

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E
BIOLOGIA MOLECULAR
TESE DE DOUTORADO**

**IDENTIFICAÇÃO E MAPEAMENTO DE DOMÍNIOS DE
LIGAÇÃO DE HOMÓLOGOS DO FATOR eIF4G DE
INICIAÇÃO DA TRADUÇÃO DE *Leishmania major*.**

CHRISTIAN ROBSON DE SOUZA REIS

VIRTUS IMPAVIDA

**RECIFE
2009**

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR
TESE DE DOUTORADO**

**IDENTIFICAÇÃO E MAPEAMENTO DE DOMÍNIOS DE
LIGAÇÃO DE HOMÓLOGOS DO FATOR eIF4G DE
INICIAÇÃO DA TRADUÇÃO DE *Leishmania major*.**

CHRISTIAN ROBSON DE SOUZA REIS

Tese apresentada ao Programa
de Pós-graduação em Genética e
Biologia Molecular da
Universidade Federal de
Pernambuco como requisito para
obtenção do grau de doutor em
Genética pela UFPE

Orientador: Dr. Osvaldo Pompílio de Melo Neto

**RECIFE
2009**

Reis, Christian Robson de Souza

Identificação e mapeamento de domínios de ligação de homólogos do fator eIF4G de iniciação da tradução de Leishmania major/ Christian Robson de Souza Reis. – Recife: O Autor, 2009.

198 folhas : il., fig., tab.

Tese (doutorado) – Universidade Federal de Pernambuco. CCB. Departamento de Genética e Biologia Molecular, 2009.

Inclui bibliografia e anexos.

1. Leishmania major – tradução – genética. 2. Tripanossomatídeos. 3. EIF4F. I Título.

572.645

CDD (22.ed.)

UFPE/ CCB – 2009- 198

UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM
GENÉTICA E BIOLOGIA MOLECULAR

Parecer da comissão examinadora da dissertação/tese de:
CHRISTIAN ROBSON DE SOUZA REIS

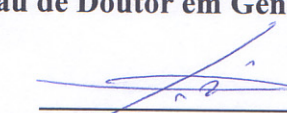
intitulada:
IDENTIFICAÇÃO E MAPEAMENTO DE DOMÍNIOS DE LIGAÇÃO DE
HOMÓLOGOS DO FATOR eIF4G DE INICIAÇÃO DA TRADUÇÃO DE
Leishmania major.

A comissão examinadora considera o presente trabalho

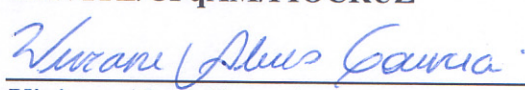
APROVADO

Portanto, cumpridas todas as exigências regimentais, Christian Robson de Souza Reis
faz jus ao grau de Doutor em Genética e Biologia Molecular pela UFPE. Recife,

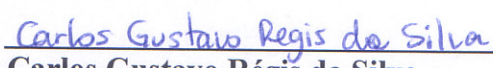
30/07/2009


1º Examinador: Rafael Dhalia

LAVITE/CPqAM/FIOCRUZ


2º Examinador: Viviane Alves Gouveia

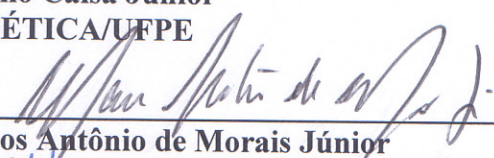
MICROBIOLOGIA/CPqAM/FIOCRUZ


3º Examinador: Carlos Gustavo Régis da Silva

IMUNOLOGIA/CPqAM/FIOCRUZ


4º Examinador: Tercilio Calsa Junior

GENÉTICA/UFPE

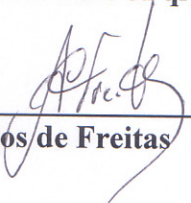

5º Examinador: Marcos Antônio de Moraes Júnior

GENÉTICA/UFPE


Orientador:

Osvaldo Pompílio de Melo Neto

MICROBIOLOGIA/CPqAM/FIOCRUZ


Coordenador: Antônio Carlos de Freitas

AGRADECIMENTOS

A Dr. Osvaldo Pompílio de Melo Neto, por ter ajudado de forma fundamental a consolidar minha formação científica, e hoje em dia essa relação ter como resultado prático minha participação como colega de Departamento e de Instituição.

A Dr. Rafael Dhália pela amizade, pelo aprendizado, pela criteriosa revisão desta tese.

Em especial aos amigos Franklin Magalhães, Éden Freire, Tamara Lima, Danielle Moura, Mariana Marques, Francisco Cariri, colegas mais antigos, cuja convivência frutífera no dia a dia permitiu, e continuará permitindo meu progresso científico, pessoal e intelectual.

Aos colegas atuais e egressos do Departamento de Microbiologia: Mariana Palma, Ana Paula, Amaranta Malvezzi, Rodrigo Pontes, Marília Barbosa, Bruna Caiado, Wellington Marques, Adelino Lima, Kellyanne Carvalho, Eduardo Nunes, Diego Tavares, Janaína Nascimento e Mariana Nunes.

Ao corpo técnico: Isaac Martins, Cláudio Eduardo, Silvana Santos, Yara Nakasawa, Bruna Lima.

As instituições que possibilitaram a realização deste trabalho: CPqAM, CNPq, CAPES, FACEPE.

Aos meus familiares que me ajudaram muito a alcançar meus objetivos e a família da minha noiva Verônica Padilha.

DEDICATÓRIA

*A meus Pais (Seu Reis e Dona Dalva)
e a minha noiva Verônica*

LISTA DE FIGURAS

Figura 1. Ciclo de transmissão do <i>Trypanosoma brucei</i>	20
Figura 2. Ciclo de transmissão do <i>Trypanosoma cruzi</i>	22
Figura 3. Ciclo de transmissão da <i>Leishmania sp</i>	24
Figura 4. Transcrição policistrônica nos tripanossomatídeos	28
Figura 5. Representação esquemática do <i>trans-splicing</i>	31
Figura 6. Processamento de mRNAs nos tripanossomatídeos	33
Figura 7. Representação esquemática da iniciação da tradução de mRNAs típicos em eucariotos	38
Figura 8. Estrutura do eIF4E de camundongo	40
Figura 9. Estrutura tridimensional do eIF4A de levedura	43
Figura 10. Esquema do eIF4GI humano e seus principais sítios de ligação a outras proteínas.	45
Figura 11. Motivos em comum entre eIF4GI, eIF4GII e proteínas da família eIF4G	48
Figura 12. Interações do eIF4G, eIF4H e RNA com o eIF4A	51
Figura 13. Esquema do eIF4GI de levedura ressaltando o domínio de ligação ao eIF4E, também presente nas 4E-BPs	53
Figura 14. Esquema dos homólogos ao eIF4G de <i>L. major</i> e seus principais domínios	61
Figura 15. Esquema da técnica de <i>Pull-Down</i> usando a proteína isca (I) expressa em bactéria e a proteína presa (P) expressa em sistema de tradução de mamíferos (marcada com ³⁵ S) <i>in vitro</i>	76
Figura 16. Esquema das proteínas <i>LmEIF4Gs</i> e dos vários fragmentos utilizados nesse trabalho destacando possíveis domínios de ligação a outras proteínas	81
Figura 17. Alinhamento no Clustal W comparando as seqüências polipeptídicas do MIF4G dos cinco homólogos do eIF4G de <i>L. major</i> com o domínio equivalente no homem, levedura, planta e na proteína humana PAIP-1	82
Figura 18. Análise de interação entre os homólogos ao eIF4A e os homólogos ao eIF4G de <i>L. major</i>	84
Figura 19. Análise de interação entre o eIF4E de <i>Xenopus laevis</i> e eIF4G humano	85
Figura 20. Análise de interação entre os homólogos ao eIF4E e os homólogos ao eIF4G de <i>L. major</i>	86
Figura 21. Análise de interação entre o homólogo <i>LmEIF4G4</i> com o homólogo <i>LmEIF4E3</i> de <i>L. major</i>	87
Figura 22. Análise de interação entre os homólogos do eIF4A e eIF4E com diferentes domínios dos homólogos EIF4G3-4 de <i>L. major</i>	89

Figura 23. Alinhamento múltiplo por Clustal W comparando as seqüências dos homólogos <i>LmEIF4G3</i> e <i>LmEIF4G4</i> com seus ortólogos de <i>T. brucei</i> e <i>T. cruzi</i>	92
Figura 24. Análise da interação entre a proteína mutante <i>LmEIF4G3</i> _{LNK/AAA} com a proteína <i>LmEIF4AI</i> selvagem	95
Figura 25. Análise da interação entre mutantes da proteína <i>LmEIF4G4</i> com as proteínas <i>LmEIF4AI</i> e <i>LmEIF4E3</i>	96
Figura 26. Análise da interação entre a proteína <i>LmEIF4AI</i> e as diferentes proteínas <i>LmEIF4G3</i>	98
Figura 27. Análise de interação entre as proteínas <i>LmEIF4G3-4</i> e as proteínas <i>LmEIF4E1-4</i> (ensaio reverso)	101
Figura 28. Análise da interação entre a proteína <i>LmEIF4E4</i> e as diferentes proteínas <i>LmEIF4G3</i>	102
Figura 29. Análise da interação entre a proteína <i>LmEIF4E3</i> e as diferentes proteínas <i>LmEIF4G4</i>	103
Figura 30. Análise de interação entre a PABP e o eIF4G humanos	106
Figura 31. Análise de interação entre os homólogos PABP1-2 com homólogos do eIF4G de <i>L. major</i>	107
Figura 32. Análise da interação entre os homólogos da PABP e o homólogo <i>LmEIF4G3</i> de <i>L. major</i> na presença de caudas poli-C e poli-A sintéticos	108
Figura 33. Análise da interação entre as proteínas <i>LmEIF4G3-4</i> com <i>LmPABP1</i>	110
Figura 34. Análise da interação entre a proteína <i>LmPABP1</i> e as diferentes proteínas <i>LmEIF4G3</i>	111
Figura 35. Reconstituição, “ <i>in vitro</i> ”, do complexo eIF4F de <i>Leishmania major</i>	113
Figura 36. Análise da interação entre as proteínas EIF4G3-4 e as proteínas EIF4E3-4 de tripanossomatídeos	116
Figura 37. Candidatos a motivos de ligação ao eIF4E encontrados nos homólogos ao EIF4G3-4 nos tripanossomatídeos	122
Figura 38. Os dois possíveis complexos homólogos ao eIF4F nos tripanossomatídeos, com base nos ensaios de <i>pull-down</i> em <i>L. major</i>	125
Figura 39. Quantificação dos níveis intracelulares das proteínas <i>TbEIF4AI-III</i> nas formas procíclicas e sanguíneas de <i>T. brucei</i>	128
Figura 40. Interferência do RNA <i>TbEIF4AIII</i>	130
Figura 41. Análise de interação entre os homólogos ao eIF4A e o homólogo Mago de <i>T. brucei</i>	131

LISTA DE TABELAS

Tabela I. Organização genômica dos três principais tripanossomatídeos	26
Tabela II. Sumário das sequências dos homólogos eIF4G de <i>L. major</i> descrevendo as identidades (similaridades) entre as sequências de <i>L.major</i> / humano e <i>L.major</i> / <i>T.brucei</i> .	62
Tabela III. Oligonucleotídeos usados para amplificação e clonagem dos genes ou fragmentos gênicos nos vetores plasmidiais utilizados neste trabalho.	67
Tabela IV. Características dos oligonucleotídeos desenhados para reações de mutagênese sítio-dirigida seguindo recomendações do fabricante.	71
Tabela V. Descrição dos oligonucleotídeos utilizados para mutagênese sítio-dirigida dos genes <i>LmEIF4G3</i> e <i>LmEIF4G4</i> .	72

LISTA DE ABREVIATURAS E SIGLAS

AA – Aminoácidos

ATP – Adenosina Trifosfato

AREs – Regiões ricas em repetições dos nucleotídeos AU (*AU-rich elements* – AREs)

AUG – Códon de iniciação da tradução

BLAST – Ferramenta de comparação de seqüências por alinhamento local (*Basic Local Alignment Search Tool*)

BSA – Albumina sérica bovina

BSF – Forma sanguínea de *Trypanosoma brucei* (*bloodstream form*)

°C – grau Celsius

EJC – Complexo de Junção entre Éxons (*Exon Junction Complex*)

DEPC – dietilpirocarbonato

DMSO – dimetilsulfóxido

DNA – ácido dextrorribonucleico

DTT – Ditionitrito

ECL – Detecção por quimioluminescência (*Enhanced Chemiluminescence*)

EDTA – ácido etilenodiaminotetracético

eIF – Fator de iniciação da tradução eucariótico (*Eukaryotic Initiation factor*)

