MOLECULAR AND CELLULAR BIOLOGY, Mar. 1994, p. 2021–2028
0270-7306/94/$04.00+0

Vol. 14, No. 3

RNA Polymerase III-Mediated Transcription of the Trypanosome U2 Small Nuclear RNA Gene Is Controlled by both Intragenic and Extragenic Regulatory Elements

ANTONIO FANTONI,† AMOS O. DARE, AND CHRISTIAN TSCHUDI*

Department of Internal Medicine, Section of Infectious Diseases, Yale University School of Medicine, New Haven, Connecticut 06520-8022

Received 21 September 1993/Returned for modification 9 November 1993/Accepted 3 December 1993

Transcription of U2 small nuclear RNA (snRNA) genes in eukaryotes is executed by RNA polymerase II and is dependent on extragenic *cis*-acting regulatory sequences which are not found in other genes. Here we have mapped promoter elements of the *Trypanosoma brucei* U2 snRNA gene by transient DNA expression of mutant constructs in insect form trypanosomes. Unlike other eukaryotic U2 snRNA genes, the *T. brucei* homolog is transcribed by an RNA polymerase III-like enzyme on the basis of its sensitivity to the inhibitors α-amanitin and tagetitoxin. Thus, the trypanosome U2 snRNA provides a unique example of an RNA polymerase III transcript carrying a trimethylated cap structure. The promoter of this gene consists of three distinct elements: an intragenic sequence close to the 5′ end of the coding region, which is probably required to position the polymerase at the correct transcription start site; and two extragenic elements, located 110 and 160 nucleotides upstream, which are essential for U2 snRNA gene expression. These two elements closely resemble both in sequence and in distance from each other the A and B box consensus sequences of the internal control regions of tRNA genes.

Historically, eukaryotic transcription units have been subdivided into three classes, each transcribed by a different RNA polymerase. This classification is based on the differential sensitivities of RNA polymerase I, II, and III to the drug α-amanitin. In animal cells, RNA polymerase I is insensitive, RNA polymerase II is generally very sensitive, and RNA polymerase III is only moderately sensitivity to α-amanitin. The uridylic acid-rich small nuclear RNAs (U-snRNAs) represent a unique gene family in that some genes, such as the U2 snRNA genes, are transcribed by RNA polymerase II and others, such as the U6 snRNA genes, are transcribed by RNA polymerase III (reviewed in reference 11). In addition, recently it was shown that the plant U3 snRNA genes are transcribed by RNA polymerase III and not by RNA polymerase II as in other eukaryotic organisms (15). Thus, the U3 snRNA represents the first example of an RNA that is synthesized by different RNA polymerases in different organisms.

Transcription of U-snRNA genes has been studied most extensively in vertebrates (reviewed in reference 11). There are two *cis*-acting regulatory elements which are shared between the RNA polymerase II- and RNA polymerase III-transcribed genes: the distal sequence element (DSE) and the proximal sequence element (PSE), located at approximately −200 to −250 and −50 to −60, respectively. The DSE has properties similar to an enhancer, whereas the PSE is required for selection of the transcription start site. In addition to these common sequences, the RNA polymerase III-transcribed U6 snRNA gene has a TATA box, centered around position −27, which determines the RNA polymerase specificity (21). This TATA box constitutes a dominant signal: when it is introduced into the 5′-flanking region of the U2 snRNA gene, the specificity of the U2 promoter switches from RNA polymerase II to RNA polymerase III.

Promoter elements of invertebrate U-snRNA genes have not yet been analyzed in such detail. Expression of the sea urchin U1 snRNA gene requires a DSE located between −318 and −300 and a PSE centered at −55 (41). On the other hand, the sea urchin U2 snRNA gene contains four *cis*-acting elements, including a TATA box at −25 and a PSE at −55 (30). Since there is no similarity between the regulatory elements required for expression of the sea urchin U1 and U2 snRNA genes, it appears that each sea urchin U-snRNA gene utilizes distinct promoter elements.

In contrast, in higher plants the promoter elements for the RNA polymerase II- and RNA polymerase III-transcribed U-snRNA genes are identical (38, 40). They contain a TATA box at −30 and an upstream sequence element. Interestingly, the RNA polymerase specificity is determined by the spacing between these two elements (39).

Progress in understanding the mechanisms of transcription of U-snRNA genes in trypanosomes has been delayed by the lack of efficient and accurate expression systems. Indeed, very few promoters of any type have been characterized in trypanosomes. The only examples are RNA polymerase I promoters of the procyclic acidic repetitive protein (PARP) genes (7, 28), a variant surface glycoprotein gene expression site (43), and the rRNA genes (27, 42). Immunoprecipitations with antibodies against the trimethylated cap structure of U-snRNAs have identified a number of possible candidates for U-snRNAs in trypanosomes (25, 33). Of these, four have been characterized in more detail, and on the basis of structural homology with other eukaryotic U-snRNAs, three have been identified as U2, U4, and U6 snRNAs (25, 32, 33). The fourth RNA, snRNA B, appears to be the U3 snRNA homolog (10). Similar to what is known in other organisms, the U2 and U4 snRNAs contain a trimethylated cap structure (25, 33) whereas the U6 snRNA contains a γ-monomethylphosphate cap (22), and the U-

* Corresponding author. Mailing address: Department of Internal Medicine, Section of Infectious Diseases, Yale University School of Medicine, P.O. Box 20822, 333 Cedar Street, New Haven, CT 06520-8022. Phone: (203) 785-7332. Fax: (203) 785-3864.

† Permanent address: Department of Human Biopathology, University of Rome, La Sapienza, 00100 Rome, Italy.

snRNAs are involved in the processing of other RNA molecules. The U2 and U4/U6 snRNPs are essential cofactors in the processing of pre-mRNA in trypanosomes, namely the addition of the spliced leader sequence to the 5′ end of every pre-mRNA by *trans* splicing (35). However, at present nothing is known about the RNA polymerases that transcribe these U-snRNA genes and about regulatory elements essential for trypanosome U-snRNA gene expression.

