A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in *Trypanosoma brucei*

Bernd Schimanski[1], Jens Brandenburg[1], Tu Ngoc Nguyen[1], Melissa Jo Caimano[2] and Arthur Günzl[1,3,*]

[1]Department of Genetics and Developmental Biology, [2]Department of Medicine and [3]Department of Molecular, Microbial and Structural Biology, University of Connecticut Health Center, Farmington, CT 06030, USA

ABSTRACT

The lack of general class II transcription factors was a hallmark of the genomic sequences of the human parasites *Trypanosoma brucei, Trypanosoma cruzi* and *Leishmania major*. However, the recent identification of TFIIA as part of a protein complex essential for RNA polymerase II-mediated transcription of *SLRNA* genes, which encode the *trans* splicing-specific spliced leader RNA, suggests that trypanosomatids assemble a highly divergent set of these factors at the *SLRNA* promoter. Here we report the identification of a trypanosomatid TFIIB-like (TFIIB_{like}) protein which has limited overall sequence homology to eukaryotic TFIIB and archaeal TFB but harbors conserved residues within the N-terminal zinc ribbon domain, the B finger and cyclin repeat I. In accordance with the function of TFIIB, *T.brucei* TFIIB_{like} is encoded by an essential gene, localizes to the nucleus, specifically binds to the *SLRNA* promoter, interacts with RNA polymerase II, and is absolutely required for *SLRNA* transcription.

INTRODUCTION

Initiation of transcription by eukaryotic RNA polymerase (pol) II is a highly conserved process mediated by a set of general transcription factors (1–3). The core factors required for this process are the TATA-binding protein (TBP) which binds to the TATA box, nucleating the formation of a pre-initiation complex, and the transcription factor (TF)IIB, which interacts with promoter DNA, TBP, the general factor TFIIF and RNA pol II to recruit the polymerase to the correct transcription initiation site (2). TFIIB harbors three distinct domains: an N-terminal zinc ribbon, which binds to the dock domain of RNA pol II and consists of a zinc binding motif and three anti-parallel beta strands (4–6), the B finger which

reaches into the active center of RNA pol II, and a C-terminal core domain which consists of two cyclin repeats facilitating the interaction with TBP and the promoter DNA (7–9). TBP and TFIIB have conserved orthologs in archaean prokaryotes in which the single RNA polymerase is directed to the transcription initiation site according to the same principle (10,11). Trypanosomatids, a family of unicellular flagellated parasites, appear to represent an exception to the above scheme. While these organisms harbor a single TBP homologue termed TBP-related factor 4 (TRF4) (12), analysis of the completed genomes from the human parasites *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*, also known as the Tritryp parasites, revealed no orthologs of the general transcription factors TFIIA, TFIIB, TFIIE and TFIIF (13–15). Trypanosomatid protein coding genes are arranged in tandem in directional gene clusters which are transcribed in a polycistronic manner (16,17). This leads to large precursor RNAs from which individual mRNAs are processed by spliced leader (SL) *trans* splicing and polyadenylation. In trans splicing, the same SL, which is derived from the 5′ terminus of the small nuclear SL RNA, is fused to the 5′ end of each mRNA. Hence, 5′ ends of trypanosomatid mRNAs are determined by the splice site, not the transcription initiation site. Accordingly, a specific RNA pol II initiation site has not been determined for protein coding gene transcription, suggesting that these parasites recruit RNA pol II to the DNA by a novel mechanism which is independent of general transcription factors.

However, trypanosomatids utilize RNA pol II also for the synthesis of SL RNA and, in contrast to the protein coding genes, SL RNA genes (*SLRNAs*) are transcribed monocistronically from a concrete transcription initiation site (18,19). SL RNA synthesis is of crucial importance to the parasites because maturation of each mRNA results in the consumption of one SL RNA molecule. To accommodate a high synthesis rate, SL RNA is encoded in ∼100 tandemly arranged *SLRNAs* per haploid genome. The *SLRNA* promoter is conserved within trypanosomatids and consists of a bipartite upstream sequence

---

element (USE) and an initiator element (20–23). Recently, a transcription factor complex was characterized in *T.brucei* which specifically binds the USE, is essential for *SLRNA* transcription *in vitro*, and consists of TRF4, the small nuclear RNA activating complex (SNAPc) and a highly divergent TFIIA (24,25). In the human system, the same factors are involved in RNA pol II-mediated transcription of small nuclear RNA genes [reviewed in (26)]. However, this enzymatic activity also requires other general transcription factors including TFIIB (27). Hence, the parallels to the human system and the identification of TFIIA in *T.brucei* suggest that, in contrast to the predictions of the genome analyses, trypanosomatids do harbor a highly divergent set of general transcription factors. Here we report the identification of a trypanosomatid TFIIB-like (TFIIB$_{like}$) protein which, as would be anticipated for TFIIB, is an essential transcription factor of *T.brucei* *SLRNA* transcription interacting with RNA pol II.

MATERIALS AND METHODS

Plasmids

The transfection vector pTFIIB$_{like}$-RNAi is a derivative of the *T.brucei* stem–loop RNAi vector (28,29) containing the *TFIIB$_{like}$* coding region from position 534 to 1008 as inverted repeats around a stuffer sequence. pTFIIB$_{like}$-PTP-NEO was derived from pC-PTP-NEO (30) by fusing 827 bp of the C-terminal *TFIIB$_{like}$* coding sequence to the PTP sequence via restriction sites ApaI and NotI. For the dot blot analysis, pTZ18U-derived plasmids were generated in which the complete coding regions of SL RNA, GPEET procyclin, α tubulin, HSP70, 18 S rRNA, U2 snRNA or U6 snRNA were cloned into the SmaI site of the polylinker.

Cells

Procyclic cell culture, targeted integration of linear DNAs into cells by electroporation, and the generation of stable cell lines by selection and limiting dilution were described in detail previously (31–33). For *TFIIB$_{like}$* silencing, *T.brucei* 29.13.6 cells (34) were transfected with 10 µg of the SacII-linearized construct pTFIIB$_{like}$-RNAi and cloned by limited dilution in the presence of 50 µg/ml of hygromycin, 15 µg/ml of G418 and 2.5 µg/ml of phleomycin. Silencing of *TFIIB$_{like}$* expression was induced by adding doxycycline to the culture medium at a final concentration of 10 µg/ml. Cells were counted daily and diluted to 2×10^6 cells/ml. The clonal cell line TbH8 which exclusively expressed PTP-tagged TFIIB$_{like}$ was generated in a first step by integrating StuI-linearized plasmid TFIIB$_{like}$-PTP-NEO into one *TFIIB$_{like}$* allele. In a second step, the remaining *TFIIB$_{like}$* allele was replaced by a DNA amplification product of the hygromycin phosphotransferase coding region which was flanked by 100 nt of *TFIIB$_{like}$* untranslated region on either side. TbH8 cells were resistant to 40 µg/ml of G418 and 20 µg/ml of hygromycin. For each cell line, correct construct integration was confirmed by PCR and Southern analysis.