GST – Glutathione S-Transferase

GTP – Guanosina Trifosfato

IPTG – Isopropil β -D-1 tiogalactopiranosídeo

kDa – kiloDáton

KV – kiloVolt

LB – meio Luria Bertani

mRNA – ácido ribonucleico mensageiro

Met – Metionina

MAP– Proteínas quinases ativadas por mitógenos (*Mitogen Activated Protein Kinases*)

Mnk1 – proteína quinase MAP de integração de sinal (*MAP-kinase signal-integrating kinases*)

NMR – Ressonância Nuclear Magnética (*Magnetic Nuclear Resonance*)

ORF – Seqüência aberta de leitura (*Open Reading Frame*)

OD – Densidade Óptica (*Optic Density*)

pb – pares de base

PABC – Domínio C-terminal da PABP (*PABP C-terminal Domain*)

PABP – Proteína de ligação à cauda poli-A (*Poly-A binding protein*)

PBS – Solução salina tamponada com fosfato (*Phosphate Buffered Saline*)

P-bodies: corpos de processamento (*Processing bodies*)

PCF – Forma procíclica de *Trypanosoma brucei* (*procyclic form*)

PCR – Reação em cadeia da polimerase (*Polymerase Chain Reaction*)

RNAP – RNA polimerase

RNase – Ribonuclease

rpm – rotações por minuto

RRMs – Motivos de reconhecimento de RNA (*RNA recognition motif*)

rRNA – Ácido ribonucléico ribossômico

SDS-PAGE – Eletroforese em gel de poliacrilamida com dodecil sulfato de sódio (*Sodium Dodecyl Sulfate – Polyacrilamide Gel Electrophoresis*)

SL – Sequência *spliced leader* ou mini-éxon

VSG – Glicoproteína variante de superfície (*Variant Surface Glycoprotein*)

URE – Regiões ricas em repetições dos nucleotídeos U (*U rich instability elements*)

UTR – Região não traduzida (*Untranslated Region*)

Tag – Marca de purificação

TAP – *Tandem Affinity Purification*

TBS - Solução salina tamponada com tris (*Tris Buffered Saline*)

TE – Tampão Tris-EDTA

tRNA – Ácido ribonucléico transportador

WHO – Organização Mundial da Saúde (*World Health Organization*)

AMINOÁCIDOS							
A	Alanina	G	Glicina	M	Metionina	S	Serina
C	Cisteína	H	Histidina	N	Asparagina	T	Treonina
D	Ácido aspártico	I	Isoleucina	P	Prolina	V	Valina
E	Ácido glutâmico	K	Lisina	Q	Glutamina	W	Triptofano
F	Fenilalanina	L	Leucina	R	Arginina	Y	Tirosina

RESUMO

Os tripanossomatídeos são caracterizados por processos moleculares atípicos como o controle pós-transcricional da expressão gênica, através da degradação de mRNA e controle da tradução. Na maioria dos eucariotos, a iniciação da tradução começa com a ligação do complexo eIF4F (composto pelas subunidades eIF4A, eIF4E e eIF4G) ao cap monometilado presente na região 5' dos mRNAs, permitindo seu reconhecimento pelo ribossomo e outros eIFs participantes. A atividade do eIF4F é reforçada pela proteína de ligação a cauda poli-A (PABP), que atua através de uma interação direta com o eIF4G. Vários homólogos às subunidades do eIF4F e PABP foram previamente identificados nos tripanossomatídeos, como cinco homólogos do eIF4G (EIF4G1-5) e múltiplos homólogos das demais subunidades (EIF4E1-4, EIF4AI e III e PABP1-2) e suas funções vem sendo investigadas em *Leishmania major* e *Trypanosoma brucei*. Como parte deste estudo, nós demonstramos que tanto o *LmEIF4G3* como o *LmEIF4G4* interagem com o *LmEIF4AI*, embora mutações nos aminoácidos LNK destas proteínas abolem suas interações com o *LmEIF4AI*. As proteínas *LmEIF4G3* e *LmEIF4G4* também se ligam respectivamente aos homólogos *LmEIF4E4* e *LmEIF4E3* e diferentes aminoácidos foram associados a estas interações. A substituição dos resíduos FSL no *LmEIF4G3* abole sua interação com o *LmEIF4E4* e a mutagênese dos resíduos IL no *LmEIF4G4* abole sua ligação ao *LmEIF4E3*. Adicionalmente, nós demonstramos que as proteínas *LmEIF4G3* e *LmEIF4G4* ligam-se à *LmPABP1* e nós também sugerimos funções distintas para o *TbEIF4AI* e *TbEIF4AIII* em ambas as formas procíclica e sanguínea de *T. brucei*. Em síntese, nossos resultados indicam uma maior complexidade na iniciação da tradução nos tripanossomatídeos e provavelmente a mesma é regulada por pelo menos dois diferentes complexos eIF4F.

Palavras chave: Tripanossomatídeos, eIF4F, Iniciação da tradução.

ABSTRACT

The trypanosomatids are characterized for unusual molecular process as pos-transcriptional control of gene expression, through mRNA degradation and translation control. In the majority of the eukaryotes, translation initiation starts with the binding of the translation initiation complex eIF4F (formed by the subunits eIF4A, eIF4E and eIF4G) to the monomethylated cap present on the 5'end of the mRNAs allowing it's recognition by the ribosome and others eIFs participants. The eIF4F activity is enhanced by the poly(A) binding protein (PABP), which acts through a direct interaction with eIF4G. Several homologues to the eIF4F subunits and PABP have been previously identified in trypanosomatids, as five homologues of eIF4G (EIF4G1-5) and multiple homologues to the others subunits (EIF4E1-4, EIF4AI-III and PABP1-2) and their functions have been investigated in *Leishmania major* and *Trypanosoma brucei*. As a part of this work, we showed that both *LmEIF4G3* and *LmEIF4G4* interact with *LmEIF4AI*, however mutations in LNK residues of these proteins abolish their interactions with *LmEIF4AI*. The *LmEIF4G3* and *LmEIF4G4* also respectively bind to *LmEIF4E4* and *LmEIF4E3* homologues and different residues were associated to these interactions. Substitution of FSL residues in *LmEIF4G3* abrogates your interaction with the *LmEIF4E4* and IL residues mutagenesis in *LmEIF4G4* abolishes it's interaction to the *LmEIF4E3*. Additionally, we showed that *LmEIF4G3* and *LmEIF4G4* binding to *LmPABP1* and we also suggest distinct functions for *TbEIF4AI* and *TbEIF4AIII* in both procyclic and bloodstream forms of *T. brucei*. To sum up, our findings indicate the major complexity of translation initiation in trypanosomatids and probably it is regulated by at least two different eIF4F complexes.

Keywords: Trypanosomatids, eIF4F, translation initiation.

SUMÁRIO

Lista de Figuras

Lista de Tabelas

Resumo

Abstract

1 - INTRODUÇÃO	16
2 - REVISÃO DA LITERATURA	18
2.1-Tripanossomatídeos	18
2.1.1 - Gênero <i>Trypanosoma</i> e as Tripanossomíases	19
2.1.2 - Gênero <i>Leishmania</i> e as Leishmanioses	22
2.2 - Biologia Molecular dos tripanossomatídeos	25
2.2.1 - Organização genômica dos tripanossomatídeos	25
2.2.1.1 - Análise comparativa dos genomas de diferentes espécies de <i>Leishmania</i>	26
2.2.2 - RNA polimerases, promotores e transcrição policistrônica	27
2.2.3. - Regulação da expressão gênica nos tripanossomatídeos	29
2.2.3.1 - Processamento em <i>trans</i> e poliadenilação dos mRNAs	30
2.2.3.2 - Mecanismos de degradação de mRNAs e seu controle nos tripanossomatídeos	33
2.2.3.3 - Grânulos de processamento e de estresse	35
2.3 - O processo de síntese protéica, tradução, em eucariotos	36
2.3.1 - Iniciação da tradução em eucariotos	36
2.3.2 - O complexo eIF4F de iniciação da tradução	39
2.3.2.1 - eIF4E, a proteína de ligação ao cap	39
2.3.2.2. - A RNA helicase eIF4A	41
2.3.2.3. - eIF4G, uma proteína modular que estrutura o complexo eIF4F	44
2.3.2.3.1 - Estrutura do eIF4G	44
2.3.2.3.2 - Família eIF4G	46
2.3.2.3.3 - Modificações pós-traducionais do eIF4G	49
2.3.2.3.4 - Bases moleculares da interação eIF4G-eIF4A	50
2.3.2.3.5 - Modo de ligação do eIF4E ao eIF4G	52
2.3.2.3.6 - O sítio de ligação para eIF3 no eIF4G	54
2.3.3 - A proteína de ligação ao poli-A (PABP), parceira do eIF4F	55
2.4 - Complexo eIF4F e PABP nos tripanossomatídeos	57
2.4.1 - PABP nos tripanossomatídeos	57
2.4.2 - eIF4AI e eIF4AIII nos tripanossomatídeos	58

2.4.3 - Isoformas do eIF4E em <i>Leishmania major</i>	60
2.4.4 - Candidatos a eIF4G em <i>Leishmania major</i>	60
3 - OBJETIVOS	63
3.1 - Objetivo Geral	63
3.2 - Objetivos Específicos	63
4 - MATERIAIS E MÉTODOS	64
4.1 - PCR e clonagem/subclonagem dos homólogos ao eIF4F de <i>L. major</i> e <i>T. brucei</i>	64
4.2 - Sequenciamento das construções plasmidiais	68
4.3 - Expressão e purificação de proteínas recombinantes	69
4.4 - Mutagênese sítio-dirigida	70
4.4.1 - Desenho dos oligonucleotídeos	71
4.4.2 - Reação de mutagênese	72
4.5 - Transcrição e expressão de mRNAs sintéticos dos mRNAs	74
4.6 - Ensaio de interação proteína-proteína (<i>Pull-Down</i>)	75
4.9 – Ensaio de <i>Western-blot</i>	77
5 - RESULTADOS	78
5.1- Estudo da interação entre homólogos do eIF4G com as demais subunidades do complexo eIF4F e PABP de <i>L. major</i>	79
5.1.1. Estudo da interação entre os homólogos do eIF4G/eIF4A em <i>L. major</i>	83
5.1.2. Estudo da Interação entre os homólogos eIF4G/eIF4E em <i>L. major</i>	85
5.1.3. Estudo dos domínios de interação entre os homólogos eIF4G/eIF4A e eIF4G/eIF4E em <i>L. major</i> .	87
5.1.4. Alinhamento múltiplo das seqüências <i>LmEIF4G3-4</i> e identificação de motivos conservados	90
5.1.5. Estudo da interação entre os mutantes <i>LmEIF4G3-4</i> com homólogos das subunidades do eIF4F de <i>L. major</i>	93
5.1.6. Estudo reverso da interação entre os mutantes <i>LmEIF4G3-4</i> com homólogos das subunidades do eIF4F de <i>L. major</i>	97
5.1.7. Estudo da interação entre os homólogos de eIF4G e da PABP de <i>L. major</i>	104
5.1.7.1 Estudo da interação entre os homólogos de eIF4G e de PABP de <i>L. major</i>	104
5.1.7.2 Ensaio reverso da interação entre homólogos de eIF4G e de PABP de <i>L. major</i>	109
5.1.8. Formação <i>in vitro</i> do complexo eIF4F de <i>L. major</i>	111
5.2 Estudo comparativo da interação entre os homólogos EIF4G3-4 de <i>L. major</i> com os homólogos EIF4E3-4 e <i>Trypanosoma brucei</i>	114

5.2.1 - Ensaios para investigar a conservação das interações entre homólogos EIF4G3-4 e homólogos EIF4E3-4 em <i>L. major</i> e <i>T. brucei</i>	114
6 - DISCUSSÃO	117
6.1 - Identificação de dois complexos eIF4F em <i>L. major</i> .	117
6.2 - Mapeamento dos motivos de ligação eIF4G/eIF4E	119
6.3 - Análise da interação eIF4G/PABP	122
6.4 - Considerações finais	123
7- RESULTADOS COMPLEMENTARES OBTIDOS DURANTE O DOUTORADO	126
7.1 - Estudo complementar acerca do papel funcional de homólogos de eIF4A em <i>T. brucei</i>	126
7.1.1 - Análise da expressão das proteínas TbEIF4AI-III nas fases procíclicas e sanguíneas de <i>T. brucei</i>	127
7.1.2 Efeito da interferência de RNA nos níveis intracelulares da proteína TbEIF4AIII.	129
7.1.3 - Interação entre homólogos do EIF4A com o homólogo à proteína Magoh em <i>T. brucei</i>	130
7.2 – Discussão dos resultados complementares obtidos durante o Doutorado	132
8 - CONCLUSÕES	134
9 - REFERÊNCIAS BIBLIOGRÁFICAS	135
10 - ANEXOS (ARTIGOS CIENTÍFICOS)	152

1. INTRODUÇÃO

A família Trypanosomatidae, da Ordem Kinetoplastida, inclui as espécies de protozoários flagelados de importância médico-veterinária encontradas nos gêneros *Leishmania* e *Trypanosoma*. Estes parasitas causam as doenças conhecidas como Leishmanioses e as Tripanossomíases que afetam milhões de pessoas em várias regiões do mundo. As espécies mais estudadas de tripanossomatídeos do ponto de vista molecular são *Leishmania major*, o *Trypanosoma cruzi* e o *Trypanosoma brucei*.

