Pelle R, Murphy NB 1993. Stage-specific differential polyadenylation of mini-exon derived RNA in African trypanosomes. *Mol Biochem Parasitol 59*: 277-286.

Rudenko G, Chung HM, Pham VP, Van Der Ploeg LH 1991. RNA polymerase I can mediate expression of CAT and neo protein-coding genes in *Trypanosoma brucei*. *Embo J 10*: 3387-3397.

Schnare MN, Collings JC, Spencer DF, Gray MW 2000. The 28S-18S rDNA intergenic spacer from *Crithidia fasciculata*: repeated sequences, length heterogeneity, putative processing sites and potential interactions between U3 small nucleolar RNA and the ribosomal RNA precursor. *Nucleic Acids Res 28*: 3452-3461.

Schumacker MA, Lau AO, Johnson PJ 2003. Structural basis of core promoter recognition in a primitive eukaryote. *Cell 115*: 413-424.

Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GA 2009. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev 23*: 1063-1076.

Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GA 2010. Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. *Nucleic Acids Res 38*: 4946-4957.

Simpson AG, Stevens JR, Lukes J 2006. The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol 22*: 168-174.

Slomovic S, Laufer D, Geiger D, Schuster G 2006. Polyadenylation of ribosomal RNA in human cells. *Nucleic Acids Res 34*: 2966-2975.

Thomas S, Green A, Sturm NR, Campbell DA, Myler PJ 2009. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. *BMC Genomics 10*: 152.

Tyler-Cross RE, Short SL, Floeter-Winter LM, Buck GA 1995. Transient expression mediated by the *Trypanosoma cruzi* rRNA promoter. *Mol Biochem Parasitol 72*: 23-31.

Uliana SR, Fischer W, Stempliuk VA, Floeter-Winter LM 1996. Structural and functional characterization of the *Leishmania amazonensis* ribosomal RNA promoter. *Mol Biochem Parasitol 76*: 245-255.

Uliana SR, Nelson K, Beverley SM, Camargo EP, Floeter-Winter LM 1994. Discrimination amongst *Leishmania* by polymerase chain reaction and hybridization with small subunit ribosomal DNA derived oligonucleotides. *J Eukaryot Microbiol 41*: 324-330.

Worthey EA, Martinez-Calvillo S, Schnaufer A, Aggarwal G, Cawthra J, Fazelinia G, Fong C, Fu G, Hassebrock M, Hixson G, Ivens AC, Kiser P, Marsolini F, Rickel E, Salavati R, Sisk E, Sunkin SM, Stuart KD, Myler PJ 2003. *Leishmania major* chromosome 3 contains two long convergent polycistronic gene clusters separated by a tRNA gene. *Nucleic Acids Res 31*: 4201-4210.

Zomerdijk JC, Kieft R, Borst P 1991. Efficient production of functional mRNA mediated by RNA polymerase I in *Trypanosoma brucei*. *Nature 353*: 772-775.