

Sequence Elements in both the Intergenic Space and the 3' Untranslated Region of the *Crithidia fasciculata* *KAP3* Gene Are Required for Cell Cycle Regulation of *KAP3* mRNA

Nuraly K. Avliyakov, Jane C. Hines, and Dan S. Ray*

Molecular Biology Institute and Department of Microbiology, Immunology and Molecular Genetics,
University of California—Los Angeles, Los Angeles, California 90095-1570

Received 3 January 2003/Accepted 12 May 2003

mRNA levels of several *Crithidia fasciculata* genes involved in DNA metabolism have previously been found to cycle as cells progress through the cell cycle. Octamer consensus sequences in the 5' untranslated regions (5' UTRs) of these transcripts were shown to be required for cycling of these mRNAs. The *KAP3* gene encodes a kinetoplast histone H1-like DNA binding protein, and its mRNA levels cycle in parallel with those of the kinetoplast DNA topoisomerase (TOP2), dihydrofolate reductase-thymidylate synthase (DHFR-TS), and the large subunit of the nuclear single-stranded DNA binding protein (RPA1). *KAP3* mRNA contains two octamer consensus sequences in its 3' UTR but none in its 5' UTR. Mutation of these octamer sequences was not sufficient to prevent cycling of a sequence-tagged *KAP3* mRNA expressed from a plasmid. Mutation of an octamer sequence contained on the precursor transcript but not on the mRNA, in addition to mutation of the two octamer sequences in the 3' UTR, was necessary to abolish cycling of the mRNA. The requirement for a sequence not present on the mature mRNA indicates that regulation of the mRNA levels by the octamer sequences occurs at or prior to splicing of the transcript. Incompletely processed RNAs containing octamer sequences were also found to accumulate during the cell cycle when the mRNA levels were lowest. These RNA species hybridize to both the *KAP3* coding sequence and that of the downstream drug resistance gene, indicating a lack of processing within the intergenic region separating these genes. We propose a cell cycle-dependent interference in transcript processing mediated by octamer consensus sequences as a mechanism contributing to the cycling of such transcripts.

Gene expression in trypanosomes and other kinetoplastid protozoa differs significantly from that in most other eukaryotes. These unicellular parasites are among the earliest diverging organisms containing a mitochondrion, and many of their biological features are unique (21). One unusual feature of these organisms is the arrangement of genes along their chromosomes. The sequence of the *Leishmania major* chromosome I has revealed a 1.6-kb region lacking open reading frames (16). Predicted open reading frames to the left of this region are all oriented toward the left telomere, and those to the right are all oriented toward the right telomere. Also, in contrast to genes in most other eukaryotes, genes in kinetoplastids lack introns, with only a single exception (14). Transcription of genes on the *Leishmania* chromosome I appears to be bidirectional from within the 1.6-kb region. Similar studies of other chromosomes indicate that both promoters and terminators are present in the regions in which the direction of transcription changes (P. Myler, cited in reference 3).

Constitutive transcription of protein-coding genes generates polycistronic transcripts, which are processed to produce mature monocistronic mRNAs (reviewed in reference 3). Maturation of the polycistronic transcripts involves cleavage within intergenic spaces, *trans*-splicing of a 39-nucleotide (nt) spliced

leader sequence to the 5' end of all mRNAs, and polyadenylation of 3' ends. *trans*-splicing of precursor RNAs occurs by mechanisms similar to those for *cis*-splicing in higher eukaryotes and involves U2, U4, and U6 small nuclear RNAs (24). Processing of the precursor transcripts occurs rapidly, and polycistronic intermediates are generally not observed. Consequently, regulation of genes expressed by RNA polymerase II occurs predominantly at the posttranscriptional level.

Gene regulation in trypanosomes has been studied most extensively for species that undergo morphological and physiological transformations during their life cycles. In the case of *Trypanosoma brucei*, the causative agent of African sleeping sickness, variant surface glycoprotein gene mRNA shows a high degree of stage-specific expression. Sequence elements within the 3' untranslated regions (3' UTRs) of these genes have been shown to influence the mRNA levels (19, 25). A U-rich 26-mer in the 3' UTRs of the surface protein genes *EP* and *GPEET* in the insect form of *T. brucei* has been found to be important for the developmental regulation of these genes (9, 20). These U-rich elements appear to be similar to the destabilizing AU-rich elements present in the 3' UTRs of mammalian mRNAs involved in growth and differentiation (15). Similar sequence elements present in the mRNAs of *Trypanosoma cruzi* genes encoding mucin-like proteins have also been shown to regulate mRNA levels (5).

An octamer consensus sequence [(C/A)AUAGAA(G/A)] that is present in several mRNAs of the trypanosomatid

* Corresponding author. Mailing address: 301A Paul D. Boyer Hall, 611 Charles Young Dr. East, UCLA, Los Angeles, CA 90095-1570. Phone: (310) 825-4178. Fax: (310) 206-7286. E-mail: danray@ucla.edu.

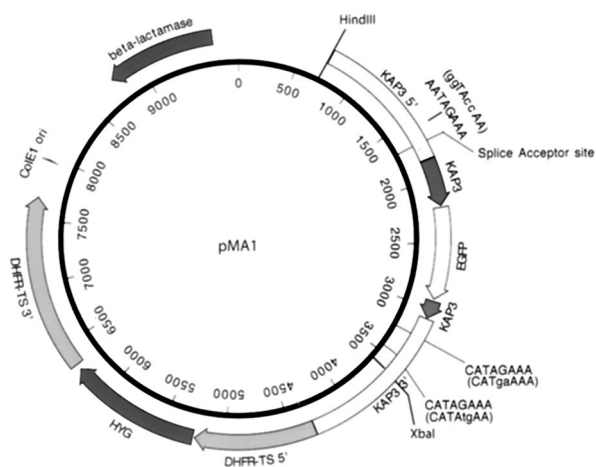


FIG. 1. Physical map of the KAP3-GFP expression plasmid pMA1. Full-length *EGFP* was fused in frame near the 3' end of the *KAP3* coding sequence as a tag to distinguish the chimeric transcript from that of the *KAP3* gene. The vector contains the HYG selectable marker flanked by the *L. major* 5' and 3' flanking regions of the *DHFR-TS* gene for expression of the HYG drug resistance marker. The 5' and 3' intergenic sequences flanking the *C. fasciculata KAP3* gene are indicated. The locations of the *KAP3* splice acceptor site and octameric sequences are also indicated. Mutated sequences are given in parentheses.