RNA analysis

For steady state analysis of SL RNA and U snRNA abundance in cells in which *TFIIB$_{like}$* silencing had been induced, total RNA was prepared from induced and non-induced cells and assayed by primer extension using the 5′-^{32}P-end labeled oligonucleotides SLf and U2f (35), and the Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. For nascent RNA labeling, cells were permeabilized with the detergent lysolecithin as detailed elsewhere (36), mixed into a transcription cocktail containing 20 mM potassium L-glutamate, 20 mM HEPES–KOH, pH 7.7, 3 mM MgCl$_2$, 1 mM DTT, 25 mM creatine phosphate, 0.6 mg/ml creatine kinase, 10 mg/ml leupeptin, 10 mg/ml aprotinin, 2 mM ATP, 1 mM GTP, 1 mM CTP, 5 µM UTP and 100 µCi of [α-^{32}P]UTP (3000 Ci/mmole); and incubated for 8 min at 28°C. A small aliquot of total RNA prepared from these cells was separated on 6% polyacrylamide–50% urea gels and visualized by autoradiography. For dot blot assays, denatured plasmid DNA (5 µg per dot) was blotted onto positively charged nylon membrane and hybridized to the remainder of the radio-labeled RNA for 16 h at 55°C. Blots were then washed twice with 2× SSC/0.1% SDS for 5 min at room temperature and twice with 0.2× SSC/0.1% SDS for 15 min at 65°C. Signals were visualized by autoradiography and quantified by scintillation counting.

Immunofluorescence

Cells were harvested, washed once with phosphate-buffered saline (PBS, pH 7.4), and resuspended in 60 µl of PBS. Eight microliters of the cell suspension were fixed onto a glass slide overnight in methanol at −20°C. The slides were incubated with rat anti-TFIIB$_{like}$ polyclonal serum, or the corresponding pre-immune serum diluted 1:100 in PBS/1% BSA. After washing with PBS containing 0.1% Tween 20 (Sigma, St Louis, MO), 4,6-diamidino-2-phenylindol (DAPI, 2 µg/ml) and a 1:400 dilution of the secondary anti-rat IgG conjugated with fluorescein isothiocyanate (FITC) were added. The rinsed samples were mounted in Vectashield (Vector laboratories); images were made using a Zeiss Axiovert 200 microscope with identical exposure settings and analyzed with AxioVision software.

Protein analysis

TFIIB$_{like}$-PTP was purified from crude TbH8 extract by IgG affinity chromatography, TEV protease treatment and anti-ProtC immunoaffinity chromatography exactly as described previously (30). Purified proteins were separated on a 10–20% SDS–polyacrylamide gradient gel and stained with Pierce Gelcode Coomassie stain (Pierce, Rockford, IL). Protein bands were excised and analyzed by MALDI-TOF mass spectrometry. For the expression and purification of TFIIB$_{like}$ in *Escherichia coli*, the complete coding sequence of this protein was C-terminally fused to ProtC, a thrombin cleavage site and the 6× His tag in the expression vector pET100/D-TOPO (Invitrogen). The recombinant protein was expressed in BL21Star (DE3) cells upon induction with 1 mM IPTG at 37°C for 4 h. rTFIIB$_{like}$ was purified by TALON metal affinity chromatography (BD Biosciences), thrombin cleavage and anti-ProtC immunoaffinity chromatography. For functional analysis, rTFIIB$_{like}$ was directly eluted into transcription buffer in the presence of 0.5 mg/ml of ProtC peptide. Monospecific polyclonal antiserum directed against TFIIB$_{like}$ was generated by hyper-immunizing female Sprague Dawley

rats with purified rTFIIB$_{like}$ protein mixed 1:1 with Freunds complete adjuvant (Sigma) which was administered intraperitoneally. The rats were boosted after 3 and 6 weeks with the same antigen mixed 1:1 with Freunds incomplete adjuvant (Sigma). Pre-immune sera was collected from each animal prior to immunization to serve as controls. All animal handling and procedures were performed according to protocols approved by the University of Connecticut Animal Care Committee. For the bioinformatic analysis of the TFIIB$_{like}$ zinc ribbon domain, the Superfamily server was used and human TFIIB protein sequence served as the query for the ten Hidden Markov models in the zinc beta-ribbon superfamily. The trypanosomatids sequences were then aligned against these models.

Functional *in vitro* analysis

Promoter pull-downs were carried out as described previously (37) except that the salt concentration in the binding reaction and in the washing buffers was increased to 80 mM potassium chloride. The *in vitro* transcription system has been detailed elsewhere (31,38). Briefly, reactions of 40 µl containing 8 µl of extract, 20 mM potassium L-glutamate, 20 mM KCl, 3 mM MgCl$_2$, 20 mM HEPES–KOH, pH 7.7, 0.5 mM of each nucleoside triphosphate (NTP), 20 mM creatine phosphate, 0.48 mg/ml of creatine kinase, 2.5% polyethylene glycol, 0.2 mM EDTA, 0.5 mM EGTA, 4 mM DTT, 10 mg/ml leupeptin, 10 mg/ml aprotinin, 12.5 µg/ml vector DNA, 20 µg/ml GPEET-trm template and 7.5 µg/ml SLins19 template were incubated for 1 h at 27°C. In addition, the ProtC peptide was added to a final concentration of 0.1 mg/ml in each reaction to eliminate a nonspecific effect in reconstitution assays performed with peptide-eluted protein. Specific transcripts from GPEET-trm and SLins19 were analyzed by extension of ^{32}P-end-labeled primers Tag_PE (31) and SLtag (21) which hybridize to unrelated oligonucleotide tags of GPEET-trm and SLins19, respectively.

RESULTS AND DISCUSSION

Identification of a trypanosomatid TFIIB$_{like}$ protein

Homology searches of trypanosomatid genome databanks with human or yeast TFIIB amino acid sequences yielded only one significant match to an ortholog of the TFIIB-related protein 1 (BRF1), a subunit of transcription factor TFIIIB involved in RNA pol III-mediated transcription initiation (data not shown). By analyzing all sequences from the homology search for the presence of one of the key features of the TFIIB family of proteins, namely, an N-terminal zinc binding motif, we identified a candidate TFIIB protein which we tentatively designated as TFIIB$_{like}$. The corresponding gene is present in each of the Tritryp parasites and has been annotated to encode a hypothetical conserved protein. The *T.brucei* TFIIB$_{like}$ consists of 345 amino acids and has a predicted mass of 37.6 kDa and a theoretical p*I* of 6.28 (accession numbers: GeneDB Tb09.160.4220, cDNA AM235385). These attributes are similar to those of other eukaryotic TFIIBs (data not shown). All three trypanosomatid sequences contain the zinc-binding motif and conserved residues in most positions of the zinc ribbon beta strands (Figure 1A). Using the Superfamily Server (39), the trypanosomatid proteins were

predicted to form a corresponding beta sheet in this domain (data not shown). In TFIIB/TFB, the B finger is located adjacent to the zinc ribbon. The trypanosomatid sequences contain a highly conserved sequence element at this position, but it deviates substantially from the B finger consensus of other organisms. However, the B finger key residues (Glu51 and Arg66 in the human sequence), which form a salt bridge and are essential for the integrity of this structure, are conserved in the Tritryp parasites (Figure 1A) [Refs (6,40)]. Moreover, in the TFIIB core domain, the trypanosomatid sequences exhibit many conserved residues within the first of the two cyclin repeats (Figure 1A). On the other hand, cyclin repeat 2 is much less well conserved and most of the mapped contact residues of human TFIIB with TBP and promoter DNA (7) are absent in the trypanosomatid sequences (Figure 1A and data not shown).

