Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 309 (2003) 183–188
BBRC
www.elsevier.com/locate/ybbrc

A novel type of single-stranded nucleic acid binding protein recognizing a highly frequent motif in the intergenic regions of *Trypanosoma cruzi*

María Ana Duhagon,[a] Bruno Dallagiovanna,[a] Martín Ciganda,[a] William Ruyechan,[b] Noreen Williams,[b] and Beatriz Garat[a,*]

[a] *Laboratorio de Interacciones Moleculares, Facultad de Ciencias, Universidad de la República, Montevideo 11400, Uruguay*
[b] *Department of Microbiology, 251 Biomedical Research Building, State University of New York at Buffalo, Buffalo, NY 14214, USA*

Received 1 August 2003

Abstract

Regulation of gene expression in trypanosomatids is not yet well understood. Genes are organized in long polycistronic transcriptional units separated by intergenic regions that may contain the signaling information for nucleic acid processing. Polydinucleotides are frequent in these regions and have been proposed to be involved in regulation of gene expression. Previously, we have reported that [dT–dG] are highly frequent, asymmetrically strand distributed, and constitute targets for specific protein binding [Biochem. Biophys. Res. Commun. 287 (2001) 98]. Here, we present the purification and characterization of a new type of single stranded nucleic acid binding protein (Tc38) that recognizes specifically the motif poly[dT–dG] in this parasite. The protein has a deduced molecular weight of 38 kDa and its salient characteristics include an isoelectric point of 9.34, a high frequency of Ser, Leu, and di-amino acids. Neither compositional nor architectural conserved domains could be detected in database searches. Recombinant Tc38 was expressed as a GST fusion protein, purified, and used to analyze target specificity by electrophoretic mobility shift assays. The unusual characteristics of the protein together with the peculiar features of the specific nucleic acid target suggest the existence of a novel event that may be involved in the mechanisms of gene expression in trypanosomatids.

Keywords: Trypanosomatids; DNA–protein interactions; Dinucleotide repeats; Gene expression; Single-stranded DNA binding proteins; Poly[dT–dG]

Trypanosoma cruzi is a protozoan responsible for Chagas' disease. It constitutes a major health problem affecting several million people. Its endemic area extends throughout South and Central America and Mexico. The parasite has a life-cycle that includes a vertebrate and an insect vector, with at least three different morphological and functional forms.

Molecular studies have shown several peculiarities in the regulatory mechanisms of gene expression in trypanosomatids. Protein-coding genes are arranged in large polycistronic transcriptional units from which individual mRNAs are generated by 5′ *trans*-splicing, involving the addition of a small conserved RNA called spliced leader (SL) and 3′ polyadenylation. Trypanosome genes typically lack introns, though a single exception for *Trypanosoma brucei* and *T. cruzi* has recently been described [1]. In *Leishmania major* chromosome 1 and in the longest reported *T. cruzi* contig, all protein coding genes are primarily organized in long divergent transcriptional units [2,3]. Mechanisms of transcription initiation for protein coding genes are not yet completely understood. No typical RNA polymerase II promoter for this type of gene has been detected. In addition, intergenic regions containing non-canonical promoter sequences can drive transcription [4]. Though the switch region of *L. major* chromosome 1 has been considered a potential putative promoter, recently it was reported that it is not required for the expression of all genes present in the transcriptional units [5]. It has been

* Corresponding author. Fax: +598-2-525-86-17.
E-mail address: bgarat@fcien.edu.uy (B. Garat).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0006-291X(03)01561-4

suggested that RNA polymerase II initiation has very low specificity and that transcription could result from initiation at many sites along the polycistronic unit through transient DNA melting, which could allow RNA polymerase II assembly [6]. In this context, single-stranded DNA binding proteins should be relevant factors in the molecular mechanisms governing transcription. Since individual genes belonging to a single polycistronic unit show differential or stage specific expression, regulation of gene expression is considered to occur mainly at post-transcriptional stages [7]. 3′ UTR sequences have been shown to confer stage specificity and protein interactions with targets therein have been identified [8,9].

