THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 27, Issue of July 2, pp. 19361–19367, 1999

Printed in U.S.A.

The Role of Intron Structures in *trans*-Splicing and Cap 4 Formation for the *Leishmania* Spliced Leader RNA*

(Received for publication, September 4, 1998, and in revised form, March 17, 1999)

Nancy R. Sturm‡ and David A. Campbell§

From the Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, California 90095-1747

A 39-nucleotide leader is *trans*-spliced onto all trypanosome nuclear mRNAs. The precursor spliced leader RNA was tested for *trans*-splicing function *in vivo* by mutating the intron. We report that in *Leishmania tarentolae* spliced leader RNA 5′ modification is influenced by the primary sequence of stem-loop II, the Sm-binding site, and the secondary structure of stem-loop III. The sequence of stem-loop II was found to be important for cap 4 formation and splicing. As in *Ascaris*, mutagenesis of the bulge nucleotide in stem-loop II was detrimental to *trans*-splicing. Because restoration of the *L. tarentolae* stem-loop II structure was not sufficient to restore splicing, this result contrasts the findings in the kinetoplastid *Leptomonas*, where mutations that restored stem-loop II structure supported splicing. Methylation of the cap 4 structure and splicing was also dependent on both the Sm-binding site and the structure of stem-loop III and was inhibited by incomplete 3′ end processing. The critical nature of the *L. tarentolae* Sm-binding site is consistent with its essential role in the *Ascaris* spliced leader RNA, whereas in *Leptomonas* mutation of the Sm-binding site and deletion of stem-loop III did not affect *trans*-splicing. A pathway for *Leishmania* spliced leader RNA processing and maturation is proposed.

Kinetoplastid nuclear gene expression is dependent on the *trans*-splicing process. The common substrate for all *trans*-splicing reactions is the spliced leader (SL)[1] RNA, also known as the mini-exon derived RNA, whose first 39 nt constitute the 5′ ends of both mono- and polycistronically synthesized mRNAs (1). The polycistronic pre-mRNAs require *trans*-splicing to acquire the specialized "cap 4" structure on the SL RNA. The cap 4 consists of a ^{7}mG attached to the first nucleotide (2), in addition to methylation of the first four and sixth nucleotides of the SL RNA (3–5). These modifications are made to the primary SL RNA and spliced onto the mRNA as part of the 39-nt exon. The cap 4 may have roles in mRNA *trans*-splicing, transport, stability and translation.

The SL RNA contains two functional domains as follows: the exon and the intron or snRNA-like domain (6). The exon sequence is conserved among 38 different members of the order Kinetoplastida (7). Positions 1–9 and 20–39 of the exon are nearly identical, whereas positions 10–19 are relatively heterogeneous and characteristically A/T-rich. This conservation cannot be ascribed to internal promoter location in *Leishmania* (8, 9), as found in *Ascaris* (10). It was surprising that mutations within positions 20–39 permitted accurate *trans*-splicing in *Leishmania tarentolae* and did not lower splicing efficiency (11) because these results contrasted with findings in *Leptomonas* (12). Thus, the results in *L. tarentolae* more closely resemble the findings in worms as follows: in *Ascaris*, exon sequences are not necessary for *trans*-splicing *in vitro* (13); in *Caenorhabditis elegans*, length, primary sequence, and composition of the SL are not critical parameters for essential embryonic function, although certain nucleotides may be essential for *in vivo* splicing of the SL1 RNA (14, 15).

The primary sequence of the SL RNA intron is not conserved among the trypanosomatids (7); however, the secondary structure is consistent (16). This structure has been confirmed by physical-chemical and enzymatic studies (17, 18) and examined by mutagenesis (12). An equivalent structure is also conserved in the nematode SL RNAs (16, 19). The intron contains a putative Sm-binding site (16), an element found in the small nuclear RNAs of higher eukaryotes but apparently lacking in all U-RNAs of kinetoplastids (20) except U5 RNA (21, 22). The Sm-binding element is required for SL RNA *trans*-splicing in *Ascaris* (23) but not in *Leptomonas* (12).

We demonstrated recently that the T tract downstream of the SL RNA gene is a transcription termination element and that staggered T tract termination products are processed via nucleolytic cleavage to the base of stem-loop III (24). The signals for 3′-processing begin in the Sm-binding site at position 76 and include the structure, but not content, of stem-loop III. Studies in *Leptomonas seymouri* demonstrated that mutation of a variety of elements in the intron was acceptable for *trans*-splicing (12), whereas in *Leptomonas collosoma* the loop portions of stem-loops II and III were tolerant to insertions but not to replacement with the *Trypanosoma brucei* intron (25). By contrast, the bulge of stem-loop II was critical for *trans*-splicing in *Ascaris* (26).

In this paper we report that methylation of nucleotides in the cap 4 structure of the *Leishmania* SL RNA is influenced by formation of stem-loop III, the Sm-binding site, and specific sequences in stem-loop II. The methylation of the cap 4 structure correlates with correct 3′ end formation; defects in 3′ end processing and cap 4 formation result in failure of the mutated SL RNA to undergo *trans*-splicing. However, correct maturation of the SL RNA is not sufficient to obtain a positive splicing phenotype since mutation of nucleotides in the stem I region of the intron can also result in loss of function. Our data from *L. tarentolae* broadly reflect the results obtained *in vitro* in the nematode *Ascaris*, where nucleotides in stem-loop II and the

* This work was supported in part by National Institutes of Health Grant AI34536. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Postdoctoral trainee supported by Microbial Pathogenesis Training Grant 2-T32-AI-07323.

§ To whom correspondence should be addressed: Dept. of Microbiology and Immunology, UCLA School of Medicine, 10833 Le Conte Ave., Los Angeles, CA 90095-1747. Tel.: 310-825-4195; Fax: 310-206-3865; E-mail: dc@ucla.edu.