Phylogenetic analysis also suggested that the trypanosomatid sequences are more divergent than their putative eukaryotic and archaeal counterparts (Figure 1B). This result was not due to insertions or C-terminal extensions present in the trypanosomatid TFIIB$_{like}$ sequences because their exclusion from the analysis did not significantly change the phylogenetic tree (data not shown). The accuracy of the tree is further supported by the correct clustering of archaeal, trypanosomatid, fungal and metazoan species. Taken together, these data suggest that trypanosomatid parasites harbor a highly divergent member of the TFIIB family of proteins.

Silencing of *T.brucei* TFIIB$_{like}$ is lethal and affects the synthesis of SL RNA

To evaluate whether TFIIB$_{like}$ is a factor of *SLRNA* transcription, we silenced *T.brucei* TFIIB$_{like}$ expression *in vivo* by conditional RNAi. For this, we employed the tetracycline-inducible RNAi system in *T.brucei* which involves the stable transfection of a construct for double-stranded RNA (dsRNA) synthesis under the control of tetracycline repressor into trypanosomes constitutively expressing this bacterial repressor protein (34). After transfection of a stem-loop vector (28,29) harboring both strands of the *TFIIB$_{like}$* coding region from position 534 to 1008, three clonal cell lines were obtained and raised by limiting dilution. In the absence of antibiotic, the cell growth in those lines was logarithmic. However, when *TFIIB$_{like}$* dsRNA synthesis was induced by adding doxycycline to the medium, cells began to die 48 h after induction, leading to a cessation of cell growth (Figure 2A). The partial growth recovery beginning at day six is a phenotype commonly observed in trypanosomes and attributed to the emergence of RNAi-defective cells (41). Northern blot analysis confirmed that *TFIIB$_{like}$* mRNA was quantitatively degraded within 48 h after induction of *TFIIB$_{like}$* dsRNA synthesis (Figure 2B). Similarly, immunodetection of TFIIB$_{like}$ protein with a polyclonal antibody raised against the purified recombinant protein showed that only trace amounts of TFIIB$_{like}$ were detectable 48 and 54 h after induction (Figure 2C). We concluded, therefore, that TFIIB$_{like}$ is essential in trypanosomes.

Since TFIIB$_{like}$ may be involved in SL RNA synthesis, we first determined the abundance of SL RNA in total RNA preparations from cells in which *TFIIB$_{like}$* expression was silenced. A primer extension assay revealed that, in
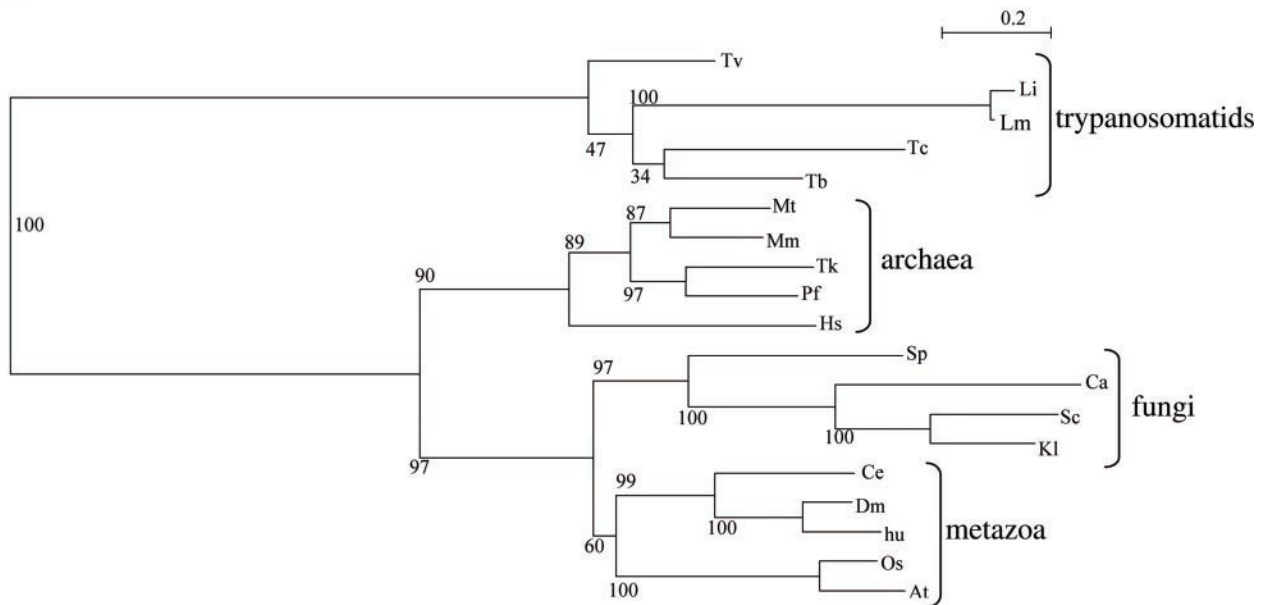
Figure 1. Trypanosomatid TFIIB$_{like}$ is highly divergent to eukaryotic TFIIB and archaeal TFB. (**A**) Conserved TFIIB domains were aligned manually (B finger) or with ClustalW. TFIIB sequences are from human (hu, accession number AAH21000), *Drosophila melanogaster* (Dm, P29052), *Caenorhabditis elegans* (Ce, O16991), *Arabidopsis thaliana* (At, NP_181694), *Saccharomyces cerevisiae* (Sc, NP_015411) and *Schizosaccharomyces pombe* (Sp, O13749), TFIIB$_{like}$ sequences from *T.brucei* (Tb), *T.cruzi* (Tc, Tc00.1047053505983.30) and *L.major* (Lm, LmjF15.1170); the TFB sequence is from *Thermococcus kodakarensis* (Tk, P58109). Identities and similarities are shaded in black and gray, respectively. Only positions with a minimum of six identical or conserved residues are shaded. Parasite-specific identities are shaded in red. Asterisks indicate the zinc-binding residues in the zinc ribbon, as well as key B finger residues, while arrows denote the beta-strands of the zinc ribbon. TFIIB residues which were shown previously to interact with TBP and promoter DNA (7) are marked with 't' and 'd', respectively. (**B**) Phylogenetic tree constructed with the ClustalX program which uses the neighbor-joining method. The bootstrap values obtained by sampling a thousand replicates are indicated as percentages. At the top right corner is the scale bar for branch length. In addition to the sequences listed above, the following sequences were analyzed: Trypanosomatidae—*Leishmania infantum* (Li, accession number LinJ15.1140); *Trypanosoma vivax* (Tv, tviv1354f09.q1k_4); archaea—*Halobacterium salinarum* (Hs, Q9HRE6), *Methanosarcina mazei* (Mm, Q977U3), *Methanothermococcus thermolithotrophicus* (Mt, Q9P9I7), *Pyrococcus furiosus* (Pf, P61998); fungi—*Candida albicans* (Ca, XP_710164), *Kluyveromyces lactis* (Kl, Q05959); metazoa—*Oryza sativa* (Os, Q8W0W3).
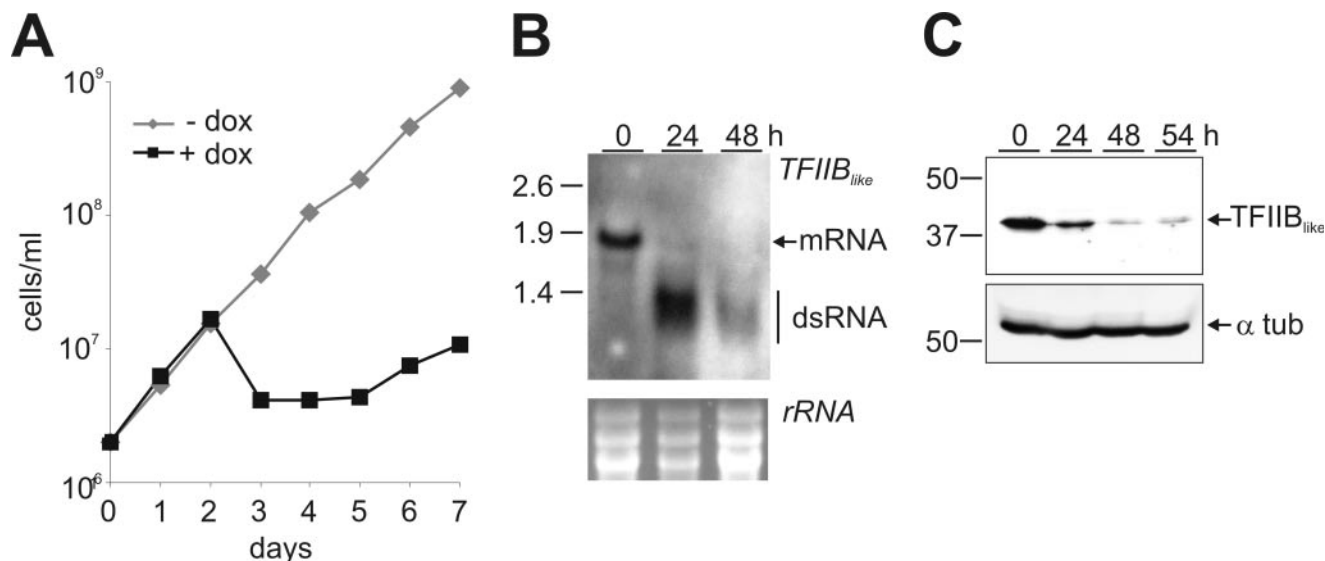
Figure 2. *T.brucei TFIIB*$_{like}$ is an essential gene. (**A**) Growth curve of a clonal procyclic *T.brucei* cell line in which the addition of doxycycline induces the expression of *TFIIB*$_{like}$ dsRNA. Cells were diluted to a density of 2×10^6 cells/ml each day and culture growth was measured in the absence (−dox, grey line) or presence (+dox, black line) of doxycycline at specified times. (**B**) Northern blot of *TFIIB*$_{like}$ mRNA before and 24 and 48 h after induction of *TFIIB*$_{like}$ dsRNA expression. Ethidium bromide-stained *rRNA* bands served as loading controls. RNA marker sizes are indicated on the left in kilobases. (**C**) Corresponding immunoblot analysis with antibodies directed against TFIIB$_{like}$ and α-tubulin (α tub). Protein size markers in kDa are specified on the left.

