ELSEVIER

Molecular and Biochemical Parasitology 110 (2000) 207–221

MOLECULAR AND BIOCHEMICAL PARASITOLOGY

www.elsevier.com/locate/parasitology

PPB1, a putative spliced leader RNA gene transcription factor in *Trypanosoma cruzi*☆

Li-Min Wen, Ping Xu, Gauri Benegal, Maria Ruth M. Carvalho, Gregory A. Buck *

Department of Microbiology and Immunology, Medical College of Virginia Campus, Box 980678, Virginia Commonwealth University, Richmond, VA 23298-0678, USA

Received 16 February 2000; received in revised form 17 May 2000; accepted 17 May 2000

Abstract

In trypanosomatids, the spliced leader RNA, or SL RNA, donates its 5′ 39 nucleotides to mature nuclear mRNAs in a process termed trans-splicing. We have previously characterized the SL RNA gene from *Trypanosoma cruzi* and identified its transcription promoter, including a 14 nt proximal sequence element, or PSE, that binds a putative transcription factor and activates transcription of the gene. Herein, we describe establishment of a yeast one-hybrid system using the 14 nt PSE as bait, and use this system to select *T. cruzi* cDNAs encoding a putative transcription factor that activates transcription of the SL RNA gene. The cDNA was selected from a normalized library and encodes an ~45 kDa putative PSE promoter-binding protein, PPB1. PPB1 in vitro translated or overexpressed in and isolated from transformed *E. coli*, showed PSE-specific binding activity by electrophoretic mobility shift assays. Finally, overexpression of PPB1 in *T. cruzi* led to increased expression of the SL RNA gene as well as reporter genes in episomal constructs under the control of the SL RNA gene promoter. These observations suggest that PPB1 is a transcription factor that plays an important role in SL RNA gene expression.

Keywords: cDNA library normalization; DNA-protein interaction; One-hybrid system; PSE promoter; SL RNA; Transcription factor

Abbreviations: 3-AT, 3-aminotriazole; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift assays; GFP, green fluorescent protein; PPB1, proximal sequence element promoter-binding protein 1; PSE, proximal sequence element; SL, spliced leader..

☆ *Note*: Nucleotide sequence data reported in this paper is available in the GenBank™, EMBL and DDBJ databases under the accession number AF213249.

* Corresponding author. Tel.: +1-804-8282318; fax: +1-804-8281397.

E-mail address: buck@hsc.vcu.edu (G.A. Buck).

0166-6851/00/$ - see front matter

PII: S0166-6851(00)00271-1

1. Introduction

The mechanism of mRNA maturation in the kinetoplast protozoa is distinct from the unimolecular or *cis*-splicing pathways of higher eukaryotes in its requirement for a bio-molecular or *trans*-splicing event which joins a 39-nt spliced leader (SL) to the 5′ terminus of the mature transcript [1]. The small abundant SL RNAs that donate their capped 5′ 39 nt SL to mature mRNAs in the trans-splicing reaction are transcribed from clusters of tandem direct repeats of SL RNA genes. Transcription promoters have been difficult to identify in trypanosomes because of the polycistronic nature of the transcripts of protein-coding genes. However, in all trypanosomatids, each SL RNA gene repeat is presumed to bear a transcription promoter [2–5]. We [6] and others [7–9] have previously identified putative proximal sequence elements (PSE) of the transcription promoters of the SL-RNA genes from *Trypanosoma cruzi* and several other trypanosomatids. Because of the differences between the transcription mechanisms of the trypanosomatids and other eukaryotes, we anticipate that the transcription factors may also differ substantially.

Luo and Bellofatto [10] recently described identification of several factors, PBP-1 and PBP-2, that bind to and co-purify with the SL RNA gene promoter of *Leptomonas seymouri*. Allthough PBP-1 and PBP-2 were subsequently shown to be required for in vitro transcription from the SL RNA promoter *L. seymouri* [11], these factors have not yet been characterized at the genetic level. A more complete genetic dissection of transcription factors of these kinetoplastid protozoa is necessary to provide insight into the transcriptional mechanisms and provide possible clues to the development of new chemotherapeutic strategies.

The yeast one-hybrid system has been developed and used to study specific DNA-protein interactions [12–19]. This system has been particularly useful for isolation of presumptive transcription factors that activate genes by binding known promoter elements [20–30]. Usually, the one-hybrid system is used to screen cDNA libraries. However, the mRNAs encoding many transcription factors are present only at very low copy numbers in eukaryotic cells [31], and abundant messages for highly transcribed genes, particularly those encoding nucleic acid binding proteins such as the many abundant ribosomal proteins and histones, etc., are often selected as 'false positives'. To obviate this problem, the cDNAs library can be 'normalized' prior to the screening; i.e. abundant messages are selectively removed, or low-abundance messages are selectively amplified. Normalization has resulted in the successful cloning of many rare eukaryotic transcripts [32,33].

Our objective in this study was to identify transcription factors that activate the PSE of the SL RNA gene in *T. cruzi*. Thus, we constructed normalized cDNA libraries from *T. cruzi* mRNA and developed and used a yeast one-hybrid system to screen these libraries for *T. cruzi* cDNAs that encode proteins that bind to and activate PSE of the SL RNA.

2. Materials and methods

2.1. T. cruzi culture, nucleic acid isolation and initial cDNA library construction

Epimastigotes of *T. cruzi* CL Brener were grown essentially as previously described [34] in LIT medium supplemented to 10% with fetal calf serum at 28°C under gentle agitation. Genomic DNA was isolated from exponential phase *T. cruzi* as previously described [35]. Total RNA from *T. cruzi* was isolated by LiCl precipitation essentially as previously described [36]. Poly(A)$^+$ RNA was purified by oligo(dT) cellulose affinity chromatography in an mRNA Separator kit (Clontech). The initial non-normalized *T. cruzi* cDNA library was constructed from poly(A)$^+$ RNA using a HybriZAP® Two-hybrid cDNA Synthesis Kit (Stratagene).

2.2. Conversion of normalized single-stranded cDNA into double-stranded plasmids

The single stranded normalized library was converted to double strand with Sequenase (USB)

essentially as previously described [37], using ~100 ng of the single-stranded normalized cDNA library mixed with ~0.5 μg of each of the following primers (N3′, 5′-GAGATCGAATTCGGCACGAG-3′; AD5′-*Eco*RV; 5′-GCTTCATCGGAGATGATATCA-3′; AD5′-*Not*I, 5′-TAACGCTTACAATTTACGCGC-3′; AD5′-*Dar*II, 5′-AAAAATAGGCGTATCACGAGG-3′). The double-stranded plasmids were transformed into *E. coli* XL-1 blue bacteria by electroporation.

2.3. Plasmid construction and selections for yeast one-hybrid screening

As described in the Results (below), two vectors, pHISi which carries the *his*3 reporter gene, and pLacZi which carries the *lacZi* reporter gene, were used to generate new yeast strains containing five copies of the target element upstream from the relevant reporter gene. Double stranded pentameric oligonucleotides bearing five tandem copies of the PSE promoter element at −60 to −31 upstream from the SL RNA gene were generated by ligation of 5′-phosphorylated p55 ds oligonucleotides; i.e. hybridized p55 BB 5′, 5′-GATCCAAGTGCCGCGAAGGACCCCTCATCAAAATA-3′ and p55 BB 3′, 5′-GATCTATTTTGATGAGGGGTCCTTCGCGGCACT-TG-3′. These pentamers were purified by electrophoresis in a 2% agarose gel, electroeluted, digested and ligated into the *Bam*HI site of pUC19 vector (Stratagene). After verification by sequence analysis, the pentamer oligo was excised from pUC19 by digestion with *Eco*RI/*Xba*I or *Eco*RI/*Sal*I and ligated upstream of the *his* and *β-gal* genes in the pHISi or pLacZi vectors linearized with *Eco*RI/*Xba*I or *Eco*RI/*Sal*I, resp., to generate pHISi-5 × PSE and pLacZi-5 × PSE. These target-reporter constructs were linearized by *Xho*I or *Nco*I digestion at the *his*3 or *ura*3 loci, resp, to facilitate homologous integration, transformed into the YM4271 (MATa, *ura*3-52, *his*3-200, *ade*2-101, *lys*2-801, *leu*2-3, 112, *trp*1-903, *tyr*1-501) by electroporation, and selected for complementation of the auxotrophic marker (*his3* or *ura3*) essentially as previously described [37] to generate pHISi-5 × PSE/YM4271 and pLacZi-5 × PSE/YM4271. Plasmid integration in the genome of yeast strains was confirmed by Southern analysis using a [^{32}P]labeled p55 oligonucleotide.

