The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct

Rafael Dhalia, Nina Marinsek¹, Christian R. S. Reis, Rodolfo Katz, João R. C. Muniz², Nancy Standart¹, Mark Carrington³ and Osvaldo P. de Melo Neto*

Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Avenue Moraes Rego s/n, Campus UFPE, Recife PE 50670-420, Brazil, ¹Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK and ²Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, São Carlos SP 13560-970, Brazil

Received February 1, 2006; Revised March 4, 2006; Accepted April 6, 2006

ABSTRACT

Protozoan parasites belonging to the family *Trypanosomatidae* are characterized by an unusual pathway for the production of mRNAs via polycistronic transcription and *trans*-splicing of a 5' capped mini-exon which is linked to the 3' cleavage and polyadenylation of the upstream transcript. However, little is known of the mechanism of protein synthesis in these organisms, despite their importance as agents of a number of human diseases. Here we have investigated the role of two *Trypanosoma brucei* homologues of the translation initiation factor eIF4A (in the light of subsequent experiments these were named as *Tb*EIF4AI and *Tb*EIF4AIIi). eIF4A, a DEAD-box RNA helicase, is a subunit of the translation initiation complex eIF4F which binds to the cap structure of eukaryotic mRNA and recruits the small ribosomal subunit. *Tb*EIF4AI is a very abundant predominantly cytoplasmic protein (over 1 × 10⁵ molecules/cell) and depletion to ~10% of normal levels through RNA interference dramatically reduces protein synthesis one cell cycle following double-stranded RNA induction and stops cell proliferation. In contrast, *Tb*EIF4AIIi is a nuclear, moderately expressed protein (~1–2 × 10⁴ molecules/cell), and its depletion stops cellular proliferation after approximately four cell cycles. Ectopic expression of a dominant negative mutant of *Tb*EIF4AI, but not of *Tb*EIF4AII, induced a slow growth phenotype in transfected cells. Overall, our results suggest that only *Tb*EIF4AI is involved in protein synthesis while the properties and sequence of *Tb*EIF4AIIi indicate that it may be the orthologue of eIF4AIII, a component of the exon junction complex in mammalian cells.

INTRODUCTION

The flagellate protozoan parasites belonging to the family *Trypanosomatidae* include a number of important pathogens responsible for diseases of worldwide impact such as the Sleeping Sickness (*Trypanosoma brucei*), Chagas’ Disease (*Trypanosoma cruzi*) and the various forms of Leishmaniasis (*Leishmania* sp.) (www.who.int/tdr). These organisms are unusual in a number of processes necessary for mRNA synthesis and maturation; transcription is polycistronic and monocistronic mRNAs arise after *trans*-splicing of a capped short exon on to the 5’ end and cleavage and polyadenylation at the 3’ end [reviewed in (1,2)]. As a result of *trans*-splicing, the 5’ ends of mature trypanosomatid mRNAs all share the same 39 nt leader sequence with a modified cap 4 structure (3). To date, little is known about how these mRNAs are translated, if major differences exist within the process of protein synthesis when compared with other eukaryotes and whether the common leader sequence influences how the mRNAs are recruited for translation.

In eukaryotes, protein synthesis is a complex process which requires a myriad of different macromolecules including RNAs and proteins. The critical initiation step requires a number of translation initiation factors (eIFs) whose activity can be highly regulated [for reviews see (4–7)]. Paramount within these factors is the heterotrimeric eIF4F complex, which is required for the recruitment of the small ribosomal subunit to the 5’ end of the mRNA. eIF4F is composed of the RNA helicase eIF4A, the cap-binding protein eIF4E and the large scaffolding protein eIF4G which mediates interactions between eIF4F and other translation factors as well as the small ribosomal subunit [reviewed in (8)].

*To whom correspondence should be addressed. Tel: 55 81 2101 2636; Fax: 55 81 3453 2449; Email: opmn@cpqam.fiocruz.br

© The Author 2006. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org
elf4A is the prototype member of the DEAD-box family of RNA helicases which includes several proteins mainly involved in RNA metabolism. These proteins are classified within the superfamily II of a much larger group of related RNA and DNA helicases (9). The RNA helicases couple the hydrolysis of ATP to various activities relevant for RNA function such as rearrangement of inter- or intra-molecular RNA structures, dissociation of RNA–protein complexes and RNA unwinding. The DEAD-box family members are characterized by nine sequence motifs (I, Ia, Ib, II, III, IV, V and VI and the Q motif), as well as several individual amino acids, conserved among the various proteins assigned to this family [for reviews see (10,11)]. Structurally, elf4A assumes a ‘dumbbell’ shape with two globular domains connected by a flexible linker (12). Comparison with the structure of related RNA and DNA helicases and the Methanococcus jannaschii DEAD-box protein, similar in size to elf4A (13), have confirmed the basic overall structure of the core helicase domains. The various conserved motifs are positioned in the interface between the two domains and have been involved in RNA binding and ATP binding and hydrolysis. However, little is known about the molecular basis for RNA specificity and helicase function [reviewed in (11,14)]. In translation initiation, elf4A binds to the central region of elf4G, via the elf4G HEAT domain (15) and, in mammals at least, also to the elf4G C-terminus (16,17). elf4A seems to be responsible for melting secondary structures along the mRNA 5'-untranslated region (5'-UTR), facilitating the binding of the small ribosomal subunit and the scanning of the leader region to locate the initiation codon (18,19) [reviewed in (4,6)].

In mammals three different isoforms of elf4A have been described. Both elf4AI and II (90% identity between the two proteins) are able to reconstitute the elf4F subunit and presumably have similar roles in translation (20,21). In contrast, elf4AIII, only 66% identical to mammalian elf4A1, is functionally distinct. While elf4AIII exhibits RNA-dependent ATPase activity and ATP-dependent RNA helicase activity, it does not support binding of the small ribosomal subunit to the mRNA, and inhibits translation in vitro (22). elf4AIII localizes to the nucleus (23) and recent reports indicate that it may act as an anchoring factor for the exon junction complex (EJC), and is essential for nonsense-mediated decay (NMD) in mammals (24–30).