comparison to the RNA pol III-transcribed U2 snRNA, the amount of SL RNA was dramatically reduced 48 h after the addition of doxycycline (Figure 3A). To analyze whether this reduction was caused by a defect in synthesis, we labeled nascent RNA using a permeabilized cell system (36) in which newly synthesized SL RNA can be directly visualized when labeled, nascent RNA is separated by denaturing PAGE (21,36). As shown in Figure 3B, silencing of *TFIIB*$_{like}$ did not significantly affect labeling of either tRNA or high molecular weight RNA within the 54 h time frame of the experiment, whereas the level of labeled SL RNA was reproducibly decreased 2- to 3-fold after 48 h of induction with doxycycline. The labeled RNA was also used in a dot blot analysis to assess the effect of *TFIIB*$_{like}$ silencing on specific genes (Figure 3C and D). These experiments confirmed that silencing of *TFIIB*$_{like}$ reduced SL RNA synthesis nearly 3-fold 54 h after induction. In contrast, transcription of the U2 and U6 snRNA genes, as well as RNA pol I-mediated synthesis of 18 S rRNA, was not greatly affected within this same time frame. The RNA pol I of *T.brucei* is unique since it transcribes not only the large ribosomal gene unit (*RRNA*), but also the gene units encoding the major cell surface antigens *GPEET* and *EP* procyclin and variant surface glycoprotein (*VSG*) (33). Interestingly, the dot blot analysis suggested that *TFIIB*$_{like}$ silencing affected mRNA synthesis by RNA pol I and II differently. While labeling of GPEET procyclin mRNA was not greatly affected, RNA pol II–mediated synthesis of α-tubulin and *HSP70* mRNA was significantly reduced 54 h after induction, albeit not as strongly as was transcription of *SLRNA*. Taken together, these data strongly suggested that TFIIB$_{like}$ plays an important role in *SLRNA* transcription, and in addition, they raised the possibility that this protein may also function in RNA pol II-mediated polycistronic transcription of protein coding genes.

Interestingly, a recent study on archaean TFB revealed a stimulatory transcription effect of this factor which is recruitment-independent (42). Hence, it is possible that even if in trypanosomatids RNA pol II recruitment to protein coding genes is non-specific, TFIIB$_{like}$ may enhance successful transcription initiation by interacting with the RNA polymerase.

TFIIB$_{like}$ localizes to the nucleus and interacts with RNA pol II and the *SLRNA* promoter

As a transcription factor, TFIIB$_{like}$ should predominantly localize to the nucleus. To test this, we used polyclonal antisera to detect TFIIB$_{like}$ in fixed procyclic cells by indirect immunofluorescence. As expected, the antiserum exclusively stained the nucleus whereas there was no labeling with the pre-immune serum (Figure 4).