Previously, we have reported that in *T. cruzi*, [dT–dG] repeats are highly frequent and show asymmetric strand distribution in the vicinity of coding sequences. We have shown that poly[dT–dG] and its complementary sequence constitute specific targets for nuclear proteins [10]. Here we present the purification of one such protein. Data obtained by the microsequence analysis of the purified protein enabled the cloning and sequencing of the complete gene using RT-PCR and library screening. Finally, the ability of the recombinant protein to recognize the repetitive target was verified by competition with other repeating dinucleotides. The identification of a new type of single-stranded nucleic acid binding protein in the context of the unusual features in frequency, location, and strand distribution of the specific target, certainly constitutes a first step towards the comprehension of trypanosomatids genomics.

Materials and methods

Cell growth. The *T. cruzi* CL Brener strain [11] was used. Epimastigotes were cultured in liver infusion tryptose (LIT) liquid medium supplemented with 10% heat inactivated fetal calf serum (Sigma) at 28 °C.

Protein purification. A nuclear enriched fraction prepared as reported previously [10] was homogenized and cleared from nucleic acids by high salt extractions and dialyzed overnight at 4 °C against 150 mM KCl, 50 mM Tris–HCl, pH 7.5, 5 mM EDTA, and 0.5 mM DTT [12]. The [dT–dG]-binding activity contained in the clarified fraction (usually, 10 mg of protein) was further purified by two affinity chromatographic steps. The first was based on the protein's affinity for single stranded DNA agarose (Gibco-BRL) and the other based on the specific DNA target (poly[dT–dG]$_{15}$–biotin)–(streptavidin–agarose) (Sigma). For the latter, the binding buffer was the same as described below for EMSA. The columns were blocked using BSA and salmon sperm DNA. In both cases, the elution was performed using a step gradient with increasing KCl concentrations. The relative protein concentration was determined by absorbance at 280 nm measurements with an UV monitor. The fractions were extensively dialyzed against 150 mM KCl buffer and later concentrated with Centricon-10 concentrators to a final volume of approximately 50 μl. Electrophoretic mobility shift assay (EMSA) was used to determine binding activities of the fractions.

Binding reaction, electrophoretic mobility shift assays, and Southwestern blot. Binding reactions and electrophoretic mobility shift assays were performed as previously described [10] using oligodeoxynucleotides [dT–dG]$_{15}$ or [dT–dG]$_{20}$ as a probe. For competition MP1 (CGCCGGGCGTGTGCTGCCAAACCGCTGACC), poly[dT]$_{15}$, Sigma poly[dA–dT], poly[dC–dG], poly[dI–dC], and poly[G–U] were used. Southwestern blots were done as reported by Pitula et al. [12].

Library screening. Based on the sequence of the EST TENU2471, two primers MAD1 (5′-GCCCGCAGCAGTCCAAGTAATG-3′) and MAD2 (5′-TTCCTCGATCCGCCAGAGACTG-3′) were designed to obtain a specific probe. Genomic DNA from *T. cruzi* epimastigotes was prepared by phenol extraction [13] and was used to specifically amplify the partial sequence by PCR (cycle conditions: 94 °C-30 s, 50 °C-60 s, and 72 °C-60 s). The amplified product was cloned into pGEM-T-easy vector (Promega) and sequenced. The probe was labeled by PCR in the presence of [α-^{32}P]dCTP.

A λEMBL3 genomic library from *T. cruzi* (kindly provided by Dr. Samuel Goldenberg, IBMP, Brazil) was screened with the MAD1–MAD2 labeled probe. Positive clones were purified and sequenced using an ABI Prism 377 sequencer. The sequence reported here has been submitted to GenBank (AY357259).

RT-PCR. Total RNA was obtained by standard methods and used for RT-PCR assays. cDNA was synthesized using either the MAD2 or poly[dT]-Anchor (5′-CTCTCAAGGATCTTACCGCTTTTTTTTTT TTTTTTT-3′) primers. PCR amplification was performed using either a spliced leader based primer (5′-CGCTATTATTGATACAG TTTCTG-3′) or MAD1 and the primers used for the first strand. Two other primers MAD3 (5′-TAACGTGGAACAGCTTGAAGATC-3′) and MAD4 (5′-TTGTCCAGCAGCAACTGAGTCGG-3′) were also used for RT-PCR. Bands were excised from the gel, cloned into pGEM-T-easy vector (Promega), and sequenced.