Previously, we reported the cloning of the U2 snRNA gene from three different and evolutionarily distant trypanosomatid protozoa (36). By comparing the U2 upstream regions, we identified one sequence motif centered at position −156, which was highly conserved and was capable of binding to a protein factor(s) derived from a *Trypanosoma brucei* nuclear extract. In this report we provide evidence that, remarkably, the trypanosome U2 snRNA gene is transcribed by an RNA polymerase III-like enzyme. Thus, the trypanosome U2 snRNA provides a unique example of an RNA polymerase III transcript carrying a trimethylated cap structure. The promoter of this gene consists of three distinct elements. An intragenic sequence close to the 5′ end of the gene plays a role in determining the transcription start site, and two elements in the upstream region resemble tRNA-like A and B boxes.

MATERIALS AND METHODS

Construction of mutant plasmids. The U2 snRNA gene (148 nucleotides [nt]) and flanking sequences were excised from the parental plasmid pSPR3A1 (33) by digestion with *Pvu*II and *Bam*HI and cloned into pBSIIKS$^-$ between the *Sma*I and *Bam*HI sites. Deletion mutants with mutations in the 5′-flanking region and the coding region were prepared by using convenient restriction enzyme sites. The linker insertion of the marked U2 gene at nt 21 (plasmid TU2∇21) contains a unique *Xho*I restriction site. We used this site in combination with a *Sty*I site at nt 80 and a *Nru*I site at nt 149 (at or very close to the 3′ end of the U2 snRNA) to generate deletions of the 5′ and 3′ halves, respectively, of the U2 coding region. Single-stranded template for oligonucleotide-directed mutagenesis was prepared by superinfection with the helper phage VCSM13 (Stratagene). Mutagenesis was performed essentially as described previously (16). Mutants were identified by DNA sequencing. Purified plasmid DNA was prepared by Triton lysis and CsCl gradient centrifugation.

RNA synthesis and hybridization. [α-^{32}P]RNA was synthesized in permeabilized procyclic trypanosome cells as described previously (37). For inhibitor studies, cells were preincubated on ice for 5 min with either α-amanitin (Boehringer Mannheim) or tagetitoxin (Epicentre Technologies), and transcription was initiated by the addition of the transcription cocktail. Hybridizations to dot blots of plasmid DNAs (5 μg per dot) were performed as described previously (37). The different plasmid clones contained the following inserts: 5S, a 0.74-kb genomic fragment with one repeat of the 5S rRNA genes (20); tRNA, a 0.4-kb fragment with a tyrosine tRNA gene (31a); 7SL, a 0.6-kb genomic fragment coding for 7SL RNA (24); tubulin (Tub), a 1.1-kb *Eco*RI-*Sal*I restriction fragment from the tubulin gene cluster (14); calmodulin (Cal), a 0.45-kb *Eco*RI-*Hin*dIII fragment from the calmodulin-translated sequence (34); U2, the 144-bp U2 coding region; and U6, the 90-bp U6 coding region. Both the U2 and U6 coding regions were generated by PCR.

DNA transfection and RNA isolation. Exponentially growing procyclic *T. brucei* cells (10^8 cells per transfection) were transfected with 50 μg of test plasmid by electroporation in a total volume of 0.5 ml of Zimmerman postfusion medium. Two consecutive pulses of 1.5 kV each and 25 μF were applied by using a Bio-Rad gene pulser apparatus. The test plasmid was cotransfected with 50 μg of a plasmid carrying the bacterial chloramphenicol acetyltransferase (CAT) gene under the control of the PARP promoter (plasmid pJP44 [29]) to control for transfection efficiency and RNA recovery. Transfected cells were transferred to 10 ml of modified Cunningham's medium (34), and total RNA was isolated 4 to 6 h posttransfection by the guanidinium thiocyanate method (6). Briefly, cells from 10-ml cultures were washed twice with a buffer containing 100 mM NaCl, 3 mM $MgCl_2$, and 20 mM Tris-HCl (pH 7.4) and lysed with 250 μl of solution D (4 M guanidinium thiocyanate, 25 mM sodium citrate [pH 7.0], 0.5% *N*-lauroylsarcosine). The lysate was extracted once with water-saturated phenol and then precipitated with 3 volumes of ethanol. To remove residual amounts of DNA, nucleic acids were treated with RNase-free DNase I (0.1 U/μl; Boehringer Mannheim). The reaction was stopped by the addition of an equal volume of solution D, and RNA was precipitated with ethanol.

Primer extension. Total RNA was analyzed by primer extension with 5′-end-labeled synthetic DNA oligonucleotide primers. The following U2-specific primers were used: U2B-17, complementary to nt 45 to 62; U2-F, complementary to nt 87 to 103; and TRU2-17, complementary to nt 106 to 122. The CAT mRNA was primed with CAT-5, which is complementary to nt 26 to 39 of the translated region. In a typical experiment, 5 μg of total RNA was hybridized to 1 ng of γ-^{32}P-labeled primer in 15 μl of 83 mM Tris-HCl (pH 8.3)–125 mM KCl–5 mM $MgCl_2$ at 65°C for 10 min and at 37°C for 20 min. The reaction volume was then increased to 20 μl by the addition of dithiothreitol and the four deoxynucleoside triphosphates to 10 mM and 0.5 mM final concentrations, respectively. The reaction was started by the addition 100 U of Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories) and incubated at 43°C for 45 min. At the end of the incubation, 20 μl of 20 mM EDTA–0.6 N NaCl was added and the mixture was precipitated with ethanol. An aliquot of the CAT extension products was first separated on 6% polyacrylamide–7 M urea gels and exposed to X-Omat AR5 film with an intensifying screen. The bands corresponding to the CAT cDNA were quantitated by scanning autoradiograms on a Bioimage Visage 2000 system. A second gel fractionation was then performed by loading equal amounts of the CAT cDNA, and the samples of the U2-specific primer extension were normalized to the CAT cDNA reference band.