For a functional *in vitro* analysis, we generated the procyclic cell line TbH8 which exclusively expressed *TFIIB*$_{like}$ C-terminally fused to the PTP tag (TFIIB$_{like}$-PTP). The PTP tag is an epitope combination for tandem affinity purification and consists of two protein A moieties, a tobacco etch virus protease site and the protein C epitope (30). Exclusive expression of TFIIB$_{like}$-PTP in TbH8 cells was achieved by replacing one allele of the single copy *TFIIB*$_{like}$ gene with the hygromycin phosphotransferase gene and by targeted integration of construct pTFIIB$_{like}$-PTP-NEO into the second allele (Figure 5A). Immunoblot analysis of extracts derived from wild type and TbH8 cells using antibodies directed against the PTP tag or TFIIB$_{like}$ showed that TbH8 cells expressed TFIIB$_{like}$-PTP but not the untagged protein (Figure 5B). Since *TFIIB*$_{like}$ is an essential gene and TbH8 cells did not exhibit a growth defect (data not shown), we inferred that the C-terminal tag did not critically interfere with TFIIB$_{like}$ function. We first used this cell line to purify TFIIB$_{like}$-PTP from crude TbH8 extract (Figure 5C). The final eluate contained several proteins with sizes ranging from 200 to 10 kDa. Most of these proteins were present in approximately stoichiometric
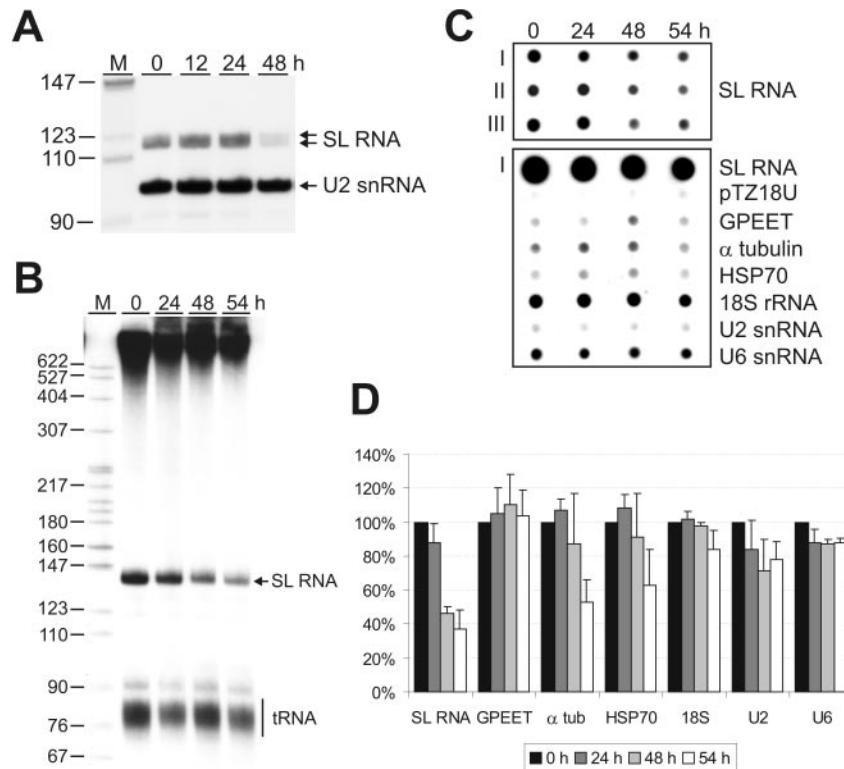
Figure 3. Silencing of $TFIIB_{like}$ expression affects SL RNA synthesis *in vivo*. (**A**) Total RNAs prepared from RNAi-induced cells were analyzed by primer extension assays using $5'$-^{32}P-end labeled oligonucleotides specific for SL RNA and U2 snRNA. The lower of the two extension products for the SL RNA is a result of the hyper-methylated SL RNA cap, which prematurely terminates reverse transcription. DNA marker (M) sizes are indicated on the left. (**B**) Nascent RNAs were labeled with $[\alpha$-^{32}P]UTP in permeabilized cells in which expression of $TFIIB_{like}$ dsRNA was induced for the times specified. The RNAs were separated on a 6% polyacrylamide/50% urea gel and visualized by autoradiography. SL RNA and tRNA are indicated on the right and DNA marker (M) sizes on the left. (**C**) Labeled, nascent RNA was used to probe dot blots containing the complete coding regions of the SL RNA, GPEET procyclin, α-tubulin, heat shock protein 70 (HSP70), 18S ribosomal RNA, U2 snRNA and U6 snRNA. The vector pTZ18U served as a control. Shown are low exposures of experiments I–III for the SL RNA and a long exposure of experiment I for all other RNAs. (**D**) Quantification of the dot blot signal strengths from three independent experiments. The signal of non-induced cells was set to 100%.

amounts and the apparent sizes of the top band doublet and of the band beneath correlated well with the predicted molecular masses of the two largest RNA pol II subunits RPB1 (197 kDa) and RPB2 (135 kDa). Excision of these bands and mass spectrometric identification of the proteins confirmed that both bands of the doublet contained RPB1 and that the lower band consisted of RPB2 (data not shown). Hence, $TFIIB_{like}$ interacts with RNA pol II in a manner which is unaffected by tandem affinity purification.

To assess whether $TFIIB_{like}$ binds to the *SLRNA* promoter, we employed a promoter pull-down assay using the TbH8 cell line and linear promoter DNA fragments which comprised complete promoter regions to allow the formation of a stable pre-initiation complex. The promoter DNAs were labeled with biotin, immobilized on streptavidin beads, and incubated with transcriptionally active cell extract. The proteins which remained bound to the DNAs after several washing steps were analyzed by immunoblotting (Figure 6). In comparison to pull-downs with unrelated control DNA, $TFIIB_{like}$–PTP clearly interacted with the *SLRNA* promoter but not with the class I *GPEET*, *RRNA* and *VSG* expression site promoters (Figure 6, compare lane 2 with lanes 3, 4, 6 and 8). In addition to the proximal *RRNA* promoter (*RRNA* $-84/+126$) which is sufficient for maximal transcription *in vitro* (31), we analyzed the promoter fragment *RRNA* $-257/-162$ which harbors the

most distal domain of the *RRNA* promoter. The relevant sequence elements of this domain resemble an oppositely oriented *SLRNA* USE and bind SNAPc and TRF4 (37,38) (data not shown) but not $TFIIB_{like}$ (Figure 6, lane 5). This result is in accordance with the finding that the region upstream of the *RRNA* promoter is transcriptionally silent (37). In addition to these characterized promoters, we analyzed the intergenic region of *HSP70* genes 2 and 3 (*H23*) which was identified previously as a potential promoter region for RNA pol II-mediated transcription of protein coding genes (43). A pull-down assay performed using *H23* DNA reproducibly yielded a signal slightly above background (Figure 6, lane 7). Given that this DNA fragment was the longest fragment used in these analyses, it is likely that the larger DNA length caused an increased background signal. Collectively, our data suggest that $TFIIB_{like}$ specifically binds to the *SLRNA* promoter, but does not recognize either class I promoters or the *H23* intergenic region.

