

# Functional Genomics and Insights into *Trypanosoma cruzi* Gene Expression Regulation

Andréa Rodrigues Ávila\* and Samuel Goldenberg\*

Instituto Carlos Chagas – ICC, Fiocruz-Paraná, Rua Prof. Algacyr Munhoz Mader 3775, Curitiba, PR, Brazil

**Abstract:** The regulation of gene expression in trypanosomatids is predominantly post-transcriptional. Polycistronic transcripts are processed by the addition of a common 5'-spliced leader and polyadenylation. However, the processed mRNAs are not necessarily functionally related, suggesting the existence of mechanisms for the degradation or storage of untranslatable mRNAs. Determination of the TriTryps (*Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*) genome sequences has allowed the identification of genes encoding potential regulatory proteins. This review discusses some of the mechanisms and regulatory elements involved in cytoplasmic gene expression regulation in *Trypanosoma cruzi*. We also discuss how functional genomic tools have contributed toward determining the role played by RNA binding protein complexes, supporting the concept of "post-transcriptional RNA operons" or "RNA regulons". This suggests the existence of interconnected regulatory networks in the parasite, in which RNA granules act as protagonists in cytoplasmic mRNA metabolism.

**Keywords:** *Trypanosoma cruzi*, RNA binding protein, gene expression regulation, functional genomics.

## TRANSCRIPTION OF MESSENGER RNA

*Trypanosoma cruzi* (*T. cruzi*) is the protozoan parasite causing Chagas disease, a major disease endemic to Latin America. It belongs to the Kinetoplastida order and has a complex life cycle, alternating between insect vectors and mammalian hosts. In both hosts, *T. cruzi* goes through morphological and functional changes, creating non-infective and infective forms. Those changes resulted from selective adaptation and are the major consequence of differential gene expression. Being a single cell, it needs to quickly regulate the synthesis of several proteins for rapid adaptation to a different environment. Its protein synthesis has particular characteristics in comparison with higher eukaryotes. For example, the protein-coding genes are organized into polycistronic transcription units; however, in contrast to what occurs in bacterial operons, the polycistronic units must be transcriptionally processed together before translation. In this case, polycistronic RNAs are converted into monocistronic messenger RNAs (mRNAs) through two coupled events: *trans*-splicing and polyadenylation. During *trans*-splicing, a 39-nt spliced leader (SL) RNA molecule is joined to the 5' terminus of the mature transcript, while a poly(A) tail is polymerized at the 3' terminus. The same polycistronic RNA can contain functionally unrelated genes, which yield very different steady-state mRNA levels and/or which are differentially expressed during the life cycle [see review in [1]].

Another trypanosome transcription characteristic is the absence of defined RNA polymerase II (pol II) promoters. There are a few exceptions, however, such as the genes

encoding heat shock protein 70 [2] in *Trypanosoma brucei* (*T. brucei*) and gene19 encoding GARP in *Trypanosoma congolense* [3]. It seems that the genes encoding spliced leader RNA molecules (SLRNA) in *T. brucei* are transcribed by RNA pol II; however, the promoter elements have different sequences to those related to the basal transcription elements identified in other eukaryotic genes [4]. As yet, no promoters have been defined for *T. cruzi* coding-genes, even though transcription is modulated by RNA pol II during its life cycle [5]. Interestingly, it has been described that histone modifications (acetylation and methylation) might play a role in defining transcription initiation sites in trypanosomatids indicating the importance of epigenetic processes in gene expression regulation [6]. Previous findings suggest that RNA pol II machinery controls overall transcription, but it is clear that polycistronic transcriptional organization prevents the independent control of individual genes. Thus, it is very likely that the differential expression of most (if not all) coding genes in *T. cruzi* is mostly mediated by post-transcriptional mechanisms.

## DIFFERENTIAL EXPRESSION OF CODING GENES DURING LIFE-CYCLE

Much of the evidence for post-transcriptional control of gene expression come from studies on genes differentially expressed during the *T. cruzi* life cycle; the characterization of these stage-specific genes has been the major goal of several research groups. The identification of stage-specific genes has been used to identify virulence factors [7] and to elucidate differences in drug-resistance among various parasite strains [8]. At the same time, it has been useful to obtain data on post-transcriptional regulation in this parasite. Various groups have shown that the accumulation of mRNA during a specific stage of *T. cruzi* development is due to differential mRNA stability and not to increased transcription of the gene [9-14]. Some developmental forms of the parasite

\*Address correspondence to these authors at the Instituto Carlos Chagas – ICC, Fiocruz-Paraná, Rua Prof. Algacyr Munhoz Mader 3775, Curitiba, PR, Brazil; Tel: +55-41-33163230; Fax: +55-41-33163267; E-mails: sgoldenb@fiocruz.br, aravila@fiocruz.br

have similar steady-state mature transcript levels; however, differences in protein levels suggest that protein production is regulated at the translational level [15]. Other findings suggest that gene expression is dependent on the movement of mRNA to the polysomes [16] or mechanisms that involve the blocking of translation elongation and/or termination, as suggested by studies on rRNA processome proteins [17].

The investigation of factors involved in the regulation of mRNA stability or translation revealed that elements present in mRNA untranslated regions (UTR) are responsible for differential gene expression. Data relating to these elements were obtained mostly using reporter gene approaches [18]. For example, amastin steady state mRNA levels are 68 times more abundant in amastigotes than in epimastigotes; moreover, luciferase activity is 7-fold greater in amastigotes and 5-fold lower in epimastigotes if the 3'-UTR of amastin mRNA is inserted downstream from the luciferase coding region in a chimeric plasmid for transient transfections [19]. Subsequent studies have identified a 630-nt *cis*-element in the 3' UTR responsible for amastin mRNA up-regulation [10].

