Molecular & Biochemical Parasitology 176 (2011) 25–36

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties

Eden R. Freire[a], Rafael Dhalia[a], Danielle M.N. Moura[a], Tamara D. da Costa Lima[a], Rodrigo P. Lima[a], Christian R.S. Reis[a], Katie Hughes[b], Regina C.B.Q. Figueiredo[a], Nancy Standart[b], Mark Carrington[b], Osvaldo P. de Melo Neto[a,*]

[a] Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Av. Moraes Rego s/n, Campus UFPE, Recife, PE 50670-420, Brazil
[b] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

ARTICLE INFO

Article history:
Received 16 August 2010
Received in revised form 15 October 2010
Accepted 16 November 2010
Available online 24 November 2010

Keywords:
Translation
eIF4E
eIF4G
Trypanosomatid
Leishmania

ABSTRACT

Translation initiation in eukaryotes requires eIF4E, the cap binding protein, which mediates its function through an interaction with the scaffolding protein eIF4G, as part of the eIF4F complex. In trypanosomatids, four eIF4E homologues have been described but the specific function of each is not well characterized. Here, we report a study of these proteins in *Trypanosoma brucei* (*Tb*EIF4E1 through 4). At the sequence level, they can be assigned to two groups: *Tb*EIF4E1 and 2, similar in size to metazoan eIF4E1; and *Tb*EIF4E3 and 4, with long N-terminal extensions. All are constitutively expressed, but whilst *Tb*EIF4E1 and 2 localize to both the nucleus and cytoplasm, *Tb*EIF4E3 and 4 are strictly cytoplasmic and are also more abundant. After knockdown through RNAi, *Tb*EIF4E3 was the only homologue confirmed to be essential for viability of the insect procyclic form. In contrast, *Tb*EIF4E1, 3 and 4 were all essential for the mammalian bloodstream form. Simultaneous RNAi knockdown of *Tb*EIF4E1 and 2 caused cessation of growth and death in procyclics, but with a delayed impact on translation, whilst knockdown of *Tb*EIF4E3 alone or a combined *Tb*EIF4E1 and 4 knockdown led to substantial translation inhibition which preceded cellular death by several days, at least. Only *Tb*EIF4E3 and 4 were found to interact with *T. brucei* eIF4G homologues; *Tb*EIF4E3 bound both *Tb*EIF4G3 and 4 whilst *Tb*EIF4E4 bound only to *Tb*EIF4G3. These results are consistent with *Tb*EIF4E3 and 4 having distinct but relevant roles in initiation of protein synthesis.

1. Introduction

The trypanosomatids, including the various *Leishmania* and *Trypanosoma* species, are pathogenic protozoans well known not only for the diseases they cause, but also for their complex biology and the unusual molecular mechanisms required for their gene expression. Transcription of protein coding genes occurs polycistronically [1,2] and processing to monocistronic mRNAs occurs through coupled *trans*-splicing and polyadenylation (reviewed in [3]). The result is that mRNAs have a common 39nt long spliced-leader (or mini-exon) sequence at the distal end of the 5′ UTR, which is identical for all mRNAs of a given species. At the 5′ end of the spliced-leader sequence lies the 7-methyl-GTP cap nucleotide (m^7G), followed by four methylated nucleotides in the spliced-leader sequence, the cap4 structure [4]. Regulation of gene expression in trypanosomatids is accomplished mainly through post-transcriptional mechanisms such as control of mRNA stability and possibly translation (for reviews see [5–7]). However, in contrast to the considerable characterization of their mechanisms of mRNA synthesis and processing, relatively little is understood regarding protein synthesis and how it can be regulated.

