RNA polymerase I promoter and splice acceptor site recognition affect gene expression in non-pathogenic *Leishmania* species

Tereza Cristina Orlando/*, Mário Gustavo Mayer, David A Campbell***, Nancy R Sturm***, Lucile Maria Floeter-Winter/+**

Departamento de Fisiologia, Instituto de Biociências, Rua do Matão, travessa 14/101, 05508-900 São Paulo, SP, Brasil *Departamento de Ciências Biológicas, Universidade Federal de Alfenas, Alfenas, MG, Brasil **Laboratório de Genética, Instituto Butantan, São Paulo, SP, Brasil ***Department of Microbiology, Immunology and Molecular Genetics, University of California at Los Angeles, Los Angeles, USA

Leishmania (Sauroleishmania) tarentolae *has biotechnological potential for use as live vaccine against visceral leishmaniasis and as a system for the over expression of eukaryotic proteins that possess accurate post-translational modifications. For both purposes, new systems for protein expression in this non-pathogenic protozoan are necessary. The ribosomal RNA promoter proved to be a stronger transcription driver since its use yielded increased levels of recombinant protein in organisms of both genera* Trypanosoma *or* Leishmania*. We have evaluated heterologous expression systems using vectors with two different polypyrimidine tracts in the splice acceptor site by measuring a reporter gene transcribed from* L. tarentolae *RNA polymerase I promoter. Our data indicate that the efficiency of chloramphenicol acetyl transferase expression changed drastically with homologous or heterologous sequences, depending on the polypyrimidine tract used in the construct and differences in size and/or distance from the AG dinucleotide. In relation to the promoter sequence the reporter expression was higher in heterologous lizard-infecting species than in the homologous* L. tarentolae *or in the mammalian-infecting* L. (Leishmania) amazonensis.

Key words: intergenic spacer - heterologous protein system - polypyrimidine tract - trans-splicing acceptor site

The non-pathogenic lizard-infecting protozoan *Leishmania* (*Sauroleishmania) tarentolae* is used currently as a safe model for mammal-infecting *Leishmania* species. It is also a useful alternative organism for the expression of accurately-modified heterologous eukaryotic proteins. A short doubling-time and simple nutrient requirements (Simpson & Braly 1970) make *L. tarentolae* an attractive host for high level production of heterologous proteins containing N-linked oligosaccharide modifications similar to those found in mammals (Breitling et al. 2002, Kushnir et al. 2005), and for generating isotopically-labeled protein for nuclear magnetic resonance studies (Niculae et al. 2006). In addition, *L. tarentolae* can be used as a live vaccine capable of eliciting protective immune response against *L. (Leishmania) donovani,* opening a new avenue for vaccination against other species involved in cutaneous leishmaniasis without the risk of disease development in humans (Breton et al. 2005). These preliminary results open up the perspective of dramatic improvements in the immunity response by vaccination with a nonpathogenic organism expressing an antigenic recombinant protein. Thus, efficient vectors and expression strategies are essential tools for the employment of *L. tarentolae* as recipient vessels for production of recombinant proteins with prophylactic promise.

A particular biological feature of Kinetoplastidae species is the polycistronic transcription of protein-coding genes followed by *trans*-splicing, and regulation of gene expression almost exclusively at the post-transcriptional level (Clayton 2002, Campbell et al. 2003, Mayer & Floeter-Winter 2005). Unlike most eukaryotes, almost all trypanosomatid protein-coding genes are transcribed constitutively as polycistronic pre-mRNAs from a few non-conventional RNA polymerase (RNA pol) II promoter regions (Palenchar & Bellofatto 2006). Two linked events in the intergenic region convert polycistronic to monocistronic mRNA. The *trans*-splicing of a pre-capped 39-nt spliced leader (a.k.a. the mini-exon) onto the pre-mRNA is initiated at an AG downstream of a polypyrimidine tract; the second event is the polyadenylation of the upstream mRNA. As a result, a large number of mature mRNAs are produced from a single transcription initiation event (Mayer & Floeter-Winter 2005). Since transcription and pre-mRNA processing are uncoupled events, exogenous protein-coding genes can be transcribed by the endogenous RNA pol I (Tyler-Cross et al. 1995, Uliana et al. 1996) or by an introduced polymerase, e.g. T7 RNA pol (Wirtz et al. 1998). This particularity allows the use of strong RNA pol I promoters for the production of recombinant proteins in a variety of kinetoplastids (Martinez-Calvillo et al. 1997, Downey and Donelson 1999). The ribosomal RNA promoter imparts stability to plasmid inheritance in *Leishmania* (Boucher et al. 2004) and may thus be beneficial to long-term protein expression.

