Review

Untranslated regions of mRNA and their role in regulation of gene expression in protozoan parasites

SHILPA J RAO[1,2], SANGEETA CHATTERJEE[3] and JAYANTA K PAL[1,*,†]

[1]*Cell and Molecular Biology Laboratory, Department of Biotechnology, and*

[2]*Interdisciplinary School of Health Sciences, Savitribai Phule Pune University, Pune 411 007, India*

[3]*Centre for Converging Technologies, University of Rajasthan, Jaipur 302 004, India*

**Corresponding author (Email, jkpal@hotmail.com)*

[†]*Present address: Dr. D.Y. Patil Biotechnology & Bioinformatics Institute, Tathawade, Pune 411033, India*

Protozoan parasites are one of the oldest living entities in this world that throughout their existence have shown excellent resilience to the odds of survival and have adapted beautifully to ever changing rigors of the environment. In view of the dynamic environment encountered by them throughout their life cycle, and in establishing pathogenesis, it is unsurprising that modulation of gene expression plays a fundamental role in their survival. In higher eukaryotes, untranslated regions (UTRs) of transcripts are one of the crucial regulators of gene expression (influencing mRNA stability and translation efficiency). Parasitic protozoan genome studies have led to the characterization (*in silico*, *in vitro* and *in vivo*) of a large number of their genes. Comparison of higher eukaryotic UTRs with parasitic protozoan UTRs reveals the existence of several similar and dissimilar facets of the UTRs. This review focuses on the elements of UTRs of medically important protozoan parasites and their regulatory role in gene expression. Such information may be useful to researchers in designing gene targeting strategies linked with perturbation of host-parasite relationships leading to control of specific parasites.

[Rao SJ, Chatterjee S and Pal JK 2017 Untranslated regions of mRNA and their role in regulation of gene expression in protozoan parasites. *J. Biosci.* **42** 189–207]

1. Introduction

Survival is ingrained as the basic instinct of all living entities, be it microscopic or gigantic. Protozoan parasites are one of the few members of the elite groups who have endured it all, from freezing climates to hot humid climes in their thousands of years of existence. Additionally, being parasitic in nature, they have also adapted to frequent biochemical and biophysical environmental changes as they go vector-host hopping. Thus, it is evident that such specialists have captured human interest since antiquity (as evident from ancient remains and the scriptures in Aryan, Greek, Egyptian, Aztec, Mayan and other cultures). However, they still continue to bring significant number of deaths and woe to human population. Migrations, slave trade, colonization, globalization and, more recently, the spread of immunodeficiency diseases (e.g. HIV-AIDS) and associated immunodepression have brought parasitic protozoan diseases to the forefront in world health policy and philanthropic efforts (Cox 2002). Several species of these parasites, viz. *Leishmania*, *Trypanosoma*, *Plasmodium*, *Entamoeba, Toxoplasma*, *Giardia* and *Trichomonas*, which were once identified

Keywords. 3′-UTR; 5′-UTR; AU-rich elements; gene expression; protozoan parasites; upstream open reading frames

Abbreviations used: ARE, AU-rich element; IRE, iron responsive element; IRP, iron responsive protein; HSP, heat shock protein; PABP, poly(A)-binding protein; RBP, RNA binding protein; SIDER, short interspersed degenerated retrotransposons; SL, spliced leader; TbZFP3, *T. brucei* zinc finger protein 3; TcSMUG, *T. cruzi* small mucin gene; TcUBP, *T. cruzi* U-rich binding protein; TR, translationally repressed; uORFs, upstream open reading frames; UTR, untranslated region; VSG, variable surface glycoprotein

DOI: 10.1007/s12038-016-9660-7

Published online: 11 January 2017

as causative agents of various diseases of the Third World Countries, are now being revisited. Due to lack of effective drugs, these parasites exert a large toll on endemic population including loss of ability to attend school or work, growth retardation, impairment of cognitive skills and development in young children, in addition to the serious economic burden on the individual as well as the society on the whole. Therefore, efforts are on to develop appropriate therapeutic and public health awareness measures. However, any therapeutic intervention requires detailed knowledge of their complex life cycle and identification of peculiarities in their metabolism *vis-a-vis* that of their respective hosts. In this context, comprehension of gene regulation in both the parasites and their hosts is of relevance.

In a cell, gene regulation takes place at both the levels of transcription and translation. Translation-mediated gene regulation occurs at both global and mRNA-specific levels. In eukaryotes (e.g. yeast) as much as 30% of genes are regulated at the level of translation. Furthermore, in higher organisms, translation-mediated gene regulation dominates during many specialized processes and conditions such as fertilization, early embryogenesis (differentiation), stress response and cancer (Chatterjee and Pal 2009). Untranslated regions (UTRs) of transcripts play a significant role in mRNA-specific translation regulation, in various processes including host–parasite relationships. UTRs are sequences present at the 5′- and 3′-ends of a transcript that do not code for protein(s) but carry features that allow differential regulation of a gene. These *cis*-acting elements influence the stability and translation efficiency of an mRNA with the help of a cohort of *trans*-acting proteins. In humans, perturbations in the 5′- and 3′-UTRs result in deregulation of genes, leading to various diseases or susceptibility to diseases (Chatterjee and Pal 2009). Gene regulation through UTRs serves as an important means of controlling various cellular events as the protozoan parasites shuttle from vector to host. UTR-mediated translation regulation facilitates appropriate response of the parasites to the periodic changes in the environment (as many of them have multiple hosts) by fine-tuning their gene expression to ensure their survival and propagation.

The current review focuses on various aspects of gene regulation by the UTRs in protozoan parasites. Further, various lacunae in our current understanding regarding their pathogenesis with respect to post-transcriptional regulation of gene expression are also discussed. This information might therefore highlight the potential applications of UTR-mediated gene regulation in infectious protozoan parasites and open up new avenues for drug targeting.

2. Regulation of gene expression in protozoan parasites

Different hierarchies of gene regulation, including transcriptional, post-transcriptional and translational control, orchestrate the gene expression profiles of organisms to varying extents. Detailed study of parasitic protozoans such as kinetoplastids (*Leishmania* and *Trypanosoma*), apicomplexans (e.g. *Plasmodium* and *Toxoplasma*), *Trichomonas*, *Giardia* and *Entamoeba* shows that they have diverged early in evolution from other eukaryotes as evident from their allegiance to different supergroups in the domain Eukaryota. This has led to the evolution of many alternate metabolic pathways, tweaking of core pathways or dispensing of an entire metabolic pathway essential in a free living organism (Ginger 2006). They have in their arsenal, distinct genes for parasitic subversion of host and they share several genes with their niche neighbors and their hosts as a result of repeated horizontal gene transfers among themselves during the course of evolution. We find that each of these parasites have developed unique *modus operandi* of gene expression as they fight to carve a niche for themselves (Ginger 2006).

Post-transcriptional mechanisms are major contributors to rapidly changing mRNAs and proteomic profiles in trypanosomatids, while transcription and post-transcriptional regulations are predominant in apicomplexans (table 1). Unusual features, viz. lack of RNA polymerase II promoters, transcription factors, formation of polycistronic transcripts and *trans*-splicing in trypanosomatids, are suggestive of the absence of transcription regulation and presence of post-transcriptional mechanisms in them (Clayton and Shapira 2007; Lahav *et al.* 2011). A significant delay between maximum mRNA abundance and peak protein expression in a study relating to the total life cycle of the *Plasmodium falciparum* indicates the existence of post-transcriptional regulation (Le Roch *et al.* 2004). A genome-wide study using RNA-seq experiments elucidated translation control as one of the key mechanisms contributing to regulation of gene expression in the erythrocytic cycle of *P. falciparum* (Bunnik *et al.* 2013). The presence of relatively few transcription regulators, the regulatory role of the UTRs (Le Roch *et al.* 2004; Hall *et al.* 2005), stable storage of translationally silent transcripts in P-granules, and hard-wired intermediate metabolism unresponsive to specific lethal metabolic perturbations (Ganesan *et al.* 2008) further suggest the importance of post-transcriptional regulation in apicomplexans. Furthermore, the presence of uniquely long UTRs in *Toxoplasma gondii* indicate their regulatory roles in this apicomplexan as well (Ramaprasad *et al.* 2015).

A clear picture regarding gene regulation in early branching protists, *Entamoeba histolytica* and *Trichomonas* spp. is yet to emerge. Available data indicates the existence of both transcriptional and post-transcriptional regulation, but the extent of their influence on gene expression is uncertain. *Entamoeba histolytica* has an AT-rich highly repetitive genome with few introns, short 3′-UTRs and an unusual RNA polymerase resistant to α-amanitin (Lioutas and Tannich 1995; Shrimal *et al.* 2010; Hon *et al.* 2013). Further, the presence of alternative splicing, polyadenylation, P-body-

Table 1. Gene regulation by UTRs in protozoan parasites

Supergroup	Phylum/Order	Parasitic Protozoa	Disease	Gene regulation	Length of UTRs in nucleotides	Possible mediators of posttranscriptional regulation through UTRs
Excavata	Kinetoplastida	*Trypanosoma*	*T. cruzi* – Chagas Disease *T. brucei* - Sleeping sickness	Posttranscriptional	*T. cruzi* 5′-UTR: 10-400 3′-UTR: 17-2800 (Brandao and Jiang 2009)	Cap 4, Alternative Trans-splicing and polyadenylation, miRNA, uORFs, AREs, RNA editing, Stabilizing/ destabilizing element in 3′-UTR
	Kinetoplastida	*Leishmania*	*L. donovani*, *L. infantum*, Visceral *L. chagasi* Leishmaniasis *L. tropica*, *L. major*, Cutaneous *L. mexicana*, Leishmaniasis *L. ethopica*	Posttranscriptional	*L. major* 5′-UTR: 547 3′-UTR: 729 (median) Dillon *et al.* 2015	Cap 4, Alternative Trans-splicing and polyadenylation, uORFs, AREs, Retrotransposons, Stabilizing/destabilizing element in 3′-UTR
	Sarcomastigophora	*Giardia*	*G. lamblia* - Giardiasis	*	5′-UTR: 0-14 3′-UTR: 10-30 (Adam 2001)	Cap 0, miRNA and Retrotransposons
	Metamonada	*Trichomanas*	*T. vaginalis* – Trichomanasis	*	5′-UTR: 10 3′-UTR: 263 (Davis-Hayman *et al.* 2000)**	Cap 1, ARE, Poly(A) signals, IRE/IRP and miRNA
Chromoalveolates	Apicomplexa	*Plasmodium*	*P. falciparum,* *P.vivaz,* *P. ovale,* *P. malariae*	Transcriptional and posttranscriptional	P. falciparum: 5′-UTR: 604-1040 (Caro *et al.* 2014) 3′-UTR: 523 (avg) (Siegel *et al.* 2014)	Cap 0, uORF, Stabilizing/ destabilizing element in 3′-UTR, IRE/IRP and Poly(A) signals
	Apicomplexa	*Toxoplasma*	*T. gondii* – Toxoplasmosis	Transcriptional and posttranscriptional	5′-UTR: 969.2 (avg) 3′-UTR: 934.5 (avg) (Ramaprasad *et al.* 2015)	miRNA
Unikonts	Amoebozoa	*Entamoeba*	*E. histolytica* - Amoebiasis	*	5′-UTR: 5-20 3′-UTR: 21 (avg) (Bruchhaus *et al.* 1993)	Retrotransposons, miRNA and Poly(A) signals

*Existence of transcriptional and/or post-transcriptional regulation, but extent of influence unknown. Please see text for additional information.