The pHISi-5 × PSE/YM4271 yeast strain was electroporated with the normalized *T. cruzi* cDNA library and selected on SD His$^-$Leu$^-$ plates containing 60 mM 3-AT to prevent growth of colonies due to leaky histidine expression. Plasmid was isolated from presumptive positives using Zymoprep™ yeast plasmid minipreparation kits (ZYMO RESEARCH) and re-transformed into the pLacZi-5 × PSE/YM4271 strain to screen for β-galactosidase expression. Transformant colonies were lifted on to Whatman No. 5 filter paper and analyzed for β-galactosidase expression as described by the manufacturer (Clontech).

2.4. RT-PCR, primer extension and genome walking

For RT-PCR, total *T. cruzi* RNA was treated with RNase-free DNase (Promega). First strand synthesis was performed on ~5 μg of this RNA with Superscript II reverse transcriptase (Gibco-BRL) according to the manufacturer's protocols. The full length 5′ end of the PPB1 cDNA was amplified using *Pfu* DNA polymerase (Stratagene) with the SL sequence as a 5′ primer (SL5′: 5′-AACTAACGCTATTATTGATACAG-3′) and the PPB1 cDNA sequence as 3′ primer (H3′: 5′ - CCCAAGCTTTCAAAAAATATTTCTACATCGGA-3′) under the conditions: 94°C/1 min, 55°C/2 min, 72°C/3 min for 30 cycles. For primer extension, ~60 μg total *T. cruzi* RNA was mixed with ~10^6 cpm [^{32}P]labeled PPB1-top oligo (5′-ATGGCTCAAGAGGCGTTTCC T-3′) in RTase buffer and extended with AMV Reverse transcriptase (USB) essentially as previously described [4].

Genomic DNA fragments containing the full-length PPB1 gene were amplified by nested PCR from *T. cruzi* genomic DNA libraries using the Universal Genome Walker™ Kit (Clontech) essentially as described by the manufacturer using four gene-specific primers (GSP1-5′: 5′-ACGCCTGAGCCATACTCGCAATGCAG-3′; GSP2-5′: 5′ - GGAATTCGCCGTGAACGGGGGAATCC-

AGAGGCGTTC-3′; GSP1-3′: 5′-GGCCCTAG-CATCATCCCTGAGAAGGACTC-3′; GSP2-3′: 5′ - GGAATTCGCGGAACGTACTGAACCAT-GGGGAGTTTG-3′). The amplified fragments were cloned into pGEM-3Zf (+) vector (Promega) using standard protocols [37].

2.5. Expression of PPB1 in reticulocyte lysates and E. coli

An *Eco*RI-*Xba*I fragment containing the full-length coding sequence of PPB1 excised from the pADGAL4-PPB1 was ligated into the *Eco*RI-*Xba*I site of the pGEM-3Zf (+) vector using standard protocols to generate pGEM-PPB1. This plasmid was incubated in the $T_NT^{®}$ T_7 quick coupled transcription/translation system (Promega) in the presence of [^{35}S] methionine (Amersham) to generate in vitro translated protein which was revealed by electrophoresis on sodium dodecyl sulfate (SDS)-polyacrylamide gels and autoradiography following standard technology [37].

For expression of PPB1 in *E. coli*, a PPB1 cDNA fragment was PCR amplified with primers containing either a *Sph*I or a *Hin*dIII site, S5′: 5′ - ACATGCATGCATGGCTCAAGAGGCGT-TT-3′, and H3′: 5′-CCCAAGCTTTCAAAAA-AATATTTCTACATCGGA-3′, resp., using *Pfu* DNA Polymerase (Stratagene), with cycling conditions of 94°C/1 min, 55°C/2 min, 2°C/3 min, for 30 cycles. The amplified fragments were digested and cloned into *Sph*I-*Hin*dIII site of the pQE-30 vector (Qiagen) to generate pQE-30-PPB1. pQE-30-PPB1 was transformed into *E. coli* strain M15 (pREP4).

For purification of the histidine-tagged PPB1 protein, transformed bacteria were grown and induced by IPTG at a final concentration of 0.5 mM for 3 h. The bacteria were harvested by centrifugation and all subsequent procedures were performed at 4°C on ice. Cells were resuspended in buffer A (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 5 mM imidazole, 5 mM 2-mercaptoethanol) containing 1 mM phenylmethylsulfonyl fluoride, treated with 1 mg/ml lysozyme at 4°C for 1 h, lysed by sonication, and centrifuged at 10 000 × *g* for 40 min. The clear lysate was loaded onto a Ni-NTA column (Qiagen) and chromatography was performed at a flow rate of 0.5 ml/min. The loaded column was washed first with buffer B (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 40 mM imidazole, 5 mM 2-mercaptoethanol, 0.1% Tween 20), then with buffer A to remove the excess Tween. The bound protein was eluted with buffer C (50 mM NaH_2PO_4 [pH 8.0], 300 mM NaCl, 1.0 M imidazole, 5 mM 2-mercaptoethanol).

2.6. Electrophoretic mobility shift assays (EMSA)

A target double stranded oligonucleotide homologous to the −60 to −31 region of the SL RNA gene and containing the PSE promoter element was prepared by annealing oligos p55, 5′ - CAAGTGCCGCGAAGGACCCCTCATCA-AAAT-3′; and p55′, 5′-ATTTTGATGAGG-GGTCCTTC GCGGCACTTG-3′, labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP, and used for EMSA as previously described [6]. Similarly prepared double strand TATA oligos from 5′-GCAGAGCATATAAGGTGAGGTAGGA-3′ and 5′-TCCTACCTCACCTTATATGCTCTGC-3′, were used as controls. Approximately 2 μl of the in vitro translated protein or 0.5–1 μg of the purified histidine-tagged PPB1 protein was incubated with ∼15 000 cpm of each 5′-end labeled oligonucleotide in a volume of 20 ul of binding buffer in the presence of 2 μg of poly (dI:dC) (Sigma). For competitions, the indicated quantities of unlabeled oligonucleotide were added to the reaction mixture prior to the addition of protein.

2.7. Construction of plasmids pPATAN-PPB1 and pPATAN-GFP/PPB1

The PPB1 gene was amplified by PCR from pADGAL4-PPB1, digested with *Xba*I/*Nde*I, and cloned into the *Xba*I/*Nde*I site of the pPATAN *T. cruzi* expression vector (D.R. Butler and G.A. Buck, unpublished) to generate pPATAN-PPB1. To generate pPATAN-GFP/PPB1, the PPB1 gene was amplified from pADGAL4-PPB1 as above, digested with *Bgl*II/*Hind*III, and cloned into the *Bgl*II/*Hind*III site of pEGFP-C3 C-Terminal

Protein Fusion Vector (Clontech). The GFP-PPB1 fusion fragment in the latter construct was reamplified, digested with *Xba*I/*Nde*I, and cloned into *Xba*I/*Nde*I site of pPATAN to generate pPATAN-GFP/PPB1.

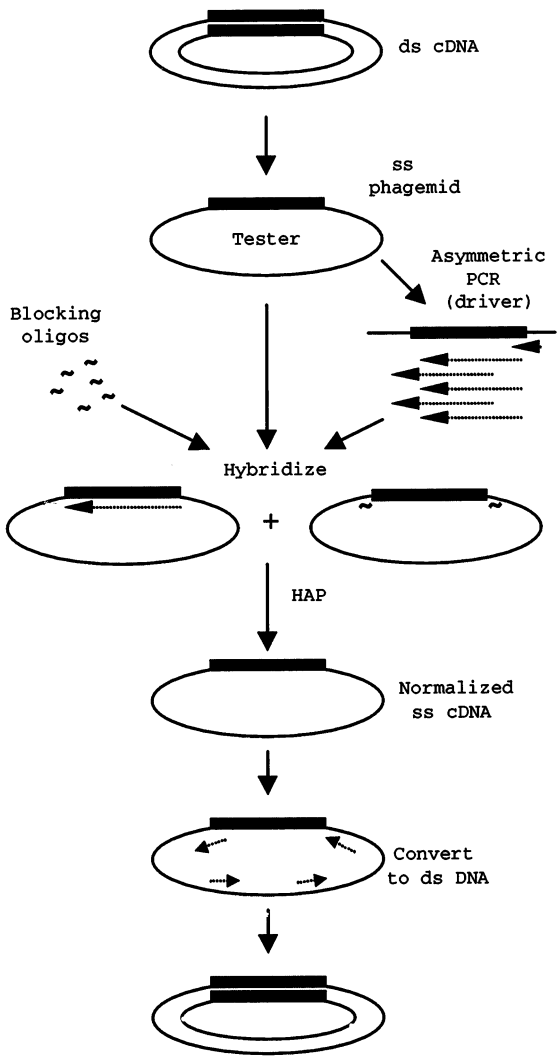

Fig. 1.