Financial support: FAPESP, CNPq, NIH AI34536 (to DAC and NRS).
+Corresponding author: lucile@ib.usp.br
Received 18 July 2007
Accepted 22 October 2007

The generally-accepted paradigm is that ribosomal RNA promoters show species specificity in their ability to direct transcription by RNA pol I (Jacob & Ghosh 1999, Grummt 1999, Paule & White 2000). Studies of RNA pol I promoter sequences in *L. (L.) amazonensis* showed that these sequences were better recognized by heterologous host species *L. (L.) mexicana* and *L. (L.) major* (de Andrade Stempliuk & Floeter-Winter 2002). In this regard, we have proposed the existence of regulation carried out by a repressor in the region upstream of the promoter for the considered species. In addition, differential recognition of RNA processing signals could influence expression levels. The mRNA maturation signals in trypanosomatids consist of a polypyrimidine tract upstream of an AG dinucleotide, marking the point for mini-exon addition by *trans*-splicing (Mayer & Floeter-Winter 2005). It is also known that the size of the tract and its distance to the AG are essential to the efficiency of the process (Curotto de Lafaille et al. 1992, Benz et al. 2005, Siegel et al. 2005).

To evaluate the effect of the polypyrimidine tract and cross recognition of the RNA pol I promoter, we constructed two expression vectors using chloramphenicol acetyl transferase (CAT) as reporter, directed by the RNA pol I promoter region from *L. tarentolae,* but differing only in relation to the polypyrimidine tract in the splice acceptor site. The expression of CAT was also evaluated in relation to the host receiving both episomes. Pathogenic and nonpathogenic *Leishmania* species were transfected, with *L. tarentolae* the homologous host and other *Leishmania* species as heterologous hosts.

Initially, we constructed the vector pAB3.0stCAT, using an approximately 3-kb genomic fragment from the ribosomal DNA *L. tarentolae* intergenic spacer and external transcribed spacer (IGS/ETS) that encompasses the RNA pol I promoter region (Orlando et al. 2002), followed by a synthetic splice acceptor site (st) based on *L. major hsp70* gene (Uliana et al. 1996) to drive expression of the CAT gene reporter (Fig. 1). To determine the host range of promoter activity, transient transfection experiments and CAT activity assays were performed as described (Uliana et al. 1996) using lizard-infecting host cells *L. tarentolae* and *L. (S.) hoogstraali*, and the pathogenic mammalian infecting *L. amazonensis*. Surprisingly, no CAT activity was detected in any *Leishmania* species analyzed (Fig. 2). Interestingly, the same synthetic acceptor site 'st' sequence had already been efficiently used for *L. amazonensis* with the *L. amazonensis* ribosomal RNA promoter (Uliana et al. 1996, de Andrade Stempliuk & Floeter-Winter 2002). Although likely to be functional in the *L. major* hsp70 pre-mRNA, we suspected that the short polypyrimidine tract used in pAB3.0stCAT (3 from the ETS plus 13 from the st, see Fig. 1C-2) could be responsible for the failure of gene reporter expression by abolishing or lowering *trans*-splicing of the CAT gene. It is notable that in the 3' end of *L. amazonensis* ETS the number of pyrimidines is higher (9, instead of 3, see Fig. 1C-1), as possible explanation for the efficiency of that construct. We constructed another vector, pAB3.0tbCAT, replacing the synthetic acceptor site by the splice acceptor site (tb)