**given length is of Heat Shock Protein 70 (HSP70). Additional information about other genes unknown.

like structures and RNA interference (RNAi) indicates post-transcriptional gene regulation (Lopez-Rosas *et al.* 2012). Similarly, *Trichomonas vaginalis* with an AT-rich genome contains several repeats and transposable elements. Genes are transcribed by an α-amanitin-resistant RNA polymerase II. The core promoter element consists of a highly conserved Inr element located in the 5′-UTR of ~75% of transcripts. Relatively, very few genes (65 out ~45,000) contain introns. In trichomonads, the presence of capped and polyadenylated transcripts with IRE, and IRE/IRP-like iron regulatory system, indicates post-transcriptional regulation (Torres-Romero and Arroyo 2009). *Giardia lamblia*, a deep branching protist, has a small genome (12 Mb) and most of the genes lack introns. Most of the known transcription factors are missing; however, existence of a single highly divergent TATA-binding protein has been reported (Li and Wang 2004). *Giardia* transcripts have unusually short 5′- and 3′-UTRs, suggesting the absence of canonical ribosomal scanning (table 1). Although a definitive picture of gene regulation in *Giardia* is unavailable, the presence of microRNA (miRNA)-mediated translation repression and thus the existence of post-transcriptional regulation has been reported (Li and Wang 2004). In view of the above discussion, a comprehensive understanding of gene regulation in the context of evolutionary disparity and similarity with the hosts, which can be effectively exploited for therapeutic intervention, is of high significance.

3. Role of 5′-UTR in the regulation of gene expression in parasitic protozoa

5′-UTR refers to the region of a transcript upstream of the start codon. Regulation of translation through various determinants in the 5′-UTR is well characterized in higher eukaryotes. The readers may refer to a few excellent reviews (Chatterjee *et al.* 2010; Barrett *et al*. 2012) for detailed information. A few of these determinants along with some unconventional characteristics, unique to the group of protozoan parasites, are discussed below with specific reference to their inter-generic divergence.

3.1 *Length of 5′-UTR, 5′-cap structure and internal ribosome entry site*

The length of the 5′-UTR influences the translation efficiency by determining the energy a navigating ribosome needs to reach the AUG through a highly structured 5′-UTR (Chatterjee and Pal 2009). The length of 5′-UTR in trypanosomatids (Brandao and Jiang 2009; Smandi *et al.* 2012) is comparable to that of other eukaryotes, and it is longer in apicomplexans (Watanabe *et al.* 2002; Yamagishi *et al.* 2012). However, it is unusually short in *Entamoeba* (Bruchhaus *et al.* 1993; Hon *et al.* 2013), *Giardia* (Adam 2001) and *Trichomonas* (Davis-Hayman *et al.* 2000) (table 1). Initiation of protein synthesis requires binding of an eIF4F complex comprising eIF4E (cap binding protein), eIF4G (a scaffolding protein) and eIF4A1 (DEAD box containing RNA helicase) to the 5′-cap of the transcripts. The scanning complex containing small subunit of ribosomes with the initiator-Met $tRNA_i$ and several initiation factors (eIFs) typically binds at the 5′-end of an mRNA and scans for the initiation codon, AUG. However, as in the case of *Giardia*, unusually short or even missing 5′-UTR and lack of eIF4B, eIF4H and eIF4G homologs indicate the absence of ribosomal scanning for translation initiation. Such short 5′-UTRs possibly indicate the presence of a unique and simplified protein synthesis machinery in deep branching protists (Li and Wang 2004). In *Trypanosoma brucei*, a reduction in the abundance of reporter transcripts was observed on introduction of secondary structure or lengthening of the 5′-UTR (for translation inhibition), linking translation rate with mRNA stability (Delhi *et al.* 2011).

The 5′ 7-methylguanosine (m^7G) cap affects cellular processes like transcription, splicing, polyadenylation, nuclear export of mRNA, mRNA stability and translation. Various cap structures, namely, cap 0, cap 1, cap 2, cap 3 and cap 4 have been characterized in eukaryotes (Cowling and Cole 2010). A yeast-like capping apparatus suggests the presence of cap 0 structure in *P. falciparum* and *Giardia* (Jagus *et al.* 2012). Although the exact cap structure of *Giardia* transcripts has not been determined yet, exogenously introduced m7GpppN-capped mRNA was found to express well (Li and Wang 2004), indicating thereby the possible presence of cap 0 structure. Transcripts of *Trichomonas vaginalis* bear cap 1 structure (Simoes-Barbosa *et al.* 2010) while unusually hypermethylated cap 4 is exclusively found in trypanosomatids (figure 1) (Jagus *et al.* 2012). In *T. brucei,* gene deletions of three cap 2′-*O*-ribose methyltransferases, TbMTr1, TbMTr2 and TbMTr3 revealed that individual deletions of gene pairs yielded viable cells, but they grow slowly as compared to normal cells, indicating the minimal mRNA ribose methylation is necessary for trypanosome viability (Zamudio *et al.* 2009). Authors suggest that cap 4 structure may be preferred during mRNA processing and translation while early cap modifications on splice leader (SL) RNA intermediates could function particularly in SL RNA biogenesis. A different level of mRNA cap ribose methylation also occurs in *Xenopus laevis* and *Caenorhabditis elegans* wherein it provides additional level of post-transcriptional regulation (enhancement of transcript stability, processing, trafficking and translation efficiency) mediated by interaction with nuclear and cytoplasmic cap binding proteins (Zamudio *et al.* 2009). Whether the same applies true for trypanosomatids remains to be determined.

The unique cap structure has facilitated evolutionary variation in the cap-binding protein eIF4E, and it has been an

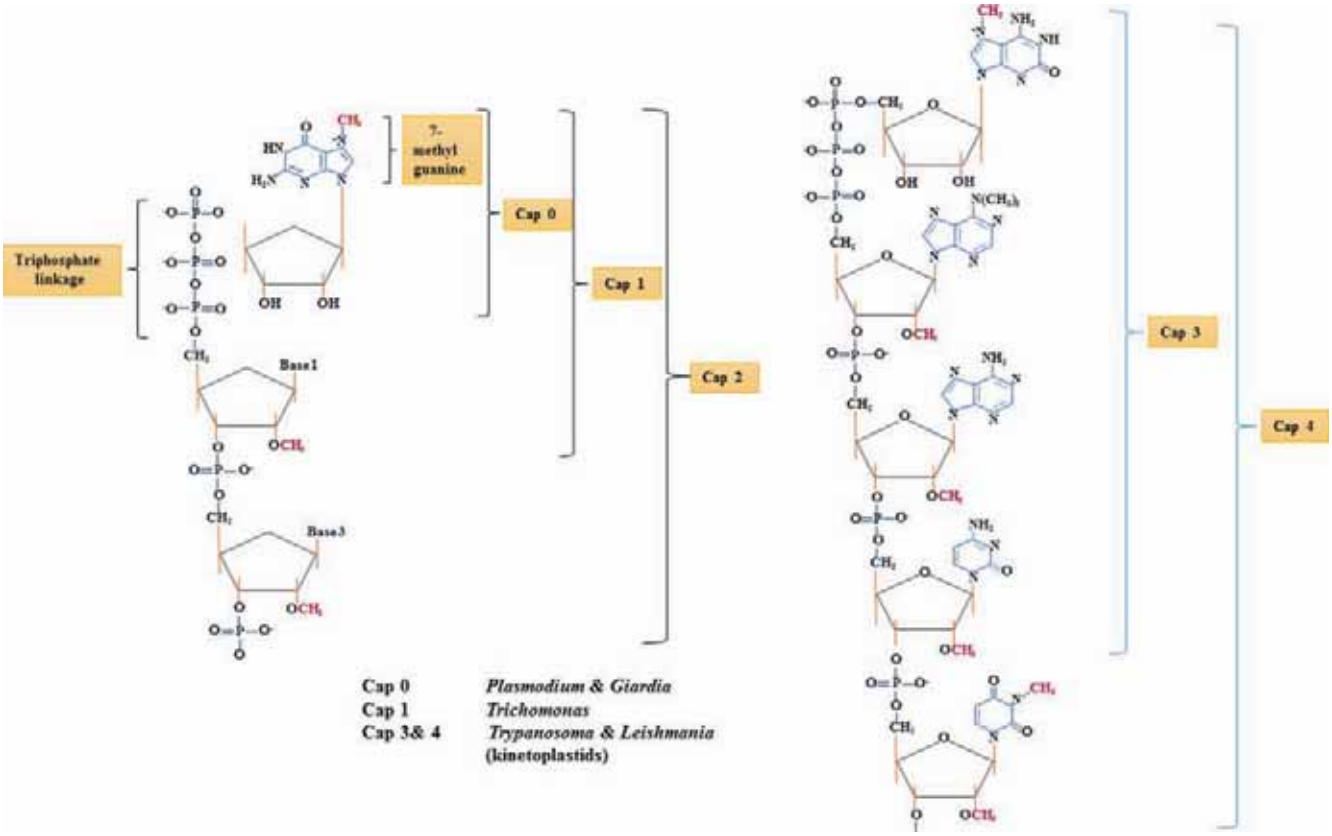

Figure 1. 5′-cap structures present in eukaryotes. An m^7G attached by a 5′-5′ triphosphate linkage to the first nucleotide of a transcript results in cap 0 structure. Methylation at the 2′-*O*-ribose of the first nucleotide or first and second nucleotides results in cap 1 or cap 2 structure respectively. Further methylation at the 2′-*O*-ribose of the third and fourth nucleotides and base methylations at the first (m_2^6A) and fourth (m^3U) positions results in a hypermethylated cap 4 structure.

area of extensive research, especially in trypanosomatids. Trypanosomatids encode multiple paralogs of eIF4E (4) and eIF4G (6) and these show a high functional divergence from their higher eukaryotic counterparts (Zinoviev and Shapira 2012). Interestingly, in promastigotes of *Leishmania*, Leish4E-4 and Leish4G-3 are a part of typical eIF4F complex, but the binding of Leish4E-4 to Leish4G-3 and to the cap 4 structure is abolished upon exposure to mammalian-specific temperature (Yoffe *et al.* 2009). Another cap-binding paralog, Leish4E-1, found to express in both promastigotes and amastigotes of *Leishmania* interacts with a novel 4E-Interacting Protein (4E-IP). On exposure to host temperature, 4E-IP probably gets hyperphosphorylated and releases Leish4E-1, which might promote cap binding (Zinoviev *et al.* 2011). However, in the absence of binding of Leish4E-1 to any Leish4G, its role seems to involve a novel interaction with a translation initiation factor Leish4E-3, although the significance of this binding needs to be fully resolved (Meleppattu *et al.* 2015). In contrast, in *T. brucei*, TbIF4E-4 and TbIF4E-1 were found to be relevant to the bloodstream forms rather than to the procyclic form (Zinoviev and Shapira 2012). Leish4E-3 shows no affinity to cap 4 structure but binds with Leish4G-4, under normal conditions in promastigotes. During nutrient starvation, Leish4E-3 is probably modified, released by Leish4G-4 and enters stress granules, although its role is still unclear (Zinoviev *et al.* 2012). A similar localization of TbIF4E-3 in stress granules was observed in *T. brucei*, during heat shock (Kramer *et al.* 2008). The essentiality of TbIF4E-3 for *T. brucei* as assessed by silencing experiments and unique interacting motifs in Leish4E-3, makes 4E-3 a potential drug target for trypanosomatids.