$TFIIB_{like}$ is essential for *SLRNA* transcription *in vitro*

Finally, we investigated the function of $TFIIB_{like}$ in an *in vitro* system in which *SLRNA* and RNA pol I-mediated transcription are active (31). We used the TbH8 cells because the high affinity interaction of the protein A domains within
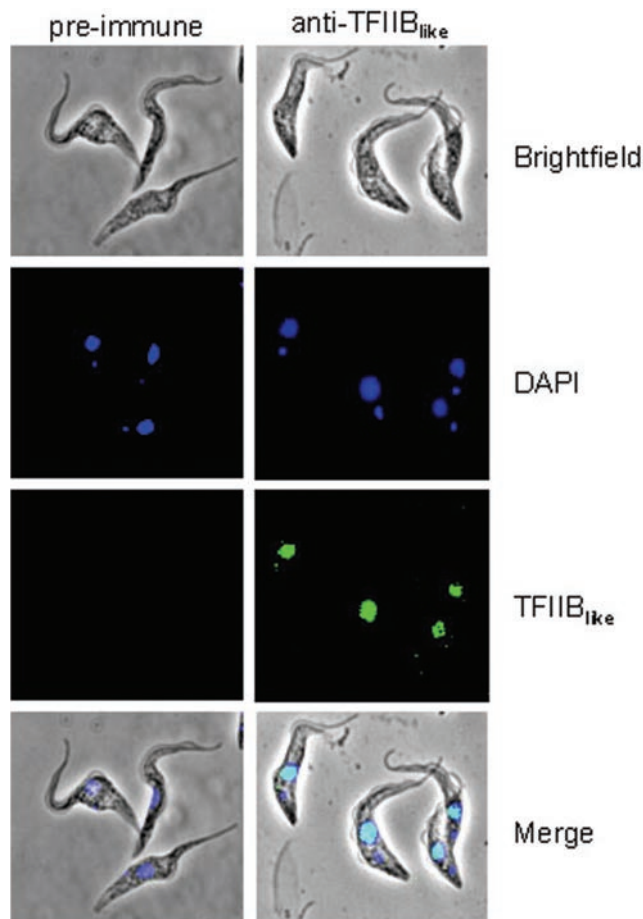
Figure 4. TFIIB$_{like}$ is localized to the nucleus. Wild-type cells were fixed and stained with DAPI. TFIIB$_{like}$ was detected using either a polyclonal antisera or, as a control, with the corresponding pre-immune serum and a FITC-labeled anti-rat IgG secondary antibody.

Figure 5. PTP purification of TFIIB$_{like}$. (**A**) Illustration of the two *TFIIB$_{like}$* alleles in the clonal cell line TbH8. In one allele, the *TFIIB$_{like}$* coding region was replaced by the hygromycin phosphotransferase gene (HYG-R) and in the second allele, the PTP coding region was fused to the 3′ end by targeted insertion of pTFIIB$_{like}$-PTP-NEO. The endogenous *TFIIB$_{like}$* coding region is represented by open boxes, the PTP tag by a closed box, and homologous sequences in the construct by checkered boxes. Resistance marker coding regions are represented by striped boxes and introduced gene flanks for RNA processing signals by smaller gray boxes. (**B**) Immunoblot analysis of TFIIB$_{like}$ expression in wild-type (427) and TbH8 cells. Total cell lysates were analyzed with antibodies against the protein A domains of the PTP tag (α ProtA), TFIIB$_{like}$ (α TFIIB$_{like}$) and α tubulin (α tub). On the right, arrows indicate positions of the proteins and on the left, protein marker sizes are provided in kDa. (**C**) Coomassie blue staining of purified proteins. The total eluate from a TFIIB$_{like}$-PTP tandem affinity purification was separated on a 10–20% SDS–polyacrylamide gradient gel and stained with Coomassie blue. For comparison, 0.002% of the input material (INP) and 5% of the TEV protease eluate (TEV) were loaded. Arrows on the right mark TFIIB$_{like}$-P and the two largest RNA pol II subunits RPB1 and RPB2 which were identified by mass spectrometry. On the left, sizes of marker proteins in kDa are indicated.

the PTP tag with immobilized IgG allowed efficient and rapid depletion of TFIIB$_{like}$-PTP from the extract (Figure 7A). Reactions were carried out in which the *SLRNA* promoter template SLins19 and, as a control, the class I *GPEET* promoter template GPEET-trm were co-transcribed. Both of these constructs harbor unrelated sequence tags downstream of the transcription initiation site which allow for specific detection of the corresponding RNAs by primer extension analysis. As shown in Figure 7B, depletion of TFIIB$_{like}$ completely abolished *SLRNA* transcription without affecting *GPEET* promoter transcription (lane 3), while a mock depletion had only a slight effect on the transcription on both templates (compare lanes 1 and 2). This result is in close agreement with our earlier finding that TFIIB$_{like}$ did not bind to the *GPEET* promoter in the pull-down assay. To confirm that TFIIB$_{like}$ was responsible for the effect on *SLRNA* transcription, we attempted to reconstitute the transcriptional activity in the depleted extract by adding back purified protein. The final eluate of purified TFIIB$_{like}$-PTP (see above, Figure 5C) reconstituted *SLRNA* transcription in a dose-dependent manner (Figure 7B, lanes 6–8). A control eluate from the PTP-tagged and purified largest subunit of RNA pol I did not affect *SLRNA* transcription (lane 4), demonstrating
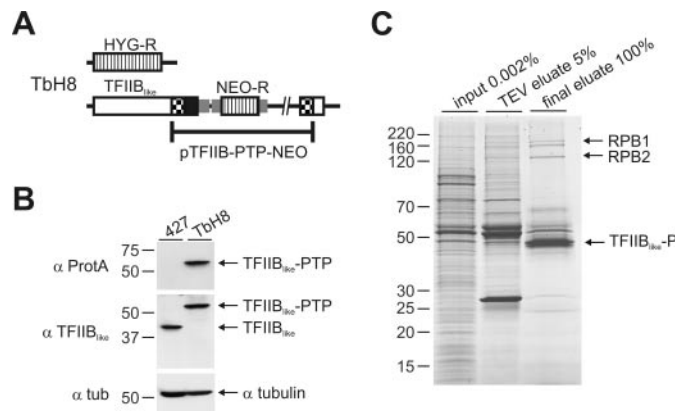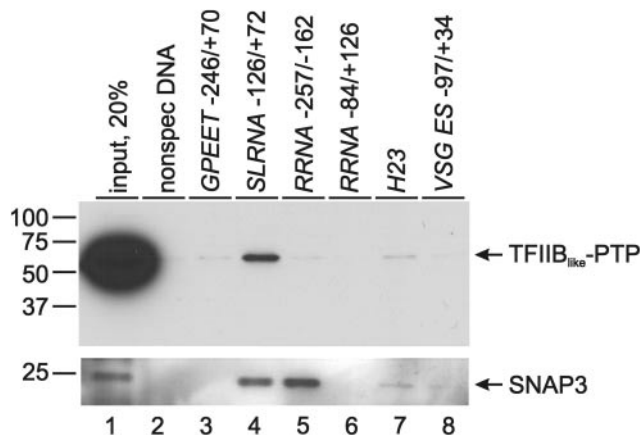

Figure 6. TFIIB$_{like}$ specifically binds to the SLRNA promoter. Promoter pull-down assays were carried out with TbH8 extract and immobilized DNAs of *GPEET, SLRNA, RRNA* and *VSG* expression site (*VSG ES*) promoters. The numbers indicate the end positions of the DNA fragments relative to the transcription initiation sites. In addition to these promoter DNAs, a 222 bp non-specific DNA and the complete intergenic region of *HSP70* genes 2 and 3 (*H23*) were analyzed. Proteins which bound to the DNAs were analyzed by immunoblotting using the PTP-specific PAP reagent and a polyclonal antibody against the SNAPc subunit SNAP3.

