

Published in final edited form as:

*Mol Biochem Parasitol*. 2010 February ; 169(2): 115–119. doi:10.1016/j.molbiopara.2009.10.003.

## The role of local secondary structure in the function of the trans splicing motif of *Brugia malayi*

Canhui Liu<sup>1</sup>, Chitra Chauhan<sup>1</sup>, and Thomas R. Unnasch<sup>2</sup>

Global Health Infectious Disease Program, Department of Global Health, College of Public Health, University of South Florida, Tampa, FL

### Abstract

A 7nt motif (the trans-splicing motif or TSM) was previously shown to be necessary and sufficient to direct trans-splicing of transgenic mRNAs in transgenic *B. malayi* embryos. Insertion of the TSM into two genes lacking a TSM homologue resulted in-trans splicing of transgenic mRNAs from one transgene but not the other, suggesting that local sequence context might affect TSM function. To test this hypothesis, constructs inserting the TSM into different positions of two *B. malayi* genes were tested for their ability to support trans-splicing of transgenic mRNAs. Transgenic mRNAs derived from constructs in which the insertion of the TSM did not result in a perturbation of the local predicted secondary structure were trans-spliced, while those in which the TSM perturbed the local secondary structure were not. These data suggest that local secondary structure plays a role in the ability of the TSM to direct trans-splicing.

### Keywords

filariasis; transfection; trans-splicing; spliced leader

mRNA 5' trans-splicing is a process found in a variety of organisms, including kinetoplasts, many nematodes, flatworms and a few species in the phylum chordata [1]. It involves the addition of a small leader RNA sequence, known as the spliced leader, to the 5' end of many pre-mRNAs. In nematodes, two types of trans-splicing have been noted. The most conserved form involves the addition of a 22 nt sequence, known as the spliced leader 1 (SL1), to the 5' end of the nascent mRNAs [2]. In *C. elegans* and some other nematodes, the downstream genes of operons are resolved through the addition of a distinct SL sequence, known as the SL2 [3]. However, in other nematodes, including the human filarial parasite *B. malayi*, SL2 trans-splicing appears not to exist, and all transcripts, including those located downstream in operons receive an SL1 [4].

Recently, a transient transfection system for the parasitic nematode *Brugia malayi* has been used to explore sequence requirements for trans splicing of exogenously introduced genes [5–7]. Studies employing constructs derived from the *B. malayi* gene encoding the parasite homologue of the 70 kDa heat shock protein (BmHSP70), whose native mRNA is trans-spliced, demonstrated that sequences located in the first intron of the native gene were necessary to

<sup>2</sup>Address of corresponding author: Thomas R. Unnasch, Global Health Infectious Disease Program, 3720 Spectrum Blvd, Suite 304, Tampa, FL 33612, Telephone: 1-813-974-0507, Telefax: 1-813-974-4838, tunnasch@health.usf.edu.

<sup>1</sup>These authors contributed equally to this work

**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.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

permit trans-splicing of transgenic messages transcribed from the BmHSP70 promoter [5]. Subsequent studies identified a 7nt motif encoded in the first intron of the BmHSP70 gene as the minimal motif necessary to permit trans-splicing of transgenic mRNAs, which has been designated the trans-splicing motif (TSM) [6]. Confirmatory evidence for the role of the TSM came from similar studies of constructs derived from the genes encoding the 12 kDa small subunit ribosomal protein gene (BmRPS12) and an abundant RNA binding protein (BmRBP1) from *B. malayi*. While the native mRNAs from both of these genes are trans-spliced, the first introns of both genes lack a TSM homologue. When the promoter, first exon, first intron and part of the second intron from both of these genes were inserted upstream of a luciferase reporter gene and the corresponding constructs transfected into *B. malayi* embryos, transgenic mRNAs were detected that were correctly cis-spliced and correctly translated [7]. However, these transgenic mRNAs were not trans-spliced, suggesting that the TSM was required to direct trans-splicing. In support of this, when the TSM was inserted into the intron of the construct containing the BmRPS12 promoter, first exon, first intron and part of the second exon (BmRPS12(E1-I1-E2)/luc), embryos transfected with this construct produced mRNA that was trans-spliced [7]. This suggested that the TSM was capable of transforming an mRNA that was not normally trans-spliced into one that was trans-spliced. However, when the TSM was introduced into the intron of the corresponding construct derived from the BmRBP1 gene (BmRBP1(E1-I1-E2)/luc) and the construct transfected into *B. malayi* embryos, the resulting transgenic mRNAs were not trans-spliced [7]. This finding suggested that the TSM, while capable of directing trans-splicing, might be functional in only certain contexts. In this regard, previous studies have shown that intron secondary structure can affect cis-splicing in some circumstances [8,9]. It was therefore possible that the secondary structure surrounding the BmHSP70 TSM may play a role in its activity in the trans-splicing process. The overall goal of the experiments presented here were to test this hypothesis.