[1] The abbreviations used are: SL, spliced leader; tSL, tagged SL; Arl, ADP-ribosylation factor-like; PCR, polymerase chain reaction; RT, reverse transcriptase; TV, transversion; WT, wild type; bp, base pair; kb, kilobase pair; nt, nucleotide(s); snRNA, small nuclear RNA.

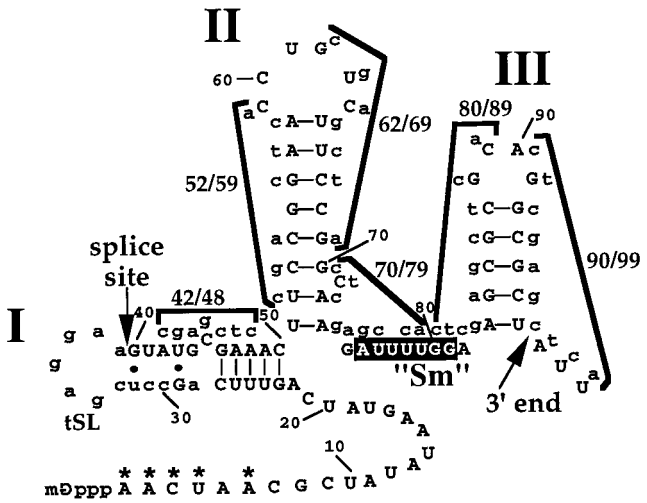

FIG. 1. **A linker scan series was generated through the intron of the *L. tarentolae* SL RNA gene.** The secondary structure of *L. tarentolae* SL RNA is based on the *L. collosoma* structure that predominates *in vivo* (17, 18); numbering is relative to the start of transcription. The stem-loop structures are labeled *I, II,* and *III*, and the exon-intron junction is indicated by the *arrow* after nucleotide 39. The Sm-binding site sequence AUUUUGG and the mature 3′ end of the SL RNA are indicated. The ^{7}mG cap is shown at the 5′ end, along with the methylated nucleotides (*) that comprise the cap 4 structure (3, 4). All constructs contained an exon tag (tSL) sequence at position 28–39, indicated by *lowercase letters*. The scan sequence, CTCGAGCTCA, included *Xho*I and *Sac*I sites; positions within the 10-base pair blocks that are altered are indicated in *lowercase letters* adjacent to the WT position. Three additional scan mutations, 100/109, 110/119, and 120/129, are not shown but are described elsewhere (24).

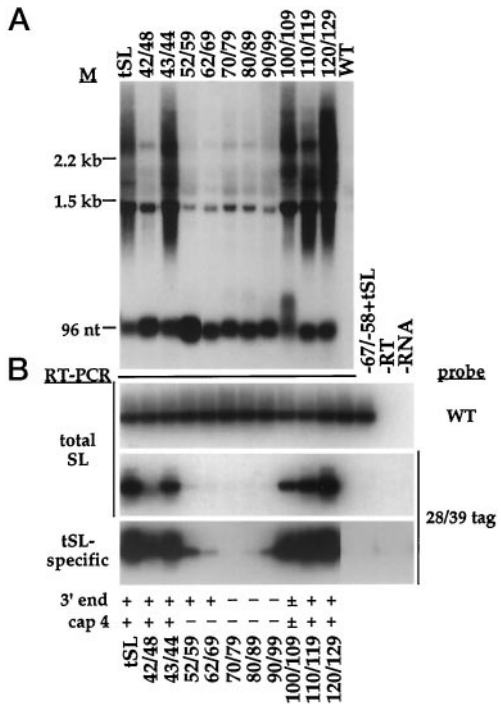

FIG. 2. **SL RNA intron mutations affect *trans*-splicing.** *A*, transcription and variable *trans*-splicing of mutated SL RNAs. Total RNA from WT, the tSL, and the double-site mutants (containing tSL and mutations detailed in Fig. 1) were electrophoresed through a 1.1% formaldehyde-agarose gel, blotted, and hybridized with an oligonucleotide (28/39 tag) directed against the exon tag. Two of the ribosomal RNA exclusion zones (2.2 and 1.5 kb) and the SL RNA (96 nt) served as size markers (*M*). *B*, low levels of *trans*-splicing detected by RT-PCR. RT-PCR assays were performed on the mutant RNA populations with either nonspecific SL RNA primer (*top* and *middle*) or 28/32 tag-specific SL RNA primer (*bottom*) using the *L. tarentolae* Arl mRNA (29) as the query template. Control reactions included RNA from a mutant containing tSL coupled with an inactive promoter, *−67/−58 + tSL* (11), *-RT,* and *-RNA* reactions. Additional 5′ and 3′ end phenotypes (24) are summarized as follows: + = WT, ± = WT, and − = T tract for 3′ end formation, and + = WT, ± = 50%, and − = < 5% for cap 4 methylation.

Sm-binding site are necessary for splicing. Our data broadly contrast the results obtained in *Leptomonas*, where the structure and not the primary sequence of stem-loop II was necessary for splicing, and where the Sm-binding site and stem-loop III were not required for splicing. A model summarizing the features of the *Leishmania* SL RNA involved in maturation and *trans*-splicing is presented.

EXPERIMENTAL PROCEDURES

Generation of Mutations and Transfectants—Mutagenesis was performed using the Sculptor Mutagenesis kit (Amersham Pharmacia Biotech) or using PCR to generate mutagenized DNA fragments for subcloning into the transfection vector. Mutated fragments were cloned for transfection into a pX plasmid (27) containing an SL RNA gene (9).

Transfections were performed by electroporation as described (11).

Nucleic Acid Isolation and Gel Analysis—RNA was purified using TriZOL reagent (Life Technologies, Inc.) and was electrophoresed through 1.1% agarose-formaldehyde, blotted, and hybridized as described previously (28). Quantitation was performed using a PhosphorImager (Molecular Dynamics).

