Expression Studies on Clustered Trypanosomatid Box C/D Small Nucleolar RNAs*

Yu-xin Xu‡§, Li Liu§, Carlos Lopez-Estraño¶, and Shulamit Michaeli††

From the ‡Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel and the §Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan 52900, Israel

We analyzed three chromosomal loci of the trypanosomatid Leptomonas collosoma encoding box C/D small nucleolar RNAs (snoRNAs). All the snoRNAs that were analyzed here carry two sequences complementary to rRNA target sites and obey the +5 rule for guide methylation. Studies on transgenic parasites carrying the snoRNA-2 gene in the episomal expression vector (pX-neo) indicated that no promoter activity was found immediately adjacent to this gene. Deleting the flanking sequences of snoRNA-2 affected the expression; in the absence of the 3'-flanking (but not 5'-flanking) sequence, the expression was almost completely abolished. The snoRNA genes are transcribed as polycistronic RNA. All snoRNAs can be folded into a common stem-loop structure, which may play a role in processing the polycistronic transcript. snoRNA B2, a member of a snoRNA cluster, was expressed when cloned into the episomal vector, suggesting that each gene within a cluster is individually processed. Studies with permeable cells indicated that snoRNA gene transcription was relatively sensitive to α-amanitin, thus supporting transcription by RNA polymerase II. We propose that snoRNA gene expression, similar to protein-coding genes in this family, is regulated at the processing level.

The nucleolus of eukaryotic cells contains numerous small nucleolar RNAs (snoRNAs) that play a regulatory role in rRNA biogenesis (1–5). The snoRNAs can be divided into two major groups: the box C/D snoRNAs and H/ACA snoRNAs (6–8). Although some of the snoRNAs are involved in pre-rRNA cleavage, most function in site-specific modification, i.e. 2'-O-ribose methylation and pseudouridine formation.

Ribose methylations are guided by box C/D snoRNAs that contain two sequence motifs, box C (5'-PuUGAUGA-3', where Pu is purine) and box D (5'-CUGA-3') (7, 9, 10), which are required for their processing and stability and are important for binding the nucleolar protein fibrillarin (11, 12). The box C/D snoRNAs contain long stretches (>10 nucleotides) that perfectly complement the universal core regions of the mature rRNA. Studies in both yeast and mammals indicate that the fifth nucleotide (+5) upstream from box D or D' within the domains of interaction between the guide rRNA and the snoRNA target site is methyalted; this is known as the +5 rule (7, 10, 13).

snoRNA genes have been of interest because of their distinct mode of expression. The essential and abundant snoRNAs such as U3, U8, and U13 are expressed from their own promoters, whereas most vertebrate and some yeast snoRNAs are encoded by introns of host genes (1, 5). In mammalian cells, most of the snoRNAs are encoded by introns of host genes that encode proteins involved in ribosome biosynthesis and function (14–16). However, in yeast, only a few snoRNAs are encoded by introns, and most of them are independently transcribed (1, 4).

The maturation of most of the intron-encoded snoRNAs involves debranching of the lariat, followed by exonucleolytic trimming (17). The self-transcribed snoRNAs are processed from a precursor by endonucleolytic cleavage and exonucleolytic trimming (18). In yeast, the exonucleases Rat1p and Xrn1p were shown to play a role in 5' to 3' trimming (18–20). In addition, the processing of the independently transcribed box C/D or H/ACA snoRNAs in yeast requires the endonuclease Rnt1p, which is the yeast ortholog of bacterial RNase III. Interestingly, Rnt1p specifically cleaves the duplex region of pre-snoRNAs (21, 22). The cleaved precursor is then processed by exonuclease trimming. The splicing-independent processing mechanism also functions in processing the clustered plant snoRNAs, which are processed from a precursor molecule (23).

Protein factor(s) were found to bind the snoRNA termini, including boxes C and D (24, 25), which protect the snoRNAs from further degradation by the processing exonucleases.

Trypanosomatids are protozoan parasites that diverged very early in the eukaryotic lineage (26). These organisms possess some unique RNA-processing pathways, including pre-mRNA trans-splicing (27, 28) and RNA editing (29). So far, there is no evidence that polymerase (pol) II-regulated promoters exist; these genes are regulated post-transcriptionally by mRNA stability (30, 31). Trypanosomatid rRNA also has a peculiar processing pathway; the large subunit rRNA (28S) is further processed into two large and six small rRNA fragments (32).

Very little is known so far about ribosome biogenesis in trypanosomatids. Previously, we reported the identification of the first box C/D snoRNA, which was termed snoRNA-2 in Leptomonas collosoma (33). We suggested that trypanosomatids obey the canonical +5 rule for snoRNA-mediated methylation. The snoRNAs located within the spliced leader-associated (SLA1) RNA loci were studied in several trypanosomatids, including Trypanosoma brucei, Trypanosoma cruzi, and Leishmania tarentolae. It was found that the trypanosomatid snoRNAs do not obey the general methylation rules. More specifically, the methylation sites corresponding to the duplex formed between the snoRNA and the target site are located either 1 or 6 nucleotides upstream from box D or D' (34). More recently, 17 box
Expression of Trypanosomatid snoRNAs

C/D snoRNAs were analyzed from T. brucei (35) by immunoprecipitation using antibodies against the cloned T. brucei fibrillarin protein. Mapping of the methylation sites potentially guided by eight snoRNAs suggested that the T. brucei snoRNAs obey the +5 rule.