The preparation of the two parental constructs used in these studies (BmRPS12(E1-I1-E2)/luc and BmRBP1(E1-I-E2)/luc) has been previously described [7]. BmRPS12 E1-I1-E2/luc consists of the 562 nt upstream of the initiating codon of the ORF of the BmRPS12 gene (encompassing the BmRPS12 promoter [10]), the first exon, first intron and 8 nt of the second exon cloned upstream of the firefly luciferase reporter gene in the reporter vector pGL3 basic (Promega, Madison WI). BmRBP1(E1-I-E2)/luc consists of the 463 nt upstream of the BmRBP1 ORF (containing the promoter [11]), the first exon, the single intron and 15 nt of the second exon of the BmRBP1 gene, cloned upstream of the firefly luciferase gene of pGL3 basic. Both constructs were engineered so that the residual ORFs derived from the first and second exons would be in frame with the luciferase reporter gene in transgenic mRNAs that were correctly cis-spliced.

The RNA structures of the first introns of RBP1 and RPS12 genes were analyzed using mfold version 3.2 [12]. The native intron sequences were extracted and run through the mfold webserver (<http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi>) using the default parameters for folding. The program gives several probable structures for a given sequence as the output. The structure with the lowest final free energy was considered the most stable, and was reported below. Modified intron structures containing the TSM were then generated in silico by inserting or replacing nucleotides at several positions in the native intron sequences, and the effect of the TSM addition on the intron structure was again analyzed through mfold.

TSM motifs were inserted into the parental constructs using the GeneTailor™ site-directed mutagenesis system (Invitrogen, Carlsbad, CA) as previously described [7]. The sequence of all mutated constructs was confirmed by DNA sequence analysis.

Isolated *B. malayi* embryos were biolistically transiently transfected as previously described [5]. In brief, embryos were isolated from gravid female parasites and transfected with the

experimental DNA driving the expression of firefly luciferase mixed with a constant amount of an internal standard, consisting of the BmHSP70 promoter fragment driving the expression of renilla luciferase. Transfected embryos were maintained in culture for 48 hours before being assayed for transgene activity. Firefly luciferase activity was normalized to the amount of renilla luciferase activity in each sample to control for variations in transfection efficiency. Firefly/renilla activity ratios for each sample were further normalized to the activity ratio found in embryos transfected in parallel with a construct consisting of the corresponding promoter alone driving the expression of the luciferase reporter. All constructs were assayed in triplicate.

RNA prepared from 48-hour transfected embryos was used as the template in hemi-nested RT-PCRs to detect evidence of trans-splicing, as previously described [7]. In brief, total RNA was purified from each batch of transfected embryos using Trizol (Invitrogen) following the manufacturer's instructions. A total of 1 µg of this purified RNA was used as a template in a single tube RT-PCR amplification reaction, employing the Qiagen One Step reagent (Qiagen, Valencia, CA), following the manufacturer's protocol. The one step reaction employed primers derived from the spliced leader (SL1) sequence and the luciferase gene encoded in the pGL3 basic vector (luc nc490) [7]. A total of 5 µl of this amplification reaction was used as a template in a 50 µl hemi-nested PCR reaction, employing Pfu Turbo DNA polymerase (Invitrogen, Carlsbad, California) using the buffer and conditions recommended by the manufacturer. The hemi-nested PCR reaction employed the SL primer and a nested primer derived from the luciferase gene (luc nc238) [7]. The amplification reactions were analyzed by agarose gel electrophoresis. The identities of all amplicons were confirmed by DNA sequence analysis.