RESULTS

The trypanosome U2 snRNA gene is transcribed by RNA polymerase III. Inspection of the genomic sequence for the trypanosome U2 snRNA gene did not reveal any similarity to the regulatory elements generally associated with the corresponding vertebrate genes (11). However, the U2 snRNA coding region of 148 nt is followed by a run of T residues reminiscent of an RNA polymerase III termination signal (9). This led us to investigate which RNA polymerase transcribes the *T. brucei* U2 snRNA gene. On the basis of the α-amanitin sensitivity of different transcription units, the trypanosome RNA polymerases can be divided into three classes which resemble those in animal cells (17, 37), the only exception being the PARP and variant surface glycoprotein genes which are transcribed by an α-amanitin-resistant RNA polymerase (7, 28, 43).

We used lysolecithin-permeabilized trypanosome cells, which efficiently and accurately transcribe endogenous genes (37), and incubated them with various concentrations of

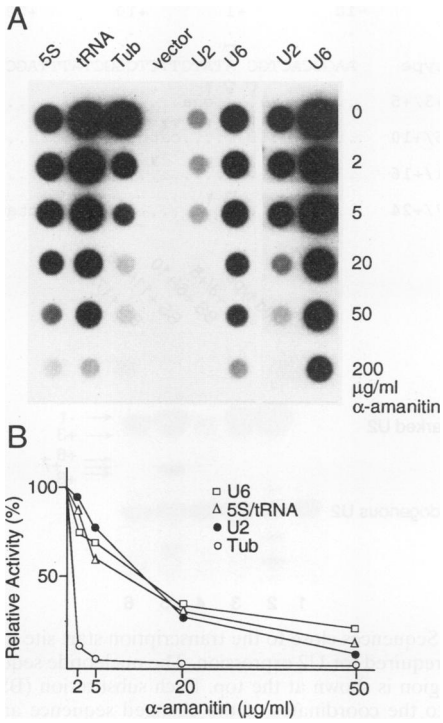

FIG. 1. α-Amanitin sensitivity of trypanosome U2 snRNA gene transcription. (A) [α-^{32}P]RNA was synthesized in permeable insect form trypanosomes in the presence of increasing concentrations of α-amanitin as indicated and hybridized to cloned DNAs immobilized onto nitrocellulose filters. The hybridization to the U2 and U6 genes is shown with two different exposure times. (B) Quantitation of the results as shown in panel A. Signals were quantitated by counting the areas of the dots in a scintillation counter. Since the values for 5S and tRNA genes were almost identical, they are represented by one curve (5S/tRNA). Tub, tubulin gene probe (see Materials and Methods for a description of the plasmids).

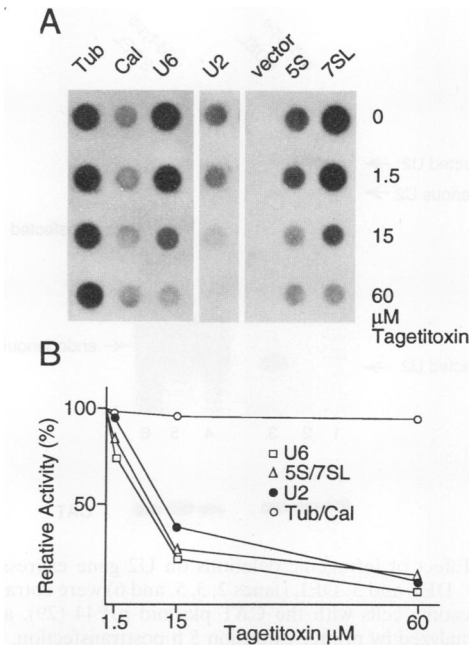

FIG. 2. Effect of tagetitoxin on U2 snRNA gene transcription. (A) Dot blot hybridization as described in the legend to Fig. 1 and in Materials and Methods. The hybridization to the U2, 5S, and 7SL RNA genes was exposed four times longer than that to the tubulin (Tub), calmodulin (Cal), and U6 genes. (B) Quantitation of the results as shown in panel A. Since the values for tubulin and calmodulin genes were almost identical, they are represented by one curve (Tub/Cal).

α-amanitin in the presence of [α-^{32}P]UTP. Synthesis of various RNAs was analyzed by hybridizing the newly synthesized [^{32}P]RNA to different gene probes immobilized on nitrocellulose filters (Fig. 1A). As previously reported (17), low concentrations of α-amanitin (2 to 5 μg/ml) inhibited RNA polymerase II-mediated transcription of the tubulin genes to 5% of the control value (Fig. 1B). Under the same conditions, α-amanitin had only a moderate effect on transcription of the U2 snRNA gene: 2 and 5 μg of α-amanitin per ml inhibited synthesis by 6 and 23%, respectively. Overall, the α-amanitin inhibition curve obtained for the U2 snRNA gene resembled the one obtained for the 5S rRNA genes and a tyrosine tRNA gene, which represent typical RNA polymerase III transcription units (Fig. 1A [longer exposure] and B) (17). The experiments shown in Fig. 1 also established that the trypanosome U6 snRNA gene is transcribed by an RNA polymerase III-like enzyme, similar to what is known to occur in other eukaryotic organisms. To corroborate the α-amanitin results, we used tagetitoxin, a specific inhibitor of RNA polymerase III in other eukaryotic systems (31). In permeabilized trypanosome cells, tagetitoxin preferentially inhibited canonical RNA polymerase III transcription (5S and 7SL RNA) but had no significant effect on RNA polymerase II-mediated transcription of the tubulin and calmodulin genes (Fig. 2). In this analysis, transcription of the U2 snRNA gene was clearly sensitive to tagetitoxin. On the basis of all these results, we therefore conclude that synthesis of the trypanosome U2 snRNA gene is mediated by an RNA polymerase III-like enzyme and not by RNA polymerase II.

Effect of U2 coding regions on promoter activity. To identify DNA elements essential for expression of the U2 snRNA gene in vivo, we analyzed transcription of different plasmid constructs following transfection of insect-form trypanosome cells. To distinguish endogenous transcripts from transcripts originating from the transfected gene, we marked the incoming U2 snRNA gene by insertion of a short linker sequence in the coding region. Thus, primer extension analysis with an oligonucleotide complementary to U2 snRNA sequences downstream of the linker insertion will give two extension products and the cDNA derived from the marked gene will be longer by the length of the linker insertion. Since the U2 snRNA gene is transcribed by an RNA polymerase III-like enzyme (see above), the possibility existed that regulatory elements were located within the coding region of 148 nt. We therefore constructed two different marked genes: one contained an insertion of 13 nt at position 21 (TU2∇21), and the other had 14 nt inserted at position 83 (TU2∇83). The two constructs, when provided with 3.8 kb of upstream sequences and 250 bp of downstream sequences, gave rise to similar levels of the marked U2 snRNA (data not shown), suggesting that neither linker insertion had compromised the activity of putative intragenic control sequences.