The mechanisms of translation initiation are virtually unknown in trypanosomatids. A Leishmania elf4A homologue (called LeiF) was first described in Leishmania braziliensis and Leishmania major as a 45.3 kDa antigen, expressed in both insect and mammalian stages of the parasite life cycle, but its role in translation was not investigated (31,32). Recently, our group has identified multiple L.major homologues for the three elf4F subunits, all of which are conserved in T.brucei (33). We characterized two putative L.major elf4A homologues, LmElf4A1 (LeiF) and LmElf4A2, with 59 and 52% identities to human elf4A1, respectively. When assayed with isoform specific antibodies these two factors differ significantly in abundance in L.major promastigotes. LmElf4A1 is very abundant with over 10^5 molecules/cell whilst LmElf4A2 is either absent or present at levels below 10^4 molecules/cell. Furthermore, only LmElf4A1 was found to bind specifically to the HEAT domain of one of the Leishmania elf4G homologues (33).

In this paper we take advantage of the genetic tools available for the study of gene function in T.brucei to extend this analysis of the two trypanosomatid elf4A homologues. Initially, the mRNA and protein levels of the two T.brucei elf4A orthologues were analysed during the life cycle. Their intracellular localization was identified through overexpression of enhanced yellow fluorescent protein (EYFP) fusions and their role for parasite viability investigated through RNA interference and overexpression of dominant negative mutants. Our results show that the T.brucei orthologue of LmElf4A1 (named as TbElf4A1) is the functional homologue of the elf4A present in elf4F. As for the orthologue of LmElf4A2, it seems to be the functional homologue of the nuclear elf4AIII present in higher eukaryotes and has been named here as TbElf4AIII.

MATERIALS AND METHODS

Sequence analysis and molecular modeling

BLAST searches were carried out with the T.brucei genome sequences available at the Gene DB website of the Sanger Institute Pathogen Sequencing Unit (www.genedb.org). Further sequence searches, Clustal W alignments and molecular modeling were done as described previously (33).

PCR and cloning methods

The TbElf4A1 coding sequence was amplified from T.brucei Lister 427 genomic DNA (5'primer, AAG CTG CCG CCA CCA TGG CCC AAC AAG GAA AG; and 3'primer, GGA TCC AGA ACC CTC ACC AAG GTA GGC AGC; added restriction sites used in cloning are underlined) resulting in the entire open reading frame (ORF) flanked by sites for the enzymes HindIII and BamHI. The same strategy was used for the amplification of the TbElf4AIIII sequence (5'primer, AAG CTT CCG CCA CCA TGA CAG CAA CCG CAA GG; and 3'primer, GGA TCC AGA ACC CCA GAA CTG TTC ACC GAC GTT TG). The amplified fragments were then cloned into the vector pGEM-T Easy (Promega) and sequenced. In order to express N-terminal His-tagged fusion proteins both fragments were then recovered by digestion with HindIII and BamHI and subcloned into the same sites of a modified pET15b vector. To generate the TbElf4A1-EYFP and TbElf4AIIII-EYFP constructs, the two elf4A fragments were cloned into the HindIII and BamHI sites of p2215, a modified form of pLEW82 (34). To make p2215, the EYFP ORF (Clontech) was obtained as a BamHI/BglII fragment and inserted into the BamHI site of the RNAi experiments the same two elf4A DNA fragments were also subcloned into the HindIII/BamHI sites of pGEM-T Easy (Promega) and sequenced. In order to express N-terminal His-tagged fusion proteins both fragments were then recovered by digestion with HindIII and BamHI and subcloned into the same sites of a modified pET15b vector. To generate the TbElf4A1-EYFP and TbElf4AIIII-EYFP constructs, the two elf4A fragments were cloned into the HindIII and BamHI sites of p2215, a modified form of pLEW82 (34). To make p2215, the EYFP ORF (Clontech) was obtained as a BamHI/BglII fragment and inserted into the BamHI site of pGEM-T Easy (Promega) and sequenced. In order to express N-terminal His-tagged fusion proteins both fragments were then recovered by digestion with HindIII and BamHI and subcloned into the same sites of a modified pET15b vector. To generate the TbElf4A1-EYFP and TbElf4AIIII-EYFP constructs, the two elf4A fragments were cloned into the HindIII and BamHI sites of p2215, a modified form of pLEW82 (34). To make p2215, the EYFP ORF (Clontech) was obtained as a BamHI/BglII fragment and inserted into the BamHI site of
Parasite growth, transfection and RNAi

Procyclic form *T. brucei* Lister 427 cells were used throughout. RNAi and ectopic expression of eIF4A were performed using *T. brucei* Lister 427 29–13, containing integrated copies of pLEW 29 and pLEW13 (34). Procyclic *T. brucei* forms were propagated in SDM-79 medium at 27°C, supplemented with 10% fetal calf serum (FCS). For the 29–13 cell line, cultures were also supplemented with G418 (15 mg/ml) and hygromycin (25 mg/ml). Parasite growth was monitored microscopically every 24 h. Mid-log phase cultures (10^6–10^7 cells/ml) were then used for transfection and total protein extract production. Bloodstream forms (Lister 427) were cultivated in HMI-9 medium (37) at 37°C, 5% CO₂, supplemented with 10% FCS. Cultures grown to mid-log phase cultures (10^5–10^6 cells/ml) were also harvested for the production of total protein extract.

Plasmids were linearized with NotI prior to electroporation and stable DNA integration was selected using phleomycin (2.5 mg/ml). For the RNAi experiments 1 mg/ml of tetracycline was added to mid-log phase cultures of transfected cells.