Recombinant protein expression and purification. The complete coding sequence was amplified by PCR and cloned into the pGEX-4T vector (Amersham), generating a glutathione-*S*-transferase (GST) fusion and transformed in *Escherichia coli* BL21 (Novagen). Cultures were induced with IPTG for 4 h at 37 °C. The recombinant protein was purified using GST–SepharoseFF columns (Amersham–Pharmacia-Biotech) following the manufacturer's instructions.

For Western blot analysis, samples were fractionated on SDS–PAGE and transferred to Protran nitrocellulose membranes (Schleicher and Schuell) as indicated. Membranes were probed with an anti-GST polyclonal antibody (Amersham–Pharmacia-Biotech) and developed using a Goat ExtrAvidin Peroxidase Staining Kit (Sigma).

Results and discussion

Purification of a single-stranded nucleic acid binding protein from T. cruzi epimastigotes

In order to purify proteins that specifically recognize the repetitive dinucleotide [dT–dG] contained in the enriched nuclear fraction, a combination of high salt extraction and two steps of affinity chromatography was used. The activity was followed by EMSA using the target as a probe.

First, the enriched nuclear fraction was applied to a single stranded DNA–agarose column. Elution was performed using a step gradient with increasing KCl concentrations ranging from 0.2 to 2.0 M. Binding activity was observed in the 500 mM KCl eluates in several independent experiments. Interestingly, binding activity for both probes: poly[dC–dA]$_{15}$ and [dT–dG]$_{15}$, co-eluted in the same fraction, suggesting that the same

protein(s) could be responsible for the recognition of both probes (Figs. 1A and B). Binding activity recovery was improved when the ionic strength of the initial binding buffer was lowered to 50 mM KCl (data not shown). In the 500 mM KCl fraction, a prominent band of approximately 40 kDa was observed (Fig. 1C). To determine whether the binding activity correlated with this observed band, a southwestern blot was performed. One strong signal co-migrating with the purified protein was found using both the poly[dT–dG]$_{15}$ (Fig. 1D) and poly[dC–dA]$_{15}$ targets (not shown). A weaker and faster migrating signal was also observed. Second, a poly[dT–dG]$_{15}$–biotin–streptavidin–agarose affinity chromatography was used. In this case, binding activity was recovered both in the 0.5 and 1 M KCl fraction, suggesting a higher affinity for the target sequences (Fig. 1E). A weak single protein band of approximately 38 kDa was observed in silver stained SDS–PAGE of this fraction (not shown).

This band was excised from the gel and analyzed by mass spectrometry following trypsin digestion (Borealis Biosciences). The fragmentation pattern correlated with the expected tryptic peptide masses from the TENU2471 *T. cruzi* EST (dbEST Id: 1805863). This EST showed no homology with characterized proteins or conserved protein domains. However, an orthologous gene in *L. major* (NCBI Accession No. CAB71224) and in *T. brucei* (TIGR_67_26A17) was found in the databases.