3. Results

3.1. Construction of a normalized cDNA library from T. cruzi epimastigotes

Standard cDNA libraries contain an excess of clones from highly expressed genes with abundant messages, and these abundant messages often generate false positives in functional screens [38–40]. Therefore, we used a normalized cDNA library to screen for proteins that bind to the SL RNA PSE. The library was generated from a standard *T. cruzi* cDNA library constructed with the HybriZAP® Two-hybrid cDNA Synthesis Kit (Stratagene) as outlined in the Materials and methods. The cDNAs were fused with the GAL4 activation domain in the pAD-GAL4 phagemid vector (Stratagene), electroporated into *E. coli* XL-1 blue, and single stranded DNA was isolated. The single stranded DNA was fractionated by hydroxyapatite (HAP) chromatography to eliminate contaminating double-stranded DNA.

The cDNA library was normalized by PCR amplification, hybridization and hydroxyapatite chromatography as described in Fig. 1. The level of normalization achieved was examined by comparing the levels of known high and low abun-

Fig. 1. Normalization of the cDNA library. A double strand (ds) cDNA library was constructed from *T. cruzi* poly(A)$^+$ RNA using the HybriZAP® Two-hybrid cDNA Synthesis Kit (Stratagene). Single stranded (ss) phagemid DNA was prepared from the library and inserts were PCR amplified using primers (N5′: 5′-AGAGTCGACC CGGGCTCGAG-3′, N3′: 5′-GAGATCGAATTCGGCACGAG-3′) targeted to sequences flanking the cloning sites of the pAD-GAL4 cloning vector. Fragments larger than 400 bp were purified after electrophoresis in a low-melting temperature agarose, single strand antisense DNA was produced by 20 cycles of single primer amplification using only the N3′ primer, and these single stranded PCR products were gel purified. The single stranded antisense cDNA PCR products was hybridized to a calculated $C_o t$ of ~ 18 [40] in a 5:1 mass ratio with the single-stranded 'sense' cDNA library in the presence of 5′ and 3′ blocking oligos (5′-CTCGAGCCCGGGTCGACTCT-3′, 5′-GAGATCGAATTCGGCACGAG-3′). These products were fractionated by hydroxyapatite chromatography to eliminate 'high-$C_o t$' fragments, and the remaining single-stranded cDNAs representing the normalized library were converted into double stranded plasmids by primer extension as described in the Materials and methods and used to transform *E. coli*.

dance mRNAs represented in the initial non-normalized and final libraries in two ways (data not shown). First, probes for the *T. cruzi* ribosomal protein P-JL5, histone H3, and the topoisomerase II genes, representing abundant, intermediate and rare transcripts, resp. [41], were hybridized to slot blots of cDNAs isolated from the libraries before and after normalization, and the results were quantified by radiofluorography. The results showed greater than 50% relative reduction in hybridization of the ribosomal protein P-JL5 probe to the normalized library compared to the non-normalized library. Similarly, the histone probe exhibited a $>25\%$ decrease in hybridization in the normalized library. In contrast, the topoII probe showed no decrease in hybridization in the normalized library. In parallel, the normalized and non-normalized libraries were used to transform *E. coli*, and the resultant colonies were hybridized to the same three radiolabeled probes. The results paralleled those of the slot blots, i.e. the number of ribosomal protein and histone clones in the normalized library were reduced greater than 50 and 25%, respectively, whereas the number of topoisomerase II clones remained essentially constant. Thus, the proportional representation of the topoII low abundance transcript was increased in the normalization process, while the representation of high abundance mRNAs decreased, indicating that the library had been successfully partially normalized.

3.2. One-hybrid system identifies a putative PSE promoter-binding factor

The normalized cDNA library was generated in vector pADGAL4 so that the cloned genes are fused with the GAL4 activation domain (Fig. 2A). Vector pHIS-5 × PSE bearing five copies of a 30 bp double stranded oligo containing the ~14

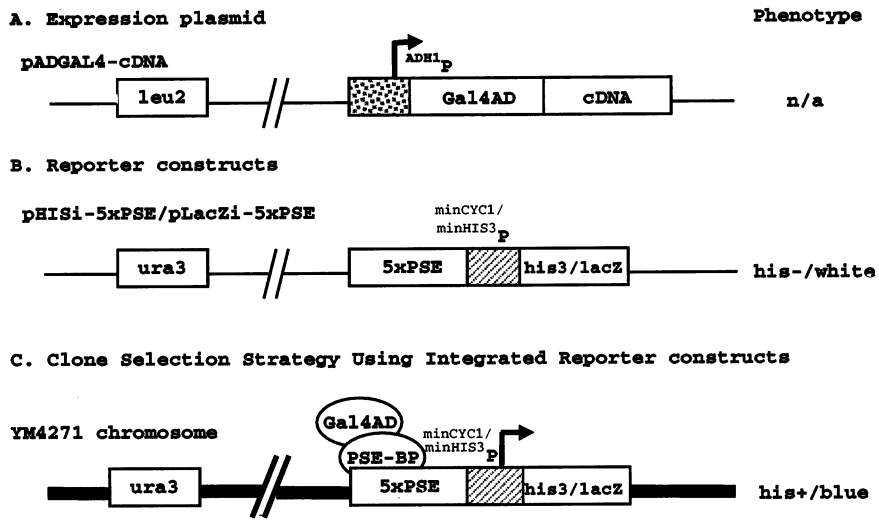

Fig. 2. One-hybrid system for isolation of PSE-binding protein. (A) Structure of the pADGAL4 expression plasmid. As described in the text, a cDNA library was constructed using the HybriZAP® Two-hybrid cDNA synthesis kit (Stratagene) according to manufacturer's instructions. Vector sequences are indicated as thin lines, the selectable marker *leu*2, and the Gal4AD/cDNA hybrid bearing the activation domain of the Gal4 activation factor fused to the *T. cruzi* cDNA, and the yeast ADH1 promoter sequence are indicated. (B) Structure of pHIS1-5xPSE and pLacZi-5xPSE for integration into yeast chromosomes. The selectable marker *ura*3, and the hybrid reporter genes are indicated with the 5 × PSE denoting the multimeric PSE sequence inserted upstream from the minimal *his*3 and *cyc*1 gene promoters linked to the *his*3 and *lac*Z genes, respectively. There is no transcription from either pHISi-5xPSE or pLacZi-5xPSE after integration with 60 mM 3-AT in the medium, and the cells are phenotypically his^- or white, respectively. (C) Schematic of overall strategy. If a cDNA encodes a protein (PSEBP) able to recognize the PSEs, the binding of this factor will lead to activation via the Gal4 activation domain and cells will express *his*3 or *lac*Z and therefore will be his^+ or blue, respectively.

bp SL RNA gene PSE in the promoter region driving transcription of the yeast *his*3 gene (Fig. 2B) was integrated into the chromosome of yeast strain YM4271 as described in the Materials and methods. The yeast were transformed with the normalized *T. cruzi* cDNA library and plated on selective medium. Of $\sim 5 \times 10^6$ independent clones plated, 11 independent clones, representing ten discrete mRNAs, formed colonies indicating upregulation of the selectable marker gene presumably via activation by a *T. cruzi* protein fused to the GAL4 activation domain (Fig. 2C). To confirm these presumptive positives, plasmids were isolated from each of these clones and used to transform a yeast strain integrated with vector pLacZi-5 × PSE (Fig. 2B) containing five copies of the PSE-bearing oligo in the regulatory element of chromosomal *lacZ* reporter gene, and β-galactosidase activity induced via interaction of the fusion peptide with the *T. cruzi* PSE (Fig. 2C) was assessed. Two independent clones, S4 and T6, from the same gene, PPB1, that exhibited β-gal expression were selected for further investigation.