To further elucidate the trypanosomatid snoRNA structure-function relationship and the mode of transcription and processing, we specifically studied four box C/D snoRNAs encoded by three chromosomal loci. Our analyses indicate that there is no irregularity in the guide methylation rule of the trypanosomatid snoRNAs and that the L. collosoma box C/D snoRNAs obey the +5 rule. The genes we analyzed were transcribed as polycistronic precursors. Proper expression of a tagged snoRNA-2 gene cloned into the episomal vector px-neo requires at least 20-nucleotide flanking sequences. Expression, although at a lower level, was detected even in the absence of an upstream sequence, suggesting the lack of a conventional promoter adjacent to the gene. However, we cannot exclude the possibility that a promoter may lie upstream from the gene cluster. Transcriptional studies performed in permeable cells suggest that snoRNA gene transcription is relatively sensitive to α-amanitin, similar to pol II-transcribed genes. The data in this study suggest that the expression of trypanosomatid snoRNAs is mainly regulated at the processing level.

EXPERIMENTAL PROCEDURES

Oligonucleotides—The oligonucleotides for snoRNA-2 (33) were as follows: 16865, 5'-ATGACATGCGCTAATCTGTT-3', antisense from positions 67 to 83 of the coding region; 22075, 5'-CGGATCTGACGAGGGA-3', sense upstream from positions 282 to 300 carrying a BamHI site at the 5'-end (underlined); 22076, 5'-CGGATCT-CCAGTGGATTGTCCCTG-3', antisense downstream from positions 85 to 104; 22178, 5'-CTGGTTGAAGGAAAGAAGGTGAC-3', sense from positions 35 to 66 of the coding region with an EcoRI site in the insertion between positions 52 and 53; 24358, 5'-CGGATCTGACGGCGGTTCG-3', sense from positions -241 to -260; 24357, 5'-CGGATCTCAGTGGACACAGTTCA-3', sense from positions -190 to -209; 24356, 5'-CGGATCTCATCTGTCATGCTG-3', sense from positions -120 to -138; 24355, 5'-CGGATCTCAGTGGACACAGTTCA-3', antisense from positions -51 to -63; 22176, 5'-CTGGTTGAAGGAAAGAAGGTGAC-3', antisense downstream from positions 241 to 260; 24357, 5'-CGGATCTCAGTGGACACAGTTCA-3', antisense from positions -10 to -28; 25709, 5'-CGGATCTGACGAGGGA-3', sense from positions 1 to 15 of the coding region; 17853, 5'-CGGATCTGACGAGGGA-3', sense from positions 182 to 204; 29279, 5'-CGGATCTGACGAGGGA-3', antisense from positions 63 to 85 of the coding region; and 30355, 5'-CGGATCTGACGAGGGA-3', antisense upstream from positions -39 to -19.

The oligonucleotides for snoRNA B2 were as follows: 26533, 5'-CGGATCTGACGAGGGA-3', antisense downstream from positions 46 to 65; 26680, 5'-CGGATCTGACGAGGGA-3', antisense downstream from positions 63 to 82; 26554, 5'-CGGATCTGACGAGGGA-3', antisense downstream from positions 63 to 82; 22652, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26555, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26556, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26557, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26558, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26559, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82; 26560, 5'-GACGAGCAGGTGCTCCTGACAA-3', antisense downstream from positions 63 to 82.

The oligonucleotides for snoRNA TS1 were as follows: 20935, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

The oligonucleotides for snoRNA G2 were as follows: 20935, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

The oligonucleotides for snoRNA TS2 were as follows: 20935, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

The oligonucleotides for snoRNA TS2 were as follows: 20935, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to 28 S mRNA from positions 3746 to 3756 for mapping the methylation site guided by snoRNA G2. The oligonucleotides for snoRNA TS2 were as follows: 20935, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.

Other oligonucleotides were as follows: 21995, 5'-ATGACATGCGCTAATCTGTT-3', antisense to the +5-end of the coding region from positions 67 to 82; 22051, 5'-AGGATGACGACAGGATGTTCAACACACATGTAAC-3', antisense to 28 S mRNA from positions 958 to 977 for mapping the methylation site guided by snoRNA TS1.
Expression of Trypanosomatid snoRNAs

**Northern and Primer Extension Analyses**—The RNA samples (5 µg) were fractionated on a 10% polyacrylamide gel and electroblotted onto a nylon membrane (Hybond, Amersham Pharmacia Biotech). Hybridization with labeled oligonucleotides was performed at 42 °C in 5× SSC, 0.1% SDS, 5× Denhardt’s solution, and 100 µg/ml salmon sperm DNA. Primer extension was performed using end-labeled oligonucleotide (100,000 cpm/pmol). After annealing at 60 °C for 15 min, the sample was kept on ice for 1 min; 1 unit of reverse transcriptase (Expand RT, Roche Molecular Biochemicals) and 1 unit of RNase inhibitor (Promega) were added; and extension was performed at 42 °C for 90 min. The reaction was analyzed on a 6% denaturing polyacrylamide gel next to a DNA sequencing reaction with the same primer. Primer extension sequencing of the RNA was carried out as previously described (41).

