Experimental Parasitology 124 (2010) 346–349

Contents lists available at ScienceDirect

Experimental Parasitology

journal homepage: www.elsevier.com/locate/yexpr

Research Brief

Trypanosoma cruzi TBP shows preference for C/G-rich DNA sequences *in vitro*

Pamela Cribb [a], Luis Esteban [b], Andrea Trochine [a], Javier Girardini [a], Esteban Serra [a,*]

[a] Instituto de Biología Molecular y Celular de Rosario, CONICET, Facultad de Ciencias Bioquímicas y Farmacéuticas, UNR, Suipacha 351, 2000 Rosario, Argentina
[b] Facultad de Ciencias Médicas, UNR, Santa Fe 3100, 2000 Rosario, Argentina

ARTICLE INFO

Article history:
Received 30 July 2009
Received in revised form 28 September 2009
Accepted 12 November 2009
Available online 16 December 2009

Keywords:
Trypanosoma cruzi
TATA-binding protein
SELEX

ABSTRACT

Recent findings associate transcription start in trypanosomatids with chromatin regions containing modified and variant histones. TATA-binding protein (TBP) and other fundamental transcription factors have been also found at these Transcription Start Sites (TSS). Results of Systematic Evolution of Ligands by Exponential Enrichment (SELEX) experiments show that *Trypanosoma cruzi* TBP (TcTBP) has an *in vitro* binding preference for G-rich sequences. This finding correlates with the presence of G-rich stretches at the Strand Switch Regions (SSR) and at some putative internal TSS in *Trypanosoma brucei* and *Leishmania*. A scanning study of partially assembled *T. cruzi* genomic contigs determined the presence of G-rich stretches in the coding strands. TcTBP affinity for G-rich sequences suggests that this factor could play a role in locating the initiation complex in the right TSS, probably by "sensing" the G-content on the strand to be transcribed.

1. Introduction

Trypanosomatid parasites are a group of early divergent protozoa that cause severe diseases in humans, including leishmaniasis, sleeping sickness and Chagas' disease. In these parasites transcription is polycistronic and seems to be regulated only globally by chromatin-mediated epigenetic events. Nuclear run-on assays demonstrated that RNA pol II transcription starts in both directions within the Strand Switch Region (SSR) (Martínez-Calvillo et al., 2003, 2004). In addition to the RNA polymerases, a reduced number of very divergent transcription factors, an ortholog of TATA-binding protein (TBP) among them, also participate in the transcription process.

In *Trypanosoma brucei*, TBP showed to be essential for RNAP I, II, and III transcription and is recruited to the SL RNA-gene promoter, as well as to RNAP I-transcribed procyclic acidic repetitive genes and RNAP III-transcribed U-rich snRNA and 7SL RNA genes (Ruan et al., 2004; Das et al., 2005). Recruitment of trypanosomatid TBP to the SL RNA-gene promoter was also demonstrated for *Leishmania tarentolae* and *Trypanosoma cruzi* (Thomas et al., 2006; Cribb, unpublished results). In this promoter, TBP is part of a complex that includes three divergent components of the SNAP complex, at least one TFIIA orthologous subunit and a very divergent version of TFIIB (Das and Bellofatto, 2003; Ruan et al., 2004; Schimanski et al., 2005, 2006; Das et al., 2005; Palenchar et al., 2006; Cribb and Serra, 2009).

2. Materials and methods

2.1. Expression of recombinant TcTBP

Trypanosoma cruzi TBP coding region from CL-Brener strain (Tc00.1047053503809.149) was amplified using *Pfu* polymerase with specific oligonucleotides, cloned into pGEMT-easy plasmid (Promega) and sequenced. TcTBP coding sequence was cloned in pGEX4-T3 and pQE30 (QIAGEN) expression plasmids. Recombinant GST-fusion and His-tagged proteins were purified by affinity chromatography using standard procedures.

2.2. Determination of TcTBP recognition sequence

Selection of recognition sequence for TcTBP was performed by SELEX (Systematic Evolution of Ligands by Exponential Enrichment) assay using recombinant TBP, a 54-mer double stranded oligonucleotide containing 12 completely degenerated nucleotides (CHRI, 5′-gATgAAgCTTCCTggACAAT-$(N)_{12}$-gCAgTCACTgAAgAATTCTg-3′) and two primers complementary to the conserved regions on the random primer (CHRIS-1: 5′-gATgAAgCTTCCTggACAAT-3′; CHRIS-2: 5′-CAgAATTCTTCAgTgACTgC-3 (Oliphant et al., 1989). Before the first binding, CHRI was converted to double stranded DNA with DNA polymerase I Klenow fragment, using CHRIS2 as primer. Agarose-immobilized GST–TcTBP was incubated with 20 pmol of double stranded random oligonucleotide for 30 min at 4 °C in 200 µl binding buffer (20 mM Hepes pH 7, 5 mM $MgCl_2$, 5 mM DTT, 150 mM NaCl, 0.1% Triton X-100, 10% glycerol). After centrifugation, agarose was washed 5 times at 4 °C with 1 ml binding

* Corresponding author. Fax: +54 341 4390465
E-mail addresses: serra@ibr.gov.ar, eserra@fbioyf.unr.edu.ar (E. Serra).

0014-4894/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2009.11.003

buffer containing 0.1 mg/ml sonicated salmon sperm DNA, resuspended in 50 μl TE and boiled. Two microliters of supernatant were used for amplification with CHRIS-1 and CHRIS-2 and amplification products were analyzed by polyacrylamide gel electrophoresis. Further cycles were performed in the same way, using 2 μl of amplification product from the precedent binding assay. After six cycles, the products of amplification reaction were cloned in pGEMT-easy plasmid (Promega). A control assay was performed in parallel using purified GST immobilized to the agarose beads, instead of the fusion protein.

