Molecular & Biochemical Parasitology 164 (2009) 137–146

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Functional characterization and protein–protein interactions of trypanosome splicing factors U2AF35, U2AF65 and SF1

Martin P. Vazquez[a,b], David Mualem[c], Natalia Bercovich[a,b], Michael Zeev Stern[c], Benson Nyambega[a,d], Omer Barda[c], Dan Masiga[d], Sachin Kumar Gupta[c], Shulamit Michaeli[c], Mariano J. Levin[a,e,*]

[a] Laboratorio de Biología Molecular de la Enfermedad de Chagas, INGEBI- CONICET, Vuelta de Obligado 2490 2P, 1428, Buenos Aires, Argentina
[b] Departamento de Fisiología y Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, University of Buenos Aires, Argentina
[c] The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan 52900, Israel
[d] International Center for Insect Physiology and Ecology, P.O.Box 30772-00100, Nairobi, Kenya
[e] Institut Cochin, Departement Maladies Infectieuses, INSERM U567, Paris F-75014, France

ARTICLE INFO

Article history:
Received 17 August 2008
Received in revised form
24 November 2008
Accepted 30 December 2008
Available online 15 January 2009

Keywords:
Trypanosoma cruzi
Trypanosoma brucei
Trans-splicing
U2 auxiliary factor

ABSTRACT

Early in the assembly of the spliceosome of eukaryotes the branch-point binding protein (BBP, also called SF1) recognizes the branch point sequence, whereas the heterodimer U2AF consisting of a 65 and a 35 kDa subunit, contacts the polypyrimidine tract and the AG splice site, respectively. Herein, we identified, cloned and expressed the *Trypanosoma cruzi* and *Trypanosoma brucei* U2AF35, U2AF65 and SF1. Trypanosomatid U2AF65 strongly diverged from yeast and human homologues. On the contrary, trypanosomatid SF1 was conserved but lacked the C-terminal sequence present in the mammalian protein. Yeast two hybrid approaches were used to assess their interactions. The interaction between U2AF35 and U2AF65 was very weak or not detectable. However, as in other eukaryotes, the interaction between U2AF65 and SF1 was strong. At the cellular level, these results were confirmed by fractionation and affinity-selection experiments in which SF1 and U2AF65 co-fractionated in a complex of approximately 400 kDa and U2AF65 was affinity-selected with TAP tagged SF1, but not with TAP tagged U2AF35. Silencing of the three factors affected growth and *trans*-splicing in the first step of this reaction. Trypanosomes are the first described example of eukaryotic cells in which the interaction of two expressed U2AF factors seemed to be very weak, or even undetectable.

1. Introduction

The genome sequences of the three trypanosomatid species suggest that in these parasites entire chromosomes are transcribed as long polycistronic transcripts that are maturated by the concerted action of *trans*-splicing and polyadenylation [1,2]. Through *trans*-splicing, the first 39-nt (miniexon or spliced leader, SL) of a 139 nt long capped RNA, the SL RNA, is joined at appropriate sites to each of the protein-coding exons of the primary transcripts. Thus, SL addition serves two functions: together with polyadenylation, it dissects mRNAs from polycistronic primary transcripts, and it provides the cap structure to all mRNAs. In contrast to *cis*-splicing, *trans*-splicing joins exons derived from two independently transcribed RNAs. Mechanistically, however, *trans*- and *cis*- splicing share several common features, both require the same motifs, the GU at the 5′ splice site (SS), the adenosine branch point, a polypyrimidine tract (Py) and the AG at the 3′ SS [1,2].

Early in the assembly of the spliceosome of higher eukaryotes the branch point sequence is recognized by the branch-point binding protein (BBP, also called SF1) [3]. Human U2AF was initially identified as an activity necessary for the recruitment of U2 snRNP to the branchpoint [4] and consists of a 65 [5] and a 35 kDa subunit [6]. The large subunit contacts the Py [5], while the 35 kDa subunit (U2AF35) binds to the 3′SS AG [7]. Thus, splicing at the first AG downstream of the branch point is a consequence of the combined interactions between these two factors. In eukaryotes, U2AF65 and U2AF35 form a heterodimeric complex in which the RRM of U2AF35 and a central polyproline segment of U2AF65 interact via reciprocal "tongue in groove" involving tryptophan residues [8]. The latter is composed by an N-terminal Arg-Ser (RS) rich domain followed by three RRM motifs. The N-terminal domain helps to strengthen the interaction between the U2 snRNP and the branch point region [9], whereas the first two RRM domains mediate the interaction with the Py tract [10]. The third RRM domain, RRMIII, is engaged in a strong unidirectional "tongue in groove" interaction with a tryptophan residue of the N-terminal portion of SF1

