RNA Recognition Motif-type RNA-binding Proteins in *Trypanosoma cruzi* Form a Family Involved in the Interaction with Specific Transcripts in Vivo*

Javier G. De Gaudenzi‡, Iván D’Orso§, and Alberto C. C. Frasch¶

From the Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús, CONICET-UNSAM, 1650 San Martín, Provincia de Buenos Aires, Argentina

Trypanosomes, protozoan parasites from the order Kinetoplastida, have to deal with environmental changes during the interaction with their hosts. *Trypanosoma cruzi*, the causative agent of Chagas’ disease, uses post-transcriptional mechanisms to regulate gene expression. However, few RNA-binding proteins involved in mRNA turnover control have been identified to date. In this work, an RNA recognition motif (RRM)-type RNA-binding protein family named *T. cruzi* RNA-binding protein (TcRBP) was composed of at least six members was identified. The genomic organization of four members revealed a head to tail arrangement within a region of 15 kilobase pairs. TcRBP members have a common RRM and different auxiliary domains with a high content of glycine, glutamine, and histidine residues within their N- and C-terminal regions. TcRBPs differ in their expression patterns as well as in their homoribopolymer binding interaction in vitro, although they preferentially recognize poly(U) and poly(G) RNAs. An interesting observation was the relaxed RNA-binding interactions with several trypanosome transcripts in vitro. In contrast, co-immunoprecipitation experiments of TcRBP-containing ribonucleoprotein complexes formed in vivo revealed a highly restricted binding interaction with specific RNAs. Several TcRBPs containing complexes are stage-specific and, in some cases, bear the poly(A)-binding protein TcPABP1. Altogether, these results suggest that TcRBPs might be modulated in vivo, to favor or preclude the interaction with specific transcripts in a developmentally regulated manner.

*This work was supported by grants from the World Bank/United Nations Development Program/World Health Organization Special Program for Research and Training in Tropical Diseases and the Agencia Nacional de Promoción Científica y Tecnológica (Argentina). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AT188446 (Trbp3), AT209185 (Tcrbp4), AT209189 (Tcrbp5), AT209187 (Tcrbp6), and AT210411 (contig containing Trebp3, Trebp2, and Trebp1).

‡ Research fellow from the Comisión de Investigaciones Científicas, Provincia de Buenos Aires, Argentina.
§ Research fellow from the National Research Council (CONICET), Argentina.
¶ Researcher from CONICET, Argentina, supported in part by an International Research Scholars Grant from the Howard Hughes Medical Institute and a fellowship from the John Simon Guggenheim Memorial Foundation. To whom correspondence should be addressed: IIB-UNSAM INTI, Av. Gral. Paz s/n, Edificio 24, Casilla de Correo 30, 1650 San Martín, Provincia de Buenos Aires, Argentina. Tel.: 54-11-4580-7255; Fax: 54-11-4752-9639; E-mail: cfrasch@ibb.unsam.edu.ar.

*Trypanosoma cruzi*, the etiological agent of Chagas’ disease, is a unicellular digenetic protozoan parasite with a complex life cycle that alternates between an insect and a vertebrate host, with different replicative and infective stages in both organisms. The parasite has a wide range of mammalian hosts, from humans to wild and domestic animal species, that act as reservoirs (1). *T. cruzi* infection is established in the mammal by the insect-derived stage metacyclic trypomastigote that differentiates into the intracellular replicative amastigote stage. This parasite form differentiates to the infective trypomastigote in the mammal, which is released from cells and circulates in blood, infecting other cells and being eventually ingested by the insect with its blood meal. Ingested trypomastigotes differentiate to epimastigotes (the replicative insect stage) that migrate along the digestive tract until differentiation into metacyclic trypomastigotes that are eliminated with the feces closing the cycle. Because the parasite suffers continuous environmental changes, it needs to regulate the expression of many proteins to allow its rapid adaptation. In contrast to other eukaryotic organisms, trypanosomatids do not regulate gene expression at the classical level of transcription initiation (2), although an RNA polymerase II transcriptional complex was recently identified (3). Transcription in these organisms is polycistronic, and the main point of regulation of gene expression is at the post-transcriptional level (4).

In order to understand the post-transcriptional regulatory mechanisms, it is necessary to individualize the cis-elements within the mRNA required for this purpose. Sequences within untranslated regions (UTR)¹ of eukaryotic cells are essential for the correct expression of many genes under diverse circumstances. The 3′-UTR is not under the same selective pressure as the 5′-UTR, which possesses a necessary structural requirement to accommodate the translation machinery. Then, evolutionary pressure could take advantage of the greater degree of freedom of 3′-UTRs to control messengers stability (5) through the presence of specific regulatory elements. This portion of the transcript is involved in events such as localization, stability, transport, and translation, due to the recruitment of specific protein complexes in a precise moment of the cellular cycle and/or developmental stages (6).

Previous work from our laboratory demonstrated the presence of AU-rich elements (AREs) in the 3′-UTR of *T. cruzi* small mucin genes (Tcsmug) mRNAs. These sequences are cis-acting

Received for publication, February 19, 2003
Published, JBC Papers in Press, March 13, 2003, DOI 10.1074/jbc.M301756200

1 The abbreviations used are: UTR, untranslated region; ARE, AU-rich element; TcUBP, *T. cruzi* uridine-binding protein; TcRBP, *T. cruzi* RNA-binding protein; RRM, RNA recognition motif; RNP, ribonucleoprotein; hnRNP, heterogeneous nuclear ribonucleoprotein; GST, glutathione S-transferase; RT, reverse transcription; contig, group of overlapping clones; tssa, trypanosome small surface antigen; Tcsmug, *T. cruzi* small mucin genes; PABP, poly(A)-binding protein.
motifs that affect Tcsmug mRNA turnover in specific developmental stages (7). The results obtained suggest the existence of different trans-acting factors that could bind to Tcsmug transcripts and selectively regulate its stability throughout the parasite life cycle. Recently, we have identified a gene denominated Tcvbp1, for U-rich RNA-binding protein, that encodes a protein able to bind ARE sequences from certain mRNAs (8). A second gene, named Tcvbp2, was also shown to be present in the T. cruzi genome (9). Its product, TcUBP2, is part of an RNA-protein complex together with TcUBP1 and the polya(t)-binding protein (10).