**Determination of 2-O-Methylated Nucleotides**—The primer extension reaction for mapping the 2’-O-methylation nucleotides was performed in the presence of different concentrations of dNTPs using 5’-end-labeled oligonucleotides complementary to the 3’-end downstream from each methylated position (42, 43). Total RNA (0.5 µg/µl) was annealed with end-labeled primer at 60 °C for 5 min; the annealing mixture was chilled on ice. The reaction was carried out in buffer (50 mM Tris-Cl (pH 8.6), 60 mM NaCl, 9 mM MgCl2, and 10 mM dithiothreitol), 1 µg/ml of RNA, 10,000 cpm labeled primer, and 1 unit/µl reverse transcriptase (Expand RT) at 0.05–0.5 mM dNTP. After incubation at 42 °C for 1 h, the reaction was analyzed on 6% polyacrylamide gel with 7 M urea gel next to primer extension RNA sequencing with the same primer. The same experimental procedure was used to determine the 2’-O-methylated sites in yeast (43).

**RT-PCR**—Total RNA (20 µg) was extensively treated with RQ1 RNase-free DNase I (5 units; Promega), and the RNA was extracted with phenol/chloroform and precipitated in ethanol. One-tenth of the precipitated RNA was used to prepare cDNA as described above, except that two unlabeled oligonucleotides (10 pmol) complementary to the sequence downstream of the snoRNA-2 and G2 coding region (oligonucleotides 22076 and 22078, respectively) were used. One-tenth of the sample was used as templates for the PCRs with primers located in the intergenic or coding region.

**Transcription Analysis in Permeable Cells**—L. collosoma cells were made permeable generally as previously described (44), except that the transcription buffer contained an additional 20 mM potassium glutamate. The radiolabeled RNA was purified with TRIzol reagent and was used for hybridization. RNA was hybridized with a labeled plasmid DNA (using a Bio-Rad dot-blot apparatus). Hybridization was performed in 5× saline/sodium phosphate/EDTA, 0.1% SDS, 5× Denhardt’s solution, 50% formamide, and 100 µg/ml salmon sperm DNA at 45 °C. The blots were washed at 55 °C in 2× SSC and 0.1% SDS.

**RESULTS**

Sequence, Structure, and Genomic Organization of Two Loci Encoding Four snoRNAs—We previously described the enrichment of L. collosoma small ribonucleoproteins by fractionation on a DEAE-Sephacel column and succrose gradient (33, 39). RNA enriched by these fractions was separated on a denaturing gel, and size-selected RNAs were labeled and used as probes to clone the corresponding genes from a genomic library. Using such a scheme, we previously identified and isolated the...
Expression of Trypanosomatid snoRNAs

The genomic organization of the b2 and g2 loci is illustrated in Fig. 1. The b2 locus carries two putative snoRNAs that were termed B2 and TS1. The g2 locus carries four snoRNAs. The first two are shared with the b2 locus and were therefore termed B2' and TS1'; the other two were designated G2 and TS2. Fig. 1A presents the sequence of the 1.2-kb g2 genomic locus. The coding regions of B2 and B2' as well as TS1 and TS1' are identical. However, the sequences flanking the genes are different (Fig. 1A), suggesting that the two cloned genes represent two different chromosomal loci. Indeed, Southern blot analyses using g2 and b2 as probes showed that these genes are duplicated in the genome since digestion with restriction enzymes of a 6-bp recognition site always resulted in at least two fragments, and partial digestion with Sau3AI produced a duplicated ladder (data not shown). Southern blot analyses also indicated that the gene is repeated at least five times in each locus, as previously demonstrated for the snoRNA-2 gene (33). The genomic organization of the b2 and g2 loci is illustrated in Fig. 1B. One characteristic of these genomic loci was the presence of runs of GT and GC dinucleotides in the regions flanking the snoRNA genes (Fig. 1A).

To demonstrate the existence of these snoRNA transcripts, we used an antisense oligonucleotide to each of the snoRNAs in Northern analyses and primer extensions. The results, presented in Fig. 5, indicate that the sizes of B2, G2, TS1, and TS2 are 92, 84, 91, and 90 nucleotides, respectively. Position +1 of the RNA (Fig. 1A) was obtained from primer extension (data not shown). The intensity of the hybridization signals reflects the abundance of the snoRNAs; snoRNA-2 is more abundant than the other snoRNAs examined.

