tbCPSF30 Depletion by RNA Interference Disrupts Polycistronic RNA Processing in Trypanosoma brucei*

Received for publication, March 7, 2003, and in revised form, April 15, 2003
Published, JBC Papers in Press, May 13, 2003, DOI 10.1074/jbc.M302405200

Edward F. Hendriks, Ammar Abdul-Razak, and Keith R. Matthews‡

From the School of Biological Sciences, Division of Biochemistry, 2.205 Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, United Kingdom

Gene expression in eukaryotes requires the post-transcriptional cleavage of mRNA precursors into mature mRNAs. In Trypanosoma brucei, mRNA processing is of particular importance, since most transcripts are derived from polycistronic transcription units. This organization dictates that regulated gene expression is promoter-independent and governed at the post-transcriptional level. We have identified tbCPSF30, a protein containing five CCCH zinc finger motifs, which is a homologue of the cleavage and polyadenylation specificity factor (CPSF) 30-kDa subunit, a component of the machinery required for 3′-end formation in yeast and mammals. Using gene silencing of tbCPSF30 by RNA interference, we demonstrate that this gene is essential in bloodstream and procyclic forms of Trypanosoma brucei. Interestingly, tbCPSF30-specific RNA interference results in the accumulation of an aberrant tbCPSF30 mRNA species concomitant with depletion of tbCPSF30 protein. tbCPSF30 protein depletion is accompanied by the accumulation of unprocessed tubulin RNAs, implicating tbCPSF30 in polycistronic RNA processing. By genome data base mining, we also identify several other putative components of the T. brucei cleavage and polyadenylation machinery, indicating their conservation throughout eukaryotic evolution. This study is the first to identify and characterize a core component of the T. brucei CPSF and show its involvement in polycistronic RNA processing.

The African trypanosome (Trypanosoma brucei) is the causative agent of sleeping sickness in humans and nagana in cattle. Trypanosomes are evolutionarily ancient organisms in which numerous unusual aspects of eukaryotic biology have been discovered, for example the possession of glycosomes, extensive antigenic variation mediated by genome recombination, and the RNA editing of mitochondrial transcripts. Also, the trypanosomatid family of protozoan parasites does not use the normal eukaryotic arrangement for gene expression (for a review, see Ref. 1). Instead, genes are organized into polycistronic transcription units, whereby many genes may be transcribed by an upstream promoter. Precursor RNAs are then processed into mRNA by the addition of a 39-nucleotide capped RNA (spliced leader) through a trans-splicing event (2, 3) and by cleavage and polyadenylation. A consequence of this arrangement is that the regulation of genes within such transcription units is not governed by promoter activity but by mRNA processing and stability. These organisms, therefore, represent a regulatory extreme in which the genome is almost exclusively controlled post-transcriptionally. In consequence, T. brucei is an interesting model for the study of the regulation of gene expression at the RNA level.

Analyses of RNA processing in trypanosomatids have largely focused on trans-splicing. Trans-splicing has been shown to be mechanistically similar to cis-splicing of introns in yeast and higher eukaryotes, and the basic components of both processes are conserved (for a review, see Ref. 4). In contrast, little is known about the process of trypanosome mRNA 3′-end formation, and no recognizable conserved motifs, such as the AAUAAA sequence of higher eukaryotes, are present upstream of trypanosome polyadenylation sites. Evidence also suggests that maturation of T. brucei polycistronic pre-mRNAs involves a temporal and mechanistic relationship between trans-splicing and 3′-end formation. It has been shown that the pyrimidine-rich sequences immediately upstream of the α-tubulin trans-splice site are a major determinant for β-tubulin mRNA 3′-end formation (5). Therefore, the sequence signals important for trans-splicing of the downstream gene are believed to contribute to 3′-end formation of the upstream gene in a polycistron. This is consistent with the possibility that the polypyrimidine tract is recognized by both the trans-splicing and polyadenylation machineries, either sequentially or simultaneously (6). The nematode Caenorhabditis elegans also undergoes trans-splicing during mRNA maturation, and, as in trypanosomes, this organism is believed to couple trans-splicing and polyadenylation to process precursor RNAs. Although several studies have concentrated on the identification of the components of the trans-splicing machinery in both T. brucei and C. elegans (7, 8), there have been no studies in either of these organisms concerning the cleavage and polyadenylation machinery to date. Thus, the mechanistic links between the trans-splicing machinery and 3′-end formation in polycistronic RNA processing have only been the result of investigating polyadenylation as a consequence of perturbations in the signals/machinery for trans-splicing (5, 6, 9).