Crithidia fasciculata and is required for the cell cycle-dependent regulation of the mRNAs containing it has been described previously (2, 11, 13, 18). The mRNAs of genes encoding the large subunit of the nuclear single-stranded DNA-binding protein (RPA1), dihydrofolate reductase-thymidylate synthetase (DHFR-TS), the kinetoplast type II DNA topoisomerase (TOP2), and a histone H1-like kinetoplast DNA-binding protein (KAP3) all cycle in parallel as cells progress through the cell cycle. Levels of all these mRNAs are maximal just prior to S phase and then decline as DNA synthesis is completed. Episomal constructs bearing truncated *RPA1* or *TOP2* genes and their 5' UTRs allowed the identification of sequences in the 5' UTRs of these genes that are required for periodic accumulation of the mRNAs. Mutation of octamer consensus sequences present in the 5' UTRs of these genes abolished the periodic accumulation of the mRNAs and resulted in mRNA levels close to the maximum levels attained by the cycling transcripts and showing little variation throughout the cell cycle (2, 11). The central hexamer (AUAGAA) in the consensus sequence is absolutely conserved in the mRNAs that have been observed to cycle so far, and single-base changes in the hexamer have been shown to abolish cycling of the mRNA levels.

In the present work we have examined the sequence requirements for cycling of the *C. fasciculata KAP3* gene, a gene encoding a histone H1-like protein associated with kinetoplast DNA (8, 26). A plasmid construct containing the *KAP3* gene with 5' and 3' flanking sequences was used to insert the green fluorescent protein (GFP) gene in frame within *KAP3* for use as a sequence tag to permit easy discrimination of the expressed transcript from that of the chromosomal *KAP3* mRNA during the cell cycle. We show here that cycling of the *KAP3-GFP* transcript requires octamer sequence elements both in the *KAP3* 3' UTR and in the sequence 5' of the *KAP3* splice acceptor site.

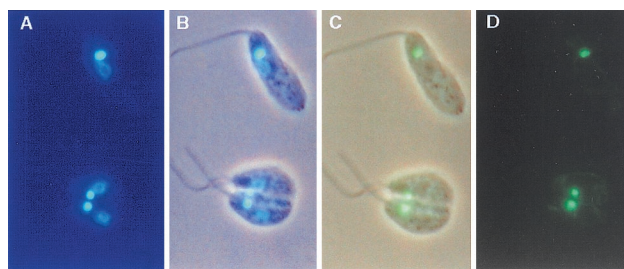


FIG. 2. Expression and mitochondrial targeting of the KAP3-GFP fusion protein in *C. fasciculata* cells transformed with plasmid pMA1. (A) DAPI staining; (B) phase-contrast image with DAPI staining; (C) phase-contrast image with GFP fluorescence; (D) GFP fluorescence.

MATERIALS AND METHODS

Isolation and cloning of the *KAP3* gene with 5' and 3' flanking sequences. A λ GEM11 genomic library of *C. fasciculata* was screened by using a labeled fragment of the *KAP3* coding region (8). Positive clones were picked and digested with *XhoI*. To construct plasmid pJCH1, a 2.1-kb gel-purified *XhoI* fragment of a λ GEM11 clone containing the *KAP3* gene, 1.0 kb of 5' flanking

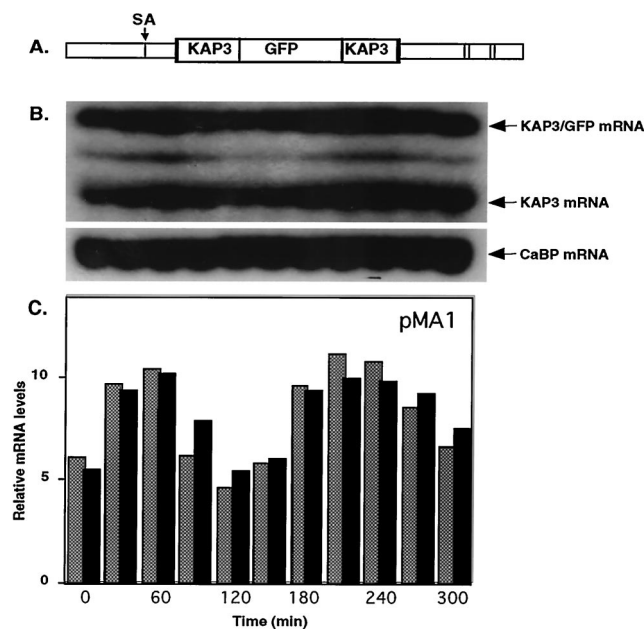


FIG. 3. Expression of endogenous *KAP3* and chimeric *KAP3-GFP* transcripts varies during the cell cycle. (A) Schematic structure (not to scale) of the *KAP3-GFP* gene and flanking sequences in pMA1. SA, the *KAP3* splice acceptor site. Unlabeled open rectangles indicate the locations of the CATAGAAA octamer sequences. (B) Northern blot analysis of total RNA isolated from synchronized *C. fasciculata* cells carrying the pMA1 plasmid. Total RNA was isolated at the time of release from hydroxyurea arrest (0 min) and at 30-min intervals thereafter. Total RNA (8 μ g per lane) was subjected to Northern blot analysis by hybridization to the *KAP3* gene coding sequences to detect both chimeric *KAP3-GFP* transcripts and endogenous *KAP3* transcripts. The same Northern blot was stripped and rehybridized to a 260-bp probe derived from a partial cDNA encoding the *C. fasciculata* homolog to the *T. cruzi* flagellar calcium binding protein (CaBP). (C) PhosphorImager quantitation of transcript levels in the Northern blot. Endogenous *KAP3* (solid bars) and chimeric *KAP3-GFP* (dark shaded bars) RNA transcript levels have been normalized at each time point relative to the level of the *CaBP* transcript, which is expressed at a constant level during the cell cycle and serves as a loading control.

sequence, and 0.7 kb of 3' flanking sequence was cloned into the *SaI* site of pX63HYG (4). The pX63HYG plasmid contains the hygromycin selectable marker flanked by the *L. major* 5' and 3' flanking regions of the *DHFR-TS* gene for expression of the drug resistance marker (hygromycin B phosphotransferase [HYG]). An additional ~1.2-kb *XbaI-Sau3AI* fragment containing the 3' flanking sequence of the *KAP3* gene was cloned into the *XbaI-BamHI* sites of pGEM7Z(+) (Promega) to create the pJL1 plasmid. The pJL1 plasmid was subsequently digested with *XbaI* and *SpeI*, and the resulting ~0.69-kb fragment was cloned into the *XbaI* site of pJCH1 to yield plasmid pJCH2.