Another example of 3' UTR and coding region elements that may be involved in modulating the stability of mRNAs relates to the surface antigen FL-160 gene of *T. cruzi*. Steady-state mRNA levels of this gene are 80- to 100-fold higher in trypomastigotes than in epimastigotes, even though transcription rates are equivalent between the life cycle stages. The presence of the 3' UTR and the intergenic region (IR) following the coding sequence of the FL-160 gene was sufficient to cause an 8-fold increase in luciferase production in trypomastigotes in comparison with epimastigotes [20]. In addition, recent results with tubulin genes indicate that sequences located within the 3'-UTR and coding region of the alpha-tubulin gene are involved in modulating the stability of these transcripts, in response to changes in the dynamics of *T. cruzi* microtubules [13].

Some elements may also control both stability and the translation of mRNAs, and may be responsible for controlling mRNA stability and improving translational efficiency. These elements correspond to AU-rich sequences, and they are similar to ARE motifs known to cause mRNA instability in higher eukaryotes by interacting with specific RNA-Binding Proteins (RBP) [21, 22]. Elements from the 5' UTR may also be involved in gene expression control. Analyses of Tuzin 5' UTR mRNA indicate that elements present in this region post-transcriptionally suppress the expression of this gene in *T. cruzi*, by both inefficient RNA processing and poor translation initiation [9].

The mechanisms involving mRNA degradation/stabilization seem to be the main control features during *T. cruzi* gene expression. In fact, active deadenylation systems have been found in trypanosome cells [23,24], including the description of the enzymes triggering mRNA degradation. The mechanism involves mRNA degradation from both ends, by XRN1-related exoribonucleases (5'-3' direction) and the exosome (3'-5' direction), after removal of the poly(A) tail and the 5' Cap [25,26]. However, data related to the exosome machinery were reported for *T. brucei* alone. This mechanism may occur in *T. cruzi*. Accordingly, genes have been annotated in the *T. cruzi* Gene Database as components of the exosome complex. Yet, most findings

point to *cis*-elements in the 5' and 3' untranslated regions or *trans*-elements, such as RNA-binding proteins (more details are described below), that together with mRNAs form messenger-ribonucleoprotein (mRNP) complexes directing post-transcriptional regulation.

## GENOMIC AND POST-GENOMIC APPROACHES

The WHO Special programme for Research and Training in Tropical Diseases (TDR) launched a consortium in 1994 to analyze the genomes of parasites, such as filarial nematodes, *Schistosoma*, *Leishmania*, *Trypanosoma brucei* and *Trypanosoma cruzi*. Since the publication of the first complete sequence of a chromosome from a kinetoplastid parasite [27], enormous progresses have been made on trypanosomatids genomics. The whole *T. cruzi* genome sequence was published in 2005 and revealed that the diploid genome contains 22,570 gene-encoded predicted proteins, of which 12,570 represent allelic pairs [28]. Genome assembly consisted of 5489 scaffolds (containing 8740 contigs) totaling 67 Mb. Based on the assembly results, the *T. cruzi* diploid genome size was estimated to be between 106.4 and 110.7 Mb, which is larger than the previous estimate (87 Mb) based on pulse-field gel-electrophoresis combined with densitometric analysis of chromosomal DNA. At least 50% of the *T. cruzi* genome is composed of repetitive sequences, consisting mostly of large gene families of surface proteins, retrotransposons, and subtelomeric repeats.

Publication of the three TriTryp genomes (*T. cruzi*, *T. brucei*, and *L. major*) in 2005 was also an important step in advancing our understanding of other related kinetoplastids. However, a published genome is not necessarily a complete genome. Incompleteness within the TriTryp genomes is particularly problematic with regards to *T. cruzi*, which was published without a complete chromosome. This results from its highly repetitive nature and the fact that the reference strain (CL Brener) is a hybrid of two different *T. cruzi* lineages. The *T. brucei* and *L. major* genomes projects were much more successful, providing not only nearly complete chromosomes but also synteny maps linking most of the core (and conserved) genes in the *T. cruzi* genome to *T. brucei* and *L. major* chromosomes [29,30]. However recently, most contigs and scaffolds assemble into pairs of homologous chromosomes based on predicted parental haplotype, inferences from TriTryp synteny maps, and the use of end sequences from *T. cruzi* BAC libraries. The validity of the process was supported experimentally via Southern blot analysis of a set of chromosomes, whereby individual genes are used as probes to confirm the predicted organization of each chromosome [31]. Thus, the assembled chromosomes provide a physical platform for studying gene function and gene function variation in *T. cruzi*. This is particularly useful for planning and confirming gene knockouts, thus determining the function of hypothetical genes or confirming the function of annotated genes. The assembled chromosomes, including the mapped BAC isolates used to construct them, are available on TriTrypDB, a trypanosome database of the EuPathDB organization [Eukaryotic Pathogen Database Resources [http://eupathdb.org/]].

Notwithstanding the limitations on the annotation of *T. cruzi* genome, it has been possible to gain further insight into genome organization or gene structure by *in silico* analysis

using sequences of the TriTryp genomes. In particular, a super-family of more than 100 RNA recognition motif-type (RRM-type) RNA-binding proteins was discovered in the *T. cruzi* genome [32]. Some of these proteins are involved in alternative splicing processes, mRNA stabilization/degradation, polyadenylation or translational control. However, most do not have a clear homolog in other species, even though they are highly conserved in kinetoplastids, demonstrating the presence of specific factors for gene expression regulation in this order. Comparative analysis allowed the identification of alternative sequences for RNA polymerase subunits that are likely to be important features in the unusual function of the trypanosome RNA polymerases [33].