A major goal of investigating the mechanisms of pre-mRNA 3′-end processing is the identification and functional characterization of the different trans-acting factors involved in the reaction. In a previous study, we identified a family of proteins in T. brucei, possessing a CCCH zinc finger motif (10), characteristic of RNA-binding proteins. In this study, we have used genome data base mining to identify a protein containing five CCCH fingers and show it to be the cleavage and polyadenylation specificity factor 30-kDa subunit of T. brucei (tbCPSF30). Here, we specifically perturb tbCPSF30 by RNA interference and show that tbCPSF30 is essential for T. brucei survival in
both the mammalian bloodstream and insect procyclic forms of the parasite. Interestingly, unprocessed tubulin transcripts accumulate when tbCPSF30 is depleted by RNA interference (RNAi), as has been seen previously when trans-splicing is disrupted. This implies that the integrity of the polyadenylation machinery as well as the trans-splicing machinery is required for polycistrionic RNA processing. The significant sequence conservation of yeast, worm, fly, zebrafish, mammalian, and now T. brucei CPSF30 homologues, suggests that all of these are members of a family of proteins conserved from the earliest eukaryotes to vertebrates. Moreover, identification of other core component homologues of the yeast and mammalian CPSF in the T. brucei CPSF complex, by data base mining, suggests a basic conservation of the CPSF complex among eukaryotes with diverse mechanisms for regulating gene expression.

**EXPERIMENTAL PROCEDURES**

Trypanosoma Cultures and Transfection—Monomorphic bloodstream form T. brucei rhodesiense EATRO 2340 were cultured in HMI-9 at 37 °C, 5% CO₂. For RNAi analysis, cultured T. brucei 427 “single marker T7 RNAP/TETR” bloodstream form cells (11) (SMB) were used. These cells have been engineered to express the T7 polymerase and a tetracycline repressor. RNAi analysis in procyclic forms was carried out in T. brucei 427 29-13 cells, which have also been engineered to express the T7 polymerase and the tetracycline repressor protein (11).

Transfection of both bloodstream and procyclic form parasites was carried out as described in Ref. 10 using 10 μg of Nodl (pHD451) or EcoRV (p27Ti) linearized DNA electroinjected into 500 μl of cells at 4 × 10⁷ cells ml⁻¹. Transfected cells were recovered overnight in 10 ml of HMI-9 at 37 °C, 5% CO₂ (bloodstream forms), or SDM-79 medium at 27 °C (procyclic forms) before being diluted to 1 × 10⁷ cells ml⁻¹ and subjected to drug selection in 24-well plates. Drug concentrations used for selection were hygromycin (2 μg ml⁻¹ for bloodstream forms, 20–50 μg ml⁻¹ for procyclic forms), G418 (2.5 μg ml⁻¹ for bloodstream forms, 15 μg ml⁻¹ for procyclic forms), or phleomycin (0.5–2.5 μg ml⁻¹ for bloodstream forms, 5 μg ml⁻¹ for procyclic forms). Selected cells were cloned by limiting dilution under drug selection.

Transgenic Plasmid Constructs—The trypanosome expression vector pH451 (12) was modified to contain the T71 epitope tag fused to the C terminus of genes cloned into the HindIII and BamHI sites. This vector was designated pH451-TY1. The tbCPSF30 coding region was amplified by PCR using primers CPSF30.F and CPSF30.R (CPSF30.F is CC AAC CCT ATG TTT ACT GAC AAC GCT GCC C; CPSF30.R is GGG ATC CTT GCC TTT CGG TTA CAT CAC) and cloned into HindIII/BamHI-digested pH451-TY1. This plasmid was then used to direct tbCPSF30 transgene expression in either bloodstream or procyclic form parasites that had been previously engineered to express the tetracycline repressor by stable transfection with the expression construct pH1449 (12). The tbCPSF30 RNAi construct was made using the vector p27Ti (13), which allows for tetracycline-inducible expression of double-stranded RNA from a T7 promoter in T. brucei 427 SMB or 427 29-13 procyclic cells. The tbCPSF30 coding region was amplified by PCR using primers CPSF30.F and CPSF30.R and cloned into the HindIII and BamHI sites of p27Ti. Further details of all constructs used in this paper are available upon request.

**Total RNA Isolation, cDNA Synthesis, and Northern Analysis**—Trypanosome RNA was isolated using Qiagen RNA-easy kit (by the manufacturer’s method), and Northern blotting was performed as described previously (14). Blots were hybridized at 60 °C in 0.1 × SSC. Blot detection was by chemiluminescence using CDP-star as a reaction substrate. The αβ tubulin intergenic region riboprobe was generated from a construct previously engineered in pBluescript II (5).