**RNA analysis**

RNA extraction and Northern blots were performed using standard methods (38,39). DNA fragments containing complete ORFs were used as probes for *Tb* EIF4As and EP procyclin. A genomic repeat containing both α- and β-tubulin genes was used to detect tubulin mRNA.

**Recombinant protein expression, antibody production and western blots**

His-tagged *Tb*EIF4A1 and *Tb*EIF4AIII were expressed in *Escherichia coli* BL21 Star (DE3) using pET15b derived plasmids. The recombinant polypeptides were insoluble after lysing the cells using a French Press. The polypeptides were purified by preparative SDS–PAGE and the bands corresponding to the recombinant proteins were then excised and sent for the production of total protein extract.

Plasmids were linearized with NotI prior to electroporation and stable DNA integration was selected using phleomycin (2.5 mg/ml). For the RNAi experiments 1 mg/ml of tetracycline was added to mid-log phase cultures of transfected cells.
with known concentrations of BSA (data not shown). After quantification they were then used in western blots with the respective antisera and compared with serial dilutions of total protein extract from both procyclic and bloodstream forms of \textit{T. brucei}. The endogenous protein levels were then estimated by the densitometric analysis of the western blot results as described elsewhere (33).

Fluorescence microscopy

For the indirect immunofluorescence assay, wild-type procyclic cells grown to mid-log phase (5 × 10^6/ml) were harvested, washed with phosphate-buffered saline (PBS)/10 mM glucose and adsorbed to polylysine coated slides. The cells were then fixed in 100% methanol at −20°C/15 min. Antibody detection of \textit{Tb}EIF4AI and III followed standard procedures. DNA was stained using Hoechst 33258. For the analysis of the cells expressing \textit{Tb}EIF4AI-EYFP and \textit{Tb}EIF4AIII-EYFP, aliquots of 5 × 10^5 cells were harvested, washed with PBS/10 mM glucose and fixed in 0.1% formaldehyde for 5 min. In this case, DNA was stained using Hoechst 33342.

Metabolic labelling

To measure the rate of protein synthesis, [35S]methionine (10 µCi/ml) was added to mid-log cultures which were incubated for 1 h prior to the determination of trichloroacetic acid precipitable incorporation into protein. Parallel incubations in the presence of 50 µg/ml cycloheximide were used to estimate incorporation of radiolabel by processes other than cytoplasmic protein synthesis. For metabolic labeling, cultures were washed twice with methionine-free RPMI 1640 medium and then resuspended at 1 × 10^7 cells/ml in methionine-free RPMI 1640 containing 50 µCi/ml [35S]methionine and incubated for 1 h at 28°C prior to harvesting and analysis by SDS–PAGE and autoradiography.

RESULTS

Identification of the \textit{T. brucei} eIF4A homologues

The \textit{T. brucei} homologues of eIF4A were identified in searches of the genome sequence using human eIF4A1 as well as the two \textit{Leishmania} eIF4A sequences. At the amino acid level, the two \textit{T. brucei} proteins, \textit{Tb}EIF4AI and \textit{Tb}EIF4AIII, are very similar to their \textit{L. major} orthologues with identities of 88% for the \textit{Tb}EIF4AI and 82% for the \textit{Tb}EIF4AIII pair (to avoid confusion and in view of the data presented below the \textit{L. major} eIF4A homologues, previously called \textit{Lm}EIF4A1 and 2, have also been renamed to \textit{Lm}EIF4AI and \textit{Lm}EIF4AIII and this nomenclature will be used when needed). The assignment as \textit{Tb}EIF4AI homologues reflects the fact that the two sequences are the closest matches in the two parasite protein databases to human eIF4AI and both share identities of over 50% with the human protein. The third nearest eIF4A homologue in both \textit{T. brucei} and \textit{L. major} databases has been assigned to another group of RNA helicases, Dhh1 (33), with an identity of only ~40% to human eIF4AI. In order to analyse the conservation of the putative eIF4A homologues within an additional member of the family \textit{Trypanosomatidae}, we performed similar searches using the \textit{T. cruzi} genome database. Again, orthologues to both proteins could be found in \textit{T. cruzi} with the third nearest match to human eIF4AI being Dhh1.

Figure 1 shows a sequence alignment comparing the two eIF4A sequences from \textit{T. brucei} with those from \textit{T. cruzi} and \textit{L. major}. Highlighted in the figure are the various conserved motifs typical of eIF4A and related proteins which have been shown to be required for different aspects of the RNA helicase activity. Motifs I, II, VI and the recently identified Q motif (41) have been implicated in ATP binding and hydrolysis; motif III may link nucleotide hydrolysis to helicase function; motifs Ia, Ib, IV and V may be involved in RNA binding [reviewed in (11,14)]. Several conserved arginine residues, which have also been implicated as important for eIF4A/helicase function in yeast eIF4A (12), as well as a conserved N-terminal phenylalanine residue are also shown (42). Overall the alignment confirms the close similarity between the various homologues. In general the N-terminal half of the protein is less conserved than the C-terminal half but only in the very N-terminus are significant differences in the sequences observed. A few conserved differences between the three \textit{Tb}EIF4AI and three \textit{Tb}EIF4AIII homologues with potential significance for their function can be identified, such as the replacement of a conserved F46 E47 doublet within the Q motif of eIF4AI by YK in eIF4AIII proteins. Other individual substitutions conserved between the eIF4AI and eIF4AIII homologues can be seen within motifs Q (S50T, S51A), I (Q71S), IV (A275C), V (K499R) and VI (G495T). So far however the functional significance of these substitutions is unknown.