TcUBP1 and TcUBP2 can be grouped within the proteins having an RNA-binding domain also referred to as the RNA recognition motif (RRM) (11). The RRM comprises about 90 amino acid residues whose sequence is evolutionarily conserved and also presents a characteristic motif, denominated RNP-1. This motif is an octapeptide formed by the sequence (KR/G/F/Y)(G/A)(F/V/X)(F/Y), and its presence is a strong indicator for a function in RNA recognition. A second RNP motif present within the RRM is a hexapeptide denominated RNP-2. Both RNP-1 and RNP-2 constitute the RNP consensus sequences that characterize these RRM-type proteins. The N-terminal domain of the C-terminal portion of RRM-containing proteins have, in general, auxiliary domains that are rich in glycine and charged residues as glutamine and lysine involved in protein-protein interactions (12). The combination of RRM domains with different auxiliary domains confers to these proteins a modular structure similar to the ones found in transcription factors (13). RRM-type proteins regulate gene expression by post-transcriptional mechanisms, including 1) alternative splicing events, 2) mRNA stabilization, and 3) translational control. For example, the Drosophila sex determination and maintenance pathway includes alternative splicing processes that are mediated by RRM-type proteins (11), and HuR, a member of the Hu family of RNA-binding proteins, binds AREs and is involved in selective mRNA stabilization (14). Finally, the RRM of yeast translation initiation factor Tif3 is required for translational activity in vitro (15). In other species, a large number of RNA-binding protein families were described. Examples are RRM-containing proteins in Arabidopsis thaliana, Caenorhabditis elegans, and Drosophila melanogaster (16).

In this work, we characterized an RRM-type gene family in T. cruzi, composed of at least six members that encode proteins having signature RNP motifs and clearly different auxiliary domains. We named them TcRBP, for T. cruzi RNA-binding proteins. This family includes the two previously described U-rich RNA-binding proteins TcUBP1 and TcUBP2 and four new members. TcRBPs have a higher specificity for RNA sequences in vivo than when tested in vitro, suggesting an in vivo modulation of the auxiliary domains. The developmentally regulated expression pattern of some TcRBP members along with RNP complexes composition might be related with stage-specific mRNA turnover of specific transcripts and, thus, required for the survival of the parasite under different environments.

**EXPERIMENTAL PROCEDURES**

**Parasite Cultures**—T. cruzi CL-Brener cloned stock (17) was used. Different forms of the parasites were obtained as previously described (18).

**Southern Blot Analysis**—Genomic DNA of the T. cruzi CL-Brener clone was digested with the indicated restriction enzymes (New England Biolabs). DNA fragments were separated by electrophoresis in 0.8% agarose gel, transferred by capillarity (19) to Zeta Probe Nylon-N membranes (Bio-Rad), and UV-cross-linked. Filters were hybridized with the probes described below using hybridization solution containing 0.5 M NaH2PO4, 7% SDS, 1 mM EDTA, and 1% bovine serum albumin at 62°C. Probes were made from TcRbps clones by PCR and were labeled with [α-32P]dCTP (PerkinElmer Life Sciences) using the following primers: probe RRM, with NH2-tcubp1 and RNP2; probe NH2-TcUBP2, with NH2-tcubp2 and NH2-tcubp2/2AS; probe TcRBP3, with NH2-tcrbp3 and RRM-AD/AS; probe NH2-TcUBP4, with NH2-tcrbp4 and TENG0110/AS; probe TcRBP5, with NH2-tcrbp5 and COOH-tcrbp5; and probe COOH-TcRBP6, with 5′-tcrbp6 and COOH-tcrbp6 (see Table I).

**Cloning DNA Fragments from a Cosmid Library**—Filters of a genomic DNA T. cruzi library constructed in the Lawrast 7 cosmid vector (20) were hybridized with Tcvbp1 probe in the same conditions described for Southern blot analysis. Cosmids from different positive clones were digested with several restriction enzymes, and DNA fragments were processed for Southern blot. Those bands displaying positive signals were cloned into pBS (-) vector (Stratagene) and sequenced.

**Cloning of TcRbps**—Searches in the Genome Sequence Survey and Expressed Sequence Tag T. cruzi databases were performed using the TcUBP1 RRM motif as query. Four new sequences were identified containing partially RRM motif genes: GenBank accession numbers AI0250633, AI0350500, AI0771474, and AA556992. A RT-PCR was performed with the primer oligo(dT)18-anchor (see primer sequences in Table I). antisense primers were designed to be used in a PCR with the sense TcME primer (common 39-nucleotide sequence added by transsplicing and present in all trypanosomatid transcripts) in order to identify S-ends. Full-length products were obtained by PCR using especially primed NH2-tcubp1 and anchor primers and cloned into pGEM-T Easy vector (Promega). Sequencing was performed either manually using Sequenase 2.0 (Amersham Biosciences) or by dye terminator cycle sequencing chemistry in an ABI PRISM 373 DNA Sequencer (PerkinElmer Life Sciences).

**Computer Analysis**—Computer analysis of sequences was done on http://www.ncbi.nlm.nih.gov/entrez/query.fcgi following the manufacturer’s instructions. Northern blot was carried out as described (19). Zeta Probe nylon membranes (Bio-Rad) were used for all blotting. Probe NH2-TcUBP1 was made from the Tcvbp1 clone by PCR using oligonucleotides NH2-tcubp1 and NH2-tcubp1/AS (see Table I) and was labeled with [α-32P]dCTP (PerkinElmer Life Sciences). The other probes used were previously described for Southern blotting.

**Fusion Protein Expression and Purification**—Tcubps and TcRbps overexpressing frames were amplified by PCR and cloned into the BamHI and EcoRI sites of pGEX-2T vector (Amersham Biosciences), generating a glutathione S-transferase (GST) fusion protein and translated into E. coli DH5α 1 Fq. Cultures were induced with 0.1 mM isopropyl β-d-thiogalactopyranoside for 3 h at 37°C. GST fusion proteins were purified using a glutathione-agarose column (Sigma). Dakhydrazide-Agarose RNA Cross-linking—Homoribopolymers (A, C, G, U) were oxidized with NaIO4 and cross-linked to adipic acid dihydrazide-agarose beads (Sigma) as previously indicated (22). Purified proteins were incubated with RNA-cross-linked beads for 1 h at room temperature and washed. Elution was done with 2× Laemmli buffer, and samples were resolved by electrophoresis in SDS-PAGE gels and Coomassie Blue-stained.

**Protein Extract Preparation**—Cytosolic protein extracts and subcellular fractionation were done according to previous protocols (8).

**Western Blot Analysis**—Protein samples fractionated on SDS-PAGE gels were transferred to Hybond C nitrocellulose membranes (Amer- sham Biosciences), probed with primary antibodies, and developed using an enhanced chemiluminescence-conjugated anti-mouse or anti-rabbit antibodies and Supersignal® West Pico Chemiluminescent Substrate (Pierce) according to the manufacturer’s instructions.