**Anti-peptide Antibody Production, Protein Isolation, and Western Blotting**—Anti-peptide antibody to T. brucei CPSF30 was generated using peptides comprising the sequence LINERADDPSFNKNTAC and QQWGGHRDAGTGRQ (see Fig. 1A). Both were coupled to keyhole limpet hemocyanin and used to immunize rabbits (Eugene, Belgium), the resulting serum being affinity-purified against each of the peptide antigens. Preimmune serum and immune serum were screened against whole cell trypanosome proteins. Preimmune serum did not show any reaction with T. brucei proteins.

The proteins were prepared as described in Ref. 15 and resolved on 11% polyacrylamide gels. Blotted proteins were incubated with either the anti-TV1 epitope tag monoclonal antibody (BB2, diluted 1:20) or with the affinity-purified rabbit anti-tbCPSF30 antibody (diluted 1:20). Western blots were processed using horseradish peroxidase-conjugated secondary antibody and visualized by chemiluminescence (Amersham Biosciences).

**Cell Growth, Microscopy, and Computational Analysis**—tbCPSF30 RNAi was induced by the addition of tetracycline (1 μg ml⁻¹), and growth was monitored over a period of 2 weeks. At intervals, ~2 × 10⁶ cells were harvested, spread onto microscope slides, and then air-dried for 10 min. The cells were then fixed in methanol at −20 °C for at least 20 min. The cells were rehydrated in PBS, and then slides were incubated with 4,6-diamidine-2, phenylindole for 5 min. Finally, the cells were washed in PBS three times and mounted in MOWIOL (Harlow Chemical Co., Kent, UK) containing phenylene diamine (1 mg ml⁻¹). Slides were examined on a Zeiss Axioscope 2 microscope, and images were captured using Scion Image version 1.62. Figures were processed using Adobe Photoshop 6.0. The tbCPSF30 locus was mapped using the T. brucei database at the Sanger Institute (available on the World Wide Web at www.sanger.ac.uk/Projects/T_brucei/) and at the Institute for Genomic Research (available on the World Wide Web at www.tigr.org/db/mlb/dbd/index.shtml). These same databases were screened, using Blast analysis (16), against both the yeast and mammalian proteins known to be involved in 3′-end formation.

**RESULTS**

tbCPSF30—As part of our interest in the family of proteins containing the RNA-binding CCCH zinc finger, we searched the trypanosome genome data base for genes encoding proteins containing this motif. As a search motif, we used the CCCH zinc finger sequence from tbZFP2, a molecule implicated in T. brucei cellular differentiation (10). This identified a sequence predicted to encode part of a protein containing five CCCH zinc finger motifs as well as two zinc knuckle motifs conforming to the CX₂CX₅CX₇C structure (Fig. 1A). tbZFP2 has a CX₂CX₅CX₇H zinc finger, typical of the family, whose best studied member is Trietetraprolin (TTP, also known as TIS11 and NUP475) (17, 18). In contrast, the five CCCH motifs in the newly identified gene had a degenerate spacing between the first and two coordinating cysteine residues in the sequence CX₂CX₅CX₇H and CX₂CX₅CX₇H; see Fig. 1B). Consequently, the gene, although a member of the tbZFP family of proteins as defined by the presence of the CCCH motif (Fig. 1B), represents a distinct subclass of this family.

The complete gene sequence was obtained by a combination of clone walking (within the data base) and PCR between the intervening sequence of contiguous clones. The resulting predicted sequence was then verified by sequencing the entire gene isolated from genomic DNA by PCR amplification. tBlastx analysis of the complete coding region demonstrated strong homology to the protein sequence for the cleavage and polyadenylation specificity factor 30-kDa subunit (CPSF30) from several organisms (34% identity and 65% similarity to yeast, 35% identity and 58% similarity to cow) (Fig. 1A). In addition to the conservation of the overall zinc finger motif structure (Fig. 1C), there were several further highly conserved regions of identity within the molecule and members of the CPSF30 family. On the basis of its strong homology to CPSF30 from several organisms, the gene was designated as the T. brucei cleavage and polyadenylation specificity factor 30-kDa subunit, tbCPSF30.