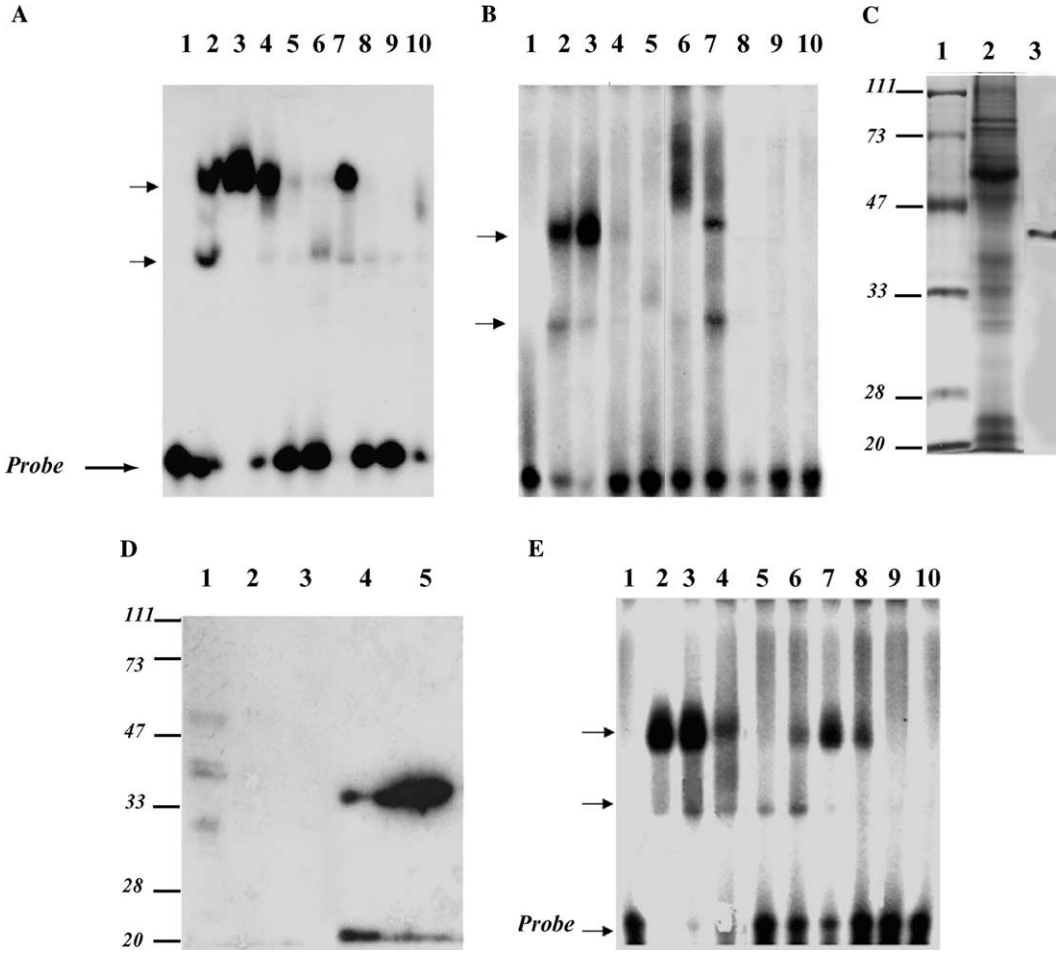

Fig. 1. Purification analysis. (A) Poly[dC–dA]$_{15}$ was radiolabeled and used as a probe for EMSA with protein fractions of the ss-DNA–agarose affinity chromatography. (B) Same as (A) using radiolabeled poly[dT–dG]$_{15}$ as a probe. In both cases: lane 1, free probe; lane 2, enriched nuclear fraction before high salt extraction; lane 3, fraction applied to the column; lane 4, fraction not bound to the column (loading conditions 150 mM KCl); and lanes 5–10, fractions eluted with 150, 300, and 500 mM and 1, 1.5, and 2 M KCl, respectively. Complexes are indicated by arrows. (C) SDS–PAGE electrophoresis (15%) of different protein fractions of the ss-DNA–agarose affinity chromatography stained with Coomassie blue. Lane 1, molecular weight markers (kDa); lane 2, 10 μg high salt extracted nuclear enriched fraction (affinity chromatography input); and lane 3, fraction eluted at 500 mM KCl (2 μg). (D) Southwestern blot assay. Poly[dT–dG]$_{15}$ was radiolabeled and used as a probe. Lanes 1–3, nuclear enriched fraction 20, 2, and 0.2 μg, respectively. Lanes 4 and 5, fractions from a ss-DNA–agarose affinity chromatography eluted with 300 and 500 mM KCl, respectively. (E) Poly[dT–dG]$_{15}$ was radiolabeled and used as a probe for EMSA with protein fractions of the poly[dT–dG]$_{15}$–biotin–avidin–agarose chromatography. Lane 1, free probe; lane 2, enriched nuclear fraction before high salt extraction; lane 3, fraction applied to the column; lane 4, fraction not bound to the column (loading conditions 150 mM KCl); and lanes 5–10, fractions eluted with 150, 300, and 500 mM and 1, 1.5, and 2 M KCl, respectively. Complexes are indicated by arrows.