* Corresponding author at: Laboratorio de Biología Molecular de la Enfermedad de Chagas, INGEBI- CONICET, Vuelta de Obligado 2490 2P, 1428, Buenos Aires, Argentina. Tel.: +54 11 4783 2871; fax: +54 11 4786 8578.
E-mail address: mlevin@dna.uba.ar (M.J. Levin).

0166-6851/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.molbiopara.2008.12.009

[10]. The U2AF35-65 heterodimer has been described in humans [11], *Caenorhabditis elegans* [12], *Drosophila melanogaster* [13] and *Schizosaccharomyces pombe* [14], and in these species its existence is essential for life. However, in *S. cerevisiae* no homologue of U2AF35 exists, and U2AF65 is not essential [14].

In trypanosomes, no homologue of U2AF65 has been yet identified and the only protein of the U2AF complex that has been characterized is the U2AF35 of *Trypanosoma cruzi* [15]. This protein shared 38% identity with its human homologue, and immunolocalization demonstrated its distribution in nuclear speckles [15]. Interestingly, the central RRM motif of TcU2AF35 was surrounded by two CCCH zinc finger motifs (Cx8Cx5Cx3H) and instead of the classical C-terminal SR domain it accommodated a third and peculiar zinc finger motif, CCHC. Residues Thr 45, Leu 47 and Tyr 114, known to be directly involved in RNA recognition were conserved. On the contrary, only two of the seven residues that were essential for the interaction with U2AF65 were present in the *T. cruzi* protein. Notably, Trp 134, necessary for the reciprocal "tongue in groove" heterodimerization with U2AF65, was changed to Lys, suggesting that, in trypanosomes, the U2AF35–U2AF65 interaction could be altered.

In this study we identified, cloned and expressed the *T. cruzi* and *T. brucei* U2AF65 and SF1, together with the U2AF35 of *T. brucei*. As predicted from sequence analyses, the interaction between U2AF35 and U2AF65 was very weak or not detectable. However the typical U2AF65, unidirectional "tongue in groove" interaction with SF1 was as strong as the one described for other organisms. Silencing of all three factors affected growth and *trans*-splicing in the first step of this reaction [2,16,17,18]. Trypanosomes are the first described example of eukaryotic cells in which both U2AF factors are expressed but do not engage in a strong protein–protein interaction. This feature may be related to the very flexible distance requirements between the 3′ splice site and branch point of trypanosome pre-mRNAs [19].

2. Materials and methods

2.1. Cloning of Tc/Tb U2AF35, Tc/Tb U2AF65 and Tc/TbSF1 in the gateway system

The coding sequences of these genes were PCR amplified from 100 ng of total genomic DNA of *T. cruzi* CL-Brener and *T. brucei* 927 strain using the primers specified in Supplemental Material.

PCR products were cloned into pGEM-T (Promega), subsequently digested with *Kpn*I and *Eco*RI, *BamH*I and *Not*I or *BamH*I and *Eco*RI, respectively and subcloned in frame into the Gateway entry vector pENTR2B (Invitrogen).

Regions and domains were amplified from the corresponding full protein pGEM-T clones. Amino-terminal U2AF65 was made by using primer 65up and 65-Nterm 5′CTAATCGCGTGGCCGACGGATTG 3′. U2AF65RRM2 was made by using primer 65RRM2up 5′GGATCCCAGACGCAAGGTTGTTATC 3′ and 65RRM2down 5′CT AGCTGCCGTACAACGTTGA 3′. pRRM3 was made by using primer 65pRRMup 5′GGATCCCACGGTGCCTATTTACCCCA and 65down.*T. brucei* coding sequences were amplified from gDNA using the following primers: TbSF1 with 5′ GCTAGCGATGGGGGAGAACCGTCGAC 3′ up containing a NheI site and 5′ GCGGCCGCCTACAACTCGTCCAGAAA 3′ down containing a NotI site for direct cloning into the Binding Domain vector pDBLeu (Invitrogen). TbU2AF35 with 5′GGATCCCATGTATCAAGACCGTTGC 3′ up containing BamHI and 5′ TTATTTTAAGGGGCATTCGCG 3′ down. TbU2AF65 with 5′ GGATCCAATGGGGCGTGATAGTCGCG 3′ up containing BamHI and 5′ TTAACCGTCAATACCTGCGATAC 3′ down with NotI site.