Other studies have analyzed the similarities and differences in the composition of TriTryp genomes in further detail. A comparative study of codon/amino acid usage in TriTryp revealed that gene expressivity and GC-bias play key roles in shaping the genetic composition of all three parasites [34]. Another study showed that the sequence patterns related to base J within subtelomeric regions are developmentally regulated in *T. cruzi* [35] and similar to those in *T. brucei*, in which base J has a role in regulating telomeric genes involved in pathogenesis. The analysis of DNA sequences immunoprecipitated with J antiserum in this study aided in localizing base J within the subtelomeric regions; these regions are rich in stage-specific surface glycoprotein genes involved in the pathogenesis of this parasite.

By contrast, comparative analysis of the sequences and genomic organization of tRNA genes and other Pol III-transcribed genes showed largely weak synteny in TriTryp genomes, indicating that the location of Pol III genes in the genome was less important than the location of protein-coding genes. Some of the differences between isodecoder genes occur in the internal promoter elements suggesting that the expression of some isoacceptor tRNA genes in Trityps is differentially controlled [36].

*In silico* analysis has also been very useful at identifying regulator elements present in trypanosome RNAs. A genome-wide analysis of RNA processing signals in *T. brucei* and *Leishmania* [37] showed that, in *T. brucei*, trans-splicing reactions generally occur at the first AG dinucleotide after an 8- to 25-nt poly-pyrimidine (Y) sequence, generating mRNAs with 5' UTRs of an average length of 68 nucleotides. In addition, it was shown that polyadenylation occurs at positions with one or more adenine residues located between 80 and 140 nt from a downstream poly(Y) motif. This results in 3' UTRs as short as 21 nt or up to 5040 nt long, with the median length of 3' UTRs being 348 nt. These results were then used to develop an algorithm that enabled the prediction of *trans*-splicing and polyadenylation sites for most protein coding genes in *T. brucei*. A similar whole genome screening process was performed to identify the sequences involved in mRNA processing in *T. cruzi*. Using DNA sequences derived from the *T. cruzi* genome [28] and all available expressed sequence tags (ESTs) and cDNAs containing poly(A) tails and/or spliced leader (SL) sequences, [38] uncovered a pattern for mRNA processing in this parasite and compared the elements involved in RNA processing events in *T. cruzi* and *T. brucei*. They determined

that the median lengths of 5' UTR and 3' UTR sequences are 35 and 264 nucleotides, respectively. They also reported that the median distance between SL addition sites and a poly-pyrimidine motif is 18 nucleotides, whereas the median distance between poly(A) addition sites and the closest poly-pyrimidine-rich sequence is 40 nucleotides. This type of *in silico* prediction of *T. cruzi* UTRs should allow the rapid identification of a large number of putative *cis*-regulatory elements, whose role in modulating gene expression could be experimentally validated. In addition, the information provided in this study may also contribute to studies on the optimization of vectors currently used to manipulate gene expression in this parasite.

The information provided by the genome network has also improved studies on target identification and the functional characterization of regulatory proteins (*trans*-elements). In this case, TcUBP1 and TcRBP3, members of the RRM family, showed that various auxiliary domains display different RNA binding characteristics [39]. Noé and co-workers systematically identified cellular groups of targets for each protein and reported the structural RNA elements conserved in each group of transcripts. These signature RNA motifs were then used to predict novel TcUBP1 and TcRBP3 target mRNAs from the trypanosome database [39].

The *T. cruzi* genome project also provided several challenges for those studying kinetoplastids. The obvious challenge is whether to investigate overall gene function or whether to move more progressively and selectively. The *in silico* transcriptome has benefitted from the development of new technologies, which reveal overall gene expression patterns, providing further insight into gene expression regulation. One example of this kind of technology is microarray analysis. The first description of a DNA microarray system in *T. cruzi* was for a group of partially sequenced *T. cruzi* isolates [40]. This study investigated overall gene expression profiles during trypomastigote to amastigote transformation. In this study, PCR product probes, selected specifically from *T. cruzi* open reading frames (ORF) and randomly from genomic survey sequencing (GSS) libraries, were competitively hybridized with fluorescent cDNA targets generated from total RNA samples of trypomastigotes and transitional parasites. Although the authors observed few changes in mRNA abundance, they report on a number of genes that were up-regulated in amastigotes. They basically showed that the *T. cruzi* microarray system is a valuable resource for future studies on the regulation of gene expression in this parasite. More recently, Mining *et al.* used a whole genome oligonucleotide microarray to determine the overall extent to which *T. cruzi* regulates the relative abundance of mRNA over the course of its complete life-cycle [41]. In contrast to previous microarray studies on kinetoplastids, transcript abundance for over 50% of the genes detected on the *T. cruzi* microarrays was significantly regulated during the *T. cruzi* life-cycle, supporting the conclusion that transcript abundance is an important step of gene expression regulation in this parasite.

Undoubtedly, microarrays seem to be valuable screening tools for identifying stage-regulated *T. cruzi* genes and metabolic pathways. They are useful in the analysis of gene expression of isolates presenting high genetic diversity [42].

The DNA microarray based on CL Brener expressed sequence tags (ESTs) was shown to be appropriate for investigating differential gene expression in parasite strains. It also provided evidence for post-transcriptional regulation, including evidence for significant changes to mRNA abundance in *T. cruzi*. More recently, a DNA microarray was used to compare transcript profiles of six human isolates [43]: three from asymptomatic and three from cardiac patients. NADH dehydrogenase subunit 7 (ND7) and other electron transport pathway genes are thought to be valuable targets for PCR assays, particularly in diagnosing differential expression in infective strains. Thus, this represents a feasible strategy for identifying specific molecular markers that will be used in the association between parasite genetics and disease pathology. Note that array technologies have also been important in providing further insight into relevant aspects of *trans*-elements, particularly the role of RNA-binding proteins.

