ELSEVIER

Molecular and Biochemical Parasitology 98 (1999) 117–129

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

Trypanosoma brucei poly(A) binding protein I cDNA cloning, expression, and binding to 5′ untranslated region sequence elements☆

Tracy L. Hotchkiss, George E. Nerantzakis, Sarah C. Dills, Limin Shang, Laurie K. Read *

Department of Microbiology and Center for Microbial Pathogenesis, SUNY Buffalo School of Medicine, 138 Farber Hall, Buffalo NY 14214, USA

Received 24 July 1998; received in revised form 12 October 1998; accepted 12 October 1998

Abstract

Poly(A) binding protein I (PABPI) is a highly conserved eukaryotic protein that binds mRNA poly(A) tails and functions in the regulation of translational efficiency and mRNA stability. As a first step in our investigation of the role(s) of mRNA poly(A) tails in posttranscriptional gene regulation in *Trypanosoma brucei*, we have cloned the cDNA encoding PABPI from this organism. The cDNA predicts a protein homologous to PABPI from other organisms and displaying conserved features of these proteins, including four RNA binding domains that span the N-terminal two-thirds of the protein. Comparison of northern blot data with the cDNA sequence indicates an unusually long 3′ untranslated region (UTR) of approximately three kilobases. The 5′ UTR contains both A-rich and AU repeat regions, the former being a ubiquitous property of PABPI 5' UTRs. *T. brucei* PABPI, expressed as a glutathione-*S*-transferase fusion protein, bound to RNA comprised of its full length 5′ UTR in UV cross-linking experiments. This suggests that PABPI may play an autoregulatory role in its own expression. Competition experiments indicate that the A-rich region, but not the AU repeats, are involved in this binding.

Keywords: *Trypanosoma*; Polyadenylation; RNA binding; Translation; mRNA stability

Abbreviations: GST, Glutathione-S-transferase; IPTG, isopropyl-1-thio-B-D-galactopyranoside; kb, kilobases; nt, nucleotides; PABPI, poly(A) binding protein I; RACE, rapid amplification of cDNA ends; UTR, untranslated region; RBDs, RNA binding domains; VSGs, variant surface glycoproteins.

* Corresponding author. Tel.: +1-716-829-3307; fax: +1-716-829-2158; e-mail: lread@acsu.buffalo.edu.

☆ Nucleotide sequence data reported in this paper is available in the GenBank database under the accession number AF042190.

0166-6851/99/$ - see front matter
PII: S0166-6851(98)00156-X

1. Introduction

In contrast to most eukaryotic organisms, gene expression in trypanosomes is primarily controlled by posttranscriptional events (for review see Ref. [1]). Trypanosome genes are often unusually tightly spaced, both in the nucleus and mitochondrion, and few promoters have been identified. Consistent with this observation, both nuclear and mitochondrial genomes are transcribed into long, polycistronic RNAs [2–7]. Nevertheless, expression of genes that are polycistronically transcribed is often differentially regulated [2–6]. Levels at which gene expression has been shown or suggested to be regulated in trypanosomes include RNA processing (*trans*-splicing and polyadenylation in the nucleus; editing, cleavage and polyadenylation in the mitochondrion), RNA stability, nuclear export, translational efficiency, and protein degradation. For example, mRNA stability plays a role in the stage-specific expression of both procyclin [8] and variant surface glycoproteins (VSGs) [9]. In transfection experiments, procyclin and VSG 3′ untranslated region (UTR) elements that positively or negatively regulate reporter RNA half-life have been identified [10–12]. Translational efficiency of reporter constructs also is regulated by at least two different procyclin 3′ UTR elements [12,13]. Moreover, translational regulation has been implicated for several trypanosome genes based on discrepancies between mRNA and protein ratios in bloodstream and procyclic forms. These include the developmentally regulated protein kinase, NRK [14], and the 34/37 kDa nucleic acid binding proteins [15]. While sequence elements that confer increased or decreased RNA stability, translational efficiency, or RNA processing to reporter constructs have been identified, the precise mechanisms that govern these effects are not currently understood.

One protein that has been demonstrated to play a critical role in posttranscriptional regulation of eukaryotic gene expression is the multi-functional poly(A) binding protein I (PABPI; for reviews see Refs [16,17]). PABPI binds mRNA poly(A) tails, and is the most abundant RNA binding protein in messenger ribonuclear protein particles. It is ubiquitous among those eukaryotes which have been examined, and is essential for yeast viability [18]. PABPI structure is highly conserved, consisting of four RBD-type RNA binding domains [19] that comprise the N-terminal two-thirds of the protein, and a less well conserved C terminus [18]. One generally accepted function of PABPI is the mediation of poly(A)-dependent events in translation initiation. Genetic and biochemical evidence suggests a role for poly(A)-PABPI in both 60S ribosomal subunit joining [20,21] and 40S ribosomal subunit addition [22,23]. Poly(A)-associated PABPI has also been suggested to mediate the poly(A) effect on translation via its recently discovered interaction with the eIF-4G subunit of the cap binding protein complex, eIF-4F, which recruits the 40S ribosomal subunit to mRNA [24]. A second widespread function of PABPI is in the regulation of mRNA degradation, although its role in mammalian cells differs from that in yeast (for review see Ref. [25]). It is thought that destabilizing sequences present in mRNA 3′ UTRs stimulate deadenylation, which is a prerequisite for mRNA degradation [26]. In cell-free extracts of mammalian cells, PABPI protects transcripts from degradation possibly by blocking nuclease digestion of the poly(A) tail [27]. In contrast, in yeast PABPI promotes mRNA degradation through activation of a PABPI-dependent poly(A) nuclease [28,29]. Most recently, PABPI has been shown to associate with the yeast cleavage and polyadenylation machinery, where it functions in regulation of poly(A) tail length during the polyadenylation reaction [30,31]. Finally, filter binding [32] and UV cross-linking [33] experiments have revealed binding of PABPI to its own A-rich 5′ UTR, suggesting the protein plays an autoregulatory function in its own expression. In support of this hypothesis, it has been shown that recombinant PABPI represses translation of PABPI mRNA in reticulocyte lysate [33].

In kinetoplastid parasites, mRNAs encoded in both the nucleus and mitochondrion are polyadenylated [34–36]. The machinery which catalyzes polyadenylation in kinetoplastids has not been identified. The function of polyadenylation in kinetoplastids is also unknown, and poly(A) tail length has not been reported for most

nuclearly encoded mRNAs. However, it was noted that an increase in the size and heterogeneity of procyclin transcripts followed induction of *Trypanosoma brucei* to procyclic forms [37]. This was suggested to be due to an increase in poly(A) tail length, which might result in increased translation of procyclin mRNA [37]. In kinetoplastid mitochondria, many RNAs exhibit extreme size heterogeneity that has been demonstrated to be a result of differential poly(A) tail lengths [3,36,38,39]. Intriguingly, in many cases, the edited status of an RNA is correlated with the length of its poly(A) tail [3,36,40] (K. Militello and L. Read, unpublished data).