Trypanosomatid Box C/D Methylation snoRNAs Obey the Canonical +5 Rule—Computer sequence analysis was performed to determine whether these snoRNAs belong to the family of box C/D methylation snoRNAs by searching for the potential base-pairing interaction between the snoRNAs and the rRNA sequences. Fig. 3 presents the results. snoRNA B2 has the potential to guide ribose methylation of C\textsuperscript{3403} and C\textsuperscript{3404} on 28 S rRNA, TS1 to guide ribose methylation of U\textsuperscript{910} and A\textsuperscript{921} on 28 S rRNA, G2 to guide ribose methylation of A\textsuperscript{3697} and A\textsuperscript{3709} on 28 S rRNA, and TS2 to guide ribose methylation of G\textsuperscript{892} and G\textsuperscript{9891} on 28 S rRNA. In each case, the potential for base-pairing presented perfect complementarity across a duplex of 10–12 bp. One exception was that the sequence downstream of box C in snoRNA G2 can potentially interact with 18 S rRNA by perfect base pairing (Fig. 3). However, no methylation sites were found in the region of rRNA that forms base pairing with the snoRNA G2 sequence (data not shown). Our data suggest that there are methylated sites on rRNA that are conserved through trypanosomatids, yeast, and mammans. The methylation of adenosine at position 921, proposed to be guided by TS1, is guided by U32/U51 in humans and snR39/snR59 in yeast, and the methylation of adenosine at position 3709, proposed to be guided by G2, is guided by human U29 and by yeast snR71. These yeast/human snoRNAs carry exactly the same guiding sequences as the trypanosomatid homologs. However, the other methylation sites that we analyzed are trypanosomatid-specific. Of interest is the potential of these snoRNAs to form a very conserved structure with their adjacent target sites on the rRNA (Fig. 3). In this proposed model, boxes C and C' were found in proximity.

The finding that the methylated positions are conserved among trypanosomatids, yeast, and mammals supports the notion that the same guiding rule is shared by these organisms. Indeed, our previous studies using the mapping of partial alkali hydrolysis supported the +5 rule since the methylated site on 5.8 S rRNA was located at position 5 upstream from box D of snoRNA-2 (33). Here, we determined the 2'-O-methylated nucleotides by a reverse transcriptase primer extension assay in the presence of dNTPs from low to high concentrations (42), and the results agree well with the previous data detected by partial alkali hydrolysis. In the presence of low dNTPs, the reverse transcriptase stops 1 nucleotide before the methylation site. The same method was used to map modified nucleotides in yeast (43). Here, we present the mapping of modified nucleotides potentially guided by snoRNA B2 using the primer extension assay (Fig. 4). The data, presented in Fig. 4, support the +5 rule. Moreover, we have mapped all the methylation sites proposed to be guided by the snoRNAs in this study, and all obey the +5 rule (Fig. 3 and data not shown).

Structural Requirement for Expression of snoRNA-2 and B2 Genes in the pX-neo Episomal Vector—As a first step toward

---

2 T. M. Lowe, personal communication.
Expression of Trypanosomatid snoRNAs

Understanding the elements that control the expression of snoRNA genes, we analyzed the snoRNA-2 gene (33). This snoRNA-2 gene is encoded by a reiterated repeat unit and is flanked by 300-bp upstream and 279-bp downstream sequences (Fig. 1B). The gene was tagged by inserting a linker of 8 nucleotides in a position located between boxes D' and C'. The tagged gene was cloned into the pX-neo epimosal vector, which carries a neo gene upstream from the cloning site and confers resistance to neomycin. The stable transgenic parasites carrying the vector were obtained. The expression of the tagged gene compared with the wild-type transcription was assayed by primer extension using an oligonucleotide complementary to the 3'-end of the RNA. The results, presented in Fig. 5B (lane 1), indicate that tagged snoRNA-2 was highly expressed. To further determine the elements that control the expression of the gene and especially to find out whether a specific promoter drives the transcription of this snoRNA gene, we generated constructs carrying variable upstream and downstream sequences and cloned them into the pX-neo vector in two orientations with respect to the neo gene on the vector (Fig. 5A). The results demonstrate that high expression of the tagged gene was dependent on its orientation in the vector. Efficient expression was obtained only when the direction of transcription of the snoRNA-2 gene coincided with the neo gene (Fig. 5B, lanes 1–9). When the snoRNA gene was transcribed in the opposite orientation, the expression was very poor (Fig. 5C, lanes 1–9). The expression of the tagged gene was reduced only when the entire upstream sequence was deleted (Fig. 5A, construct 7), but was unaffected when only 10 nucleotides of upstream sequence were present, suggesting that snoRNA genes may lack conventional upstream promoters adjacent to the gene. Deletion of the 3'-flanking sequence had a more profound effect on expression; expression was reduced when only 15 nucleotides of downstream sequence were present, but was completely abolished in the absence of downstream sequence (Figs. 5, B and C, lanes 8 and 9). These results suggest that proper expression of the snoRNA gene is dependent on both the 5'- and 3'-sequences immediately flanking the coding region of the gene. These flanking sequences (10–20 nucleotides) were probably needed to form the stem structure that could potentially be generated with sequences flanking the snoRNA gene (see below). The poor expression of the tagged gene in the opposite orientation with respect to the neo gene was likely to depend on transcription activity derived from the opposite strand. It has been known for some time that transcription, at different rates, takes place from both strands of the Leishmania episomal vector (36).

The level of the wild-type snoRNA was dependent on the expression of the tagged RNA since the same amount of RNA was used in the primer extension assay (based on the U6 control) (Fig. 5B), yet the level of the wild-type snoRNA-2 was reduced by 30–60% (compare lane W with lanes 1–6). This demonstrates that the overexpression of the tagged RNA gene repressed the level of the wild-type transcript. This repression could originate from competition for an RNA-binding protein (for example, fibrillarin) that binds to this snoRNA. Examining the level of other snoRNAs such as snoRNA B2 in these cell lines suggests that the effect was gene-specific since there was little effect on the level of snoRNA B2 (Fig. 5B). The repression therefore did not originate from a reduction in the level of a common binding protein for all snoRNAs.