As a first step in testing the hypothesis that local secondary structure in the introns played a role in this process, the predicted secondary structure of the native intron of BmRBP1 was compared to that of the intron into which the TSM had been inserted. These analyses revealed that the predicted secondary structure of the intron in the construct into which the TSM has been inserted at position 405 relative to the 5' end of the first intron (Construct BmRBP1(E1-I1(TSM 405)-E2)/luc) was substantially different from that of the native intron (Figure 1, Panel A). When BmRBP1(E1-I1(TSM 405)-E2)/luc was transfected into *B. malayi* embryos, the amount of luciferase reporter activity seen was not significantly different from that seen in embryos transfected with the native construct BmRBP1(E1-I1-E2)/luc (Figure 1, Panel B). This finding suggested that transcription and cis-splicing in BmRBP1(E1-I1(TSM 405)-E2)/luc transfected embryos had not been affected by the insertion of the TSM. However, as previously shown [7], hemi-nested SL mediated RT-PCR analysis of the transgenic mRNAs produced no product, suggesting that the transgenic mRNA was not trans-spliced (Figure 1, Panel C).

Secondary structure models were then constructed of putative introns into which the TSM was inserted at a number of different positions, in order to identify a position into which insertion of the TSM resulted in no large scale changes in the predicted secondary structure. Insertion of the TSM at position 189 downstream of the start of the intron appeared to result in a construct in which the local structure was predicted to remain basically unchanged (Figure 1, Panel A). A construct was therefore prepared in which the TSM was inserted at this position (construct BmRBP1(E1-I1(TSM 189)-E2)/luc) was prepared and transfected into *B. malayi* embryos. The luciferase reporter activity in embryos transfected with this construct was not significantly different from embryos transfected with either the native construct, or construct BmRBP1(E1-I1(TSM 405)-E2)/luc (Figure 1, Panel B). However, in contrast to either embryos transfected with the native construct or BmRBP1(E1-I1(TSM 405)-E2)/luc, hemi-nested SL mediated RT-PCR analysis of the transgenic mRNA from embryos transfected with BmRBP1(E1-I1(TSM 189)-E2)/luc produced a product suggesting that the transgenic mRNA was trans-spliced (Figure 1, Panel C). DNA sequence analysis of this amplicon revealed that it was derived from a correctly cis- and trans-spliced mRNA.

In contrast to BmRBP1, a construct derived from BmRPS12 into which the TSM was inserted at position 37 relative to the start of the intron (construct BmRPS12(E1-I1(TSM 37)-E2)/luc, when transfected into *B. malayi* embryos produced transgenic mRNA that was correctly cis- and trans-spliced ([7] and Figure 2, Panel C). An analysis of the predicted secondary structures of the native RPS12 intron and that of BmRPS12(E1-I1(TSM 37)-E2)/luc revealed that insertion of the TSM at position 37 did not substantially alter the predicted secondary structure of the intron (Figure 2, Panel A). This finding suggested that insertion of the TSM into the BmRPS12 intron in a position which disrupted the predicted secondary structure should result in transgenic mRNA that was not trans-spliced. An analysis was therefore conducted placing the TSM at different positions in the BmRPS12 first intron to identify a position where insertion of the TSM would substantially alter the predicted secondary structure of the intron. It was found that insertion of the TSM at position 21 relative to the start of the intron resulted in a substantial change to the predicted secondary structure (Figure 2, Panel A). A construct was prepared inserting the TSM into this position (BmRPS12(E1-I1(TSM 21)-E2)/luc), and this construct was transfected into *B. malayi* embryos. The level of luciferase reporter activity in these embryos was not significantly different from that seen in embryos transfected with the native construct BmRPS12(E1-I1-E2)/luc or with BmRPS12(E1-I1(TSM 37)-E2)/luc (Figure 2, Panel B). However, hemi-nested SL mediated RT-PCR analysis of the transgenic mRNA derived from these embryos did not produce a product, suggesting that the transgenic mRNA was not trans-spliced (Figure 2, Panel C).