### RBP COMPLEXES AND GENE EXPRESSION REGULATION

The bidirectional polycistronic transcription units in trypanosomatids, different to those found in bacteria, are formed by genes not necessarily functionally related, highlighting the clear need for a post-transcriptional regulatory system allowing the expression of stage-specific genes based on the cell-cycle. There is now a general consensus that proteins mediate the regulatory pathways. The degradation or accumulation of mature mRNAs in the nucleus may be a possibility, although supporting evidence has yet to be presented. Alternatively, it may be that all processed transcripts are transported to the cytoplasm, to undergo fine regulation. As described previously, most data indicate that mRNA

turnover and translation are the major regulation control points in trypanosomatids. Several studies using chimeras of reporter genes, including 5'- and 3'-UTR regions, have shown the role these sequences play in rendering a given mRNA stable [10,44,45]. However, in addition to nucleotide sequences, proteins specifically expressed during a given developmental stage play a role in the stabilization process.

The TriTryp genome project [28,29,46] allowed, in addition to the characterization of regulatory sequences and gene distribution along the genome, the possibility of identifying potential genes based on their homology to other organisms through *in silico* searches. The data presented in Table 1 list some RNA-binding proteins described in *T. cruzi*, but there are many other proteins or hypothetical proteins that have been described as RBPs in trypanosomatids [32].

The characterization of RBPs should lead to the determination of their role in gene expression regulation. In organisms with a relaxed transcription and little evidence of significant nuclear processed mRNA turnover, it is very likely that most of the processes regulating gene expression occur in the cytoplasm. Upon reaching the cytoplasm, mRNAs can be directly translated, degraded or stored as free-mRNP and thus do not undergo translation. The amount of evidence supporting the presence of stable non-translated mRNAs in the cytoplasm is increasing. Some of the first evidence supporting the existence of a stable pool of non-translated RNA in *T. cruzi* came from studies showing the presence of mRNAs not associated with polysomes that could be actively translated in a rabbit reticulocyte *in vitro* system [47]. These results were further corroborated by northern blot analysis of stage-specific genes comparing various developmental stages of *T. cruzi*; no differences were observed after com-

Table 1.

RNA Binding Protein	RNA Binding Motif	Description	Reference
TcRBP 28	ND	Reacts with antibodies from chagasic patients .	[71]
XB1	ND	Homology with yeast splicing factor PRP31p. Interacts with SL RNA <i>in vitro</i> .	[72]
TcPUF6	Pumilio	Member of PUF Family. Binds specifically to transcripts up-regulated in meta-cyclic trypomastigotes.	[51]
TcRBP1-6	RRM	Characterized as member of the <i>T. cruzi</i> RRM family. Has a common RRM and different auxiliary domains.	[73]
TcPUF1	Pumilio	Member of PUF family. Interacts with Nanos response elements (NRE). The targets include genes for mitochondrial proteins and protein kinases.	[50]
TcUBP1	RRM	Involved in mRNA destabilization <i>in vivo</i> through binding to AU-rich elements at the 3'-UTR of SMUG mucin mRNAs. Part of complex containing TcPUF2 and TcPABP1.	[22]
TcUBP2	RRM	Part of a ribonucleoprotein complex containing TcPABP1 and TcUBP1.	[74]
TcRBP3	RRM	Share target transcripts with TcUBP1, although binds to specific sets of mRNA.	[39]
TcRBP19	RRM	Shows target selectivity because binds preferentially to polyC.	[75]
TcRRM1 e TcRRM2	RRM	Similar to Tbp34 and Tbp37 of <i>Trypanosoma brucei</i> , respectively.	[76]
TcZFP1	Zinger finger (CCCH)	Recognizes C-rich ribopolymers and seems to be involved in modulating stage-specific mRNA decay pathways and accessibility to the translation machinery.	[57]

ND – Not assigned.

paring total RNA populations, whereas expression level differences were evident after polysomal RNAs were compared [12,16].

The recent development of functional genomics tools has led to a wider understanding of the role played by individual RBP and their association with other RBPs in the regulation of gene expression. The use of microarray analysis, next generation sequencing and high-throughput proteomics have strengthened ribonomic approaches, in which a given RBP can be biologically characterized through the identification of its molecular targets. This has led to the concept of “post-transcriptional RNA operons” or “RNA regulons” [48], i.e. trans-acting factors (proteins and eventually small RNAs) associate to coordinate the expression of multiple mRNAs. This may be a suitable model for gene expression regulation in trypanosomatids, as polycistronic transcription does not allow the temporal co-expression of functionally related genes.

In fact, there are some recent examples of RBPs that may act as post-transcriptional operons upon binding to specific mRNAs in *T. cruzi*. An example was reported during the characterization of members from the Puf (Pumilio) family, which have been extensively studied in trypanosomatids [49-52]. TcPUF6 over-expression in *T. cruzi* and a microarray approach allowed a comparative analysis of RNA from transfected and control parasites; this study showed that TcPUF6 had a significant effect on overall gene expression. Interestingly, some of the affected genes encode RNA-binding proteins, various kinases, and proteases, pointing to the existence of complex and interconnected post-transcriptional and post-translational regulatory networks in *T. cruzi* [51]. A given RBP may also display a different biological function depending upon the proteins associated in a mRNP complex. Accordingly, this study also suggested that interactions between TcPUF6 and TcDhh1, a DEAD-box helicase protein associated with P-bodies in *T. cruzi* [53], result in the destabilization and hence degradation of mRNAs in replicative epimastigote forms. By contrast, the absence of interaction between TcPUF6 and TcDhh1 stabilizes associated mRNAs in metacyclic trypomastigotes.