Primer extension analysis using a PPB1-specific primer and *T. cruzi* total RNA showed that the two cDNA clones differed only in the length of their overlapping sequences and together lacked only ~100 5′ bases (data not shown). The full-length genomic clone was isolated as described in the Materials and methods (Fig. 3). The genomic sequence exhibits a 1239 bp single open reading frame that encodes a 413 amino acid protein of ~45 kDa (Fig. 3). Northern analysis (not shown) showed that the full length mRNA is about 1.4 kb, and a sequence resembling trypanosomatid 3′ splice sites [42]; i.e. an AG dinucleotide preceded by a short polypyrimidine track, is present immediately upstream of the putative 3′ splice site of the PPB1 transcript (Fig. 3). Southern analyses indicated that the PPB1 gene is a single copy gene (data not shown).

The sequence of PPB1 yielded no significant 'hits' in homology searches against the non-redundant GenBank, SwissPlus or PIR databases (GCG) using BLASTX or BLASTN [43]. However, the protein exhibits a leucine-zipper like motif (bZIP) [31,44] at amino-acid positions 9–52, a relaxed leucine repeat in which any of the leucines is replaced by hydrophobic amino acids Met, Ile or Val. Secondary structure analysis of the PPB1 protein predicted that amino acids 52–71 form a helix–turn–helix (HTH) motif immediately followed by a Bzip–Basic region (amino acids 72–97), and two helix–loop–helix (HLH) motifs at amino acids 299–315 and 348–364. These structures are all consistent with a DNA-binding function for PPB1 [45]. The leucine-zipper like motif (bZIP), HTH motif and Bzip–Basic region are juxtaposed, which is consistent with the observation that leucine-zippers frequently occur together with a DNA binding basic region or an HLH domain, and are a common composite motif in the DNA-binding domains of transcriptional activators [46]. Together, these observations are consistent with a role of PPB1 as a bZIP-type transcription factor.

3.3. PPB1 binds to the SL RNA PSE in yeast

To confirm that PPB1 has PSE-specific binding activity, we performed several control experiments. First, we verified the ability of PPB1 to activate transcription of *his*3 and *lac*Z in the presence of the integrated PSE 'bait'. Thus, the S4 and T6 GAL4 fusion cDNA plasmids were re-transformed into pHISi-5 × PSE/YM4271 and pLacZi-5 × PSE/YM4271 containing the PSE pentamer integrated into the *his3* or the *lacZ* promoter regions, respectively. Both fusions reproducibly activated expression of both *his*3 and *lac*Z, verifying that activation was not due to second site mutants. Second, the plasmids were transformed into the pHISi/YM4271 and pLacZi/YM4271 parental yeast strains lacking the integrated PSE pentamer. When the former transformed yeast were plated on medium lacking histidine, no growth was observed (data not shown). Similarly, no β-galactosidase activity was observed in the transformed pLacZi/YM4271 strain (data not shown). Together, these results indicate that the PPB1 protein encoded by the S4 and T6 clones only activates transcription of the *his*3 or *lac*Z reporters when the PSE is present in the promoter site, implying a specific interaction between PPB1 and the PSE.

```
TTGAAGATCCCCAGGAGGTCGAGCCGTGTCAGCAAGAATCTTCACTTTATGCCATTGGTC
TTTCGGCCGCAAGGGAAATGTCAGGATGCGAAGAACACGTCTCTAGGGCTCGTGGAGGAT
TGACGGTTCGAGATATCGATACCCATTTGGCGTCTATTAGACAGAACGCCTCTGGATTCC
CCCGTTCACGGCATTCTTTCACATCGAGGAAAGGCTCTCGTTGAAAACTGCATTGCGAGT
ATGGCTCAAGAGGCGTTTCCTCGGCTTCGTTCAAAACACCTTCAGGCTCTTCTTCATGAG
M  A  Q  E  A  F  T  R  L  R  S  K  H  L  Q  A  L  L  H  E     20
TGCTGTGGGCTTTTGGCGTGTGGACGCGTAGCGCTTCGACTAGGATTTGCGGACATGTGT
C  C  G  L  L  A  C  G  R  V  A  L  R  L  G  F  A  D  M  C     40
AATGTGAGTGGGGAAATAAGTATGTGGAGAGAGCTGAATGTATTGACGGAACGGTTTAAC
N  V  S  G  E  I  S  M  W  R  E  L  N  V  L  T  E  R  F  N     60
ATTGCACGCCGCAAAGCAGCTGCACATTTGCAGCGCATTGAGAAGGGAGTGCCACAGCAG
I  A  R  R  K  A  A  A  H  L  Q  R  I  E  K  G  V  P  Q  Q     80
CGACGGTGGTATTGGCGGGGAGTTCTTCGTTCTGCCGCCGGGAGACTGAAACTTTTTCCT
R  R  W  Y  W  R  E  V  L  R  S  A  A  G  R  L  K  L  F  P    100
GTACACCAAGAAAATTTAATTCCACGGATGGAGTGGATACGGGGCTCCGTGTTTGCCTTT
V  H  Q  E  N  L  I  P  R  M  E  W  I  R  G  S  V  F  A  F    120
GTCGCCTTTATTGAAATGGCGGAGCAGTCCGGTGATGGGGTTCGAGTGGAGCCGTTGATC
V  A  F  I  E  M  A  E  Q  S  G  D  G  V  R  V  E  P  L  I    140
AGGAAGTTGTTTTGCAGTCAGTTGGGATCCTTCCTACACGTGTCTCAGATACTGGACAAT
R  K  L  F  C  S  K  L  G  S  F  L  H  V  S  Q  I  L  D  N    160
GCACAAAGACTGACGGAGGAACTTCTGCAATACAAGGCGACTTCTGAAAATGACAAGACG
A  Q  R  L  T  E  E  L  L  Q  Y  K  A  T  S  E  N  D  K  T    180
TCGTCCCCTCCAGGGATTTACAGTGATGTTCATGAGCGGCAGACATTAGCGAATTGCATT
S  S  T  T  G  I  Y  S  D  V  H  E  R  Q  T  L  A  N  C  I    200
AAGGAGGAACTGGAGAAACTTCGACACATGACCGAAGATGAATTGGAGGTGTTGTCCGAG
K  E  E  L  E  K  L  R  H  M  T  E  D  E  L  E  V  L  S  E    220
CGGACGCATCCGTTAAATACCCGTGCAGATGCTCGTCGGGATCAGTTTGCATCTAGCGGT
R  T  H  P  L  N  T  R  A  D  A  R  R  D  Q  F  A  S  S  G    240
GGGGCACATGCAACGCAAAGTCTCCGTGATGACATCTTTGATGATTACCGGGTCACTCGA
G  A  H  A  T  Q  S  L  R  D  D  I  F  D  D  Y  R  V  T  R    260
CATCACAAGGAGGCACCCCCTGCTGTGTTACATTGCATACGCCCTCAAAATGCCCTTAAA
H  H  K  E  A  P  P  A  V  L  H  C  I  R  P  Q  N  A  L  K    280
GAAGCTCAAATTATTCTTCGTTATGATCCGAATGTGCCCGAAGTGTTGTCTTCTGCACCG
E  A  Q  I  I  L  R  Y  D  P  N  V  P  E  V  L  S  S  A  P    300
ATCGACGTGGATCAGGTGGTTCGTCGCGTTACGGAGGTGCATGAAACCAATCCATATGCA
I  D  V  D  Q  V  V  R  R  V  T  E  V  H  E  T  N  P  Y  N    320
TCATTGCATTCCCGACAACAGTTTAGGCAGGTGGAGTACACGGTGTGTCGACTGTATGCG
S  L  H  S  R  Q  Q  F  R  Q  V  E  Y  T  V  C  R  L  Y  A    340
GGAACGCGTTGTATTGGCCAAGGCAAGGGCGAGACGTTGATGGAAGCCATGAATGAGGCG
G  T  R  C  I  G  Q  G  K  G  E  T  L  M  E  A  M  N  E  A    360
GCGCAACATACATTGCTGAACTATTACCTTAAAAAAGGCCCTAGCATCATCCCTGAGAAG
A  Q  H  P  L  L  N  Y  Y  L  K  K  G  P  S  I  I  P  E  K    380
GACTCAACAACTGTTTTTGAGGGAACATTTGTGGATGCGGAACGTACTGAACCATGGGGA
D  S  T  T  V  F  E  G  T  F  V  D  A  E  R  T  E  P  W  G    400
GTTTGTGTAAATAAATCATCCGATGTAGAAATATTTTTTTGATTTCATCTTCTCTCACAC
V  C  V  N  K  S  S  D  V  E  I  F  F  *                      413
AATTTCCGCGTCGGGTGGCCGGTGATTGCTAAGGCTACGGAGGATTCTGAAGACTCACTG
CGTGTGTGTATGTGTGTGGTTTTATGTTTTATGTACTGTTTTTGCTCTTCTTTTATTTTG
CAGATAATTGTCCGTGACGCGTGGGAAGACACGACGACTGCCGCACACACACACACACAC
AAATACGCACTGAAAACAAAATAAAAGGAATAAAACAATGGAGGGTCCGGTAACGGCGGT
```