The importance of intragenic sequences for expression of the U2 snRNA gene was first investigated by generating two deletions (5′ DEL and 3′ DEL). In 5′ DEL, sequences from nt 21 to 80 were deleted in TU2∇21. Upon transfection and RNA isolation, transcripts were primer extended with oligonucleotide U2-F, which is complementary to sequences downstream

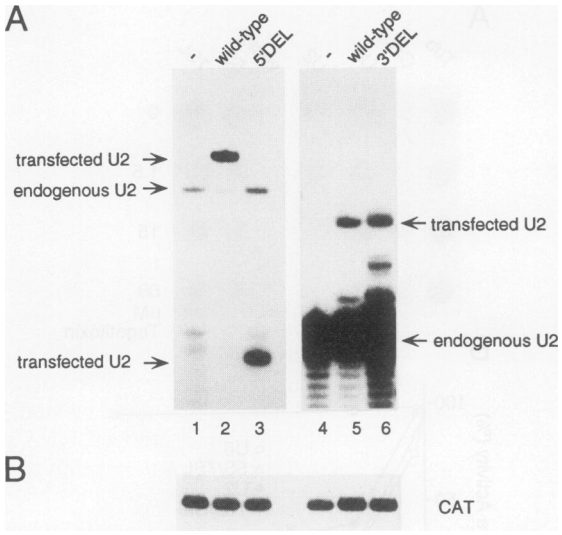

FIG. 3. Effect of intragenic deletions on U2 gene expression. U2 constructs 5′ DEL and 3′ DEL (lanes 2, 3, 5, and 6) were cotransfected into trypanosome cells with the CAT plasmid pJP44 (29), and total RNA was analyzed by primer extension 5 h posttransfection. (A) The U2 snRNA was reverse transcribed with U2-F, complementary to nt 83 to 103 (lanes 1 to 3), and U2B-17, complementary to nt 45 to 62 (lanes 4 to 6). The shorter transcripts in lane 6 were not consistently obtained. (B) Expression of CAT mRNA from pJP44 served as an internal control for all experiments. CAT cDNA was generated with an oligonucleotide complementary to nt 26 to 39 of the translated region, and samples shown in panel A were normalized to the CAT cDNA band. Note that the ratio between the cDNAs derived from transfected and endogenous U2 transcripts is highly variable when different oligonucleotides are used as primers. This we believe is the result of a different priming efficiency of the oligonucleotides on endogenous U2 snRNA versus the mutant constructs.

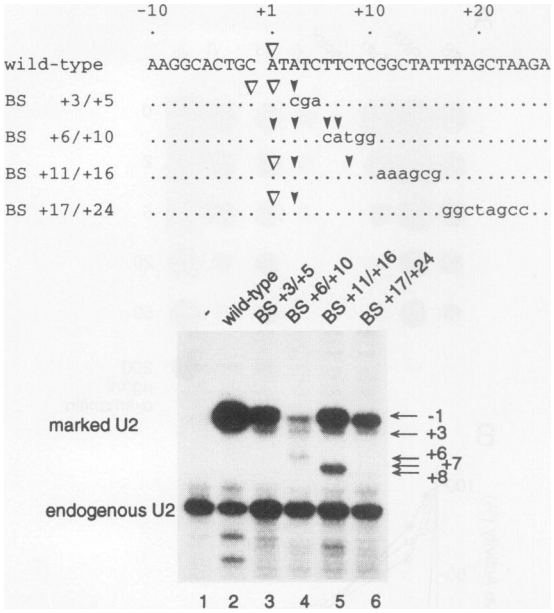

FIG. 4. Sequences close to the transcription start site in the coding region are required for U2 expression. The nucleotide sequence of the mutated region is shown at the top. Each substitution (BS) is named according to the coordinates of the changed sequence and is shown aligned with the wild-type sequence. Only substituted nucleotides are indicated. Major 5′ ends are indicated above each sequence by open arrowheads, and minor 5′ ends are indicated by solid arrowheads. The template activity of the various constructs is shown below. Mutations were introduced in the marked plasmid TU2∇83, and transcripts were analyzed by primer extension with oligonucleotide U2-F. Transcripts with 5′ ends either shorter or 1 nt longer than the wild-type transcript are indicated by arrows. The exact 5′ ends were determined by electrophoresing the samples next to a dideoxy sequencing ladder obtained by extension of the same primer on the wild-type template (data not shown).

of the deletion end point (nt 83 to 103). Thus, the cDNA derived from the 5′ DEL transcript will be smaller than that derived from the parent transcript TU2∇21. To generate a deletion of the 3′ half of the U2 snRNA gene coding region (3′ DEL), nt 80 to 149 were removed in TU2∇21. In this case, primer extension was done with an oligonucleotide complementary to sequences upstream of the deleted region (U2B-17, complementary to nt 45 to 62). Thus, the cDNAs derived from the 3′ DEL and parent transcripts will have the same size. Since we noted that the levels of endogenous U2 snRNA varied a lot when different oligonucleotides were used for primer extension (Fig. 3A), we cotransfected a plasmid carrying the CAT gene under the control of the PARP promoter to serve as an internal control (Fig. 3B). Transfection of 5′ DEL and 3′ DEL into trypanosomes showed that the level of expression from these two constructs was comparable to that of the parent construct TU2∇21 (Fig. 3), suggesting that sequences from nt 21 to the end of the coding region are not required for transcription of the U2 snRNA gene in vivo.