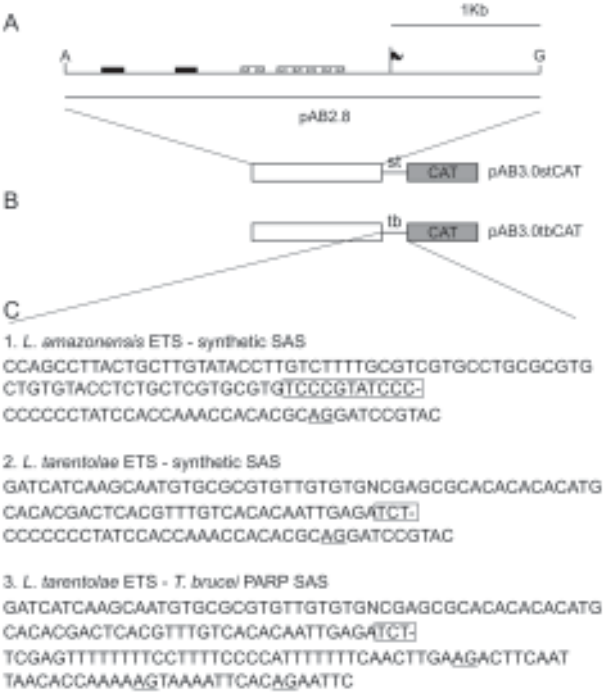

Fig. 1: plasmids used in this study, showing ETS sequence and splice acceptor sites in each construct. A: map of the *L. (Sauroleishmania) tarentolae* ribosomal DNA IGS/ETS region encompassing the RNA pol I promoter. Restriction enzymes flanking the AB fragment A, *Apa*I and G, *Bgl*II; the Flag maps the transcription start point; the black rectangles indicate inverted octanucleotide elements and stippled and white rectangles correspond to the 63bp repetitive elements (Orlando et al. 2002); B: schematic of plasmids: the two constructions containing the *L. tarentolae* IGS/ ETS (white box) upstream to the synthetic (st) or to *T. brucei* derived (tb) splice acceptor site followed by CAT coding sequence (gray box); C: sequence of ETS/acceptor site junction of the plasmid constructions. 1: construct pLaΔ14ASCAT: 3'end of *L. (Leishmania) amazonensis* ETS-synthetic splice acceptor site (Uliana et al., 1996); 2: construct pAB3.0stCAT: 3' end of *L. tarentolae* ETS-synthetic splice acceptor site; 3: construct pAB3.0tbCAT: 3'end of *L. tarentolae* ETS-*T. brucei* derived splice acceptor site (Sherman et al. 1991). Normal characters correspond to the ETS sequence; boxes indicate additional run of pyrimidines in the 3'ETS; letters after the dash indicate the polypyrimidine tract and splice acceptor site with AG indicated by underlining.

of *Trypanosoma brucei* EP1 (or PARP B) gene (Sherman et al. 1991), which contains a longer polypyrimidine tract (3 from the ETS plus 28 from tb, see Fig. 1C-3), positioned upstream the CAT gene. In addition, the tb sequence presents more than one AG that can assure the occurrence of *trans*-splicing, a fact already observed for other trypanosomatid genes (Mayer & Floeter-Winter 2005). Transient transfection of the same *Leishmania* species with pAB3.0tbCAT showed positive expression results (Fig. 2).

Since the most obvious difference between the plasmids is the length of the polypyrimidine tracts, we speculate that additional runs of pyrimidines (Fig 1B, boxed) in the *L. amazonensis* promoter region may have enabled expression in the homologous transfection. Moreover, CAT activity in heterologous lizard-infecting species *L. hoogstraali* was 4.5 fold higher than observed in the homologous species, indicating that expression driven

by the RNA pol I promoter is up-regulated in heterologous lizard-*Leishmania* species. In contrast, the CAT expression detected in the distant relative human-pathogenic *L. amazonensis* transfected with that construct was approximately half that in the homologous species (Fig. 2).

We extended the study of heterologous expression systems by transfecting two other lizard-infecting nonpathogenic *Leishmania* with construct pAB3.0tbCAT. *L. (S.) gymnodactyli* and *L. (S.) adleri* presented a CAT expression higher than the CAT level determined for the homologous *L. tarentolae* (Fig. 3). A more recent isolate of *L. tarentolae,* LEM 125, also showed an expression level similar to that of the *L. amazonensis*, which was lower than the level of expression detected in the homologous *L. (S.) tarentolae,* ATCC 30267strain (Fig. 3). This observation may indicate the presence of well-regulated RNA pol I machinery with lineage-specific transcription factors that could interact with the promoter region in each species. Another explanation could be that repressor recognition is stronger in closely-related isolates.