As a first step in our investigation of poly(A) tail function in *T. brucei*, we have cloned the cDNA encoding PABPI from this organism. The predicted *T. brucei* PABPI protein is conserved in structure compared with other PABPI proteins. We find that the PABPI 5′ UTR contains both A-rich and AU repeat elements. Using wild type and mutant 5′ UTRs in UV cross-linking competition experiments, we demonstrate binding of PABPI protein to its own 5′ UTR and identify an element involved in this binding.

2. Materials and methods

2.1. Cells and nucleic acids

Procyclic form *T. brucei brucei* strain EATRO 164 was grown as previously described [41]. Total cellular RNA was isolated using the guanidinium isothiocyanate–phenol–chloroform method of Chomczynski and Sacchi [42]. Genomic DNA was isolated as described by Carrington [43].

2.2. Oligonucleotides

The oligonucleotides used in this study are shown below with restriction sites incorporated at their 5′ ends underlined:

ESL-22	GCGAATTCGCTATTATTAG AACAGTTTCTG
T7	GTAATACGACTCACTATAG GGC
T7-SLS	GTAATACGACTCACTATAG GGGCTATTATTAGAACAGT TTCTG
XSC-$(dT)_{17}$	GACTCGAGTCGACATACGA TTTTTTTTTTTTTTTT
PABPI-1	GCGAATTCGG(T/C) TA(T/C)GG(T/C)TA(T/C) GT(G/A/T/C)AA(T/C)TT
PABPI-2	GCGGATCCTTCTT(C/A) AC(G/A)TA(G/C/A) AG(G/A)TT(G/A/T/C)GT
PABPI-3	GCGGATCCTCAAGCGCCTT CTCAGCATC
PABPI-6	CCATTACCGTGGAACGCTA C
PABPI-7	GCGGATCCCAT(G/A/T/C) CC(G/A/T/C)GT(A/C) ACCTT(A/T/C/G)GC
PABPI-8	GCGAATTCTACATTGTAC GAAACAATCCC
PABPI-14	GTATACAAGTACCTTGTTC TC
PABPI-15	AGTTTAT-TAAAAATATAAACCAGG
PABPI-17	GTATCCAGTGCCACTTAG
PABPI-18	AGAGGCTGGACAAAGAC
PABPI-19	ACGAAACAATCCCTCCTTC G
PABPI-d2	GATTTCTCTGGTTGTTCC
PABPI-5′ exp1	GCGGATCCATGGCTGCATT TGCTGCTGCG
PABPI-3′ exp2	ACGCGTCGACTACATGCCA ATGTGACGGTTG

2.3. Cloning of PABPI cDNA

Fig. 1 shows the positions and orientations of oligonucleotides used in PABPI cDNA cloning. Initially, an internal fragment of the *T. brucei* PABPI cDNA was amplified by nested PCR. The first reaction used XSC-$(dT)_{17}$-primed procyclic form cDNA as a template and oligonucleotides ESL-22 and PABPI-2. Five μl of this reaction was then used as a template for amplification with PABPI-1 and PABPI-2, and the product was di-

gested and ligated into the *Eco*RI/*Bam*HI site of pBluescript II SK − . Based on the sequence of this fragment, oligonucleotide PABPI-3 was designed and used in conjunction with ESL-22 to amplify the 5′ portion of the cDNA using PABPI-2-primed cDNA as template, and the product was cloned as described above. A large portion of the 3′ end of the cDNA was amplified from XSC-$(dT)_{17}$-primed cDNA using PABPI-6 and PABPI-7 oligonucleotides, and the product filled and phosphorylated as described [44], and ligated into the *Eco*RV site of pBluescript II SK − . The extreme 3′ end of the cDNA was obtained by nested inverse PCR using a mass excised λZAPII procyclic form cDNA library as template [45]. The first PCR was done using oligonucleotides PABPI-3 and PABPI-8. Five μl of this reaction was then used as a template for amplification with PABPI-19 and T7 oligonucleotides, and the product of this reaction was filled and phosphorylated [44] and cloned into the *Eco*RV site of pBluescript II SK − .

For each PCR reaction, two clones were sequenced in both directions. DNA sequencing was performed by Taq cycle sequencing using an ABI automated sequencer at the SUNY Buffalo CAMBI nucleic acids facility. Sequences were analyzed using the GCG software package [46] and sequence comparisons were done with CLUSTAL W [47].

2.4. Gel electrophoresis and hybridizations

For the northern hybridization, 10 μg of total procyclic form RNA was electrophoresed on a 1.5% formaldehyde-agarose gel and transferred to Nytran by capillary action. A radiolabeled RNA probe was synthesized using T3 RNA polymerase from an *Eco*RI-digested clone which was generated using oligonucleotides PABPI-1 and PABPI-2 (see above). The resulting riboprobe is complementary to 418 nucleotides of coding sequence near the 5′ end of the open reading frame. Filters were prehybridized, hybridized and washed as previously described [48] except that the prehybridization and hybridization temperatures were 65°C.

For Southern hybridization, 3.5 μg of genomic DNA was digested with *Xho*I, *Eco*RI, *Eco*RV, or *Dra*I, electrophoresed on a 1.5% agarose gel, denatured, and transferred to Nytran by capillary action. A radiolabeled RNA probe was generated from the same template used for the northern blot except that the template was digested with *Bam*HI and transcribed with T7 polymerase using the Ambion Strip-EZ™ T7 kit. Prehybridization and hybridization conditions were identical to those used for the northern blot.

2.5. Bacterial expression of GST-PABPI

The full length PABPI open reading frame was amplified by PCR from procyclic form *T. brucei* XSC-$(dT)_{17}$-primed cDNA using oligonucleotides PABPI-5′ exp1 and PABPI-3′ exp2. The PCR product was digested and ligated into the *Bam*HI/*Sal*I site of pGEX-4T1 in frame with the glutathione-*S*-transferase (GST) gene, and the ligation products transformed into *Escherichia coli* DH5α competent cells. Expression was induced and fusion protein was purified as described [49], except that 0.3 mM isopropyl-1-thio-B-D-galactopyranoside (IPTG) was used. GST protein, used as a negative control in UV cross-linking experiments, was generously provided by Arvind Thakur and Phil LoVerde.