Expression of \textit{Tb}EIF4A mRNAs in \textit{T. brucei} bloodstream and procyclic forms

To begin the functional characterization of \textit{T. brucei} eIF4A homologues and assay their expression at the mRNA level, the two genes were amplified, cloned and used as probes in northern blots of RNA from \textit{T. brucei} procyclic and bloodstream forms (Figure 2A). The membranes were also probed for the constitutively expressed tubulin, to confirm that equal amounts of mRNA were loaded in each lane, and for the procyclic-specific EP procyclin mRNA to verify the stage specificity of both sets of mRNAs (43).

The two \textit{T. brucei} eIF4A mRNAs were readily detected and found to be at constant levels throughout the parasite life cycle. However, according to the northern blot, the \textit{Tb}EIF4AI mRNA produces a much stronger signal than \textit{Tb}EIF4AIII. Since both probes used were of similar specific activity, and the exposures times for the films were similar as well, it seems that the \textit{Tb}EIF4AIII mRNA is far less abundant than \textit{Tb}EIF4AI. Remarkably, although the ORFs for both proteins are similar, 1215 versus 1206 bp for \textit{Tb}EIF4AI and \textit{Tb}EIF4AIII, respectively, their mRNAs differ significantly, with the \textit{Tb}EIF4AI message, at ~3 kb, being nearly twice the length of \textit{Tb}EIF4AIII (~1.6 kb), probably reflecting a considerable difference in the length of the 3'-UTR.

\textit{Tb}EIF4AI is about 10-fold more abundant than \textit{Tb}EIF4AIII in both procyclic and bloodstream forms

Recombinant His-\textit{Tb}EIF4AI and III were expressed in \textit{E. coli}, purified from inclusion bodies and used to produce antisera. Affinity purification and depletion was used to produce specific antibodies for each protein (see below, Figure 6A). The
Expression analysis of \( T.brucei \) procyclic (PCF) and bloodstream (BSF) forms was separated on denaturing gels and used in northern blot assays to detect the expression of \( T.brucei \) EIF4AI and III. One of the blots was probed with tubulin (ubiquitously expressed) and EP procyclin (expressed in procyclins only) as controls. The migration of RNA size markers is indicated on the left in kilobases. (B) Quantification of \( T.brucei \) EIF4AI and EIF4AIII in the procyclin and bloodstream forms of \( T.brucei \). Recombinant His-tagged EIF4AI and III were quantified, diluted to defined concentrations (in fmol) and ran on SDS–PAGE gels with whole parasite extract obtained from known numbers of cells from both procyclin and bloodstream forms (1.25 \( \times \) 10\(^5\)–2 \( \times \) 10\(^5\) \( T.brucei \) EIF4AI and 1.25 \( \times \) 10\(^3\)–2 \( \times \) 10\(^3\) \( T.brucei \) EIF4AIIII). The proteins samples were then transferred to Immobilon-P membranes followed by incubation with the affinity purified isoform specific antisera and goat anti-rabbit IgG conjugated with peroxidase, and detection by ECL. The values obtained for the abundance of the two proteins in fentomoles/cell are shown in Table 1. Overall we conclude that both proteins are expressed constitutively and that only the \( T.brucei \) EIF4AI levels are compatible with a role in translation.

Subcellular localization of \( T.brucei \) EIF4AI and III

To determine the subcellular localization of the \( T.brucei \) EIF4AI homologues we used two different experimental approaches. First the \( T.brucei \) EIF4AI and EIF4AIII ORFs were cloned into the vector p2215 and the construct integrated into the non-transcribed spacer of a ribosomal RNA gene locus in the procyclin cell line Lister 427 29-13. This resulted in a tetracycline-inducible transgene encoding the EIF4AI fused at the C-terminus to EYFP. Expression of both constructs was first verified by western blotting and similar levels of expression were observed for both \( T.brucei \) EIF4AI and III-EYFP fusion proteins (data not shown). The fluorescent proteins were visualized by microscopy (Figure 3) and strikingly, the two proteins localize differentially within cells. \( T.brucei \) EIF4AI-EYFP is found predominantly in the cytoplasm, whilst \( T.brucei \) EIF4AIIII-EYFP is only found in the nucleus. These results were confirmed for the endogenous proteins through indirect immunofluorescence using isoform specific antibodies: again \( T.brucei \) EIF4AI was mainly found in the cytoplasm whilst \( T.brucei \) EIF4AIIII was only detected in the nucleus (Figure 3).

RNAi of \( T.brucei \) EIF4AI and EIF4AIII

The function of the two eIF4AI homologues was then investigated by knock down of expression through RNA interference. First, both ORFs were subcloned into the vector p2T7-177 vector (35) and the constructs integrated into the procyclin cell line Lister 427 29-13 resulting in cell lines with tetracycline-inducible expression of double-stranded RNA.