Taken together, these data provide support for the hypothesis that local secondary structure plays a role in determining if the TSM will be capable of directing trans-splicing. Thus, when the TSM is inserted into positions that result in a substantial change to the secondary structure of the intron, the TSM was not capable of directing trans-splicing. However, when inserted into positions in which the predicted secondary structure was not substantially changed, the TSM was capable of directing trans-splicing.

The local secondary structural changes initiated by insertion of the TSM did not appear to affect cis-splicing of the introns, as the amount of reporter gene activity obtained from all constructs (those in which the TSM was active and those in which it was not) were not significantly different from that obtained from the parental constructs lacking the TSM. This would not be expected if cis-splicing was affected by the insertion of the TSM, as multiple in frame stop codons are present in each of the intron sequences, which would have resulted in pre-mature translational termination in non-cis spliced mRNAs. Termination of translation would have occurred following the first exon of the native gene in these constructs, a point that is upstream of the start of the luciferase reporter gene sequence (data not shown). In support of this, previous studies have demonstrated that the transgenic mRNAs derived from intron-containing constructs that do not contain a functional TSM and are therefore not trans-spliced are still correctly cis-spliced [7].

From the data presented above, it does not appear that the functionality of the TSM was associated with its location within a particular type of local secondary structure. For example, in BmRBP1(E1-I1(TSM 189)-E2)/luc, the TSM is present in a double stranded stem in the predicted structure, while in BmRPS12(E1-I1(TSM 37)-E2)/luc, it is present in a loop in the predicted structure. The TSM was capable of directing trans-splicing in both of these contexts, suggesting that it is functional in both single stranded and double stranded contexts.

The mechanism through which the TSM mediates trans-splicing remains obscure. Previous studies have conclusively demonstrated that trans-splicing involves the recruitment of specialized proteins to the spliceosome ribonucleoprotein complex [13,14]. It is tempting to hypothesize that the TSM is responsible for recruiting a similar trans-splicing factor to the spliceosome that promotes the trans-splicing process. Alternatively, it is possible that the TSM

acts through some sort of a small RNA mediator, such as an miRNA. In support of this hypothesis, mutation of any one of the 7 nt making up the TSM has been shown to eliminate TSM activity, a property that is consistent with an interaction with an miRNA [15]. Further experiments will be necessary to differentiate these possibilities.

## Acknowledgments

We would like to thank Dr. Julian Rayner for initially suggesting that secondary structure might play a role in the functionality of the TSM, and Dr. Naomi Lang-Unnasch for critical reading of the manuscript. This work received financial support from a grant from the United States National Institute of Allergy and Infectious Diseases (Project # R01AI048562) to TRU.