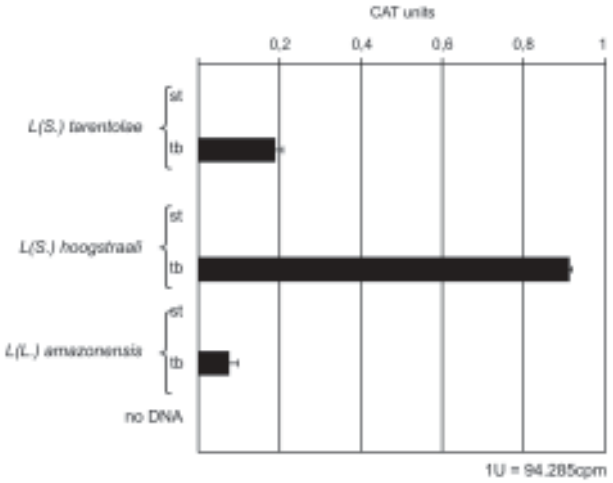

Fig. 2: CAT expression in *Leishmania* species transfected with constructions directed by *L. (S.) tarentolae* RNA pol l I promoter and different splice acceptor sites. Transfected organisms are indicated as well as the construction used pAB 3.0stCAT (st) or pAB 3.0tbCAT (tb), or the negative control with no DNA. Values are related to 1 U of *E. coli* CAT determined in the assays.

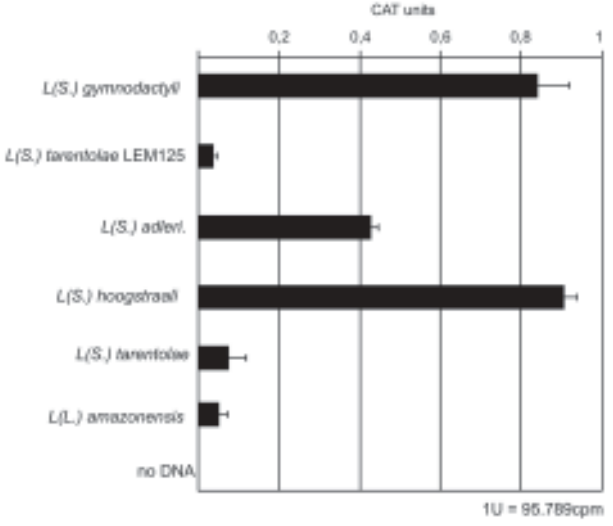

Fig. 3: CAT expression in *Leishmania* species transfected with pAB 3.0tbCAT. Organisms are indicated as well as the negative control with no DNA. Values are related to 1 U of CAT from *E. coli*.

Expression of the CAT reporter that we describe can be credited to a sum of features present in the vector. Our results show that acceptor site features can affect expression in nonpathogenic *Leishmania* species analyzed, and depending on the acceptor site choice, expression can be almost completely abolished even using a strong RNA pol I promoter region in a heterologous organism (e.g. Fig. 2, *L. hoogstraali*). The use of a heterologous promoter to drive transcription proved to be a good strategy for the lizard-infecting group, and correlates with our previous observation that RNA pol I driven expression in heterologous but closer species is higher than in homologous species, as for some species in the *Leishmania* group (Uliana et al. 1996, de Andrade Stempliuk & Floeter-Winter 2002). Although when considering heterologous but distantly related species, such as lizard and mammalian-infecters (Orlando et al. 2002, Yu et al. 2002), lack of recognition for the RNA pol I promoter sequence results in a lower level of expression than in the homologous species.

In conclusion, considering non-pathogenic *Leishmania* as important hosts for the production of proteins of biotechnological and immunological importance, the optimization of *trans*-splicing acceptor sites and heterologous RNA pol I-driven expression in closer species are important parameters for enhanced protein expression.