**tbCPSF30 Is Expressed in Bloodstream and Procyclic Life Cycle Stages**—In Drosophila, the CPSF30 homologue (Clipper, or CLP) is not ubiquitously expressed; it is absent during embryonic development (19). Since tbCPSF30 and CLP shared...
the greatest structural similarity (see Fig. 1C), the expression pattern of tbCPSF30 was assessed during the trypanosome developmental cycle. Total RNA was prepared from bloodstream and procyclic form parasites and hybridized with a riboprobe made to the coding region of tbCPSF30. This analysis revealed that tbCPSF30 was expressed at both these life cycle stages in similar abundance (Fig. 2A). The message was 1100 nucleotides, in agreement with the predicted size of the coding region (831 bp) plus 200–300 nucleotides of 5'- and 3'-untranslated regions. Southern analysis showed that tbCPSF30 was a single copy gene (data not shown).

In the case of the Drosophila CLP, the protein and RNA expression pattern do not coincide due to developmental regulation at the post-transcriptional level (20). To verify that the expression pattern observed for tbCPSF30 at the RNA level also reflected protein expression, an antibody was generated to two peptide antigens (LIERADDPSFNKNATC and QQWGGHRHGDATGRQ) derived from the C terminus of tbCPSF30, away from the conserved five CCCH zinc fingers and two zinc knuckle motifs. Protein samples were prepared from both bloodstream and procyclic forms and subjected to Western analysis with the specific tbCPSF30 anti-peptide antibody. The purified tbCPSF30 antibody detected two bands of 32–34 kDa (Fig. 2B). This double band was always observed in several different parental or wild type cell lines, although the relative proportion of each did vary between samples. We believe that both bands are tbCPSF30 and that they represent different forms of the protein (i.e., they may represent a processed or
Transfected cells were induced with tetracycline into bloodstream SMB cells and procyclic 29-13 cells, which had been transfected with a copy of tbCPSF30. Two bands were always detected with this antibody; the reason for this is not known. The protein is ~32-34 kDa in size and is equally expressed in bloodstream and procyclic forms. Preimmune tbCPSF30 serum did not detect any proteins (data not shown). Equal cell numbers were used to prepare each sample, and equivalent loading was confirmed by analysis of the Ponceau S-stained blot (not shown).

Gene Silencing of tbCPSF30 Causes Cell Death—The polyadenylation machinery comprises a complex of proteins of precise stoichiometry. We therefore assessed the functional consequences of ectopic overexpression and RNAi-mediated depletion of tbCPSF30. Initially, we used the trypanosome expression vector pH451 (12) to generate bloodstream and procyclic forms that overexpressed a TY-1 epitope-tagged copy of tbCPSF30. Transgenic tbCPSF30-TY could be clearly detected in both life cycle stages (Fig. 3A), being overexpressed 2-fold in bloodstream forms and 3–5-fold in the procyclic form with respect to the endogenous protein (data not shown). These cell lines were morphologically identical to the parental lines and grew at an identical rate (Fig. 3B), confirming that tbCPSF30 is expressed equally in both bloodstream and procyclic life cycle stages, contrasting with the developmentally regulated expression demonstrated by Drosophila CLP.

Northern analysis of tbCPSF30 expression in monomorphic bloodstream SMB and procyclic 29-13 forms, using the affinity-purified anti-peptide antibody to tbCPSF30 (αtbCPSF30). Two bands were always detected with this antibody; the reason for this is not known. The protein is ~32-34 kDa in size and is equally expressed in bloodstream and procyclic forms. Preimmune tbCPSF30 serum did not detect any proteins (data not shown). Equal cell numbers were used to prepare each sample, and equivalent loading was confirmed by analysis of the Ponceau S-stained blot (not shown).

Post-transcriptionally modified form of the protein. This has not been further examined; however, it is interesting to note that in vitro purification of mammalian polyadenylation complexes also indicated the presence of two forms of CPSF30 (28 and 30 kDa) (21). Moreover, both bands disappear upon tbCPSF30-specific transcript ablation (see below). Fig. 2B confirms that tbCPSF30 is expressed equally in both bloodstream and procyclic life cycle stages, contrasting with the developmentally regulated gene silencing demonstrated by Drosophila CLP.

Trypanosome RNA Processing
samples). We excluded the possibility that this novel transcript represented double-stranded RNA transcribed from the tbCPSF30 RNAi vector by hybridization with an antisense-specific riboprobe. This revealed that no detectable levels of tbCPSF30 double-stranded RNA were present in either the bloodstream or procyclic RNAi cell lines (data not shown). All of these facts strongly suggested that the observed tbCPSF30 RNA was tbCPSF30 RNAi-specific and did not represent a cryptic tbCPSF30 RNA derived from the tbCPSF30 RNAi 2T7i construct.