Translation initiation is the most complex stage of protein synthesis and the one which can vary more significantly between different taxonomic groups [8]. A major player in this process in eukaryotes, and also a major target for translation control, is eIF4E (eukaryotic initiation factor 4E), the m^7G cap binding protein. eIF4E is a small polypeptide (24–25 kDa in mammals) which, apart from protein synthesis, has been implicated in a number of processes involved in mRNA metabolism such as transport and control of its stability (reviewed in [9–12]). In translation it is part of the heterotrimeric complex eIF4F, which also includes the RNA helicase eIF4A and the large scaffolding protein eIF4G [13]. eIF4F facilitates the recruitment of the small ribosomal subunit to the mRNA, which is accomplished with the help of another translation initiation complex, eIF3. eIF4E mediates the binding of the complex to the mRNA and, as part of eIF4F, can bind simultaneously to both the cap and

* Corresponding author at: Departamento de Microbiologia, Centro de Pesquisas Aggeu Magalhães/Fiocruz, Av. Moraes Rego s/n, Campus UFPE, Recife, PE 50670-420, Brazil. Tel.: +55 81 2101 2636; fax: +55 81 2101 2647.
E-mail address: opmn@cpqam.fiocruz.br (O.P. de Melo Neto).

0166-6851/
doi:10.1016/j.molbiopara.2010.11.011

eIF4G, which can then mediate interactions with other translation factors [14–16]. Despite its well described function in translation in the cytoplasm, substantial levels of eIF4E localize to the nucleus (reviewed in [17]), where it promotes the nuclear export of selected mRNAs [10,18,19].

The structure of eIF4E bound to 7-methyl-GDP [20,21] is shaped like a cupped hand, with the cap analogue located in a narrow cap-binding slot on the concave side of the protein. eIF4E is characterized by eight tryptophan residues located at conserved positions along the protein. Cap recognition is mediated by, among other interactions, base sandwich-stacking between W56 and W102 (mammalian eIF4E-1 numbering) whilst W73 is involved in the interaction with eIF4G [9,13,20]. The activity of eIF4E can be regulated in a number of ways such as through the action of eIF4E interacting proteins, eIF4E-BPs, or directly through phosphorylation (reviewed in [22–25]).

In the yeast *Saccharomyces cerevisiae* only one eIF4E homologue, essential for viability, is found [26]. More usually, multiple forms of this protein are present in any specific organism (reviewed by [27]). In mammals, three different classes of eIF4E have been described: eIF4E-1, the prototypical cap-binding protein; eIF4E-2 (also known as 4EHP), which binds the cap less well, does not interact with eIF4G and has a regulatory role [28,29]; and eIF4E-3, which binds both cap and eIF4G but has a restricted, tissue specific, expression pattern [30]. An extensive analysis of hundreds of eIF4E-related sequences has indicated that different events of duplication of the eIF4E gene may have occurred in distinct eukaryotic groups after divergence, resulting in new forms of the protein with novel or modified functions [31].

In trypanosomatids, the initial description of a single putative eIF4E homologue [32] was followed by the identification of four homologues in genome sequences from *L. major*, all of which are also conserved in *Trypanosoma* species [33]. The four *Leishmania* eIF4E homologues (*Lm*EIF4E1 through 4, LeishIF4E1 through 4 or simply EIF4E1 through 4) were shown to diverge in binding affinity for different synthetic cap analogues, in expression levels and in polysomal association in sucrose sedimentation gradients. None could rescue the growth of a yeast strain lacking a functional eIF4E and none could confidently be implicated as having a predominant role in translation or its control [33,34].

Here we describe work focused on the four *Trypanosoma brucei* proteins (here named *Tb*EIF4E1 through 4). First, sequence analysis identified features which distinguish *Tb*EIF4E1 and 2 from *Tb*EIF4E3 and 4. For all four homologues, the cap-binding affinities, subcellular localization, and the expression in two stages of the *T. brucei* life cycle were characterized. The effect of depletion, through RNAi, on cellular growth and translation was determined and the interaction between putative eIF4E/eIF4G homologues was also analyzed. The results presented allow the separation of the four proteins into two distinct groups. The first comprises *Tb*EIF4E1 and 2, which localize both to the nucleus and the cytoplasm, do not seem to be directly involved in translation but perform functions essential for cellular viability. The second group is formed by the *Tb*EIF4E3 and 4 homologues, more abundant, strictly cytoplasmic proteins which seem to be required for translation and take part in the formation of distinct eIF4F-like complexes.