Reverse Transcriptase-PCR trans-Splicing Assay—To assay for *trans*-splicing of mutant-SL RNAs, complementary oligonucleotides were hybridized to Arl mRNA (29) and extended by Moloney murine leukemia virus-reverse transcriptase (RT) to produce templates for PCR analysis (RT-PCR (30)) with a second, SL-specific oligonucleotide. The following oligonucleotides were used: Arl(−)68, 5′-TGCGGATCGCCTTCTGGCCACC; LtSL5′RI, 5′-GGGAATTCGCTTTCAACTAACGCTAT; 30/39–5′HI, 5′-GGGATCCTGTATCAGTTTCAGCCT. Amplification products were run in 1.5% agarose gels, denatured, blotted, and hybridized with 28/39-tag oligonucleotide (5′-ACTTCCTCGAGGCTGAA) to detect the exon tag or wild-type (5′-CAATAAAGTACAGAAACTGA) exon sequence (corresponding to positions 20–39). The expected product size is ~300 bp.

5′ End Analyses—5′ cap 4 methylation was assayed by primer extension of total RNA with an exon tag-specific oligonucleotide (28/39-tag) as described previously (9, 11).

RESULTS

Intron Mutations Affect trans-Splicing—To localize specific elements within the intron of the SL RNA that play a role in the process of *trans*-splicing, a systematic mutagenesis approach was adopted (Fig. 1). To differentiate mutated, episomally derived SL RNA from the endogenous WT SL RNA population, an exon tagged at positions 28 and 30–39 (28/39), which was previously shown to *trans*-splice accurately and efficiently (11), was used as a molecular tag (tSL RNA) for detection by hybridization. A series of linker scan (CTCGAGCTCA) mutations in the tSL RNA gene was created for transfection to *L. tarentolae*. Two mutations in the 40–49 region were created as follows: a mutant with alterations in bases 43 and 44 (43/44 = GT *versus* TG in WT) tested a postulated SL RNA-U6 snRNA interaction (31), and a second mutant, altered at positions 42–48 (42/48), changed all but the splice donor site with the linker scan sequences. Subsequent intron mutations continued from position 50 (52/59) and proceeded through the end of the intron. Three mutations lay downstream of the mature 3′ end of the SL RNA transcript (position 96) and were included to identify potential adjacent expression elements (100/109, 110/119, 120/129; WT sequence not shown).

Analysis of total RNA from the transfectants demonstrated tSL RNA (~96 nt) in all the samples by low resolution formaldehyde-agarose gel blotting (Fig. 2*A*). A broadened size range from the wild-type (WT) 96 nt to at least 175 nt was noted for the 100/109 tSL RNA, consistent with discrete higher molecular weight bands visible in higher resolution gel analyses (24). The presence of tagged precursor SL RNA indicated that all the mutants have the potential to *trans*-splice. An artifactual tran-