Os tripanossomatídeos são eucariotos que apresentam características celulares e moleculares distintas em relação a outros eucariotos, por terem divergido precocemente da linhagem evolutiva que deu origem aos demais. Entre estas, destacam-se a conservação e sintenia dos genes, a transcrição policistrônica, o processamento em *trans* dos pré-mRNAs (*trans-splicing*) e o controle da expressão gênica pós-transcricional através da regulação da meia vida e tradução de mRNAs. Embora a tradução constitua um dos principais processos biológicos dos seres vivos pouco se sabe a seu respeito nos tripanossomatídeos.

A tradução é o processo pelo qual a informação contida na molécula de RNA é utilizada para a síntese de proteínas. Este evento pode ser dividido em Iniciação, Elongação e Terminação. A iniciação da tradução é a etapa mais complexa e o principal alvo de mecanismos de regulação. Ela tem início com a formação do complexo eIF4F constituído pelas subunidades eIF4E (proteína que se liga ao cap dos mRNAs na sua extremidade 5'), eIF4A (RNA helicase) e eIF4G (proteína mantenedora do complexo). Adicionalmente, a proteína eIF4G interage diretamente com a proteína de ligação a cauda Poli-A (PABP) presente na extremidade 3' dos mRNAs, o que permite verificar a integridade destes antes de sua tradução. O eIF4G ainda atua diretamente facilitando o reconhecimento dos mRNAs pelo ribossomo. Homólogos das subunidades do eIF4F e PABP foram identificados em tripanossomatídeos, entre eles cinco homólogos do eIF4G (EIF4G1-5) e múltiplos homólogos

das demais proteínas (EIF4E1-4, EIF4AI e III e PABP1-2) e suas funções vêm sendo investigadas em *Leishmania major* e *Trypanosoma brucei*.

Este trabalho estudou de forma detalhada a interação dos diferentes homólogos de eIF4G de *L. major* com homólogos das demais subunidades do eIF4F, e também da PABP, encontrados nesses organismos. Para este propósito foram realizados ensaios de interação proteína-proteína do tipo *pull-down*, e uma vez identificadas as interações, os domínios foram mapeados e aminoácidos-chaves identificados a partir de ensaios de mutagênese sítio-dirigida. Adicionalmente, algumas das interações obtidas em *L. major* foram confirmadas em *T. brucei* e resultados complementares auxiliaram no estabelecimento de diferentes funções para os ortólogos do eIF4A em *T. brucei*. Os resultados alcançados constituem uma contribuição significativa à compreensão do modo de atuação das subunidades do complexo eIF4F na iniciação da tradução em tripanossomatídeos.

Parte dos resultados obtidos durante esse trabalho já foram incluídos em um artigo publicado na revista *Nucleic Acids Research* e os demais devem fazer parte de três novas publicações, em diferentes estágios de redação, mas cuja submissão deverá ocorrer nos próximos meses.

2. REVISÃO DA LITERATURA

2.1 - Tripanossomatídeos

Os kinetoplastídeos são um grupo de protozoários flagelados que se distinguem pela presença de uma região rica em DNA chamada kinetoplasto, presente em sua grande e única mitocôndria (Stuart *et al.*, 2008). A ordem kinetoplastida compreende desde protozoários de vida livre a patógenos de invertebrados, vertebrados e plantas (Simpson *et al.*, 2006). Nesta ordem se destacam os tripanossomatídeos formados por nove gêneros de parasitas obrigatórios, com destaque para os gêneros *Trypanosoma* e *Leishmania* (Simpson *et al.*, 2006). Estes protozoários apresentam uma organização genômica, estruturas celulares e processos biológicos semelhantes, contudo se destacam por apresentarem mudanças morfológicas significantes durante seu ciclo de vida, visto que habitam dois diferentes nichos biológicos, sendo transmitidos aos humanos por diferentes insetos vetores (Stuart *et al.*, 2008). As principais doenças humanas causadas por estes protozoários são as tripanossomíases, doença de Chagas e doença do Sono, e as diversas formas de Leishmaniose sendo causados respectivamente pelo *Trypanosoma cruzi*, *Trypanosoma brucei* e várias espécies do gênero *Leishmania* (Stuart *et al.*, 2008).

2.1.1 - Gênero *Trypanosoma* e as Tripanossomíases

Do ponto de vista médico, o gênero *Trypanosoma* pode ser representado pelas espécies *T. brucei* e *T. cruzi*, responsáveis respectivamente por duas parasitoses de grande impacto para saúde humana: a Doença do sono (Tripanossomíase Africana) e a Doença de Chagas (Tripanossomíase americana). A Doença do sono é encontrada em 36 países da África Subsaariana, apresentando-se sob duas formas: no Leste da região a subespécie *T. brucei rhodensiense* causa a forma aguda da doença e à Oeste o *T. brucei gambiense* causa a forma crônica da doença (Kennedy, 2008). A Tripanossomíase Africana é fatal se não for tratada. O primeiro estágio da doença (hemolinfático) é acompanhado de sintomas não específicos como febre, dor de cabeça, perda de peso, fraqueza entre outros. O segundo estágio (encefálico) é acompanhado por severos sintomas neurológicos incluindo distúrbios mentais e do sono (Kennedy, 2008). Cerca de 60 milhões de pessoas vivem em áreas de transmissão, com uma prevalência de 300 a 500 mil novos casos e taxa de mortalidade estimada em 30.000 mortes anualmente (Stuart *et al.*, 2008). O ciclo de infecção tem início quando as moscas Tsé-Tsé (gênero *Glossina* – 20 diferentes espécies), ao se alimentarem do hospedeiro humano ou animal parasitado, adquirem as formas tripomastigotas sanguíneas. No trato digestivo do inseto as formas tripomastigotas se transformam em tripomastigotas procíclicas, se dividem por divisão binária e depois se diferenciam em epimastigotas. Estas últimas migram para as glândulas salivares, onde se multiplicam assexuadamente e em seguida se diferenciam em tripomastigotas metacíclicas que são inoculadas no hospedeiro vertebrado durante um novo repasto sanguíneo. No hospedeiro definitivo, ocorre a transformação das formas tripomastigotas metacíclicas para as formas tripomastigotas sanguíneas, que se multiplicam assexuadamente em vários fluidos corpóreos como o sangue, a linfa e o fluido cérebro-espinhal. Estas formas ao serem adquiridas novamente pelo inseto recomeçam um novo ciclo de infecção (Figura 1) (Barrett *et al.*, 2003; Stuart *et al.*, 2008).

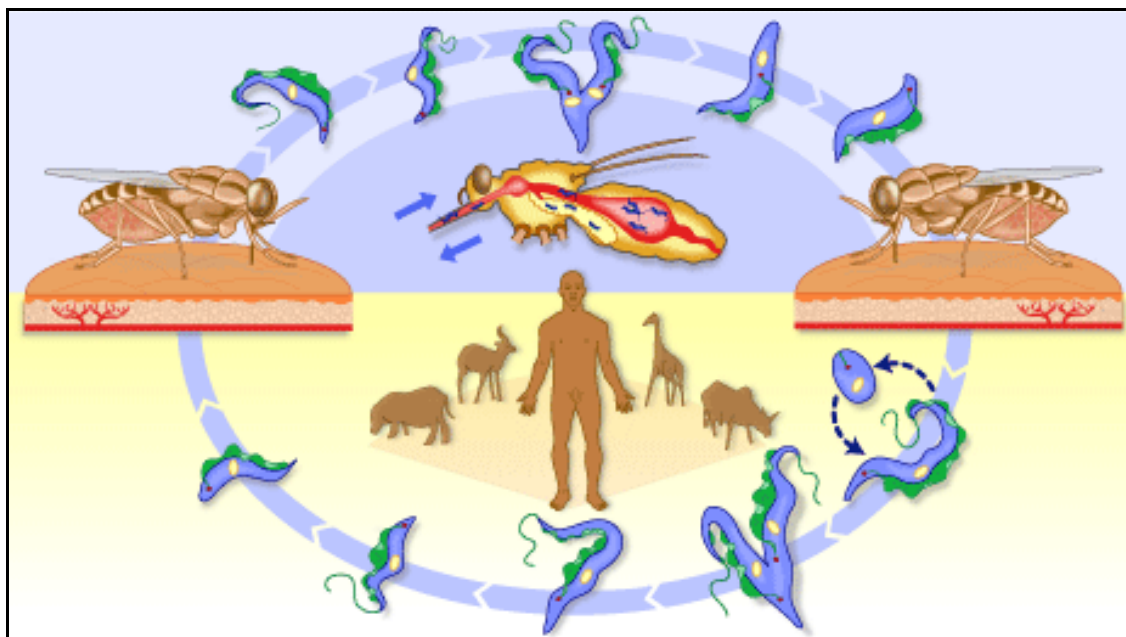


Figura 1. Ciclo de transmissão do *Trypanosoma brucei*. Parasitas no estágio de tripomastigotas metacíclicos, liberados pela mosca Tse-Tse durante seu repasto sanguíneo, entram em contato com o hospedeiro vertebrado se diferenciando na forma tripomastigota sanguínea. As formas tripomastigotas sanguíneas podem ser ingeridas pelos insetos vetores, onde se diferenciam nas formas procíclicas. No interior do hospedeiro intermediário as formas procíclicas migram para as glândulas salivares onde se diferenciam novamente em tripomastigotas metacíclicos, fechando o ciclo de transmissão (<http://www.who.int/tdr>).

O *T. cruzi* é o agente etiológico da doença Chagas, que afeta 18 países (distribuídos pela América do Sul, América Central e México), sendo responsável por uma prevalência de infecção de aproximadamente 11 milhões de pessoas, com uma média de 700 mil novos casos e 14000 mortes anuais (Stuart *et al.*, 2008). A doença de Chagas de acordo com as manifestações clínicas pode ser classificada em duas fases: A fase aguda e a fase crônica. Na fase aguda da Doença de Chagas muitas pessoas desconhecem que estão infectadas pelo *T. cruzi* e desenvolvem sintomas benignos e inespecíficos da patologia como febre e inflamação nos linfonodos e, ocasionalmente, reação inflamatória no sítio da picada ou inchaço no olho.

Após anos ou décadas, aproximadamente 30% dos pacientes desenvolvem a fase crônica da doença que se manifesta principalmente por complicações cardíacas e/ou gastrointestinais e, devido ao severo grau de debilitação que a patologia conduz, pode levar nos casos não tratados a morte (Tarleton *et al.*, 2007; Stuart *et al.*, 2008). Estes parasitas são transmitidos ao homem principalmente por insetos vetores, hemípteros da subfamília *Triatominae* (vulgarmente conhecidos como barbeiros), que durante seu repasto sanguíneo liberam parasitas (tripomastigotas metacíclicos) em suas fezes que penetram no local da picada. Uma vez em contato com a corrente sanguínea, os tripomastigotas metacíclicos invadem vários tipos celulares incluindo os macrófagos e se diferenciam na forma amastigota replicativa. Em seguida, as células hospedeiras são lisadas liberando os amastigotas, que podem tanto invadir novas células como se diferenciar novamente na forma tripomastigota e permanecer no fluido sanguíneo (extracelular). Durante o repasto sanguíneo, as tripomastigotas são ingeridas pelo inseto vetor e, no interior de seu trato digestivo, se diferenciam nas formas epimastigotas que se replicam assexuadamente e se diferenciam nas formas tripomastigotas metacíclicas fechando o ciclo de transmissão (Figura 2) (Barrett *et al.*, 2003; Stuart *et al.*, 2008). Alternativamente, o *T. cruzi* pode ser transmitido por transfusão sanguínea, transplante de órgãos, transmissão vertical e pela ingestão de comida ou bebida contaminada (Stuart *et al.*, 2008).