To directly link the observed growth phenotype with tbCPSF30 expression, we analyzed the protein levels for this molecule. Fig. 6 shows protein samples from induced and uninduced bloodstream tbCPSF30 RNAi cell lines probed with the tbCPSF30 anti-peptide antibody. These samples show 80–90% reduction in the level of tbCPSF30 protein by day 3–4 upon induction with tetracycline. This confirmed that gene silencing was operating in an efficient and inducible manner. Consistent with the observation that the bloodstream cells return to normal growth by day 10, the levels of tbCPSF30 protein returned to wild type levels by day 14 (Fig. 6). From this analysis, it is clear that there is a strict correlation between the presence of tbCPSF30 protein and the ability of *T. brucei* to survive in culture. It can therefore be concluded that tbCPSF30 is essential for viability in both the bloodstream and procyclic forms of *T. brucei*.

**tbCPSF30 Is Required for Polycistronic RNA Processing**—To investigate the consequences of tbCPSF30 depletion on trypanosome polycistronic RNA processing, we examined the resolution of pre-mRNA in the bloodstream and procyclic tbCPSF30 RNAi lines. In the case of the bloodstream tbCPSF30 RNAi cells, RNA was prepared 2 days postinduction, and in the case of the procyclic tbCPSF30 RNAi lines, samples were taken 3 days postinduction. This corresponds to the period before cell death in each population. To assay RNA processing, each RNA sample was hybridized with a riboprobe detecting the αβ-tubulin intergenic region. This probe detects the mature αβ-tubulin mRNA and unprocessed RNAs spanning the αβ-tubulin dicistron. It has been reported (5, 23) that perturbation of the trans-splicing machinery results in an accumulation of dicistronic αβ-tubulin RNA due to disruption of processing of the tubulin gene pre-mRNA. Our results in Fig. 7A clearly show the presence of an additional transcript in both the bloodstream and procyclic tbCPSF30 RNAi-induced cell lines, which
is consistent with the size of a larger unprocessed, tubulin pre-mRNA (dicistron). Significantly, this additional transcript is not seen in either the wild type or uninduced cell lines for either of the life cycle stages of the parasite.

Further confirmation of the specificity of this result was provided by analysis of the bloodstream samples that were tbCPSF30 RNAi-resistant (i.e. those isolated 12 days after induction of tbCPSF30 RNAi). In bloodstream tbCPSF30 RNAi samples taken 0, 2, and 12 days postinduction with tetracycline, the increase in the unprocessed α/β-tubulin intergenic region was specific to the samples with depleted tbCPSF30, with the outgrowing RNAi-resistant line behaving as wild type cells (Fig. 7B). This confirms that the perturbation of the processing of the tubulin polycistronic RNA was a direct result of the gene silencing of tbCPSF30.

In order to eliminate the possibility that the appearance of polycistronic RNA was simply the result of a general “dying cell” phenomenon, we allowed the parental RNAi bloodstream line (SMB) to reach maximal growth levels in culture (typically 5 × 10⁶ cells/ml). At this stage, cells are still intact and motile, although if left for 48 h without passage, these cultures overgrow and rapidly die. Therefore, these cultures were used as a population of cells undergoing several mechanisms (starvation, cell cycle defects, cytotoxic effects) resulting in cell death. Fig. 7C shows RNA from the cultures hybridized with a probe for the α/β-tubulin intergenic region. It is clear that in these dying cells there was no specific increase in unprocessed tubulin RNA. Rather, the accumulation of polycistronic RNA was specific to the depletion of tbCPSF30, demonstrating that this molecule contributes to the efficient processing of tubulin transcripts in *T. brucei*.

Although trans-splicing is ubiquitous for all genes in *T. brucei*, it was recently discovered that at least one gene (poly(A)-polymerase (PAP)) also undergoes cis-splicing. This raised the question as to whether this RNA processing event might also be disrupted in tbCPSF30-depleted cells. Thus, cis-splicing was assayed in these cells using RT-PCR across the PAP intron. No change in cis-splicing efficiency for the PAP transcript was detectable (data not shown).