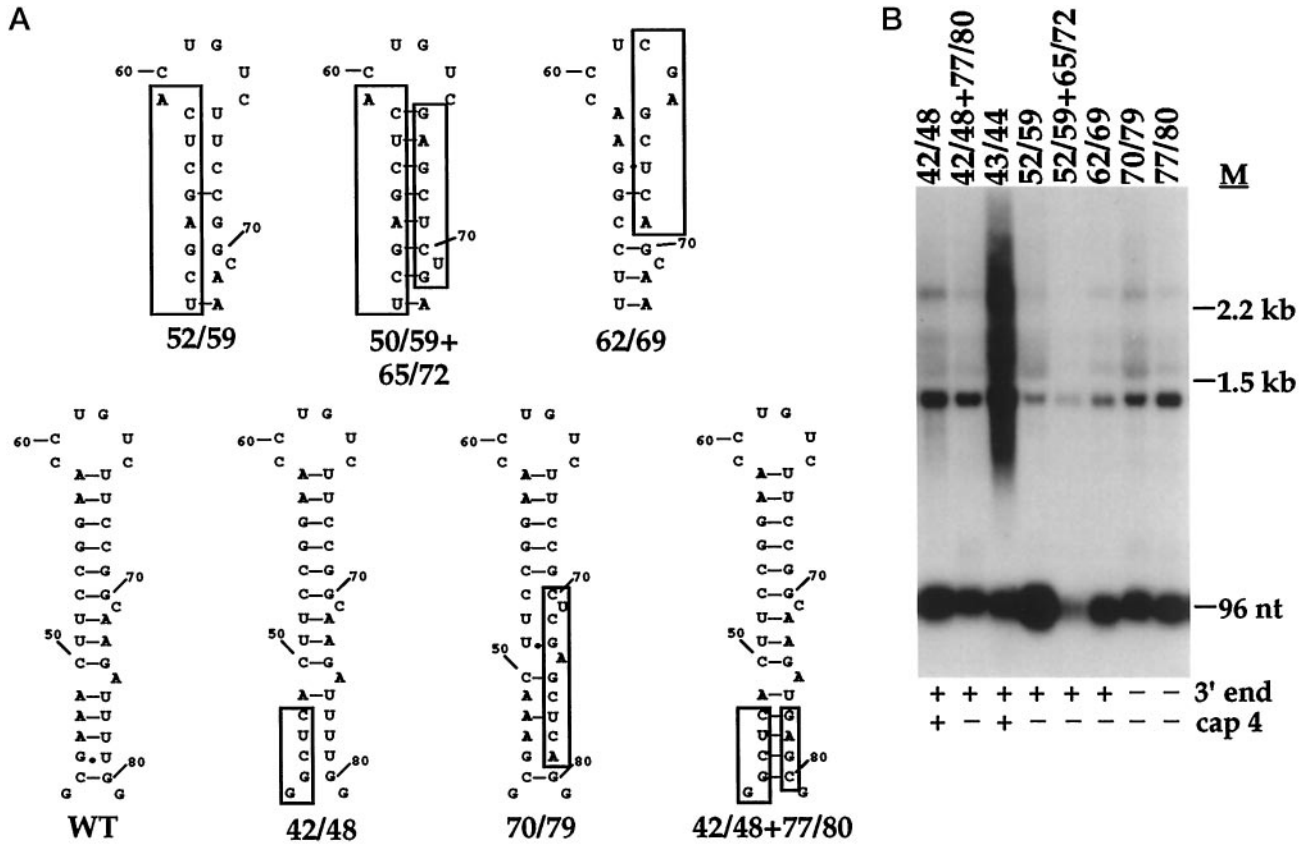

FIG. 3. **Stem-loop II structure and primary sequence are necessary for *trans*-splicing.** *A*, disruptions of stem-loop II and compensatory mutations. The disruptions in stem-loop II caused by 52/59 and 62/69 are shown, with the structural compensatory mutation 52/59 + 65/72. All mutated positions are *boxed*. *B*, the compensatory changes do not rescue *trans*-splicing. Examination of total RNAs from the original and compensatory mutants on a blot of a 1.1% agarose-formaldehyde gel hybridized with the 28/39 tag oligonucleotide. 43/44 is included as a positive splicing control. Markers (*M*), 3′ end, and cap 4 phenotypes are as described for Fig. 2 or in the text.

script (~1.45 kb) that accumulated in each sample provided an internal control for transfection, should stability be disrupted. The presence of the exon tag in a range of high molecular weight RNA species (500 nt to 9 kb) in the tSL, 43/44, 100/109, 110/119, and 120/129 samples suggested that active *trans*-splicing was occurring in these transfectants (11). Splicing of 43/44 is consistent with results from a similar study in *Leptomonas* (12). Conversely, splicing of the tagged exon was impaired in the 42/48, 52/59, 62/69, 70/79, 80/89, and 90/99 mutants, where only substrate tSL RNA and the artifactual transcripts accumulated. The levels of accumulating tSL RNAs varied relative to the artifactual transcripts and the episomally encoded drug selectable marker mRNA NEO (data not shown). The 52/59 mutant in particular showed an increased accumulation of substrate molecules relative to other non-splicing mutants.

In addition, *trans*-splicing was assayed by RT-PCR (11, 29) to detect low levels of splicing (Fig. 2*B*). When "total" SL primer (*i.e.* will amplify from both WT and tSL exons) was used in the amplification, all samples showed the positive control WT amplification products, but the 28/39-tag oligonucleotide hybridized only to the tSL, 43/44, 100/109, 110/119, and 120/129 products, consistent with the total RNA blot analysis. However, using a tSL-specific primer for amplification, some level of splicing was detected in all but the 70/79 mutant. These experiments included WT *L. tarentolae* RNA, no reverse transcriptase, and no RNA reactions as negative controls for contamination and dependence on the use of RNA templates. Furthermore, a promoter knockout in combination with tSL (−67/−58 + tSL (11)) was used as a control for spurious PCR amplification; this cell line resulted in an artifactual ~1.45-kb transcript containing the 28/39-tag sequence but no mature tSL RNA (shown in Fig. 5*B*) and did not yield a tSL PCR product. Previously determined phenotypes for cap 4 and 3′ end formation (24) are also indicated in Fig. 2. Thus, *trans*-splicing was adversely affected in mutants 42/48, 52/59, 62/69, 80/89, and 90/99 and appeared to be abolished in mutant 70/79.

Structural analyses of the SL RNA predict three stem-loop structures and a single-stranded region containing the Sm-binding site (Ref. 17; Fig. 1). Previously, it was demonstrated that stem-loop I is not required for *trans*-splicing in *L. tarentolae* (11). Because *trans*-splicing was reduced or abolished in mutants 52/59, 62/69, 70/79, 80/89, and 90/99, we considered the effects of mutations in stem-loop II, the Sm-binding site, and stem-loop III on SL RNA *trans*-splicing with regard to the structural or sequence elements. The mutations described below are organized with regard to both these elements and the linker scan mutation results in the following order: structural features of stem-loop II, fine analysis of the 70–81 region which includes part of stem-loop II and the Sm-binding site, and features of stem-loop III.

Sequence and Structure of Stem-Loop II Are Necessary for trans-Splicing—Two mutations, 52/59 and 62/69, disrupted stem-loop II (Fig. 3*A*) and were not efficiently *trans*-spliced (Fig. 2). To address the importance of stem-loop II, 52/59 was further mutated to restore base pairing (52/59 + 65/72; Fig. 3*A*); this replaced the stem structure but with a different sequence content than WT. A further mutation was designed (42/48 + 77/80) to restore a possible extension of stem-loop II in the 45–48 region, which was disrupted by mutations 42/48 and 70/79 (Fig. 3*A*). 77/80 was also tested for independent effects due to its disruption of the conserved Sm-binding site. Neither

of the compensatory base pairing mutations restored *trans*-splicing (Fig. 3*B*). 77/80 alone or in combination with 42/48 resulted in extended, heterogeneous 3′ end formation (data not shown) consistent with the 70/79 phenotype (24), whereas 52/59 and 52/59 + 65/72 possessed correct 3′ ends (data not shown). 42/48 + 77/80, 52/59 + 65/72, and 77/80 showed undermethylated cap 4 structures (data not shown), as did 42/48, 52/59, and 70/79 (24).

Both the structure and sequence content of stem-loop II are thus important features in the maturation of the SL RNA precursor. The structure alone is not sufficient to direct either cap 4 methylation or splicing. The stem-loop II extension structure may play an intermediate role in the splicing pathway, but it is not sufficient to restore processing or splicing.