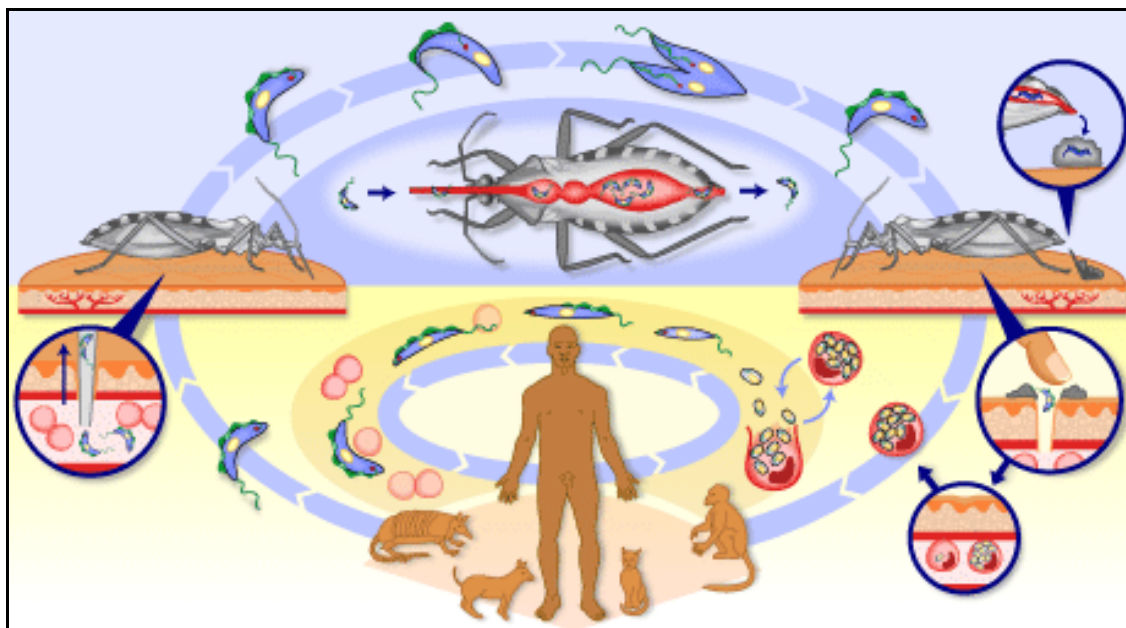


Figura 2. Ciclo de transmissão do *Trypanosoma cruzi*. Parasitas tripomastigotas metacíclicos, liberados nas fezes de vetores da família *Triatominae*, entram em contato com a corrente sanguínea do hospedeiro pelo local da picada, invadindo vários tipos celulares, onde se diferenciam na forma amastigota. Após replicações sucessivas, as células são lisadas liberando estas formas na corrente sanguínea. As amastigotas tanto podem invadir células não infectadas como se diferenciar em formas tripomastigotas. Os tripomastigotas podem ser ingeridos pelos insetos vetores, onde se diferenciam na forma epimastigota. No interior do vetor os epimastigotas se replicam no intestino, e em seguida migram para o reto, onde se diferenciam em tripomastigotas metacíclicos, fechando o ciclo de transmissão (<http://www.who.int/tdr>).

2.1.2 - Gênero *Leishmania* e as Leishmanioses

O gênero *Leishmania* é representado por mais de 20 espécies que são transmitidas ao homem pela picada de aproximadamente 30 espécies de vetores, insetos dípteros da subfamília Phlebotominae (Chappuis *et al.*, 2007). Estes protozoários são os agentes etiológicos das doenças conhecidas como Leishmanioses, um grave problema de saúde

pública. De maneira geral, as Leishmanioses podem ser divididas em dois grandes grupos de doenças: Leishmaniose Cutânea (LC) e Leishmaniose Visceral (LV). A LC representa a maioria dos casos de Leishmaniose registrados, geralmente não é fatal, se limita a pele e pode evoluir para cura espontânea. Esta é a forma da doença causada pela *L. major*. Concomitantemente com a LC pode se estabelecer uma forma mais severa de Leishmaniose que atinge também as mucosas, sendo denominada assim de Leishmaniose Mucosa Cutânea (LMC). A LMC é caracterizada por lesões não cicatrizantes que podem deformar os indivíduos infectados, é potencialmente fatal, e tem como principal agente etiológico a *L. braziliensis* (Reithinger *et al.*, 2007). A LV é uma doença sistêmica crônica caracterizada por febre, hepatoesplenomegalia, linfadenopatia, pancitopenia, astenia, fraqueza e, se não for tratada, leva a morte dos indivíduos parasitados (Reithinger *et al.*, 2007). Um dos principais agentes etiológicos da LV é a *Leishmania infantum*. A incidência anual das Leishmanioses é de 1.5 a 2 milhões de casos, sendo registradas acima de 51.000 mortes por ano. Cerca de 350 milhões de pessoas vivem em áreas de risco, com uma prevalência estimada de infecção em torno de 12 milhões de casos (Reithinger *et al.*, 2007; Reithinger, 2008). As Leishmanioses são encontradas em 88 países (WHO, 2008), sendo 90% dos casos de LC notificados no Afeganistão, Argélia, Irã, Peru, Arábia Saudita, Síria e Brasil e 90% dos casos de LV estão distribuídos em Bangladesh, Índia, Nepal, Sudão e Brasil. Estes protozoários são transmitidos ao homem por fêmeas de insetos dípteros hematófagos do gênero *Phlebotomus* no velho mundo, e pelo gênero *Lutzomyia* no novo mundo (Santos *et al.*, 2008).

O ciclo biológico da *Leishmania* tem início quando mosquitos infectados picam o hospedeiro vertebrado e regurgitam formas promastigotas do parasito no interior da sua pele. Estes são internalizados por macrófagos onde, dentro dos fagolisossomos, se transformam na forma amastigota. Após sucessivos ciclos de replicação por divisão binária, eventualmente os amastigotas podem lisar as células do hospedeiro e infectar outros macrófagos do mesmo tecido ou de tecidos diferentes. Durante o repasto sangüíneo, mosquitos não contaminados

tornam-se infectados ao adquirirem os amastigotas livres ou no interior dos macrófagos. Estes, por sua vez, se transformam em promastigotas no intestino do inseto vetor e, após divisões celulares e diferenciação, promastigotas infectantes migram para probóscide completando o ciclo (Figura 3). Além do homem, outros mamíferos constituem reservatórios para diferentes espécies de *Leishmania sp*, tendo um papel importante na epidemiologia da doença.

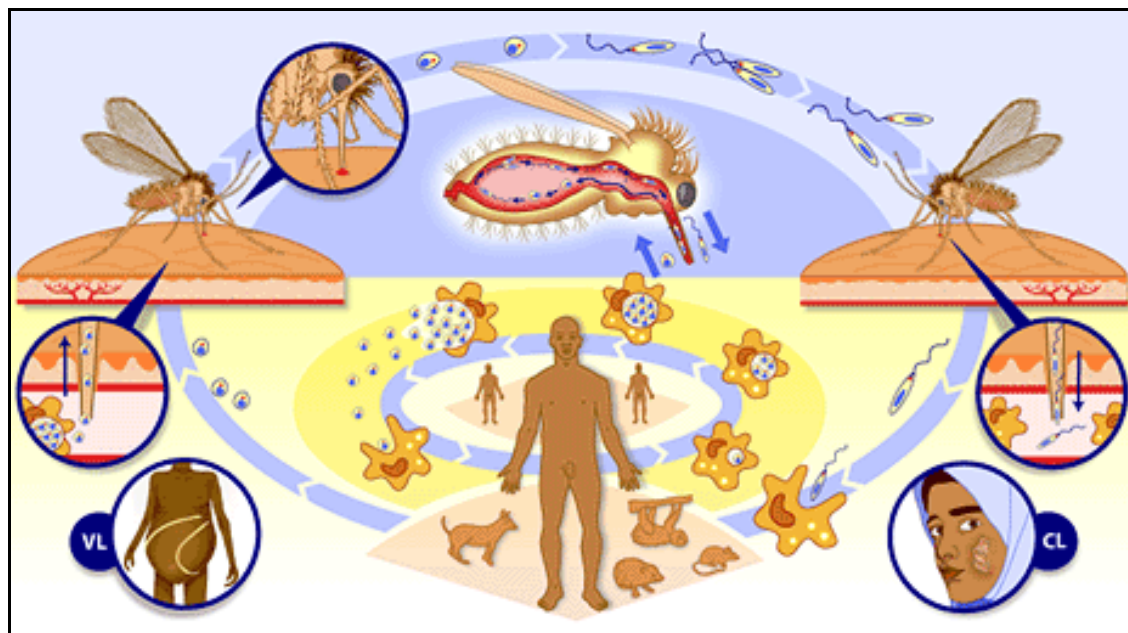


Figura 3: Ciclo de transmissão da *Leishmania sp*. Promastigotas, regurgitados por vetores dos gêneros *Phlebotomus* e *Lutzomyia*, entram em contato com a corrente sanguínea humana e invadem macrófagos, onde se diferenciam na forma amastigota. Após sucessivas replicações das formas amastigotas, os macrófagos são lisados liberando-as e estas podem invadir novos macrófagos ou serem ingeridas por insetos vetores, onde se diferenciam em promastigotas, fechando o ciclo de transmissão (<http://www.who.int/tdr>).

2.2 - Biologia Molecular dos tripanossomatídeos

2.2.1 - Organização genômica dos tripanossomatídeos

A publicação dos três genomas dos tripanossomatídeos *Trypanosoma. brucei*, *Trypanosoma cruzi* e *Leishmania major* (Tritryps) foi um avanço científico de extrema relevância por impulsionar pesquisas que permitam a compreensão da biologia destes protozoários, as relações vetor – parasita – hospedeiro, assim como o estudo de novos quimioterápicos e desenvolvimento de testes diagnósticos e vacinas para as doenças causadas pelos mesmos (Kissinger, 2006). Acredita-se que os tripanossomatídeos tenham divergido a partir de um ancestral comum há cerca de 200 a 500 milhões de anos, e apesar deste fato, estas espécies apresentam elevada sintonia (conservação da ordem dos seus genes), assim como um grande número de genes em comum (Ginger, 2005; El-Sayed *et al.*, 2005a; Kissinger, 2006). O genoma dos Tritryps contém entre 25 e 55 megabases, distribuídos de 11 a 36 cromossomos diplóides. Respectivamente os genomas de *T. brucei*, *T. cruzi* e *L. major* contém 8100, 12000 e 8300 genes que codificam para proteínas. Destes 6158 são comuns para as três espécies (Tabela 1) e estão organizando em agrupamentos de genes sintênicos, que fornecem as unidades para transcrição policistrônica de dezenas de genes não relacionados (Ginger, 2005; El-Sayed *et al.*, 2005a). A identidade de aminoácidos entre proteínas das três espécies tem uma média de 57% entre *T. brucei* e *T. cruzi*, e de 44% entre *L. major* e as duas espécies de *Trypanosoma*, refletindo suas relações filogenéticas (El-Sayed *et al.*, 2005a), entretanto dos 1617 domínios de proteína encontrados, apenas 5% são exclusivos de cada espécie. Elementos móveis de DNA, RNAs estruturais e muitos genes espécie-específicos são encontrados nas regiões de quebra de sintonia (Kissinger, 2006). Apesar das grandes semelhanças existentes, há substanciais diferenças que provavelmente

refletem adaptações espécie-específicas a diferentes modos de sobrevivência de cada organismo de acordo com os nichos que eles exploram (Ivens *et al.*, 2005; El-Sayed *et al.*, 2005a). Em *T. cruzi*, estima-se que 50% do genoma seja constituído por seqüências repetitivas representadas por retrotransposons, seqüências teloméricas e famílias gênicas de proteínas de superfície. As maiores famílias gênicas são as proteínas de superfície associadas à mucina (MASP), trans-sialidases, mucinas e a família da glicoprotease 63, as quais são geralmente específicas de *T. cruzi* e constituem cerca de 18% dos genes codificadores de proteínas (El-Sayed *et al.*, 2005b), o que poderia justificar seu conteúdo gênico relativamente maior em relação a *T. brucei* e *L. major*.

Tabela I. Organização genômica dos três principais tripanossomatídeos. (Adaptado de Ginger, 2005). *Incluindo pseudogenes. ** Podem variar em outras espécies de *Leishmania*, de 34 a 36, devido a fusões cromossômicas.