To address the importance of the first 21 nt, we introduced four clustered point mutations in TU2∇21 (Fig. 4). All four mutations had two effects: they reduced the accumulation of wild-type-size U2 transcripts and in addition gave rise to RNA molecules either smaller or 1 nt longer than the wild-type transcript. We did not quantitate precisely the relative levels of expression, since transcripts with aberrant 5′ ends might be unstable in vivo. The majority of the additional transcripts had 5′ ends which mapped downstream of the normal start site (+1), with BS +6/+10 showing the greatest heterogeneity of 5′ ends located between nt +1 and +8 (lane 4). Only mutant BS +3/+5 (lane 3) produced a transcript with a 5′ end upstream (−1) of the normal start site, suggesting that this transcript arose from the use of a new transcription start site. On the other hand, at present we cannot distinguish whether the shorter transcripts were generated by RNA degradation or by alternate transcription initiation.

Transcription of the U2 snRNA gene requires regulatory elements in the 5′-flanking region. Next we made progressive deletions in the U2 snRNA gene 5′-flanking region. Removal of sequences up to position −174 had no effect on the expression of the U2 snRNA gene (Fig. 5A, lane 4). However, further deletion of sequences up to position −112 reduced expression to undetectable levels (lane 5): no U2 transcripts could be detected even after longer exposures of the autoradiogram. To map the regulatory element(s) upstream of the start site of transcription, we introduced a series of block substitutions which span the region between positions −188 and −2 (Fig. 6). Figure 7A shows the activity of the different constructs upon transfection into procyclic trypanosome cells (band labeled "marked U2"). The accumulation of U2 snRNA was reduced to undetectable levels by BS −160/−155 (lane 6) and was reduced to a lesser extent by BS −168/−161 (lane 5).

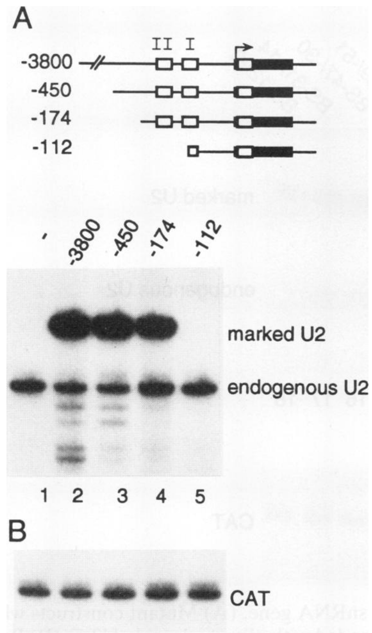

FIG. 5. Transcription of the U2 snRNA gene requires the 5′-flanking region. (A) The structure of the different truncated constructs is shown at the top. Drawings are not to scale. An arrow indicates the U2 transcription start site, and numbers refer to the number of base pairs present upstream of the U2 gene. Following transfection, U2-specific transcripts were primer extended with oligonucleotide U2-F. (B) Primer extension analysis of CAT mRNA.

BS −154/−147 consistently gave about 50% reduction in transcription efficiency (lane 7). The only other mutation that significantly reduced U2 expression in vivo was BS −116/−107 (lane 11). One substitution flanking this sequence, BS −106/−102, also reduced expression by 30% (lane 12). Although in the experiment shown the accumulation of U2 snRNA in mutant BS −95/−67 was decreased by approximately 40% (lane 14), this result was not consistently obtained. In summary, the BS series of clustered point mutations defined two distinct promoter elements in the 5′-flanking region of the U2 snRNA gene. One element, region II, is located between positions −168 and −147, and a second element, or region I, is found between positions −116 and −102. None of the other sequence substitutions affected U2 snRNA gene expression at a detectable level. In addition, sequences between the transcription start site and position −102 do not appear to provide promoter function for the trypanosome U2 snRNA gene.

The upstream regulatory elements resemble the internal control regions of tRNA genes. Since the trypanosome U2 snRNA gene is transcribed by an RNA polymerase III-like enzyme, we examined the identified promoter elements for possible sequence similarities with characterized RNA polymerase III regulatory elements. This search produced the surprising result that the complementary sequences of regions I and II closely resemble the intragenic A and B box promoter elements of tRNA genes. In tRNA genes, the A box is found 8 to 25 bp downstream of the transcription start site and the B box is located 30 to 90 bp downstream of the A box. Figure 8 shows the comparison between the A and B box consensus sequences and their counterparts in regions I and II of the trypanosome U2 snRNA gene 5′-flanking region. The B box in region II has almost perfect homology to the consensus sequence, whereas the A box homology is somewhat degenerate, with three mismatches relative to the consensus sequence. These tRNA-like A and B boxes are positioned and oriented relative to each other in the way they are normally found in tRNA coding regions (Fig. 8B). However, in this case they do not appear to be part of a bona fide tRNA gene, since the surrounding sequences cannot be folded into a convincing cloverleaf structure (data not shown). In addition, they direct transcription in the opposite direction to that in a tRNA gene (Fig. 8B).

```
                                      region II                                               region I              relative
              -190      -180      -170      -160      -150      -140      -130      -120      -110      -100         expression
                .         .         .         .         .         .         .         .         .         .
wild-type     CGGGGGAAAAACGGTATACGTACGGAGGCTCGAACCTGCGGCCACCAAGACGCAAGCACAGAAGTGCAGTTAACGACACCCGGGCCATTACAGAG   1
BS -188/-179  ..ttttcccca....................................................................................   1
BS -178/-169  ............ttgcgcatgc.........................................................................   0.8
BS -168/-161  ......................attcttag.................................................................   0.1
BS -160/-155  ..............................atgcat...........................................................   0.01
BS -154/-147  ....................................gtattaac...................................................   0.5
BS -146/-137  ............................................aacctcatac.........................................   1
BS -136/-127  ......................................................ctacactcct...............................   0.9
BS -126/-117  ................................................................gtactggcca.....................   1
BS -116/-107  ..........................................................................tcacaaattt...........   0.1
BS -106/-102  ....................................................................................aacaa......   0.7
BS -101/-96   .........................................................................................cactct   1

                   -90       -80       -70       -60       -50       -40       -30       -20       -10     +1
                    .         .         .         .         .         .         .         .         .      .
wild-type     CCGGGGCCCGCCCGAATATATAATTACTTATTGAATTTATTCTTTGGTATTTCCTAGCTTGTTGCAGCCTACGGAACTTTTGGACAAGGCACTGCA
BS -95/-67    tt.catt..atggtatggcgcc.g.ggaa...................................................................   0.6
BS -69/-51    ..........................t.cca..gta.ggcg.cag...................................................   1
BS -47/-30    ................................................tccatggtatggcgccag..............................   0.9
BS -31/-14    ................................................................ttc.a.ggtatggcgcca..............   0.9
BS -12/-2     ...................................................................................gagctcttaga..   1
```

FIG. 6. Structure and relative level of transcription of different U2 snRNA gene mutants. The first line shows the U2 5′-flanking sequence from positions −190 to +1. The underlined sequences represent the two identified promoter elements, labeled as region I and region II. For each mutant, the substitutions are indicated by lowercase letters and unchanged nucleotides are indicated by dots. On the right, the relative levels of transcription in vivo are shown, as obtained from the results displayed in Fig. 7.