The Stem-Loop II Single Nucleotide Bulge and Sm-binding Site Affect trans-Splicing and Cap 4 Methylation—Because 70/79 altered most of the Sm-binding site and resulted in no *trans*-splicing and defects in both 5′ and 3′ end formation, we examined the area in finer detail. A 2-bp transversion (TV) series was created from position 70 to 81; in addition, 70/79 and 75/81 TV mutations were made (Fig. 4*A*). It should be noted that 70/71 and 72/73 comprise part of stem-loop II (see Fig. 1) and that 74/75 TV may extend the Sm sequence (AAUCUUUUGG).

The total RNA of these transfectants revealed a variety of phenotypes for *trans*-splicing and methylation. By formaldehyde-agarose gel analysis, only the 74/75-tSL RNA was an efficient *trans*-splicing substrate, with low levels of splicing evident in 72/73 (Fig. 4*B*). The presence of tSL RNA but lack of the 1.45-kb artifact RNA in mutants 76/77, 70/79 TV, and 75/81 suggested additional increased stability phenotypes. Primer extension revealed an intriguing gradient of SL RNA cap 4 methylation in the 70/71, 72/73, and 74/75 mutants (Fig. 4*C*), which showed low (5%), medium (40%), and normal (75%) methylation, respectively, and were *trans*-spliced proportional to their methylation state. Thus, as a component of the Sm-binding consensus, A^{75} does not appear to be an essential nucleotide in *Leishmania*; alternatively, the 74/75 mutation is a biologically acceptable extension of the Sm site.

The Stem-Loop III Formation Affects trans-Splicing and Cap 4 Methylation—Mutations 80/89 and 90/99 disrupted stem-loop III (Fig. 5*A*) and resulted in 3′-extended, undermethylated tSL RNAs (11) that were not efficiently *trans*-spliced. To examine further the role of stem-loop III structure, a series of mutations was created that disrupted and then replaced 1 or 3 bp of the stem and that altered the loop sequence (Fig. 5*A*). The single base disruptions were expected to disrupt only one rung of the stem, and thus lead to a minor size difference in the tSL RNA, whereas the triple base mutations were designed to disrupt the stem completely.

Formaldehyde-agarose gel analysis of total RNA showed that the single base pair-disrupted SL RNAs (83, 96) could be *trans*-spliced, but the triple base pair-disrupted SL RNAs (83/85, 94/96) were not *trans*-spliced efficiently (Fig. 5*B*). The base pairing compensated SL RNAs (83 + 96, 83/85 + 94/96) were both efficient splicing substrates. The 88/91 loop mutation SL RNA appeared to be spliced with lowered efficiency based on the reduced mRNA smear relative to the tSL RNA substrate levels. These experiments show that the structure, but not the primary sequence, of stem-loop III is necessary for *trans*-splicing.

Primer extension analysis indicated that the cap 4 methylation patterns (Fig. 5*C*) correlated with the levels of *trans*-splicing. 83, 96, and 83 + 96 were methylated efficiently and *trans*-spliced relative to the abundance of free tSL RNA (Fig. 5*B*) and by total SL RT-PCR assays; 83/85 and 94/96 showed

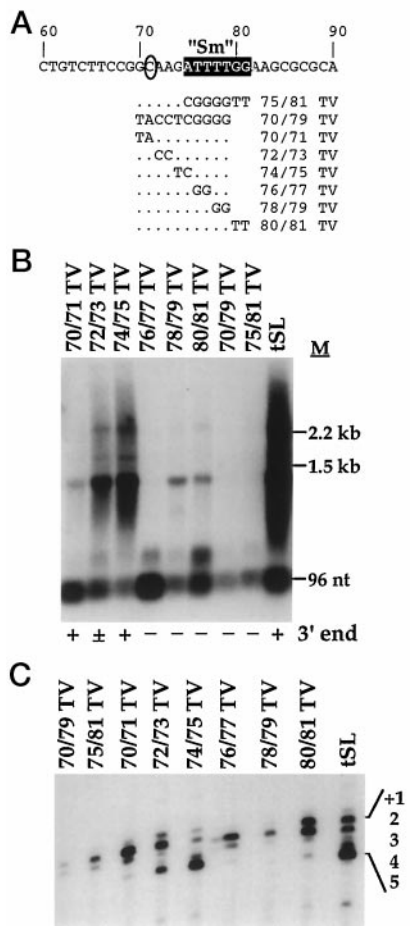

FIG. 4. **The stem-loop II bulge and Sm-binding site affect *trans*-splicing and methylation.** ***A*,** mutations in the 70/81 region. Nucleotide transversions were used to maximize structural disruption. The Sm-binding site (*black background*) and stem-loop II bulge nucleotide (*oval*) are indicated. *B*, the *trans*-splicing ability of the mutants described in *A* is shown in a blot of a 1.1% agarose-formaldehyde gel hybridized with the 28/39 tag oligonucleotide. tSL is included as a positive splicing control. Size markers (*M*) and 3′ end phenotypes are as described in Fig. 2. *C*, variable methylation in the 70–81 mutants. Primer extension analysis was performed. tSL is shown as a representative of wild-type cap 4 formation. The +1 to +5 RT termination sites are marked.

less than 5% methylation and had splicing that was only detectable by the mutation-specific RT-PCR assay (data not shown). In mutant 83/85 + 94/96, the tSL RNA cap 4 was methylated to WT levels, thus the compensating mutations, which restored stem-loop III, also restored a structural signal for the cap 4 methylase. 88/91 showed approximately 50% methylation (Fig. 5*C*) and displayed reduced splicing (Fig. 5*B*); the intron tag previously used to follow SL RNA transcription (9) was inserted into this loop and does not interfere with cap 4 methylation of tagged SL RNA (11). Thus, methylation is linked to the formation of a minimum of 4/5 bp stem in stem-loop III.