**Components of the *T. brucei* RNA Processing Machinery**—Although there have been numerous studies of the components of both yeast and mammalian cleavage and polyadenylation complexes (for reviews, see Refs. 24 and 25), to date there have been no studies on the CPSF or associated factors in *T. brucei*. Since the *T. brucei* genome is organized into polycistrons, it is interesting to speculate that the core components of the polyadenylation and cleavage machinery might be different from other eukaryotes using monocistronic transcription. Consequently, by genome data base mining, we searched for components of the cleavage and polyadenylation machinery conserved in *T. brucei* (Table I). This revealed the presence of the previously cloned poly(A)-polymerase (26), a poly(A)-polymerase-binding protein homologue (PAPB1), and at least one other component of the cleavage and polyadenylation complex, CPSF73 (Table I). On the basis of sequence homology, other more putative component homologues of either yeast or mammalian proteins included CPSF100, CPSF160, CstF-50, CstF-68, PFS2, MPE1, HRP1, and PAN2/3 (Table I). The PAP, CPSF30, CPSF73, CPSF100, and CPSF160 form core components of the cleavage and polyadenylation complex in yeast and mammals. Thus, our survey indicates that several components of the cleavage and polyadenylation machinery are already detectable in the incomplete *T. brucei* genome data base. This suggests that even in this evolutionarily ancient protozoan parasite, with a very usual mechanism of gene expression, the basic machinery for the processing of transcripts is similar to other lower eukaryotes as well as higher organisms.

**DISCUSSION**

Gene regulation is vital to cell viability, growth, and development in all eukaryotic organisms. One aspect of gene regulation is the requirement for pre-mRNA processing into mature RNA species. This aspect is particularly important in trypanosomes whose genes are organized into polycistronic transcription units. Here, we report the first identification of a core component of the *T. brucei* cleavage and polyadenylation complex, tbCPSF30, and demonstrate that depletion of this protein following RNAi causes aberrant processing of tubulin polycistronic RNAs at two stages in the parasite life cycle. Following this phenotype, cell death occurs, indicating either that tbCPSF30 is itself essential in *T. brucei* or that inefficient or inappropriate RNA processing upon tbCPSF30 depletion is lethal. tbCPSF30 is, therefore, a key component in trypanosome gene expression.
ant bloodstream forms. The existence of this RNA can be rationalized in terms of recent understanding of the mechanism of RNA interference (27). This process is believed to be a cytoplasmic phenomenon effected by the action of siRNAs on mRNA, this being mediated by a complex of proteins conserved among species able to undergo RNAi. Recently, it has been suggested that RNAi operates at the level of the translational apparatus, whereby polyribosomal RNA is the target of degradation (28). In this scenario, the novel tbCPSF30 RNA that accumulates might represent nontranslated RNA protected from degradation by RNAi. This is directly supported by our observation that the aberrant RNA does not result in tb-CPSF30 protein, which is specifically depleted at the same time that the novel RNA is maximally expressed. One interesting model would be that the novel tbCPSF30 RNA is compartmentalized within the cell, for example being contained within the nucleus, and accumulates as the cell attempts to compensate for tbCPSF30 protein depletion.

In eukaryotes, the transcription of specific genes and the formation of mature mRNAs are processes that occur in the nucleus. Prior to nuclear export, mRNAs require several post-transcriptional modifications, which are tightly coupled. The precise mechanism that couples these post-transcriptional events is still poorly understood. In Leishmania and Trypanosoma, there is strong evidence that supports the model that polycistronic pre-mRNA polyadenylation requires active trans-

### Table 1

| Yeast protein | Mammalian protein | Expect value for nearest T. brucei match | T. brucei assignment |
|---------------|-------------------|------------------------------------------|----------------------|
| YTH1          | CPSF30            | 1.1 x 10^{-35} (M)                      | tbCPSF30             |
| BRR5/YSH1     | CPSF73            | 9 x 10^{-105} (M)                       | tbCPSF73             |
| CFT2/YDH1     | CPSF100           | 3.9 x 10^{-17} (M)                      | Possible CPSF100     |
| CFT1/YHH1     | CPSF160           | 9 x 10^{-24} (M)                        | Possible CPSF160     |
| PTA1          | None              |                                         |                      |
| FIP1          | None              |                                         |                      |
| PFS2          | None              |                                         |                      |
| PAP1          | None              |                                         |                      |
| PAB1          | None              |                                         |                      |
| PAN2          | None              |                                         |                      |
| PAN3          | None              |                                         |                      |
| RNA14         | CstF-77           | None                                     |                      |
| RNA15         | CstF-64           | None                                     |                      |
| MPE1          | None              | 9 x 10^{-10} (Y)                        | Possible MPE1        |
| CLP1          | None              |                                         |                      |
| PCF11         | None              |                                         |                      |
| HRP1          | None              | 4.4 x 10^{-13} (Y)                      | Possible HRP1        |
| -             | CstF-50           | 9.5 x 10^{-26} (M)                      | Possible CstF-50     |
| -             | CFI-68            | 1.1 x 10^{-23} (M)                      | Possible CFI-68      |
| -             | None              |                                         |                      |