Construction of modified proteins. The TcSF1PPRR was made by PCR amplification using a sense primer containing the mutations CCG (Pro) to CGG (Arg) and CCT (Pro) to CGT (Arg), SF1PPRR 5′GGATCCGGAGGCGAAACGGAGGCG TTCGC 3′. The TcSF1SRDA was made by PCR amplification where the sense primer contain the mutation CGC (Arg) to GAC (Asp) and TGG (Trp) to GCG (Ala), RWDA 5′GGATCCGGCGAAACCGAGGCCTTCGGACGCGAGCAAAGAG3′.
SF1down was used as antisense primer in both amplifications. TcU2AF35WF was generated by SOEing PCR [20] as follows, two complementary internal primers were designed containing the mutations AAG (Lys) to TGG (Trp) and TTA (Leu) to TTT (Phe) in the RRM domain, tc35WF-s 5′GGAACTAAAGGCGAAATGGTTTAACGAAAT 3′ and tc35WF-as 5′ATTTCGTT AAACCATTTCGCCTTTAGTTCC 3′. These primers were used in combination with the external primers 35down and 35up, respectively, to produce the PCR hemi-products. These hemi-products were gel purified and combined in equal amounts to perform a second PCR with the external primers to generate the final product that was Gateway cloned as described.

The deletion mutant pRRM3Δnors was also generated by SOEing PCR using the internal complementary primers pRRMdel-up 5′CCGCCGTGCCCCACCACCGATGCATCCAATTCACGGTGGTTTTGG 3′ and pRRMdel-down 5′CATCGGTGGTGGGG CACGGCGG 3′ and the external primers 65down and 65pRRMup, respectively.

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

2.2. Yeast two-hybrid pair wise analysis

The ProQuest Yeast Two-Hybrid Gateway compatible System (Invitrogen) was used for protein–protein interaction analysis as previously described [21].

Identification of proteins that interact with U2AF35 and U2AF65 by yeast two-hybrid (Y2H) screening. Bait cloning and Y2H screening were performed by Hybrigenics, S.A., Paris, France (http://www.hybrigenics.com/services.html).

DNA from *Trypanosoma brucei* 927 strain, prepared by Dr. M. Turner, Glasgow Biomedical Research Centre, University of Glasgow, Scotland, UK, was randomly sheared and used to construct genomic library into the Y187 yeast strain. The library contained 7.5 million independent fragments, and was used for screening [22]. Forty-four million interactions were actually tested with U2AF35 and 97 million with U2AF65. After selection on medium lacking Leu, Trp and His, 190 positive clones were picked from U2AF35 full-scale screen and 36 positive clones from U2AF65 full-scale screen.

2.3. Protein extract preparation

For total protein extract preparation, parasites were resuspended in lysis buffer (5% NP40, 2 M $MgCl_2$, 500 mM EDTA, 10 mM DTT, 250 mM Tris–HCl (pH 7.6), 100% glycerol, 0.25 M sucrose) supplemented with protease inhibitors as described in Ref. [23].

2.4. RNA silencing and transfection of T. brucei

The constructs were generated following protocols described in Ref. [16] and cloned into the pZJM and pLew100 vector. To generate the *T. brucei* transgenic cell lines, strain 29–13 was used for transfection and clonal population of the transgenic parasites were obtained as described [16]. Northern analysis and primer extension to monitor the silencing and splicing defects were as previously described [16]. A list of all the oligonucleotides used in these assays is provided as Supplemental Material.

TAP-tag purification of T. brucei U2AF35, U2AF65, SF1. To generate the TAP-tagged versions, the genes were amplified with primers specified in Supplemental Material.