DISCUSSION

We have made a series of mutations in the region downstream of the exon in the SL RNA gene to examine effects on *trans*-splicing. We have assayed for the ability of mutated SL RNA to *trans*-splice, and we have correlated this with correct cap 4 formation, transcription termination, and 3′ end processing as determined here and elsewhere (11, 24). A summary of nine phenotypes associated with the intron mutations is presented in Table I. In the majority of mutants, reduction or loss

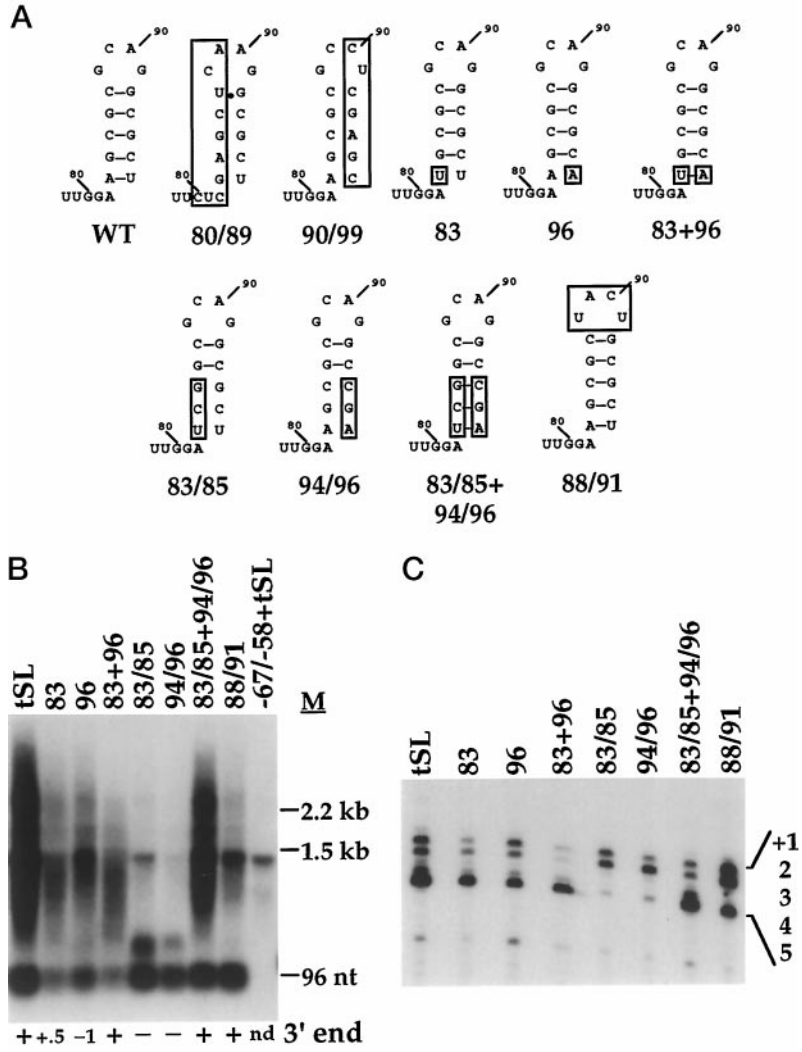

FIG. 5. **Stem-loop III formation is necessary for *trans*-splicing.** *A*, the WT stem-loop III structure and its disruption by various mutations. Single (83 and 96) and triple (83/85 and 94/96) base pair stem disruptions and their compensatory stem structure mutations (83 + 96 and 83/85 + 94/96) are shown. A loop transversion mutation was also created (88/91). All mutated positions are *boxed*. *B*, single and compensatory mutants *trans*-splice. A 1.1% agarose-formaldehyde gel blot was hybridized with 28/39 tag. tSL is a positive splicing control, and −67/−58 + tSL is a negative splicing control (11), size markers (*M*), and 3′ end formation are as indicated in Fig. 2. Additional 3′ end phenotypes +0.5 and −1 refer to discrete transcript size changes, and *nd* = not done. *C*, stem-loop III disruption reduces methylation. Primer extension analysis was performed on the stem-loop III mutant SL RNAs. The SL RNA nucleotide positions corresponding to the product sizes are indicated.

TABLE I
Summary of mutant phenotypes

The symbols used are: +++, WT activity; +, reduced activity; −, ≤5% activity.

trans-Splicing		5′ end, cap 4	3′ end		Mutants, +tSL
Agarose	RT-PCR		stlpIII[a]	T track[b]	
+++	+++	+++	+++	+++	tSL (≅WT), 43/44, 110/119, 120/129, 74/75, 83, 83/85 + 94/96, 106/107[c]
+++	+++	+	+	−	100/109, 102/107,[c] 104/107[c]
+	+++	+	+++	+++	88/91
+	+++	+	+	+++	72/73
−	+	+++	+++	+++	42/48
−	+	−	+++	+++	52/59, 62/69, 52/59 + 65/72, 70/71
−	+	−	−	+++	80/89, 90/99, 83/85, 94/96
−	−	−	+	+++	80/81
−	−	−	−	+++	70/79, 75/81, 76/77, 78/79, 80/81

[a] Nucleolytic processing to base of stem-loop III.
[b] Staggered T track termination products.
[c] See Ref. 24.

of *trans*-splicing correlates with defects in cap 4 methylation and 3′ end maturation. Where cap 4 methylation and 3′ end formation are WT, primary catalytic elements may have been mutated (*e.g.* 42/48).

The mutant phenotypes have allowed us to evaluate structures and elements that may be important for SL RNA maturation and to propose a possible pathway for discrete processing steps in *L. tarentolae* (Fig. 6). In this model, the T tract functions as a transcription termination element (24). The Sm-binding site and stem-loop III structure are required to allow precise 3′ end maturation. Formation of the mature 3′ end, along with elements within stem-loop II, are required for cap 4 synthesis, as is the 10-29 region of the exon (11). Nucleotides within the intron region of stem I are likely to be required for splicing catalysis.