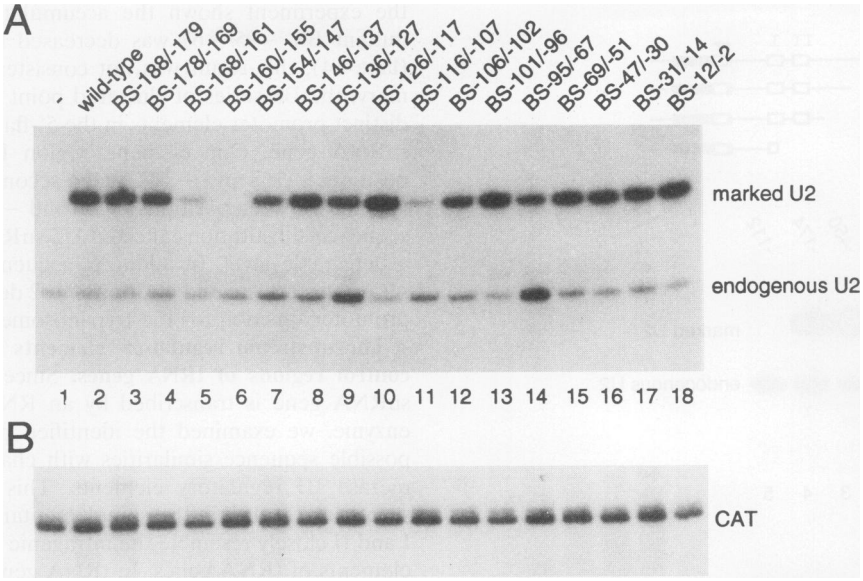

FIG. 7. Two distinct promoter elements are located in the 5′-flanking region of the U2 snRNA gene. (A) Mutant constructs whose sequences are shown in Fig. 6 were transfected into trypanosome cells, and transcripts were primer extended with oligonucleotide U2-F. (B) Primer extension analysis of CAT mRNA.

DISCUSSION

Here we report that synthesis of the *T. brucei* U2 snRNA is mediated by an RNA polymerase III-like enzyme and not by RNA polymerase II as in other organisms. Since the U2 snRNA carries a trimethylated cap structure (25, 33), this makes it, to the best of our knowledge, the first RNA polymerase III transcript with a trimethyl guanosine cap. This conclusion is based on (i) the effect of α-amanitin and tagetitoxin on transcription and (ii) the structural features of the U2 snRNA gene promoter. Our experiments showed that the U2 snRNA gene promoter is composed of both extragenic and intragenic regulatory sequences. Two upstream elements resemble the intragenic A and B boxes which function in tRNA gene expression, and a third element is located in the coding region close to the transcription start site.

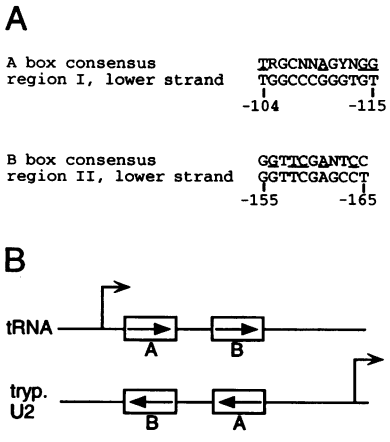

FIG. 8. tRNA-like A and B boxes in trypanosome U2 snRNA gene expression. (A) Comparison of tRNA A and B box consensus sequences and U2 extragenic elements. Invariant positions in the A and B box consensus sequences are underlined. The complementary sequence of the U2 promoter elements (regions I and II) indicated in Fig. 6 is shown. (B) Structure of the tRNA gene promoter and the trypanosome (tryp.) U2 snRNA gene promoter. Arrows within boxes indicate the orientation of the A and B boxes, and the transcription start site is shown by a large arrow.

Until a few years ago, it appeared that RNA polymerase III-mediated gene expression was a relatively straightforward matter in that transcription was mediated largely by short intragenic promoter elements, the A and B box in the case of tRNA genes, that bound a limited number of transcription factors. However, recent work has revealed a new complexity of class III promoters, as exemplified by the vertebrate U6 snRNA genes (reviewed in reference 11), the gene encoding selenocysteine $tRNA^{Ser}$ (5, 18, 26), and the genes encoding RNase P RNA in humans (1) and *Saccharomyces cerevisiae* (19). The common emerging theme is that expression is regulated primarily by extragenic sequence elements. Of interest for our studies is the finding that the yeast RNase P RNA gene requires a tRNA gene-like A and B box located in the 5′ leader sequence (19). In this case, although these regulatory elements are downstream of the transcription start site, like those in more traditional RNA polymerase III promoters, they do not impose constraints on the sequence of the mature functional RNA since they are removed during maturation. Furthermore, analysis of the yeast U6 snRNA gene promoter revealed a conventionally positioned A box within the coding sequence and a functional B box downstream of the coding region (2, 4, 8). In the trypanosome U2 snRNA gene promoter we found an A box-B box combination upstream of the transcription start site. Comparison with consensus sequences showed that region II has an excellent match to the B box, whereas region I has several differences from the A box, most notably A-to-G and G-to-T changes at invariant positions (Fig. 8A). In tRNA genes, the B box is the primary binding site for transcription factor TFIIIC (see reference 13 and references therein). It is also likely that in the trypanosome U2 promoter the B box homology serves as a binding site for the transcription factor(s). Indeed, we have previously identified a nuclear protein(s) that specifically binds to the B box upstream of the