Identification of the complete coding sequence of Tc38

By analogy with the orthologous genes, this EST represented up to 50% of the total coding sequence. RT-PCR and library screening was performed in order to obtain the complete coding sequence. A genomic clone was isolated containing a 1014 pb ORF coding for a predicted protein of 338 amino acids (MW 38381.1 Da) named Tc38. The genomic Southern blot pattern suggests that *Tc38* is a single copy gene (data not shown). RT-PCR allowed the identification of the 5′ splicing aceptor site which was located 36 nt upstream of the AUG (Fig. 2). Analysis of the predicted primary structure of the Tc38 protein indicated a basic isoelectric point (p*I* of 9.34). No putative conserved domains, either compositional or architectural, were found. The protein is characterized by a high frequency of Ser, Leu, and diamino acids. This peculiarity has also been reported for a putative transcription factor in *Leptomonas seymouri* [14,15]. These features suggest that Tc38 may belong to a new family of nucleic acid binding proteins. The orthologous genes in *L. major* and in *T. brucei* (coding for proteins of 347 and 338 amino acids) show amino acid identities of 70% and 79%, and similarities of 83% and 89%, respectively.

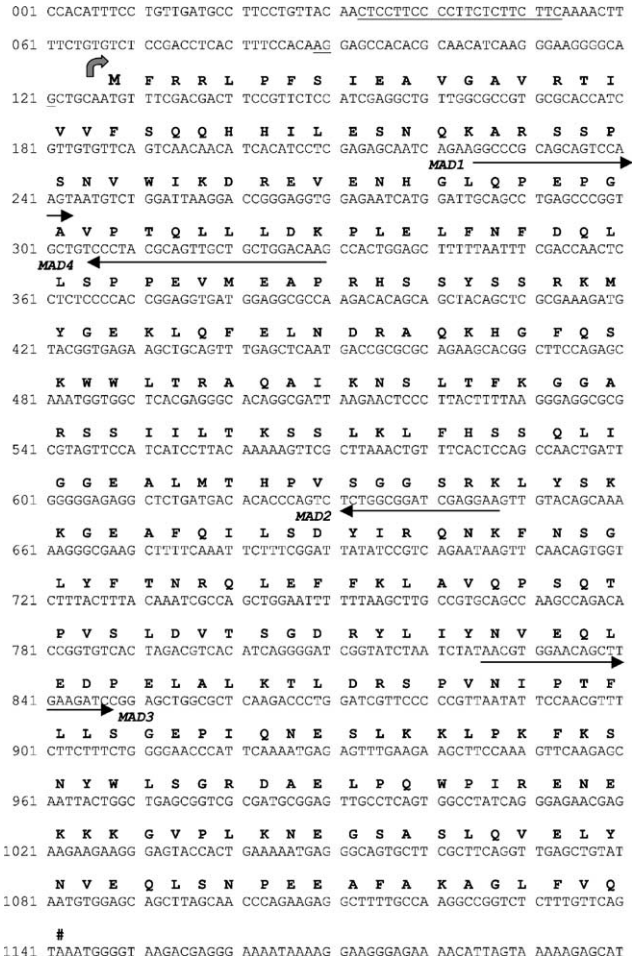

Fig. 2. Sequence analysis. The open reading frame is shown in bold. RT-PCR enabled the location of the trans-splicing acceptor AG (boxed) and the polypyrimidine tract preceding the spliced leader acceptor (underlined). The different primers used in this work are indicated by arrows. Stop codon (#).

A putative mitochondrial targeting sequence is observed in the amino terminal portion of the protein [16]. Recently, the orthologous genes in *T. brucei* and *L. tarentolae* have been reported as mitochondrial RNA binding proteins [17]. These data suggest that Tc38 may also have a mitochondrial localization. Further experiments are needed to demonstrate the actual localizations.