Cell proliferation was reduced within 24 h and ceased around 48 h after induction of \( T.brucei \) EIF4AI RNAi (Figure 4A) and the cell density increased by ~3-fold during this time. Western blotting over a time course after addition of tetracycline showed that the level of the protein decreased to <10% of the starting level but expression was not completely ablated (Figure 4B). Protein synthesis after induction of \( T.brucei \) EIF4AI RNAi was monitored in two ways: (i) metabolic labelling to identify any alterations in the complement of polypeptides synthesized, and (ii) the rate of total protein synthesis was measured. There were no substantial changes in the profile of proteins synthesized although a small number of polypeptides appear to be relatively less affected by \( T.brucei \) EIF4AI depletion. The overall rate of protein synthesis had halved by ~22 h, the time at which cell proliferation ceased and was reduced to <20% of the uninduced control by 48 h as shown in Figure 4C.
The phenotype of cells after RNAi ablation of *TbEIF4AIII* was different. These cells only showed a dramatic reduction in the rate of proliferation 3 days after induction of RNAi, during which time the cell density increased ~20-fold (Figure 5A). Levels of *TbEIF4AIII* fell dramatically during the first 24 h of RNAi (Figure 5B), and the protein was only just detectable in extracts derived from cells at the 48 h time point. It is possible that, owing to its low abundance even in wild-type cells, residual levels of *TbEIF4AIII* persist longer than 48 h in the cells after RNAi although they are not detected by the western blotting assay. These residual levels would be responsible for the delayed onset of the growth phenotype. The western blotting results also confirm that lack of *TbEIF4AIII* is not involved in the phenotype induced by the depletion of *TbEIF4AII* since no reduction in levels of *TbEIF4AII* was observed in cells after RNAi although they are not detected by the western blotting assay. These residual levels would be responsible for the delayed onset of the growth phenotype. The western blotting results also confirm that lack of *TbEIF4AIII* is not involved in the phenotype induced by the depletion of *TbEIF4AII* since no reduction in levels of *TbEIF4AII* was observed in cells submitted to the *TbEIF4AII* RNAi procedure (Figure 5B). Likewise the *TbEIF4AIII* RNAi does not lead to any reduction in the levels of *TbEIF4AII* (Figure 4B). These results are compatible with *TbEIF4AIII* being required only at very low levels so that many cell cycles are required after addition of tetracycline to impair cell growth. In contrast, the levels of *TbEIF4AII*, despite its abundance in wild-type cells, are much more sensitive to RNAi mediated depletion, consistent with a role in overall protein synthesis.

**Expression of dominant negative mutants of *TbEIF4AII* and *TbEIF4AIII* in transfected procyclic cells**

The helicase activity of eIF4A is essential for protein synthesis and viability and some mutations that abolish its activity can

---

**Figure 3.** Subcellular localization of *TbEIF4AII* and III in *T. brucei* procyclic forms. Subcellular localization of the *TbEIF4AII* and III/YFP fusion proteins in transfected *T. brucei* cells was examined with a fluorescence microscope. The localization of native *TbEIF4AII* and III was also confirmed in wild-type procyclic cells (WT 427) by indirect immunofluorescence using the *TbEIF4AII* or *TbEIF4AIII* specific antibodies followed by incubation with the fluorescein-conjugated secondary antibody. Where indicated, the cells were counterstained to locate the nuclear and kinetoplast DNA. Note lack of *TbEIF4AIII* staining of the kinetoplast.

---

**Figure 4.** RNAi of *TbEIF4AII*. Procyclic *T. brucei* cells were transfected with the p2T7-177 derived plasmid containing the *TbEIF4AII* gene. Transfected cells were selected after growth in the presence of phleomycin and RNA interference induced after tetracycline addition. At regular intervals, cellular growth was monitored by counting the number of viable cells, expression of *TbEIF4AII* assayed and total protein synthesis investigated by [35S]methionine incorporation. (A) Cell density of transfected cultures with and without tetracycline addition. (B) Western blot analysis of the time course. Note the various dilutions of total cell extract for comparison (1–1/32 cell equivalent—1 cell equivalent equals to 10^6 cells and was used in the various RNAi lanes). *TbEIF4AII* was detected with the affinity purified antiserum and anti-BiP was used as a loading control. The same blot was probed with both antibodies. Equivalent extracts of cells transfected with the p2T7-177/*TbEIF4AIII* construct (see also Figure 5) were also used in the blot to monitor for *TbEIF4AII* levels. (C) [35S]methionine incorporation profile in transfected cells grown without tetracycline or 24 and 48 h after its addition. Total protein synthesis was estimated after RNAi for *TbEIF4AII* by incubating aliquots of the cells in the presence of [35S]methionine for 1 h followed by TCA precipitation, quantitation of the incorporated radioactivity or SDS–PAGE followed by autoradiography of the selected samples.
act as dominant negative mutants. Wild-type *T. brucei* helicases and equivalent DEAD-box mutant transgenes, in which the glutamic acid residue in the DEAD motif II (Figure 1) was substituted with a glutamine (DEAD→DQAD), were expressed using a tetracycline-inducible promoter. This mutation induces a dominant negative phenotype in mammalian eIF4AI, resulting in potinhibition of protein synthesis and is widely used to abrogate the function of DEAD-box proteins (11,36). The transgenes encoded a C-terminal triple myc tag to distinguish the expression of the transgene from the endogenous protein. Wild-type and mutant versions of the two proteins were then expressed in procyclic Lister 427 29-13 cells and analysed by western blotting (Figure 6). The tetracycline regulation of expression was effective and the wild type and mutant proteins were expressed at similar levels (Figure 6A). The level of expression relative to the endogenous protein varied; the expression from the *TbEIF4AI* transgenes was lower than expression from the endogenous gene whereas expression from the *TbEIF4AI* transgene that had any effect on growth was the mutant form of *TbEIF4AI* (Figure 6C), all others grew at the same rates as the control cultures without tetracycline (data not shown). At 18–51 h after the addition of tetracycline, the expression levels of the *TbEIF4AI* transgenes were readily detectable and the mutant, but not the wild type, produced a slowing of growth. As the expression of the transgene reduced, the culture returned to the same rate of growth as the no tetracycline control. In contrast, the significantly overexpressed *TbEIF4AI* mutant transgene had no effect on growth of the culture. Interestingly, the expression of the *TbEIF4AI* transgenes, but not the *TbEIF4AI* transgenes, resulted in increased levels of the endogenous protein. Overall, the results are compatible with the RNA helicase activity of *TbEIF4AI* being strictly required for growth. As for *TbEIF4AI*, the lack of a slow growth phenotype when the dominant negative mutant is expressed suggests that either its RNA helicase activity is not required for the protein function or it is not affected by the DEAD→DQAD mutation. Either option strongly indicates that *TbEIF4AI* is not active in translation.