**Immunoprecipitations**—A cytosolic extract corresponding to 400 μg of parasite proteins was incubated with either mouse or rabbit polyclonal antibody, anti-TcUBP1, anti-TcUBP2, anti-TcRBP4, anti-TcRBP5, or anti-TcRBP6 followed by cross-linking at 4°C with gentle mixing. 50 μl of 50% protein A-Sepharose slurry were added to the mixture and incubated for 2 h at 4°C. The mixture was centrifuged at 10,000 × g for 1 min, and the wash was resuspended four times with 500 μl of Tris-buffered saline, 0.1% Tween. After that, proteins were resolved by electrophoresis in SDS-PAGE gels and analyzed by Western blotting.
were treated for co-immunoprecipitation and isolation of RNA-protein complexes or eluted with 2°C complexes or eluted with 2°C protocol described above was utilized. After the washing steps, RNA extraction from RNP complexes or eluted with 2°C was done. Using the RRM motif as query, four new sequences were obtained (Table I) and named Tcrbp3 (from T. cruzi RNA-binding protein). Tcrbp4, Tcrbp5, and Tcrbp6. Two positive cosmid DNA clones of the mentioned library were used to sequence a contig containing four of the RRM coding genes (see below).

A comparison of the sequences obtained revealed that they all have an RRM motif with a high degree of conservation that is restricted to the canonical octapeptide RNP-1 and hexapeptide RNP-2 motifs. The auxiliary domains are all different among members of the family (Fig. 1B). The UTRs are quite different among genes. On average, the 3′-UTRs only share about 25% identity, whereas in the 5′-UTR most of the sequences showed insufficient similarity for alignment (data not shown). A phenogram was inferred from a multiple alignment of the complete TcRBP protein sequences (Fig. 1C), and sequence distances between TcRBP members were calculated (Table II). This phylogenetic analysis demonstrated that the most related proteins are TcUBP1 and TcUBP2 (about 71.4% similarity), TcRBP3 has around 50% similarity with these two members. TcRBP4 presents ~48% similarity with TcUBP2 and TcRBP3 but only ~30% with TcUBP1. TcRBP5 possesses ~50% similarity with TcUBP2, TcRBP3, and TcRBP4 and 40% similarity with TcUBP1. TcRBP6, on the other hand, is the protein that mostly diverges from the rest of the TcRBP family members with a similarity index of about 10% with each of the other members (Table II).

**RRM-type Proteins Are Conserved among Different Parasites**—Several RNA-binding proteins were found in trypanosomatids and other species. Fig. 2 lists some proteins in which the RRM1 consensus sequence is present. Apart from TcRBP1, homologues from other species were found. TcRBP1 is the most related protein among members with a similarity index of about 10% with each of the other members.

### RESULTS

**TcRBPs, a Family of RRM-containing RNA-binding Proteins in T. cruzi**—In previous work, we demonstrated the existence of two RNA-binding proteins having RRMs in T. cruzi, named TcUBP1 and TcUBP2 (8, 10). In order to elucidate whether they form part of a protein family, we carried out hybridization experiments on a high density filter containing a cosmid library (20), with a probe encompassing the RRM motif of TcUBP1. This experiment revealed that Tcubp1 and Tcubp2 belong to a family of genes encoding RRM-containing proteins (data not shown). Based on the library's representation, 25 genome equivalents, and an estimated size of 55 megabase pairs per haploid T. cruzi genome, there are about 20 genes per haploid genome encoding RRM-type proteins (data not shown). In addition, a search in trypanosome data base in the Sanger Center was done. Using the RRM motif as query, four new sequences were obtained. We then used sequence information from the Expressed Sequence Tag and Genome Sequence Survey data bases and RT-PCR strategy to identify and clone TcRbpb full-length homologues. Thus, complete cDNAs encoding for all RNA-binding proteins were obtained (Fig. 1A) and named Tcrbp3 (from T. cruzi RNA-binding protein), Tcrbp4, Tcrbp5, and Tcrbp6. Two positive cosmid DNA clones of the mentioned library were used to sequence a contig containing four of the RRM coding genes (see below).

**Oligonucleotides used in this work**

| Primer            | Sequence (5′ to 3′)                       |
|-------------------|------------------------------------------|
| NH2–tcrbp1        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| RNP2              | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp2        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp2/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp3        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp3/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp4        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp4/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp6        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp6/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp6        | cgggatccATGTCTGTGGAACGGGG                |
| NH2–tcrbp6/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
| NH2–tcrbp1/AS     | TGGTGTTCTGGCTGCTGCTGAC                  |
TcRBP members are present in all life cycle stages, although not equally distributed, as illustrated in Fig. 4. The mRNAs for TcRBP members were assayed by Northern blot analysis using DNA probes previously described for Tcubp1 and Tcubp2 (8, 10). Two cosmids DNA clones picked up during the screening process on the arrayed library and were numbered 8 and 23. Both clones were digested with restriction enzymes, and fragments from 2 to 4 kbp cross-reacting with the RRM probe were cloned (see “Experimental Procedures”). The sequence of these clones allowed the identification of part of the genomic organization of three members of this family, Tcrbp3, Tcrbp2, and Tcubp1. Within a genomic fragment that comprises ~11 kb, these genes are located from 5’ to 3’ of the sequence (Fig. 3B). Interestingly, another Tcrbp member, Tcrbp4, is located next to this genomic region. This one yielded a positive signal in a hybridization experiment with the cosmid DNA clone number 8. A PCR over this clone confirms that this gene is located ~4 kbp upstream of Tcrbp3 (Fig. 3B). The contig containing Tcrbp3, Tcrbp2, and Tcubp1 genes was completely sequenced and deposited at GenBankTM (see “Experimental Procedures”).

Sequence and comparison of the contig revealed that intergenic regions are quite different, but all share the distance among them of about 3.3 kbp, with the exception of the intergenic region between Tcrbp4 and Tcrbp3 that comprises about 4 kbp as revealed by PCR experiments.

Identification of Tcrbps Transcripts Larger than the Expected Size for Mature mRNAs—We next analyzed the expression of Tcrbps mRNAs and proteins. A, scheme of Tcrbp mRNA structures. The 5′-UTR, open reading frame, 3′-UTR, and poly(A) tail are indicated. The length of each region is expressed in nucleotides (nt). B, scheme of TcRBPs showing the RRM, Gln-rich (GLY), Gly-rich (GLY), His-rich (HIS), acidic, basic, and neutral regions. RNP-1 and RNP-2 motifs within the RRM are indicated. C, phenogram generated from a multiple alignment of the six family proteins using ClustalW (21) and displayed with DRAWGRAM program. SL, mRNA 5′-spliced leader.