Construction of a *KAP3-GFP* expression plasmid. To generate plasmid pMA1, expressing *KAP3-GFP* transcripts, the full coding region of the enhanced GFP (EGFP) gene was PCR amplified from plasmid pEGFP-1 (BD Clontech, Palo Alto, Calif.) with oligonucleotides G10 (5'-GCAGCCAAGACGGCCATGGTGAGCAAGGGCGAG-3') and G11 (5'-GAAGGCCGTCTTGGCGATATCCTGTACAGCTCGTCCATG-3') (underlining indicates added *BglI* and *EcoRV* sites), digested with *BglI*, and ligated into *BglI*-digested pJCH2.

Immunolocalization of the *KAP3-GFP* protein. *C. fasciculata* cells expressing the pMA1 plasmid at approximately 2×10^7 cells per ml were harvested by centrifugation (at $1,000 \times g$ for 5 min), washed twice in an equal volume of phosphate-buffered saline (PBS), and resuspended to 4×10^7 cells per ml. The cell suspension (50 μ l) was applied to a microscope slide which had been coated previously with 0.1% poly-L-lysine solution (Sigma-Aldrich). Slides were placed in a humid chamber for 30 min to allow cells to adhere. The slides were washed in PBS, immersed in 4% paraformaldehyde in PBS for 10 min, and then quenched by addition of glycine to a final concentration of 0.1 M. Slides were washed three times in PBS for 5 min, drained, and incubated with 100 μ l of a solution containing 0.2 μ g of 4',6'-diamidino-2-phenylindole (DAPI) per ml of PBS for 3 min. Finally, slides were rinsed with PBS, drained, and mounted in Vectashield (Vector Laboratories, Inc., Burlingame, Calif.). Cell images were captured on an Axioscope 2 microscope (Carl Zeiss, Jena, Germany), and images were processed with Adobe PhotoShop (version 6; Adobe Systems, Mountain View, Calif.).

Oligonucleotide mutagenesis. To create pMA1.1, plasmid pMA1 was digested with *HindIII* and *XbaI*, and a 2.75-kb *HindIII-XbaI* fragment containing the *KAP3-GFP* gene (see Fig. 1) was cloned into *HindIII*- and *XbaI*-digested pGEM11Z(+) (Promega) for mutagenesis. Mutation of the first conserved octameric sequence in the pMA1.1 plasmid was carried out by oligonucleotide mutagenesis using oligonucleotide E45 (5'-GCAATCGGCTTGTCATgaAAAA CATACAAATAC-3' [the octamer sequence is underlined and nucleotide substitutions within it are shown in lowercase]) and the Promega Gene Editor kit according to the manufacturer's instructions. Plasmid pMA1.1 was cut with *HindIII* and *XbaI*, and the 2.75-kb fragment released was then subcloned into the *HindIII-XbaI* sites of pGEM11Z (+) to generate the pMA1.2 plasmid. The pMA1.2 plasmid was used as a template in mutating the second octameric sequence with oligonucleotide F30 (5'-CAATGTGGAGTTCGTATTcaTATGTGATGGGCAACA-3') as described above. The base substitutions in pMA1.1 and pMA1.2 were confirmed by DNA sequence analysis. The pMA1.2 plasmid was digested with *HindIII* and *XbaI*, and the 2.75-kb fragment was used to replace the same *HindIII-XbaI* fragment of pMA1, creating pMA2.

To construct pMA5.4, we employed the cloning strategy described above. The 2.75-kb sequence was excised from plasmid pMA2 as a *HindIII-XbaI* fragment and inserted into *HindIII*- and *XbaI*-digested pGEM11Z(+) to produce plasmid pMA5.3. This plasmid was used to mutate the octameric (AATAGAAA) sequence upstream of the splice acceptor site of the *KAP3* gene to GGTACCAA by oligonucleotide mutagenesis using oligonucleotide H6 (5'-CATTCGGCGCA CAACAGCCACACAGCTTTGGTACCATCTCTCAAAGAAACACCCT-3') as described above. The base substitutions were confirmed by DNA sequencing. A 2.75-kb *HindIII-XbaI* fragment was released from pMA5.3 and then cloned into *HindIII*- and *XbaI*-digested pMA2 to yield pMA5.4.

Plasmid pMA5.1 was constructed as follows: plasmid pMA5.4 was digested with *EcoRV* and *XbaI*, and the 0.54-kb fragment was replaced with the corresponding *EcoRV-XbaI* sequence from plasmid pMA1. All plasmid constructs were electroporated into wild-type *C. fasciculata* cells as described previously (17).

Isolation and analysis of RNA. Hydroxyurea synchronization of *C. fasciculata* cells and isolation of total RNA were performed as described previously (17). Total RNA (8 to 10 μ g) was fractionated on 1.2% agarose-formaldehyde gels and transferred to nylon membranes (Hybond-XL; Amersham Biosciences, Piscataway, N.J.) by vacuum blotting in $20 \times$ SSC, pH 7.0 ($1 \times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate). RNA was cross-linked to the membrane by UV irradiation by use of a Stratilinker (Stratagene) according to the manufacturer's instructions. The membrane was prehybridized at 42°C in a solution containing 50% formamide, 5% Denhardt's reagent (EM Science), 1% sodium dodecyl