Transcription of the SL RNA gene *in vivo* terminates in a T tract of greater than six T residues (Fig. 6, *step 1*). Mutation of the Sm-binding site (*e.g.* 76/77) and stem-loop III (*e.g.* 90/99) yielded mutants with defects in the nucleolytic formation of the mature 3′ end of the SL RNA, demonstrating a cooperative

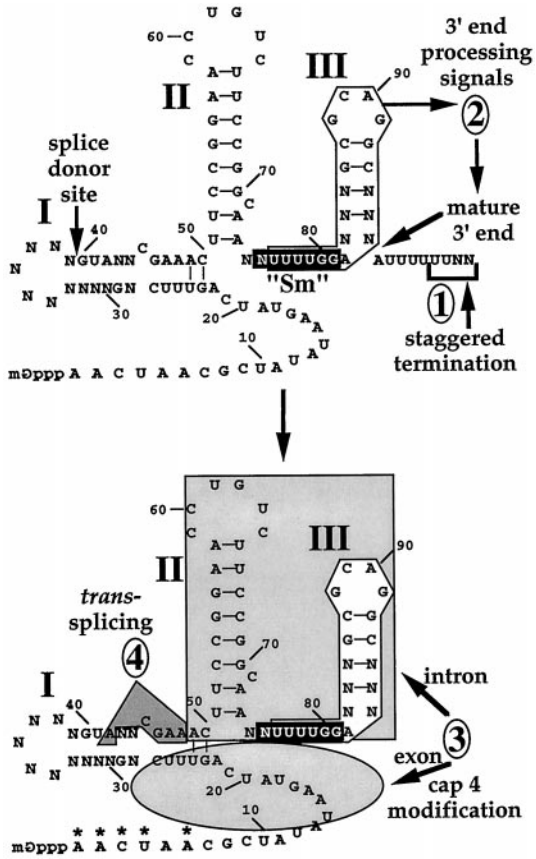

FIG. 6. **A summary of SL RNA elements important for 3′ and 5′ end formation and *trans*-splicing.** The *L. tarentolae* SL RNA structure shown in Fig. 1 has been modified according to the data presented in this paper and elsewhere (24). The 3′ end of the SL RNA precursor is shown by the "staggered termination" in the T tract, and the mature 3′ end is indicated by an *arrow* at the base of stem-loop III. Nucleotides that have been examined by mutagenesis and do not disrupt normal SL RNA maturation or splicing have been changed to "*N*"; proposed base pairing interactions that are not necessary are deleted. The structures and sequences proposed to interact with maturation components, especially for cap 4 methylation, are included in the *shaded oval* and are discussed in the text.

function between these two elements (*step 2*). Structure rather than primary sequence of the stem-loop III stem was required for *trans*-splicing. Consistent with this, an 8-nucleotide insertion in the loop of stem-loop III in *L. collosoma* did not affect *trans*-splicing (25); however, in *L. seymouri* deletion of stem-loop III resulted in an actively *trans*-spliced and normally methylated SL RNA (12). Mutants that do not terminate accurately due to the disruption of their downstream T tract show an intermediate cap 4 phenotype (*e.g.* 100/109) that we interpret as indicative of a temporal order of 3′-processing (*step 2*) followed by cap 4 methylation (*step 3*). Methylation alone is not sufficient to confer splicing potential, since the 42/48 mutant is normally methylated but a marginal *trans*-splicer (a phenotype similar to the *Leptomonas* Δstl II mutant (12)). This suggests that the intron region of stem I may be involved in catalytic steps of *trans*-splicing (*step 4*). Positions 43/44 are not implicated in the 42/48 splicing defect. Nucleotides 42 + 45–48 may interact with other splicing entities such as SLA1 RNA (32) or U5 snRNA (21, 22). Nucleotides +7 and +8 of the intron (equivalent to positions 46–47 in *L. tarentolae*) in *L. collosoma* can be mutated without affecting *trans*-splicing (22).

The essential nature of the Sm-binding site for *in vivo trans*-splicing in *L. tarentolae* agrees with data from *Ascaris*, where *in vitro* studies showed that the Sm-binding site was required for SL RNA *trans*-splicing (23). We are aware of the limitations in comparing *in vitro trans*-splicing assays with *in vivo* splicing phenotypes (12). We generally interpret lack of splicing phenotypes as due to splicing catalysis or, when they are detected, to maturation-related defects, but at this level of analysis our studies cannot exclude other explanations, for example impaired nucleus-cytoplasm-nucleus shuttling of the SL RNA. In contrast, a *Leptomonas* Sm-binding site mutant ("sub-Sm" (12)) that closely approximated a non-splicing, 3′-extended *L. tarentolae* counterpart (78/79) was viable for ribonucleoprotein assembly and splicing. Splicing in the 74/75 mutant, which has a transversion of the A of the Sm site, may reflect flexibility within the conserved Sm-binding site, as found in the U5 snRNA of *Saccharomyces* (33). An additional experimental difference to be considered between the two studies in trypanosomatids is that the exon tag in *L. tarentolae* consisted of 11 mutated nucleotides, whereas that in *Leptomonas* consisted of one mutated nucleotide. The contradictory results for Sm-binding site and stem-loop III in the kinetoplastids may be informative in interpreting our results as follows: given that stem-loop III does not contain primary sequence necessary for *trans*-splicing in *Leptomonas*, our non-splicing phenotypes may be secondary effects (*e.g.* additional 3′-extended sequences may inhibit the folding of stem-loop II).

As indicated by mutant 70/71, elements in stem-loop II are required for the intron component of cap 4 formation. In *Leptomonas*, deletion or substitution of stem-loop II above the bulge position did not affect cap 4 methylation (12), suggesting that some of the methylation phenotypes that we observed may be secondary effects due to interference with secondary or tertiary structure formation within the SL RNA itself or between the SL RNA and other splicing components. Similar to *L. tarentolae*, nucleotides in stem-loop II of the *Ascaris* SL RNA (positions 39–42 and 61–65) are essential for *trans*-splicing and include a single nt (U, position 62) bulge (26). Consistent with the *L. tarentolae* results and contrasting the *Leptomonas* results, deletion of nucleotides 59–68 in stem-loop II of the *Leishmania amazonensis* SL RNA (Δ1) resulted in either inefficient or no *trans*-splicing (8).