2.6. Creation of substrates for UV cross-linking

Clone TbPABPI-5′-2, derived from the amplification described above using ESL-22 and PABPI-3 and containing the entire 5′ UTR and a small region of coding sequence, was used as a template for creation of substrates for UV cross-linking assays. Template for transcription of full length 5′ UTR RNA was generated by PCR amplification of clone TbPABPI-5′-2 with oligonucleotides T7-SLS and PABPI-d2. Template for transcription of dA RNA, which lacks 77 nucleotides of A-rich sequence near the 5′ end of the 5′ UTR, was generated by inverse PCR of clone TbPABPI-5′-2 with oligonucleotides PABPI-17 and PABPI-18. The resulting PCR product was filled and phosphorylated [44] and circularized with T4 DNA ligase. This plasmid was then PCR amplified with

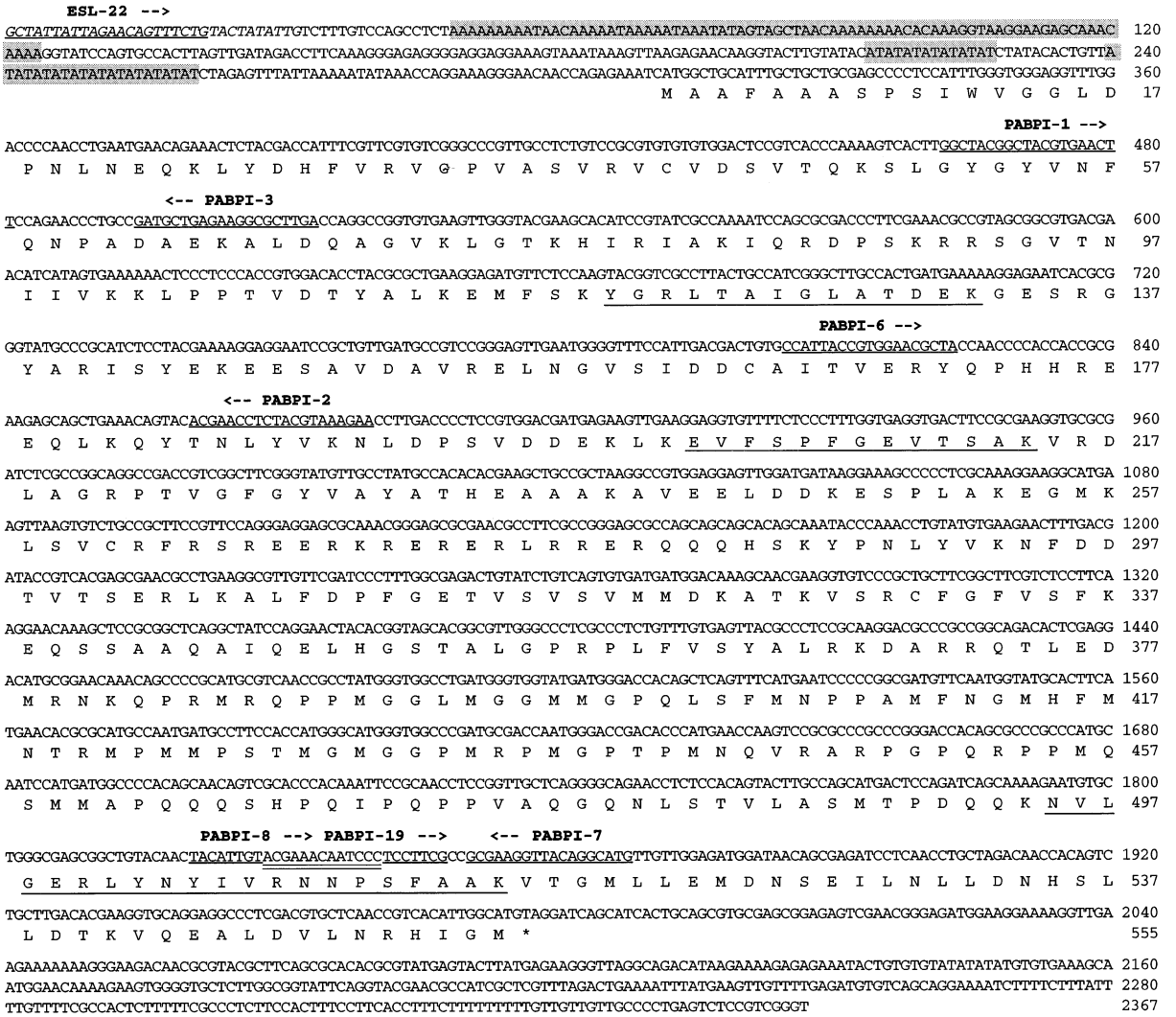

Fig. 1. cDNA and deduced amino acid sequences of the *T. brucei* PABPI. Positions and orientations of oligonucleotides used to amplify portions of the PABPI cDNA are indicated by underlining and arrows. Double underlining denotes a region of overlap between the PABPI-8 and PABPI-19 oligonucleotides. The portion of the spliced leader sequence present in cDNA clones is italicized. A 76 nt A-rich stretch near the 5′ end of the 5′ UTR and two AT repeat regions near the 3′ end of the 5′ UTR are shaded. Deduced amino acid sequences in agreement with the three tryptic peptide sequences of a previously reported *T. brucei* poly(A) binding protein ([58]; see text) are underlined.

oligonucleotides T7-SLS and PABPI-d2 to provide the transcription template. Template for transcription of dAU RNA, which lacks 51 nucleotides of primarily AU repeats, was generated by an identical strategy as the dA template, except that the inverse PCR was done with oligonucleotides PABPI-14 and PABPI-15. RNAs were transcribed with T7 polymerase, and radiolabeled RNAs were internally labeled by incorporation of [α^{32}P]UTP.

2.7. UV cross-linking assays

UV cross-linking assays were carried out in a volume of 15 μl containing 750 fmols of GST-PABPI, 7.5 fmols radiolabeled wild type 5′ UTR RNA, 10 mM magnesium acetate, 80 mM potassium chloride, and 100 mM Tris (pH 8.5). Competition reactions contained 10-, 50-, 100-, or 500-fold molar excess unlabeled wild type, dA, or dAU 5′ UTR RNA. UV cross-linking was carried out as previously described [50].