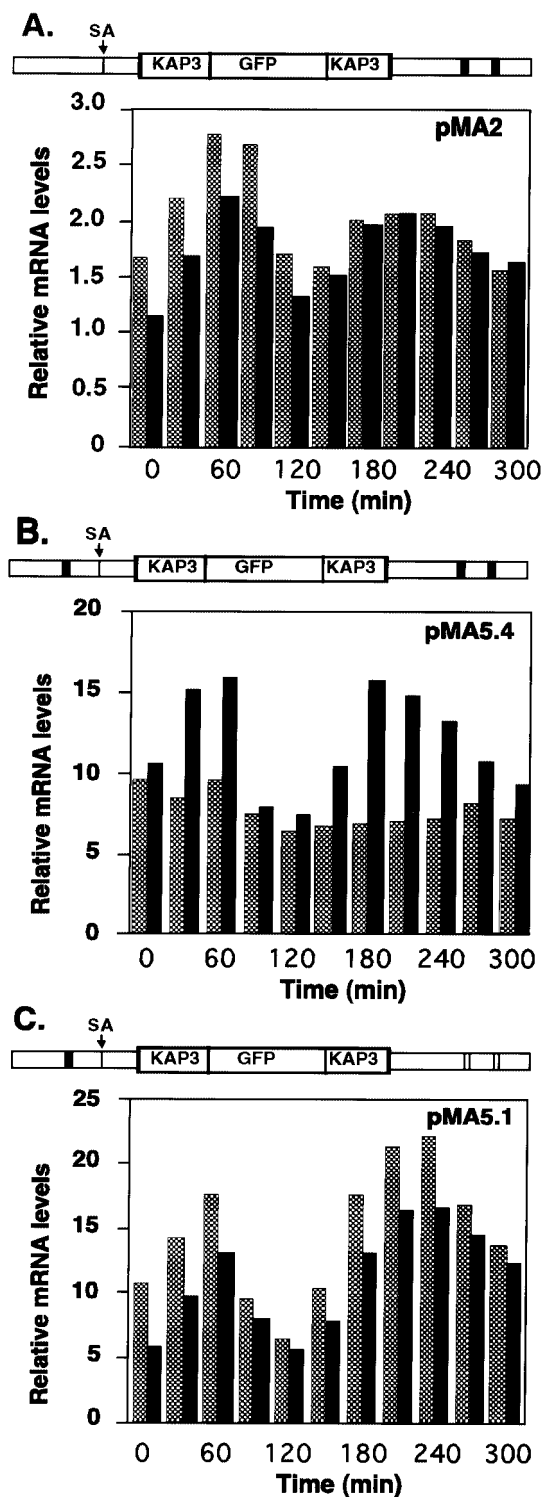


FIG. 4. Sequence elements required for cycling of *KAP3-GFP* mRNA levels. Shown is a PhosphorImager quantitation of Northern blots of total RNA isolated from synchronized *C. fasciculata* cells carrying either pMA2 (A), pMA5.4 (B), or pMA5.1 (C). RNA samples were isolated at 30-min intervals after release from hydroxyurea arrest and analyzed as for Fig. 3. A schematic diagram of the relevant region of each plasmid, with open rectangles representing wild-type consensus octamer sequence elements (CATAGAAA) and filled rectangles representing mutated octamer sequences, is shown above the respective bar graph. Solid bars, *KAP3* mRNA; dark shaded bars, *KAP3-GFP* mRNA.

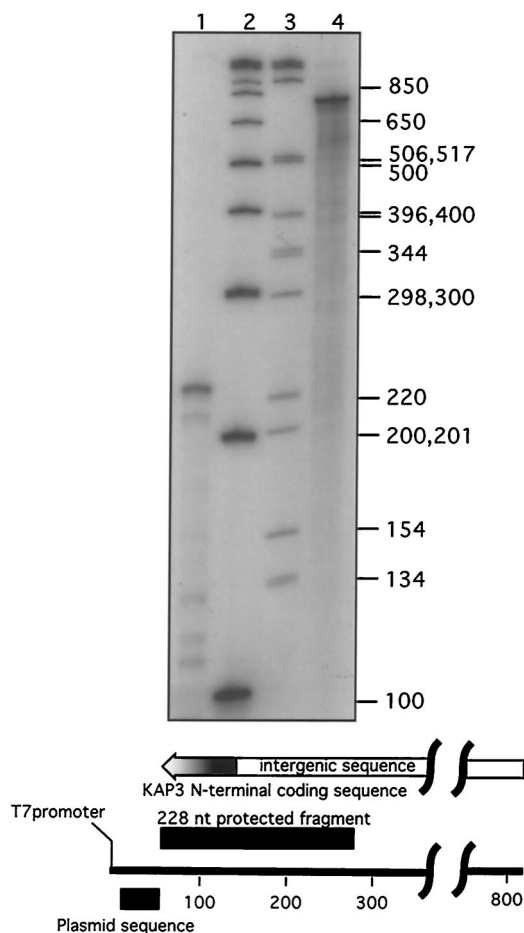


FIG. 5. RNase protection mapping of the *KAP3* splice acceptor site. *C. fasciculata* total-cell RNA (15 μ g) was hybridized to a 32 P-labeled RNA probe and digested with RNase A and RNase T1. The protected hybridized probe fragment was resolved on a 4% polyacrylamide gel containing 8 M urea. Lane 1, RNase A- and RNase T1-digested RNA; lanes 2 and 3, 32 P-labeled 1-kb Plus and 1-kb DNA markers (Invitrogen); lane 4, RNA probe. The probe and protected fragment are diagrammed underneath the autoradiogram.

sulfate (SDS), 0.75 M NaCl, 50 mM Na_2HPO_4 , 5 mM EDTA, 100 μ g of sonicated salmon sperm DNA/ml, and 100 μ g of yeast tRNA/ml for 3 h. A 405-bp DNA fragment encoding the full *KAP3* coding sequence was PCR amplified with oligonucleotides A49 (5'-CAATGCTGCGCCGTAGC-3') and A100 (5'-CGGG AATTCTACTTGCGCGCTGCCGTTGTC-3') and gel purified. The PCR product was labeled with the Random Primer Labeling beads (Amersham Bioscience, Piscataway, N.J.) and purified on a Sephadex G-50 column (Amersham Pharmacia Biotech). The membrane was hybridized at 42°C for 15 to 17 h and then washed twice with $2\times$ SSC-0.1% SDS at 60°C and twice with $0.1\times$ SSC-0.1% SDS at 60°C. To generate the *HYG* probe (a 1,026-bp DNA fragment carrying the *HYG* coding sequence), oligonucleotides K96 (5'-ATGAAAAAG CCTGAACCTCAC-3') and K97 (5'-CTATTCTTTGCCCTCGGAC-3') were used in a PCR with 20 ng of plasmid pX63HYG (4). The membrane was subsequently stripped and hybridized with the *CaBP* probe as a loading control. The *CaBP* probe was a partial cDNA of the putative *C. fasciculata* homolog of the *T. cruzi* 1F8 protein (17). Blots were quantitated by using a PhosphorImager and ImageQuant software (Molecular Dynamics).