## Abbreviations

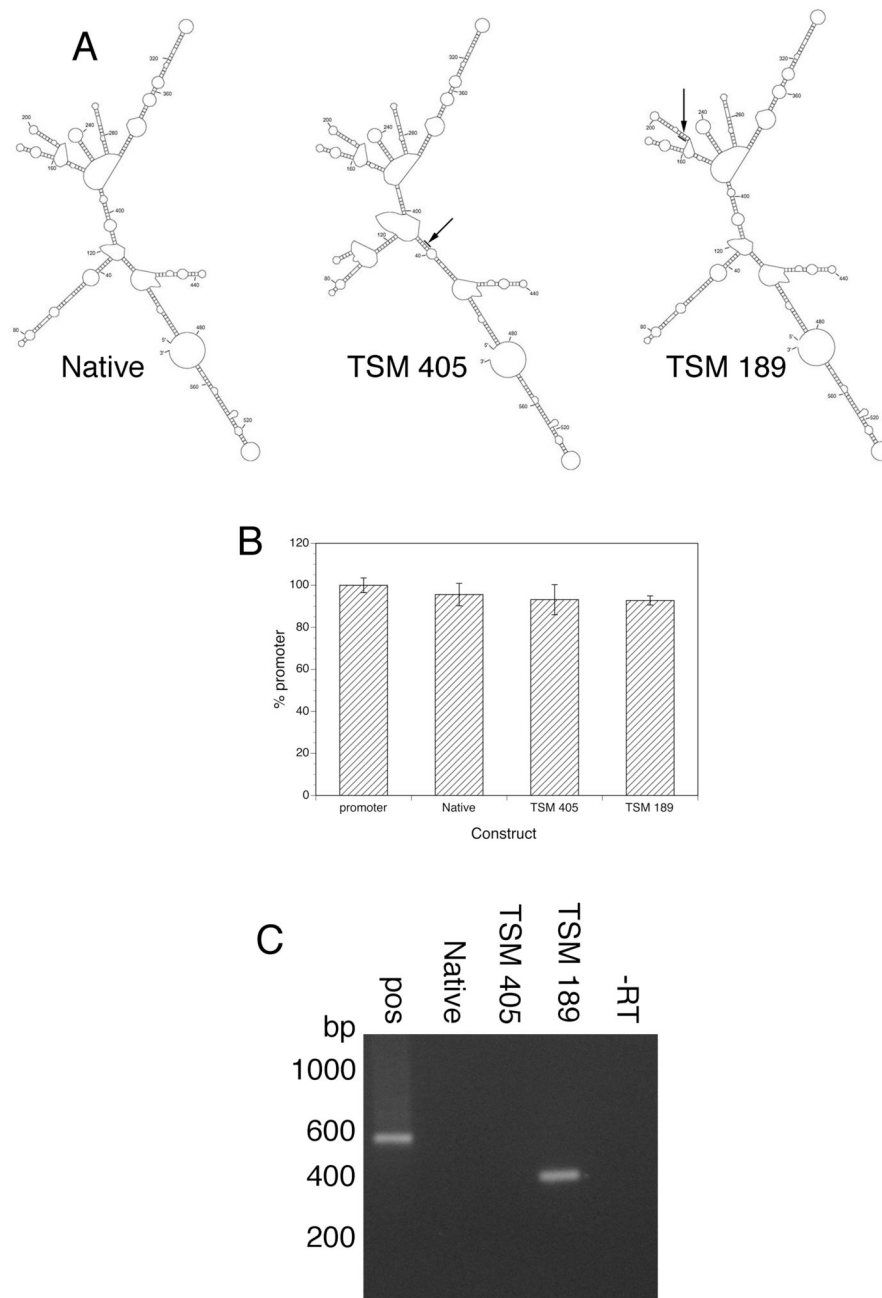
HSP70	70 kDa heat shock protein
RBP1	RNA binding protein 1
RPS12	ribosomal small subunit 12 kDa protein
nt	nucleotides
SL	spliced leader
TSM	trans-splicing motif

## References

- Hastings KE. Sl trans-splicing: Easy come or easy go? Trends Genet 2005;21:240–7. [PubMed: 15797620]
- Nilsen TW. Trans-splicing of nematode premessenger RNA. Annu Rev Microbiol 1993;47:413–40. [PubMed: 8257104]
- Blumenthal T, Evans D, Link CD, et al. A global analysis of *Caenorhabditis elegans* operons. Nature 2002;417:851–4. [PubMed: 12075352]
- Guiliano DB, Blaxter ML. Operon conservation and the evolution of trans-splicing in the phylum Nematoda. PLoS Genet 2006;2:e198. [PubMed: 17121468]
- Shu L, Katholi C, Higazi T, et al. Analysis of the *Brugia malayi* hsp70 promoter using a homologous transient transfection system. Mol Biochem Parasitol 2003;128:67–75. [PubMed: 12706798]
- Higazi TB, Unnasch TR. Intron encoded sequences necessary for trans splicing in transiently transfected *Brugia malayi*. Mol Biochem Parasitol 2004;137:181–84. [PubMed: 15279965]
- Liu C, de Oliveira A, Higazi TB, et al. Sequences necessary for trans-splicing in transiently transfected *Brugia malayi*. Mol Biochem Parasitol 2007;156:62–73. [PubMed: 17727976]
- McAlinden A, Havlioglu N, Liang L, et al. Alternative splicing of type ii procollagen exon 2 is regulated by the combination of a weak 5' splice site and an adjacent intronic stem-loop cis element. J Biol Chem 2005;280:32700–11. [PubMed: 16076844]
- Glazov EA, Pheasant M, Nahkuri S, et al. Evidence for control of splicing by alternative RNA secondary structures in dipteran homothorax pre-mrna. RNA Biol 2006;3:36–9. [PubMed: 17114937]
- de Oliveira AD, Katholi CR, Unnasch TR. Characterization of the promoter of the *Brugia malayi* 12kda small subunit ribosomal protein (rps12) gene. Int J Parasitol 2008;38:1111–19. [PubMed: 18364245]
- Higazi TB, DeOliveira A, Katholi CR, et al. Identification of elements essential for transcription in *Brugia malayi* promoters. J Mol Biol 2005;353:1–13. [PubMed: 16154590]
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–15. [PubMed: 12824337]
- Denker JA, Zuckerman DM, Maroney PA, et al. New components of the spliced leader RNP required for nematode trans-splicing. Nature 2002;417:667–70. [PubMed: 12050671]