3. Results and discussion

Degenerate primers (PABPI-1 and -2) corresponding to conserved PABPI amino acid sequences were designed with a bias toward the *T. cruzi* sequence [51] and used in a nested PCR amplification of *T. brucei* procyclic form cDNA as outlined in Materials and methods (Fig. 1). A primer based on the resulting sequence (PABPI-6) was then used in conjunction with a third degenerate primer (PABPI-7) to amplify a large portion of the open reading frame. Finally, additional primers based on the sequences of the previously generated PCR products were used to amplify the extreme 5′ and 3′ regions of the cDNA. The 5′ region was amplified from procyclic form cDNA by RACE with oligonucleotide PABPI-3 and an oligonucleotide containing spliced leader RNA sequence, and the 3′ region was amplified from an in vivo-excised λZAP II cDNA library by nested inverse PCR [45] using oligonucleotides PABPI-3 and PABPI-8 in the first round and PABPI-19 and T7 in the second round.

The full length cDNA (Fig. 1) encodes an open reading frame of 1665 nt. The 5′ UTR, exclusive of the spliced leader sequence, differs between the two clones sequenced, measuring either 275 or 279 nt. All PABPI mRNAs described to date contain 50-70 nt A-rich regions in their 5′ UTRs [32,51–55]. Likewise, the *T. brucei* PABPI mRNA contains a 76 nt region consisting of 70% A residues near its 5′ end (shaded in Fig. 1). An AT repeat region comprised of two stretches of AT repeats (shaded in Fig. 1), one of 14 nts and one of 18 or 22 nts (depending on the clone), is present near the 3′ end of the 5′ UTR. The 5′ UTR of *T. brucei* PABPI is similar to that of *T. cruzi* PABPI cDNA [51] in that both contain a A-rich region; however, AT repeats were not identified in *T. cruzi*. Since the entire 5′ UTR of the *T. cruzi* sequence was not obtained, AT repeats may be present upstream of the region sequenced. Nevertheless, the *T. brucei* and *T. cruzi* PABPI 5′ UTRs differ in organization since in *T. brucei* the AT repeats are downstream of the major A-rich region.

The 3′ UTRs of both *T. brucei* clones measure 388 nt. However, northern blot analysis of *T. brucei* procyclic form RNA reveals one PABPI mRNA species of ≈ 5 kb (data not shown). Since our 5′ sequence includes the spliced leader, together these data indicate that the 3′ UTR is over 3 kb, significantly longer than the length deduced from the cDNA sequence. Thus, the cDNA was likely primed at an internal A-rich stretch in the 3′ UTR. Similar results were obtained in *T. cruzi*; however, in this case it could not be confirmed that the discrepancy in length was due to the 3′ UTR since the complete 5′ UTR was not obtained [51]. Unusually long 3′ UTRs of ≈ 1.5 kb have also been reported for human and *Drosophila* PABPI mRNAs [53]. The significance of such long PABPI 3′ UTRs is unknown, but they are presumably involved in posttranscriptional regulation of gene expression. Their presence in an mRNA encoding an RNA binding protein suggests autoregulation via 3′ UTR binding, but this remains to be experimentally tested.

The deduced amino acid sequence of *T. brucei* PABPI is 555 amino acids in length (Figs. 1 and 2) with a calculated molecular mass of 62.2 kDa and isoelectric point of 10.1. This size agrees with that of PABPI proteins from other organisms which have been reported to be 62–73 kDa [32,51–55]. The predicted protein is 86.4% identical and 88.7% similar to the *T. cruzi* PABPI. The C-terminal one-third is the most divergent region between these two species, and contains a five amino acid insert in *T. brucei* relative to *T. cruzi* (Fig. 2). Outside of the Kinetoplastida, the *T. brucei* PABPI sequence is most homologous to the *Caenorhabditis elegans* PABPI (40.6% identity and 50.3% similarity). Homologies to other PABPIs range from 32.0–38.9% identity and 41.3–48.5%

similarity. The predicted *T. brucei* protein contains amino acid motifs common to all PABPI proteins (Fig. 2; [32,51–55]). For example, the N-terminal two-thirds of the protein comprises four RNA binding domains (RBDs; [19]), each containing an RNP-1 and an RNP-2 element (shaded in Fig. 2). The most N-terminal RNP-1 motif is the most highly conserved of all the motifs. This RNP-1 motif may be critical for poly(A) binding since the two N-terminal RBDs have been shown to be essential for this activity [18]. The C-terminal third of the protein is not highly conserved between species. However, like PABPI proteins from other species, a portion of the C-terminal region of the *T. brucei* protein is proline-rich (amino acids 383–474; 23% proline). In addition, an 11 amino acid sequence near the extreme C terminus is well-conserved among all species (boxed in Fig. 2). Interestingly, a similar sequence has been identified in the HTLV-I envelope protein gene [56] and a 100 kDa rat protein structurally related to the U1 snRNP 70 kDa protein and possibly involved in sexual maturation [57]. While the function of this motif is unknown and it is has been shown to be dispensable for binding of PABP1 to poly(A) [18], its conservation suggests it plays an important role.

Purification of a 65 kDa poly(A) binding protein from *T. brucei* by affinity chromatography on single stranded DNA-agarose and poly(A)-Sepharose has previously been reported [58]. Three peptide sequences were obtained from a tryptic digest of this protein. The protein was presumed to be PABPI based on the sequence conservation of these peptides, the strong affinity and specificity of the protein for poly(A), and the protein degradation pattern [58]. Since the PABPI cDNA described here predicts a protein of 62.2 kDa containing exact matches to all three of the previously reported peptide sequences (Fig. 1), the data indicate that we have obtained the cDNA sequence encoding the previously described 65 kDa *T. brucei* poly(A) binding protein.

Southern blot analysis was carried out on genomic DNA which had been digested with *Xho*I, *Eco*RV, *Eco*RI, or *Dra*I (data not shown). A *Xho*I restriction site is present in the PABPI open reading frame, while no sites for the other enzymes are present in the cDNAs sequenced. When probed with a radiolabeled fragment corresponding to sequence upstream of the *Xho*I site, two bands are observed in the *Xho*I digest, and one band is observed in all other lanes, indicating the presence of two copies of the PABPI gene as was observed in *T. cruzi* [51].

The full length PABPI open reading frame was cloned into pGEX-4T1 in-frame with GST. The recombinant plasmid was transformed into *E. coli*, and fusion protein expression induced with IPTG (Fig. 3). Affinity purification of GST-PABPI fusion protein on glutathione-Sepharose revealed a major protein of ≈ 91 kDa, the expected size based on the reported apparent molecular mass of *T. brucei* PABPI (65 kDa; [58]) plus the added 26 kDa of GST. Several smaller minor bands were also observed. These presumably represent GST-PABPI degradation products, since similar, discrete degradation products were observed during the purification of PABPI from both *T. brucei* and *T. cruzi* [51,58]. The minor species we observe at apparent molecular masses of 75 and 61 kDa correspond in size to the major breakdown products reported for the kinetoplastid proteins, taking into account the added size of the GST moiety. In addition, we observed binding of A-rich RNA to both the full size and 75 kDa protein (see below), indicating that the smaller protein contains PABPI sequence. This is consistent with the results of Pitula et al. [58], who reported that the largest *T. brucei* PABPI breakdown product retains the capacity to bind poly(A).