Mapping of the *KAP3-GFP* splice acceptor site by PCR. The splice acceptor site of *KAP3-GFP* transcripts was determined by reverse transcription-PCR (RT-PCR). Total RNA was isolated from *C. fasciculata* cells expressing plasmid pMA1 by using RNeasy columns according to the manufacturer's directions (Qiagen). Total RNA (8 μ g) was treated with 6 U of DNase I (Ambion) for 15 min at 37°C. DNase I was inactivated for 10 min at 65°C. cDNA was synthesized

by using 0.3 μ g of DNase I-treated total RNA with the GFP-specific primer G11 (5'-GAAGGCCGTCTTGCGGATATCCTTGTACAGCTCGTCCATG-3') and Superscript II reverse transcriptase (Gibco BRL). The cDNA was used in PCRs with the spliced leader oligonucleotide B88 (5'-CCGGTACCATGGATCCA ACGG TATATAAGTATCAGTTTCTG-3') and the *KAP3* gene-specific oligonucleotide B46 (5'-TCGTTTTGCGAGGTACCATTCTTGG-3'). Thirty cycles of PCR consisting of 1 min at 94°C, 1 min at 58°C, and 1 min at 72°C were performed with *Taq* polymerase (Invitrogen), followed by 10 min at 72°C. PCR products were cloned using a Topo TA cloning vector (Invitrogen), and several clones were sequenced by dideoxynucleotide chain termination with modified T7 DNA polymerase (Sequenase 2.0) according to the manufacturer's instructions (U.S. Biochemicals).

Mapping of the *KAP3-GFP* splice acceptor site by RNase protection. An RNase protection assay was performed by using an Ambion RPAIII kit according to the manufacturer's instructions. *C. fasciculata* DNA-free total RNA (15 mg) was isolated by using an RNeasy kit (Qiagen) and hybridized with an 819-bp 32 P-labeled RNA probe (2.5×10^5 cpm; ca. 2 fmol) containing 55 nt of nonhybridizing sequence followed by 90 nt complementary to the *KAP3* coding sequence and 674 nt 5' of the *KAP3* start codon. RNA hybrids were digested with RNase A and RNase T1 and were electrophoresed on a 4% polyacrylamide gel containing 8 M urea. 32 P end-labeled DNA ladders (Invitrogen) were used as approximate size markers. The gel was dried and exposed to X-ray film.

RESULTS

The expression plasmid pMA1 (Fig. 1) was constructed in order to examine the sequence requirements for cycling of *KAP3* mRNA levels during the cell cycle. The plasmid contains the *KAP3* gene with 5' and 3' flanking sequences. The EGFP coding sequence was inserted in frame within the *KAP3* coding sequence to distinguish the plasmid-encoded *KAP3* mRNA from that of the endogenous gene. Sequence analysis of the chimeric gene and kinetoplast localization of the green fluorescence (Fig. 2) confirmed that the gene sequences were fused as expected and that a correctly localized protein product resulted.

C. fasciculata cells carrying plasmid pMA1 were synchronized by release from a hydroxyurea block in order to compare the expression of the chimeric transcript with that of the endogenous *KAP3* gene (8). Northern blot analysis of whole-cell RNAs isolated at 30-min intervals showed that the *KAP3-GFP* mRNA cycled in parallel with the endogenous *KAP3* mRNA as cells progressed through the cell cycle (Fig. 3). It should be noted that while these transcripts only vary about twofold during the cell cycle, the actual variation is likely much greater because the cells are not perfectly synchronous. A minor band migrating between those of *KAP3* and *KAP3-GFP* mRNAs is not observed in cells lacking the plasmid and may reflect a minor splicing event within the GFP sequence. The cycling of the *KAP3* and *KAP3-GFP* mRNAs is similar to that of the *C. fasciculata* *TOP2* and *RPA1* genes, which have been studied previously (2, 17, 18). Interestingly, both *TOP2* and *RPA1* mRNAs contain multiple copies of the consensus octamer sequence in their 5' UTRs. Cycling of the mRNAs of reporter constructs containing these 5' UTRs was shown to require the octamer sequences. In the *KAP3* gene, consensus octamer sequences are present only in the 3' UTR, which has two such sequences (Fig. 1). However, mutation of both of these sequences in the *KAP3-GFP* chimeric gene had no effect on the cycling of the mRNA (Fig. 4A).

Examination of the nucleotide sequence 5' of the identified *KAP3* splice acceptor revealed the sequence AATAGAAA 108 nt 5' of the splice acceptor site. Although this sequence conforms to the octamer consensus sequence in the mRNAs that