U2 snRNA gene (36), and in the present work we show that point mutations in the B box severely reduce in vivo expression. The same mutations, when tested by the mobility shift assay, abolished binding (data not shown). At present we do not know the identity of the transcription factor(s) interacting with the U2 5′-flanking region. By analogy with the yeast RNase P system, in which TFIIIC has been shown to bind to the A and B boxes in the 5′ leader sequence (19), we would anticipate that TFIIIC is also involved in trypanosome U2 snRNA gene expression. However, the assembly of factors, including TFIIIC, on the reversed A and B boxes of the U2 snRNA gene would be inconsistent with polymerase placement over the transcription initiation site, since this assembly normally occurs in an orientation-dependent manner. A more likely function for the upstream elements is suggested by the recent finding of a new role for TFIIIC in gene activation (3). TFIIIC binding to the B box downstream of the yeast U6 snRNA gene prevents transcriptional repression by chromatin rather than playing a role in transcription complex assembly.

Until recently, it was assumed that each of the three RNA polymerases in eukaryotic cells had its own set of associated factors. This general belief has been challenged by the discovery that the TATA-binding protein is a ubiquitous transcription factor for all three RNA polymerases (reviewed in reference 12). This finding, together with the notion that the large subunits of the three eukaryotic RNA polymerases derive from a common progenitor (23), underscores the possibility that modern promoter structures were built from modular elements and that the variety of modern promoters is a result of different combinations of these modular structures. Because trypanosomes are ancient eukaryotes, it is tempting to speculate that the unusual arrangement of the U2 snRNA gene promoter represents a molecular fossil in the evolution of U-snRNA gene promoters. The finding that elements corresponding to the internal control regions of tRNA genes serve as promoter elements for transcription of the U2 snRNA gene could suggest that the RNA polymerase III intragenic promoters of tRNA genes provided the first modular elements for assembly of extragenic promoters. This idea is further strengthened by our recent finding that functional tRNA genes upstream of the U6 snRNA, 7SL RNA, and snRNA B genes of *T. brucei* provide promoter elements for expression of the corresponding downstream genes (unpublished data).

ACKNOWLEDGMENTS

We are grateful to Nouria Hernandez, Valerian Nakaar, and Elisabetta Ullu for critical comments on the manuscript, to Katherine Rogers for excellent technical assistance, and to Philippe Male for photography.

This work was supported by Public Health Service grant AI-28798 from the National Institutes of Health to Elisabetta Ullu and by the John D. and Catherine T. MacArthur Foundation to C.T.

REFERENCES

1. **Baer, M., T. W. Nilsen, C. Costigan, and S. Altman.** 1989. Structure and transcription of a human gene for H1 RNA, the RNA component of human RNase P. Nucleic Acids Res. **18:**97–103.
2. **Brow, D. A., and C. Guthrie.** 1990. Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. Genes Dev. **4:**1345–1356.
3. **Burnol, A.-F., F. Margottin, J. Huet, G. Almouzni, M.-N. Prioleau, M. Méchali, and A. Sentenac.** 1993. TFIIIC relieves repression of U6 snRNA transcription by chromatin. Nature (London) **362:**475–477.
4. **Burnol, A.-F., F. Margottin, P. Schultz, M.-C. Marsolier, P. Oudet, and A. Sentenac.** 1993. Basal promoter and enhancer element of yeast U6 snRNA gene. J. Mol. Biol. **233:**644–658.
5. **Carbon, P., and A. Krol.** 1991. Transcription of the *Xenopus laevis* selenocysteine $tRNA^{(Ser)Sec}$ gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. EMBO J. **10:**599–606.
6. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
7. **Clayton, C. E., J. P. Fueri, J. E. Itzhaki, V. Bellofatto, D. R. Sherman, G. S. Wisdom, S. Vijayasarathy, and M. R. Mowatt.** 1990. Transcription of the procyclic acidic repetitive protein genes of *Trypanosoma brucei*. Mol. Cell. Biol. **10:**3036–3047.
8. **Eschenlauer, J. B., M. W. Kaiser, V. L. Gerlach, and D. A. Brow.** 1993. Architecture of a yeast U6 RNA gene promoter. Mol. Cell. Biol. **13:**3015–3026.
9. **Geiduschek, E. P., and G. P. Tocchini-Valentini.** 1988. Transcription by RNA polymerase III. Annu. Rev. Biochem. **57:**873–914.
10. **Hartshorne, T., and N. Agabian.** 1993. RNA B is the major nucleolar trimethylguanosine-capped small nuclear RNA associated with fibrillarin and pre-rRNAs in *Trypanosoma brucei*. Mol. Cell. Biol. **13:**144–154.
11. **Hernandez, N.** 1992. Transcription of vertebrate snRNA genes and related genes, p. 281–313. *In* S. McKnight and K. R. Yamamoto (ed.), Transcriptional regulation. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. **Hernandez, N.** 1993. TBP, a universal eukaryotic transcription factor? Genes Dev. **7:**1291–1308.
13. **Kassavetis, G. A., B. R. Braun, L. M. Nguyen, and E. P. Geiduschek.** 1990. *S. cerevisiae* TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. Cell **60:**235–245.
14. **Kimmel, B. E., S. Samson, J. Wu, R. Hirschberg, and L. R. Yarbrough.** 1985. Tubulin genes of the African trypanosome *Trypanosoma brucei rhodesiense*: nucleotide sequence of a 3.7-kb fragment containing genes for alpha and beta tubulins. Gene **35:**237–248.
15. **Kiss, T., C. Marshallsay, and W. Filipowicz.** 1991. Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and *in vitro*. Cell **65:**517–526.
16. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.
17. **Laird, P. W., J. M. Kooter, N. Loosbroek, and P. Borst.** 1985. Mature mRNAs of *Trypanosoma brucei* possess a 5′ cap acquired by discontinuous RNA synthesis. Nucleic Acids Res. **13:**4253–4265.
18. **Lee, B. J., S. G. Kang, and D. Hatfield.** 1989. Transcription of *Xenopus* selenocysteine $tRNA^{Ser}$ (formerly designated opal suppressor phosphoserine tRNA) gene is directed by multiple 5′-extragenic regulatory elements. J. Biol. Chem. **264:**9696–9702.
19. **Lee, J. Y., C. F. Evans, and D. R. Engelke.** 1991. Expression of RNase P RNA in *Saccharomyces cerevisiae* is controlled by an unusual RNA polymerase III promoter. Proc. Natl. Acad. Sci. USA **88:**6986–6990.
20. **Lenardo, M. J., D. M. Dorfman, L. V. Reddy, and J. E. Donelson.** 1985. Characterization of the *Trypanosoma brucei* 5S ribosomal RNA gene transcript: the 5S rRNA is a spliced-leader-independent species. Gene **35:**131–141.
21. **Lobo, S. M., and N. Hernandez.** 1989. A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. Cell **58:**55–67.
22. **Lund, E., and J. E. Dahlberg.** 1992. Cyclic 2′,3-phosphates and nontemplated nucleotides at the 3′ end of spliceosomal U6 small nuclear RNAs. Science **255:**327–330.
23. **Memet, S., W. Saurin, and A. Sentenac.** 1988. RNA polymerases B and C are more closely related to each other than to RNA polymerase A. J. Biol. Chem. **263:**10048–10051.
24. **Michaeli, S., D. Podell, N. Agabian, and E. Ullu.** 1992. The 7SL RNA homologue of *Trypanosoma brucei* is closely related to mammalian 7SL RNA. Mol. Biochem. Parasitol. **51:**55–64.
25. **Mottram, J., K. L. Perry, P. M. Lizardi, R. Lührmann, N. Agabian, and R. G. Nelson.** 1989. Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4, and U6 RNA analogs. Mol. Cell. Biol. **9:**1212–1223.