	Nº de pares de cromossomos	Tamanho do genoma Haplóide (Mpb)	Nº de genes (Haplóide)	Nº de genes específicos da espécie	Nº de genes em comum
<i>Trypanosoma brucei</i>	11	25	9068*	1392	6158
<i>Trypanosoma cruzi</i>	~28	55	~12000	3736	6158
<i>Leishmania major</i>	36**	35	8311*	910	6158

2.2.1.1 - Análise comparativa dos genomas de diferentes espécies de *Leishmania*

Os genomas das espécies *Leishmania major*, *Leishmania braziliensis* e *Leishmania infantum*, respectivos agentes etiológicos das LC, LMC e LV, já estão disponíveis em bancos públicos de dados (Peacock *et al.*, 2007). Estima-se que a divergência entre as espécies do gênero *Leishmania* tenha ocorrido de 15 a 20 milhões de anos atrás (Lukes *et al.*, 2007), e apesar da distância evolutiva estas espécies apresentam uma alta conservação da arquitetura genômica e do repertório gênico (Smith *et al.*, 2007). Os três genomas têm em média 33 Mpb

distribuídos em 36 cromossomos para *L. major* e *L. infantum* e 35 cromossomos para *L. braziliensis* e a sintonia dos blocos gênicos entre as espécies é de 99%. De um total de aproximadamente 8.300 genes, apenas 200 são diferencialmente distribuídos entre as três espécies. A *L. braziliensis*, a mais divergente, possui 47 genes ausentes nas outras espécies, enquanto *L. infantum* tem 27 genes específicos e *L. major* apenas 5 genes (Peacock *et al.*, 2007). Assim poucos genes podem estar relacionados com as diferentes patologias desenvolvidas, e ao tropismo específico para determinados tecidos. O nível de conservação dos aminoácidos nas diferentes proteínas é de 92% entre *L. major* e *L. infantum* e de 77% quando se compara *L. braziliensis* com as demais (Peacock *et al.*, 2007). O genoma de *Leishmania* é assim altamente estável e não tem sofrido grandes mudanças durante a evolução, como demonstrado para *L. major* e *L. infantum*, principalmente devido à perda de elementos móveis (Smith *et al.*, 2007). Contudo, *L. braziliensis*, a mais divergente espécie de *Leishmania* seqüenciada, apresenta elementos transponíveis únicos (Smith *et al.*, 2007).

2.2.2 – RNA polimerases, promotores e transcrição policistrônica.

Nos tripanossomatídeos, além da RNAP mitocondrial que é envolvida na transcrição de seus maxicírculos de kDNA, são encontradas, como em outros eucariotos, três distintas RNA polimerases (RNAPI, II e III). A RNAP I transcreve os genes que codificam para RNAs ribossômicos (rRNA) e em *T. brucei*, promove ainda a transcrição de genes que codificam antígenos de superfície estágio específicos (VSG e PARP). A RNAP II transcreve genes codificantes de proteínas, além de pequenos RNAs nucleares (snRNAs) que funcionam no processamento dos pré-mRNAs em mRNAs maduros. Já a RNAP III transcreve algumas classes de pequenos RNAs, entre eles tRNAs, rRNA 5S e também snRNAs (Campbell *et al.*, 2003; Das *et al.*, 2008). Alguns homólogos de subunidades das três RNAs polimerases estão ausentes nos tripanossomatídeos e diferenças importantes são observadas nas subunidades

presentes (Ivens *et al.*, 2005), possivelmente refletindo o modo como elas desempenham suas funções. De maneira geral, há uma escassez de homólogos aos fatores de transcrição nos genomas dos tripanossomatídeos. Embora a proteína de ligação ao motivo TATA (*TATA-Box*), importante fator de transcrição presente em diversos organismos, é preservada nos tripanossomatídeos e tem papel essencial para o funcionamento de suas polimerases (Ruan *et al.*, 2004).

Muito pouco é conhecido sobre o mecanismo da iniciação da transcrição dos genes codificadores de proteínas nos tripanossomatídeos, e poucos promotores têm sido funcionalmente caracterizados (Clayton, 2002). Seus genes em geral não possuem íntrons, exceto em apenas quatro exemplos (Ivens *et al.*, 2005), sendo separados um dos outros por regiões intergênicas e reunidos em grupos de genes contíguos que seguem uma mesma orientação numa mesma fita do DNA. Estes grupos podem conter dezenas de genes que codificam proteínas não-relacionadas, do ponto de vista funcional (Figura 4) (El-Sayed *et al.*, 2005a). A transcrição desses grupos de genes é realizada de forma contínua em eventos únicos de transcrição pela RNAP II, gerando RNAs mensageiros precursores policistrônicos (transcrição policistrônica) (Palenchar e Bellofatto, 2006). Grupos de genes vizinhos, em um mesmo cromossomo, muitas vezes são codificados por fitas distintas da dupla hélice de DNA e estão posicionados em orientação oposta.

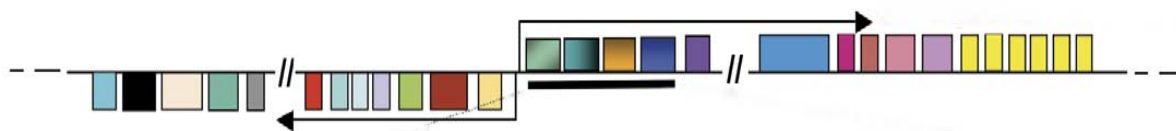


Figura 4. Transcrição policistrônica nos tripanossomatídeos. Blocos gênicos não-relacionados estão agrupados em fitas distintas do DNA e são bidirecionalmente transcritos pela ação da RNAP II, a partir de regiões de troca de fitas (Adaptado de Haile e Papadopolou, 2007).

Apesar de poucos promotores terem sido caracterizados nos tripanossomatídeos, alguns promotores para RNAP I e para RNAP III vem sendo identificados (Clayton, 2002; Campbell *et al.*, 2003). Apenas o promotor do gene *Spliced Leader* foi descrito para RNAPII. Em situações onde dois grupos de genes estão posicionados lado a lado em um mesmo cromossomo em orientações opostas nas duas fitas do DNA, evidências experimentais sugerem que a transcrição policistrônica se inicie bidirecionalmente em sentidos opostos, na região entre os grupos denominada de troca de fitas (*strand-switch* – Martinez-Calvillo *et al.*, 2004; Ivens *et al.*, 2005). A finalização da transcrição se daria ao final dos grupos onde se co-localizam genes para tRNA, rRNA e ou snRNA (Ivens *et al.*, 2005). Acredita-se assim que a RNAP II transcreva os genes constitutivamente em níveis equivalentes (Campbell *et al.*, 2003) e que qualquer controle da expressão destes genes ocorra apenas após sua transcrição (Clayton, 2002; Ivens *et al.*, 2005).

2.2.3 - Regulação da expressão gênica nos tripanossomatídeos

A transcrição em unidades policistrônicas dos tripanossomatídeos, onde um único promotor serve a vários genes, tem como consequência a síntese de mRNAs não associados funcionalmente em níveis equivalentes. A ausência de regulação pela atividade individual do promotor, associada a níveis diferenciais de expressão gênica, indica que nos tripanossomatídeos a regulação desta expressão é principalmente pós-transcricional (Stiles *et al.*, 1999; Campbell *et al.*, 2000). Entre os diversos mecanismos que podem estar sujeitos ao controle pós-transcricional, nos tripanossomatídeos, se destacam o processamento dos pré-mRNAs ainda no núcleo (processamento em *trans*, ou *trans-splicing*, e poliadenilação), o controle da estabilidade de mRNAs maduros no citoplasma e o controle da sua tradução (Clayton e Shapira, 2007).

2.2.3.1 – Processamento em *trans* e poliadenilação dos mRNAs

Nos eucariotos em geral, os genes são formados por seqüências codificantes para proteína (éxons) e por seqüências não codificantes (íntrons). Normalmente, uma RNAP II se liga a seqüência promotora, que antecede o gene, e inicia o processo de transcrição. Ao transcrito primário é adicionado, na sua extremidade 5' o nucleotídeo cap (m^7GTP ou 7-metilguanosina) e na sua extremidade 3' uma cauda de poli-adenosina (poli-A), durante o processo de transcrição. Os transcritos primários são processados por um complexo enzimático formado por pequenos RNAs e proteínas (*spliceossomo*) em um processamento em *cis* (*cis-splicing*), onde os íntrons são retirados e os éxons unidos para obtenção dos RNAs mensageiros maduros que serão utilizados na tradução (Moore e Proudfoot, 2009).

A maturação de mRNAs individuais a partir do transcrito policistrônico, nos tripanossomatídeos, necessita da ocorrência de dois mecanismos acoplados de processamento de RNA (LeBowitz *et al.*, 1993): o processamento em *trans* (*trans-splicing*) e a poliadenilação. Ambas as reações são dirigidas por seqüências comuns ricas em polipirimidinas, presentes nas regiões intergênicas (Benz *et al.*, 2005). O processamento em *trans* é o processo onde uma seqüência de 39 a 41 nucleotídeos (seqüência *spliced leader*, SL, ou mini-éxon), proveniente da clivagem do RNA SL precursor, é adicionada a extremidade 5' de todos os mRNAs dos tripanossomatídeos (Liang *et al.*, 2003). A seqüência SL é comum a todos os mRNAs de um mesmo organismo, porém pode ter pequenas variações entre espécies de tripanossomatídeos. O mecanismo de processamento em *trans* se dá através de reações de *trans*-esterificação, onde o sítio GU doador presente no RNA SL precursor funde-se a uma região próxima ao sítio AG aceptor do mRNA policistrônico, denominada de ponto de ramificação (normalmente uma adenosina), formando uma estrutura intermediária em forma de Y. Esse sinal é reconhecido pelo complexo de processamento (*spliceossomo*) encontrado nestes protozoários, sendo as estruturas intermediárias removidas e a seqüência SL fusionada

a extremidade 5' de cada mRNA. O sinal acceptor para o processamento no pré-mRNA propriamente é o dinucleotídeo AG quando localizado imediatamente após uma região rica em pirimidinas (Figura 5) (Liang *et al.*, 2003). O processo de *trans-splicing* também serve para adicionar o cap característico de tripanossomatídeos, presente na extremidade 5' da SL, denominado de cap4. Além do radical 7-metilguanossina, o cap4 é formado ainda pelos quatro nucleotídeos da SL (AACU) também modificados por metilação nas suas riboses e em algumas de suas bases nitrogenadas (Bangs *et al.*, 1992).

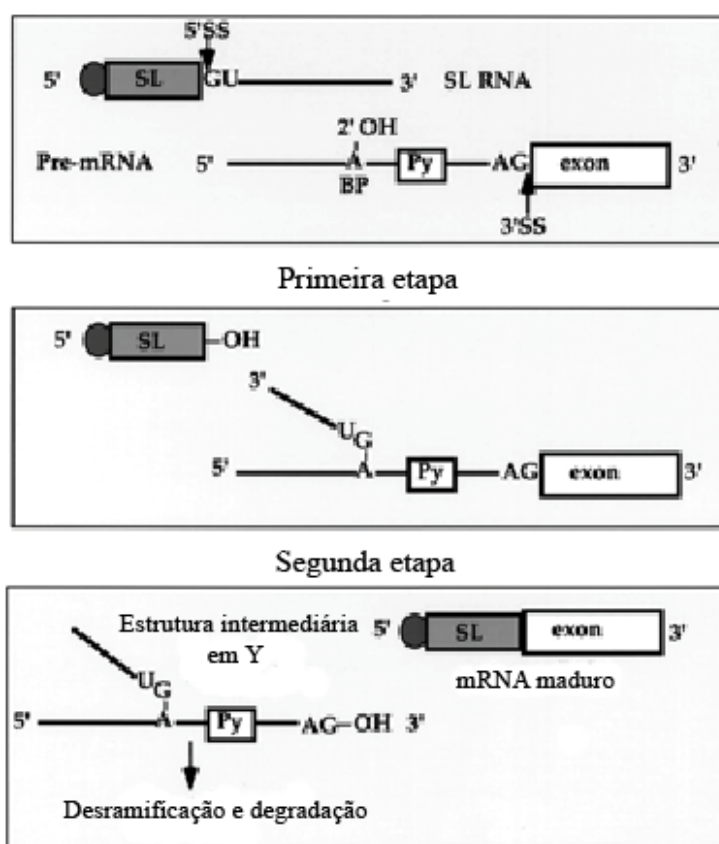


Figura 5. Representação esquemática do trans-splicing. Demonstrando os sítios aceptores de splicing da SL (GU) e do mRNA (AG), ponto de bifurcação (BP) seguida da região rica em pirimidinas (Py). Ao final do processo a porção 5' do mRNA recebe a SL, caracterizando o mRNA maduro (Liang *et al.*, 2003).