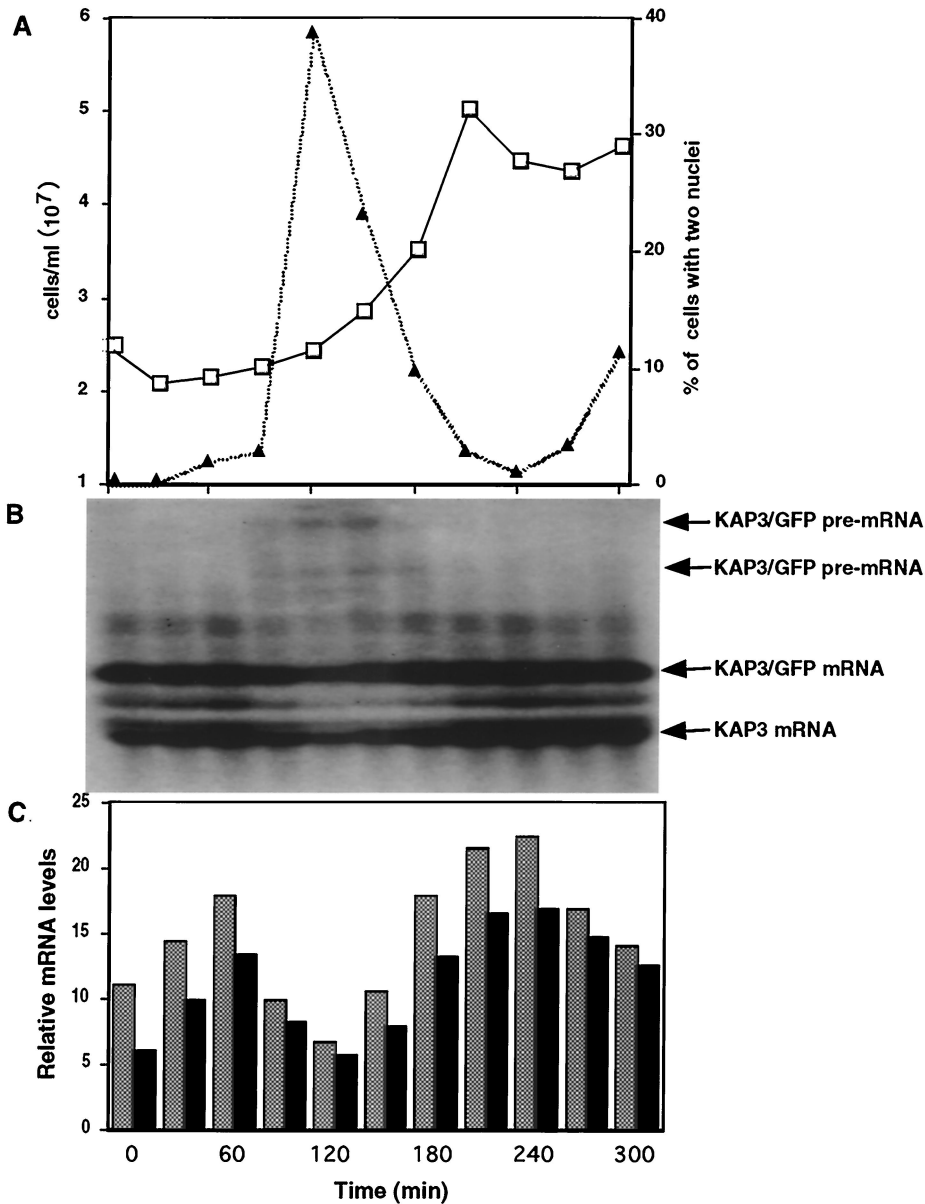


FIG. 6. Accumulation of pre-mRNA species of the *KAP3-GFP* chimeric transcript during progression through the cell cycle in *C. fasciculata* cells carrying pMA5.1. (A) The number of cells (squares) and the percentage of cells with two nuclei (triangles) were measured every 30 min following release from the hydroxyurea block to examine synchrony and cell cycle position. (B) Northern blot analysis of total RNA isolated at 30-min intervals. The blot was probed with *KAP3* coding sequences and then stripped and hybridized with the *CaBP* probe. (C) PhosphorImager quantitation of relative levels of chimeric *KAP3-GFP* and endogenous *KAP3* transcripts. Transcript levels have been normalized at each time point relative to the level of the *CaBP* transcript, which is expressed at a constant level throughout the cell cycle.

have been found to cycle (18), sequences required for cycling have so far been found only in UTRs. This additional octamer sequence was therefore mutagenized to determine whether it might have a role in the cycling of mRNA. In plasmid pMA5.4 all three octamer sequences were mutated (Fig. 4B). In this case, cycling of the chimeric transcript was abolished. Analysis of chimeric transcripts in cells containing plasmid pMA5.1, in which only the octamer sequence 5' of the splice acceptor site had been mutated, showed that disruption of this octamer sequence alone is not sufficient to abolish cycling of the transcript (Fig. 4C). Cycling of the mRNA was inhibited only when all three octamer sequences were mutated (Fig. 4B).

Since all previous octamer sequences required for cycling of an mRNA were found on the mature mRNA, the finding of an octamer sequence required for mRNA cycling that is present only on the pre-mRNA has important implications for the cycling mechanism. Because of the importance of this result, we questioned whether there might be an undiscovered splice acceptor site 5' of the octamer sequence located distal to the single identified splice acceptor. The known splice acceptor site was identified in several independent cDNA clones earlier by RT-PCR (26). Similar experiments have shown that the *KAP3-GFP* chimeric RNA is also spliced at this site in several independent isolates (data not shown). To confirm the exist-

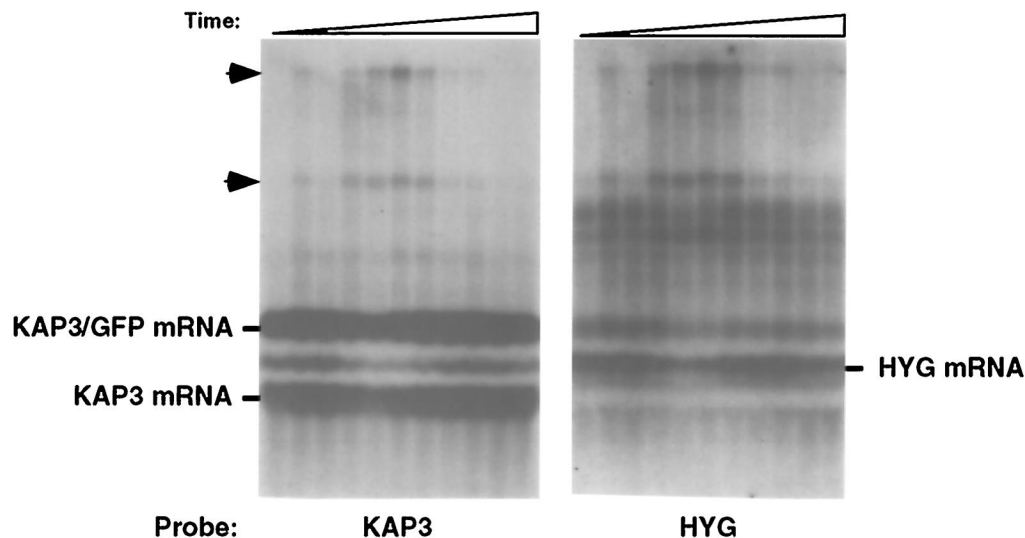


FIG. 7. Northern blot analysis of total RNA isolated at 30-min intervals from hydroxyurea-synchronized *C. fasciculata* cells carrying pMA5.1. (Right) The Northern blot was hybridized with *HYG* coding sequences. (Left) The same Northern blot was stripped and reprobed with *KAP3* coding sequences as for Fig. 6. The endogenous *KAP3*, chimeric *KAP3-GFP*, and *HYG* mRNA bands are indicated on the sides of the autoradiogram. Arrowheads indicate incompletely processed transcripts.