In vitro binding activity of the recombinant protein

The complete coding sequence was amplified and cloned in the pGEX-4T *E. coli* expression vector to obtain a recombinant protein fused to GST. A protein of approximately 64 kDa was detected in protein extracts prepared after induction of *E. coli* BL21 bacterial cells. About half of the recombinant protein remained in the soluble protein fraction after bacterial sonication. The protein was purified by chromatography on a glutathione–Sepharose column and eluted with 10 mM, pH 8, reduced glutathione. The purification yielded a major band of 64 kDa. A Western blot analysis was performed using a polyclonal anti-GST antiserum, showing that the purified protein corresponds to the Tc38-GST fusion protein (Fig. 3). The recombinant fusion protein was tested in EMSA using the [dT–dG] repeat target. This allowed for the observation of a complex yielding a clear signal (Fig. 4A). The detection of a second complex was also evident in some preparations (Fig. 4A). This second complex is also observed when using large amount of the recombinant protein. Analysis of the EMSA gels by silver staining and Western blot, using the GST-antibody, indicated that most of the protein appears as high molecular weight multimers or aggregates which do not bind the target. Nevertheless, two bands colocalized with the observed complexes (data not shown). Control experiments using GST protein were negative. Competition experiments using different polydeoxynucleotides clearly demonstrate the specificity of Tc38 for the poly[dT–dG] target (Fig. 4B). Furthermore, competition experiments using a poly[U–G] repetitive sequence showed that Tc38 presents a higher affinity for the deoxynucleotide repeats (Fig. 4B).

A novel type of single-stranded nucleic acid binding protein recognizing specifically the motif poly[dT–dG], highly frequent in the intergenic regions of *T. cruzi*, has been identified.

Single stranded nucleic acid binding proteins are involved in numerous processes in genome dynamics,

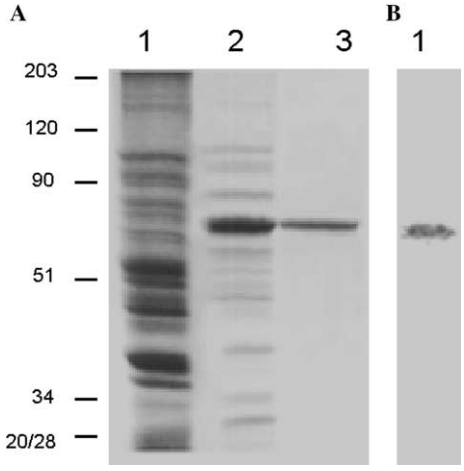

Fig. 3. Purification of the recombinant fusion protein. (A) SDS–PAGE analysis in 10% acrylamide gels. Lane 1, *E. coli* Bl21 bacterial extracts from uninduced cells; lane 2, *E. coli* Bl21 bacterial extracts from 0.5 mM IPTG induced cells; lane 3, recombinant fusion protein purified by glutathione–Sepharose chromatography. (B) Lane 1, the purified recombinant protein was electrotransferred onto Protran membranes and incubated with a 1/2000 dilution of the polyclonal anti-GST serum. The molecular mass standards (kDa) are indicated on the left.

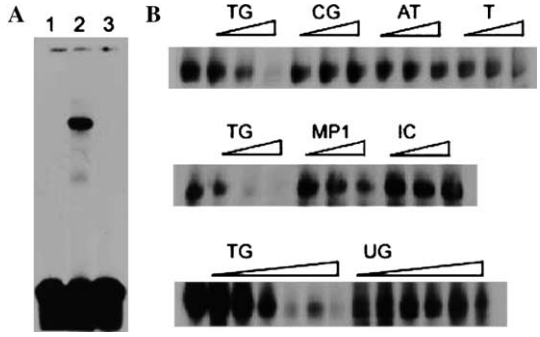

Fig. 4. EMSA with the recombinant protein. Poly[dT–dG]$_{20}$ was radiolabeled and used as a probe for EMSA with the recombinant Tc38-GST protein. (A) Lane 1, free radiolabeled probe; lane 2, 200 ng of recombinant Tc38-GST; and lane 3, 200 ng of recombinant GST. Complexes are indicated by arrows. (B) Competition assays with increasing amounts of the unlabeled nucleic acids competitors. The first lane of each gel shows the reaction without competitor. Increasing amounts of the competitors are indicated above each triangle (left to right). Upper and middle panel: 1-, 10-, and 100-fold molar excess. Lower panel: 1-, 5-, 10-, 50-, and 100-fold molar excess of the competitor. TG, [dT–dG]40; CG, poly[dC–dG]; AT, poly[dA–dT]; T, poly[dT]; MP1, 30 mer oligonucleotide; IC, poly[dI–dC]; and UG, poly[U–G].

transcription, RNA maturation, and *post*-transcriptional regulation.