Fig. 3. Sequence of the PPB1 gene and deduced amino acids. The genomic sequence of the PPB1 gene is shown. The underlined 'AG' dinucleotide 32 nt upstream from the ATG translation start point is the 3′ splice acceptor site for *trans*-splicing of the SL. The leucine-zipper like motif (bZIP) at amino acid positions 9–52 and the helix–turn–helix (HTH) motif at amino acid positions 52–71 are underlined; the 'leucines' of the putative zipper are bold. The bZIP-basic region at amino-acids positions 72–97 are indicated in bold type. HLH motifs at amino-acids positions 299–315 and 348–364 are underlined. The sequence data have been submitted to the GenBank/EMBL/DDBJ databases under accession number AF213249.

3.4. In vitro DNA-binding activity of PPB1

PPB1 was expressed in vitro in two ways as described in the Materials and methods. First, the protein was expressed directly from the plasmid using a reticulocyte lysate mediating coupled transcription and translation of genes inserted downstream from a bacteriophage T7 promoter. This

reticulocyte lysate expressed protein was generated in the presence of [^{35}S] methionine and revealed by SDS-polyacrylamide gel electrophoresis and autoradiography as a product of ~45 kDa (Fig. 4A). Second, the protein was over-expressed as a histidine-tagged fusion protein in *E. coli* (Fig. 4B), and purified as described in Materials and methods yielding a product of the expected ~45 kDa molecular mass (Fig. 4C).

Both the purified histidine-tagged PPB1 and the in vitro translated PPB1 proteins were used in gel shift experiments to demonstrate their abilities to specifically bind the *T. cruzi* SL RNA gene PSE (Fig. 5). The histidine-tagged PPB1 protein yielded a single major shifted band when incubated in the presence of the double stranded PSE-bearing oligo, whereas nonspecific competitor had no effect on the shifted band (Fig. 5A). In contrast, incubation of PPB1 in the presence of random oligos or oligos bearing sequences of other promoter elements yielded no shifted bands (data not shown). The shifted band from the in vitro translated PPB1 protein was strongly depleted in the presence of excess unlabeled PSE-bearing oligo competitor (Fig. 5B, lanes 3–6), whereas excess random oligos (not shown) or TATA-box containing oligos (Fig. 5B, lanes 7–8) had no effect on the shifted band. Thus, PPB1 specifically binds the *T. cruzi* PSE and is a strong candidate for a SL RNA gene-specific transcription factor.

3.5. Expression of PPB1 in T. cruzi enhances activity of the SL RNA gene promoter

To begin a functional analysis of PPB1, we constructed vectors pPATAN-PPB1 and pPATAN-GFP/PPB1 in which the protein could be overexpressed in *T. cruzi*. Both plasmids bear the selectable neomycin phosphotransferase gene

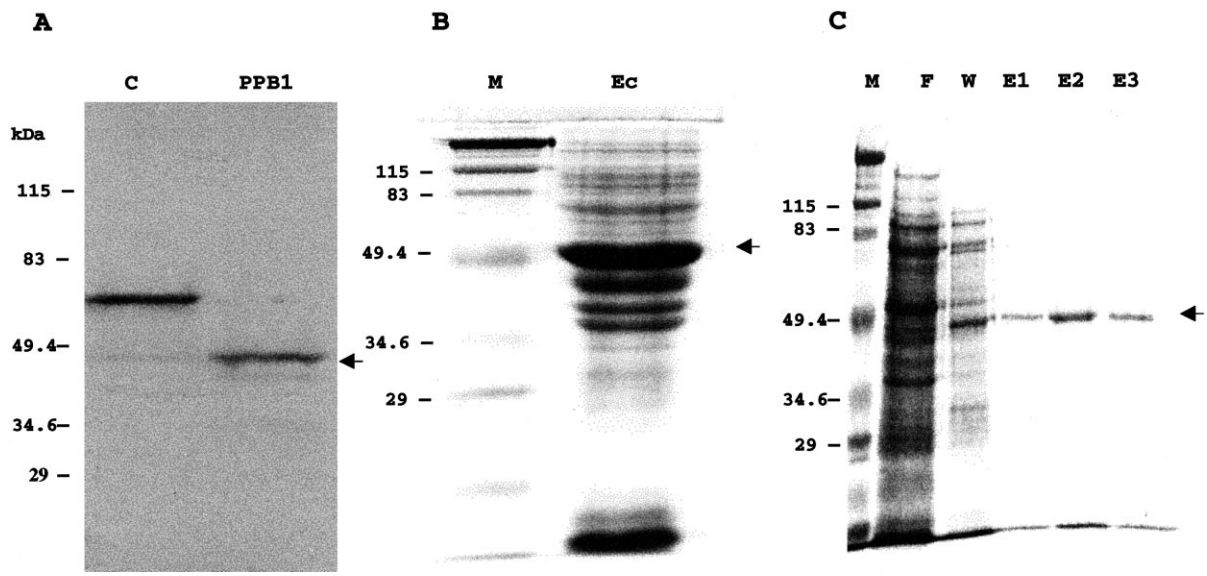

Fig. 4. In vitro translated and *E. coli* expressed PPB1 protein. (A) The PPB1 gene was cloned into pGEM-3Zf(+) vector and expressed in a coupled transcription/translation reticulocyte system in the presence of [^{35}S]methionine as described in the Materials and methods. The products were electrophoresed in an SDS-polyacrylamide gel and revealed by autoradiography. The lane marked PPB1 shows the in vitro translated ~45 kDa PPB1 protein (arrow). The lane marked C shows the 61 kDa luciferase translation control. (B) The PPB1 gene was cloned into vector pQE-30, transformed into *E. coli*, induced with IPTG, and electrophoresed in an SDS-polyacrylamide gel, and stained with Coomassie Blue as outlined in the Materials and methods. M indicates a lane containing molecular size standards. The lane marked Ec contains the unpurified PPB1 protein (arrow) overexpressed in the crude *E. coli* extract. (C) The his-tagged fusion protein was purified from *E. coli* lysates by chromatography on an Ni-NTA column as described in the Materials and methods. Lanes: M, molecular size standards; F, flow-through fractions; W, wash fractions; E1, elution fractions 1; E2, elution fractions 2; E3, elution fractions 3. The arrow indicates the purified fusion protein.