We next wanted to investigate the RNA binding properties of the GST-PABPI fusion protein. PABPI proteins from yeast [32] and humans [33] have been shown to bind the 5′ UTRs of their own mRNAs, presumably through interaction with A-rich stretches present in these RNAs. Binding affinity to 5′ UTR RNA is comparable with that observed for poly(A) [32]. Such binding may form an autoregulatory loop, and one study showed that in in vitro translation reactions addition of recombinant PABPI protein repressed translation of PABPI mRNA [33]. Like all reported PABPI mRNAs, the 5′ UTR of the *T. brucei* mRNA contains an A-rich stretch (Fig.

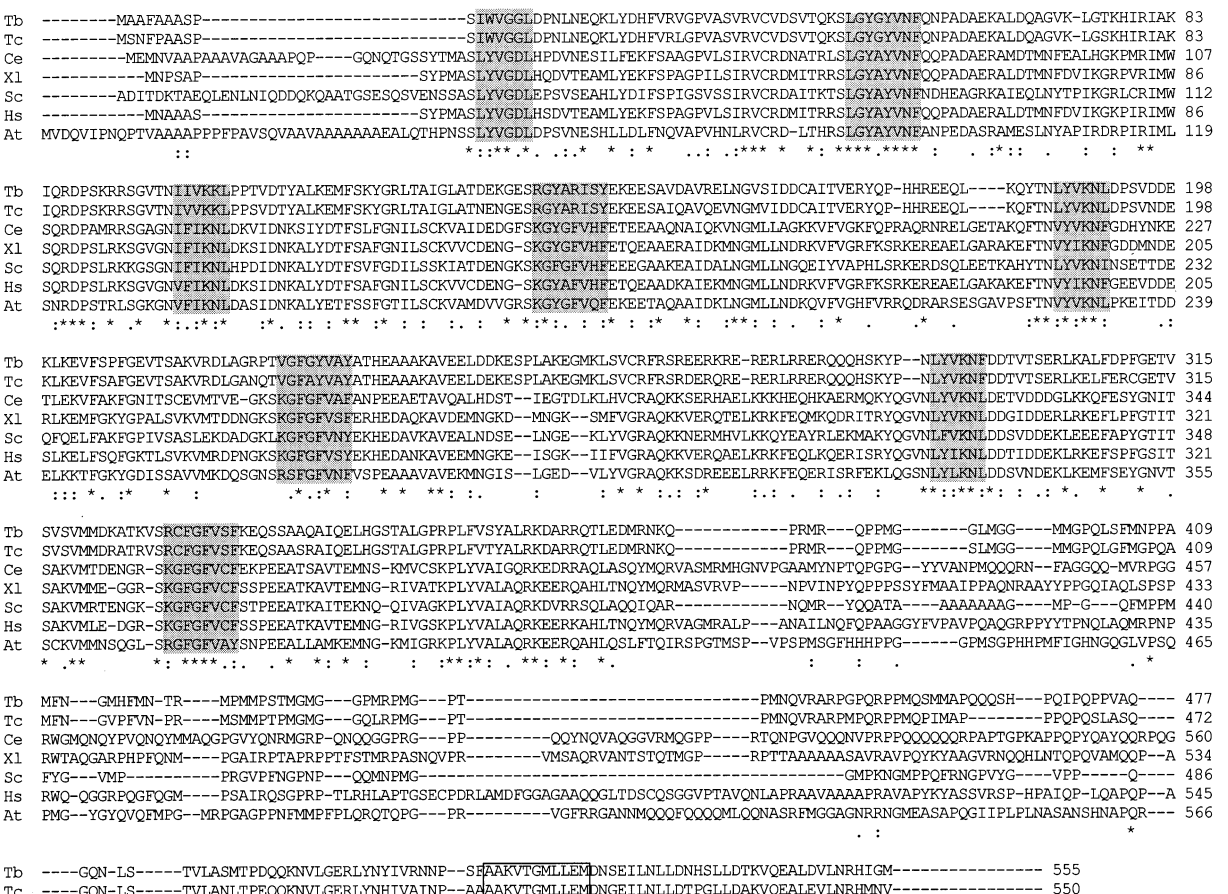

Fig. 2. Sequence comparison of PABPI proteins. PABPI amino acid sequences from *T. brucei* (Tb; GenBank accession number AF042190), *T. cruzi* (Tc; U06070), *Caenorhabditis elegans* (Ce; U24123), *Xenopus laevis* (Xl; X57483), *Saccharomyces cerevisiae* (Sc; P04147), *Homo sapiens* (Hs; U33818), and *Arabidopsis thaliana* (At; Q05196) were aligned using CLUSTAL W. * Indicates identical amino acids, : indicates conserved substitutions, and . indicates semi-conserved substitutions. RNP-1 (8-mer) and RNP-2 (6-mer) motifs are shaded. A conserved 11 amino acid motif in the C termini of all PABPI proteins and present in two non-PABPI proteins ([56,57]; see text) is boxed.

4(A)). Also present in this region are two closely spaced sets of AU repeats (Fig. 4(A)), and we thought it possible that the *T. brucei* PABPI might show affinity for AU-rich RNAs, since poly(A) tails of mitochondrial RNAs in kinetoplastids frequently contain substantial numbers of U residues [3,40,59–61]. In order to determine whether the *T. brucei* PABPI binds to its own 5′ UTR, we carried out UV cross-linking experiments. The wild type 5′ UTR RNA used in these experiments contained 311 nts comprising exclusively the PABPI 5′ UTR sequence, including the majority of the spliced leader sequence (Fig. 4(A)). Both the 91 kDa GST-PABPI protein and the 75 kDa breakdown product exhibited binding to radiolabeled wild type 5′ UTR RNA (Fig. 4(B), lane N). No binding was observed to GST alone (data not shown). To determine if the A-rich or