### Table II

| Protein   | TcUBP1 | TcUBP2 | TcRBPs | TcRBP4 | TcRBP5 | TcRBP6 |
|-----------|--------|--------|--------|--------|--------|--------|
| TcUBP1   | 100.0  | 71.4   | 42.9   | 32.5   | 39.8   | 12.4   |
| TcUBP2   | 100.0  | 50.5   | 48.3   | 58.4   | 12.8   |        |
| TcRBPs  | 100.0  | 47.4   | 48.3   | 13.6   |        |        |
| TcRBP4  | 100.0  | 53.5   | 13.2   |        |        |        |
| TcRBP5  |        |        |        |        |        | 8.7    |

previously described for Tcubp1 and Tcubp2 (8, 10). Two cosmid DNA clones picked up during the screening process on the arrayed library and were numbered 8 and 23. Both clones were digested with restriction enzymes, and fragments from 2 to 4 kbp cross-reacting with the RRM probe were cloned (see “Experimental Procedures”). The sequence of these clones allowed the identification of part of the genomic organization of three members of this family, Tcrbp3, Tcrbp2, and Tcubp1. Within a genomic fragment that comprises ~11 kb, these genes are located from 5’ to 3’ of the sequence (Fig. 3B). Interestingly, another Tcrbp member, Tcrbp4, is located next to this genomic region. This one yielded a positive signal in a hybridization experiment with the cosmid DNA clone number 8. A PCR over this clone confirms that this gene is located ~4 kbp upstream of Tcrbp3 (Fig. 3B). The contig containing Tcrbp3, Tcrbp2, and Tcubp1 genes was completely sequenced and deposited at GenBankTM (see “Experimental Procedures”).

Sequence and comparison of the contig revealed that intergenic regions are quite different, but all share the distance among them of about 3.3 kbp, with the exception of the intergenic region between Tcrbp4 and Tcrbp3 that comprises about 4 kbp as revealed by PCR experiments.

Identification of Tcrbps Transcripts Larger than the Expected Size for Mature mRNAs—We next analyzed the expression of Tcrbps mRNAs and proteins. A, scheme of Tcrbp mRNA structures. The 5′-UTR, open reading frame, 3′-UTR, and poly(A) tail are indicated. The length of each region is expressed in nucleotides (nt). B, scheme of TcRBPs showing the RRM, Gln-rich (GLY), Gly-rich (GLY), His-rich (HIS), acidic, basic, and neutral regions. RNP-1 and RNP-2 motifs within the RRM are indicated. C, phenogram generated from a multiple alignment of the six family proteins using ClustalW (21) and displayed with DRAWGRAM program. SL, mRNA 5′-spliced leader.

### Table II

| Protein   | TcUBP1 | TcUBP2 | TcRBPs | TcRBP4 | TcRBP5 | TcRBP6 |
|-----------|--------|--------|--------|--------|--------|--------|
| TcUBP1   | 100.0  | 71.4   | 42.9   | 32.5   | 39.8   | 12.4   |
| TcUBP2   | 100.0  | 50.5   | 48.3   | 58.4   | 12.8   |        |
| TcRBPs  | 100.0  | 47.4   | 48.3   | 13.6   |        |        |
| TcRBP4  | 100.0  | 53.5   | 13.2   |        |        |        |
| TcRBP5  |        |        |        |        |        | 8.7    |

the mRNA steady-state levels are different between stages in the case of Tcrbp6 (Fig. 4). In most of the cases, the size of the band detected in Northern blots is larger than the one deduced from cDNA sequencing. For Tcubp1 and Tcubp2, the Northern blot shows band of about 5 kb. For Tcrbp3, a transcript of the size of the mature mRNA (~1 kb) was observed (see the arrow in Fig. 4), but larger and more abundant RNA bands were also detectable. Tcrbp4 probe also revealed larger bands than expected, with a stronger signal of about 2.4 kb. Tcrbp5 yielded bands of 1.5 kb that duplicate the expected size of the mature mRNA. Finally, Tcrbp6 showed a single band greater than 2.5 kb preferentially present in trypomastigote and amastigote stages. Similar observations were obtained for other transcripts of regulatory proteins (e.g. TcPABP1 among others) (25).3 We conclude that, with the exception of Tcrbp6, Tcrbp mRNA steady-state levels are quite similar between different developmental stages of the parasite.

TcRBPs Have a Common RRM and Different Auxiliary Domains—TcRBPs present a modular structure. They have a single conserved RNA-binding domain and one or two auxiliary domains present in the 5′- or 3′-ends from the protein (see Fig. 1B). These N- and C-terminal regions present a marked deviation in sequence composition. Auxiliary domains of this type are believed to be involved in protein-protein interaction and therefore present an important functional relevance. TcUBP1 and TcUBP2 present Gln-rich regions in their N-terminal ends and Gly-rich regions in their C-terminal ends. TcUBP1 also possesses a Gln-rich random coil region in the C terminus (8). TcRBP3 presents an acidic N-terminal region of 67 amino acids, due to the high content of glutamic acid, whereas in the C-terminal portion it has a basic domain of 40 residues, where the more distinctive are neutral and basic amino acids, such as alanine and arginine respectively. TcRBP4 presents a neutral N-terminal region of 38 amino acids with glycine, serine, and isoleucine interspersed between basic residues such as arginine. TcRBP5 has an RNA-binding domain with a basic 32-amino acid C-terminal extreme. Finally, TcRBP6 has a large N-terminal auxiliary domain with two clearly distinctive regions.
substrate specificity, we performed RNA-binding reactions between RRM and auxiliary domains, evidenced by the sequence comparison (Fig. 5), might result in different RNA binding specificities.

To further analyze whether the difference in binding specificity might contain a short amphiphobic region and a His-rich region (Figs. 1B and 5A). The phylogenetic tree was displayed by DRAGRAM software.

If we look at the RRM, it can be appreciated that the RNP signature motifs are highly conserved among proteins. RNP-1 and -2 motifs are located in the β2- and β1-sheets, respectively, and they participate in the direct contact with RNA (26). The variable region of the RRM domain, mainly the amino acids present in the loops of the β-sheets, might give specificity for RNA recognition (11). It is in these regions where a greater variability is observed in the RRM, suggesting that these RNA-binding proteins might have different specificities for RNA targets (see below). Secondary structures were obtained automatically using the PSIPred protein structure prediction server (available on the World Wide Web at bioinf.cs.ucl.ac.uk/psipred). The secondary consensus motifs shared between all proteins were displayed together with multiple sequence alignments (Fig. 5A).

