

Identification of SL addition *trans*-splicing acceptor sites in the internal transcribed spacer I region of pre-rRNA in *Leishmania (Leishmania) amazonensis*

Mario Gustavo Mayer¹, Lucile Maria Floeter-Winter^{2/+}

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

Trypanosomatidae is a family of early branching eukaryotes harbouring a distinctive repertoire of gene expression strategies. Functional mature messenger RNA is generated via the *trans*-splicing and polyadenylation processing of constitutively transcribed polycistronic units. Recently, *trans*-splicing of pre-small subunit ribosomal RNA in the 5' external transcribed spacer region and of precursor tRNA^{sec} have been described. Here, we used a previously validated semi-nested reverse transcription-polymerase chain reaction strategy to investigate internal transcribed spacer (ITS) I acceptor sites in total RNA from *Leishmania (Leishmania) amazonensis*. Two distinct spliced leader-containing RNAs were detected indicating that *trans*-splicing reactions occur at two AG acceptor sites mapped in this ITS region. These data provide further evidence of the wide spectrum of RNA molecules that act as *trans*-splicing acceptors in trypanosomatids.

Key words: RNA polymerase I - spliced leader RNA - spliceosome

Trypanosomatidae is a family of flagellated protozoans. A variety of *Trypanosomatidae* species from the *Leishmania* and *Trypanosoma* genera are responsible for various human diseases that lead to millions of deaths in developing countries. These organisms branch early in the phylogenetic tree and exhibit a single, conserved mechanism for gene expression (Simpson et al. 2006). Protein coding genes are polycistronically transcribed by RNA polymerase II from only a small number of promoter regions and individual nuclear messenger RNA (mRNA) is then generated *via* two coupled reactions: *trans*-splicing and polyadenylation (Lebowitz et al. 1993, Mayer & Floeter-Winter 2005). Although *trans*-splicing is a processing event that involves two independently transcribed RNA molecules, spliced leader (SL) RNA and pre-mRNA, the acceptor and donor sequences are identical to those in the majority of *cis*-spliced transcripts (Liang et al. 2003, Mayer & Floeter-Winter 2005). *Trans*-splicing processing adds a 39-nt mini-exon (ME) (SL) sequence derived from SL RNA to the 5' end of virtually all mRNAs. The SL RNA donor sites invariably consist of a GU dinucleotide. This processing occurs at transcribed intergenic regions and the acceptor sites are generally composed of an AG dinucleotide preceded by a polypyrimidine tract (Liang et al. 2003, Mayer & Floeter-Winter 2005, Orlando et al. 2007). The polyadeny-

lation of an upstream gene appears to be coupled to the *trans*-splicing of a second gene that is typically 200-500 nts downstream (Lebowitz et al. 1993, Hug et al. 1994). High-throughput analyses have shown that the locations of both the *trans*-splicing and the polyadenylation acceptor sites are quite variable in trypanosomatids and a considerable number of alternative sites have been described (Siegel et al. 2011).

Until recently, SL addition *trans*-splicing was exclusively associated with pre-mRNA processing and other RNA molecules were not considered to be involved in this processing mechanism. However, polyadenylation of non-protein coding genes has been detected in both prokaryotes and eukaryotes (Fleischmann et al. 2004). In trypanosomatids, it has been shown that ϵ ribosomal RNA (rRNA) fragment precursors (the 3'-most portion of the fragmented large rRNA subunit) are polyadenylated in *Leishmania (Viannia) braziliensis* and *Leishmania (Leishmania) donovani* (Decuyper et al. 2005). Moreover, in *Trypanosoma brucei*, a small number of RNA polymerase II-transcribed tRNA^{sec} precursor transcripts have been shown to exhibit SL ME sequences attached to their 5' ends (Aeby et al. 2010). Finally, we have identified pre-rRNA molecules that are processed by *trans*-splicing in the 5' external transcribed spacer (5'ETS) region in four trypanosomatid species [*T. brucei*, *Trypanosoma cruzi*, *Leishmania (Leishmania) amazonensis* and *Crithidia fasciculata*] and by 5'ETS polyadenylation in *L. (L.) amazonensis* (Mayer et al. 2012). These results show that such processing events are not restricted to pre-mRNAs in these organisms.