Nos mamíferos, os pré-mRNAs possuem um sinal típico para poliadenilação como a sequência AAUAAA. O complexo de proteínas necessárias para as etapas de clivagem são os fatores específicos de clivagem e poliadenilação (CPSF), o fator estimulador de clivagem (CstF), os fatores de clivagem Im e IIm (CFIm e CFIIIm) e a RNAP II. A poli-A polimerase (PAP), PABP (Proteína de ligação à cauda poli-A) e a FIP1 (proteína de ligação a PAP) estão envolvidas na reação de poliadenilação. A proteína FIP1 tem um papel fundamental em controlar a atividade da síntese da cauda poli-A, através de múltiplas interações dentro do complexo de poliadenilação com as proteínas PAP e CPSF30. Nos tripanossomatídeos nenhuma sequência típica de poliadenilação foi encontrada, sendo este processo sinalizado pela região rica em pirimidinas, previamente citada, localizada entre dois éxons adjacentes. Dessa forma, esta região serve tanto de sinal para a poliadenilação do éxon anterior, como para a adição da sequência SL no éxon posterior (Figura 6). Entretanto foram encontrados homólogos funcionais para a CPSF30, PAP, PABP e FIP1 nos *Trytryps* (Mair *et al.*, 2000; Bates *et al.*, 2000; Hendriks *et al.*, 2003; Ivens *et al.*, 2005; Bercovich *et al.*, 2009a), indicando que apesar do processo de poliadenilação dos tripanossomatídeos apresentar-se peculiar, algumas semelhanças aos outros eucariotos, como mamíferos e leveduras, provavelmente são funcionalmente conservadas. Os demais membros do complexo CPSF foram encontrados nos *TryTrips*, entretanto apenas um homólogo para uma das subunidades do CstF, a subunidade CstF50, foi encontrada (Ivens *et al.*, 2005; Benz *et al.*, 2005). Nos mamíferos, CstF50 interage com o domínio C-terminal da RNAP II e promove uma conexão entre a poliadenilação e a transcrição. A ausência do domínio C-terminal da RNAP II dos tripanossomatídeos, pode ser refletido pela ocorrência da transcrição policistrônica e resultante desacoplamento da transcrição e poliadenilação (Ivens *et al.*, 2005).

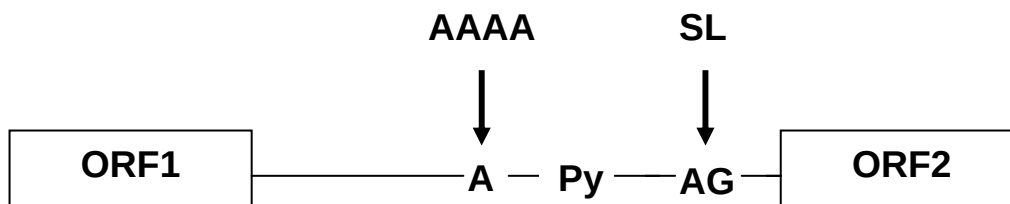


Figura 6. Processamento de mRNAs nos tripanossomatídeos. O trato de polipirimidinas (Py) serve com sinal de *trans-splicing* para o mRNA contendo a ORF2, e de sinal de poliadenilação para o mRNA contendo a ORF1 (Benz et al., 2005).

2.2.3.2 – Mecanismos de degradação de mRNAs e seu controle nos tripanossomatídeos

Em mamíferos e leveduras, a degradação de mRNAs é iniciada com o mecanismo de desadenilação, onde enzimas conhecidas como desadenilases promovem a redução e/ou eliminação da cauda poli-A. Em seguida, o mRNA pode ser degradado por duas vias. Na primeira via, a desadenilação estimula a remoção do cap (*deccaping*) pelas enzimas Dcp1 e Dcp2, permitindo a ação da exonuclease Xrn1 que degrada o mRNA na direção 5'-3' (Parker e Song, 2004). Na outra via, o mRNA pode ser degradado na direção 3'-5' através de um complexo enzimático, exossomo, contendo onze proteínas com diversas funções na degradação de mRNAs (Buttner *et al.*, 2006). Muitos dos processos envolvidos na degradação do mRNA parecem ocorrer nos tripanossomatídeos. Não foram encontrados homólogos para as enzimas Dcp1 e Dcp2, todavia a atividade de *deccaping* foi verificada em extratos de *Leptomonas seymouri* (Milone *et al.*, 2002). A desadenilação também foi demonstrada em extratos de *L. seymouri* e ocorre durante a degradação de RNA *in vivo* (Haile *et al.*, 2003; Li *et al.*, 2006). O exossomo de *T. brucei* tem uma composição e organização semelhante àqueles encontrados em leveduras e mamíferos (Estevez *et al.*, 2001). Já a exonuclease Xrn1, é representada em tripanossomatídeos por quatro homólogos possíveis: XRNA, XRNB,

XRNC e XRND. As funções de XRNB e XRNC ainda são desconhecidas, enquanto XRND é uma exonuclease nuclear, e XRNA pode ser encontrada tanto no núcleo como no citoplasma e está envolvida na degradação de mRNAs (Li *et al.*, 2006).

Nos mamíferos, a rápida degradação de mRNAs que codificam para citocinas ou reguladores do ciclo celular é crítico na regulação da resposta imune e do crescimento celular. Estes mRNAs contêm elementos ricos em AU (*AU-rich elements* – AREs) nas suas regiões 3'-UTR. Os AREs são reconhecidos por proteínas reguladoras que estabilizam ou desestabilizam o mRNA por inibirem ou promoverem a desadenilação. Assim mRNAs que contêm AREs podem ser degradados nas direções 3'-5' e 5'-3' (Stoecklin *et al.*, 2006). Nos tripanossomatídeos têm sido identificadas seqüências 3'-UTR, envolvidas em regulação estágio-específicas, similares aos AREs. Estas regiões são elementos de instabilidade ricos em U (*U rich instability elements* – UREs). Em *T. brucei* uma URE de 26 nucleotídeos foi associada à instabilidade do mRNA codificante para a proteína prociclina EP nas formas sanguíneas, sendo o mesmo mRNA estável na forma procíclica (Clayton e Shapira, 2007). Similarmente vários mRNAs em *T. cruzi* (D'Orso *et al.*, 2003) e o mRNA da amastina de *L. infantum* (Haile *et al.*, 2008) são regulados via UREs. Adicionalmente, em *L. major* foram identificadas em regiões 3'-UTR de mRNAs duas classes de pequenos retroposons (SIDERs – *Short Interspersed Degenerated Retroposons*) (Bringaud *et al.*, 2007). O SIDER1 participa na regulação estágio-específica de alguns mRNAs em *L. major*, como por exemplo o mRNA da amastina, e o SIDER2 foi demonstrado ter um papel global da degradação de mRNAs, provavelmente desempenhando uma modulação constitutiva da expressão gênica (Haile e Papadopoulos, 2007).

Proteínas que ligam a AREs e influenciam a taxa de degradação do mRNA em mamíferos têm motivos de reconhecimento de RNA do tipo RRM (*RNA recognition motif*), motivos CCCH *zinc fingers* ou domínios KH (Barreau *et al.*, 2005). Consistente com a predominância de mecanismos pós-transcricionais, o genoma de tripanossomatídeos tem

revelado um grande número de candidatos a proteínas de ligação a RNA. Entre estes podem ser citados genes que codificam proteínas contendo RRM's (De *et al.*, 2005), domínio Pumilio (PUF) e domínios CCCH *zinc fingers* (Ivens *et al.*, 2005). Entretanto, muitos problemas têm surgido na tentativa de purificar estas proteínas utilizando as seqüências regulatórias presentes na região 3'-UTR, e assim poucos exemplos foram demonstrados experimentalmente. Entre estes podem ser citados as proteínas UBP1 e UBP2 relacionadas à regulação estágio-específica de um pequeno mRNA da mucina em *T. cruzi* (D'Orso e Frasch, 2002).

2.2.3.3 - Grânulos de processamento e de estresse

Nos eucariotos, a formação de grânulos de RNA é um importante mecanismo de controle pós-transcricional da expressão gênica (Anderson e Kedersha, 2006). Estes podem ser classificados em grânulos de processamento (*Processing bodies* ou *P-bodies*) e grânulos de estresse. Os *P-bodies* são complexos dinâmicos, proporcionais em número e tamanho à quantidade de mRNAs não-traduzidos acumulados, onde se co-localizam enzimas envolvidas no *decapping* e processamento/degradação de mRNAs. A partir destes grânulos é que se define quando e se mRNA estocados serão silenciados, degradados ou usados para tradução (Eulalio *et al.*, 2007; Balagopal e Parker, 2009). Os grânulos de estresse são definidos por apresentarem mRNAs não-traduzidos complexados com um subconjunto de fatores de iniciação da tradução (eIF4E, eIF4G, eIF4A, eIF3, eIF2 e PABP), e a subunidade ribossomal 40S (Anderson e Kedersha, 2008). Eles normalmente são visualizados em condições de estresse, quando a iniciação da tradução é inibida (Balagopal e Parker, 2009). Várias observações demonstram que os mRNAs citoplasmáticos deslocam-se ciclicamente entre, *P-bodies*, grânulos de estresse e polissomos (Balagopal e Parker, 2009) e assim transcritos podem ser selecionados para degradação, estocagem ou tradução. Grânulos de mRNA estão presentes nos tripanossomatídeos, todavia parece haver diferenças em relação a outros

eucariotos, e assim, o papel destes grânulos precisa ser melhor investigado (Cassola *et al.*, 2007; Holetz *et al.*, 2007; Kramer *et al.*, 2008).

2.3 - O processo de síntese protéica, tradução, em eucariotos

A síntese de proteínas ou tradução é um importante e complexo mecanismo que é imprescindível a todos os seres vivos, tendo um importante papel no controle da expressão gênica, e constitui o alvo principal de interesse deste trabalho. A regulação da tradução é encontrada em eventos celulares tais como crescimento, proliferação e desenvolvimento (Gingras *et al.*, 1999). O processo de tradução pode ser dividido em quatro etapas: iniciação, alongação, terminação e reciclagem dos ribossomos (Ganoza *et al.*, 2002; Marintchev e Wagner, 2004). Destas a etapa mais complexa, e aquela mais sujeita a mecanismos de regulação, que definem entre outros aspectos quando a tradução de um mRNA ocorre, é a iniciação da tradução.

2.3.1 - Iniciação da tradução em eucariotos

Em eucariotos a síntese de proteínas se inicia primariamente por um mecanismo que requer o nucleotídeo cap no mRNA, e depende de proteínas acessórias que recrutam e posicionam o ribossomo corretamente no codon de iniciação da tradução, AUG. Este processo é conhecido como iniciação cap-dependente da tradução. Entretanto alguns mRNAs virais e celulares usam como alternativa seqüências nas quais o ribossomo liga-se diretamente ao mRNA (IRES- *internal ribosome entry sites* – seqüências internas de entrada para ribossomo), substituindo a utilização de muitas proteínas e da estrutura cap, num mecanismo denominado por essa razão de iniciação cap-independente da tradução (Kieft, 2008). Neste trabalho será abordada a forma clássica de iniciação da tradução, ou seja, a iniciação cap-dependente da tradução.

A iniciação da tradução é o conjunto de eventos que conduz a formação do ribossomo 80S no mRNA, no qual o códon AUG é pareado com o anticódon UAC do Met-tRNAi^{met} (RNA transportador da metionina) no sítio peptidil (P) do ribossomo. Em eucariotos superiores este processo requer as duas subunidades ribossomais (40S e 60S), envolve pelo menos 12 fatores de iniciação eucarióticos (eIFs – *Eukariotic Initiation Factors*) e a ligação e hidrólise de ATP e GTP (Pestova *et al.*, 2007). Os eIFs (eIF1, eIF1A, eIF2, eIF2B, eIF3, eIF4F, eIF4B, eIF4H, eIF5, eIF5A, eIF5B e eIF6) participam diretamente das diferentes etapas do processo, são o alvo principal de mecanismos de regulação e constituem aqueles elementos mais sujeitos a variações entre grandes grupos de organismos. O processo de iniciação da tradução propriamente dito está então ilustrado na Figura 6. Resumidamente, ele começa com a formação do complexo ternário Met-tRNAi-eIF2-GTP (1). Com a ajuda dos eIF1, eIF1A e eIF3, o complexo ternário se liga à subunidade ribossomal 40S resultando na formação do complexo 43S (2). Neste estágio, o complexo eIF4F (formado pelas subunidades eIF4E, eIF4A e eIF4G), em conjunto com a proteína de ligação à cauda poli-A (PABP – não incluída entre os eIFs clássicos) e o eIF3, se liga ao mRNA (3) e o direciona para o complexo 43S resultando no complexo 48S (4). Este último desloca-se no sentido 5'-3' pela região 5' não-traduzida do mRNA (5'-UTR) (5), uma etapa facilitada pelo eIF4F junto com os fatores eIF4B e eIF4H, até encontrar o códon de iniciação da tradução (6). Em seguida, o fator eIF5 estimula a hidrólise do GTP ligado ao eIF2 e o eIF5B-GTP promove a liberação dos fatores ligados à subunidade 40S e sua associação à subunidade 60S, reconstituindo o ribossomo 80S (7). Após hidrólise do GTP ligado ao eIF5B (8), o ribossomo está apto para dar início a fase de elongação da tradução (Marintchev e Wagner, 2004; Pestova *et al.*, 2007; Acker *et al.*, 2009). Por fim, após a síntese protéica, o fator eIF6 atua no ribossomo promovendo a dissociação das subunidades 60S e 40S que são recicladas para novo ciclo de iniciação da tradução (Dever, 1999).