Theoretical models of three-dimensional structure of TcRBP family members were obtained after submitting the sequences to the automated homology-modeling server from Swiss Model. This preliminary information suggests that the overall conformation of all TcRBPs members is similar (Fig. 5B). Their structure seems also similar to those of other RRM-containing RNA-binding proteins, such as Sex lethal from D. melanogaster (27), murine hnRNP A1 (28), and PABP from Saccharomyces cerevisiae (29). TcRBP models are even similar to theoretical predictions of both RRMs of PABP from D. melanogaster (29). TcRBP models are even similar to those of other RRM-containing RNA-binding proteins, such as Sex lethal from D. melanogaster (27), murine hnRNP A1 (28), and PABP from Saccharomyces cerevisiae (29). TcRBP models are even similar to theoretical predictions of both RRMs of PABP from D. melanogaster (29).

To sum up, all TcRBP family members have a central core with four antiparallel β-sheets packed against the two hydrophobic α-helices at correct angle orientation. Although computer modeling of the auxiliary domains does not predict accurate secondary structures, the C-terminal region of TcRBP3 might contain a short α-helix, a structure known to have a function in protein-protein interactions (31).

Characterization of TcRBP RNA-binding Properties Using Homoribopolymers—To further analyze whether the difference between RRM and auxiliary domains, evidenced by the sequence comparison (Fig. 5), might result in different RNA substrate specificity, we performed RNA-binding reactions in vitro. For this purpose, all GST fusion proteins were incubated separately with homoribopolymers (poly(A), poly(C), poly(G), and poly(U)), and their capacity to recognize them were tested (Fig. 6). The results showed that, although the binding specificities to homoribopolymers are not equal among proteins, they are relatively similar. TcUBP1 and TcUBP2 showed similar specificity and preferentially bound poly(U) and poly(G), poly(U) being better recognized than poly(G) in the case of TcUBP1 (Fig. 6). TcRBP3 displayed a different recognition pattern, having specificity for poly(A), poly(C), and poly(G). TcRBP4 and TcRBP5 had a pattern similar to that observed with TcUBP1, the first one having greater affinity to poly(G) homoribopolymer and the second to poly(U). TcRBP6 bound the four ribopolymers with an increased specificity for poly(G) (Fig. 6).

The previous results indicated that TcRBPs indeed have RNA binding activity and that, although they possess different specificities, they mostly share the capabilities to bind poly(G) and poly(U) sequences, except for TcRBP3 (Fig. 6).

Protein Expression and Subcellular Localization—Western blots were performed in order to characterize the subcellular localization and stage of expression of TcRBPs. These experiments revealed that TcUBP1 is a cytoplasmic protein of 27 kDa expressed in all of the stages of T. cruzi. TcUBP2 is also a cytoplasmic protein of 18 kDa preferentially expressed in epimastigote and, at a lower level, in the amastigote stage (Fig. 7 and 8) (see also Refs. 8 and 10). TcRBP3 is expressed in the epimastigote stage and presents an apparent molecular mass of 22 kDa (Fig. 7). A subcellular fractionation revealed that it is a cytoplasmic protein, like the other TcUBPs mentioned before (Fig. 8). For the case of TcRBP4, TcRBP5, and TcRBP6, several bands can be observed. Thus, immunoprecipitation assays were performed to facilitate the identification of proteins. TcRBP4 is expressed in epimastigote, where three bands were detectable. The protein is the 32-kDa band, since it was also immunoprecipitated with anti-TcRBP4 antibodies (Fig. 7). Likewise, this band was observed in the cytosolic fraction (Fig. 8). Since the amino acid sequence indicates that the protein has a molecular mass of 16 kDa, a possible explanation for the observed size could be found in the fact that a larger protein might be produced from the larger mRNA band identified in the Northern blot analysis (see Fig. 4). Anti-TcRBP5 antibodies detected a 29-kDa band in all trypanosome stages (Fig. 7), together with a 14-kDa band in the cytoplasmic fraction (Fig. 8). Our experiments demonstrated that the 29-kDa band, which is about twice the size of the one expected, can be immunoprecipitated using anti-TcRBP5 antibodies (Fig. 7). Finally, using anti-
TcRBP6 antibodies in Western blot experiments, we detected a 24-kDa band in the cytosolic fraction (Fig. 8) and in all the stages of the parasite life cycle (Fig. 7). This band was also immunoprecipitated using anti-TcRBP6 antibodies (Fig. 7). The slight difference observed in size after comparison with the expected molecular mass (about 5 kDa) might be due to its particular amino acid composition within the N-terminal region (see Fig. 5A).

The 45-kDa protein band observed with different antibodies in some of the Western blot experiments might correspond to another member of the RRM family not identified yet, which might have homology within the RRM motifs of TcUBP1, TcRBP4, and TcRBP6 proteins. In summary, TcRBPs are mainly cytoplasmic proteins, which differ in the life cycle stages of the parasite in which they are expressed but share the same localization within the cell (Table III).

TcRBPs Have the Potentiality to Recognize Different Trypanosome Transcripts in Vitro—To test whether these proteins can interact with different T. cruzi transcripts in vitro, mRNA binding assays were performed (Fig. 9). RNA poly(A)+ was purified and incubated with GST (as negative control) or GST-TcRBP immobilized on glutathione-agarose matrix. Eluted mRNAs were reverse-transcribed, and the cDNA obtained was tested in PCRs with primers for several T. cruzi mRNAs bearing AU- or GU-rich sequences within their 3′-UTRs. The transcripts tested were as follows: Tcsmug, trypanosome small surface antigen (tssa), tryparedoxin, cruzipain, epimastigote trans-sialidase, vacuolar-type proton-translocating pyrophosphatase 1, RGG-containing RNA-binding protein, 30-kDa translation elongation factor, T. cruzi glycoprotein of 63 kDa, and amastin (Fig. 9A).

Tcsmug mRNA was used as a positive control from the experiment, because its ARE sequence is a site for TcUBP1 recognition in vivo (8). First, Tcsmug, tssa, and cruzipain yielded positive results with all TcRBPs (Fig. 9A). Second, RGG-containing RNA-binding protein and tryparedoxin gave positive results only with TcUBP2 and TcRBP5 proteins. Third, the 30-kDa translation elongation factor was recognized by TcUBP2, TcRBP4, TcRBP5, and TcRBP6. Fourth, PCR products from vacuolar-type proton-translocating pyrophosphatase 1 were only detected with TcUBP2 protein, and amastin product was only obtained with TcUBP1. Finally, epimastigote trans-sialidase as well as T. cruzi glycoprotein of 63 kDa gave negative results in all cases tested (Fig. 9A). The results are summarized in the chart of Fig. 9B and suggest that most recombinant TcRBPs have a relaxed binding specificity for a wide variety of ARE-containing mRNAs in vitro.