In this study, we showed that pre-rRNA *trans*-splicing is not restricted to the 5'ETS region. In addition, we demonstrated that the internal transcribed spacer (ITS) I region of transcripts in *L. (L.) amazonensis* is also a possible acceptor region for *trans*-splicing.

Financial support: FAPESP, CNPq

+ Corresponding author: lucile@ib.usp.br

Received 20 April 2012

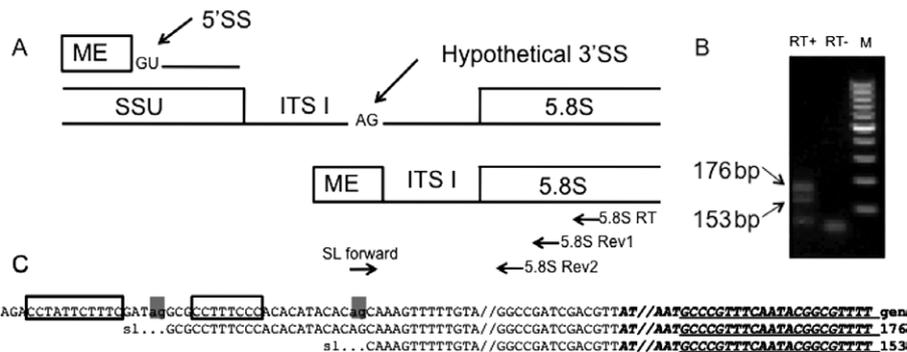
Accepted 28 July 2012

A reverse transcription-polymerase chain reaction (RT-PCR) strategy similar to that previously used for the detection of 5'ETS *trans*-splicing acceptor sites (Mayer et al. 2012) was employed for the detection of *L. (L.) amazonensis* SL addition *trans*-splicing in the ITS I region (A in Figure). *L. (L.) amazonensis* (MHOM/BR/73/M2269) promastigote total RNA was purified using the TRIzol reagent (Invitrogen, Grand Island, NY, USA). The total RNA was initially digested with RNase-free DNase (Fermentas Life Sciences, Burlington, Ontario, Canada) and re-extracted with the TRIzol LS reagent (Invitrogen). DNA-free total RNA (5 µg) was reverse transcribed (Superscript III, Invitrogen) using 2 pmol of an oligonucleotide primer complementary to 5.8S (5.8S RT primer, 5'AATAGGAAGCCAAGTCATCCATC3') at 50°C, according to the manufacturer's protocol. An identical mock reaction (with the reverse transcriptase omitted) was included to control for the presence of DNA in the total RNA preparations. The obtained cDNA (1/20 of the RT reaction volume) was amplified (using DNA *Taq* polymerase, Fermentas) with an *L. (L.) amazonensis* SL-specific forward primer (SL forward, 5'GCTATATAAGTATCAGTTTC3') and a 5.8S rRNA reverse primer (5.8S Rev 1, 5'GACACGTTATGTGAGCCGTTATCC3') (A in Figure). Finally, the PCR products were diluted 100X and amplified in a second round of PCR using the SL-specific primer and a nested 5.8S rRNA primer (5.8S Rev 2, 5'AAAACGCCGTATTGAAACGGGC3') (A in Figure). The cycling conditions for the first round of PCR were 94°C for 1 min, followed by 35 cycles of 94°C for 1 min, 52.2°C for 1 min and 72°C for 30 s, with a final extension at 72°C for 10 min in a 50 µL reaction volume containing 1.5 mM MgCl₂, 1X KCl buffer and 10 pmol of each primer. The cycling conditions for the second round of PCR were 94°C for 1 min, followed by 30 cycles of 94°C for 1 min, 54°C for 1 min and 72°C for 30 s, with a final extension at 72°C for 10 min, using the same reaction components as in the first round of PCR. The obtained PCR

products were analysed in a 1.5% agarose gel stained with ethidium bromide. Two bands between 100-200 bp that were not present in the mock reaction were identified in the PCR products, indicating the presence of two types of SL-ITS I *trans*-spliced molecules (B in Figure). These bands were purified using the Silica Bead Gel Extraction kit (Fermentas) and cloned into the pGEM-T Easy cloning vector (Promega, Madison, WI, USA) according to the manufacturer's protocol. The plasmids were isolated from selected clones *via* the alkaline lysis method and then sequenced using the T7 and SP6 sequencing primers and the Big Dye Terminator v3.1 Cycle Sequence Kit (Applied Biosystems, USA). Sequence analysis was performed with an ABI PRISM 3100 Genetic Analyzer.