The [dT–dG] repeats are highly frequent and asymmetrically strand distributed in *T. cruzi* [10]. This seems not to be restricted to trypanosomatids, as similar findings were observed also in the human, mouse, *Caenorhabditis elegans*, and yeast genomes, suggesting a specific function for these repeats [18]. The poly[dT–dG] · [dC–dA] tracts could adopt particular non-B DNA structures, generating single-stranded DNA tracts, favored or recognized by specific proteins [19–22]. The potential localization of Tc38 in mitochondria as well as nuclei leads us to search for the presence of these dinucleotide repetitive sequences in kinetoplast DNA. We have found some poly[dT–dG] tracts, but a statistical analysis could not be carried out due to the limited amount of data.

Though our in vitro analysis clearly establishes a specificity for the polydeoxy-dinucleotide probe, we could not eliminate the possibility that the actual physiological target could be the ribonucleotide. In fact, the polycistronic nature of transcription in trypanosomatids necessarily suggests another putative role of these repeats in RNA maturation and regulation [9]. It has been very recently reported that downregulation of the orthologous protein TbRBP38 in RNAi assays reduces the stability of total mitochondrial RNAs [17]. Another possibility could be the involvement of Tc38 in RNA maturation. The [U–G] repeat has been identified as an intron signal involved in alternative splicing [23] and as a strong splicing enhancer [24]. A splicing regulator protein that specifically recognizes the poly[U–G] has also the ability to bind [dT–dG] repeats [25].

In summary, we report here the identification of a novel single stranded nucleic acid binding protein from *T. cruzi* epimastigotes. The unusual features in frequency, location, and strand distribution of the target [10] strengthen the relevance of the recognition event. This characterization constitutes a contribution for the elucidation of trypanosome peculiarities in regulation of gene expression.

Acknowledgments

We thank Dr. R. Ehrlich for constant encouragement and helpful discussions. This work was financially supported by FIRCA No. R03 TW05665-01, PEDECIBA and Fondo Clemente Estable 7108 and 7109 DINACYT.

References

[1] G. Mair, H. Shi, H. Li, A. Djikeng, H.O. Aviles, J.R. Bishop, F.H. Falcone, C. Gavrilescu, J.L. Montgomery, M.I. Santori, L.S. Stern, Z. Wang, E. Ullu, C. Tschudi, A new twist in trypanosome RNA metabolism: *cis*-splicing of pre-mRNA, RNA 6 (2000) 163–169.

[2] P.J. Myler, L. Audleman, T. deVos, G. Hixson, C. Lemley, C. Magness, E. Rickel, E. Sisk, S. Sunkin, S. Swartzell, T. Westlake, P. Bastien, F. Guoliang, A. Ivens, K. Stuart, *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes, Proc. Natl. Acad. Sci. USA 96 (1999) 2902–2906.

[3] B. Andersson, L. Aslund, M. Tammi, A.N. Tran, J.D. Hoheisel, U. Pettersson, Complete sequence of a 93.4-kb contig from