Mapping of isoform specific amino acids

The results described above for *TbEIF4AI* are reminiscent of what is known of mammalian eIF4AI (Discussion). Human eIF4AIII, identified previously as a negative regulator of translation (22), has been shown to be a component of the EJC, with roles in mRNA export, cytoplasmic RNA localization and NMD (24–27). Pairwise sequence comparisons between *TbEIF4AI* and III (or their orthologues in *T. cruzi* and *L. major*) and the functionally divergent human eIF4AI/eIF4AIII do not show a clear match between either of the parasite homologues with the two human sequences. Indeed, the overall identity between human eIF4AI/eIF4AIII (66%) is greater than that between either protein and the two trypanosomatid eIF4As (~55–60%). The kinetoplastid eIF4AI and eIF4AIII sequences were then aligned with putative eIF4AI and eIF4AIII homologues from the major lines of eukaryotic evolution (Figure 7). The homologues from *Arabidopsis thaliana* and *Schizosaccharomyces pombe* were identified using BLAST searches of non-redundant sequence databases using the human eIF4AI or eIF4AIII sequences as queries. The alignment in Figure 7 does not show any continuous sequence of amino acids that distinguish between all putative eIF4AI or eIF4AIII homologues. However, at various positions, interspersed within the sequences common to both sets of proteins, individual amino acids can be identified which are conserved and unique either to the eIF4AI or eIF4AIII proteins. Table 1 lists 13 positions where a clear difference could be found between the two sets of sequences. Several, but not all, of these amino acid substitutions are also shared by an eIF4AIII-related protein from *Saccharomyces cerevisiae*, Fall1p, a nucleolar protein shown to be required for 40S ribosomal subunit formation (45). Fall1p, however, does not seem to be involved in EJC
formation since a search in *S. cerevisiae* for similar EJC constituents, conserved in other fungi and in plants, such as Magoh or Y14, did not produce any clear homologues.

The various amino acid substitutions listed in Table 1 (*T*beIF4AI numbering), indicated by a star in the alignment in Figure 7, discriminate between all putative eIF4AI and eIF4AIII homologues compared, including the two trypanosomatid proteins. These substitutions are located in the two globular domains present in eIF4A and related DEAD-box helicases (11, 14). Both the N- and the C-terminal domains have been shown to participate in the binding to RNA and ATP required for the helicase/ATPase activities, but few roles have been postulated for them regarding specific protein functions. The alignment results clearly show that candidate eIF4AIII homologues are present throughout the various eukaryotic lineages, although it has only been functionally characterized in metazoans. The unique substitutions are also indicative of amino acids involved in specific aspects of eIF4Al/III function in general (see below).

**Molecular modelling of *T*beIF4AI and *T*beIF4AIII**

To understand the functional implications of the observed amino acid substitutions to eIF4A function, not only in

---

**Figure 6.** Expression of myc-tagged dominant negative mutants of *T*beIF4AI and III in procyclic cells. (A) Western blot analysis of the expression of the various *T*beIF4AI and III/myc fusions in transfected cells in the absence or after exposure to tetracycline for 18 h. In each case the expression was detected using antibodies specific to each of the eIF4A homologues. The *T*beIF4AI western blot was simultaneously probed with anti-BiP as a loading control. (B) Time course expression of the different versions of *T*beIF4A-myc after tetracycline addition to the culture. The *T*beIF4AI western blot was simultaneously probed with anti-BiP as a loading control. (C) Effect of the expression of the dominant negative form of *T*beIF4AI-myc on the growth of the transfected cells in culture.
trypanosomatids but also in eukaryotes in general, we modeled the structures of both \textit{Tb}EIF4AI and \textit{Tb}EIF4AIII based on the solved structure of either the yeast \textit{S.cerevisiae} eIF4A (12) or the related DEAD-box protein Dhh1p (46). The structure of yeast eIF4A is in an open conformation with the two globular domains positioned apart and non-interacting. In contrast, Dhh1p is in a closed conformation with the two domains facing each other. Most of the conserved motifs in Dhh1p are positioned in close spatial proximity facing the cleft between the two domains. Both sets of models were validated as described (33) and found to have self-consistency in terms of sequence–structure compatibility and to be of good overall quality.

For our analysis we favored the closed conformation structure since the two domains need to interact in order to fully form the ATP- and RNA-binding sites (12,46,47). Figure 8A shows the ribbon drawing for the predicted structures of \textit{Tb}EIF4AI and \textit{Tb}EIF4AIII. Highlighted in the figure are several of the diagnostic amino acid substitutions identified in the eIF4A alignment (Figure 7). Of special interest is the V/L328W substitution in motif V, in the C-terminal domain.

Figure 7. Sequence alignment comparing \textit{Tb}EIF4AI and III with the putative eIF4AI and eIF4AIII from selected organisms. (A) Sequences were aligned as described in Figure 1 and the various DEAD-box motifs are shown as indicated previously. The predicted secondary structural elements derived from the modelling shown in Figure 8 and from Ref. (46) are indicated numbered \(a_1–a_{13}/h_1–h_4\) (alpha-helices—H) and \(b_1–b_{14}\) (beta-strands—S). Asterisk indicates amino acids which distinguish between the eIF4AI and eIF4AIII homologues. Further relevant GenBank accession numbers: human (\textit{Hs}) eIF4AI, AAX43035; human eIF4AIII (\textit{Hs}eIF4A3), P38919; \textit{S.pombe} (\textit{Sp}) eIF4A1, CAAS6772; \textit{S.pombe} eIF4A-like protein (\textit{Sp}4Alike), CAAS2238; \textit{A.thaliana} (\textit{At}) eIF4A1, NP_177417; \textit{A.thaliana} eIF4A-like protein (\textit{At}4Alike), NP_188610.
Motif V lies in a loop positioned in the interface between the two domains and, in Dhh1p, several amino acids in this motif are seen to make direct interactions with specific amino acids in motifs I and Q, positioned in the N-terminal domain (46). In the models shown here both the V and W residues in TbEIF4AI and III, respectively, are protruding from the main polypeptide backbone in the direction of a cleft in the proteins’ N-terminal domain. To investigate the likelihood of either amino acid interacting with neighboring chains, atoms in these chains were first identified which are positioned within a radius of 4 Å from the two residues. These are the only ones capable of forming non-covalent interactions to atoms in either amino acid and the full set of potential interactions are shown in Figure 8B as dotted lines. The substantially larger W residue in TbEIF4AIII is capable of making a number of interactions with neighboring amino acids in both the N- and C-terminal domains, as well as with the polypeptide backbone. In contrast, the V residue in TbEIF4AI is very limited in the number of interactions it can establish. It is possible then that the presence of the W residue in TbEIF4AIII, and other EIF4AIII homologues, can enhance the interaction between the helices’ two domains as compared to the V/L residue in the eIF4AI and even Dhh1p proteins.