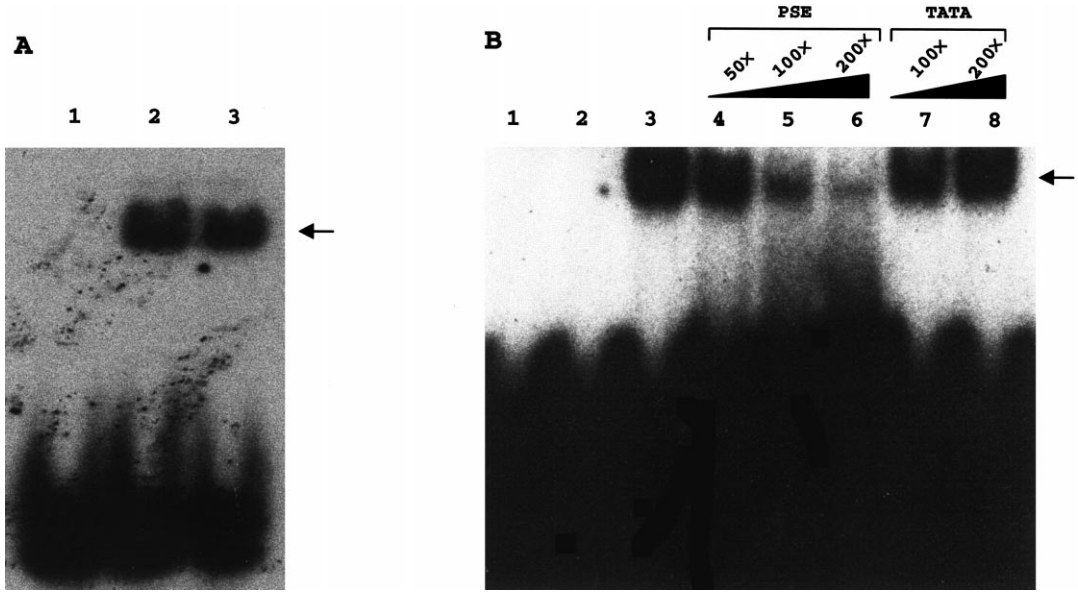

Fig. 5. EMSA of recombinant PPB1 protein. EMSA was performed using PPB1 expressed in reticulocyte lysates or purified from *E. coli* as described in the Materials and methods. (A) EMSA with His-tagged PPB1 purified from *E. coli* and double stranded PSE bearing oligo p55/p55′. Lanes: 1, free [5′-^{32}P] p55/p55′; 2, labeled p55/p55′ incubated with 0.5 μg of His-tagged PPB1; 3, same as lane 2 but with 100 fold excess of unlabeled double stranded TATA oligo. The arrow indicates the gel shift. (B) EMSA with in vitro translated PPB1 and double stranded PSE bearing oligo p55/p55′. Lanes: 1, free [5′-^{32}P] p55/p55′; 2, labeled p55/p55′ incubated with reticulocyte lysate protein without added plasmid (negative control); 3, labeled p55/p55′ incubated with reticulocyte lysate protein translated from pGEM-PPB1; 4, same as 3 plus a 50 fold excess of cold p55/p55′; 5, same as 3 plus a 100 fold excess of p55/p55′; 6, same as 3 plus a 200 fold excess of cold p55/p55′; 7, same as 3 but with a 100 fold excess of the double stranded TATA oligo instead of p55/p55′; 8, same as 7 but with a 200 fold excess of the double stranded TATA oligo. The arrow indicates the position of the gel shifted band.

(*neo*^{r}). pPATAN-PPB1 also expresses native PPB1, and pPATAN-GFP/PPB1 expresses PPB1 fused to the 3′ terminus of the green fluorescent protein gene. Expression from both plasmids is driven from the *T. cruzi* ribosomal RNA (rRNA) gene promoter [47]. These plasmids were transfected into *T. cruzi* epimastigotes and stable transfectants were selected in the presence of 500 μg/ml G418 (Sigma). After at least 3 weeks of continuous growth and selection, the strain transfected with pPATAN-GFP/PPB1 was shown by fluorescence microscopy to express the GFP/PPB1 fusion gene in 100% of the cells, indicating efficient expression of both *neo*^{r} and GFP/PPB1 in the transfected trypanosomes (data not shown). In these trypanosomes, GFP was concentrated in the nucleus although signal was also detected throughout the cytoplasm. Northern blots indicated that the level of PPB1 mRNA in epimastigotes transfected with pPATAN-PPB1 was ~3 fold higher than in non-transfected epimastigotes (not shown).

To examine the impact of overexpression of PPB1 on expression from the SL RNA gene promoter, levels of endogenous SL RNA in the *T. cruzi* bearing pPATAN-PPB1 were examined by Northern Hybridization (Fig. 6). These experiments showed that endogenous SL RNA was upregulated greater than 10 fold compared to non-transfected *T. cruzi* (compare lanes 1 to lanes 2). These pPATAN-PPB1-transfected *T. cruzi* were also super-transfected with 100 μg of DNA from the Δ3 vector that expresses CAT from the PSE-bearing SL RNA gene promoter [47] to examine the effects on a protein coding gene. All transient transfections and CAT assays were performed multiple times with each data point represented in triplicate to ensure the significance of

the results, essentially as previously described [47,48]. The results showed an ~4-fold increase in CAT activity in the presence of episomally expressed PPB1 from pPATAN-PPB1 vector (strain *T. cruzi* CL/ pPATAN-PPB1) compared with that in the absence of episomally expressed PPB1 from Δ3 (Fig. 7). In contrast, the non-PSE negative control (Δ3PSE$^-$), the non-PPB1/CAT negative control (pPATAN) and the non-PPB1 overexpressed negative control (*T. cruzi* CL/ pPATAN) exhibited no increase of CAT expression (Fig. 7). These results are consistent with PPB1 playing the role of a DNA binding transcription factor.

4. Discussion

4.1. cDNA library normalization

The mRNAs of a typical mammalian cell are distributed in three frequency classes: (i) abundant, consisting of about 10–15 mRNAs that together represent 10–20% of the total mRNA; (ii) intermediate, consisting of 1000–2000 mRNAs that comprise 40–45% of the total; and (iii) rare, consisting of 15 000–20 000 or more mRNAs that comprise the remainder [49,50]. Protozoa are expected to have a similar distribution of mRNAs, although the overall number may be smaller. Thus, in a typical *T. cruzi* epimastigote, the mRNAs of the abundant and intermediate frequency classes may comprise as much as 50–65% of the total mRNA mass, but still represent an absolute minority of the total number of discrete mRNAs. Most known transcription factors seem to be translated from mRNAs in the rare class [31]. Thus, cloning of transcription factor genes is often impeded by the presence mRNAs from other nucleic acid binding proteins; e.g. ribosomal proteins or histones, which are among the abundant and intermediate classes of mRNAs. In fact, we have experienced this problem in attempts to isolate SL RNA gene PSE binding transcription factors using the one-hybrid system with non-normalized cDNA libraries. We have selected many false positive clones that were later shown to be ribosomal proteins or histones (data not shown). To obviate this problem, we applied a normalization procedure to a cDNA library from *T. cruzi* epimastigotes and used it with the yeast one-hybrid system to select potential SL RNA gene PSE binding proteins. Using this approach, none of the eleven clones initially selected were in the known abundant class of mRNAs. Thus, normalization of our library was probably essential to the success of this project.

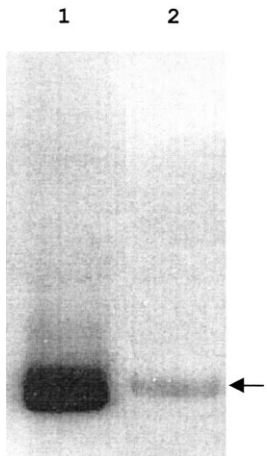

Fig. 6. Induction of the SL RNA in *T. cruzi* overexpressing PPB1. Total RNA was isolated as previously described [4] from *T. cruzi* epimastigotes transfected with pPATAN-PPB1 and selected in 500 μg/ml neomycin as described in the Materials and methods. Approximately 10 ug of total RNA from transfected (lanes 1) or non-transfected (lanes 2) *T. cruzi* was electrophoresed in a 8% polyacrylamide urea gel, electroblotted to nylon filters, and hybridized to 5′-[γ-^{32}P]labeled oligonucleotide SL33 (5′-GGGTTCGTGGACCCCGGTCAAAAGAAATTGACCAAACATAGCCGCCAGAGGCGGGTT-3′) essentially as previously described [4].

Our normalization process varied from previously published protocols [33,38,39,51–53] at several critical points. First, we directly synthesized the initial cDNA using reverse transcription instead of PCR [54] to avoid mutations introduced by thermostable polymerases. Second, we elected to perform directional instead of non-oriented cloning to enhance the probability of obtaining an in frame gene fusion. Third, we normalized the library by hybridization with single stranded PCR products derived from amplification of the library itself. Finally, the normalized library was generated as single stranded phagemid DNA, and this