14. MacMorris M, Kumar M, Lasda E, et al. A novel family of *C. Elegans* snRNPs contains proteins associated with trans-splicing. *RNA* 2007;13:511–20. [PubMed: 17283210]
15. Carthew RW, Sontheimer EJ. Origins and mechanisms of mirnas and siRNAs. *Cell* 2009;136:642–55. [PubMed: 19239886]



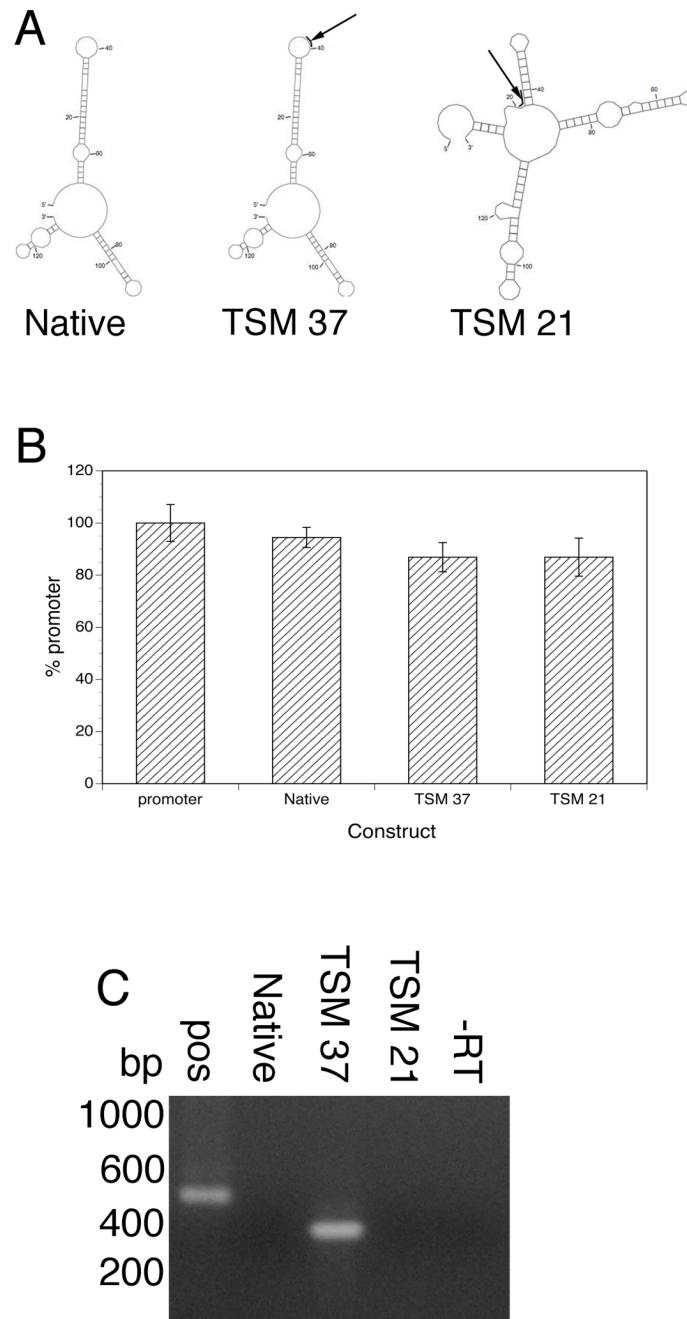


**Figure 1.**

The effect of position of the TSM on its ability to direct trans-splicing in the BmRBP1 gene: Panel A: The predicted secondary structure of the native BmRBP1 intron and introns containing TSM insertions at different positions. Secondary structure predictions were carried out using the mfold algorithm as described in the text. Arrows and bold lines indicate the position of the TSM insert in BmRBP1(E1-I1(TSM 405)-E2)/luc (TSM 405) and in BmRBP1(E1-I1(TSM 189)-E2)/luc (TSM 189). Panel B: luciferase activity in embryos transfected with BmRBP1 constructs containing the TSM inserted at different positions in the first intron. Promoter = activity from embryos transfected with a construct containing the BmRBP1 promoter alone driving the expression of the luciferase reporter. Native = activity in embryos transfected with

parental construct BmRBP1(E1-I1-E2)/luc. TSM 405 = activity in embryos transfected with BmRBP1(E1-I1(TSM 405)-E2)/luc. TSM 189 = activity in embryos transfected with BmRBP1(E1-I1(TSM 189)-E2)/luc. Bars are the mean and error bars the standard deviation of three independent transfections. Panel C: Hemi-nested SL-mediated RT-PCR analysis of the transgenic mRNAs isolated from embryos transfected with the constructs shown in Panel B. pos = RT-PCR positive control (embryos transfected with BmHSP70 (E1-I1-E2)/luc [7]). Native = RT-PCR of embryos transfected with parental construct BmRBP1(E1-I1-E2)/luc. TSM 405 = RT-PCR of embryos transfected with BmRBP1(E1-I1(TSM 405)-E2)/luc. TSM 189 = RT-PCR of embryos transfected with BmRBP1(E1-I1(TSM 189)-E2)/luc. -RT = minus reverse transcriptase negative control.