tence of only a single splice acceptor site, we have also performed RNase protection assays using an 819-nt labeled probe containing 90 nt of *KAP3* coding sequence and extending into the intergenic region to within 22 nt of the polyadenylation site of the upstream *KAP2* gene (Fig. 5). Splicing at the known splice acceptor site, at position -138 relative to the start codon, should yield a protected fragment of 228 nt. The strong band in Fig. 5, lane 1, is consistent with splicing at the single known splice acceptor site. No larger protected fragments are observed, in agreement with our earlier finding of only a single splice acceptor site by RT-PCR.

The requirement for an octamer sequence contained on the primary transcript beyond the splice acceptor site implies that the regulatory role of the sequence in mRNA cycling is performed prior to or in the process of the maturation of the transcript. Examination of longer exposures of Northern blots of cycling mRNAs showed that higher-molecular-weight pre-mRNA species accumulated at a time in the cell cycle when the mature-mRNA level was at its lowest. Figure 6 shows a Northern blot analysis of *KAP3-GFP* transcripts in cells carrying the pMA5.1 plasmid. The mature-mRNA level was at a minimum at 120 min after release from the hydroxyurea block, a time corresponding approximately to G_2/M . Higher-molecular-weight species were observed around the same time but were not detectable when the mRNA levels were highest, just prior to S phase. These higher-molecular-weight species have also been shown to hybridize to a GFP probe (data not shown). Similar accumulation of higher-molecular-weight pre-mRNAs has also been observed in cells carrying pMA1.

To further identify the higher-molecular-weight RNA species, we repeated the experiment for which results are shown in Fig. 6 except that the Northern blot was also separately probed with the coding sequence of the *HYG* gene located downstream of the *KAP3-GFP* chimeric gene (Fig. 1). The higher-molecular-weight RNAs detected by a *KAP3* probe also hybridized to a probe for the *HYG* gene (Fig. 7). This result

indicates that splicing and polyadenylation within the intergenic region separating these genes is inefficient at the time in the cell cycle when the cycling mRNA levels are lowest. We also note that the *HYG* mRNA levels cycle and that although the *HYG* mRNA does not contain a perfect match to the consensus octamer sequence, it does contain four closely related sequences in the *DHFR-TS* 5' and 3' UTR sequences used for expression of the *HYG* mRNA. These octamer sequences have a base substitution in the last nucleotide of the highly conserved central hexamer and may possibly be responsible for cycling of this transcript.

DISCUSSION

The work described here has identified sequence elements required for cycling of plasmid-encoded *KAP3-GFP* mRNA during the cell cycle. Consensus octamer sequences present in the 5' UTRs of the *C. fasciculata* *TOP2* and *RPA1* genes have previously been shown to be required for cycling of corresponding plasmid-encoded gene constructs containing their respective 5' flanking sequences and an N-terminal fragment of coding sequence (2, 18). In both cases a fragment within the 5' UTR of each gene was found to be capable of conferring cycling on a reporter gene. The *TOP2* fragment contained two copies of the consensus octamer sequence (C/A)AUAGAA(G/A). Mutation of both octamer sequences was necessary in order to completely abolish cycling of the mRNA level (18). Similarly, a fragment from the 5' UTR of *RPA1* containing two octamer sequences conferred cycling on a reporter transcript and functioned equally well when inserted 3' of coding sequences as when inserted into the 5' UTR. Mutation of the two octamer sequences substantially reduced cycling of the transcript (2).

The *KAP3* mRNA cycles in parallel with the *TOP2* and *RPA1* mRNAs and contains two copies of the consensus octamer sequence within the *KAP3* 3' UTR. The unexpected

finding that mutation of both octamer sequences in a *KAP3-GFP* mRNA did not have a significant effect on cycling of the *KAP3-GFP* mRNA led us to examine the possible contribution of a consensus octamer sequence present 5' of the only identified *KAP3* splice acceptor site. Mutation of this octamer sequence in addition to the two contained in the 3' UTR completely abolished cycling of the mRNA.

The requirement for a pre-mRNA sequence, not present on the mature mRNA, for cycling of the mRNA would imply an involvement of that sequence prior to or in the process of the maturation of the mRNA, since the 5' cleavage of the precursor RNA and addition of the spliced leader sequence result in the loss of sequences 5' of the splice acceptor site. The detection of pre-mRNA transcripts of the *KAP3-GFP* gene only at times in the cell cycle when the *KAP3-GFP* mRNA levels are lowest suggests that the octamer sequence elements may mediate a cell cycle-dependent inhibition of transcript processing leading to degradation of the transcript. The presence of downstream gene sequences on higher-molecular-weight *KAP3-GFP* transcripts when cells are dividing indicates a failure to *trans*-splice and polyadenylate within the intergenic space between these genes at this time in the cell cycle. Interference with transcript processing may contribute to transcript degradation and thus account for the decline in mRNA levels during this time. A requirement for multiple copies of a short sequence for modulating RNA splicing and mRNA levels has been observed in other systems. In higher eukaryotes, *cis*-splicing has been observed to be affected by adjacent intron sequences (1, 6, 7), and a repeated hexamer motif in an intron in the rat fibronectin gene has been shown to be important for splice site recognition (10). The role of repeated sequences in RNA splicing has also been documented in the regulation of *Drosophila doublesex* pre-mRNA splicing (22, 23). Products of the *tra* and *tra-2* genes positively regulate female-specific *dsx* splice acceptor usage.