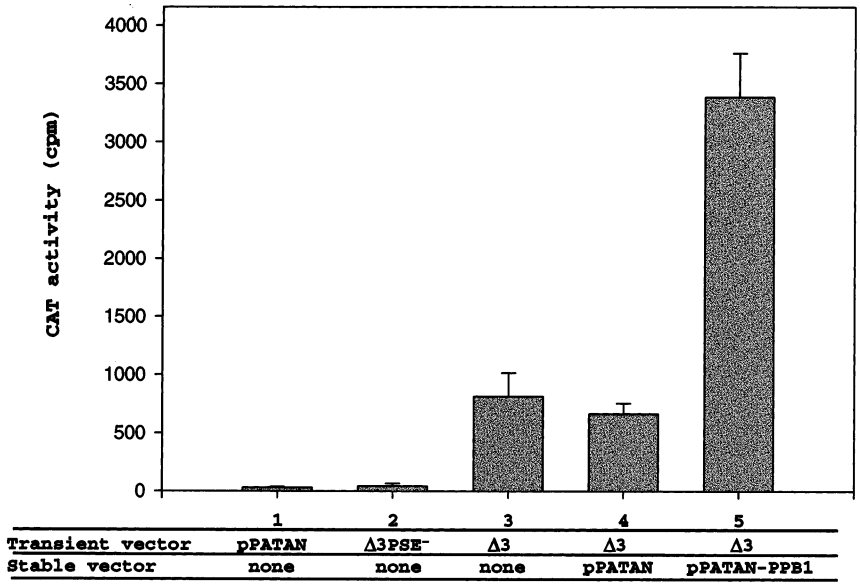

Fig. 7. The *in vivo* transcriptional activity of PPB1. *T. cruzi* epimastigotes bearing various endogenous constructs were super transfected with plasmid ┤3 which bears the bacterial CAT gene under control of the PSE-bearing *T. cruzi* SL RNA gene promoter to measure the effects of overexpression of PPB1 on activity of the SL RNA gene promoter. 1, plasmid pPATAN alone transfected into wild type *T. cruzi* epimastigotes; 2, plasmid ┤3 PSE$^-$ lacking the PSE was transfected into wild type *T. cruzi* epimastigotes; 3, plasmid ┤3 transfected into wild type *T. cruzi* epimastigotes; 4, plasmid ┤3 transfected into *T. cruzi* epimastigotes stably transfected with plasmid pPATAN; 5, plasmid ┤3 transfected into *T. cruzi* epimastigotes stably transfected with pPATAN-PPB1. CAT assays were performed essentially as previously described [45].

single stranded DNA was converted into double stranded plasmid by multiple primer extensions to enhance the transformation efficiency approximately a hundred folds. These modifications generated a high quality partially normalized cDNA library that we exploited to isolate *T. cruzi* SL RNA gene PSE binding proteins.

4.2. One-hybrid strategy to clone T. cruzi transcription factor gene

The only *T. cruzi* promoters that have been partially characterized are those of the ribosomal RNA operon [47] and the spliced leader genes [6,48]. To our knowledge, none of the cognate transcription factors of these promoters have been characterized. Similarly, few transcription promoters have been identified in other kinetoplastid protozoa, and no transcription factors have yet been characterized. In spite of significant effort, we have been unable to clone even the supposedly ubiquitous TATA binding protein from any kinetoplastid protozoa (unpublished data). Thus, herein we attempted to exploit the powerful yeast one-hybrid system to clone transcription factors that bind to and activate the SL RNA gene promoter from *T. cruzi*. This system has been successfully exploited to identify many other eukaryotic transcription factors [20–30], and therefore provided a viable alternative to traditional biochemical approaches for identifying specific transcription factors [10]. In contrast to biochemical methods, selection and screening using the yeast one-hybrid system with the SL RNA gene PSE as bait does not require large amounts of starting material.

Using this system, we selected a putative SL RNA gene PSE binding protein, PPB1, for initial study. Other selected clones that have not yet been fully characterized may also represent PSE binding proteins, as it is common that multiple proteins bind to an eukaryotic promoter element

[55]. In fact, such is the case for the PSEs of known eukaryotic processor RNAs [56]. Since the SL RNA fits most criteria for a processor RNA, we were not surprised that multiple proteins were identified in the one-hybrid screen. PPB1 is most likely one of the primary factors that bind to the SL RNA gene PSE. Multiple independent clones of PPB1 were selected on his$^-$ medium and verified using β-galactosidase screening. Both the in vitro transcribed and translated protein and the recombinant histidine-tagged protein isolated from E. coli showed strong PSE binding. This binding was specific for PSE-bearing oligonucleotides and was not competed by large excesses of other oligonucleotides.

4.3. In vivo function of PPB1

The SL RNA gene promoter in *T. cruzi* is apparently a weak promoter, at least compared to the 50–100 fold stronger rRNA gene promoter [47,48]. The reason for this weakness is not known, but may be related to the level of expression of gene-specific transcription factors. The SL RNA gene is present in the genomes of trypanosomatids in blocks of 50–100 tandem repeats, each one bearing an identical PSE [57]. Each of these would be expected to compete for transcription factors, and the level of available factor may therefore be an important rate-limiting parameter in SL RNA gene transcription. The observation that only a 3–5 fold increase in PPB1 mRNA in trypanosomes transfected with constructs in which PPB1 gene transcription is driven by the rRNA promoter results in a greater than 10 fold increase in SL RNA gene transcription supports this interpretation. Currently, we are attempting to further these investigations by identification of the active domains of PPB1 and by generation of stable PPB1 gene knockouts in *T. cruzi*.

PPB1 is $\sim$45 kDa. Recently, a similar sized protein ($\sim$46 kDa) was reported by Luo and Bellofato [10] as a member of PBP-1 putative transcription factor that binds to and activates [11] the promoter of the distantly related kinetoplast *L. seymouri*. It is interesting to speculate that $\sim$45 kDa PPB1 of *T. cruzi* could be homologous to and play a similar role as the $\sim$46 kDa protein of PBP-1 in *L. seymouri*. The latter protein seems to be a necessary subunit of PBP-1 and is required for activation of transcription in vitro [11]. These observations are consistent with the hypothesis that PPB1 in *T. cruzi* and PBP-1 in *L. seymouri* are transcription factors mediating primary recognition of the SL RNA gene promoter elements in these two kinetoplastid protozoa.

Finally, by analogy with other eukaryotic systems and as described for *L. seymouri* [10,11], we anticipate that PPB1 is also a member of a complex of transcription factors required for activation of the SL RNA gene. Clearly, some of these factors will bind the PSE and some will bind to PPB1 and other proteins of the complex. Thus, we are currently examining other cDNA clones selected in the one-hybrid screening for transcription factor-like activity, and are using PPB1 in two-hybrid screens for selection of proteins with which it interacts.

Acknowledgements

We thank the MCV-VCU DNA Core Laboratory for all oligonucleotide synthesis and DNA sequencing in this study. And we thank Tadeusz Zwierzynski, Yingping Wang, Xu Wang and Myrna Serrano for discussions and assistance in performance of the work. This work was supported by grants from National Institute of Allergy and Infectious Disease at National Institute of Health, National Chapter of the American Heart Association, the Virginia Thoracic Society, the American Lung Association, and the Jeffress Memorial Trust.

References

[1] Nilsen TW. Trans-splicing in protozoa and helminths. Infect Agents Dis 1992;1:212–8.
[2] Agami R, Aly R, Halman S, Sapira M. Functional analysis of cis-acting DNA elements required for expression of the SL RNA gene in the parasitic protozoan *Leishmania amazonensis*. Nucleic Acids Res 1994;22:1959–65.
[3] Campbell DA, Sturm NR, Yu MC. Transcription of the kinetoplastid spliced leader RNA gene. Parasitol Today 2000;16:78–82.