TcRBPs Have Restricted Binding Capabilities for RNA Recognition in Vivo—We next analyzed which of the RNAs recognized in vitro by TcRBPs are forming part of in vivo RNP

Fig. 3. Southern blot analysis and genomic organization of Tcrbp family members. A, genomic DNA was digested with the indicated restriction enzymes and hybridized at high stringency conditions with the probe that is shown below each panel. The molecular weight markers, λ DNA digested with HindIII (Invitrogen), and a 100-bp or 1-kb DNA ladder (New England Biolabs) are indicated at the sides. B, scheme of the genomic organization of four Tcrbp genes. The open reading frames are indicated as white boxes along with the corresponding portions of UTR, polyadenylation site (pA), and intergenic region. The length of each fragment is expressed in kbp. H, HindIII endonuclease site.
complexes with specific members of this protein family. For this purpose, RT-PCR was performed from total anti-TcRBP immunoprecipitates and, as an internal control, from 5% of total RNA extracted from epimastigote stage. For all mRNAs species tested, a PCR product of the expected size was observed in the controls (data not shown).

The findings indicate that the RNA-binding properties observed in vivo are different from the ones detected in vitro when using poly(A)⁺ RNA extracted from epimastigotes and recombinant GST fusion proteins (Table IV). In conclusion, TcUBP1, TcUBP2, and TcRBP3 did not seem to interact with cruzipain mRNA in vivo, although the recognition is detected
in vitro. In contrast, the other mRNAs that are recognized in vitro (Tesiumg, tssa, and amastin) are efficiently co-immunoprecipitated from RNP complexes containing TcUBP1 (Fig. 10 and Table IV). The fact that TcUBP1 interacts with Tesiumg mRNA both in vitro and in vivo confirms our previous results (8) (see “Discussion”). Although TcUBP2 interacts with many mRNAs in vitro, it showed a restricted pattern of mRNA recognition in vivo, since it is exclusively bound to Tesiumg. In the case of TcRBP3, it did not interact with either of the mRNAs tested in vivo (Table IV). Co-immunoprecipitation experiments performed with anti-TcRBP4, anti-TcRBP5, and anti-TcRBP6 did not give any positive RNA interactions in vivo (data not shown).

TcUBP1 and TcUBP2 proteins have the potentiality to interact with AU- and GU-rich elements present in a great variety of trypanosome transcripts in vitro (10). However, it can be envisaged that not all of these transcripts are bound in vivo (Table IV). All together, these results suggest that in vivo modulation might be an important step in the regulation of RNA-protein interactions in trypanosomes (see “Discussion”).

Developmentally Regulated TcRBP-containing RNP Complexes—In a second approach, the in vivo co-immunoprecipitated RNAs with anti-TcUBP1, anti-TcUBP2, and anti-TcRBP1 antibodies were compared using two of the most different life cycle stages of the parasite, epimastigotes and cell-derived tryptomastigotes (Fig. 10). First, to check the RNA extraction protocol, a 5% total protein extract was used to isolate total RNA for RT-PCR analysis. Second, RT-PCR was performed using co-immunoprecipitated RNAs. Only mRNAs that gave positive results (Tesiumg, tssa, and amastin) are shown in Fig. 10.

It was previously demonstrated in our laboratory that TcUBP1 interacts in vivo with other RRM-type proteins, forming part of a ~450-kDa RNP complex only in the epimastigote stage. These proteins are TcUBP2 and TcPABP1 (10). Since this RNP complex might play a stabilizing effect on Tesiumg mRNA in epimastigotes, we further analyzed which of the proteins of this complex bound Tesiumg, tssa, and/or amastin mRNAs. This approach might give us a clue to determine whether these different protein factors make up distinct RNP complexes on specific RNA target species and in the two different life cycle stages.

A comparison of the results obtained in vivo showed some interesting results. First, Tesiumg mRNA interacts with TcUBP1, TcUBP2, and also TcPABP1, in the epimastigote but not in the tryptomastigote stage (Fig. 10A). Second, tssa mRNA also interacts with TcUBP1, maintaining the same stage-specific pattern as Tesiumg mRNA. In addition, tssa interacts with TcPABP1 in both life cycle stages (Fig. 10A). It is important to note that in these particular cases both Tesiumg and tssa are present in different parasite stages, but it is only in the epimastigote stage when they are interacting in these RBP complexes. Third, the amastin mRNA was co-immunoprecipitated with anti-TcUBP1 and anti-TcPABP1 antibodies in both parasite stages (Fig. 10A). Finally, RT-PCR of anti-TcRBP3-immunoprecipitated RNAs did not give any positive signal in both parasite stages (data not shown). The relationship among RNA-protein interactions, mRNA stabilization, and protein expression, as compared with the results from previous works (10, 32, 33), is shown in Fig. 10B and discussed below (see “Discussion”).

**DISCUSSION**

In this work, we described a novel T. cruzi RRM-type RNA-binding protein family named TcRBP that is composed of at least six members that share similar primary structures (Figs. 1 and 5). Four TcRbp members are clustered within a 15-kbp genomic arrangement, with their coding regions separated by around 3–3.5 kbp of intergenic regions (Fig. 3B) and possibly...
involved in the synthesis of common larger transcripts. We report here that both Tcubp2 and Tcubp1 members are present in the same transcription unit (Fig. 4) and probably form part of the same stable dicistronic pre-mRNA. Clusters of co-expressed genes were identified in both lower and higher eukaryote cells (34). Moreover, in the particular case of lower eukaryotes, gene clusters or operons that share certain aspects of transcriptional regulation were found in pairs (35, 36). Many C. elegans genes exist in operons in which polycistronic precursors are processed by cleavage at the 3′-ends of upstream genes and trans-splicing 100–400 nucleotides away, at the 5′-ends of downstream genes, to generate monocistronic messages (37). Operons were first described in prokaryotes. In these organisms, genes needed for one particular process are often clustered in close proximity on the genome, with the same orientation on the DNA. These operons facilitate the coordinated regulation of genes, since the clustered genes are transcribed in a single step (38).