Other potentially interesting eIF4AIII-like substitutions map in helices 5 (Q/R139E/G, Q/A146D/E) and 10 (L256F, E264D), on the N- and C-terminal domains, respectively. These helices are largely exposed to the solvent on the external side of the proteins (Table 1) and thus the amino acids involved could mediate eIF4A binding to functional partners. Indeed, recent evidence strongly supports such a hypothesis. First, the binding surface for eIF4GII has been mapped to the C-terminal domain of eIF4AI (47). A double mutation in human eIF4AI which prevents binding to eIF4GII maps to helix 10 and targets the same glutamate residue (E264) found to be unique to the eIF4AI sequences. Likewise, the two substitutions in helix 5 are included within a proposed eIF4AIII specific motif (motif C) which has just been found to constitute part of the binding site for the eIFG component MNL51 (30). Moreover, a further unique substitution identified in Figure 7, which lies in an exposed loop near the C-terminal end of the eIF4A proteins (H/N388A/S—also shown in Figure 8), lies within another proposed motif (motif H) found to be required for eIF4AI binding to spliced mRNA and to rescue NMD in eIF4AIII depleted cells (30). In summary, we have identified several individual amino acids conserved in either eIF4A1 or eIF4AIII sequences which may play significant roles in these proteins’ functions not only in trypanosomatids but also in eukaryotes in general.

**DISCUSSION**

The results presented here provide strong support that only one of the two eIF4A homologues identified in trypanosomatids is involved in the initiation of translation. The abundance of the TbEIF4AI protein, its constitutive expression during the parasite life cycle as well as its cytoplasmic localization, the effect of RNAi depletion and the dominant negative phenotype of the DEAD→DQAD mutation are all compatible with what is expected of this protein. In contrast, TbEIF4AIII does not seem to play an obvious role in protein synthesis. The nuclear localization of TbEIF4AIII, its low abundance, longer response to the RNAi induced phenotype and lack of inhibition by the dominant negative mutant all indicate an essential role in RNA metabolism in the nucleus unrelated to eIF4A function in translation. These results are also compatible with what is known of the *L. major* orthologues; *Lm*EIF4AI binds strongly to at least two eIF4G homologues whereas *Lm*EIF4AIII has a reduced binding activity [(33) and C. R. S. Reis, unpublished data]. Mammalian eIF4AIII localizes mainly to the nucleus (23), is present in levels ~10-fold lower than eIF4AI in HeLa cells and does not function in protein synthesis (22). An unusual feature of human eIF4AIII is that the DEAD→DQAD DQAD mutation has no effect on its activity in EJC formation and NMD (30). Thus, TbEIF4AIII behaves similarly to human eIF4AIII in several important aspects and, coupled with the sequence analysis data, our results are consistent with it being an eIF4AIII orthologue with functions possibly conserved along most major lines of eukaryotic organisms.

As part of the EJC, eIF4AIII binds directly to the core proteins Magoh, Y14 and MLN51 (25–30) and also to other proteins required for EJC function such as the TAP and Aly/REF proteins involved in nuclear mRNA export (25). Magoh homologues have been clearly identified in the three trypanosomatid genomes finished to date, *T. brucei* (GenBank, AAZ12053), *T. cruzi* (EAN97132) and *L. major* (CAJ06870) and possible TAP homologues can also be found. We have also tried to identify candidate Y14, MLN51 or Aly/REF

### Table 1. Summary of the amino acid substitutions identified between the putative eIF4AI/eIF4AIII homologues from the main lineages of eukaryotic evolution

<table>
<thead>
<tr>
<th>Position in eIF4AI</th>
<th>Substitution: 4AI &gt; 4AIII</th>
<th>Secondary structure</th>
<th>Domain</th>
<th>Overall position in predicted tertiary structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>93</td>
<td>K/NV &gt; R</td>
<td>Loop</td>
<td>α5</td>
<td>N-terminal Near α5/exposed</td>
</tr>
<tr>
<td>139</td>
<td>Q/R &gt; EG</td>
<td>α5</td>
<td>N-terminal Exposed</td>
<td></td>
</tr>
<tr>
<td>146</td>
<td>Q/A &gt; E/D</td>
<td>β4</td>
<td>N-terminal Buried</td>
<td></td>
</tr>
<tr>
<td>153</td>
<td>V &gt; S/A*</td>
<td></td>
<td>N-terminal Partially exposed</td>
<td></td>
</tr>
<tr>
<td>197</td>
<td>F &gt; Y</td>
<td>α8</td>
<td>N-terminal Interface/nest to Motif III/partially buried</td>
<td></td>
</tr>
<tr>
<td>213</td>
<td>M &gt; L</td>
<td>Loop</td>
<td>N-terminal Buried</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>L/V/I &gt; M</td>
<td>α9</td>
<td>C-terminal Exposed</td>
<td></td>
</tr>
<tr>
<td>226</td>
<td>R/N &gt; T</td>
<td>Loop</td>
<td>C-terminal Partially buried</td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>L &gt; F*</td>
<td>α10</td>
<td>C-terminal Exposed</td>
<td></td>
</tr>
<tr>
<td>264</td>
<td>E &gt; D*</td>
<td>α10</td>
<td>C-terminal Interface/Motif V/Buried</td>
<td></td>
</tr>
<tr>
<td>303</td>
<td>S/D &gt; P*</td>
<td>Loop</td>
<td>C-terminal Exposed</td>
<td></td>
</tr>
<tr>
<td>328</td>
<td>V/L &gt; W*</td>
<td>Loop</td>
<td>C-terminal Exposed</td>
<td></td>
</tr>
<tr>
<td>388</td>
<td>H/N &gt; A/S*</td>
<td>Loop</td>
<td>C-terminal Exposed</td>
<td></td>
</tr>
</tbody>
</table>