[4] McCarthy Burke C, Taylor ZA, Buck GA. Characterization of the spliced leader genes and transcripts in *Trypanosoma cruzi*. Gene 1989;82:177–89.
[5] Saito RM, Elgort MG, Campbell DA. A conserved upstream element is essential for transcription of the *Leishmania tarentolae* mini-exon gene. EMBO J 1994;13:5460–9.
[6] Nunes LR, Carvalho MR, Shakarian AM, Buck GA. The transcription promoter of the spliced leader gene from *Trypanosoma cruzi*. Gene 1997;188:157–68.
[7] Gunzl A, Ullu E, Dorner M, Fragoso SP, et al. Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. Mol Biochem Parasitol 1997;85:67–76.
[8] Yu MC, Sturm NR, Saito RM, Roberts TG, Campbell DA. Single nucleotide resolution of promoter activity and protein binding for the *Leishmania tarentolae* spliced leader RNA gene. Mol Biochem Parasitol 1998;94:265–81.
[9] Luo H, Gilinger G, Mukherjee D, Bellofatto V. 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 1999;274:31947–54.
[10] Luo H, Bellofatto V. Characterization of two protein activities that interact at the promoter of the trypanosomaid spliced leader RNA. J Bio Chem 1997;272:33344–52.
[11] Luo H, Gilinger G, Mukherjee D, Bellofatto V. 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 Bio Chem 1999;5(274):31947–54.
[12] Inouye C, Remondelli P, Karin M, Elledge S. Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: the one hybrid system. DNA Cell Biol 1994;13(7):731–42.
[13] Difilippantoio MJ, McMahan CJ, Eastman QM, Spanopoulou E, Schatz DG. RAG1 mediate signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell 1996;87:253–62.
[14] Li JJ, Herskowitz I. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science 1993;262:1870–4.
[15] Gstaiger M, Hovens C, Georgiev O, Knoepfer L, Schaffner W. BZLF1 (ZEBRA, Zta) protein of Epstein–Barr virus selected in a yeast one-hybrid system by binding to a consensus site in the IgH intronic enhancer: a role in immunoglobulin expression? Bio Chem 1996;377:669–73.
[16] Dowell SJ, Romanowski P, Diffley JF. Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. Science 1994;265:1243–6.
[17] Durner-Gailus V, Chintamaneni C, Wilson R, Brill SJ, Vershon AK. Analysis of a meiosis specific URS1 site: sequence requirements and involvement of replication protein A. Mol Cell Biol 1997;17:3536–46.
[18] Bukenberger M, Horn J, Dingermann T, Dottin RP, Winckler T. Molecular cloning of a cDNA encoding the nucleosome core histone H3 from *Dictyostelium discoideum* by genetic screening in yeast. Biochim Biophys Acta 1997;1352:85–90.
[19] Sieweke MH, Tekotte H, Frempton J, Graf T. MafB is an interaction partner and repressor of Ets-1 that inhibits Erythroid differentiation. Cell 1996;85:49–60.
[20] Blaiseau PL, Isnard AD, Surdin-Kerjan Y, Thomas D. Met31p and Met32p, two related zinc finger proteins, are involved in transcriptional regulation of yeast sulfur amino acid metabolism. Mol Cell Biol 1997;17(7):3640–8.
[21] Lemercier C, To RQ, Swanson BJ, Lyons GE, Konieczny SF. Mist1: a novel basic helix-loop helix transcription factor exhibits a developmentally regulated expression pattern. Dev Biol 1997;182:101–13.
[22] Hasegawa T, Takeuchi A, Miyaishi O, Isobe K, Crombrugghe B. Cloning and characterization of a transcription factor that binds to the proximal promoters of the two mouse type I collagen gene. J Bio Chem 1997;272(8):4915–23.
[23] Ulmasov T, Hagen G, Guilfoyle TJ. ARF1, a transcription factor that binds to auxin response elements. Science 1997;276:1865–8.
[24] Wang M, Reed RR. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature 1993;364:121–6.
[25] Mak KL, Longcor LC, Johnson SE, Lemercier C, To RQ, Konieczny SF. Examination of mammalian basic helix–loop–helix transcription factors using a yeast one-hybrid system. DNA Cell Bio 1996;15(1):1–8.
[26] Sugawara T, Kiriakidou M, McAllister JM, Kallen CB, Strauss JF. Multiple steroidogenic factor 1 binding elements in the human steroidogenic acute regulatory protein gene 5′-flanking region are required for maximal promoter activity and cyclic AMP responsiveness. Biochemistry 1997;36:7249–55.
[27] Yang HY, Evans T. Homotypic interactions of chicken GATA-1 can mediate transcriptional activation. Mol Cell Biol 1995;15(3):1353–63.
[28] Ogino H, Yasuda K. Induction of lens differentiation by activation of a bZip transcription factor, L-Maf. Science 1998;280:115–8.
[29] Gstaiger M, Knoepfel L, Georglev O, Schaffner W, Hovens CM. A B-cell coactivator of octamer binding transcription factors. Nature 1995;373:360–2.
[30] Tsang AP, Visvader JE, Turner CA, et al. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell 1997;90:109–19.
[31] Pabo CO, Sauer RT. Transcription factors: structural families and principles of DNA recognition. Annu Rev Biochem 1992;61:1053–95.
[32] Gudkov AV, Kazarov AR, Thimmapaya R, Axenovich SA, Mazo IA, Rononson IB. Cloning mammalian genes by expression selection of genetic suppressor elements: association of kinesin with drug resistance and cell immortalization. Proc Natl Acad Sci USA 1994;91:3744–8.

[33] Tanaka T, Ogiwara A, Uchiyama I, Takagi T, Yazaki Y, Nakamura Y. Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing. Genomics 1996;35:231–5.
[34] Zwierzynski TA, Widmer G, Buck GA. In vitro 3′ end processing and poly(A) tailing of RNA in *Trypanosoma cruzi*. Nucleic Acids Res 1989;17:4647–60.
[35] Swindle J, Ajioka J, Eisen H, Sanwal B, Jacquemot C, Browder Z, Buck G. The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. EMBO J 1988;7:1121–7.
[36] Auffray C, Rougeon F. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur J Biochem 1980;107:303–14.
[37] Maniatis T, Fritsch EF, Sambrook J. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989.
[38] Soares MB, Bonaldo MF, Jelene P, Su L, Lawton L, Efstratiadis A. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci USA 1994;91:9228–32.
[39] Bonaldo MF, Lennon G, Soares MB. Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res 1996;6:791–806.
[40] Gruber CE, Li WB. An improved subtractive hybridization method using phagemid vectors. In: Griffin AM, Griffin HG, editors. Molecular Biology: Current Innovations and Future Trends. England: Horizon Scientific Press, 1995:93–106.
[41] Fragoso SP, Goldenberg S. Cloning and characterization of the gene encoding *Trypanosoma cruzi* DNA topoisomerase II. Mol Biochem Parasitol 1992;55:127–34.
[42] Dietrich P, Soares MB, Affonso MHT, Floeter-Winter LM. The *Trypanosoma cruzi* ribosomal RNA-encoding gene: analysis of promoter and upstream intergenic spacer sequences. Gene 1993;125:103–7.
[43] Stephen FA, Madden TL, Schaffer AA, et al. Dapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402.
[44] Bornberg-Bauer E, Rivals E, Vingron M. Computational approaches to identify leucine zippers. Nucleic Acids Res 1998;26:2740–6.
[45] Baxevanis AD, Vinson CR. Interactions of coiled coils in transcription factors: where is the specificity? Curr Opin Genet Dev 1993;3:278–85.
[46] Hurst HC. Transcription factors 1: bZIP proteins. Protein Profile 1995;2:101–68.
[47] Tyler-Cross RE, Short SL, Floeter-Winter LM, Buck GA. Transient expression mediated by the *Trypanosoma cruzi* rRNA promoter. Mol Biochem Parasitol 1995;72:23–31.
[48] Lu HY, Buck GA. Expression of an exogenous gene in *Trypanosoma cruzi* epimastigotes. Mol Biochem Parasitol 1991;44:109–14.
[49] Bioshop JO, Morton JG, Rosbash M, Richardson M. Three abundance classes in Hela cell messenger RNA. Nature 1974;250:199–204.
[50] Davidson EH, Britten RJ. Regulation of gene expression: possible role of repetitive sequences. Science 1979;204:1052–9.
[51] Patanijali SR, Parimoo S, Weissman SM. Construction of a uniform-abundance (normalized) cDNA library. Proc Natl Acad Sci USA 1991;88:1943–7.
[52] Sasaki YF, Ayusawa D, Oishi M. Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. Nucleic Acids Res 1994;22:987–92.
[53] Gudkov AV, Kazarov AR, Thimmapaya R, Axenovich SA, Mazo IA, Rononson IB. Cloning mammalian genes by expression selection of genetic suppressor elements: association of kinesin with drug resistance and cell immortalization. Proc Natl Acad Sci USA 1994;91:3744–8.
[54] Neto ED, Harrop R, Oliveria RC, Wilson RA, Pena SD, Simpson AJ. Minilibraries constructed from cDNA generated by arbitrarily primed RT-PCR: an alternative to normalized libraries for the generation of ESTs from nanogram quantities of mRNA. Gene 1997;186:135–42.
[55] Reinberg D, Orphanides G, Ebright R, Akoulitchev S, Carcamo J, Cho H, et al. The RNA polymerase II general transcription factors: past, present, and future. Cold Spring Harbor Symp Quant Biol 1998;63:83–103.
[56] Smale ST, Jain A, Kaufmann J, Emami KH, Lo K, Garraway IP. The initiator element: a paradigm for core promoter hereogeneity within metazoan protein-coding genes. Cold Spring Harbor Symp Quant Biol 1998;63:21–31.
[57] Nunes LR, Teixeira MMG, Camargo EP, Buck GA. Sequence and structural characterization of the spliced leader genes and transcripts in *Phytomonas*. Mol Biochem Parasitol 1995;74:233–7.