Since the data presented in Fig. 5 suggest that the expression of the snoRNA genes may be regulated mainly at the processing level, we examined the potential for the flanking sequences to form a structure that may play a role in the processing event. Folding the flanking sequences of the snoRNA genes using the MFOLD program showed that the sequences could be folded into a stem-loop structure where the snoRNA gene either forms the loop (in the cases of snoRNA-2, B2, G2, and TS2) or participates also in the stem structure (like TS1) (Fig. 6). The minimal proposed stem is composed of 20–50 nucleotides from the 5'- and 3'-flanking sequences. Although the conservation of these stem-loop structures seems to be at the secondary structure, the sequence of all the stems is GT- and GC-rich (Fig. 1A). Based on the data presented in Fig. 5, complete disruption of the stem by deleting sequences either from the 5'- or 3'-flanking sequence reduced the production of mature RNAs, suggesting that this stem structure may serve as a recognition site for the processing machinery.

To examine whether the snoRNA gene located in a gene cluster encoding multiple snoRNAs could also be autonomously
expressed when removed from its neighboring genes, we tagged the B2 and B2\textsuperscript{9} genes by inserting an EcoRI linker between boxes D\textsuperscript{9} and C\textsuperscript{9} (Fig. 7A). Since we observed a difference in the upstream sequence between the B2 and B2\textsuperscript{9} genes (an insertion of a sequence indicated in Fig. 7A), we tagged both genes and examined their expression. The genes were flanked by 93- or 101-bp upstream and 64-bp downstream sequences. Transgenic parasites were generated, and the expression of the tagged genes was compared with that of the wild-type transcript by a primer extension assay. The results, presented in Fig. 7B, suggest that both genes were efficiently expressed, indicating that each gene harbors elements necessary for its expression in vivo when cloned in the episomal vector. However, we cannot rule out the possibility that expression of the gene in its authentic chromosomal locus is not dependent on a distant classical promoter element located in the 5'-flanking sequence of the gene cluster. As already mentioned, the overexpression of a tagged snoRNA B2 gene from the episomal vector also repressed the level of the wild-type transcript.

\textbf{snoRNA Genes Are Transcribed as Polycistronic snoRNA Precursors—}Sequences upstream from each of the snoRNA genes did not reveal common motifs that could potentially serve as internal promoters. In addition, the finding that snoRNA-2 was properly expressed in the absence of the 5'-flanking sequence may suggest that either the promoter for the entire snoRNA cluster is present upstream from the first copy or that such a promoter does not exist, like pol II-transcribed protein-coding genes in this family (31). However, in both scenarios, we would expect to detect a polycistronic snoRNA precursor that is processed to generate the individual mature snoRNAs.

An RT-PCR assay was performed to investigate whether such a precursor snoRNA might exist. An antisense oligonucleotide 85 nucleotides downstream from the snoRNA-2 gene was used to produce a cDNA that was amplified by PCR using a sense oligonucleotide located 282 bp upstream from the snoRNA-2 coding region (Fig. 8A). To avoid DNA contamination, the RNA sample was extensively treated with DNase I (RNase-free). As a control, PCR was carried out on the RNA sample without reverse transcription; and in this case, no product was produced (Fig. 8B, lane 1), suggesting that the PCR product was generated from the cDNA template. A PCR product of 490 bp was detected using the cDNA as a template (Fig. 8B, lane 5), demonstrating that the snoRNA was processed from a longer precursor carrying sequences from both the 5' and 3'-flanking sequences. The existence of a polycistronic snoRNA precursor was also observed for the g2 locus. Using a primer located in the coding region of G2 and a primer situated in the intergenic region between TS1\textsuperscript{9} and TS2 (Fig. 8A), we detected a PCR product of 407 bp (Fig. 8B, lane 4), indicating the existence of an RNA molecule carrying both coding regions of G2 and TS2 and the intergenic region between them. Likewise, a product of 389 bp was observed (Fig. 8B, lane 3) using a primer located in the coding region of TS1\textsuperscript{9} and a primer in the upstream sequence of B2\textsuperscript{9} (Fig. 8A). In addition, we could also detect a 1.0-kb product that covers the entire g2 locus using a primer located downstream from the G2 gene and a primer upstream from the B2\textsuperscript{9} gene (Fig. 8B, lane 2). Moreover, the data support the existence of a precursor snoRNA that covers the entire g2 gene cluster. The differences in intensity of the RT-PCR products reflect the length of transcripts to be extended. Weak PCR products were observed in lanes 2 and 3 because the products either cover the entire g2 locus (lane 2) or are located at a distal portion of a long cDNA (lane 3). The longer the cDNA, the weaker is the PCR product made from the template.