PUF proteins in *T. brucei* are also involved in controlling mRNA stability during the cell cycle: moreover, TbPUF9 stabilizes target transcripts during the S-phase and RNAi against PUF9 caused an accumulation of cells in the G2/M phase and destabilized PUF9 target mRNAs [52]. The presence of RNA regulons in trypanosomes [39,51,54] possibly compensates for relaxed transcription, providing a fine tuning mechanism required for gene expression regulation in eukaryotic organisms.

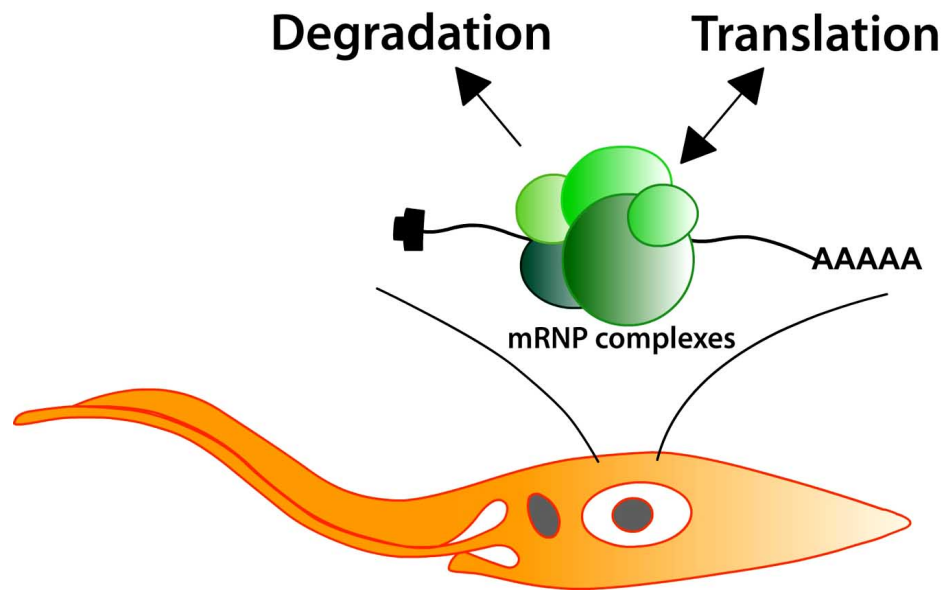
The CCCH zinc finger family is another important class of RBPs that are well characterized in trypanosomes. These proteins were first characterized in African trypanosomes, in which they were implicated in the control of differentiation to the procyclic form [55]. TriTryp genome sequencing identified several genes that potentially code for this type of protein. Several members of these small zinc finger proteins have been described in *T. cruzi* [56] and one of the proteins (named TcZFP1), homologous to a protein previously described in *T. brucei*, binds to C-rich sequences [57]. Considering that C-rich sequences are present in the 3'-UTR of sev-

eral mRNAs, this protein may be involved in regulating mRNA stability. Indeed, TbZFP3 was recently shown to regulate a procyclin isoform in *T. brucei* [58]. Interestingly, another CCCH zinc finger protein (TcZFP2) co-immunoprecipitated with particular mRNAs from *T. cruzi*, the amounts of which were greater in infective metacyclic trypomastigote forms of the parasite (Morking *et al.*, in preparation).

The mRNA molecules associate with the RBPs as they are being transcribed, forming mRNP complexes (Fig. 1). The associated proteins ultimately determine the fate of the mRNA and this association is very dynamic, i.e. the type of protein involved may change in response to various stimuli, such as stressors. These various mRNP complexes are defined based on their function or according to the presence of specific proteins in a given complex. On reaching the cytoplasm, mRNPs are either translated (polysomal mRNPs) or sequestered to the cytoplasm, after which the mRNAs are either degraded or stored as cytoplasmic-free mRNP. These free mRNPs then form microscopic foci, which may differ in function based on associated proteins or on the stimuli during their creation. One of these particles are P-bodies, which are generally associated with mRNA degradation [59]. Untranslated mRNAs can also form cytoplasmic granules known as stress-granules [60,61]. Stress-granules are usually associated with mRNAs, for which translation stalled, due to physiological stress. Thus, some initiation factors and ribosomal proteins, may be associated with these granules.

As mentioned previously, there is evidence for the existence of a pool of non-translated mRNAs in *T. cruzi* and, more recently, the presence of foci relating to P-body-like structures in the parasite [53]. Recent studies have shown that different mRNAs may be isolated from cytoplasmic granules immunoprecipitated with an antiserum against TcDhh1, a characteristic P-body protein [Holetz *et al.*, submitted]. The presence of RNP granules has also been demonstrated in *T. brucei* [62] and *T. cruzi* [63]. Findings related to *T. cruzi* showed that the amount of granules were greater in epimastigote forms starved of carbon source. Interestingly, granule protein content characterization by immunomicroscopy showed that the granules represent different RNP complexes and do not necessarily co-localize. Future studies should include a complete characterization of these RNA granules in trypanosomatids because of their importance in gene expression regulation. Proteomic and microarray analysis should help decipher RNA granule organization. It is very likely that granules with different mRNA and protein content will be found, reflecting the diversity of RNA regulons found in trypanosomes. This might shed light on the mechanisms involved in the hitherto not well-understood cytoplasmic mRNA cycle.