RRM-type proteins are conserved among different protozoan parasites. Recently, a TcUBP1 homologue was identified in T. brucei, and several clones similar to TcRBPs were also found by data base searches in other protozoan parasites (see Fig. 2). A phylogenetic analysis showed that TcRBPs are closely related among trypanosomatids in comparison with other organisms. In addition, TcUBP1, TcUBP2, and TcRBP3 proteins present the greatest similarities among proteins in the family, forming a separate phylogenetic group (Fig. 1). Overall, this evolutionary study suggests that a Tcrbp gene, probably Tcubp1, originated from a common ancestor in all trypanosomatids. Furthermore, it is possible that a gene duplication event originated both Tcubp1 and Tcubp2 before the speciation process (see Fig. 2). Interestingly, the completion of the Plasmodium genome permits us to reveal quite a few genes encoding predicted proteins that might exert a role in controlling gene transcription (39). This suggests that also in this organism the main point of gene expression regulation might be achieved through mRNA processing.

All TcRBP family members have a common protein structure and similar predicted three-dimensional folding (Fig. 5). More specifically, the RRM motif is the principal conserved region, and, particularly, the RNP-1 domain is the most similar portion of the molecule. Conversely, the auxiliary domains located

---

**Table III**

| RNA-binding Protein Family in T. cruzi |
|---------------------------------------|
| E, epimastigote stage; T, trypomastigote stage; A, amastigote stage; C, cytoplasm; a, adenosine; c, cytosine; g, guanosine; u, uridine. |
| E, T, A | E, T, A | E, T, A | E, T, A | E, T, A | E, T, A |
| RNA (stage of expression) | E, T, A | E, T, A | E, T, A | E, T, A | E, T, A | E, T, A |
| Stage of expression | E, T, A | E, T, A | E, T, A | E, T, A | E, T, A | E, T, A |
| Subcellular localization | C | C | C | C | C | C |
| Molecular mass (kDa) | 27 | 18 | 22 | 32 | 29 | 24 |
| Ribopolymer affinity | g/u | g/u | a/c/g | g/u | g/u | a/c/g/u |

**Fig. 8.** Subcellular localization of TcRBPs. Equal amounts of nuclear (N) and cytoplasmic (C) fractions of total protein extracts prepared from epimastigote stage were layered on each lane (see left panel, Coomassie staining gel). The blots were reacted with rabbit polyclonal antibodies anti-RRM, -TcRBP3, -TcRBP4, -TcRBP5, and -TcRBP6 as indicated below each panel. TcRBPs are indicated by arrows. The bands indicated with asterisks are nonspecific. H3 and H4, nuclear histones that cross-react with some of the antiserum used. Molecular mass protein standards Dalton Mark VII-L™ are indicated at the left.
in the N- and/or C-terminal regions of the RRM are markedly distinct among these proteins, suggesting their involvement in different RNA regulatory processes (see below). TcUBP1 and TcUBP2 have Gln-rich and Gly-rich regions suspected to act as protein-interacting modules. Moreover, the Gly-rich region was demonstrated to be involved in homo- and heterodimerization in both TcUBPs (10). Although TcRBPs mainly differ in their auxiliary domains, their RNA-binding properties to RNA homoribopolymers in vitro are quite comparable (Fig. 6). Thus, these findings reinforce the idea that such proteins might suffer in vivo modulation, enabling the direction of a specific RNA target recognition.

TcRBPs expression patterns revealed that all family members are cytoplasmic, and some of them, such as TcUBP2, TcRBP3, TcRBP4, and TcRBP5, are developmentally regulated proteins (Figs. 7 and 8). It is important to note that recombinant proteins produced in bacterial systems have indiscriminate binding properties to several transcripts having AU- and/or GU-rich elements within their 3’-UTRs, such as Tcsmug, tssa, amastin, and cruzipain (7, 32, 33, 40), among others. However, this observation is in contrast with the restricted mRNA species that interact with TcRBPs in vivo (Table IV). Supporting this, other well known RBPs such as HuR or hnRNP D that are structurally and functionally different factors that recognize similar ARE-containing sequences in vitro can discriminate among different classes of ARE se-

### Table IV

Comparison between in vitro and in vivo RNA-protein interactions

| mRNA   | Tcsmug | tssa  | Tryparedoxin | Cruzipain | eTS   | PPase 1 | RGG-RBP | 30-kDa EF | Tcgp63 | Amastin |
|--------|--------|-------|--------------|-----------|-------|---------|---------|-----------|--------|---------|
| **In vitro** | TcUBP1 | +     | +            | +         | +     | +       | -       | -         | +      | +       |
|        | TcUBP2 | +     | +            | +         | +     | +       | -       | -         | +      | +       |
|        | TcRBP3 | +     | +            | -         | +     | -       | -       | -         | -      | -       |
| **In vivo** | TcUBP1 | +     | +            | -         | +     | -       | -       | -         | -      | -       |
|        | TcUBP2 | +     | +            | -         | +     | -       | -       | -         | -      | -       |
|        | TcRBP3 | -     | -            | -         | -     | -       | -       | -         | -      | -       |

**Fig. 10.** *In vivo* RNA-binding interactions take place in a stage-specific manner. A, comparison of the RNA interactions detected in different *T. cruzi* stages. Agarose gels from RT-PCRs of RNA co-immunoprecipitated *in vivo* with the antibody raised against the proteins mentioned at the left of each panel. Genes tested in PCR are indicated below the panels. A reverse transcription experiment was performed with (+) or without (−) SuperScript II enzyme. RNA, reverse transcription performed with total RNA from the indicated stage; IP, reverse transcription performed with RNA co-precipitated with the antibody raised against the proteins mentioned at the left of each panel; E, epimastigote stage; T, trypomastigote stage. B, scheme of RNA-protein interactions *in vivo* comparing epimastigote and trypomastigote stages. The presence of TcUBP1, TcUBP2, and TcPABP1 binding (or not) to Tcsmug, tssa, and amastin RNA messengers is indicated. The RNP complexes and messengers might be associated with other factors not yet identified. Stage-regulated protein expression of the target mRNA is shown according to bibliographical reports. SL, spliced leader.
quences in vivo (41, 42). Thus, modulation of RNA recognition by TcRBPs in vivo might be critical for the proper development and differentiation of the parasite.