\textbf{Analysis of snoRNA-2 Precursors That Accumulate in Transgenic Parasites—}To further characterize the snoRNA precursor, we performed primer extension on RNA from wild-type cells and cell lines carrying the different snoRNA constructs (Fig. 5A). When primer extension was performed using an oligonucleotide located at the 3'-end of snoRNA-2, different extension products were detected that were dependent on the
Expression of Trypanosomatid snoRNAs

A

B

Fig. 7. Expression of the tagged snoRNA B2/B2’ genes in transgenic parasites. A, schematic representation of two different snoRNA B2 loci: snoRNA B2 (locus 1) and snoRNA B2’ (locus 2). The upstream sequence difference between the two loci is indicated. The tag sequence is marked with a flag, and the lengths of the flanking sequences are indicated in base pairs. B, primer extension analysis to determine the expression of tagged snoRNA B2/B2’ genes. Both primer extension and DNA sequencing were carried out with a 3’-end oligonucleotide for snoRNA B2 (oligonucleotide 20406). The positions of the tagged and wild-type (WT) snoRNA B2 and the DNA sequence at the 5’-end are indicated. Lane w, wild-type cells; lanes 1 and 2, cells lines carrying the constructs in A, respectively.

extent of the 5’-flanking sequence. The sizes of the extension products were always ~110 nucleotides longer than those of the upstream sequence present in the constructs (Fig. 5A). The additional sequence beyond the snoRNA 5’-upstream region originated from the vector sequence since the same pattern of extension products (shorter by 80 nucleotides) was obtained when primer extension was performed with an oligonucleotide located 20 nucleotides upstream from the gene (Fig. 5B, lanes 1–9). No extension product was detected with the wild-type RNA, which may reflect the short half-life of the precursor carrying the 5’-flanking sequence. Without an efficient system for processing the nascent snoRNA transcripts, it is difficult, at this point, to demonstrate a precursor-product relationship and to unequivocally prove that the stable snoRNAs are derived from the large precursor RNA molecules we detected in this study.

Box C/D snoRNA Genes Are Transcribed by α-Amanitin-sensitive RNA Polymerase—The results presented in Fig. 10A indicate that the expression of the snoRNA gene was dependent on the expression of the neo gene. Interestingly, as observed in Fig. 5A, the high expression of the tagged snoRNA resulted in repression of the authentic chromosomal gene. We have previously demonstrated that the elevation in neo expression is correlated with the episomal copy number (40). Because of the dependence of the tagged snoRNA gene expression on neo expression, the elevated level of neo mRNA resulted in increased expression of tagged snoRNA-2. The data suggest that snoRNA-2 can be transcribed by pol II or that transcription of an upstream pol II gene enhances snoRNA-2 transcription. However, the dependence of snoRNA-2 expression on the presence of the upstream neo gene does not unequivocally prove that pol II transcribes this gene. To gain further support for the transcription of snoRNA-2 by pol II, we used permeable cells, which efficiently and accurately transcribe endogenous genes (44). The permeable cells were incubated with various concentrations of α-amanitin, and the transcription of the different genes was examined by hybridizing the RNA with the tested genes immobilized on a filter. As presented in Fig. 10B, the transcription of the snoRNA-2 gene was compared both with the pol II-transcribed gene that codes for the SL RNA core-binding protein SmE (45) and with the pol III-transcribed gene, the U6 snRNA. In addition, we also examined and compared the transcription of the SL RNA, which is the most efficiently transcribed gene in this system. The results indicate, as in the previous study (46), that the level of inhibition of SL RNA by α-amanitin was intermediate between those of pol II- and pol III-transcribed genes. The transcription pattern of the snoRNA-2 gene resembled that of pol II-transcribed genes since its transcription was severely inhibited at 50 μg/ml (5% of the level in the absence of the drug), compared with SL RNA, whose transcription was reduced to 40% at the same concentration (Fig. 10B, panel b). Our results differ from those reported for the SL1 locus (47) since, in that study, the transcription data suggested that the SL1 locus was transcribed by a pol II that was slightly more resistant to α-amanitin than pol II which transcribes protein coding genes.

DISCUSSION

In this study, we present evidence that L. collosoma box C/D snoRNAs, proposed to function in ribose methylation, obey the canonical +5 rule established for yeast and mammals. The three snoRNA loci we have studied, which encode a single snoRNA (snoRNA-2) or clusters of snoRNAs (g2 and b2), are transcribed as polycistronic RNAs. Deletion analysis of sequences in the 5’- and 3’-flanking sequences of the snoRNA-2 gene suggests that only the immediate 20-nucleotide flanking sequences are most critical for the snoRNA expression. However, more dramatic effects on expression were observed when the entire 3’-flanking sequence was removed as opposed to the same deletion of the 5’-flanking region. Moreover, a single snoRNA gene from a cluster could be self-expressed when cloned in an epimosomal vector. The data may suggest that snoRNA genes lack a conventional promoter that lies immediately adjacent to the genes. However, we cannot exclude the possibility that a promoter exists upstream from the gene cluster. The transcription analyses suggest that the snoRNA genes are transcribed by RNA pol II.

