Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 364 (2007) 26–32
BBRC
www.elsevier.com/locate/ybbrc

Mapping of the protein-binding interface between splicing factors SF3b155 and p14 of *Trypanosoma cruzi*

M. Lara Avila [a,1], Natalia Bercovich [a,1], Gastón Westergaard [a], Mariano J. Levin [b], Martín P. Vázquez [a,*]

[a] *Departamento de Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, INGEBI-CONICET, Vta. de obligado 2490, Buenos Aires, Argentina*
[b] *Institut Cochin, INSERM U567, Paris, France*

Received 18 September 2007
Available online 1 October 2007

Abstract

SF3b155 and p14 are essential components of spliceosome core that recognize the branch point adenosine, a critical step in splicing in eukaryotes. Trypanosomes are unusual since every transcribed gene is processed by *trans*-splicing instead of *cis*-splicing. Thus, the *trans*-spliceosome emerges as an interesting anti-parasitic drug target since this process is not present in mammalian hosts. Here, we present the orthologues of these proteins in *Trypanosoma cruzi* that interact strongly with each other. To define similarities and differences with the human pair, we performed a detailed alanine scan analysis that allowed us to identify the regions and the critical amino acids of *T. cruzi* SF3b155 involved in interaction with p14. We demonstrate that the *T. cruzi* SF3b155 interface is larger and contains more complex elements than its human counterpart. Additionally, our results provide the first insights into the core of the putative mRNA processing complex of trypanosomes.

Keywords: *Trans*-splicing; Drug targets; Protein interaction

Trypanosomes are intriguing and amazing organisms in many aspects of its molecular biology [1]. The genome sequence of the three model trypanosomes (*Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*) were published in 2005 [2], thus, providing a major tool to the understanding of several of their unusual aspects. However, with so many different mechanisms between these parasites and its mammalian host there is still a lack of availability of effective anti-parasitic drugs or disease treatments [3].

RNA transcription in these parasites is polycistronic encompassing several dozen kilobases and they lack conventional promoters. Polycistronic pre-mRNAs are processed by two coupled reactions, *trans*-splicing and polyadenylation, and exported to the cytoplasm as monocistronic mRNAs [1,4] . Thus, mRNA maturation in trypanosomes differs from most eukaryotes. *Trans*-splicing involves the joining of two different molecules, the polycistronic pre-RNA and a capped 39 nucleotide sequence named Splice Leader RNA (SL-RNA). Little is known about the complex and its components that carry-out the *trans*-splicing/polyadenylation reactions. However, it is important not only to dissect the basic aspects of the mechanism but also because it could provide several interesting potential anti-parasitic drug targets.

We have found several putative splicing factors in the genome of *T. cruzi* using bioinformatics analyses, including

Abbreviations: aa, amino acids; NMR, Nuclear Magnetic Resonance; RRM; RNA Recognition Motif; 3AT, 3-amino-1,2,4-triazole. Accession Nos. AY895171 for SF3b155; AY294609 for Tcp14.

* Corresponding author. Fax: +54 11 47868578.
E-mail address: mvazquez@fbmc.fcen.uba.ar (M.P. Vázquez).

[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.09.090

those homologous to the E (early) complex of mammalian systems (U2AF65/U2AF35/SF1). The E complex of trypanosomes presents several unusual features such as a divergent U2AF65 and the lost of interaction with its U2AF35 partner [5].

In this work, we focus our attention in the protein factors homologous to the A complex. Mammalian spliceosome assembly is dynamic and the initial recognition of the 3′ splice site by the E complex is followed by the A complex where splicing factor SF1 is replaced by a multiprotein component named SF3b [6]. This particle is coordinated by splicing factor SF3b155 that establishes protein interactions with SF3b145, SF3b130, SF3b49, SF3b14b, SF3b10 and p14 [7,9]. The function of SF3b is crucial in the recognition of the branch point adenosine in the splicing reaction. The molecular architecture of human SF3b was determined by single-particle electron cryomicroscopy at a resolution of less than 10 angstroms [8], allowing identification of p14 in the central cavity of the complex. Of particular interest was the strong interaction between SF3b155 and p14 that appeared in the A complex and persisted within the fully assembled spliceosome. Human p14 is a small RRM containing protein of 125 aa that cross-link with the branch-point adenosine [9]. Crystal structure and NMR analysis of the protein interaction interface between human SF3b155 and p14 were reported by independent groups [10,11].