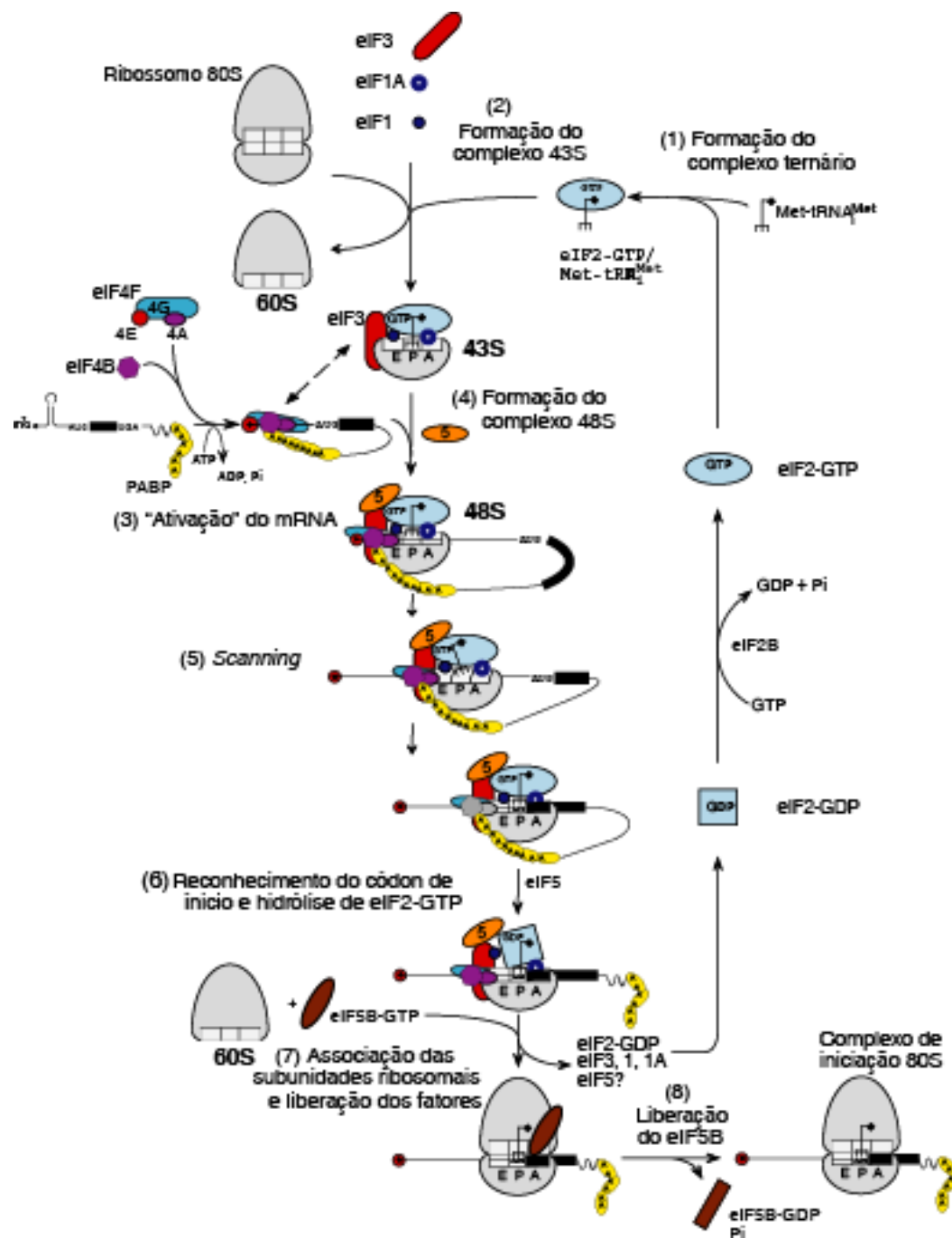


Figura 7. Representação esquemática da iniciação da tradução de mRNAs típicos em eucariotos. Este modelo divide o processo em oito diferentes estágios, mostrando os fatores de iniciação da tradução envolvidos em cada uma das etapas em que participam no processo: 1- Formação do complexo ternário; 2- Formação do complexo 43S; 3- "Ativação" do mRNA; 4- Formação do complexo 48S; 5- *Scanning*; 6- Reconhecimento de códon de início e hidrólise de eIF2-GTP; 7- Associação das subunidades ribossomais e liberação dos fatores; 8- Liberação do eIF5B. Neste momento, o ribossomo está pronto para prosseguir com a fase de elongação da síntese de proteínas (Pestova *et al.*, 2007).

2.3.2 – O complexo eIF4F de iniciação da tradução

O complexo eIF4F é o principal responsável pelo recrutamento da subunidade 40S ribossomal à extremidade 5' dos mRNAs. Este complexo é formado por três proteínas: eIF4E, (a proteína de ligação ao cap), eIF4A (uma RNA helicase e ATPase que desfaz estruturas secundárias no mRNA) e a proteína eIF4G com múltiplos sítios de ligação para parceiros como o eIF4E, eIF4A, eIF3 (que diretamente se liga a subunidade ribossomal 40S) e a PABP (Gingras *et al.*, 1999).

2.3.2.1 - eIF4E, a proteína de ligação ao cap

O eIF4E é a proteína responsável pelo reconhecimento do cap, presente nos mRNAs eucarióticos, dando início ao processo de iniciação da tradução (Sonenberg, 2008). A família eIF4E de proteínas pode ser encontrada numa grande diversidade de organismos desde protozoários a humanos, normalmente com múltiplos homólogos por espécie, sendo pelo menos um dos homólogos envolvido com tradução e os demais relacionados a outras funções celulares (Joshi *et al.*, 2005; Rhoads, 2009).

A estrutura do eIF4E, ligada ao cap, se assemelha a uma luva de *baseball* com o cap (m^7GTP) ligado à superfície côncava (Figura 7). O reconhecimento específico acontece pelo empacotamento da guanosina metilada entre os anéis aromáticos de dois triptófanos conservados (W56 e W102 em camundongo) num processo conhecido como empacotamento π - π (π - π *stacking*). Em seguida, ocorre uma estabilização mediada pelo radical metil do cap que aumenta de quatro a seis vezes a afinidade de interação com o eIF4E (Matsuo *et al.*, 1997; Marcotrigiano *et al.*, 1997).

Os parceiros mais importantes do eIF4E são os fatores eIF4G e as proteínas 4EBPs (*4E binding proteins* – proteínas de ligação ao eIF4E) que se associam ao eIF4E através de

sua região convexa, oposta ao sítio de ligação ao cap. Alguns resíduos de aminoácidos encontrados no eIF4E humano são importantes para ligação ao eIF4G e 4E-BPs: H37, P38, V69, W73, L131, E132, e L135. Destes os aminoácidos V69 e W73 são encontrados dentro de uma sequência bastante conservada nos eIF4Es, (S/T)VxxFW (Marcotrigiano *et al.*, 1997; Joshi *et al.*, 2005), e a substituição do W73 por um aminoácido não-aromático abole a interação do eIF4E com eIF4G e 4E-BPs (Pyronnet *et al.*, 1999; Ptushkina *et al.*, 1999).

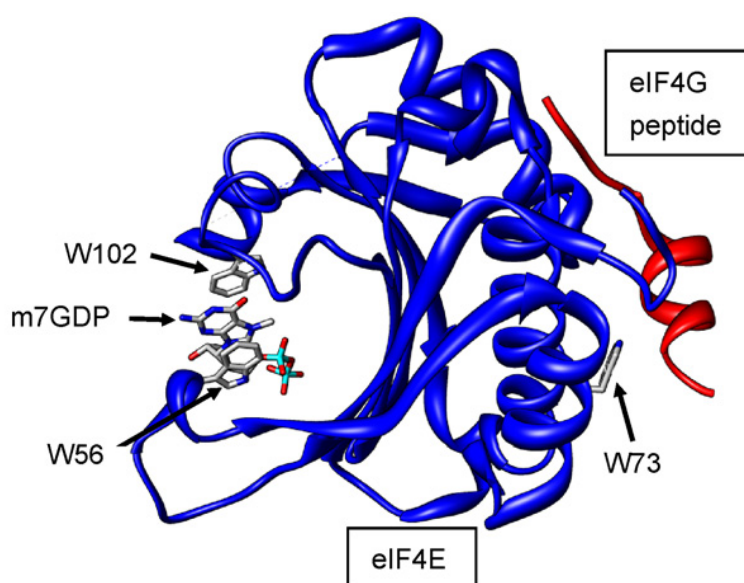


Figura 8: Estrutura do eIF4E de camundongo. Estrutura tridimensional do eIF4E de camundongo ressaltando os aminoácidos W56 e W102 envolvidos na ligação com ao cap, e o aminoácido W73 envolvido na interação com o eIF4G na superfície convexa. (Matsuo *et al.*, 1997; Goodfellow e Roberts, 2008).

A atividade da proteína eIF4E pode ser regulada, em mamíferos, pela fosforilação da Serina 209 (S209) pelas quinases Mnk1 e Mnk2 (Gingras *et al.*, 1999). As Mnks interagem com o eIF4G, através de seu domínio C-terminal, que funciona como sítio de ancoragem para as quinases ficarem próximas ao seu substrato principal, o eIF4E. Mutações no motivo de ligação ao eIF4E presente no eIF4G, ou na região de interação ao eIF4G localizado na proteína Mnk1, resulta na diminuição da intensidade de fosforilação do eIF4E. Portanto o

eIF4G parece modular a fosforilação do eIF4E (Gingras *et al.*, 1999; Pyronnet *et al.*, 1999; Parra-Palau *et al.*, 2003). A fosforilação do eIF4E aumenta a sua afinidade pelo cap (Scheper and Proud, 2002), mas não resulta em aumento da atividade traducional (Goodfellow and Roberts, 2008). O tratamento de células com fatores de crescimento, hormônios e mitógenos, estimula a fosforilação do eIF4E de mamíferos (Gingras *et al.*, 1999), mas a consequência fisiológica desse evento ainda não está bem esclarecida (Goodfellow e Roberts, 2008). Um segundo mecanismo de regulação da atividade do eIF4E é mediado pelas proteínas 4EBPs que competem com o eIF4G, pela interação com o eIF4E, e assim funcionam como inibidores da tradução. Em mamíferos foram identificadas três 4EBPs e suas funções são reguladas por fosforilação (Gingras *et al.*, 1999). Quando hipofosforiladas, as 4EBPs sequestram o eIF4E inibindo a tradução, enquanto a hiperfosforilação impede sua ligação ao eIF4E e esse, livre, pode ser utilizado no processo de iniciação da tradução.

2.3.2.2 - A RNA helicase eIF4A

O fator de iniciação da tradução eIF4A é considerado um protótipo da família de proteínas DEAD-box, um grupo de RNA helicases caracterizado por apresentar nove motivos (I, Ia, Ib, II, III, IV, V, VI e Q), dos quais o mais conservado é o domínio II conhecido como DEAD (L-D-E-A-D-X-X-L) (Linder, 2003). As RNAs helicases são encontradas tanto em organismos mais simples (procariotos), como até nos mais complexos (eucariotos superiores). Elas estão envolvidas em vários processos celulares que incluem RNA, tais como: transcrição, processamento de mRNAs, síntese de ribossomos, transporte e degradação de mRNAs, tradução (Luking *et al.*, 1998; Linder, 2003). O eIF4A é uma RNA helicase, dependente de ATP, que remove estruturas secundárias da região 5'-UTR dos mRNAs. Na iniciação da tradução sua atividade facilita tanto a ligação do complexo 43S ao mRNA como o

deslocamento da subunidade menor ribossomal, ao longo da 5'UTR (varredura ou *scanning*), até a identificação do códon de iniciação da tradução (Gingras *et al.*, 1999).