2. Materials and methods

2.1. Sequence analysis

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 and Clustal W alignments were done as previously described [33]. Relevant trypanosomatid GeneDB accessions: *Tb*EIF4E1 – Tb11.18.0004; *Tb*EIF4E2 – Tb927.10.16070; *Tb*EIF4E3 – Tb11.01.3630; *Tb*EIF4E4 – Tb927.6.1870; *Lm* (*L. major*) EIF4E1 – LmjF27.1620; *Lm*EIF4E2 – LmjF19.1500; *Lm*EIF4E3 – LmjF28.2500; *Lm*EIF4E4 – LmjF30.0450 (the annotated sequence for *Lm*EIF4E4 is missing 139 residues from what we believe is the N-terminus of the protein – they are encoded upstream of the annotated open reading frames and are homologous to the equivalent segment in the *T. brucei* orthologue).

2.2. PCR and cloning methods

All *T. brucei* eIF4E and eIF4G coding sequences were amplified from Lister 427 total genomic DNA and first cloned into the pGEM-T Easy vector (Promega). All amplified fragments were first sequenced, and the resulting sequences compared with those from the *T. brucei* genome sequencing project, prior to their use in the subcloning reactions. In order to express N-terminal His-tagged fusion proteins the *Tb*EIF4E1 through 4 full length sequences were subcloned into the modified pET15b vector as described previously for the *T. brucei* eIF4A homologues [35]. Alternatively, to generate the eYFP and HA-tagged fusions and for the RNAi experiments, the same fragments were cloned respectively into the transfection vectors p2216 [35,36], p2477 [36] or p2T7-177 [37]. To generate the plasmids used in the transcription and translation reactions they were also subcloned into the pGEM3zf+ plasmid (Promega). To express the full length proteins as N-terminal GST tagged fusions, the *Tb*EIF4G3 and 4 sequences were subcloned into the pGEX4T3 plasmid (GE Healthcare). His-tagged *Tb*EIF4G3 and 4 were also expressed, with a C-terminal tag, as fragments consisting of the proteins' N-terminus plus the central MIF4G/HEAT domain (TbEIF4G3$_{1–228}$ and TbEIF4G4$_{1–319}$ – numbers indicating the residues from the wild type protein remaining in the recombinant product) using the pET21a vector (Novagen). Supplementary Table I lists all oligonucleotides used in the amplifications reactions as well as the strategies and restrictions enzymes used in the subsequent subcloning events cited above. For the double *Tb*EIF4E1/2 construct, for the RNAi experiments, the *Tb*EIF4E2 gene was reamplified flanked by sites for *Bgl* II and *Bam*H I and cloned into the dephosphorylated *Bam*H I site of the p2T7-177-*Tb*EIF4E1 construct. For the double *Tb*EIF4E1/4 construct, first the p2T7-177-*Tb*EIF4E1 was digested with *Bam*H I, which cuts internally to the *Tb*EIFE1 gene and at its 3′ end at the site introduced after the original PCR reaction, releasing a ~360 bp fragment consisting of the second half of the target gene. This fragment was then cloned into the linearized *Bam*H I site of the p2T7-177-*Tb*EIF4E4 construct in a similar procedure as described for *Tb*EIF4E1/2. Recombinant protein expression was performed according to standard procedures using the pET15, pET21 or pGEX4T3 derived constructs after transformation into *Escherichia coli*.

2.3. Cap binding assay

The four [^{35}S]-methionine labeled *Tb*EIF4Es were synthesized after *in vitro* transcription of their respective genes cloned in the pGEM3zf+ vector, following linearization with *Xba* I (*Tb*EIF4E1 and 3) or *Bam*H I (*Tb*EIF4E2 and 4), with T7 RNA polymerase followed by translation with the nuclease treated Rabbit Reticulocyte Lysate System (Promega). For the *Lm*EIF4E4 homologue, its gene was amplified from *L. major* (MHOM/IL/81/Friedlin) genomic DNA flanked by sites for *Afl* III and *Not* I (5′ primer – CT GAC ATG TCT ACC CCT CTC GAT GTG; 3′ primer – TA TGC GGC CGC GTA GCG ACG ACG GTT CTT TTT C) and cloned into the *Nco* I/*Not* I sites of pET21D (Novagen). The resulting plasmid was sequenced and linearized with *Not* I prior to transcription and translation, as described above. Assays were performed essentially as described previously [33],