In this study we have identified how various structures within the intron of the SL RNA are interdependent in 3′ end formation and cap 4 methylation, and we provide a possible pathway to describe the processing steps. We distinguish among *trans*-splicing negative mutants that are defective for discrete steps in SL RNA maturation and a mutant that may be affected in catalytic steps. These and subsequent mutants will facilitate studies on the intracellular trafficking of SL RNA, the identification of new *trans*-spliceosomal proteins and protein-RNA interactions, and allow testing of new models of interactions with other splicing RNA/ribonucleoproteins.

Acknowledgments—We thank Steve Beverley for the pX plasmid, T. Guy Roberts, and Michael C. Yu for stimulating discussions, and Doug Black, Larry Feldman, and Dan Ray for critical reading of the manuscript.

REFERENCES

1. Agabian, N. (1990) *Cell* **61,** 1157–1160
2. Sutton, R. E., and Boothroyd, J. C. (1988) *Mol. Cell. Biol.* **8,** 494–496
3. Perry, K. L., Watkins, K. P., and Agabian, N. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 8190–8194
4. Freistadt, M. S., Cross, G. A. M., and Robertson, H. D. (1988) *J. Biol. Chem.* **263,** 15071–15075
5. Bangs, J. D., Crain, P. F., Hashizume, T., McCloskey, J. A., and Boothroyd, J. C. (1992) *J. Biol. Chem.* **267,** 9805–9815
6. Sharp, P. A. (1987) *Cell* **50,** 147–148
7. Ullu, E., Tschudi, C., and Günzl, A. (1996) in *Molecular Biology of Parasitic Protozoa* (Smith, D. F., and Parsons, M., eds) pp. 115–133, IRL Press at Oxford University Press, Oxford
8. Agami, R., Aly, R., Halman, S., and Shapira, M. (1994) *Nucleic Acids Res.* **22,** 1959–1965
9. Saito, R. M., Elgort, M. G., and Campbell, D. A. (1994) *EMBO J.* **13,** 5460–5469
10. Hannon, G. J., Maroney, P. A., Ayers, D. G., Shambaugh, J. D., and Nilsen,

T. W. (1990) *EMBO J.* **9,** 1915–1921
11. Sturm, N. R., Fleischmann, J., and Campbell, D. A. (1998) *J. Biol. Chem.* **273,** 18689–18692
12. Lücke, S., Xu, G. L., Palfi, Z., Cross, M., Bellofatto, V., and Bindereif, A. (1996) *EMBO J.* **15,** 4380–4391
13. Maroney, P. A., Hannon, G. J., Shambaugh, J. D., and Nilsen, T. W. (1991) *EMBO J.* **10,** 3869–3875
14. Xie, H., and Hirsh, D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 4235–4240
15. Ferguson, K. C., and Rothman, J. H. (1999) *Mol. Cell. Biol.* **19,** 1892–1900
16. Bruzik, J. P., Van Doren, K., Hirsh, D., and Steitz, J. A. (1988) *Nature* **335,** 559–562
17. LeCuyer, K. A., and Crothers, D. M. (1993) *Biochemistry* **32,** 5301–5311
18. Harris, K. A., Jr., Crothers, D. M., and Ullu, E. (1995) *RNA* (*NY*) **1,** 351–362
19. Nilsen, T. W., Shambaugh, J., Denker, J., Chubb, G., Faser, C., Putnam, L., and Bennett, K. (1989) *Mol. Cell. Biol.* **9,** 3543–3547
20. Mottram, J., Perry, K. L., Lizardi, P. M., Luhrmann, R., Agabian, N., and Nelson, R. G. (1989) *Mol. Cell. Biol.* **9,** 1212–1223
21. Dungan, J. M., Watkins, K. P., and Agabian, N. (1996) *EMBO J.* **15,** 4016–4029
22. Xu, Y., Ben-Shlomo, H., and Michaeli, S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 8473–8478
23. Hannon, G. J., Maroney, P. A., Yu, Y. T., Hannon, G. E., and Nilsen, T. W. (1992) *Science* **258,** 1775–1780
24. Sturm, N. R., Yu, M. C., and Campbell, D. A. (1999) *Mol. Cell. Biol.* **19,** 1595–1604
25. Goncharov, I., Xu, Y., Zimmer, Y., Sherman, K., and Michaeli, S. (1998) *Nucleic Acids Res.* **26,** 2200–2207
26. Denker, J. A., Maroney, P. A., Yu, Y. T., Kanost, R. A., and Nilsen, T. W. (1996) *RNA* (*NY*) **2,** 746–755
27. LeBowitz, J. H., Coburn, C. M., McMahon-Pratt, D., and Beverley, S. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9736–9740
28. Sturm, N. R., Kuras, R., Büschlen, S., Sakamoto, W., Kindle, K. L., Stern, D. B., and Wollman, F.-A. (1994) *Mol. Cell. Biol.* **14,** 6171–6179
29. Sturm, N. R., Van Valkenburgh, H., Kahn, R., and Campbell, D. A. (1998) *Biochim. Biophys. Acta* **1442,** 347–352
30. Sturm, N. R., and Simpson, L. (1990) *Cell* **61,** 871–878
31. Xu, G. L., Wieland, B., and Bindereif, A. (1994) *Mol. Cell. Biol.* **14,** 4565–4570
32. Roberts, T. G., Sturm, N. R., Yee, B. K., Yu, M. C., Hartshorne, T., Agabian, N., and Campbell, D. A. (1998) *Mol. Cell. Biol.* **18,** 4409–4417
33. Jones, M. H., and Guthrie, C. (1990) *EMBO J.* **9,** 2555–2561