TcUBP1, the first RBP identified in our laboratory (8), is a developmentally regulatory protein recognizing Tcsmug mRNA in vivo. Although TcUBP1 is expressed in all parasite life cycle stages, it interacts with Tcsmug mRNA only in the epimastigote stage and not in the trypomastigote stage (Fig. 10). Thus, the presence of such a protein and an mRNA in a developmental stage is not synonymous of RNA-binding interaction between them. The same stage-specific Tcsmug mRNA-binding pattern was also observed with TcUBP2 and TcPABP1. Thus, all together these results suggest that TcUBP1, TcUBP2, and TcPABP1 interact with Tcsmug mRNA in the epimastigote and not in the trypomastigote stage (see Fig. 10), allowing a correct expression pattern of mucin genes during the parasite life cycle. Likewise, TcUBP1 and TcPABP1 bind amasin mRNA in both epimastigotes and trypomastigotes (Fig. 10). This suggests that TcUBP1 could be acting in the formation of distinct RNP complexes in different developmental stages, in some cases interacting with an mRNA in a particular stage (such as Tcsmug in epimastigotes) while in other situations recognizing mRNAs in more than one parasite stage (such as amasin in both epimastigotes and trypomastigotes). It was previously demonstrated that certain factors present in the T. cruzi amastigote stage, like 30- and 36-kDa proteins, might cause up-regulation of amasin mRNA abundance (32). However, only the 30-kDa factor interacts with the amasin 3′-UTR element in epimastigotes (32). Since TcUBP1 binds this messenger in both epimastigote and trypomastigote stage, it might be tested whether it plays a destabilizing role in these particular stages (Fig. 10B). TcUBP1 also binds tssa mRNA in a stage-specific manner, and this interaction was only observed in epimastigotes. Due to TSSA expression displayed only in the trypomastigote stage (33), TcUBP1 could play a destabilizing role in this particular case, where TSSA is not detected. TcPABP1, on the other hand, recognized tssa mRNA in both epimastigotes and trypomastigotes (Fig. 10). Bandinius et al., 38 residual protein composition for the different TcRBP-containing RNP complexes and also RNP remodeling during parasite differentiation is a challenge to be solved. For this purpose, genomic arrays of endogenous RNP complexes, combined with proteomics and mass spectrometry-assisted analyses, are excellent tools.

Acknowledgments—We are indebted to Fabio Fraga for animal care, Berta Franke de Cazzulo and Liliana Sferco for parasite cultures, and Alejandro D’Antunno for helpful discussion in the phylogenetic analysis. We are also thankful to Drs. Juan J. Cazzulo and G. Getz for critical reading of the manuscript.

REFERENCES
1. Brener, Z. (1973) Annu. Rev. Microbiol. 27, 347–382
2. Lee, M. G., and Van der Ploeg, L. H. (1997) Annu. Rev. Microbiol. 51, 463–489
3. Das, A., and Bellofatto, V. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 80–85
4. Clayton, C. E. (2002) EMBO J. 21, 1881–1888
5. Conne, S., Stutz, A., and Vassalli, J. D. (2000) Nat. Med. 6, 637–641
6. Good, P. J., Chen, Q., Warner, S. J., and Herring, D. C. (2000) J. Biol. Chem. 275, 28563–28572
7. Di Noia, J. M., D’Orosio, I., Sanchez, D. O., and Frasch, A. C. (2000) J. Biol. Chem. 275, 10218–10227
8. D’Orosio, I., and Frasch, A. C. (2001) J. Biol. Chem. 276, 34801–34809
9. D’Orosio, I., and Frasch, A. C. (2001) J. Biol. Chem. 276, 15783–15793
10. D’Orosio, I., and Frasch, A. C. (2002) J. Biol. Chem. 277, 50520–50528
11. Bard, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
12. Bandinius, R. J., Swanson, M. S., and Dreyfuss, G. (1989) Genes Dev. 3, 431–437
13. Latchman, D. S. (1986) in Eukaryotic Transcription Factors, 3rd ed., pp. 271–321. Academic Press, Cambridge
14. Brennan, C., and Steitz, J. (2001) Cell Mol. Life Sci. 58, 266–277
15. Niederberger, N., Trachsel, H., and Altmann, M. (1998) RNA 4, 1259–1267
16. Larkovits, N., and Barski, A. (2002) Nucleic Acids Res. 30, 623–635
17. Zonghaling, B., Pereira, M. E., Oliveira, R. P., Almeida, I. C., and Frasch, A. C. (2000) EMBO J. 19, 4860–4979
18. Rivas, R. J., Pallevick, G. D., Xavier, M. T., Prevato, J. O., Mendoca-Prevato, L., Sanchez, D. O., and Frasch, A. C. (1996) J. Biol. Chem. 271, 32078–32083
19. Schambach, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Hanke, J., Sanchez, D. O., Henriksson, L., Aslund, L., Pettersson, U., Frasch, A. C., and Hoheisel, J. D. (1996) BioTechniques 21, 689–695
21. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
22. Capote, M., Mayed, A., Krainer, A. R., and Zahler, A. M. (1999) EMBO J. 18, 4060–4067
23. Li, J., Kinoshita, T., Pandey, S., Ng, C. K., Gherzi, R., Ong, S. E., Pruijn, G. J., and Shyu, A. B. (2002) Mol. Cell. Biol. 22, 845–854
24. Vermel, M., Guermann, B., Delage, L., Grienengerger, J. M., Marechal-Drouard, L., and Gualberto, J. M. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 5866–5871
25. Batista, J. A., Teixeira, S. M., Donelson, J. E., Kirchhoff, L. V., and de Sa, C. M. (1997) Mol. Biochem. Parasitol. 83, 261–267
26. Perez-Canadillas, J. M., and Varani, G. (2001) Curr. Opin. Struct. Biol. 11, 53–58
27. Handa, N., Nureki, O., Kuriyama, M., and Yokoyama, S. (1999) Nature 398, 579–585
28. Shames, Y., Krueger, U., Rice, L. M., Williams, K. R., and Steitz, T. A. (1997) Nature 384, 675–680
29. Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999) Cell 98, 835–845
30. Silva Junior, F. P., Veyl, F. Z., Chen, F., and De Simone, S. G. (2002) Mem. Inst. Oswaldo Cruz 97, 335–341
31. Branden, C. J., and Tooze, J. (1999) Introduction to Protein Structure, Garland Publishing, Inc., New York.
32. Cazzulo, J. J., Stoka, V., and Turk, V. (1997) EMBO J. 16, 5866–5871
33. Chen, C. Y., Xu, N., and Shyu, A. B. (2002) Mol. Cell. Biol. 22, 463–468
34. Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Raijmakers, R., Pruijn, G. J., and Shyu, A. B. (2002) Mol. Cell. Biol. 22, 797–804
35. Chen, C. Y., Xu, N., and Shyu, A. B. (2002) Mol. Cell. Biol. 22, 797–7778
RNA Recognition Motif-type RNA-binding Proteins in *Trypanosoma cruzi* Form a Family Involved in the Interaction with Specific Transcripts *in Vivo*

Javier G. De Gaudenzi, Iván D'Orso and Alberto C. C. Frasch

*J. Biol. Chem.* 2003, 278:18884-18894.
doi: 10.1074/jbc.M301756200 originally published online March 13, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301756200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 17 of which can be accessed free at http://www.jbc.org/content/278/21/18884.full.html#ref-list-1