Many transcripts implicated in stress response in *P. falciparum* are stored as uncapped, translationally quiescent form (Shaw *et al.* 2007). The authors hypothesize that these transcripts could undergo an internal ribosome entry site (IRES)-mediated translation during stress. Although IRES-mediated translation is widely reported in viruses and eukaryotes, it remains unproven in protozoan parasites. Interestingly, IRES-mediated translation has been reported in viruses infecting *Leishamania* (Maga *et al.* 1995) and *Giardia* (Garlapati and Wang 2004). IRES of *Leishmania* viral

transcripts resides in the 5′-UTR; however, in Giardia viruses, it spans sequences in the ORF in addition to a region in the 5′-UTR, forming additional structures in the IRES of the transcripts in Giardia virus unlike viruses of higher eukaryotes (Garlapati and Wang 2004, 2009). However, little is known about the presence of such uncapped mRNA in other protozoan parasites and its implication in translation regulation.

It is evident from the above discussion that these parasites have evolved distinctive and diverse approaches, viz. differently lengthed 5′-UTR, unique cap structure and cap-binding proteins which in certain cases are highly divergent from their orthologs in higher eukaryotes. In higher eukaryotes, failure in binding of the scanning apparatus to the cap at 5′-UTR leads to impaired protein synthesis and subsequent susceptibility to diseases (Sonenberg and Hinnebusch 2009). Extrapolating the same reasoning to this case, the introduction of specific inhibitors to the cap or cap binding proteins can effectively sabotage the parasite's translation machinery. Potential drug targets that could be pursued include protein–RNA and protein–protein interactions that promote assembly of the translation initiation complex. For example, interactions between the cap-binding proteins and the unique cap-4 structure, and the binding between LeishIF4E-3 and LeishIF4G-4 are of interest. TbIF4E-3, the ortholog of LeishIF4E-3, also appears to be an important drug target, as it was shown to be essential in the bloodstream form of *T. brucei* (Zinoviev and Shapira 2012). Although the data supporting this approach is currently insufficient, this approach may be feasible in future with mass screening of anti-trypanosomatid chemical libraries.

3.2 *Trans-splicing*

Absence of conventional transcription regulation in trypanosomatids demands the control of differential gene expression by post-transcriptional activities like *trans*-splicing, mRNA stability and translation (Clayton and Shapira 2007). Interestingly, *trans*-splicing is exclusively observed in trypanosomatids and is not exhibited by other protozoan parasites under discussion. However, this phenomenon was later found in nematodes, euglenoids, trematodes and chordates (Liang *et al.* 2003). Maturation of polycistronic pre-mRNAs in trypanosomatids involves addition of a 39 nt spliced leader (SL) sequence carrying cap 4 from a SL RNA to the 5′-end of the transcripts (*trans*-splicing) and polyadenylation at the 3′-end, giving rise to individual mature mRNAs (figure 2). The details of this mechanism has been reviewed in (Michaeli 2011). This section mainly deals with gene regulation facilitated by *trans*-splicing, resulting in stage-specific gene expression in trypanosomatids.

Trans-splicing of SL sequence at different sites in a pre-mRNA results in mature transcripts with varied lengths of 5′-UTR. This alternative splicing may result in inclusion or exclusion of upstream open reading frames (uORFs), alteration in the localization signals of the translated proteins or facilitate the use of alternate ORFs, permitting regulation of gene expression (Nilsson *et al.* 2010). Few examples of this mechanism have been elucidated in trypanosomatids. Two variants of the LYT1, a protein involved in the lytic pathway in *T. cruzi* (each differing by 28 amino acids), have acquired different functions and compartmentalization due to alternate *trans*-splicing in a stage-dependent manner (Benabdellah *et al.* 2007). Using a high-throughput sequencing approach, around 2500 alternative spliced events were identified in the bloodstream and insect-borne forms of *T. brucei*; more than 600 of these transcripts were found to be stage-specific (Nilsson *et al.* 2010). The dual localization (in mitochondria and cytosol) of spliced variants of isoleucyl-tRNA synthetase mRNA and the mRNAs of other predicted aminoacyl-tRNA synthetases (aaRSs) in *T. brucei* established differential *trans*-splicing as the underlying mechanism for such an event (Rettig *et al.* 2012). The evolution of dual targeting of aaRSs could be linked to the loss of mitochondrial aaRSs genes and subsequently that of mitochondrial tRNA genes which is compensated by import of cytosolic tRNAs in trypanosomatids. Further, the authors hypothesize the presence of stabilizing or destabilizing element in the 5′-UTR of transcripts that may account for differences in the isoform abundance. A similar heterogeneity in selection of SL addition sites has been mapped in many transcripts by RNA seq analysis of *L. major* (Rastrojo *et al.* 2013; Dillon *et al.* 2015).

Although genome-wide study in *Trypanosoma* mapped extensive SL addition sites, little is known about the proteins contributing to this event. A heterogeneous ribonucleoprotein (hRNP F/H) homolog, known to regulate alternative splicing in metazoan, has been shown to be a good candidate for stage-specific splicing regulation in *T. brucei* (Gupta *et al.* 2013). In higher eukaryotes, splicing efficiencies are altered by members of serine/arginine (SR)-rich protein family that bind to these regulatory elements and interact with splicing machinery. Recently two SR proteins (TSR1 and TSR1P) were found to regulate splicing in *T. brucei* (Gupta *et al.* 2014) along with other functions like mRNA stability and rRNA processing. However, a complete understanding of their interplay between these processes remains to be arrived at.

SL RNA has also been suggested as an attractive molecule for diagnosis of trypanosomatid infections (Gonzalez-Andrade *et al.* 2014). As evident from the above description, a few orthologs of components of *trans*-splicing machinery have been identified in the parasites. Nonetheless, their functional and regulatory mechanism requires further elucidation. Potential trypanocidal drugs that can target binding of SL RNA/SL RNP in pre-mRNAs or interfere with methylation of the 5′ cap are worth looking into in future.

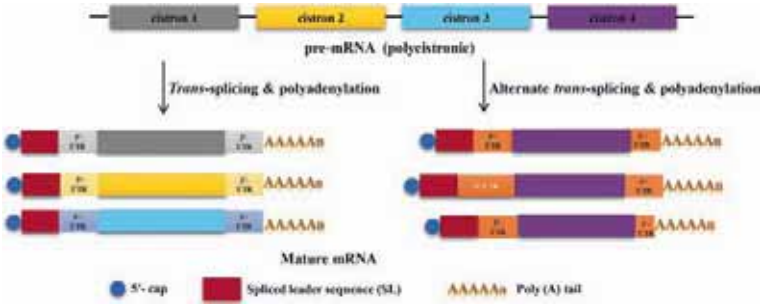

Figure 2. Processing of mRNA in trypanosomatids. Transcription results in polycistronic mRNA which undergoes maturation by addition of a capped 39 nt spliced leader (SL) sequence to the 5′-end and polyadenylation at the 3′-end, giving rise to individual mature mRNA. Alternate *trans*-splicing may occur due to multiple *trans*-splice acceptor sites in a transcript, producing proteins with diverse functions (Dillon *et al.* 2015; Nilsson *et al*. 2010).

3.3 *Upstream open reading frames*

Many genes with critical biological functions are regulated by upstream open reading frames (uORFs) in higher eukaryotes. These are present upstream of the main ORF and repress protein synthesis by several mechanisms including stalling of ribosomes, prevention of re-initiation by ribosomes and activation of non-sense-mediated decay (Chatterjee and Pal 2009).

Erythrocytes infected with *P. falciparum* randomly express a family of *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) surface molecules encoded by a family of ~60 *var* genes of the parasite that help it to adhere to the microvasculature. A *var* gene called *var2csa* encodes a PfEMP1 protein (VAR2CSA) that binds to chondroitin sulphate A (CSA) and is regulated by a unique mechanism not relevant to other members of the family. The transcripts of VAR2CSA are present in a translationally repressed state in *Plasmodium* due to the presence of a long uORF (120 amino acids) which ensures the absence of this protein in a normal malarial infection (Lavstsen *et al.* 2003; Amulic *et al.* 2009). Experiments suggest that failure of re-initiation at the downstream ORF by the scanning ribosomes could be the probable mechanism for translation repression of VAR2CSA. This strategy prevents development of memory response against VAR2CSA and allows utilization of the CSA receptor on the placenta for efficient binding and sequestration of the infected erythrocytes during pregnancy associated malaria (PAM) (Lavstsen *et al.* 2003; Amulic *et al.* 2009). The population of parasites in which the ribosomes retain initiation factors beyond uORFs may reinitiate translation of the main ORF thereby overcoming repression during PAM (Bancells and Dietsch 2013). One of the mechanisms of derepression could be phosphorylation of the initiation factor, eIF2α by eIF2α kinases. This strategy impairs formation of ternary complex (eIF2.Met-tRNA$_i^{Met}$.GTP) at the uORFs and facilitates the scanning of ribosome to reach the target ORF by bypassing the uORFs, ensuring translation of the main ORF (figure 3). Recent reports suggest the significance of these kinases in the life cycle of protozoan parasites (Moraes *et al.* 2007; Fennell *et al.* 2009; Chow *et al.* 2011); however, the role of parasitic eIF2α kinases in uORFs mediated regulation remains yet unexplored.

In the absence of transcription factors in *Plasmodium*, uORF-mediated gene regulation assumes significance in changing the proteomic profile required for sexual development of the parasite. Regulation of sexual development in *Plasmodium* is mediated through translationally repressed (TR) transcripts which are stored in cytoplasmic bodies in the female gametocytes until the blood-stage gamete precursors are ingested by a mosquito. The *cis*-acting signals in the TR transcript direct their preferential translation in such an event. Regulation through uORFs is one such signal capable of preferential translation or repression of mRNAs depending upon the molecular cue (Amulic *et al.* 2009). A similar regulation was also observed in *T. gondii* (Joyce *et al.* 2013). Few genes from *T. cruzi* (Jaeger and Brandao 2011) and *P. falciparum* (Panneerselvam *et al.* 2011) show the presence of uORFs, but their functional relevance in the regulation of the respective genes has not been analysed. Further, ribosome profiling data suggested that uORFs may be one of the important contributory mechanisms of gene regulation in *T. brucei* (Vasquezz *et al.* 2014), however the genuineness of these uORFs needs to be further validated (Jensen *et al*. 2014). uORFs in the transcripts may not be a mode of regulation for *Giardia* owing to their very short 5′-UTR.

The wealth of knowledge about uORFs may improve with deep sequencing, RNA-Seq experiments wherein hitherto unknown genome information may be unearthed; nevertheless, mutational analysis is the only way to discriminate spurious AUGs from uORFs. Identification of other genes with uORFs, the nature of uORFs and determining their mechanism of repression/derepression will help in a better understanding of this mode of regulation of gene expression in several protozoan parasites.

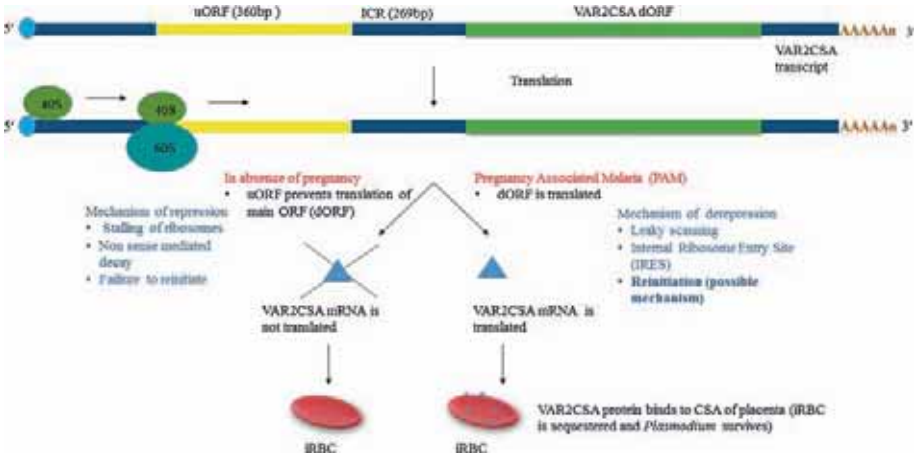

Figure 3. Regulation of translation (e.g. VAR2CSA mRNA) by uORF in *Plasmodium*. VAR2CSA mRNA contains a regulatory upstream ORF (uORF). In the absence of placenta (pregnancy), the translation of downstream ORF, VAR2CSA, is repressed by upstream ORF. However, in pregnancy associated malaria (PAM), derepression leads to expression of VAR2CSA. This Plasmodium protein is expressed on the surface of infected RBC (iRBC), which specifically binds to chondroitin sulphate (CSA) on the placenta, thereby avoiding clearance by spleen (ICR, intercistronic region).