chromosome 3 of *Trypanosoma cruzi* containing a strand-switch region, Genome Res. 8 (1998) 809–816.
[4] C.E. Clayton, Life without transcriptional control? From fly to mammals back again, EMBO J. 21 (2002) 1881–1888.
[5] P. Dubessy, C. Ravel, P. Bastien, L. Crobu, J.P. Dedet, M. Pagès, C. Blaineau, The switch region on *Leishmania major* chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential, Nucleic Acids Res. 30 (2002) 3692–3697.
[6] J. Swindle, A. Tait, Trypanosomatid genetics, in: D.F. Smith, M. Parsons (Eds.), Molecular Biology of Parasitic Protozoa, Oxford University Press, NY, 1996, pp. 6–34.
[7] S.M. Teixeira, Control of gene expression in *Trypanosomatidae*, Braz. J. Med. Biol. Res. 31 (1998) 1503–1516.
[8] T. Nosaki, G.A.M. Cross, Effects of 3′untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*, Mol. Biochem. Parasitol. 75 (1995) 55–67.
[9] I. D'Orso, J.G. De Gaudenzi, A.C.C. Frasch, RNA-binding proteins and mRNA turnover in trypanosomes, Trends Parasitol. 19 (2003) 151–155.
[10] M.A. Duhagon, B. Dallagiovanna, B. Garat, Unusual features of poly[dT–dG]·[dC–dA] stretches in CDS-flanking regions of *Trypanosoma cruzi* genome, Biochem. Biophys. Res. Commun. 287 (2001) 98–103.
[11] Z. Brener, E. Chiari, Variaçoes morfologicas observadas em diferentes amostras de *Trypanosoma cruzi*, Rev. Inst. Med. Trop. Sao Paolo 5 (1963) 220–224.
[12] J. Pitula, W.T. Ruyechan, N. Williams, *Trypanosoma brucei*: identification and purification of a poly(A)-binding protein, Exp. Parasitol. 88 (1998) 157–160.
[13] J.A. Coderre, S.M. Beverly, R.T. Schimke, D.B. Santi, Overproduction of a bifunctional Thymidylate Synthetase-Dihydrofolate Reductase and DNA amplification in methotrexate resistant *Leishmania tropica*, Proc. Natl. Acad. Sci. USA 80 (1983) 2132–2136.
[14] A. Matkin, A. Das, V. Bellofatto, The *Leptomonas seymouri* spliced leader RNA promoter requires a novel transcription factor, Int. J. Parasitol. 31 (2001) 545–549.
[15] A. Das, V. Bellofatto, RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAPcomplex-like transcription factor, Proc. Natl. Acad. Sci. USA 100 (2003) 80–85.
[16] T. Hausler, Y.D. Stierhof, J. Blattner, C. Clayton, Conservation of mitochondrial targeting sequence function in mitochondrial and hydrogenosomal proteins from the early-branching eukaryotes *Crithidia*, *Trypanosoma* and *Trichomonas*, Eur. J. Cell. Biol. 73 (1997) 240–251.
[17] S. Sbicego, J.D. Alfonzo, A. Estévez, M.A.T. Rubio, X. Kang, C.W. Turck, M. Peris, L. Simpson, RBP38, a novel RNA-binding protein from trypanosomatid mitochondria, modulates RNA stability, Eukaryotic Cell 2 (2003) 560–568.
[18] N.V. Dokholyan, S.V. Buldyrev, S. Havlin, H.E. Stanley, Distribution of dimeric tandem repeats in non-coding and coding DNA sequences, J. Theor. Biol. 202 (2000) 273–282.
[19] P.S. Ho, The non-B-DNA structure of d(CA/TG)n does not differ from that of Z-DNA, Proc. Natl. Acad. Sci. USA 91 (1994) 9549–9553.
[20] M.P. Kladde, Y. Kohwi, T. Kohwi-Shigematsu, J. Gorski, The non-B-DNA structure of d(CA/TG)n differs from that of Z-DNA, Proc. Natl. Acad. Sci. USA 91 (1994) 1898–1902.
[21] E. Biet, J.S. Sun, M. Dutreix, Conserved sequence preference in DNA binding among recombination proteins: an effect of ssDNA secondary structure, Nucleic Acids Res. 27 (1999) 596–600.
[22] C. Gaillard, F. Strauss, DNA loops and semicatenated DNA junctions, Biochem. Struct. Biol. 1 (2000) 1. Available from <http://biomedcentral.com/1471-2237/1/1>.
[23] F. Pagani, E. Buratti, C. Stuani, M. Romano, E. Zuccato, M. Niksic, L. Giglio, D. Faraguna, F.E. Baralle, Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a non evolutionary conserved intronic element, J. Biol. Chem. 275 (2000) 21041–21047.
[24] N. Gabellini, A polymorphic GT repeat from the human cardiac Na^+ Ca^{2+} exchanger intron 2 activates splicing, Eur. J. Biochem. 268 (2000) 1076–1083.
[25] E. Buratti, F.E. Baralle, Characterization aand functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of *CFTR* Exon 9, J. Biol. Chem. 276 (2001) 36337–36343.