REFERENCES

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

Boucher N, McNicoll F, Laverdiere M, Rochette A, Chou MN, Papadopoulou B 2004. The ribosomal RNA gene promoter and adjacent cis-acting DNA sequences govern plasmid DNA partitioning and stable inheritance in the parasitic protozoan *Leishmania*. *Nucleic Acids Res 32*: 2925-2936.

Breitling R, Klingner S, Callewaert N, Pietrucha R, Geyer A, Ehrlich G, Hartung R, Muller A, Contreras R, Beverley SM, Alexandrov K 2002. Non-pathogenic trypanosomatid protozoa as a platform for protein research and production. *Protein Expr Purif 25:* 209-218.

Breton M, Tremblay MJ, Ouellette M, Papadopoulou B 2005. Live nonpathogenic parasitic vector as a candidate vaccine against visceral leishmaniasis. *Infect Immun 73:* 6372-6382.

Campbell DA, Thomas S, Sturm NR 2003. Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect 5:* 1231-1240.

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

Curotto de Lafaille MA, Laban A, Wirth DF 1992. Gene expression in *Leishmania*: analysis of essential 5' DNA sequences. *Proc Natl Acad Sci USA 89:* 2703-2707.

de Andrade Stempliuk V, Floeter-Winter LM 2002. Functional domains of the rDNA promoter display a differential recognition in *Leishmania. Int J Parasitol 32:* 437-447.

Downey N, Donelson JE 1999. Search for promoters for the GARP and rRNA genes of *Trypanosoma congolense. Mol Biochem Parasitol 104:* 25-38.

Grummt I 1999. Regulation of mammalian ribosomal gene transcription by RNA polymerase I. *Prog Nucleic Acid Res Mol Biol 62:* 109-154.

Jacob ST, Ghosh AK 1999. Control of RNA polymerase I-directed transcription: recent trends. *J Cell Biochem Suppl 32-33:* 41-50.

Kushnir S, Gase K, Breitling R, Alexandrov K 2005. Development of an inducible protein expression system based on the protozoan host *Leishmania tarentolae. Protein Expr Purif 42:* 37-46.

Martinez-Calvillo S, Lopez I, Hernandez R 1997. pRIBOTEX expression vector: a pTEX derivative for a rapid selection of *Trypanosoma cruzi* transfectants. *Gene 199:* 71-76.

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.

Niculae A, Bayer P, Cirstea I, Bergbrede T, Pietrucha R, Gruen M, Breitling R, Alexandrov K 2006. Isotopic labeling of recombinant proteins expressed in the protozoan host *Leishmania tarentolae. Protein Expr Purif 48:* 167-172.

Orlando TC, Rubio MA, Sturm NR, Campbell DA, Floeter-Winter LM 2002. Intergenic and external transcribed spacers of ribosomal RNA genes in lizard-infecting *Leishmania*: molecular structure and phylogenetic relationship to mammal-infecting *Leishmania* in the subgenus *Leishmania (Leishmania). Mem Inst Oswaldo Cruz 97:* 695-701.

Palenchar JB, Bellofatto V 2006. Gene transcription in trypanosomes. *Mol Biochem Parasitol 146:* 135-141.

Paule MR, White RJ 2000. Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res 28:* 1283-1298.

Sherman DR, Janz L, Hug M, Clayton C 1991. Anatomy of the parp gene promoter of *Trypanosoma brucei. Embo J 10:* 3379-3386.

Siegel TN, Tan KS, Cross GA 2005. Systematic study of sequence motifs for RNA *trans* splicing in *Trypanosoma brucei. Mol Cell Biol 25:* 9586-9594.

Simpson L, Braly P 1970. Synchronization of *Leishmania tarentolae* by hydroxyurea. *J Protozool 17:* 511-517.

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.

Wirtz E, Hoek M, Cross GA 1998. Regulated processive transcription of chromatin by T7 RNA polymerase in *Trypanosoma brucei. Nucleic Acids Res 26:* 4626-4634.

Yu MC, Orlando TC, Sturm NR, Zhou L, Saito RM, Floeter-Winter LM, Campbell DA 2002. Two distinct functional spliced leader RNA gene arrays in *Leishmania tarentolae* are found in several lizard *Leishmania* species. *Int J Parasitol 32:* 1411-1422.