Based on the location of the methylation site thought to be guided by snoRNA-2, we have shown that L. collosoma snoRNA obeys the canonical rule for guide methylation established for yeast and mammals (33). However, a related study by Roberts et al. (34) on snoRNAs that are associated with the SL1 locus in several trypanosomatid species suggested that the selection of the specific modification site may follow an altered rule. This discrepancy led us to map all the methylation sites that have been proposed to be guided by the b2- and g2-associated snoRNAs. We always found that in L. collosoma, the +5 rule holds true. This finding is supported by the fact that two of the methylation sites (A^5921 and A^7709) identified in this study are conserved in yeast and mammals. Our conclusion regarding the +5 rule in trypanosomatids was recently confirmed by a study on T. brucei snoRNAs, in which seven methylation sites were mapped, and the corresponding guide-methylating snoRNAs were identi-
fied (35). In addition, 6 of the 15 guide snoRNAs identified in that study are potential homologs of the snoRNAs from yeast and vertebrates that obey the 15 rule. All the snoRNAs we have analyzed contain two domains complementary to the target sites and are therefore double-guide snoRNAs. This may be expected since 39 sites of 2'-O-ribose methylation were mapped in Crithidia fasciculata, which is almost as numerous as in humans (48). The need to methylate so many sites and the relatively small genome size of the parasite may have forced many more snoRNAs to guide two sites. In comparison, only 7 of 41 yeast snoRNAs were found to be double guides (43). Interestingly, as in trypanosomatids, in Archaea, numerous small RNAs appear to have the ability to guide methylation from boxes D and D' and are therefore double guides (49). In both trypanosomatids and Archaea, the predicted target sites of double-guide snoRNAs are within the same rRNA molecule and are located in adjacent sites. It still remains to be seen whether the proximity of the two sites to be

![RT-PCR analyses for the detection of snoRNA precursors](image)

**FIG. 9.** Mapping snoRNA precursors by primer extension. Primer extension was carried out on the RNAs extracted from the cell lines carrying the constructs shown in Fig. 5A. The end-labeled antisense oligonucleotides used in the primer extension analyses were oligonucleotide 16865 (specific to the 3'-coding region; A) and oligonucleotide 30355 (specific to the 5'-flanking sequence; B). The extension products are marked with short arrows. The long arrows above the panels indicate the direction of transcription (the same as the neo gene on the plasmid). Lane w, wild-type cells; lanes 1–9, cell lines with the constructs in Fig. 5A, respectively; M, end-labeled pBR322 HpaII digest. nt, nucleotides.
Expression of Trypanosomatid snoRNAs

Fig. 10. A, linkage between neo and snoRNA-2 gene expression. A cell line carrying construct 5 in Fig. 5A was established and selected for growth at the increased G418 concentrations as indicated. The expression of the snoRNA and neo genes was detected by primer extension on the RNA from the same cell culture with oligonucleotides 16865 (specific for snoRNA-2) and oligonucleotide 21995 (specific for neo mRNA). The extension products of the tagged and wild-type (WT) snoRNA-2 and neo mRNA are indicated. Panel a, in vitro transcription analyses in permeable cells in the presence of α-amanitin. Panel a, dot-blot analysis. DNA (5 μg/dot) encoding for the genes indicated was immobilized on a nylon membrane and hybridized with [α-32P]UTP-labeled RNA transcribed in permeable cells in the presence of α-amanitin. The concentrations of α-amanitin are indicated. The differential transcription efficiency of the various genes required several exposures. Panel b, quantitative analyses of transcriptional inhibition by α-amanitin. The percentage of inhibition was relative to the transcription activity without α-amanitin. The data were obtained by densitometric analyses of the dot-blot data shown in panel a.

directed by the double-guide snoRNAs, as well as the structure we proposed between the snoRNAs and their two target sites, may have any functional significance.

Studies in yeast, vertebrates, and plants suggest different strategies for box C/D snoRNA transcription and maturation (5). In vertebrates, snoRNAs encoded by introns of host genes involved in ribosome biogenesis or ribosomal functions are not transcribed by their own promoters (14, 16). Exonucleolytic degradation plays a major role in their processing by trimming the debranched lariat (24, 25). A minor alternative pathway involves endonucleolytic cleavages within the pre-mRNA intron (50). In yeast, most of the snoRNAs are dispersed as independent singlets or within clusters carrying two to seven snoRNAs (20). These snoRNAs are transcribed as polycistronic RNAs from an upstream promoter and are processed by the endonuclease Rnt1p and degraded by 5′ to 3′ exonuclease, Rat1p, and, to a lesser extent by Xrn1p (18–20, 22). In plants, the genes for both box C/D and H/ACA snoRNAs are transcribed as a polycistronic pre-snoRNA transcript from an upstream promoter (51). The plant snoRNAs are processed by endonucleolytic activity, followed by trimming (23). The three chromosomal loci of box C/D snoRNAs described in this study represent a novel organization of snoRNA genes. Although we provide evidence for the existence of pre-snoRNAs, as for several genes in yeast and plants, we have no evidence for the existence of a promoter that regulates the expression of these genes. We cannot, however, exclude the possibility that a promoter exists upstream from the snoRNA gene cluster. We favor the hypothesis that snoRNA gene expression is not regulated at the level of transcription, but, like many protein-coding genes in these organisms, is mainly regulated by RNA processing and stability (30).

This is the first study that shows that the expression of a small RNA gene in trypanosomatids is affected by its orientation in the pX-neo plasmid. The expression of SL RNA (40), 7SL RNA (52), and U snRNAs in L. collosoma (53) was shown to be dependent on extragenic promoter elements and was not affected by their orientation on the plasmid. The previous studies suggested that the SLA1 locus, which also codes for box C/D methylation snoRNA, is transcribed by a modified RNA pol II, similar to the polymerase that transcribes the SL RNA gene locus. The results presented in this study, however, demonstrate that the snoRNA genes we have analyzed are transcribed most likely by the same RNA polymerase that transcribes protein-coding genes. It is therefore possible that there are at least two types of snoRNA genes in trypanosomatids, those that are linked to other small RNAs with a different cellular function (like SLA1) and those that are flanked by protein-coding genes. It will be of great interest to determine whether the snoRNAs described here are flanked by protein-coding genes that are transcribed in the same direction.