The two fragments obtained from the semi-nested RT-PCR amplifications represent *trans*-spliced SL-ITS I molecules that have been processed at different regions to give rise to products that are 153 and 176 nts in length (B in Figure). The nucleotide sequences at the 5' ends of the two products were composed of the *L. (L.) amazonensis* SL forward primer followed by the downstream ME sequence (which was absent in the primer) linked to an ITS I sequence. These results demonstrate that the linking of the ME sequence with the ITS I sequence is not an artefact of mispriming and that the 39-nt ME sequence is added in its entirety to the ITS I sequence at two different regions. In the two analysed products, processing occurred immediately upstream of the canonical GU donor splice site from SL RNA, thereby confirming the GU donor site rule.

Identification of the ME acceptor sites was performed by comparing the obtained RT-PCR products with their corresponding genomic sequences (GenBank AB709806.1) (C in Figure). Analysis of the two types of SL-ITS I molecules showed that the canonical AG dinucleotide splice acceptor site was present. Moreover, we identified a polypyrimidine tract preceding each AG acceptor site. The polypyrimidine tracts associated with the



Detection of spliced leader (SL)-internal transcribed spacer (ITS) I molecules in *Leishmania (Leishmania) amazonensis*. A: diagrams of the mini-exon (ME) precursor RNA. A portion of the pre-ribosomal RNA encompassing the small subunit (SSU), ITS I and 5.8S regions and the processed RNA, demonstrating the semi-nested reverse transcription-polymerase chain reaction (RT-PCR) strategy used to detect ITS I *trans*-spliced molecules. RT was performed with the 5.8S RT primer and subsequently amplified with the specific *L. (L.) amazonensis* SL forward primer combined with the 5.8S Rev1 primer in the first round of PCR and with the 5.8S Rev2 primer in the second round of PCR; B: the semi-nested amplification products were analyzed by agarose gel electrophoresis. RT+ and RT- denote the presence or absence of reverse transcriptase in the RT reaction. Numbers beside the arrows correspond to the length of the obtained products, the sequences of which are indicated in part C [M: 100 bp ladder (Fermentas)]; C: comparison of the PCR product sequences with the genomic sequence (GenBank accession AB799806.1) (open box: polypyrimidine tract; grey box: AG *trans*-splicing acceptor sites; italics: 5.8S sequence; underlined: ITS Rev2 primer complementary sequence; SL: specifies the position of the inserted SL sequence detected in the PCR products).

ITS I splice acceptor sites were shorter than the tracts that have been found to be associated with the 5'ETS acceptor sites (Mayer et al. 2012). While the *L. (L.) amazonensis* 5'ETS acceptor site is 13 nts in length, the ITS I acceptor sites are 8 nts and 11 nts, with the longer site containing an A residue within its sequence. The distances between the acceptor sites and polypyrimidine tracts were less than 68 nts. These results show that the linking of the SL and ITS I molecules follows the same rules as the *trans*-splicing of pre-mRNA (Benz et al. 2005). Finally, analysis of the linkage points between the SL and ITS I sequences and the neighbouring sequences clearly indicated that an SL ME was added to the ITS I sequence *via* canonical *trans*-splicing. In addition, because we used reverse primers located in the 5.8S region, we can conclude that this processing event occurs prior to the processing between the ITS I and 5.8s RNA regions, suggesting that the *trans*-splicing at ITS I employs a pre-rRNA as a substrate, instead of a processed spacer.

Recently, we described SL addition *trans*-splicing at the 5'ETS in four trypanosomatid protozoan using a semi-nested RT-PCR approach (Mayer et al. 2012). In *L. (L.) amazonensis*, RNase protection and primer extension experiments validated the semi-nested PCR approach as a valuable tool for detecting true pre-rRNA *trans*-splicing acceptor sites. Here, we used the same approach to show that another pre-rRNA spacer region, ITS I, is also processed *via trans*-splicing in a manner following the same basic rules as pre-mRNA *trans*-splicing and 5'ETS pre-rRNA *trans*-splicing. Taken together, these results reinforce the value of this technique for exploring alternative acceptor molecules for *trans*-splicing. For example, the existence of SL-tRNA^{sec} hybrid molecules in *T. brucei* has recently been demonstrated (Aeby et al. 2010) through deep sequencing of poly A-enriched libraries and our semi-nested RT approach could be used to precisely identify the acceptor sites for these transcripts.