Another important step in gene expression regulation in eukaryotes is the transport of processed mRNAs from the nucleus to the cytoplasm. However, this pathway remains to be elucidated in trypanosomes. In higher eukaryotes, various RNA species and/or bulk mRNA subsets are exported by separated pathways, in which Exportin -1 (CRM1 in yeast) and TAP (Mex67p in yeast) [64,65] are key protagonists in transport pathways involving various proteins [for review



**Fig. (1).** The fate of the mRNA in the cytoplasm depends upon the associated proteins. Trypanosomatid mRNAs are transcribed as polycistronic transcripts and processed in the nucleus. mRNAs are transported to the cytoplasm associated to proteins and the combination of associated proteins will determine whether the mRNAs will be mobilized to the polysomes and translated. Alternatively mRNAs can be stored as free-cytoplasmic mRNPs either in equilibrium with the translated mRNA population or repressed for translation. However mRNAs can also associate to specific sets of proteins that will determine their degradation. It is important noting that these mRNP complexes are very dynamic and mRNAs can change from one compartment to another according to the associated proteins. One of the greatest challenges in the near future will be to dissect mRNP complexes and determine the synergistic interactions of mRNP proteins. This might give rise to a new genetic code: the RBP code.

see Köhler and Hurt, 2007 [66]. Cuevas and co-workers, 2005 [67] identified a CRM1 homolog in *T. cruzi* and suggested its involvement in the transport of a poly(A) +RNA subset across the nuclear pore complex (NPC) of trypanosomes. Moreover, an RBP involved in regulating the turnover of mRNAs encoding membrane proteins in *T. brucei* was also detected in the nucleus [54]. Interestingly, a RRM-type RBP from *T. cruzi* involved in mRNA stabilization has also been found in the nucleus following stress and has been suggested to be involved in the shuttling of some mRNAs from the nucleus to the cytoplasm [68].

Understanding the mechanisms involved in gene expression regulation in trypanosomes will be improved by determining the epigenetic mechanisms involved in the process (recently reviewed by [77, 78]) and by the use of functional genomics tools. Microarray analysis and new generation sequencing should make the identification of mRNAs associated with specific RNA binding proteins easier. Conversely, mass spectrometry analysis of mRNP complexes should allow the identification of all proteins associated with a given type of mRNA. Probing cells for sites of mRNA production may also foster our understanding of all aspects of gene expression regulation. Processes by which mRNAs are transported to the various cellular compartments is well documented in eukaryotes [69,70] and these processes may be important in trypanosomes as additional steps in cytoplasmic gene expression regulation.

#### ACKNOWLEDGEMENTS

We thank Itamar Crispim for the illustration. SG has a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil).

#### REFERENCES

- [1] Clayton C, Shapira M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* 200; 156: 93-101.
- [2] Lee MG. Targeting of exogenous DNA into *Trypanosoma brucei* requires a high degree of homology between donor and target DNA. *Mol Biochem Parasitol* 1996; 76: 215-29.
- [3] Graham SV, Jefferies D, Barry JD. A promoter directing alpha-amanitin-sensitive transcription of GARP, the major surface antigen of insect stage *Trypanosoma congolense*. *Nucleic Acids Res* 1996; 24: 272-81.
- [4] Gilinger G, Bellofatto V. Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. *Nucleic Acids Res* 2001; 29: 1556-64.
- [5] Elias MC, Marques-Porto R, Freymuller E, Schenkman S. Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. *Mol Biochem Parasitol* 2001; 112: 79-90.
- [6] Respuela P, Ferella M, Rada-Iglesias A, Aslund L. Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi*. *J Biol Chem* 2008; 283 (23): 15884-92.
- [7] Risso MG, Garbarino GB, Mocetti E, *et al.* Differential expression of a virulence factor, the trans-sialidase, by the main *Trypanosoma cruzi* phylogenetic lineages. *J Infect Dis* 2004; 189(12): 2250-9.
- [8] Villarreal D, Nirde P, Hide M, Barnabe C, Tibayrenc M. Differential gene expression in benzimidazole-resistant *Trypanosoma cruzi* parasites. *Antimicrob Agents Chemother* 2005; 49(7): 2701-9.
- [9] Teixeira SM, Kirchhoff LV, Donelson JE. *Trypanosoma cruzi*: suppression of tuzin gene expression by its 5'-UTR and spliced leader addition site. *Exp Parasitol* 1999; 93(3): 143-51.
- [10] Coughlin BC, Teixeira SM, Kirchhoff LV, Donelson JE. Amastin mRNA abundance in *Trypanosoma cruzi* is controlled by a 3'-untranslated region position-dependent cis-element and an untranslated region-binding protein. *J Biol Chem* 2000; 275(16): 12051-60.