The components of the RNA degradation machinery that may play a role in rRNA processing, mRNA stability, or snoRNA maturation have not so far been characterized in trypanosomatids. The yeast RNase III ortholog, Rnt1p, was shown to participate in the processing of snoRNA (U5 snRNA) (54) and polycistronic yeast snoRNAs (21, 22). The site of cleavage was always found in a double-stranded stem present upstream or downstream from or flanking the coding region of the small RNAs. Because of the potential of the trypanosomatid snoRNA gene to form a similar structure with its flanking sequences, we examined the ability of the yeast recombinant Rnt1p enzyme to cleave the trypanosomatid snoRNA substrates. In experiments performed with the yeast U5 substrate as a control, we failed to observe any specific endonucleolytic cleavages in the trypanosomatid snoRNA flanking sequences using the yeast enzyme (data not shown). This result suggests that the trypanosomatid substrates lack the sequence specificity required for Rnt1p cleavage (21).

The lack of conventional promoters for protein-coding genes...
and for the snoRNA genes adjacent to the genes may imply that snoRNA processing and mRNA stability mediated by the exosomal functions should be tightly regulated in these organisms. Another possibility for co-regulating ribosomal function and its biogenesis with the level of snoRNAs in trypanosomatids may involve genomic clustering. The answers to these questions await the completion of the trypanosomatid genome projects.

Acknowledgments—We thank Todd Lowe for help in assigning the yeast and mammalian homologs to the trypanosomatid snoRNAs reported in this study and Guillaume Chanfreau for providing the recombinant Rtt1p enzyme.

REFERENCES

1. Maxwell, E. S., and Fournier, M. J. (1995) Annu. Rev. Biochem. 53, 897–934
2. Bachellerie, J.-P., and Cavaille, J. (1997) Trends Biochem. Sci. 22, 257–261
3. Smith, C. M., and Steitz, J. A. (1997) Cell 89, 669–672
4. Tollervey, D., and Kiss, T. (1997) Curr. Opin. Cell Biol. 9, 337–342
5. Stein, L. E., and Steitz, J. A. (1999) Curr. Opin. Cell Biol. 11, 378–384
6. Balakin, A. G., Smith, L., and Fournier, M. (1996) Cell 86, 825–834
7. Kiss-Liszki, Z., Henry, Y., Bachellerie, J.-P., Caiiezergues-Ferrer, M., and Kiss, T. (1996) Cell 85, 1077–1088
8. Ganot, P., Bortolin, M., and Kiss, T. (1997) Cell 89, 789–789
9. Cavaille, J., Nicolos, M., and Bachellerie, J.-P. (1998) Nature 383, 732–735
10. Kiss-Liszki, Z., Henry, Y., and Kiss, T. (1998) EMBO J. 17, 797–807
11. Baserga, S. J., Yang, X. W., and Steitz, J. A. (1991) Nature 399, 1221–1224
12. Ruvkun, G. (1992) Science 257, 1000–1005
13. Tycowski, K. T., Smith, C. M., Shu, M. D., and Steitz, J. A. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 14480–14485
14. Leverette, R. B., Andrews, M. T., and Maxwell, E. S. (1992) EMBO J. 11, 1215–1221
15. Fraga, N., Fraga, M., Michienzi, A., Caiiezergues-Ferrer, M., and Bozzoni, I. (1993) EMBO J. 12, 2921–2928
16. Tycowski, K. T., Shu, M. D., and Steitz, J. A. (1993) Genes Dev. 7, 1176–1190
17. Oei, S. L., Samarsky, D. A., Fournier, M. J., and Boone, J. D. (1998) RNA 4, 1096–1110
18. Villa, T., Ceradini, F., Presutti, C., and Bozzoni, I. (1998) Mol. Cell. Biol. 18, 3376–3383
19. Pethakis, E., Sandoval, T., Henry, Y., and Tollervey, D. (1998) Mol. Cell. Biol. 18, 1181–1189
20. Qu, L. H., Henras, A., Lu, Y. J., Zhou, H., Zhou, W. X., Zhu, Y. Q., Zhao, J., Henry, Y., Caiiezergues-Ferrer, M., and Bachellerie, J.-P. (1999) Mol. Cell. Biol. 19, 1144–1153
21. Chanfreau, G., Legrain, P., and Jacquier, A. (1998) J. Mol. Biol. 284, 975–998
22. Chanfreau, G., Rotondo, G., Lebrin, P., and Jacquier, A. (1998) EMBO J. 17, 3726–3732
23. Leader, J. D., Clark, P. G., Watters, J., Beaven, A. F., Shaw, P. J., and Brown, J. W. (1999) Plant Mol. Biol. 39, 1091–1100
24. Watkins, N. J., Leverette, R. D., Xia, L., Andrews, M. T., and Maxwell, E. S. (1996) RNA 2, 118–132
25. Xia, L., Watkins, N. J., and Maxwell, E. S. (1997) RNA 3, 17–26