*These substitutions but not the others are present in the *S. cerevisiae* nuclear protein Fd1p [may be related to the eIF4AIII proteins (45)].
homologues but so far without success. However, both Y14 and Aly/REF are small RNA-binding proteins with single RRM s, a category which includes many proteins with unassigned functions in those three genomes (48). It may be possible that, due to the degree of evolutionary distance between trypanosomatids and animals, homologues to these two proteins cannot be clearly identified by sequence analysis alone. As for MLN51 it is poorly conserved outside the metazoans so it is unlikely also for homologues to be identified in trypanosomatids only by sequence analysis. Nevertheless the strong conservation of the Magoh sequences between the human and parasite homologues (over 50% identity) is an indication that the EJC may be present throughout the major groups of eukaryotes and that elf4AIII-like proteins may be active within this complex.

In a very recent study eight elf4AIII specific motifs (named A to H) were identified in an alignment comparing various elf4AIII homologues with the human elf4AI and II proteins. Selected amino acids in some of these motifs, as well as in the canonical elf4A motifs I, Ia and VI, were then mutated in recombinant or in vivo overexpressed elf4AIII to investigate their requirement for elf4AIII function (30). In the alignment provided here, which includes both T. brucei elf4A homologues, as well as elf4AII sequences from divergent organisms, no continuous set of amino acids were found to be typical of either elf4AI or elf4AIII proteins. However, unique amino acid substitutions were identified which distinguish elf4AIII-like proteins from elf4AI homologues in all sequences investigated so far. Several of these substitutions not only coincide with some of the proposed elf4AIII specific motifs (motifs C, E and H), but also are included in two of those motifs found to be involved in specific elf4AIII functions such as binding to the EJC partner MLN51 (motif C) and requirements for binding to spliced mRNA and for NMD (motif H) (30). However some of the unique elf4AI/elf4AIII substitutions identified here do not coincide with the remaining proposed motifs. These might be involved in mediating other aspects of elf4A function and should be considered as targets for further investigation.

Very few protein coding genes in trypanosomatids contain a cis-intron (49,50). However, every cytoplasmic mRNA is trans-spliced to form the mature 5′ end of the mRNA and this splice site is possibly the location of EJC binding. The function of the EJC in these organisms remains obscure specially considering that the splice site is always to the 5′ side of the ORF. The EJC-mediated mechanism of NMD seems to be restricted to mammalian cells [reviewed in (51,52)] and indeed there is strong evidence that NMD does not occur in trypanosomatids (38). In mammals, both the EJC and the nuclear cap-binding complex (CBC, composed of two subunits CBP20 and CBP80)

**Figure 8.** Molecular modeling of TbElf4AI and III highlighting the position of amino acids unique to the elf4AI or elf4AIII homologues. Diagrams were created with the program PyMol (http://www.pymol.org). (A) Ribbon diagrams of the overall structure of both TbElf4AI and III viewed as in (46) (upper panel) or rotated 180° about its long axis (lower panel). The structure is in a closed conformation where the two, N- and C-terminal, domains are facing each other. The arrows indicate the position of the L328W substitution which lies in the loop containing Motif V and is positioned in the interface between the two domains. The dotted circles delimit the two helices discussed in the text, α5 and α10. The H/N388A/S and M213L substitutions are also indicated (their numbering differ however from the elf4AI/elf4AIII sequences—for instance, H388 in TbElf4AI is equivalent to A387 in T9Elf4AIII and so on). (B) Balls and sticks representation showing the neighbourhood of the L328W substitution in both TbElf4AI and III. The dotted lines indicate the atoms in the neighbouring amino acid chains which are positioned within a radius of 4 Å from the atoms in either the L or W residues. In both (A and B), the relevant amino acids are listed.
bind to precursor mRNAs in the nucleus, prior to or during the splicing event, and remain bound to the mRNAs until they are transported to the cytoplasm and/or translated for the first time (53,54). In T. brucei a novel CBC has been described which consists of a CBP20 subunit (also present in yeast and humans) plus four other polypeptides, one of which is importin-α (known to associate with CBC in other eukaryotes) and three novel proteins only present in trypanosomatids. The parasite CBC has been implicated in the early steps of mRNA maturation, prior to the trans-splicing event whereas the polycistronic precursor mRNA is cleaved into mature monocistronic units (55). At this stage it still remains to be determined whether TbEF4AIII and other components of the putative EJC are also necessary for mRNA processing, export from the nucleus or even translation in trypanosomatids.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We would like to thank M. Moore for sharing unpublished data and J. Bangs for the gift of the BiP antibody. O. Thiemann and D. Souza allowed us access to the computing facilities necessary for the modelling work. The work in Cambridge was funded by the Wellcome Trust and by a Nuffield Foundation summer studentship to N. Marinsek. The Brazilian scientists and J. Bangs for the gift of the BiP antibody. O. Thiemann and

REFERENCES