*Y, yeast; M, mammalian.*

![FIG. 7. Processing of polycistronic RNA is compromised in tbCPSF30 depleted cells. A, left panel, Northern analysis of control bloodstream (BS) SMB cells and tbCPSF30 RNAi after 2 days in the presence (L) or absence of tetracycline (R) RNA samples were probed with the αβ-tubulin intergenic region. The probe will detect monocistronic α-tubulin as well as the αβ-tubulin dicistronic RNA. The same RNA analysis was performed on procyclic (PC) cell lines 3 days post-induction (right). Loading is shown by the EtBr panel. B, Northern analysis on a time course of RNA samples from bloodstream SMB cells and tbCPSF30 RNAi cells, 0, 2, and 12 days post-induction with tetracycline. These samples were probed with the αβ-tubulin intergenic region. This shows that the growth phenotype (Fig. 4, A and C) and RNAi effect correlate with the ability to process polycistronic tubulin RNA. Relative loading is shown by the EtBr panel. C, Northern analysis of RNA from SMB bloodstream cells allowed to overgrow and die. These cells were assessed under the light microscope (in terms of cell death; data not shown). Day 0 represents an overgrown culture in which cells are “viable” and intact. After 2 days of further culture, all cells were completely dead. RNA samples were probed with the αβ-tubulin intergenic region in order to determine the extent of polycistronic tubulin RNA in each sample. In no case was there an accumulation of unprocessed RNA. Loading is shown by the EtBr panel.

![TABLE I](https://example.com/table1.png)
The identification of a trypanosome homologue of CPSF30 and the results of data mining the available trypanosome genome information suggest that the basic components of the T. brucei 3′-end RNA processing complex(es) are also similar to S. cerevisiae and mammalian systems. This observation is particularly significant, given that T. brucei diverged very early in eukaryotic evolution and exhibits a gene organization that necessitates that 3′-end formation and transcript termination be unchanged. Clearly, organisms with very diverse mechanisms for gene expression conserve core elements of the machinery for cleavage and polyadenylation. Dissection of this complex will be fundamental to understanding the evolution of eukaryotic gene expression mechanisms and may provide new targets for the control of this important group of protozoan pathogens.

Acknowledgments—We thank Prof. Elisabetta Ullu and Dr. Frederick van Deursen for critical comments on the manuscript. Steve Whittaker provided excellent technical support. We are grateful to Prof. Walter Keller for yeast strains and plasmids and Dr. Barry Wilkinson for assistance with yeast complementation experiments. We are grateful to the Sanger Institute for permission to use the incomplete T. brucei genome database.