AU repeat regions of the 5′ UTR were involved in this binding, two deletion mutants were constructed. dA RNA harbored a deletion of 77 nts spanning the A-rich region, and dAU RNA harbored a deletion of 52 nts encompassing the AU repeats (Fig. 4(A)). Unlabeled wild type, dA, and dAU RNAs were then used in UV cross-linking competition experiments. GST-PABPI binding to radiolabled wild type 5′ UTR was significantly competed in the presence 10-fold excess unlabeled wild type 5′ UTR RNA, and was abolished by a 100-fold excess of this RNA (Fig. 4(B), WT). Deletion of the AU repeat region had no effect on the ability to compete for binding of wild type 5′ UTR RNA to GST-PABPI (Fig. 4(B), compare WT with dAU). Indeed, the 5′ UTR RNA lacking AU repeats appeared to show a slightly increased affinity for GST-PABPI compared with wild type. In contrast, 5′ UTR RNA from which the A-rich region had been deleted showed dramatically reduced ability to bind GST-PABPI (Fig. 4(B), dA). Almost no competition was observed in the presence of a 10-fold excess unlabeled dA RNA, and significant binding of wild type 5′ UTR RNA was observed even in the presence of 500-fold excess of dA RNA. These studies indicate that the *T. brucei* PABPI is capable of binding to the 5′ UTR of its own mRNA through an A-rich region near its 5′ end. The inability to bind to AU repeats suggests that the RNA binding properties of the *T. brucei* protein may be similar to that of PABPI proteins from *Xenopus* [62] and humans [63], which require a stretch of four or five consecutive adenine residues for binding. One such adenine stretch is apparently not sufficient for binding of *T. brucei* PABPI, however, since the dA mutant RNA contains a stretch of five consecutive adenines near its 3′ end, and yet is not efficiently bound. The position of adenine stretches within the RNA may also be important for PABPI binding to short A-rich regions since 'winners' from in vitro selection experiments with PABPI generally contain the adenine stretches near their 5′ ends [62,63].

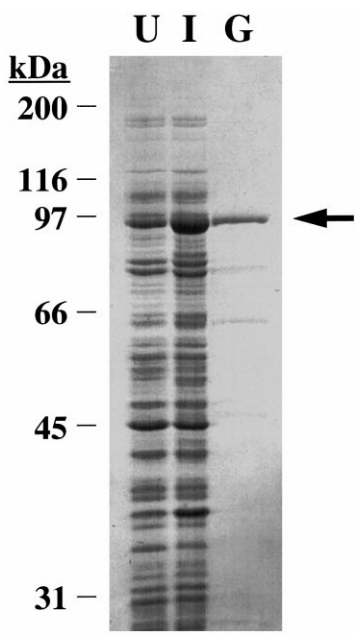

Fig. 3. Bacterial expression of GST-PABPI fusion protein. *E. coli* cells harboring the pGEX-4T1 plasmid with the *T. brucei* PABPI gene in-frame with the GST gene were treated with IPTG to induce fusion protein expression, and GST-PABPI fusion protein was affinity purified on glutathione-Sepharose 4B. Fusion protein production and purification was assessed by SDS-PAGE and Coomassie Brilliant Blue staining. Lysate of uninduced cells (U), lysate of IPTG-induced cells (I), and purified GST-PABPI (G). Arrow indicates the position of the 91 kDa GST-PABPI fusion protein. Size standards are shown on the left.

Isolation of the cDNA encoding the *T. brucei* PABPI homolog will enable future genetic and biochemical studies regarding the function(s) of this protein in trypanosomes. It is almost certain that some of these functions will be mediated by mRNA 3′ poly(A) tails. In addition, our demonstration that *T. brucei* PABPI can bind to the PABPI mRNA 5′ UTR suggests that this protein may regulate the stability and/or translation of its own RNA via 5′ UTR binding. Expression of other RNAs bearing A-rich 5′ UTRs could also potentially be affected, either positively or negatively, by PABPI-5′ UTR interactions. In vitro translation experiments demonstrated that PABPI-mediated translation repression was not restricted to PABPI mRNA [33]. Addition of a portion of human PABPI 5′ UTR or an A polymer to the 5′ end of a control RNA also resulted in repression of that RNA's translation [33]. Moreover, PABPI has recently been shown to act via 5′ UTR binding as a message-specific transla-

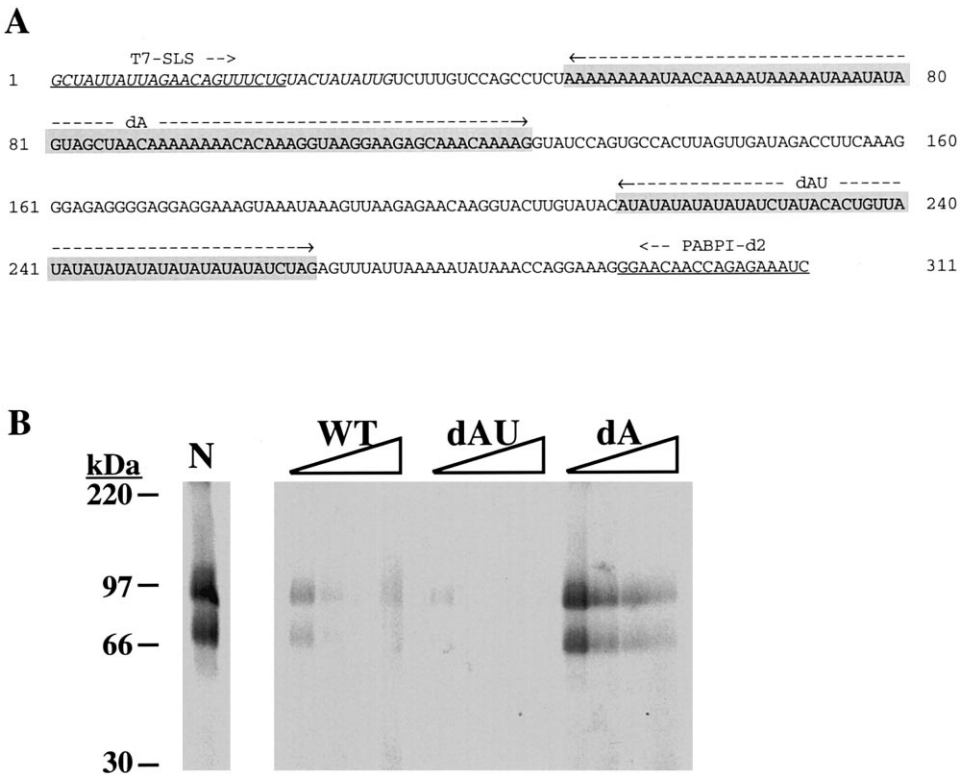

Fig. 4. Binding of GST-PABPI to PABPI 5′ UTR sequences. (A) RNAs used in UV cross-linking assays. Shown is the complete *T. brucei* PABPI 5′ UTR sequence including the portion of the spliced leader sequence present in these constructs (italics). The sequences of the oligonucleotides used to generate transcription templates are indicated by underlining, and their names and orientations shown above their sequences. The sequence shown here represents the wild type (WT) 5′ UTR RNA. The A-rich and AU repeat regions which were deleted in dA and dAU RNAs, respectively, are indicated by shading and overlining. (B) UV cross-linking competition experiments. Bacterially expressed GST-PABPI fusion protein was UV cross-linked to radiolabeled WT 5′ UTR in the presence of no competitor (N), or 10-, 50-, 100-, or 500-fold excess unlabeled RNA. Positions of size standards are shown on the left.

tional activator in *Chlamydomonas* chloroplasts [64]. Given the importance of posttranscriptional gene regulation in trypanosomatids, it is likely that PABPI will play a key role in the regulation of gene expression in these organisms.