that the TFIIB$_{like}$ portion of the fusion protein was responsible for the restoration of activity. Similarly, the SNAP2-PTP purified TRF4/SNAPc/TFIIA complex which has been shown to be functional, did not reconstitute *SLRNA* transcription
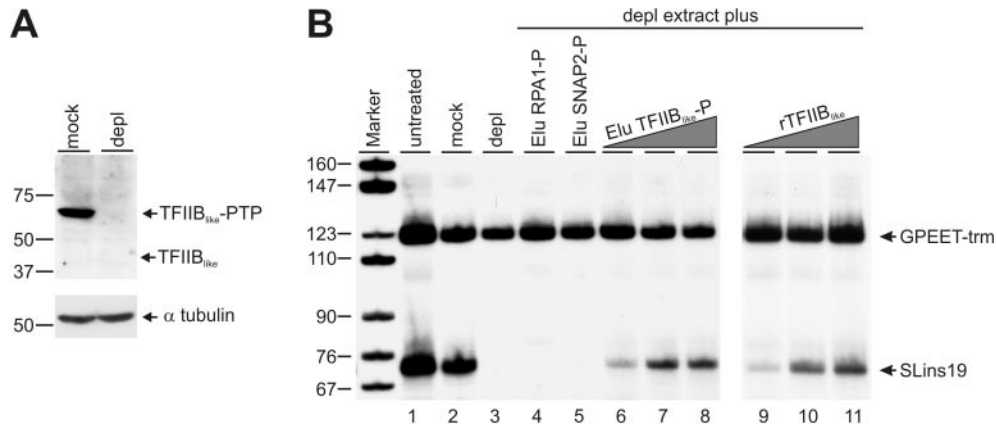
Figure 7. *SLRNA* transcription *in vitro* essentially depends on TFIIB$_{like}$. (**A**) Transcriptionally active extract of TbH8 cells was depleted of TFIIB$_{like}$-PTP by IgG affinity chromatography. Mock-treated and depleted (depl) extracts were analyzed by immunoblotting using antibodies directed against TFIIB$_{like}$ and α-tubulin. As indicated by arrows on the right, depleted extract did not contain detectable amounts of tagged or wild-type TFIIB$_{like}$. (**B**) Templates GPEET-trm and SLins19 were co-transcribed in untreated, mock-treated or depleted transcription extract. Depleted extract was reconstituted with eluates (Elu) from PTP purifications of the largest RNA pol I subunit RPA1 (RPA1-P), the SNAPc subunit SNAP2 (SNAP2-P), increasing amounts of TFIIB$_{like}$ (TFIIB$_{like}$-P), and comparable amounts of recombinant (r)TFIIB$_{like}$. GPEET-trm and SLins19 transcripts were detected by primer extension of total RNA prepared from transcription reactions with (5′ end)-^{32}P-labeled oligonucleotides TAG_PE and SLtag, respectively. Reaction products were separated on 6% polyacrylamide/50% urea gels and visualized by autoradiography. On the left, sizes of pBR322-MspI marker fragments are indicated, and on the right, GPEET-trm and SLins19 primer extension products are identified.

in the TFIIB-depleted extract (lane 5). This result also indicates that the SNAP2-P eluate did not contain TFIIB$_{like}$ and that a putative interaction of TRF4 and TFIIB$_{like}$ is not resistant to the conditions of tandem affinity purification. While these experiments strongly suggested that TFIIB$_{like}$ is essential for *SLRNA* transcription, it was still possible that RNA pol II or other proteins that co-purified with PTP-tagged TFIIB$_{like}$ were responsible for the observed transcriptional effects. To unequivocally demonstrate that these effects can be attributed to TFIIB$_{like}$ alone, we expressed and purified the protein in *E.coli*. When added back to TFIIB$_{like}$-depleted extract, recombinant TFIIB$_{like}$ was as efficient in reconstituting *SLRNA* transcription as was the eluate of the tandem affinity purification (Figure 7B, lanes 9–11). Hence, we concluded that TFIIB$_{like}$ is indispensable for *SLRNA* transcription. Since this protein was able to reconstitute *SLRNA* transcription in an extract with no detectable *SLRNA* transcription activity, we infer that TFIIB$_{like}$ is a basal transcription factor.

Although it will require a more detailed characterization to determine whether TFIIB$_{like}$ is the ortholog of TFIIB/TFB, our results indicate that this is indeed the case. Why then has TFIIB$_{like}$ diverged much more than trypanosomatid BRF1? Both factors are functionally similar and bridge promoter bound factors with RNA pol II and III, respectively. A likely explanation is that while trypanosomatids harbor a standard set of small RNA genes which are transcribed by RNA pol III and most likely require BRF1 for transcription initiation, a class II pre-initiation complex is formed at relatively few sites in the genome thus reducing the constraints on TFIIB$_{like}$ conservation. It is possible that TFIIB$_{like}$ interacts only with the *SLRNA* promoter but even if this factor is involved in recruiting RNA pol II to protein coding genes, initiation sites may be limited due to the polycistronic mode of transcription. Accordingly, strand-specific nuclear run-on assays in *L.major* suggested that protein coding gene transcription initiates predominantly at strand-switch regions between divergent gene clusters (17).

CONCLUSION

We have identified a TFIIB$_{like}$ protein in trypanosomatid parasites and our functional analysis suggests that this protein is the orthologue of eukaryotic TFIIB and archaeal TFB. If this is the case, further analysis of TFIIB$_{like}$ function may resolve the long standing question on how RNA pol II is recruited to protein coding genes in trypanosomes. Most importantly, however, is our finding that TFIIB$_{like}$ is absolutely required for SL RNA synthesis in *T.brucei*. Since this process is vital to all life cycle stages of trypanosomatid parasites and since the protein is so divergent to other members of the TFIIB family of proteins, a detailed functional and structural characterization of TFIIB$_{like}$ and its interactions with other proteins may reveal a parasite-specific target in a fundamentally important process.

REFERENCES

1. Lemon,B. and Tjian,R. (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.*, **14**, 2551–2569.

2. Hahn,S. (2004) Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Struct. Mol. Biol.*, **11**, 394–403.

3. Boeger,H., Bushnell,D.A., Davis,R., Griesenbeck,J., Lorch,Y., Strattan,J.S., Westover,K.D. and Kornberg,R.D. (2005) Structural basis of eukaryotic gene transcription. *FEBS Lett.*, **579**, 899–903.

4. Chen,H.T., Legault,P., Glushka,J., Omichinski,J.G. and Scott,R.A. (2000) Structure of a (Cys3His) zinc ribbon, a ubiquitous motif in archaeal and eucaryal transcription. *Protein Sci.*, **9**, 1743–1752.

5. Hahn,S. and Roberts,S. (2000) The zinc ribbon domains of the general transcription factors TFIIB and Brf: conserved functional surfaces but different roles in transcription initiation. *Genes Dev.*, **14**, 719–730.

6. Bushnell,D.A., Westover,K.D., Davis,R.E. and Kornberg,R.D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science*, **303**, 983–988.

7. Tsai,F.T. and Sigler,P.B. (2000) Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.*, **19**, 25–36.

8. Lagrange,T., Kapanidis,A.N., Tang,H., Reinberg,D. and Ebright,R.H. (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.*, **12**, 34–44.

9. Nikolov,D.B., Chen,H., Halay,E.D., Usheva,A.A., Hisatake,K., Lee,D.K., Roeder,R.G. and Burley,S.K. (1995) Crystal structure of a TFIIB–TBP–TATA-element ternary complex. *Nature*, **377**, 119–128.