4. Role of 3′-UTR in the regulation of gene expression in parasitic protozoa

The term 3′-UTR refers to the sequence of a transcript from the stop codon to the poly(A) tail. *Cis*-acting elements in 3′-UTR, along with their associated *trans*-acting factors, influence mRNA stability and translation efficiency in higher eukaryotes. The 3′-UTRs of transcripts may harbor AU-rich elements (AREs), polyadenylation signals, retrotransposable elements and iron-responsive elements (IREs), thereby influencing their translation (table 1). These are classical determinants of 3′-UTR-mediated gene regulation, however the protozoan parasites enjoy in having distinctive *cis*-acting elements and *trans*-acting factors, a few of which are well-understood (table 2). In this section we discuss these examples before moving onto the more familiar ones.

Gene expression in trypanosomatids is mostly post-transcriptionally regulated and is primarily mediated through the 3′-UTRs as evident from the regulation of HSPs and amastin genes in *Leishmania*. HSP70s in *L. infantum* (Quijada *et al.* 1997) and *L. braziliensis* (Ramirez *et al.* 2011) are encoded by two gene clusters, HSP70-I (genes 1-5) and HSP70-II (gene 6), that differ only with respect to their 3′-UTRs. Transcripts of HSP70-I bearing 3′-UTR I are translated at both 26°C and at elevated temperature (37°C). But, HSP70-II transcripts bearing 3′-UTR II are otherwise translationally silent until the parasites encounter heat shock temperature, indicating the existence of a different, heitherto unknown UTR-mediated regulatory mechanism (Folgueira *et al.* 2005). A riboproteomic analysis of the factors interacting with the 3′-UTRs of HSP70 mRNA in *L. braziliensis* revealed several proteins including few putative ones whose characterization may open new avenues for drug targeting, considering the significant role of HSPs in *Leishmania* biology. Computational and experimental data suggest that the translation of HSP83 transcript in *L. amazonensis* at higher mammalian temperature is facilitated due to the melting of highly probable structure of the regulatory element in its 3′-UTR (David *et al.* 2010). Further, the 3′-UTR of HSP70-I of various *Leishmania* species is remarkably conserved, but the marginal disparity has been exploited for species discrimination (Requena *et al.* 2012). A defined region of 450 nt and a downstream 100 nt in the 3′-UTR of amastin (surface glycoprotein) gene family assist in augmenting translation of the transcripts at host temperature in amastigote stage of *L. infantum* (Boucher *et al.* 2002). The other differentiation signal, i.e. low pH condition encountered in phagolysosomes in the host, enhanced the abundance of amastin transcripts. Such a bipartite regulation of genes, that aids in both increased mRNA abundance and translation efficiency, probably accounts for the drastic changes in the proteomic profiles supporting differentiation of the parasites (McNicoll *et al.* 2005).

3′-UTR-mediated regulation of transcripts has also been reported in *Trypanosoma*. Two elements in the 3′-UTR of PAD1 mRNA (encoding a surface transporter protein) of *T. brucei* have been credited to its repression in the slender form of the parasite (MacGregor and Matthews 2012). A short bifunctional element of 34 nt was mapped in the 3′-UTR of ESAG9 transcripts (encoding a group of secreted protein with cryptic function) that acts both positively (in stumpy forms) and negatively (in slender forms) probably by

Table 2. Summary of *cis*-acting elements in the UTRs of transcripts and their associated *trans*-acting factors in protozoan parasites

Organism	Transcript regulated by UTR	UTR involved	*Cis* acting element in UTR of transcript	Mode of regulation	Trans-acting factors	Reference
1. *T. cruzi*	*LYT1*	5′-UTR	**	alternate *trans*-splicing	SR proteins*	Benabdellah *et al.* 2007
	TcSMUG	3′-UTR	AU-rich element G-rich element	mRNA stability	TcUBP-1 TcUBP-2 TcPABP-1#	D'Orso and Frasch 2001a; D'Orso and Frasch 2001b; D'Orso and Frasch 2002
	Genes required for differentiation into amastigotes	3′-UTR	43 nt U-rich regions	mRNA stability	TcUBP-1*	Li *et al.* 2012
	TcRBP-19	3′-UTR	**	mRNA stability	TcRBP-19* (autoregulation)	Perez-Diaz *et al.* 2013
	GP82	3′-UTR	Multiple elements	mRNA stability	TcUBP-1##	
	TcTASV (*T. cruzi* trypomastigote alanine serine valine proteins)	3′-UTR				
	Alpha tubulin	3′-UTR		mRNA stability		
	Amastin	3′-UTR	203bp element	mRNA stability		
	HSP 70	3′-UTR	AU-rich element	mRNA stability		
2. *T. congolense*	Congolense epimastigote specific protein (cesp)	3′-UTR	AU-rich element	mRNA stability	#	Suganuma *et al.* 2013
3. *T. brucei*	aminoacyl-tRNA synthetases	5′-UTR	**	*alternate trans-splicing*	SR proteins *	Rettig *et al.* 2012
	Procyclin	3′-UTR	U-rich elements	mRNA stability	TbZFP3*	Walrad *et al.* 2009
	Hexose transporters	3′-UTR	**	mRNA stability	#	Hotz *et al.* 1995
	Phosphoglycerate kinases	3′-UTR	**	mRNA stability	#	Colasante *et al.* 2007
	Amino acid transporter 11	3′-UTR	Multiple elements	Translation repression	#	Robles and Clayton 2008
	NT8	3′-UTR	Short stem-loop	mRNA abundance	#	Fernandez-Moya *et al.* 2014
	GPEET procyclins	3′-UTR	Glycerol responsive element	Altered processing of mRNA	#	Vassella *et al.* 2004; Knusel and Roditi 2013
	HSP70	3′-UTR	Motif with AUU repeat	mRNA stability	ZCH11	Droll *et al.* 2013
	**	3′-UTR	AU-rich element	mRNA stability	DHH1	Carrington and Krammer 2014
	F-box protein (CFB1)	3′-UTR	**	mRNA stability	TbUBP1/2	Hartmann *et al.* 2007
	ESAG9	3′-UTR	34 nt element	mRNA stabilty	#	Monk *et al.* 2013
	S-adenosylmethionine decarboxylase regulatory subunit	3′-UTR		Translation efficiency	Decarboxyated S-adenosul methionine	
4. *Leishmania spp.*	Paraflagellar rod	3′-UTR	10 nt AU-rich element	mRNA stability	Homolog of TcUBP-1##	Moore *et al.* 1996
	HSP83	5′-UTR	**	Translation efficiency	#	David *et al.* 2010
	H2A	5′-UTR	**	Translation efficiency	#	Abanades *et al.* 2009

Table 2 (continued)

Organism	Transcript regulated by UTR	UTR involved	*Cis* acting element in UTR of transcript	Mode of regulation	Trans-acting factors	Reference
	TOP 2	5′-UTR	octamer consensus sequence	mRNA stabilty	#	Pasion *et al.* 1996
	HSP70 I	3′-UTR	Secondary structures	mRNA stability	#	Folgueira *et al.* 2005
	HSP70 II	3′-UTR	Secondary structures	Translation efficiency	#	Folgueira *et al.* 2005
	HSP83	3′-UTR	Polypyrimidine –rich element	Translation efficiency	#	David *et al.* 2010
	Amastin	3′-UTR	450 nt & 100 nt region	Increased translation efficiency and mRNA abundance	#	McNicoll *et al.* 2005
5. *T. vaginalis*	Transcripts involved in iron modulation	3′-UTR	AU-rich element	mRNA stability	#	Leon-Sicairos *et al.* 2004
	cysteine proteinase Tvcp12 & Tvcp4	5′ & 3′- UTR	IRE	-	α-actinin, HSP70, actin *	Calla-Choque *et al.* 2014
	Tvcp12	3′-UTR	Heterogenous poly A tail	mRNA stability	#	Leon-Sicairos *et al.* 2004
6. *Plasmodium*	VAR2CSA	5′-UTR	uORF	Translation efficiency	#	Amulic *et al.* 2009
	PAD1	3′-UTR	3′-UTR	-	#	MacGregor and Matthews 2012
	Transcripts required post-fertilization	5′-UTR	47 nt U-rich region	Translation repression	Calcium dependent protein kinase (CDPK1)*	Sebastian *et al.* 2012
	**	5′ & 3′- UTR	IRE	-	PfIRPa	Loyevsky *et al.* 2003
	Circumsporozoite	3′-UTR	Heterogeneous poly(A) tail	mRNA stability	#	Le Roch *et al.* 2004
7. *E. histolytica*	EhPgp5	3′-UTR	Heterogeneous poly(A) tail	mRNA stability	#	Lopez-Camarillo *et al.* 2003

**Unidentified *cis*-acting elements, # trans-acting factor not identified, *Regulation mediated with the help of other unidentified *trans*-acting factors, ## Experimentally not validated.

differential interaction of stage-specific RNA binding proteins (RBPs) (Monk *et al.* 2013). An RBP, TcRBP19, was reported to be involved in the down-regulation of its own transcripts by binding to the 3′-UTR in epimastigotes of *T. cruzi,* but the stage-specific factors that mediate its stabilization/destabilization need elucidation (Perez-Diaz *et al.* 2013). A short RNA element in the 3′-UTR was found to be necessary and sufficient for the expression of a purine transporter, NT8 in *T. brucei* in response to extracellular purine levels in their early logarithmic phase. The mode of its action, although speculated to be either by direct binding of nucleotide or via interaction with *trans*-acting factors, remains to be determined (Fernandez-Moya *et al.* 2014).

In *Plasmodium*, transcription is arrested post-fertilization and many earlier translationally repressed (TR) transcripts in the zygote are actively translated to support the ongoing development. A U-rich region was identified in the 3′-UTR of TR transcripts through bioinformatics approach (Hall *et al.* 2005) and its deletion led to loss of translation repression in these transcripts (Braks *et al.* 2008). The mechanism of stage-dependent derepression of TR transcripts is unknown. A study in *P. berghei* revealed that a calcium-dependent protein kinase (CDPK1) controls the derepression of TR transcripts and it requires elements in the 3′-UTR (containing 47 nt U-rich regions) as well as in the 5′-UTR for its action (Sebastian *et al.* 2012).

A systematic genome-wide study has identified conserved *cis*-regulatory RNA elements in the 3′-UTR of metabolically clustered *T. cruzi* transcripts by computational analysis. Functional groups of mRNAs were found to share RNA motifs, providing a platform to identify their associated *trans*-acting factors (De Gaudenzi *et al.* 2013). A similar study in other protozoan parasites may show conserved phylogenetic signal in the structured RNA elements. Further, a recent review by Gazestani *et al*. (2014) highlights the various approaches and challenges in deciphering the RNA regulatory elements in trypanosomatids. Thus, an understanding of the regulatory network in the protozoan parasites could open up avenues for therapeutic interventions where disparity in RNA motifs or RBPs with hosts could be targeted by designing small molecules.