Acknowledgements

We thank the members of the Read lab for helpful suggestions and critical reading of the manuscript. This work was supported in part by NIH grant GM53502 to LKR who is also a recipient of the Burroughs Wellcome Fund New Investigator Award in Molecular Parasitology.

References

[1] Van Hamme L, Pays E. Control of gene expression in trypanosomes. Microbiol Rev 1995;59:223–40.
[2] Gibson WC, Swinkels BS, Borst P. Post-transcriptional control of the differential expression of phosphoglycerate kinase genes in *Trypanosoma brucei*. J Mol Biol 1988;201:315–25.
[3] Read LK, Myler PJ, Stuart K. Extensive editing of both processed and preprocessed maxicircle CR6 transcripts in *Trypanosoma brucei*. J Biol Chem 1992;267:1123–8.

[4] Michelotti EF, Harris ME, Adler B, Torri AF, Hajduk SL. *Trypanosoma brucei* mitochondrial ribosomal RNA synthesis, processing, and developmentally regulated expression. Mol Biochem Parasitol 1992;54:31–42.
[5] Bringaud F, Baltz T. Differential regulation of two distinct families of glucose transporter genes in *Trypanosoma brucei*. Mol Cell Biol 1993;13:1146–54.
[6] Revelard P, Lips S, Pays E. Alternative splicing within and between alleles of the ATPase gene 1 locus of *Trypanosoma brucei*. Mol Biochem Parasitol 1993;62:93–102.
[7] Koslowsky DJ, Yahampath G. Mitochondrial mRNA 3′ cleavage/polyadenylation and RNA editing in *Trypanosoma brucei* and independent events. Mol Biochem Parasitol 1997;90:81–94.
[8] Clayton C, Hotz H-R. Post-transcriptional control of PARP gene expression. Mol Biochem Parasitol 1996;77:1–6.
[9] Ehlers B, Czichos J, Overath P. RNA turnover in *Trypanosoma brucei*. Mol Cell Biol 1987;7:1242–9.
[10] Berberof M, Vanhamme L, Tebabi P, Pays A, Jeffries D, Welburn S, Pays E. The 3′-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. EMBO J 1995;14:2925–34.
[11] Furger A, Schürch N, Kurath U, Roditi I. Elements in the 3′ untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. Mol Cell Biol 1997;17:4372–80.
[12] Hotz HR, Hartmann C, Huober K, Hug M, Clayton C. Mechanisms of developmental regulation in *Trypanosoma brucei*-A polypyrimidine tract in the 3′-untranslated region of a surface protein mRNA affects RNA abundance and translation. Nucleic Acids Res 1997;25:3017–25.
[13] Hehl A, Vassella E, Braun R, Roditi I. A conserved stem loop structure in the 3′ untranslated region of procyclin mRNAs regulates expression in *Trypanosoma brucei*. Proc Natl Acad Sci 1994;91:370–4.
[14] Gale M Jr, Carter V, Parsons M. Translational control mediates the developmental regulation of the *Trypanosoma brucei* Nrk protein kinase. J Biol Chem 1994;269:31659–65.
[15] Zhang J, Ruyechan W, Williams N. Developmental regulation of two nuclear RNA binding proteins, p34 and p37, from *Trypanosoma brucei*. Mol Biochem Parasitol 1998;92:79–88.
[16] Sachs A, Wahle E. Poly(A) tail metabolism and function in eucaryotes. J Biol Chem 1993;268:22955–8.
[17] Baker E. mRNA polyadenylation: Functional implications. In: Harford JB, Morris DR, editors. mRNA Metabolism and Post-Transcriptional Gene Regulation. New York: Wiley, 1997:85–105.
[18] Burd CG, Matunis EL, Dreyfuss G. The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA binding activities. Mol Cell Biol 1991;11:3419–24.
[19] Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. Science 1994;265:615–21.
[20] Sachs AB, Davis RW. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell 1989;58:857–67.
[21] Monroe D, Jacobson A. mRNA poly(A) tail, a 3′ enhancer of translational initiation. Mol Cell Biol 1990;10:3441–55.
[22] Gallie DR, Tanguay R. Poly(A) binds to initiation factors and increases cap-dependent translation *in vitro*. J Biol Chem 1994;269:17166–73.
[23] Tarun SZ, Sachs AB. A common function for mRNA 5′ and 3′ ends in translation initiation in yeast. Genes Dev 1995;9:2997–3007.
[24] Tarun SZ, Sachs AB. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J 1996;15:7168–77.
[25] Tharun S, Parker R. Mechanisms of mRNA turnover in eukaryotic cells. In: Harford JB, Morris DR, editors. mRNA Metabolism and Post-Transcriptional Gene Regulation. New York: Wiley, 1997:181–99.
[26] Muhlrad D, Parker R. Mutations affecting stability and deadenylation of the yeast MFA2 transcript. Genes Dev 1992;6:2100–11.
[27] Berstein P, Peltz SW, Ross J. The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability *in vitro*. Mol Cell Biol 1989;9:659–70.
[28] Sachs AB, Deardorff J. Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. Cell 1992;70:961–73.
[29] Boeck R, Tarun S, Rieger M, Deardorff J, Müller-Auer S, Sachs AB. The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. J Biol Chem 1996;271:432–8.
[30] Amrani N, Minet M, le Gouar M, Lacroute F, Wyers F. Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length in vitro. Mol Cell Biol 1997;17:3694–701.
[31] Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3′-end formation. Proc Natl Acad Sci 1997;94:7897–902.
[32] Sachs AB, Bond MW, Kornberg RD. A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. Cell 1986;45:827–35.
[33] de Melo Neto OP, Standart N, Martins de Sa C. Autoregulation of poly(A)-binding protein synthesis *in vitro*. Nucleic Acids Res 1995;23:2198–205.
[34] LeBowitz JH, Smith HQ, Rusche L, Beverley SM. Coupling of poly(A) site selection and *trans*-splicing in *Leishmania*. Genes Dev 1993;7:996–1007.