10. Geiduschek,E.P. and Ouhammouch,M. (2005) Archaeal transcription and its regulators. *Mol. Microbiol.*, **56**, 1397–1407.

11. Bartlett,M.S. (2005) Determinants of transcription initiation by archaeal RNA polymerase. *Curr. Opin. Microbiol.*, **8**, 677–684.

12. Ruan,J.P., Arhin,G.K., Ullu,E. and Tschudi,C. (2004) Functional characterization of a *Trypanosoma brucei* TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol. Cell. Biol.*, **24**, 9610–9618.

13. Berriman,M., Ghedin,E., Hertz-Fowler,C., Blandin,G., Renauld,H., Bartholomeu,D.C., Lennard,N.J., Caler,E., Hamlin,N.E., Haas,B. *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei. Science*, **309**, 416–422.

14. El Sayed,N.M., Myler,P.J., Bartholomeu,D.C., Nilsson,D., Aggarwal,G., Tran,A.N., Ghedin,E., Worthey,E.A., Delcher,A.L., Blandin,G. *et al.* (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*, **309**, 409–415.

15. Ivens,A.C., Peacock,C.S., Worthey,E.A., Murphy,L., Aggarwal,G., Berriman,M., Sisk,E., Rajandream,M.A., Adlem,E., Aert,R. *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major. Science*, **309**, 436–442.

16. Tschudi,C. and Ullu,E. (1988) Polygene transcripts are precursors to calmodulin mRNAs in trypanosomes. *EMBO J.*, **7**, 455–463.

17. Martinez-Calvillo,S., Yan,S., Nguyen,D., Fox,M., Stuart,K. and Myler,P.J. (2003) Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol. Cell*, **11**, 1291–1299.

18. Campbell,D.A., Sturm,N.R. and Yu,M.C. (2000) Transcription of the kinetoplastid spliced leader RNA gene. *Parasitol. Today*, **16**, 78–82.

19. Gilinger,G. and Bellofatto,V. (2001) Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. *Nucleic Acids Res.*, **29**, 1556–1564.

20. Hartree,D. and Bellofatto,V. (1995) Essential components of the mini-exon gene promoter in the trypanosomatid *Leptomonas seymouri. Mol. Biochem. Parasitol.*, **71**, 27–39.

21. Günzl,A., Ullu,E., Dörner,M., Fragoso,S.P., Hoffmann,K.F., Milner,J.D., Morita,Y., Nguu,E.K., Vanacova,S., Wünsch,S. *et al.* (1997) Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. *Mol. Biochem. Parasitol.*, **85**, 67–76.

22. Yu,M.C., Sturm,N.R., Saito,R.M., Roberts,T.G. and Campbell,D.A. (1998) Single nucleotide resolution of promoter activity and protein binding for the *Leishmania tarentolae* spliced leader RNA gene. *Mol. Biochem. Parasitol.*, **94**, 265–281.

23. Luo,H., Gilinger,G., Mukherjee,D. and Bellofatto,V. (1999) Transcription initiation at the TATA-less spliced leader RNA gene promoter requires at least two DNA-binding proteins and a tripartite architecture that includes an initiator element. *J. Biol. Chem.*, **274**, 31947–31954.

24. Schimanski,B., Nguyen,T.N. and Günzl,A. (2005) Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei. Mol. Cell. Biol.*, **25**, 7303–7313.

25. Das,A., Zhang,Q., Palenchar,J.B., Chatterjee,B., Cross,G.A. and Bellofatto,V. (2005) Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. *Mol. Cell. Biol.*, **25**, 7314–7322.

26. Hernandez,N. (2001) Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J. Biol. Chem.*, **276**, 26733–26736.

27. Kuhlman,T.C., Cho,H., Reinberg,D. and Hernandez,N. (1999) The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. *Mol. Cell. Biol.*, **19**, 2130–2141.

28. Shi,H., Djikeng,A., Mark,T., Wirtz,E., Tschudi,C. and Ullu,E. (2000) Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA.*, **6**, 1069–1076.

29. Bastin,P., Ellis,K., Kohl,L. and Gull,K. (2000) Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. *J. Cell Sci.*, **113**, 3321–3328.

30. Schimanski,B., Nguyen,T.N. and Günzl,A. (2005) Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot. Cell*, **4**, 1942–1950.

31. Laufer,G., Schaaf,G., Bollgönn,S. and Günzl,A. (1999) *In vitro* analysis of alpha-amanitin-resistant transcription from the rRNA, procyclic acidic repetitive protein, and variant surface glycoprotein gene promoters in *Trypanosoma brucei. Mol. Cell. Biol.*, **19**, 5466–5473.

32. Günzl,A., Bindereif,A., Ullu,E. and Tschudi,C. (2000) Determinants for cap trimethylation of the U2 small nuclear RNA are not conserved between *Trypanosoma brucei* and higher eukaryotic organisms. *Nucleic Acids Res.*, **28**, 3702–3709.

33. Günzl,A., Bruderer,T., Laufer,G., Schimanski,B., Tu,L.C., Chung,H.M., Lee,P.T. and Lee,M.G. (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei. Eukaryot. Cell*, **2**, 542–551.

34. Wirtz,E., Leal,S., Ochatt,C. and Cross,G.A.M. (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei. Mol. Biochem. Parasitol.*, **99**, 89–101.

35. Günzl,A., Cross,M. and Bindereif,A. (1992) Domain structure of U2 and U4/U6 small nuclear ribonucleoprotein particles from *Trypanosoma brucei*: identification of *trans*-spliceosomal specific RNA–protein interactions. *Mol. Cell. Biol.*, **12**, 468–479.

36. Ullu,E. and Tschudi,C. (1990) Permeable trypanosome cells as a model system for transcription and trans-splicing. *Nucleic Acids Res.*, **18**, 3319–3326.

37. Schimanski,B., Laufer,G., Gontcharova,L. and Günzl,A. (2004) The *Trypanosoma brucei* spliced leader RNA and rRNA gene promoters have interchangeable TbSNAP50-binding elements. *Nucleic Acids Res.*, **32**, 700–709.

38. Laufer,G. and Günzl,A. (2001) *In-vitro* competition analysis of procyclin gene and variant surface glycoprotein gene expression site transcription in *Trypanosoma brucei. Mol. Biochem. Parasitol.*, **113**, 55–65.

39. Gough,J., Karplus,K., Hughey,R. and Chothia,C. (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.*, **313**, 903–919.

40. Pinto,I., Wu,W.H., Na,J.G. and Hampsey,M. (1994) Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. *J. Biol. Chem.*, **269**, 30569–30573.

41. Chen,Y., Hung,C.H., Burderer,T. and Lee,G.S. (2003) Development of RNA interference revertants in *Trypanosoma brucei* cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. *Mol. Biochem. Parasitol.*, **126**, 275–279.

42. Werner,F. and Weinzierl,R.O. (2005) Direct modulation of RNA polymerase core functions by basal transcription factors. *Mol. Cell. Biol.*, **25**, 8344–8355.

43. Lee,M.G. (1996) An RNA polymerase II promoter in the hsp70 locus of *Trypanosoma brucei. Mol. Cell. Biol.*, **16**, 1220–1230.