4.1 *AU-rich elements*

AREs belong to a family of sequence motifs such as the AUUUA pentamer, the UUAUUUA(U/A)(U/A) nonamer, and stretches of U-rich domains that influence the stability of the transcripts (Zubiaga *et al.* 1995).

Due to their AT-rich genomes and lack of consensus sequence, prediction of AU-rich element is difficult in *Plasmodium* spp. and *E. histolytica*. To date, AREs remain undetected in *Giardia* and *Toxoplasma*. Thus, AREs characterized so far in protozoan parasites are predominately from trypanosomatids. AREs in the transcript of a small mucin gene (TcSMUG) encoding a mucin-like protein in *T. cruzi* is known to negatively regulate its stability in the trypomastigote stage of the parasite (Di Noia *et al.* 2000) by a U-rich RBP, TcUBP-1 (D'Orso and Frasch 2001b). TcUBP-2, another RBP expressed only in the epimastigote stage, forms a ribonucleoprotein complex along with TcUBP-1, *T. cruzi* poly (A)-binding protein (TcPABP1) and some unknown protein(s), thereby stabilizing the transcript. The mechanism of mRNA stabilization/destabilization involving TcUBPs remains to be elucidated. In coordination with the above regulation, the 3′-UTR of SMUG mucin transcript also harbours a novel G-rich element with two contiguous CGGGG pentamers that aid in stabilization of the transcript in the epimastigote stage (D'Orso and Frasch 2001a). Another 43 nt U-rich region has been identified in a number of genes in *T. cruzi* that interacts with TcUBP-1, along with other unidentified *trans*-acting factors and thus coordinately regulates genes required for differentiation of trypomastigotes into the intracellular amastigotes (Li *et al.* 2012).

The bloodstream forms of *Trypanosoma* express variable surface glycoprotein (VSG), whereas the procyclic forms show abundance of another coat protein, procyclin/PARP (procyclic acidic repetitive protein). The family of procyclin genes encode four major surface GPI- anchored glycoproteins, namely, GPEET, EP1, EP2 and EP3 (based on internal pentapeptide and dipeptide) which are found to be differentially expressed in the vector form of the parasite (Urwyler *et al.* 2005). The stage-specific transition from VSG to procyclin expression in *T. brucei* is a fine example of post-transcriptional gene regulation by the 3′-UTR of mRNA (Hehl *et al.* 1994; Hotz *et al.* 1997; Wilson *et al.* 1999). The expression of this gene is mediated by two positive elements (40 mer and 16 mer called LI and LIII, respectively) by counteracting the action of an internal negative regulatory element called loop II (LII) (Furger *et al.* 1997; Schurch *et al.* 1997). Later, a member of CCCH zinc finger protein family, TbZFP3, in *T. brucei* was identified as an anti-repressor of the procyclin transcripts (Walrad *et al.* 2009) which along with other proteins and procyclin mRNA forms the TbZFP3mRNP complex. A global analysis of transcripts associated with TbZFP3mRNP revealed a cohort of genes that are stabilized in the stumpy form, as a preparation for entry into the insect (Walrad *et al.* 2012), ensuring translation as soon as they perceive differentiation signals. Both the isoforms of procyclins (GPEET and EP1) are initially expressed by the procyclic form, but later the expression of GPEET is repressed, a feature that distinguishes between the early and late procyclic forms. Studies on GPEET repression revealed a 25 nt glycerol responsive element in the 3′-UTR to be responsible for repression in high glucose or glycerol deprivation condition (Vassella

et al. 2004). Interestingly, differently processed mRNAs including those with extended 3′-UTR lacking poly(A) tail, differing 3′ ends within UTR and non-templated oligo (U) tailed transcripts were detected which the authors speculate to be a contributor to GPEET repression in the late forms of the parasite (Knusel and Roditi 2013). This raises many questions regarding the identity of the proteins involved in generation of alternately processed mRNA including oligo (U) tail, the RNA turnover pathway, their target specificity and their divergence from the host counterpart.

In the bloodstream form of *T. brucei*, a zinc finger protein, ZCH311 was reported to regulate mRNAs including HSP70 mRNA translation after heat shock by preferentially binding to a consensus AUU repeat motif in the 3′-UTR and stabilization of the transcript (Droll *et al.* 2013). Recent investigations on the mechanism of stabilization suggest that in the bloodstream forms, ZCH311 serves as a platform to recruit MKT1, a post-transcriptional regulator, PBP1 (PABP-interacting protein 1), LSM12 (unknown function) and PABPs to target mRNA (Singh *et al.* 2014). A post-transcriptional regulatory network of RNA binding proteins that controls gene expression in *T. brucei* has been reviewed by Kolev *et al.* (2014) and is beyond the scope of this review.

An ARE like *cis*-acting element in the 3′-UTR, identified in the transcripts of congolense epimastigote specific protein (cesp) was found to be involved in the stabilization of transcripts and their expression in epimastigote form of *T. congolense* in a stage-specific manner (Suganuma *et al.* 2013). Although some RNA-binding proteins (RBPs) were identified by RNA-EMSA, a detailed protein characterization is still needed. Recently, an AU-rich element in the 3′-UTR was found to be responsible for DHH1 (a DEAD box RNA helicase)-mediated stabilization of specific mRNAs, possibly by displacing unknown ARE-binding protein complexes that destabilize ARE-bearing transcripts in *T. brucei* (Carrington and Krammer 2014).

A 10 nt AU-rich region in the 3′-UTR called the paraflagellar rod element (PRE) causes 10-fold down-regulation of paraflagellar rod transcripts (Moore *et al.* 1996) in the amastigote form of *Leishmania.* Later, other genes were reported to be regulated by PRE, allowing a coordinated regulation of expression of genes required during differentiation (Holzer *et al.* 2008). In *T. vaginalis*, few transcripts whose expression is governed by iron modulation bear ARE, which could be involved in destabilization of these mRNAs (Leon-Sicairos *et al.* 2004). However, further investigations are required for understanding the precise mechanism of its regulation. Further, the questions that remain unanswered are whether AREs from parasitic protozoa and their hosts are functionally similar. Do they utilize different sets of proteins for this regulation or are they analogous? This is therefore an important area of research whose answers will have a bearing on drug designing and targeting.

4.2 *Iron-responsive elements and iron regulatory proteins*

Iron metabolism being critical in most life forms, regulation of intracellular iron concentration arose early in evolution. The expression of iron-regulated proteins, namely ferritin and transferrin receptor, is modulated by interaction of IRPs with IREs in the UTRs of their transcripts in mammalian cells (Chatterjee and Pal 2009). The presence of IRE/IRP-like system is reported in *Trichomonas* and *Plasmodium*, while it remains elusive in other protozoan parasites. In *P. falciparum,* high-affinity binding of an iron regulatory protein (PfIRPa) to putative plasmodial IREs residing in 5′- and 3′-UTRs of some transcripts suggests that IRE-mediated regulation is functional in this organism (Loyevsky *et al.* 2003). In *T. vaginalis*, 5′- and 3′-UTRs of cysteine proteinase transcripts (Tvcp4 and Tvcp12, respectively) form stable IRE-like stem-loop structure, capable of binding to human IRPs in an iron-dependent manner but show little similarity in sequence and structure to known IREs (Solano-Gonzalez *et al.* 2007; Stevens *et al.* 2011). The genome of *T. vaginalis* lacks IRP-like protein encoding genes, but atypical RBPs including HSP70, α-actinin and actin were recently identified to participate in IRE mediated regulation (Calla-Choque *et al.* 2014; Figueroa-Angulo *et al.* 2015). Detailed identification and characterization of IREs and IRP-like proteins will provide novel insights into mechanisms of iron regulation in protozoan parasites and also the extent of structural and functional deviations from their host. The possibility that these parasites may borrow the IRPs from their hosts cannot be ruled out. Further experiments will unravel such phenomenon that may be a part of the host–parasite relationships.

4.3 *Polyadenylation signals*

In eukaryotes, polyadenylation occurs downstream of a hexameric motif (AAUAAA/AUUAAAA) and upstream of a U-rich sequence (Zhao *et al.* 1999). In some *Plasmodium* transcripts, a pentamer (AAUAA) and a G/U-rich sequence are required for polyadenylation (Lanzer *et al.* 1993). A tetrameric nucleotide sequence (UAAA) in *Trichomonas* spp. (Fuentes *et al.* 2012) and a pentameric motif (UAUUU) along with a polypyrimidine tract (PPT) near the stop codon in *E. histolytica* (Zamorano *et al.* 2008) have been predicted as putative polyadenylation signals by *in silico* analysis. A polyadenylation site prediction tool can be developed exclusively for protozoan parasites by compiling consensus polyadenylation signal motifs to facilitate accurate determination of 3′-UTR of their transcripts. For instance, a Web server, SLAP mapper identifies such sites in kinetoplastid genome (Fiebig et *al.* 2014). Further, the *in silico* prediction needs to be experimentally determined, RNA-seq of the concerned organism being one of the tools to detect such sites.

In higher eukaryotes, the poly(A) tails of the transcripts recruit PABP that participates in circularization of mRNA along with eIF4G (prerequisite for efficient translation), which in turn prevents their degradation. In *Leishmania*, this is achieved by binding of LeishPABP-1 to Leish4E-4 instead of LeisheIF4G, a variation from their human host and unreported in other protozoan parasites so far (Zinoviev and Shapira 2012). Multi-drug resistance in *E. histolytica* is associated with the overexpression of a 170 kDa permeability glycoprotein (EhPgp5), an energy-dependent proton pump that effluxes drugs from the cell (Descoteaux *et al.* 1995). The 264 nt long 3′-UTR in this mRNA has multiple potential polyadenylation sites (at 118 nt, 156 nt, 189 nt from the stop codon) giving rise to a heterogeneous population of transcripts (Lopez *et al.* 2000). This heterogeneity was later found to be significant at high drug concentrations wherein increased poly(A) length contributed to higher mRNA stability leading to more abundant protein levels, efficiently negating the effect of the drug on the organism (Lopez-Camarillo *et al.* 2003). Use of heterogenous polyadenylation sites is also adapted by *P. berghei* for stage-specific translation of circumsporozoite transcripts, encoding an immunodominant coat protein in the infectious stage of the parasite (Le Roch *et al.* 2004). Similarly, in *T. vaginalis*, the authors suggest that the difference in the length of the poly (A) tail of Tvcp12 transcripts observed under different iron concentrations might account for their iron-dependent modulation of protein expression (Leon-Sicairos *et al.* 2004). Whether variation in length of poly (A) tail exists in other protozoan parasites and whether it is of relevance remains to be defined.

4.4 *miRNA-mediated regulation*

Micro-RNAs (miRNAs) are 22 nt long double-stranded RNA molecules that regulate gene expression by base pairing with sequences in the 3′-UTR of specific mRNA and attenuating translation (Lee and Ambros 2001; Grosshans and Slack 2002). Computational studies have predicted the presence of miRNA in *T. brucei* (Mallick *et al*. 2008), *E. histolytica* (De *et al.* 2006) and *T. vaginalis* (Lin *et al.* 2009), which regulate genes known to be vital for establishing infection. mIR2 (Saraiya and Wang 2008) and mIR5 (Li *et al.* 2011) are the two snoRNAs-derived miRNA that were found to regulate some transcripts in *G. lamblia.* miRNA-mediated regulation in protozoan parasites therefore serves as a means to modulate gene expression not only of the parasites but probably also of the host to make the environment more conducive for survival of the parasites (Liu *et al.* 2010).