- [11] Bartholomeu DC, Silva RA, Galvao LM, *et al.* *Trypanosoma cruzi*: RNA structure and post-transcriptional control of tubulin gene expression. *Exp Parasitol* 2002; 102(3-4): 123-33.
- [12] Yamada-Ogatta SF, Motta MC, Toma HK, *et al.* *Trypanosoma cruzi*: cloning and characterization of two genes whose expression is up-regulated in metacyclic trypomastigotes. *Acta Trop* 2004; 90(2): 171-9.
- [13] da Silva RA, Bartholomeu DC, Teixeira SM. Control mechanisms of tubulin gene expression in *Trypanosoma cruzi*. *Int J Parasitol* 2006; 36(1): 87-96.
- [14] Gentil LG, Cordero EM, do Carmo MS, dos Santos MR, da Silveira JF. Posttranscriptional mechanisms involved in the control of expression of the stage-specific GP82 surface glycoprotein in *Trypanosoma cruzi*. *Acta Trop* 2009; 109(2): 152-8.
- [15] Thomas MC, Garcia-Perez JL, Alonso C, Lopez MC. Molecular characterization of KMP11 from *Trypanosoma cruzi*: a cytoskeleton-associated protein regulated at the translational level. *DNA Cell Biol* 2000; 19(1): 47-57.
- [16] Avila AR, Yamada-Ogatta SF, da Silva Monteiro V, *et al.* Cloning and characterization of the metacyclogenin gene, which is specifically expressed during *Trypanosoma cruzi* metacyclogenesis. *Mol Biochem Parasitol* 2001; 117(2): 169-77.
- [17] Nardelli SC, Avila AR, Freund A, *et al.* Small-Subunit rRNA processome proteins are translationally regulated during differentiation of *Trypanosoma cruzi*. *Eukaryot Cell* 2007; 6(2): 337-45.
- [18] Teixeira SM, daRocha WD. Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res* 2003; 2(1): 148-58.
- [19] Teixeira SM, Kirchoff LV, Donelson JE. Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. *J Biol Chem* 1995; 270(38): 22586-94.
- [20] Weston D, La Flamme AC, Van Voorhis WC. Expression of *Trypanosoma cruzi* surface antigen FL-160 is controlled by elements in the 3' untranslated, the 3' intergenic, and the coding regions. *Mol Biochem Parasitol* 1999; 102(1): 53-66.
- [21] Di Noia JM, D'Orso I, Sanchez DO, Frasch AC. AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *J Biol Chem* 2000; 275(14): 10218-27.
- [22] D'Orso I, Frasch AC. TeUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. *J Biol Chem* 2001; 276(37): 34801-9.
- [23] Milone J, Wilusz J, Bellofatto V. Characterization of deadenylation in trypanosome extracts and its inhibition by poly(A)-binding protein Pab1p. *RNA* 2004; 10(3): 448-57.
- [24] Schwede A, Ellis L, Luther J, *et al.* A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res* 2008; 36(10): 3374-88.
- [25] Li CH, Irmer H, Gudjonsdottir-Planck D, *et al.* Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation. *RNA* 2006; 12(12): 2171-86.
- [26] Estevez AM, Kempf T, Clayton C. The exosome of *Trypanosoma brucei*. *EMBO J* 2001; 20(14): 3831-9.
- [27] Myler PJ, Audleman L, deVos T, *et al.* *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc Natl Acad Sci USA* 1999; 96(6): 2902-6.
- [28] El-Sayed NM, Myler PJ, Bartholomeu DC, *et al.* The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 2005; 309(5733): 409-15.
- [29] Berriman M, Ghedin E, Hertz-Fowler C, *et al.* The genome of the African trypanosome *Trypanosoma brucei*. *Science* 2005; 309(5733): 416-22.
- [30] El-Sayed NM, Myler PJ, Blandin G, *et al.* Comparative genomics of trypanosomatid parasitic protozoa. *Science* 2005; 309(5733): 404-9.
- [31] Weatherly DB, Boehlke C, Tarleton RL. Chromosome level assembly of the hybrid *Trypanosoma cruzi* genome. *BMC Genomics* 2009; 10: 255.
- [32] De Gaudenzi J, Frasch AC, Clayton C. RNA-binding domain proteins in Kinetoplastids: a comparative analysis. *Eukaryot Cell* 2005; 4(12): 2106-14.
- [33] Kelly S, Wickstead B, Gull K. An in silico analysis of trypanosomatid RNA polymerases: insights into their unusual transcription. *Biochem Soc Trans* 2005; 33(6): 1435-7.
- [34] Chanda I, Pan A, Saha SK, Dutta C. Comparative codon and amino acid composition analysis of Tritryps-conspicuous features of *Leishmania major*. *FEBS Lett* 2007; 581(30): 5751-8.
- [35] Ekanayake DK, Cipriano MJ, Sabatini R. Telomeric co-localization of the modified base J and contingency genes in the protozoan parasite *Trypanosoma cruzi*. *Nucleic Acids Res* 2007; 35(19): 6367-77.
- [36] Padilla-Mejia NE, Florencio-Martinez LE, Figueroa-Angulo EE, *et al.* Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC Genomics* 2009; 10: 232.
- [37] Benz C, Nilsson D, Andersson B, Clayton C, Guilbree DL. Messenger RNA processing sites in *Trypanosoma brucei*. *Mol Biochem Parasitol* 2005; 143(2): 125-34.
- [38] Campos PC, Bartholomeu DC, DaRocha WD, Cerqueira GC, Teixeira SM. Sequences involved in mRNA processing in *Trypanosoma cruzi*. *Int J Parasitol* 2008; 38(12): 1383-9.
- [39] Noe G, De Gaudenzi JG, Frasch AC. Functionally related transcripts have common RNA motifs for specific RNA-binding proteins in trypanosomes. *BMC Mol Biol* 2008; 9: 107.
- [40] Minning TA, Bua J, Garcia GA, McGraw RA, Tarleton RL. Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2003; 131: 55-64.
- [41] Minning TA, Weatherly DB, Atwood J 3rd, Orlando R, Tarleton RL. The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. *BMC Genomics* 2009; 10: 370.
- [42] Baptista CS, Vencio RZ, Abdala S, *et al.* DNA microarrays for comparative genomics and analysis of gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2004; 138: 183-94.
- [43] Baptista CS, Vencio RZ, Abdala S, *et al.* Differential transcription profiles in *Trypanosoma cruzi* associated with clinical forms of Chagas disease: Maxicircle NADH dehydrogenase subunit 7 gene truncation in asymptomatic patient isolates. *Mol Biochem Parasitol* 2006; 150(2): 236-48.
- [44] Nozaki T, Cross GA. Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 1995; 75: 55-67.
- [45] D'Orso I, De Gaudenzi JG, Frasch AC. RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol* 2003; 19(4): 151-5.
- [46] Ivens AC, Peacock CS, Worthey EA, *et al.* The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 2005; 309: 436-42.
- [47] Goldenberg S, Salles JM, Contreras VT, *et al.* Characterization of messenger RNA from epimastigotes and metacyclic trypomastigotes of *Trypanosoma cruzi*. *FEBS Lett* 1985; 180(2): 265-70.
- [48] Keene JD, Lager PJ. Post-transcriptional operons and regulons coordinating gene expression. *Chromosome Res* 2005; 13(3): 327-37.
- [49] Hoek M, Zanders T, Cross GA. *Trypanosoma brucei* expression-site-associated-gene-8 protein interacts with a Pumilio family protein. *Mol Biochem Parasitol* 2002; 120: 269-83.
- [50] Caro F, Bercovich N, Atorrasagasti C, Levin MJ, Vazquez MP. *Trypanosoma cruzi*: analysis of the complete PUF RNA-binding protein family. *Exp Parasitol* 2006; 113(2): 112-24.
- [51] Dallagiovanna B, Correa A, Probst CM, *et al.* Functional genomic characterization of mRNAs associated with TcPUF6, a pumilio-like protein from *Trypanosoma cruzi*. *J Biol Chem* 2008; 283(13): 8266-73.
- [52] Archer SK, Luu VD, de Queiroz RA, Brems S, Clayton C. *Trypanosoma brucei* PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. *PLoS Pathog* 2009; 5(8): e1000565.
- [53] Holecz FB, Correa A, Avila AR, Nakamura CV, Krieger MA, Goldenberg S. Evidence of P-body-like structures in *Trypanosoma cruzi*. *Biochem Biophys Res Commun* 2007; 356(4): 1062-7.
- [54] Estevez AM. The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes. *Nucleic Acids Res* 2008; 36(14): 4573-86.
- [55] Hendriks EF, Robinson DR, Hinkins M, Matthews KR. A novel CCHC protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. *EMBO J* 2001; 20(23): 6700-11.