**Figure 2.**

The effect of position of the TSM on its ability to direct trans-splicing in the BmRPS12 gene: Panel A: The predicted secondary structure of the native BmRBP1 intron and introns containing TSM insertions at different positions. Secondary structure predictions were carried out using the mfold algorithm as described in the text. Arrows and bold lines indicate the position of the TSM insert in BmRSP12(E1-I1(TSM 37)-E2)/luc (TSM 37) and in BmRPS12(E1-I1(TSM 21)-E2)/luc (TSM 21). Panel B: luciferase activity in embryos transfected with BmRPS12 constructs containing the TSM inserted at different positions in the first intron. Promoter = activity from embryos transfected with a construct containing the BmRPS12 promoter alone driving the expression of the luciferase reporter. Native = activity in embryos transfected with

parental construct BmRPS12(E1-I1-E2)/luc. TSM 37 = activity in embryos transfected with BmRPS12(E1-I1(TSM 37)-E2)/luc. TSM 21 = activity in embryos transfected with BmRPS12(E1-I1(TSM 21)-E2)/luc. Bars are the mean and error bars the standard deviation of three independent transfections. Panel C: Hemi-nested SL-mediated RT-PCR analysis of the transgenic mRNAs isolated from embryos transfected with the constructs shown in Panel B. pos = RT-PCR positive control (embryos transfected with BmHSP70 (E1-I1-E2)/luc). Native = RT-PCR of embryos transfected with parental construct BmRPS12(E1-I1-E2)/luc. TSM 37 = RT-PCR of embryos transfected with BmRPS12(E1-I1(TSM 37)-E2)/luc. TSM 21 = RT-PCR of embryos transfected with BmRPS12(E1-I1(TSM 21)-E2)/luc. -RT = minus reverse transcriptase negative control.