Another interesting facet of miRNA-mediated gene regulation was observed in *P. falciparum* wherein the human miRNAs are translocated inside the parasite (intraerythrocytic stage). This thereby inhibits translation of parasite's transcripts which negatively affects their growth. Moreover, *P. falciparum* lacks Dicer/Ago orthologs, questioning the mode of action of these human miRNA inside the parasite. Further studies into the miRNA repertoire and their possible gene targets in the parasite would elucidate the dynamics of host-parasite interaction (LaMonte *et al.* 2012). This study thereby raises questions about the possibility of this phenomena in other intracellular parasites like *Leishmania spp.* which also lack an RNAi machinery.

5. Discussion

Regulation by UTRs of transcripts has considerable impact on the dynamics of protein expression profiles in protozoan parasites. The *cis*-acting elements along with their associated *trans*-acting factors mediate this regulation and this area has intrigued many researchers to understand their role in parasite's life cycle, especially in terms of pathogenesis. Increasing knowledge about the dependence of parasites on various *cis*-acting elements in UTRs highlights their significance; however, a clear picture of defined elements is still lacking (figure 4). This is partly contributed by our inadequate knowledge about the boundaries of individual genes and their UTRs, although the RNA-seq approach has helped in filling this gap. An UTR generally can harbour multiple elements which often work synergistically towards a common goal, and in certain cases their boundaries also overlap. Thus, a challenge in identification of such elements is that mutational/deletion analysis of one element might affect/alter the structure/function of the other. To identify these structure-disrupting mutations, *in silico* tools that predict them could be developed, thereby allowing precise analysis of numerous *cis*-acting elements in the UTRs of transcripts. A handful of RBPs have been identified in trypanosomatids and apicomplexans, and the knowledge is even more meager in other protozoan parasites. In many cases their targets are unknown and their role in translation regulation remains elusive. A genome wide analysis for the presence of common motifs in functional group of mRNAs in *T. brucei* has been carried out recently using computational tools, providing a platform for analysing the *trans*-acting factors that facilitate the coordinated regulation of genes. With the availability of genome sequence, this approach could be extended to other protozoan parasites, particularly those that share an evolutionary link amongst themselves. A single *cis*-acting element of UTRs might interact with many RBPs and vice versa, resulting in a complex RNA–protein interactome that might command the rapid changes in gene expression during the life cycle of protozoan parasites. An in-depth understanding of the

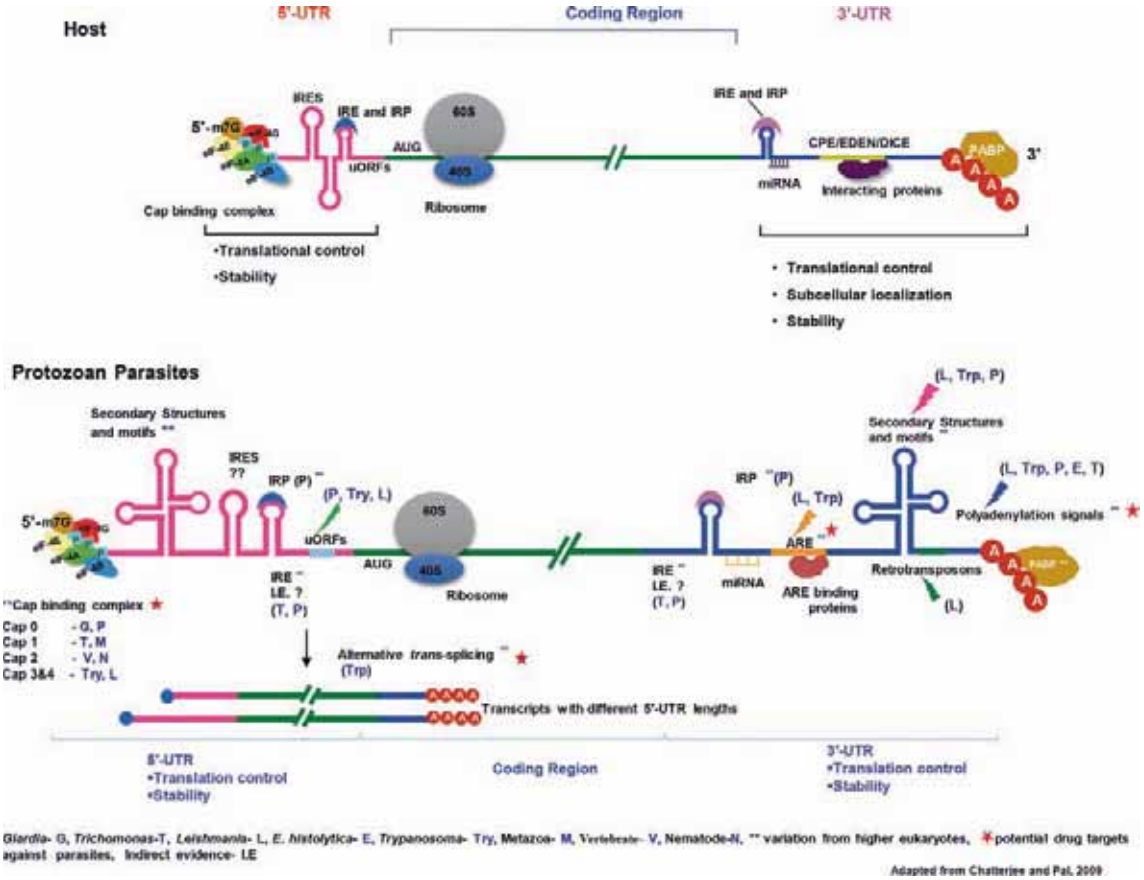

Figure 4. Comparative representation of possible translation regulation mechanisms in host and protozoan parasites. 5′-m7G, cap structure; eIF, eukaryotic initiation factor; CPE, cytoplasmic element; EDEN, embroyonic deadenylation signal; DICE, differential control element; IRE, Iron responsive element; IRES, internal ribosome entry site; PABP, poly(A)-binding protein. Possible sites of interaction of transacting factors (yet unknown) in the coding sequence: G, *Giardia*; L, *Leishmania*; M, Metazoan; N, Nematode; P, *Plasmodium*; T, *Trichomonas*; Try, *Trypanosma*; V, Vertebrates.

molecular network orchestrated by the UTRs of transcripts and strategies for disrupting it might open up new avenues for targeting these parasites.

Acknowledgements

We thank Dr Sunil Kumar Berwal, Dr Santosh Jadhav, Dr Radha Pujari and Ms Varsha Bhavnani for their critical reading and preparation of the manuscript. We would like to apologize to those whose original articles could not be cited due to space constraints and had to be replaced by reviews.

We acknowledge the financial support received from the Department of Science and Technology (DST), India (DST PURSE programme), Department Research and Development program from the Savitribai Phule Pune University, and Department of Biotechnology, India, (BT/PR5553/BRB/10/1086/2012) to JKP. SJR thanks Department of Biotechnology, India, for fellowship

References

Abanades DR, Ramirez L, Iborra S, Soteriadou K, Gonzalez VM, Bonay P, Alonso C and Soto M 2009 Key role of the 3' untranslated region in the cell cycle regulated expression of the Leishmania infantum histone H2A genes: minor synergistic effect of the 5' untranslated region. *BMC Mol. Biol.* **10** 48

Adam RD 2001 Biology of Giardia lamblia. *Clin. Microbiol. Rev.* **14** 447–475

Amulic B, Salanti A, Lavstsen T, Nielsen MA and Deitsch KW 2009 An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. *PLoS Pathog.* **5** e1000256

Bancells C and Dietsch KW 2013 A molecular switch in the efficiency of translation reinitiation controls the expression of

var2csa, a gene implicatedin pregnancy-associated malaria. *Mol. Microbiol.* **90** 472–488

Barrett LW, Fletcher S and Wilton SD 2012 Regulation of eukaryotic gene expression by untranslated gene regions and other non-coding elements. *Cell. Mol. Life Sci.* **69** 3613–3634

Benabdellah K, Gonzalez-Rey E and Gonzalez A 2007 Alternative trans-splicing of the Trypanosoma cruzi LYT1 gene transcript results in compartmental and functional switch for the encoded protein. *Mol. Microbiol.* **65** 1559–1567

Boucher N, Wu Y, Dumas C, Dube M, Sereno D, Breton M and Papadopoulou B 2002 A common mechanism of stage-regulated gene expression in Leishmania mediated by a conserved 3'-untranslated region element. *J. Biol. Chem.* **277** 19511–19520

Braks JA, Mair GR, Franke-Fayard B, Janse CJ and Waters AP 2008 A conserved U-rich RNA region implicated in regulation of translation in Plasmodium female gametocytes. *Nucleic Acids Res.* **36** 1176–1186

Brandao A and Jiang T 2009 The composition of untranslated regions in Trypanosoma cruzi genes. *Parasitol. Int.* **58** 215–219

Bruchhaus I, Leippe M, Lioutas C and Tannich E 1993 Unusual gene organization in the protozoan parasite Entamoeba histolytica. *DNA Cell Biol.* **12** 925–933

Bunnik EM, Doug Chung D, Hamilton M, Ponts N, Saraf A, Prudhhome J, Florens L and Le Roch KG 2013 Polysome profiling reveals translation control of gene expression in the human malarial parasite Plsmodium falciparum. *Genome Biol.* **14** R128

Calla-Choque JS, Figueroa-Angulo EE, Avila-Gonzalez L and Arroyo R 2014 alpha -Actinin TvACTN3 of Trichomonas vaginalis is an RNA-binding protein that could participate in its posttranscriptional iron regulatory mechanism. *Biomed. Res. Int.* **2014** 424767

Caro F, Ahyong A, Betegon M and DeRisi JL 2014 Genome-wide regulatory dynamics of translation in the Plasmodium falciparum asexual blood stages eLife. doi: 10.7554/eLife.04106

Carrington M and Krammer S 2014 An AU-rich instability element in the 3' UTR mediates an increase in mRNA stability in response to expression of a dhh1 ATPase mutant. Translation **2**. doi: 10.4161/trla.28587

Chatterjee S and Pal JK 2009 Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol. Cell.* **101** 251–262

Chatterjee S, Berwal SK and Pal, JK 2010 Pathological mutations in 5’ untranslated regions of human genes; in *Encyclopedia of life sciences* (ELS). John Wiley & Sons, Ltd: Chichester. doi: 10.1002/9780470015902.a0022408

Chow C, Cloutier S, Dumas C, Chou MN and Papadopoulou B 2011 Promastigote to amastigote differentiation of Leishmania is markedly delayed in the absence of PERK eIF2alpha kinase-dependent eIF2alpha phosphorylation. *Cell Microbiol.* **13** 1059–1077

Clayton C and Shapira M 2007 Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* **156** 93–101

Colasante C, Robles A, Li CH, Schwede A, Benz C, Voncken F, Guilbride DL and Clayton C 2007 Regulated expression of glycosomal phosphoglycerate kinase in Trypanosoma brucei. *Mol. Biochem. Parasitol.* **151** 193–204

Cowling VH and Cole MD 2010 Myc regulation of mRNA Cap methylation. *Genes Cancer* **1** 576–579

Cox FE 2002 History of human parasitology. *Clin. Microbiol. Rev.* **15** 595–612

David M, Gabdank I, Ben-David M, Zilka A, Orr I, Barash D and Shapira M 2010 Preferential translation of Hsp83 in Leishmania requires a thermosensitive polypyrimidine-rich element in the 3' UTR and involves scanning of the 5' UTR. *RNA* **16** 364–374

Davis-Hayman SR, Shah PH, Finley RW, Lushbaugh WB and Meade JC 2000 Trichomonas vaginalis: analysis of a heat-inducible member of the cytosolic heat-shock-protein 70 multigene family. *Parasitol. Res.* **86** 608–612

De Gaudenzi JG, Carmona SJ, Aguero F and Frasch AC 2013 Genome-wide analysis of 3'-untranslated regions supports the existence of post-transcriptional regulons controlling gene expression in trypanosomes. *PeerJ.* **1** e118