The detection of polyadenylated ϵ rRNA fragment precursors in both *L. (V.) braziliensis* and *L. (L.) donovani* (Decuyper et al. 2005), polyadenylated 5'ETS regions in *L. (L.) amazonensis*, *trans*-spliced *T. brucei* tRNA^{sec} precursors and *trans*-spliced 5'ETS regions in *T. brucei*, *T. cruzi*, *L. (L.) amazonensis* and *Crithidia fasciculata* (Mayer et al. 2012) shows that while polyadenylation and *trans*-splicing are essential for the generation of mature mRNAs, these processing events are not restricted to a specific class of RNA molecules in trypanosomatids.

We previously proposed a possible role for *trans*-splicing in a degradation pathway for the 5'ETS region transcribed by RNA polymerase I, although we could not rule out the less probable explanation that this processing event could be a side reaction resulting from spurious RNA pol II transcription of this region (Mayer et al. 2012). However, the detection of *trans*-splicing using another pre-rRNA spacer region, ITS I, as the substrate points to a role for this processing event in the clearance of spacer regions or at least in converting spacer regions into small fragments that are further processed by RNases. However, achieving an understanding of the exact purpose of this processing will require further investigation (Mayer et al. 2012). From an evolutionary perspective,

the identification of additional non-pre-mRNA *trans*-splicing events suggests that this processing is likely to be an ancient mechanism that is used in a broad range of RNA metabolic pathways. Because trypanosomatids regulate gene expression primarily at the post-transcriptional level *via* a single RNA metabolic pathway (Clayton 2002), *trans*-splicing could be employed as a general tool for dissecting various transcripts. Alternatively, spacer and intergenic regions may have been selected during evolution to include *trans*-splicing acceptor signals according to the following logic: if something is not prohibited, then it is "allowed" (Pirrotta 2002), which is in accordance with the variability observed in pre-mRNA *trans*-splicing acceptor sites (Siegel et al. 2011).

ACKNOWLEDGEMENTS

To RA Zampieri and Sandra M Muxel, for their technical assistance.

REFERENCES

- Aeby E, Ullu E, Yepiskoposyan H, Schimanski B, Roditi I, Muhlemann O, Schneider A 2010. tRNA^{sec} is transcribed by RNA polymerase II in *Trypanosoma brucei* but not in humans. *Nucleic Acids Res* 38: 5833-5843.
- Benz C, Nilsson D, Andersson B, Clayton C, Guilbride DL 2005. Messenger RNA processing sites in *Trypanosoma brucei*. *Mol Biochem Parasitol* 143: 125-134.
- Clayton CE 2002. Life without transcriptional control? From fly to man and back again. *Embo J* 21: 1881-1888.
- Decuyper 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.
- Fleischmann J, Liu H, Wu CP 2004. Polyadenylation of ribosomal RNA by *Candida albicans* also involves the small subunit. *BMC Mol Biol* 5: 17.
- 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.
- 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.
- Liang XH, Haritan A, Uliel S, Michaeli S 2003. *Trans* and *cis* splicing in trypanosomatids: mechanism, factors and regulation. *Eukaryot Cell* 2: 830-840.
- 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.
- Mayer MG, dos Santos MG, da Silva MFL, Floeter-Winter LM 2012. Footprints of a trypanosomatid RNA world: pre-small subunit rRNA processing by spliced leader addition *trans*-splicing. *Mem Inst Oswaldo Cruz* 107: 522-531.
- Orlando TC, Mayer MG, Campbell DA, Sturm NR, Floeter-Winter LM 2007. RNA polymerase I promoter and splice acceptor site recognition affect gene expression in non-pathogenic *Leishmania* species. *Mem Inst Oswaldo Cruz* 102: 891-894.
- Pirrotta V 2002. *Trans*-splicing in *Drosophila*. *Bioessays* 27: 988-991.
- Siegel TN, Gunasekera K, Cross GA, Ochsenreiter T 2011. *Trends Parasitol* 27: 434-441.
- Simpson AG, Stevens JR, Lukes J 2006. The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol* 22: 168-174.