- [56] Caro F, Bercovich N, Atorrasagasti C, Levin MJ, Vazquez MP. Protein interactions within the TcZFP zinc finger family members of *Trypanosoma cruzi*: implications for their functions. *Biochem Biophys Res Commun* 2005; 333(3): 1017-25.
- [57] Morking PA, Dallagiovanna BM, Foti L, *et al.* TcZFP1: a CCCH zinc finger protein of *Trypanosoma cruzi* that binds poly-C oligoribonucleotides *in vitro*. *Biochem Biophys Res Commun* 2004; 319(1): 169-77.
- [58] Walrad P, Paterou A, Acosta-Serrano A, Matthews KR. Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA cis-elements. *PLoS Pathog* 2009; 5(2): e1000317.
- [59] Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. *Mol Cell* 2007; 25(5): 635-46.
- [60] Anderson P, Kedersha N. Stress granules: the Tao of RNA triage. *Trends Biochem Sci* 2008; 33(3): 141-50.
- [61] Anderson P, Kedersha N. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* 2009; 10(6): 430-6.
- [62] Kramer S, Queiroz R, Ellis L, *et al.* Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169. *J Cell Sci* 2008; 121(18): 3002-14.
- [63] Cassola A, De Gaudenzi JG, Frasch AC. Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Mol Microbiol* 2007; 65(3): 655-70.
- [64] Strasser K, Hurt E. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* 2000; 19(3): 410-20.
- [65] Hurt E, Strasser K, Segref A, *et al.* Mex67p mediates nuclear export of a variety of RNA polymerase II transcripts. *J Biol Chem* 2000; 275(12): 8361-8.
- [66] Kohler A, Hurt E. Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* 2007; 8(10): 761-73.
- [67] Cuevas IC, Frasch AC, D'Orso I. Insights into a CRM1-mediated RNA-nuclear export pathway in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2005; 139(1): 15-24.
- [68] Cassola A, Frasch AC. An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein. *J Biol Chem* 2009; 284(50): 35015-28.
- [69] Paquin N, Chartrand P. Local regulation of mRNA translation: new insights from the bud. *Trends Cell Biol* 2008; 18(3): 105-11.
- [70] Martin KC, Ephrussi A. mRNA localization: gene expression in the spatial dimension. *Cell* 2009; 136(4): 719-30.
- [71] Pais FS, DaRocha WD, Almeida RM, *et al.* Molecular characterization of ribonucleoprotein antigens containing repeated amino acid sequences from *Trypanosoma cruzi*. *Microbes Infect* 2008; 10(7): 716-25.
- [72] Xu P, Wen L, Benegal G, Wang X, Buck GA. Identification of a spliced leader RNA binding protein from *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2001; 112(1): 39-49.
- [73] De Gaudenzi JG, D'Orso I, Frasch AC. RNA recognition motif-type RNA-binding proteins in *Trypanosoma cruzi* form a family involved in the interaction with specific transcripts *in vivo*. *J Biol Chem* 2003; 278(21): 18884-94.
- [74] D'Orso I, Frasch AC. TcUBP-1, an mRNA destabilizing factor from trypanosomes, homodimerizes and interacts with novel AU-rich element- and Poly(A)-binding proteins forming a ribonucleoprotein complex. *J Biol Chem* 2002; 277(52): 50520-8.
- [75] Perez-Diaz L, Duhagon MA, Smircich P, *et al.* *Trypanosoma cruzi*: Molecular characterization of an RNA binding protein differentially expressed in the parasite life cycle. *Exp Parasitol* 2007; 117(1): 99-105.
- [76] Gomes GG, Peter Urmenyi T, Rondinelli E, Williams N, Silva R. TcRRMs and Tcp28 genes are intercalated and differentially expressed in *Trypanosoma cruzi* life cycle. *Biochem Biophys Res Commun* 2004; 322(3): 985-92.
- [77] Figueiredo LM, Cross GA, Janzen CJ. Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* 2009; 7(7): 504-13.
- [78] Elias MC, Nardelli SC, Schenkman S. Chromatin and nuclear organization in *Trypanosoma cruzi*. *Future Microbiol* 2009; 4: 504-13.

---

Received: November 15, 2009

Revised: May 04, 2010

Accepted: May 04, 2010

© Ávila and Goldenberg; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.