26. **Myslinski, E., A. Krol, and P. Carbon.** 1991. Optimal $tRNA^{(ser)sec}$ gene activity requires an upstream SPH motif. Nucleic Acids Res. **20:**203–209.
27. **Rudenko, G., H.-M. Chung, V. P. Pham, and L. H. T. Van der Ploeg.** 1991. RNA polymerase I can mediate expression of CAT and *neo* protein-coding genes in the parasite protozoan *Trypanosoma brucei*. EMBO J. **10:**3387–3397.
28. **Rudenko, G., S. Le Blancq, J. Smith, M. G. S. Lee, A. Rattray, and L. H. T. Van der Ploeg.** 1990. Procyclic acidic repetitive protein (PARP) genes located in an unusually small α-amanitin-resistant transcription unit: PARP promoter activity assayed by transient DNA transfection of *Trypanosoma brucei*. Mol. Cell. Biol. **10:** 3492–3504.
29. **Sherman, D., L. Janz, M. Hug, and C. Clayton.** 1991. Anatomy of the *parp* gene promoter of *Trypanosoma brucei*. EMBO J. **10:** 3379–3386.
30. **Stefanovic, B., and W. F. Marzluff.** 1992. Characterization of two developmentally regulated sea urchin U2 small nuclear RNA promoters: a common required TATA sequence and independent proximal and distal elements. Mol. Cell. Biol. **12:**650–660.
31. **Steinberg, T. H., D. E. Mathews, R. D. Durbin, and R. R. Burgess.** 1990. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA polymerase III. J. Biol. Chem. **265:**499–505.
31a. **Tschudi, C.** Unpublished data.
32. **Tschudi, C., A. R. Krainer, and E. Ullu.** 1988. The U6 small nuclear RNA from *Trypanosoma brucei*. Nucleic Acids Res. **16:**11375.
33. **Tschudi, C., F. F. Richards, and E. Ullu.** 1986. The U2 RNA analogue of *Trypanosoma brucei gambiense*: implications for a splicing mechanism in trypanosomes. Nucleic Acids Res. **14:**8893–8903.
34. **Tschudi, C., and E. Ullu.** 1988. Polygene transcripts are precursors to calmodulin mRNAs in trypanosomes. EMBO J. **7:**455–463.
35. **Tschudi, C., and E. Ullu.** 1990. Destruction of U2, U4, or U6 small nuclear RNA blocks *trans* splicing in trypanosome cells. Cell **61:**459–466.
36. **Tschudi, C., S. P. Williams, and E. Ullu.** 1990. Conserved sequences in the U2 snRNA-encoding genes of *Kinetoplastida* do not include the putative branchpoint recognition region. Gene **91:**71–77.
37. **Ullu, E., and C. Tschudi.** 1990. Permeable trypanosome cells as a model system for transcription and trans-splicing. Nucleic Acids Res. **18:**3319–3326.
38. **Vankan, P., and W. Filipowicz.** 1989. A U-snRNA gene-specific upstream element and a −30 "TATA box" are required for transcription of the U2 snRNA gene of *Arabidopsis thaliana*. EMBO J. **8:**3875–3882.
39. **Waibel, F., and W. Filipowicz.** 1990. RNA-polymerase specificity of transcription of *Arabidopsis* U snRNA genes determined by promoter element spacing. Nature (London) **346:**199–202.
40. **Waibel, F., and W. Filipowicz.** 1990. U6 snRNA genes of *Arabidopsis* are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. Nucleic Acids Res. **18:**3459–3466.
41. **Wendelburg, B. J., and W. F. Marzluff.** 1992. Two promoter elements are necessary and sufficient for expression of the sea urchin U1 snRNA gene. Nucleic Acids Res. **20:**3743–3751.
42. **Zomerdijk, J. C. B. M., R. Kietf, P. G. Shield, and P. Borst.** 1991. Alpha-amanitin-resistant transcription units in trypanosomes: a comparison of promoter sequences for a VSG gene expression site and for the ribosomal RNA genes. Nucleic Acids Res. **19:**5153–5158.
43. **Zomerdijk, J. C. B. M., M. Ouellette, A. L. M. A. ten Asbroek, R. Kieft, A. M. M. Bommer, C. E. Clayton, and P. Borst.** 1990. The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. EMBO J. **9:**2791–2801.