REFERENCES

1. Clayton, C. E. (2002) EMBO J. 21, 1881–1888
2. Murphy, W. J., Watkins, K. P., and Agabian, N. (1986) Cell 47, 517–525
3. Sutton, R. E., and Boothroyd, J. C. (1986) Cell 47, 527–535
4. Agabian, N. (1990) Cell 61, 1157–1160
5. Matthews, K. R., Tschudi, C., and Ullu, E. (1994) Genes Dev. 8, 491–501
6. LeBeowitz, J. H., Smith, H. Q., Rusche, L., and Beverley, S. M. (1983) Genes Dev. 9, 996–1007
7. Palfi, Z., Gunzl, A., Cross, M., and Bindereif, A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 9087–9101
8. Thomas, J., Lea, R., Zucker-Apapron, E., and Blumenthal, T. (1990) Nucleic Acids Res. 18, 2633–2642
9. Kuersten, S., Lea, K., MacMorris, M., Spieth, J., and Blumenthal, T. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 269–274
10. Hendriks, E. F., Robinson, D. R., Hinkins, M., and Matthews, K. R. (2001) EMBO J. 20, 6700–6711
11. Wirtz, E., Leal, S., Ochatt, C., and Cross, G. A. (1999) Mol. Biochem. Parasitol. 99, 89–101
12. Biehinger, S., Wirtz, L. E., Lorenz, P., and Clayton, C. (1997) Mol. Biochem. Parasitol. 85, 99–112
13. LoCount, D. J., Barrett, B., and Donelson, J. E. (2002) J. Biol. Chem. 277, 17580–17588
14. Tasker, M., Wilson, J., Sarkar, M., Hendriks, E., and Matthews, K. (2000) Mol. Biol. Cell 11, 1905–1917
15. Tasker, M., Timms, M., Hendriks, E., and Matthews, K. (2001) Mol. Microbiol. 39, 272–285
16. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
17. Lai, W. S., Carballo, E., Strum, J. R., Kennington, E. A., Phillips, R. S., and Blacksher, P. J. (1999) Mol. Cell. Biol. 19, 4311–4323
18. Carballo, E., Lai, W. S., and Blacksher, P. J. (1998) Science 281, 1001–1005
19. Bai, C., and Tolias, P. P. (1998) Mol. Cell. Biol. 18, 6661–6667
20. Bai, C., and Tolias, P. P. (1998) Nucleic Acids Res. 26, 1597–1604
21. Verardi, K. L., Edwads-Gilbert, G., MacDonald, C. C., Wallace, A. M., and Milcarek, C. (2000) RNA 6, 768–777
22. Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 14687–14692
23. Muhich, M. L., and Boothroyd, J. C. (1988) Mol. Cell. Biol. 8, 3837–3846
24. Wahrle, E., and Ruegsaeger, U. (1999) FEBS Microbiol. Rev. 23, 277–295
25. Zhao, J., Hayman, L., and Moore, C. (1999) Microbiol. Mol. Biol. Rev. 63, 405–445
26. Mair, G., Shi, H., Li, H., Dijkstra, A., Aviles, H. O., Bishop, J. R., Falcón, F. H., Guérinelle, C., Montgomerie, J. L., Santana, M. E., Stern, L. S., Wang, Z., Ullu, E., and Tschudi, C. (2000) RNA 6, 163–169
27. Hannon, G. J. (2002) Nature 418, 244–251
28. Kennerdell, J. R., Yamaguchi, S., and Carthew, R. W. (2002) Genes Dev. 16, 1884–1889
29. Ullu, E., Matthews, K. R., and Tschudi, C. (1993) Mol. Cell. Biol. 13, 720–725
30. Evans, D., Perez, I., MacMorris, M., Leake, D., Wiles, C. J., and Blumenthal, T. (1997) Genes Dev. 11, 3209–3217
31. Niwa, M., Rose, S. D., and Berget, S. M. (1990) Genes Dev. 4, 1552–1559
32. Niwa, M., and Berget, S. M. (1991) Genes Dev. 5, 2086–2095
33. Niwa, M., and Berget, S. M. (1991) Gene Expr. 1, 5–14
34. Niwa, M., MacDonald, C. C., and Berget, S. M. (1992) Nature 360, 277–280
35. Li, Y., Chen, Z. Y., Wang, W., Baker, C. C., and Krug, R. M. (2001) RNA 7, 490–501
36. Preker, P. J., Ohnacker, M., Minvielle-Sebastia, L., and Keller, W. (1997) EMBO J. 16, 4727–4737
37. Vo, L. T., Menet, M., Schmitter, J. M., Laaroute, F., and Wyers, F. (2001) Mol. Cell. Biol. 21, 8346–8356
38. Dichtl, B., and Keller, W. (2001) EMBO J. 20, 3197–3209
39. Barabino, S. M., Ohnacker, M., and Keller, W. (2000) EMBO J. 19, 3778–3787
40. Zhao, J., Kessler, M., Helmling, S., O’Connor, J. P., and Moore, C. (1999) Mol. Cell. Biol. 19, 7733–7740
41. Zarudnaya, M. I., Kolomiets, I. M., and Hovorun, D. M. (2002) IUBMB Life 54, 27–31
42. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
43. Barabino, S. M., Hubner, W., Jenny, A., Minvielle-Sebastia, L., and Keller, W. (1997) Genes Dev. 11, 1703–1716
44. Gaiano, N., Amsterdam, A., Kawakami, K., Allende, M., Becker, T., and Hopkins, N. (1996) Nature 383, 829–832
45. Hofmann, K., Bucher, P., Falquet, L., and Bairoch, A. (1999) Nucleic Acids Res. 27, 215–219
tbCPSF30 Depletion by RNA Interference Disrupts Polycistronic RNA Processing in *Trypanosoma brucei*
Edward F. Hendriks, Ammar Abdul-Razak and Keith R. Matthews

*J. Biol. Chem. 2003, 278:26870-26878.*
doi: 10.1074/jbc.M302405200 originally published online May 13, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M302405200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 45 references, 28 of which can be accessed free at [http://www.jbc.org/content/278/29/26870.full.html#ref-list-1](http://www.jbc.org/content/278/29/26870.full.html#ref-list-1)