Here, we found orthologues of these core spliceosomal proteins in the genome of *T. cruzi* that interact strongly with each other. To define similarities and differences with the human pair, we performed a detailed alanine scan analysis that allowed us to identify the region and the critical amino acids of *T. cruzi* SF3b155 involved in interaction with p14. Our results provide the first insights into the core of the putative mRNA processing complex of trypanosomes.

Materials and methods

*Data mining.*The protein sequences of human p14 and SF3b155 were used as probes in a TBLASTN to search of a locally licensed copy of the complete nucleotide sequence of the *T. cruzi* genome (http://www.tigr.org/tdb/e2k1/tca1/). Results were correlated with the annotations in GENEDB (http://www.genedb.org). ESTs coding for both proteins were found in the NCBI database dbEST (*T. cruzi* filtered) using the blastn algorithm (http://www.ncbi.nlm.nih.gov/BLAST/).

Cloning of TcSF3b155, Tcp14 genes in the Gateway™ system. The coding sequences of TcSF3b155 and p14 were PCR amplified from 100 ng of total genomic DNA of *T. cruzi* CL-B clone using the following primers: Tc155_up: 5′ GGATCCGATGACGGACGAAGAAAGAAGC 3′; Tc155_down: 5′ CAACTACAGAATCACCTCCAG 3′; Tcp14_up: 5′ GGATCCAATGCCGGATGAACGCATTC 3′; Tcp14_down: 5′ CTCGAGTTAGCTCTTCTCCTCAGTGA 3′. PCR products were cloned into pGEM-T Easy™ (Promega), subsequently digested with BamHI and EcoRI and subcloned in frame into the Gateway™ entry vector pENTR-2B (Invitrogen).

Construction of mutant proteins. N-terminal and C-terminal deletion mutants were generated by PCR with the specific primers (see Supplementary table S1). Internal deletion mutants and Alanine scan mutants were generated by SOEing PCR [12] using the primers listed in Supplementary tables S2 and S3. All the mutant products were Gateway™ cloned as described above.

All PCRs were performed using the proofreading AccuTaq™ enzyme (Sigma–Aldrich) and all the products were sequenced on a MegaBACE 500 (Amersham Biosciences) capillary sequencer to verify the mutations introduced.

Mapping of the interaction interface using yeast two-hybrid analysis. The ProQuest™ Yeast Two-Hybrid Gateway™ compatible System (Invitrogen) was used for the protein interaction analysis as previously described [13]. Tcp14 variants were always analyzed in DB configurations and SF3b155 variants in AD configurations.

Activation of *His3* reporter gene induced by interaction pairs was studied analyzing the ability of yeast to grow on plates lacking His (SC $L^-W^-H^-$) in the presence of 25 mM of 3AT and incubated at 30 °C for up to 48 h.

Activation of *LacZ* reporter gene was monitored by a β-galactosidase activity liquid assay. Reactions were performed in duplicates and measured at $OD_{420\ nm}$. The data was normalized with the amount of yeast in the initial culture ($OD_{600\ nm}$).

Simultaneously, we processed the ProQuest's negative Control A (empty DB and AD) and positive Control E (DB—full length GAL4 protein) as standard controls.

Results

Characterization of TcSF3b155 and Tcp14

Splicing factor Tcp14 is a small RRM containing protein of 117 amino acids. It shares 42% identical and 61% conserved residues compared to its human orthologue. This homology is restricted to the 77 aa long RRM domain. There is no significant sequence conservation in the C-terminal domain but both proteins present an extended alpha helix in this region (not shown).

Splicing factor SF3b155 is a large protein of 1112 aa and it is 192 aa shorter than its human counterpart. It is a modular protein with an unfolded N-terminal domain and a helical C-terminal region bearing several HEAT repeats [7,8]. TcSF3b155 shares 33% identical and 51% conserved residues compared to its human orthologue. However, the N-terminal domain of the *T. cruzi* protein is less conserved (27%) and 161 aa shorter than the same region in HsSF3b155.

Moreover, while HsSF3b155 N-terminal domain contains 22 TP dipeptide and 5 RWD repeats, TcSF3b155 contains 5 TP and 7 GGTTP repeats lacking RWD repeats (Fig. 1).

Using Proquest yeast two hybrid assays, we demonstrated that TcSF3b155 and Tcp14 interacted strongly with each other. In fact, they activated the *His3* reporter gene in plates lacking histidine with concentrations up to 75 mM 3AT (not shown) and the *LacZ* reporter gene with a strength of interaction equivalent to 80% of the strongest Proquest positive control (see E and Tc155 in Fig. 2C).

Mapping the essential domains in TcSF3b155 for interaction with Tcp14

In order to map the minimal region of TcSF3b155 involved in protein interaction with Tcp14, we generated