2.3. Analysis of strand switch-containing contigs from T. cruzi

Assembled sequences from the *T. cruzi* genome were recovered from NCBI nucleotide database (CH473309:CH473946[PACC]). Contigs with a Strand Switch (El-Sayed et al., 2005, Supplementary material) were visualized with Artemis and their CG-skews were calculated with the same program. Scanning of the contigs in sliding windows for the calculation of the C or G content was performed with an ad hoc written Pearl script as follows: A 100-nucleotides sliding window moves along the contig, one nucleotide by step, and calculates the G (or C) composition at each position. The results of the scanning are recovered in a table where the percentage of the nucleotide obtained for each window is assigned to the first position of the window. The graphics were generated from data in the table using R (http://CRAN.R-project.org). The script is given as Supplementary material (file: CounCG.pl).

3. Results and discussion

To explore the TcTBP DNA-binding preference we performed two different SELEX assays. First, we performed a typical gel shift-based SELEX, however no shifted band was observed when GST- or His-tagged TcTBP was incubated with the pool of random oligonucleotides, under the tested conditions. Then, we performed a liquid-binding and selection SELEX, using GST–TcTBP fusion protein bound to a GS-agarose column as described in Section 2. A control assay was performed in parallel using purified GST instead of the fusion protein. The results of the amplification reactions from GST–TcTBP- and GST-binding assays were analyzed in parallel by polyacrylamide gel electrophoresis, after washing, at each cycle. No amplification product was observed in control assay after the first cycle. In following cycles amplification products from the GST–TcTBP-binding reaction were also incubated with GST. No amplification product was recovered after anyone of the binding and washing cycles in these control assays. After six rounds, amplified oligonucleotides that bounded to TcTBP were cloned in pGEM-T easy plasmid (Promega). Plasmids from 34 randomly chosen colonies were extracted and sequenced. A first analysis showed that sequences obtained were rich (67%) in C/G. Sequences were analyzed both by alignment and visual inspection, and by the Multiple Expectation Maximization for Motif Elicitation (MEME) program (Bailey and Elkan, 1994). Fig. 1 shows sequences obtained after MEME-driven alignment and a consensus motif generated GG(C/T)G(T/G/A) by the program. For the run, the program was set to align each sequence or its reversed complementary one (which is indicated as + or – strand in the figure). The final alignment is rich in G, which is reflected in the high frequency for this nucleotide at positions 1 (0.70), 2 (0.73) and 4 (0.91) of the motif represented in Fig. 1C. The fact that the interaction between TBP and DNA could not be observed directly suggests that this interaction should be very weak, and most probably mediated and/or stabilized by other factors in the complexes. This could explain why we succeeded in selecting interacting oligonucleotides in a liquid-binding-selection

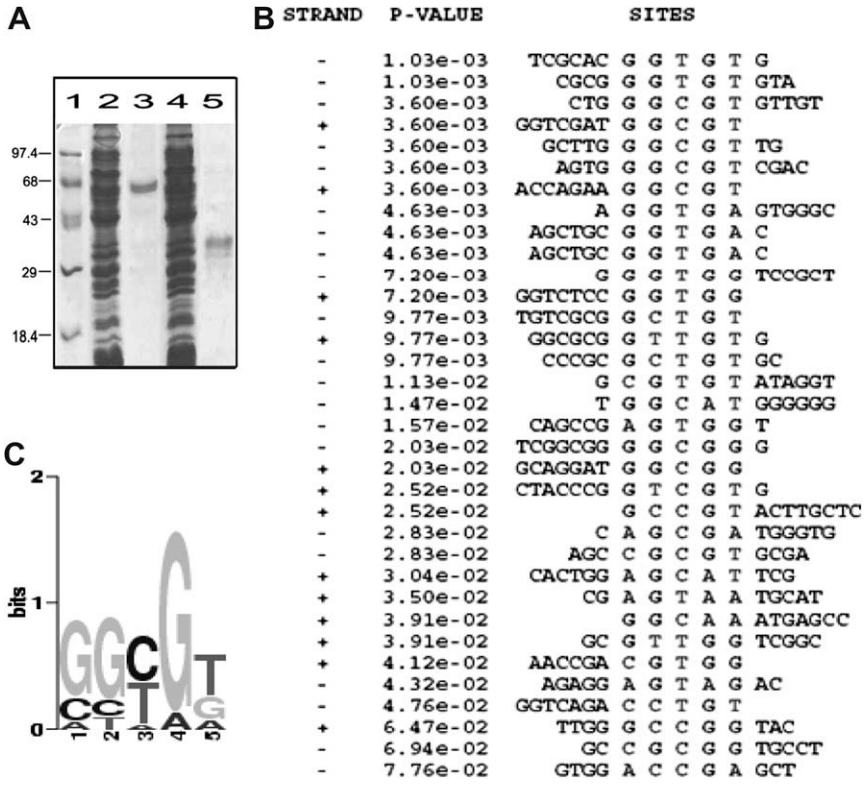

Fig. 1. (A) Recombinant GST- and His-TcTBP purified from *E. coli*. 1. Molecular weight marker. 2. *E. coli* pGEX-TcTBP total extract. 3. Purified GST–TcTBP. 4. *E. coli* pQE-TcTBP total extract. 5. Purified His-TcTBP. (B) Sequences from oligonucleotides recovered after six SELEX cycles aligned by MEME. Minus (–) means that the sequence was complementary reverted by the program during the calculations. The *P*-value generated by MEME estimates the probability for a given sequence to be aligned with the core motif by chance in a dataset of the same size. (C) Logo representation of the 5 nucleotides consensus generated by MEME.