[35] Ullu E, Matthews KR, Tschudi C. Temporal order of RNA-processing reactions in trypanosomes: rapid *trans* splicing precedes polyadenylation of newly synthesized tubulin transcripts. Mol Cell Biol 1993;13:720–5.

[36] Bhat GJ, Souza AE, Feagin JE, Stuart K. Transcript-specific developmental regulation of polyadenylation in *Trypanosoma brucei* mitochondria. Mol Biochem Parasitol 1992;52:231–40.

[37] Pays E, Hanocq-Quertier J, Hanocq F, Van Assel S, Nolan D, Rolin S. Abrupt RNA changes precede the first cell division during the differentiation of *Trypanosoma brucei* bloodstream forms into procyclic forms in vitro. Mol Biochem Parasitol 1993;61:107–14.

[38] Bhat GJ, Myler PJ, Stuart K. The two ATPase 6 mRNAs of *Leishmania tarentolae* differ at their 3′ ends. Mol Biochem Parasitol 1991;48:139–50.

[39] Read LK, Stankey KA, Fish WR, Muthiani AM, Stuart K. Developmental regulation of RNA editing and polyadenylation in four life cycle stages of *Trypanosoma congolense*. Mol Biochem Parasitol 1994;68:297–306.

[40] Souza AE, Myler PJ, Stuart K. Maxicircle *CR1* transcripts of *Trypanosoma brucei* are edited and developmentally regulated and encode a putative iron-sulfur protein homologous to an NADH dehydrogenase subunit. Mol Cell Biol 1992;12:2100–7.

[41] Stuart KD, Gobright E, Jenni L, Milhausen M, Thomashow L, Agabian N. The IsTaR 1 serodeme of *Trypanosoma brucei*: development of a new serodeme. J Parasitol 1984;70:747–54.

[42] Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidimium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987;162:156–9.

[43] Carrington M. Preparation of DNA and RNA from *Trypanosoma brucei*. In: Hyde JE, editor. Protocols in Molecular Parasitology. Totowa: Humana Press, 1993:101–11.

[44] Bhat GJ, Lodes MJ, Myler PJ, Stuart KD. A simple method for cloning blunt ended DNA fragments. Nucleic Acids Res 1991;19:398.

[45] Qi L, Nicholas RA. PCR strategy to obtain the 5′ and 3′ flanking sequences of partial clones from λZAP cDNA libraries. BioTechniques 1997;22:1038–44.

[46] Devereux J, Haebberli P, Smithies O. A set of sequence analysis programs for the VAX. Nucleic Acids Res 1984;12:387–95.

[47] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.

[48] Read LK, Wilson KD, Myler PJ, Stuart K. Editing of *Trypanosoma brucei* maxicircle CR5 mRNA generates variable carboxy terminal predicted protein sequences. Nucleic Acids Res 1994;22:1489–95.

[49] Ausabel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. Current Protocols in Molecular Biology. New York: Wiley, 1995.

[50] Read LK, Göringer HU, Stuart K. Assembly of mitochondrial ribonucleoprotein complexes involves specific guide RNA (gRNA)-binding proteins and gRNA domains but does not require preedited mRNA. Mol Cell Biol 1994;14:2629–39.

[51] Batista JAN, Teixeira SMR, Donelson JE, Kirchhoff LV, Martins de Sá C. Characterization of a *Trypanosoma cruzi* poly(A)-binding protein and its genes. Mol Biochem Parasitol 1994;67:301–12.

[52] Grange T, Martiins de Sá C, Oddos J, Pictet R. Human mRNA polyadenylate binding protein: evolutionary conservation of a nucleic acid binding motif. Nucleic Acids Res 1987;15:4771–87.

[53] Lefrere V, Vincent A, Almaric F. *Drosophila melanogaster* poly(A)-binding protein: cDNA cloning reveals an unusually long 3′-untranslated region of the mRNA, also present in other eukaryotic species. Gene 1990;96:219–25.

[54] Nietfeld W, Mentzel H, Pieler T. The *Xenopus laevis* poly(A) binding protein is composed of multiple functionally independent RNA binding domains. EMBO J 1990;9:3699–705.

[55] Belostotsky DA, Meagher RB. Differential organ-specific expression of three poly(A)-binding-protein genes from *Arabidopsis thaliana*. Proc Natl Acad Sci 1993;90:6686–90.

[56] Malik KT, Even J, Karpas A. Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia virus/human T cell leukaemia virus type I (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other members of the ATLV/HTLV-I subgroup. J Gen Virol 1988;69:1695–710.

[57] Müller D, Rehbein M, Baumeister H, Richter D. Molecular characterization of a novel rat protein structurally related to poly(A) binding proteins and the 70K protein of the U1 small nuclear ribonucleoprotein particle (snRNP). Nucleic Acids Res 1992;7:1471–5.

[58] Pitula J, Ruyechan WT, Williams N. *Trypanosoma brucei*: identification and purification of a poly(A)-binding protein. Exp Parasitol 1998;88:157–60.

[59] Campbell DA, Spithill TW, Samaras N, Simpson AM, Simpson L. Sequence of a cDNA for the ND1 gene from *Leishmania major*: potential uridine addition in the polyadenosine tail. Mol Biochem Parasitol 1989;36:197–200.

[60] Decker CJ, Sollner-Webb B. RNA editing involves indiscriminate U changes throughout precisely defined editing domains. Cell 1990;61:1011–1.

[61] Van der Spek H, Speijer D, Arts G-J, Van den Burg J, Van Steeg H, Sloof P, Benne R. RNA editing in transcripts of the mitochondrial genes of the insect trypanosome *Crithidia fasciculata*. EMBO J 1990;9:257–62.

[62] Kühn U, Pieler T. *Xenopus* poly(A) binding protein: functional domains in RNA binding and protein–protein interaction. J Mol Biol 1996;256:20–30.

[63] Gorlach M, Burd CG, Dreyfuss G. The mRNA poly(A)-binding protein: localization, abundance, and RNA binding specificity. Exp Cell Res 1994;211:400–7.

[64] Yohn CB, Cohen A, Danon A, Mayfield SP. A poly(A) binding protein functions in the chloroplast as a message-specific translation factor. Proc Natl Acad Sci 1998;95:2238–43.