De S, Pal D and Ghosh SK 2006 Entamoeba histolytica: computational identification of putative microRNA candidates. *Exp. Parasitol.* **113** 239–243

Delhi P, Queiroz R, Inchaustegui D, Carrington M and Clayton C 2011 Is there a classical nonsense-mediated decay pathway in trypanosomes? *PLoS One.* **6** e25112

Descoteaux S, Ayala P, Samuelson J and Orozco E 1995 Increase in mRNA of multiple Eh pgp genes encoding P-glycoprotein homologues in emetine-resistant Entamoeba histolytica parasites. *Gene* **164** 179–184

Di Noia JM, D'Orso I, Sanchez DO and Frasch AC 2000 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.* **275** 10218–10227

Dillon LAL, Okrah K, Hughitt VK, Suresh R, Li Y, Fernandes MC, *et al*. 2015 Transcriptomic profiling of gene expression and RNA processing during Leishmania major differentiation. *Nucleic Acids Res.* **43** 6799–813

D'Orso I and Frasch AC 2001a TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. *J. Biol. Chem.* **276** 34801

D'Orso I and Frasch AC 2001b Functionally different AU- and G-rich cis-elements confer developmentally regulated mRNA stability in Trypanosoma cruzi by interaction with specific RNA-binding proteins. *J. Biol. Chem.* **276** 15783–15793

D'Orso I and Frasch AC 2002 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.* **277** 50520–50528

Droll D, Minia I, Fadda A, Singh A, Stewart M, Queiroz R and Clayton C 2013 Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. *PLoS Pathog.* **9** e1003286

Fennell C, Babbitt S, Russo I, Wilkes J, Ranford-Cartwright L, Goldberg DE and Doerig C 2009 PfeIK1, a eukaryotic initiation factor 2 alpha kinase of the human malaria parasite Plasmodium falciparum, regulates stress-response to amino-acid starvation. *Malar. J.* **8** 99

Fernandez-Moya SM, Carrington M and Estevez AM 2014 A short RNA stem-loop is necessary and sufficient for repression of gene expression during early logarithmic phase in trypanosomes. *Nucleic Acids Res.* **42** 7201–7209

Fiebig M, Gluenz E, Carrington M and Kelly S 2014 SLaP mapper: a webserver for identifying and quantifying spliced-leader addition andpolyadenylation site usage in kinetoplastid genomes. *Mol. Biochem. Parasitol.* **196** 71–74

Figueroa-Angulo EE, Calla-Choque JS, Mancilla-Olea MI and Arroyo R 2015 RNA-binding proteins in trichomonas vaginalis: atypical multifunctional proteins. *Biomolecules* **5** 3354–3395

Folgueira C, Quijada L, Soto M, Abanades DR, Alonso C and Requena JM 2005 The translational efficiencies of the two Leishmania infantum HSP70 mRNAs, differing in their 3'-untranslated regions, are affected by shifts in the temperature of growth through different mechanisms. *J. Biol. Chem.* **280** 35172–35183

Fuentes V, Barrera G, Sanchez J, Hernandez R and Lopez-Villasenor I 2012 Functional analysis of sequence motifs involved in the polyadenylation of Trichomonas vaginalis mRNAs. *Eukaryot. Cell.* **11** 725–734

Furger A, Schurch N, Kurath U and Roditi I 1997 Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of Trypanosoma brucei by modulating RNA stability and translation. *Mol. Cell Biol.* **17** 4372–4380

Ganesan K, Ponmee N, Jiang L, Fowble JW, White J, Kamchonwongpaisan S, Yuthavong Y, Wilairat P, *et al*. 2008 A genetically hard-wired metabolic transcriptome in Plasmodium falciparum fails to mount protective responses to lethal antifolates. *PLoS Pathog.* **4** e1000214

Garlapati S and Wang CC 2004 Identification of a novel internal ribosome entry site in giardiavirus that extends to both sides of the initiation codon. *J. Biol. Chem.* **279** 3389–97

Garlapati S and Wang CC 2009 Giardiavirus internal ribosome entry site has an apparently unique mechanism of initiating translation. *PLoS One.* **4** e7435

Gazestani VH, Lu Z and Salavati R 2014 Decipheing RNA regulatory elements in trypansomatids: one piece at a time or genome-wide? *Trends Parasitol.* **30** 234–240

Ginger ML 2006 Niche metabolism in parasitic protozoa. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361** 101–118

Gonzalez-Andrade P, Camara M, IIboudo H, Bucheton B, Jamonnaeu V and Deborggraeve S 2014 Diagnosis of trypanosomatids infections: targeting the spliced leader RNA. *J. Mol. Diagn.* **16** 400–4

Grosshans H and Slack FJ 2002 Micro-RNAs: small is plentiful. *J. Cell Biol.* **156** 17–21

Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, Cohen-Chalamish S, Biswas DK, Wachtel C, *et al*. 2013 The hnRNP F/H homologue of Trypanosoma brucei is differentially expressed in the two life cycle stages of the parasite and regulates splicing and mRNA stability. *Nucleic Acids Res.* **41** 6577–6594

Gupta SK, Chikne V, Eliaz D, Tkacz ID, Naboishcchikov I, Carmi S, Walsman BH and Michaeli S 2014 Two splicing factors carrying serine-arginine motifs, TSR1 and TSR1IP, reglate splicing, mRNA stsbility and rRNA processing in Trypanosoma brucei. *RNA Biol.* **11** 715–31

Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, *et al*. 2005 A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. *Science.* **307** 82–86

Hartmann C, Benz C, Brems S, Ellis L, Luu VD, Stewart M, D'Orso I, Busold C, *et al*. 2007 Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. *Eukaryot. Cell* **6** 1964–1978

Hehl A, Vassella E, Braun R and Roditi I 1994 A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in Trypanosoma brucei. *Proc. Natl. Acad. Sci. USA* **91** 370–404

Holzer TR, Mishra KK, LeBowitz JH and Forney JD 2008 Coordinate regulation of a family of promastigote-enriched mRNAs by the 3'UTR PRE element in Leishmania mexicana. *Mol. Biochem. Parasitol.* **157** 54–64

Hon CC, Weber C, Sismeiro O, Proux C, Koutero M, Deloger M, Das S, Agrahari M, *et al*. 2013 Quantification of stochastic noise of splicing and polyadenylation in Entamoeba histolytica. *Nucleic Acids Res.* **41** 1936–1952

Hotz HR, Lorenz P, Fischer R, Krieger S and Clayton C 1995 Role of 3'-untranslated regions in the regulation of hexose transporter mRNAs in Trypanosoma brucei. *Mol. Biochem. Parasitol.* **75** 1–14

Hotz HR, Hartmann C, Huober K, Hug M and Clayton C 1997 Mechanisms of developmental regulation in Trypanosoma brucei: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. *Nucleic Acids Res.* **25** 3017–3026

Jaeger LH and Brandao A 2011 The composition of upstream open reading frames (uORF) in four genes from Trypanosoma cruzi typical strains. *Parasitol. Res.* **109** 1205–1208

Jagus R, Bachvaroff TR, Joshi B and Place AR 2012 Diversity of eukaryotic translational initiation factor eIF4E in protists. *Comp. Funct. Genomics* **2012** 134839–134860

Jensen BC, Ramasamy G, Vasconcelos EJ, Ingolia NT, Myler PJ and Parsons M 2014 Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. *BMC Genomics* **15** 911

Joyce BR, Tampaki Z, Kim K, Wek RC and Sullivan WJ 2013 The unfolded protein response in the protozoan parasite Toxoplasma gondii features translational and transcriptional control. *Eukaryot. Cell* **12** 979–989

Knusel S and Roditi I 2013 Insights into the regulation of GPEET procyclin during differentiation from early to late procyclic forms of Trypanosoma brucei. *Mol. Biochem. Parasitol*. **191** 66–71

Kolev NG, Ullu E and Tschudi C 2014 The emerging role of RNA-inding prteins in the lifecycle of Trypanosoma brucei. *Cell. Microbiol.* **16** 482–9

Kramer S, Queiroz R, Ellis L, Webb H, Hoheisel JD, Clayton C and Carrington M 2008 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.* **121** 3002–3014

Lahav T, Sivam D, Volpin H, Ronen M, Tsigankov P, Green A, Holland N, Kuzyk M, *et al*. 2011 Multiple levels of gene regulation mediate differentiation of the intracellular pathogen Leishmania. *FASEB J.* **25** 515–525

LaMonte G, Philip N, Reardon J, Lacsina JR, Majoros W, Chapman L, Thornburg Courtney D, *et al*. 2012 Translocation of sickle cell erythrocyte MicroRNAs into Plasmodium falciparum inhibits parasite translation and contributes to malaria resistance. *Cell Host Microbe* **12** 187–199

Lanzer M, Wertheimer SP, de Bruin D and Ravetch JV 1993 Plasmodium: control of gene expression in malaria parasites. *Exp. Parasitol.* **77** 121–128

Lavstsen T, Salanti A, Jensen AT, Arnot DE and Theander TG 2003 Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar. J.* **2** 27

Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, Yan SF, Williamson KC, *et al*. 2004 Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome Res.* **14** 2308–2318

Lee RC and Ambros V 2001 An extensive class of small RNAs in Caenorhabditis elegans. *Science* **294** 862–864

Leon-Sicairos CR, Leon-Felix J and Arroyo R 2004 tvcp12: a novel Trichomonas vaginalis cathepsin L-like cysteine proteinase-encoding gene. *Microbiology.* **150** 1131–1138

Li L and Wang CC 2004 Capped mRNA with a single nucleotide leader is optimally translated in a primitive eukaryote, Giardia lamblia. *J. Biol. Chem.* **279** 14656–14664

Li W, Saraiya AA and Wang CC 2011 Gene regulation in Giardia lambia involves a putative microRNA derived from a small nucleolar RNA. *PLoS Negl. Trop. Dis.* **5** e1338

Li ZH, De Gaudenzi JG, Alvarez VE, Mendiondo N, Wang H, Kissinger JC, Frasch AC and Docampo R 2012 A 43-nucleotide U-rich element in 3'-untranslated region of large number of Trypanosoma cruzi transcripts is important for mRNA abundance in intracellular amastigotes. *J. Biol. Chem.* **287** 19058–19069

Liang XH, Haritan A, Uliel S and Michaeli S 2003 trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot. Cell.* **2** 830–840

Lin WC, Li SC, Shin JW, Hu SN, Yu XM, *et al*. 2009 Identification of microRNA in the protist Trichomonas vaginalis. *Genomics* **93** 487–493

Lioutas C and Tannich E 1995 Transcription of protein-coding genes in Entamoeba histolytica is insensitive to high concentrations of alpha-amanitin. *Mol. Biochem. Parasitol.* **73** 259–261

Liu Q, Tuo W, Gao H and Zhu XQ 2010 MicroRNAs of parasites: current status and future perspectives. *Parasitol. Res.* **107** 501–507

Lopez C, Marchat LA, Luna-Arias JP and Orozco E 2000 An initial characterization of the 3' untranslated region of the EhPgp5 mRNA in Entamoeba histolytica. *Arch. Med. Res.* **31** S282–284

Lopez-Camarillo C, Luna-Arias JP, Marchat LA and Orozco E 2003 EhPgp5 mRNA stability is a regulatory event in the Entamoeba histolytica multidrug resistance phenotype. *J. Biol. Chem.* **278** 11273–11280

Lopez-Rosas I, Orozco E, Marchat LA, Garcia-Rivera G, Guillen N, Weber C, Carrillo-Tapia E, de la Cruz H, *et al*. 2012 mRNA decay proteins are targeted to poly(A)+ RNA and dsRNA-containing cytoplasmic foci that resemble P-bodies in Entamoeba histolytica. *PLoS One.* **7**, e45966