The molecular mechanism by which the *C. fasciculata* octamer consensus sequences mediate the cell cycle regulation of mRNAs containing these sequences remains to be revealed. Since a single octamer sequence would have little informational capacity to recruit *trans*-acting factors, it is not surprising that multiple copies of the octamer in the transcript might be necessary to increase cycling. Identification of *trans*-acting factors will be essential for revealing the mechanism regulating mRNA cycling. A protein complex (CSBP) that binds with high sequence specificity to the consensus octamer sequence has been described previously (11–13). This complex contains subunits encoded by the *CSBPA* and *CSBPB* genes. The CSBP complex and a second complex (CSBP II) with similar binding specificity (unpublished data) are currently being investigated regarding their possible involvement in the cycling of mRNAs containing octamer sequence elements.

ACKNOWLEDGMENTS

We thank Julius Lukeš for constructing plasmid pJL1 and for helpful discussions.

This work was supported by National Institutes of Health grant GM53254.

REFERENCES

- Black, D. L. 1991. Does steric interference between splice sites block the splicing of a short *c-src* neuron-specific exon in non-neuronal cells? *Genes Dev.* **5**:389–402.
- Brown, L. M., and D. S. Ray. 1997. Cell cycle regulation of *RPA1* transcript levels in the trypanosomatid *Crithidia fasciculata*. *Nucleic Acids Res.* **25**:3281–3289.
- Clayton, C. E. 2002. Life without transcriptional control? From fly to man and back again. *EMBO J.* **21**:1881–1888.
- Cruz, A., C. M. Coburn, and S. M. Beverley. 1991. Double targeted gene replacement for creating null mutants. *Proc. Natl. Acad. Sci. USA* **88**:7170–7174.
- Di Noia, J. M., I. D'Orso, D. O. Sanchez, and A. C. Frasch. 2000. AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *J. Biol. Chem.* **275**:10218–10227.
- Gallego, M. E., L. Balvay, and E. Brody. 1992. *cis*-acting sequences involved in exon selection in the chicken beta-tropomyosin gene. *Mol. Cell. Biol.* **12**:5415–5425.
- Helfman, D. M., R. F. Roscigno, G. J. Mulligan, L. A. Finn, and K. S. Weber. 1990. Identification of two distinct intron elements involved in alternative splicing of beta-tropomyosin pre-mRNA. *Genes Dev.* **4**:98–110.
- Hines, J. C., and D. S. Ray. 1997. Tandem arrangement of two genes encoding kinetoplast-associated H1 histone-like proteins. *Mol. Biochem. Parasitol.* **89**:41–49.
- Hotz, H. R., P. Lorenz, R. Fischer, S. Krieger, and C. Clayton. 1995. Role of 3'-untranslated regions in the regulation of hexose transporter mRNAs in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **75**:1–14.
- Huh, G. S., and R. O. Hynes. 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes Dev.* **8**:1561–1574.
- Mahmood, R., J. C. Hines, and D. S. Ray. 1999. Identification of *cis* and *trans* elements involved in the cell cycle regulation of multiple genes in *Crithidia fasciculata*. *Mol. Cell. Biol.* **19**:6174–6182.
- Mahmood, R., B. Mittra, J. C. Hines, and D. S. Ray. 2001. Characterization of the *Crithidia fasciculata* mRNA cycling sequence binding proteins. *Mol. Cell. Biol.* **21**:4453–4459.
- Mahmood, R., and D. S. Ray. 1998. Nuclear extracts of *Crithidia fasciculata* contain factors that bind to the 5' untranslated regions of *TOP2* and *RPA1* mRNAs containing sequences required for their cell cycle regulation. *J. Biol. Chem.* **273**:23729–23734.
- Mair, G., H. Shi, H. Li, A. Djikeng, H. O. Aviles, J. R. Bishop, F. H. Falcone, C. Gavrilescu, J. L. Montgomery, M. I. Santori, L. S. Stern, Z. Wang, E. Ullu, and C. Tschudi. 2000. A new twist in trypanosome RNA metabolism: *cis*-splicing of pre-mRNA. *RNA* **6**:163–169.
- Mitchell, P., and D. Tollervey. 2000. mRNA stability in eukaryotes. *Curr. Opin. Genet. Dev.* **10**:193–198.
- Myler, P. J., L. Audleman, T. deVos, G. Hixson, P. Kiser, C. Lemley, C. Magness, E. Rickel, E. Sisk, S. Sunkin, S. Swartzell, T. Westlake, P. Bastien, G. Fu, A. Ivens, and K. Stuart. 1999. *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc. Natl. Acad. Sci. USA* **96**:2902–2906.
- Pasion, S. G., G. W. Brown, L. M. Brown, and D. S. Ray. 1994. Periodic expression of nuclear and mitochondrial DNA replication genes during the trypanosomatid cell cycle. *J. Cell Sci.* **107**:3515–3520.
- Pasion, S. G., J. C. Hines, X. Ou, R. Mahmood, and D. S. Ray. 1996. Sequences within the 5' untranslated region regulate the levels of a kinetoplast DNA topoisomerase mRNA during the cell cycle. *Mol. Cell. Biol.* **16**:6724–6735.
- Pays, E., S. Lips, D. Nolan, L. Vanhamme, and D. Perez-Morga. 2001. The VSG expression sites of *Trypanosoma brucei*: multipurpose tools for the adaptation of the parasite to mammalian hosts. *Mol. Biochem. Parasitol.* **114**:1–16.
- Schurch, N., A. Furger, U. Kurath, and I. Roditi. 1997. Contributions of the procyclin 3' untranslated region and coding region to the regulation of expression in bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **89**:109–121.
- Sogin, M. L., and J. D. Silberman. 1998. Evolution of the protists and protistan parasites from the perspective of molecular systematics. *Int. J. Parasitol.* **28**:11–20.
- Tian, M., and T. Maniatis. 1992. Positive control of pre-mRNA splicing in vitro. *Science* **256**:237–240.
- Tian, M., and T. Maniatis. 1993. A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell* **74**:105–114.
- 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.
- Vanhamme, L., E. Pays, R. McCulloch, and J. D. Barry. 2001. An update on antigenic variation in African trypanosomes. *Trends Parasitol.* **17**:338–343.
- Xu, C. W., J. C. Hines, M. L. Engel, D. G. Russell, and D. S. Ray. 1996. Nucleus-encoded histone H1-like proteins are associated with kinetoplast DNA in the trypanosomatid *Crithidia fasciculata*. *Mol. Cell. Biol.* **16**:564–576.