Loyevsky M, Mompoint F, Yikilmaz E, Altschul SF, Madden T, Wootton JC, Kurantsin-Mills J, Kassim OO, *et al*. 2003 Expression of a recombinant IRP-like Plasmodium falciparum protein that specifically binds putative plasmodial IREs. *Mol. Biochem. Parasitol.* **126** 231–238

MacGregor P and Matthews KR 2012 Identification of the regulatory elements controlling the transmission stage-specific gene expression of PAD1 in Trypanosoma brucei. *Nucleic Acids Res.* **40** 7705–7717

Maga JA, Widmer G and LeBowitz JH 1995 Leishmania RNA virus 1-mediated cap-independent translation. *Mol. Cell Biol.* **15** 4884–4889

Mallick B, Ghosh Z and Chakrabarti J 2008 MicroRNA switches in Trypanosoma brucei. *Biochem. Biophys. Res. Commun.* **372** 459–463

McNicoll F, Muller M, Cloutier S, Boilard N, Rochette A, Dube M and Papadopoulou B 2005 Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in Leishmania J. *Biol. Chem.* **280** 35238–35246

Meleppattu S, Kamus-Elimelah D, Zinoviev A, Cohen-Mor S, Orr I and Shapira M 2015 The eIF3 complex of Leishmania-subunit composition and mode of recruitement to different cap-binding cmplexes. *Nucleic Acids Res.*. doi:10.1093/nar/gkv564

Michaeli S 2011 Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. *Future Microbiol.* **6** 459–474

Monk SL, Simmonds P and Matthews KR 2013 A short bifunctional element operates to positively or negatively regulate ESAG9 expression in different developmental forms of Trypanosoma brucei. *J. Cell Sci.* **126** 2294–2304

Moore LL, Santrich C and LeBowitz JH 1996 Stage-specific expression of the Leishmania mexicana paraflagellar rod protein PFR-2. *Mol. Biochem. Parasitol.* **80** 125–135

Moraes MC, Jesus TC, Hashimoto NN, Dey M, Schwartz KJ, Alves VS, Avila CC, Bangs JD, *et al*. 2007 Novel membrane-bound eIF2alpha kinase in the flagellar pocket of Trypanosoma brucei. *Eukaryot. Cell.* **6** 1979–1991

Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi I and Ochsenreiter T 2010 Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei. *PLoS Pathog.* **6** e1001037

Panneerselvam P, Bawankar P, Kulkarni S and Patankar S 2011 In silico prediction of evolutionarily conserved GC-rich elements associated with antigenic proteins of Plasmodium falciparum. *Evol. Bioinform.* **7** 235–255

Pasion SG, Hines JC, Ou X, Mahmood R and Ray DS 1996 Sequences within the 5' untranslated region regulate the levels of a kinetoplast DNA topoisomerase mRNA during the cell cycle. *Mol. Cell. Biol.* **16** 6724–6735

Perez-Diaz L, Pastro L, Smircich P, Dallagiovanna B and Garat B 2013 Evidence for a negative feedback control mediated by the 3' untranslated region assuring the low expression level of the RNA binding protein TcRBP19 in T. cruzi epimastigotes. *Biochem. Biophys. Res. Commun.* **436** 295–299

Quijada L, Soto M, Alonso C and Requena JM 1997 Analysis of post-transcriptional regulation operating on transcription products of the tandemly linked Leishmania infantum hsp70 genes. *J. Biol. Chem.* **272** 4493–4499

Ramaprasad A, Mourier T, Naeem R, Malas TB, Moussa E, Panigrahi A, et al 2015 Comprehensive evaluation of Toxoplasma gondii VEG and Neospora caninum LIV genomes with tachyzoite stage transcriptome and proteome defines novel transcript features. PLoS One **10**

Ramirez CA, Requena JM and Puerta CJ 2011 Identification of the HSP70-II gene in Leishmania braziliensis HSP70 locus: genomic organization and UTRs characterization. *Parasit. Vectors* **4** 166

Rastrojo A, Carrasco-Ramiro F, Martín D, Crespillo A, Reguera RM, Aguado B, *et al*. 2013 The transcriptome of Leishmania major in the axenic promastigote stage: transcript annotation and relative expression levels by RNA-seq. *BMC Genomics* **14** 223

Requena JM, Chicharro C, Garcia L, Parrado R, Puerta CJ and Canavate C 2012 Sequence analysis of the 3'-untranslated region of HSP70 (type I) genes in the genus Leishmania: its usefulness as a molecular marker for species identification. *Parasit. Vectors* **5** 87

Rettig J, Wang Y, Schneider A and Ochsenreiter T 2012 Dual targeting of isoleucyl-tRNA synthetase in Trypanosoma brucei is mediated through alternative trans-splicing. *Nucleic Acids Res.* **40** 1299–1306

Robles A and Clayton C 2008 Regulation of an amino acid transporter mRNA in Trypanosoma brucei. *Mol. Biochem. Parasitol.* **157** 102–106

Saraiya AA and Wang CC 2008 snoRNA, a novel precursor of microRNA in Giardia lamblia. *PLoS Pathog.* **4**, e1000224

Schurch N, Furger A, Kurath U and Roditi I 1997 Contributions of the procyclin 3' untranslated region and coding region to the regulation of expression in bloodstream forms of Trypanosoma brucei. *Mol. Biochem. Parasitol.* **89** 109–121

Sebastian S, Brochet M, Collins MO, Schwach F, Jones ML, Goulding D, Rayner JC, Choudhary JS, *et al*. 2012 A Plasmodium calcium-dependent protein kinase controls zygote development and transmission by translationally activating repressed mRNAs. *Cell Host Microbe.* **12** 9–19

Shaw PJ, Ponmee N, Karoonuthaisiri N, Kamchonwongpaisan S and Yuthavong Y 2007 Characterization of human malaria parasite Plasmodium falciparum eIF4E homologue and mRNA 5' cap status. *Mol. Biochem. Parasitol.* **155** 146–155

Shrimal S, Bhattacharya S and Bhattacharya A 2010 Serum-dependent selective expression of EhTMKB1-9, a member of Entamoeba histolytica B1 family of transmembrane kinases. *PLoS Pathog.* **6**, e1000929. doi:10.1371/journal.ppat.1000929

Siegel TN, Hon CC, Zhang Q, Lopez-Rubio JJ and Scheidig-Benatar C, *et al*. 2014 Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in Plasmodium falciparum. *BMC Genomics* **15** (1) 150

Simoes-Barbosa A, Hirt RP and Johnson PJ 2010 A metazoan/plant-like capping enzyme and cap modified nucleotides in the unicellular eukaryote Trichomonas vaginalis. *PLoS Pathog.* **6**, e1000999

Singh A, Minia I, Droll D, Fadda A, Clayton C and Erben E 2014 Trypanosoma MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. *Nucleic Acids Res.* **42** 4652–4668

Smandi S, Guerfali FZ, Farhat M, Ben-Aissa K, Laouini D, Guizani-Tabbane L, Dellagi K and Benkahla A 2012 Methodology optimizing SAGE library tag-to-gene mapping: application to Leishmania BMC Res. *Notes* **5** 74

Solano-Gonzalez E, Burrola-Barraza E, Leon-Sicairos C, Avila-Gonzalez L, Gutierrez-Escolano L, Ortega-Lopez J and Arroyo R 2007 The trichomonad cysteine proteinase TVCP4 transcript contains an iron-responsive element. *FEBS Lett.* **581** 2919–2928

Sonenberg N and Hinnebusch AG 2009 Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136** 731–745

Stevens SG, Gardner PP and Brown C 2011 Two covariance models for iron-responsive elements. *RNA Biol.* **8** 792–801

Suganuma K, Mochabo KM, Hakimi H, Yamasaki S, Yamagishi J, Asada M, Kawazu S and Inoue N 2013 Adenosine-uridine-rich element is one of the required cis-elements for epimastigote form stage-specific gene expression of the congolense epimastigote specific protein Mol. *Biochem. Parasitol.* **191** 36–43

Torres-Romero JC and Arroyo R 2009 Responsiveness of Trichomonas vaginalis to iron concentrations: evidence for a post-transcriptional iron regulation by an IRE/IRP-like system. *Infect. Genet. Evol.* **9** 1065–1074

Urwyler S, Vassella E, Van Den Abbeele J, Renggli CK, Blundell P, Barry JD and Roditi I 2005 Expression of procyclin mRNAs during cyclical transmission of Trypanosoma brucei. *PLoS Pathog.* **1** e22

Vasquezz JJ, Hon CC, Vanselow JT, Schlosser A and Siegel TN 2014 Comparative ribosome profiling reveals extensive translation complexity indifferent Trypanosoma brucei life cyle stages. *Nuclei Acids Res.* **42** 3623–37

Vassella E, Probst M, Schneider A, Studer E, Renggli CK and Roditi I 2004 Expression of a major surface protein of Trypanosoma brucei insect forms is controlled by the activity of mitochondrial enzymes. *Mol. Biol. Cell.* **15** 3986–3993

Walrad P, Paterou A, Acosta-Serrano A and Matthews KR 2009 Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA cis-elements. *PLoS Pathog.* **5** e1000317

Walrad PB, Capewell P, Fenn K and Matthews KR 2012 The post-transcriptional trans-acting regulator, TbZFP3, co-ordinates transmission-stage enriched mRNAs in Trypanosoma brucei. *Nucleic Acids Res.* **40** 2869–2883

Watanabe J, Sasaki M, Suzuki Y and Sugano S 2002 Analysis of transcriptomes of human malaria parasite Plasmodium falciparum using full-length enriched library: identification of novel genes and diverse transcription start sites of messenger RNAs. *Gene* **291** 105–113

Wilson K, Uyetake L and Boothroyd J 1999 Trypanosoma brucei: cis-acting sequences involved in the developmental regulation of PARP expression. *Exp. Parasitol.* **91** 222–30

Yamagishi J, Watanabe J, Goo YK, Masatani T, Suzuki Y and Xuan X 2012 Characterization of Toxoplasma gondii 5' UTR with encyclopedic TSS information. *J. Parasitol.* **98** 445–447

Yoffe Y, Leger M, Zinoviev A, Zuberek J, Darzynkiewicz E, Wagner G and Shapira M 2009 Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions. *Nucleic Acids Res.* **37** 3243–353

Zamorano A, Lopez-Camarillo C, Orozco E, Weber C, Guillen N and Marchat LA 2008 In silico analysis of EST and genomic sequences allowed the prediction of cis-regulatory elements for Entamoeba histolytica mRNA polyadenylation. *Comput. Biol. Chem.* **32** 256–263

Zamudio JR, Mittra B, Campbell DA and Sturm NR 2009 Hypermethylated cap 4 maximizes Trypanosoma brucei translation. *Mol. Microbiol.* **72** 1100–1110

Zhao J, Hyman L and Moore C 1999 Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63** 405–445

Zinoviev A and Shapira M 2012 Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids. *Comp. Funct. Genomics* **2012** 813718

Zinoviev A, Leger M, Wagner G and Shapira M 2011 A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways. *Nucleic Acids Res.* **39** 8404–8415

Zinoviev A, Manor S and Shapira M 2012 Nutritional stress affects an atypical cap-binding protein in Leishmania. *RNA Biol.* **9** 1450–1460

Zubiaga AM, Belasco JG and Greenberg ME 1995 The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediated mRNA degradation. *Mol. Cell. Biol.* **15** 2219–2230

MS received 12 July 2016; accepted 14 December 2016

Corresponding editor: Sudha Bhattacharya