Em mamíferos três isoformas de eIF4A podem ser encontradas: eIF4AI, eIF4AII e eIF4AIII (Weinstein *et al.*, 1997). O eIF4AI e eIF4AII, em humanos, apresentam uma alta homologia entre si (90% de identidade) e são funcionalmente equivalentes na tradução (Nielsen e Trachsel, 1988; Yoder-Hill *et al.*, 1993; Gingras *et al.*, 1999). Em contraste, o grau de identidade entre eIF4AI e eIF4AIII é relativamente menor (66% de identidade), e são consideradas funcionalmente distintas. O eIF4AIII não pode substituir o eIF4AI na ligação da subunidade menor ribossomal ao mRNA, inibe a tradução *in vitro* (Li *et al.*, 1999) e sua localização é nuclear (Holzmann *et al.*, 2000), atuando como ancoradouro para formação do Complexo de Junção entre Éxons (EJC). O EJC é um complexo de proteínas que se forma numa região de 20 a 24 nucleotídeos 5' das junções entre os éxons, nos mRNAs processados por *splicing* (Le *et al.*, 2000). O EJC participa de eventos centrais no metabolismo dos mRNAs como exportação do núcleo para o citoplasma, degradação de mRNAs terminados precocemente ou processados de uma forma errônea e tradução citoplasmática (Le *et al.*, 2003; Tange *et al.*, 2004; Lejeune e Maquat, 2005). Como parte do EJC, o eIF4AIII liga-se diretamente às proteínas Magoh, Y14 e MLN51 (Ballut *et al.*, 2005). Mutações nos seus motivos Ia, VI e H afetam a formação do EJC e impedem que a via de degradação de mRNA ocorra (Shibuya *et al.*, 2006). Recentemente foi ainda demonstrado também que o EJC está relacionado com a regulação da tradução (Le e Seraphin, 2008).

A estrutura da molécula do eIF4A se assemelha a um haltere, onde os domínios N-terminal (DNT) e C-terminal (DCT) são interligados por uma sequência flexível (Caruthers *et al.*, 2000). Ambos os domínios estão envolvidos na ligação de RNA e ATP. O eIF4A, como outras RNA helicases, alterna sua estrutura entre dois estados: “fechado” e “aberto”. No estado “fechado”, conformação ativa, os dois domínios formam uma superfície que permite a ligação do eIF4A à RNA e ATP (Andersen *et al.*, 2006)(Figura 8) de uma forma cooperativa entre os

dois domínios (Lorsch e Herschlag, 1998). A atividade de helicase do eIF4A livre é baixa, mas é estimulada quando a proteína se liga ao eIF4G (dentro do complexo eIF4F) e quando eIF4B ou eIF4H estão associados (Marintchev e Wagner, 2004; Marintchev *et al.*, 2009). Recentemente foi demonstrado que ambos os domínios do eIF4A interagem com o eIF4G e que o domínio C-terminal interage com o eIF4H. O conjunto de interações fracas multi-domínio, porém específicas, das proteínas eIF4G e eIF4H com o eIF4A na presença de ATP ou ADP, é responsável pela atividade helicase desempenhada pelo eIF4A mesma durante a iniciação da tradução (Marintchev *et al.*, 2009). Adicionalmente, em plantas foi demonstrado que o eIF4B, além de estimular a atividade helicase do eIF4A, parece também ter efeitos na seleção de mRNAs durante a iniciação da tradução (Mayberry *et al.*, 2009).

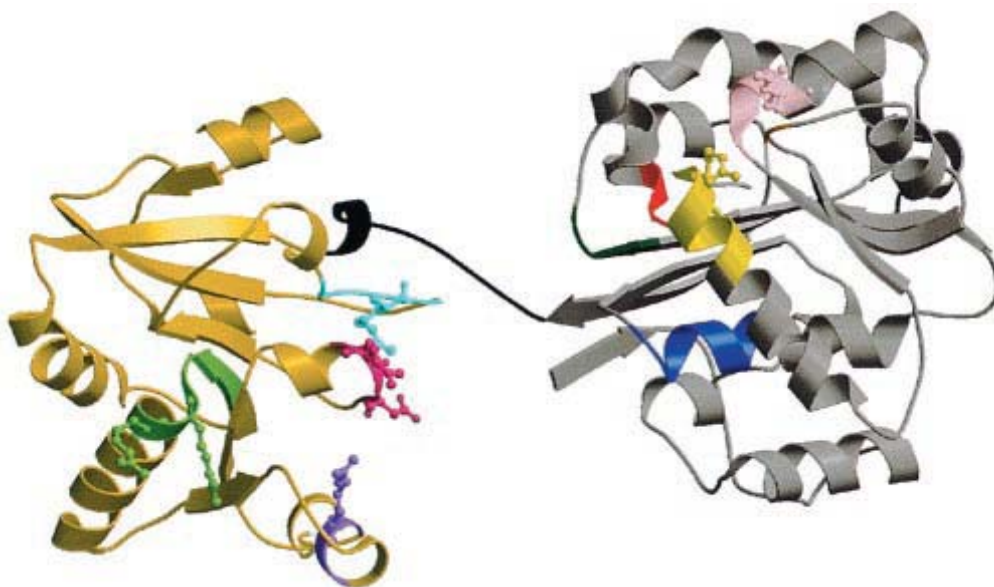


Figura 9: Estrutura tridimensional do eIF4A de levedura. A molécula do eIF4A apresenta os domínios N-terminal (DNT) e C-terminal (DCT), coloridos respectivamente de cinza e amarelo e interligados por uma sequência flexível de onze aminoácidos. Os motivos no DNT estão coloridos como se segue: motivo I, motivo Walker A ASQSGTGKT resíduos 65-72, azul; motivo Ia, PTRELA, resíduos 97-102, amarelo; GG, resíduos 125-126, laranja; TPGR, resíduos 145-148, rosa; motivo II, motivo Walker B DEAD, resíduos 169-172, vermelho; motivo III, SAT, resíduos 200-202, verde. No DCT os motivos estão coloridos da seguinte forma: motivo IV, VIFCNTRR, resíduos 263-270, verde motivo IV, conservado motivo R, resíduo R298, púrpura; motivo V, RGID, resíduos 321-324, magenta; motivo VI, HRIGRGGR, resíduos 345-352, azul ciano (Caruthers *et al.*, 2000).

2.3.2.3 - eIF4G, uma proteína modular que estrutura o complexo eIF4F

O fator de iniciação da tradução eIF4G é uma proteína de alto peso molecular cujo papel crucial é coordenar a formação do complexo de ligação ao cap (eIF4F) para, junto com a PABP, recrutar a subunidade menor ribossomal. Neste processo o eIF4G interage com eIF3, eIF4A, eIF4E, PABP, com a própria molécula de RNA, entre outros (Figura 9). O protótipo eIF4G humano, proteína de aproximadamente 1600 aminoácidos, se divide em três regiões (N-terminal, central e C-terminal) de acordo com um padrão de clivagem utilizando proteases virais (Lamphear *et al.*, 1995). O segmento N-terminal inclui sítios de ligação para o PABP e eIF4E (Prevot *et al.*, 2003). A região central é crítica para formação do complexo de iniciação da tradução, visto que esse segmento possui sítios para ligação ao eIF3 e eIF4A (Imataka e Sonenberg, 1997; Marcotrigiano *et al.*, 2001). Já a região C-terminal possui um segundo sítio para ligação ao eIF4A e um sítio para ligação à Mnk1, proteína que fosforila eIF4E (Prevot *et al.*, 2003).

2.3.2.3.1-Estrutura do eIF4G

As estruturas tridimensionais dos principais domínios do eIF4G, de mamífero, foram resolvidas por cristalografia de raio-X. Entretanto, ainda não foi possível obter a estrutura de grandes segmentos do eIF4G, não existindo portanto o modelo cristalográfico da proteína completa (Bellolell *et al.*, 2006). A região de ligação ao eIF4E, dentro da região N-terminal, modifica-se da forma não-estruturada para a forma estruturada quando ligada ao eIF4E (Marcotrigiano *et al.*, 1999; Gross *et al.*, 2003). A região central é constituída por arranjos de cinco pares de repetições HEAT (HRs), de alfa-hélices antiparalelas, denominado de domínio HEAT (Marcotrigiano *et al.*, 2001). HRs tem seu nome derivado das primeiras quatro proteínas nas quais este tipo de domínio foi encontrado, **H**untingtina – proteína envolvida na doença neurológica de Huntington, fator de **E**longação da tradução 3 (EF-3), subunidade **A** da

proteína Fosfatase 2 (PP2A) e a proteína sensível a rapamicina mTor/FRAP envolvidas em processos celulares como transdução de sinal e ciclo celular (Andrade e Bork, 1995; Gingras *et al.*, 1999; Hong e Sarge, 1999; Chakraborty, 2001). O domínio HEAT é geralmente encontrado em proteínas formadoras de complexos (Andrade e Bork, 1995). Recentemente foi demonstrado que a estrutura do eIF4G humano na realidade é composta por três domínios HEATs, o primeiro localizado na região central é denominado de HEAT-1/MIF4G, e os outros dois presentes na região C-terminal são HEAT-2/MA3 e HEAT-3/W2. O HEAT-2 e o HEAT-3 são formados por 5 e 4 repetições respectivamente, sendo essas repetições atípicas quando comparadas ao HEAT-1 (Bellolell *et al.*, 2006).

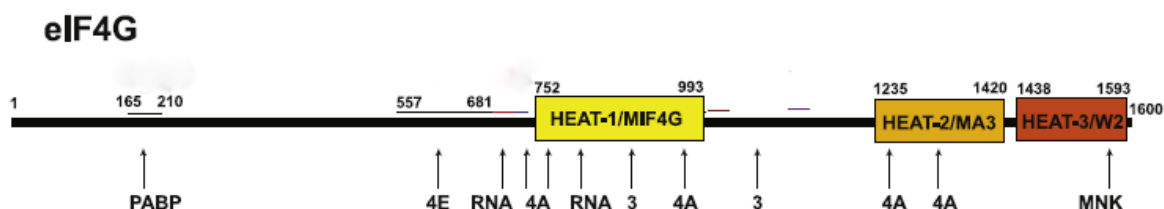


Figura 10: Esquema do eIF4GI humano e seus principais sítios de ligação para outras proteínas. O eIF4G possui na sua região N-terminal os sítios de ligação à PABP e eIF4E. Na região central desta proteína encontra-se o domínio central de ligação ao eIF4A, além de uma região de ligação a RNA (RRM) e do sítio de ligação ao eIF3. Já na sua porção C-terminal se encontram um segundo sítio de ligação ao eIF4A (apenas em mamíferos) e o sítio de ligação da proteína Mnk1 (quinase de eIF4E). O eIF4G humano, proteína protótipo da família, é formada possui três domínios HEAT: HEAT-1/MIF4G, HEAT-2/MA3 e HEAT-3/W2 (Marintchev *et al.*, 2009).

2.3.2.3.2 - Família eIF4G

O fator eIF4G é encontrado em diferentes espécies de eucariotos, desde humanos (Yan *et al.*, 1992; Gradi *et al.*, 1998), leveduras (Goyer *et al.*, 1993) e plantas (Browning *et al.*, 1992) até em protozoários (Dhalia *et al.*, 2005). Os fatores eIF4G, envolvidos com a iniciação da tradução, estão presentes sob duas isoformas em mamíferos, leveduras e plantas. Em leveduras e mamíferos há duas isoformas de eIF4G que exibem moderada conservação de sequência entre si (53% de identidade em levedura e 46% em humanos). Aparentemente ambas isoformas (eIF4GI e eIF4GII) desempenham a mesma função, considerando que uma pode substituir a outra em diferentes ensaios de tradução (Prevot *et al.*, 2003). Entretanto, foi observado que durante a diferenciação do megacariócito o eIF4GII é seletivamente recrutado para a formação do complexo eIF4F, indicando que estas proteínas podem estar relacionadas a diferentes situações (Caron *et al.*, 2004).

Diferentes isoformas do eIF4G podem co-existir na mesma célula, com expressão diferencial e ter implicações nos mecanismos que envolvem esta proteína. Nos mamíferos, foi demonstrado que diferentes isoformas podem ser geradas a partir do gene eIF4GI através da tradução diferencial do seu mRNA, a partir do uso de diferentes códons AUG de iniciação, e/ou pelo processamento alternativo de seu transcrito por *cis-splicing* (Byrd *et al.*, 2002; Bradley *et al.*, 2002; Byrd *et al.*, 2005). Desta forma, cinco isoformas do eIF4GI, de diferentes tamanhos com base no encurtamento da região N-terminal, são co-expressas em células de mamíferos (Coldwell e Morley, 2006). A isoforma eIF4GIf (proteína inteira) apesar de ser expressa em baixo níveis, é capaz de promover mais eficientemente a tradução de mRNAs celulares em relação às demais isoformas expressas em quantidades bem maiores (Coldwell e Morley, 2006).

Duas isoformas bem diferentes do eIF4G, conhecidas como eIF4G e eIFiso4G, são expressas em plantas (Browning *et al.*, 1992) e, diferentemente de mamíferos e leveduras,

ligam-se respectivamente a dois diferentes homólogos do eIF4E formando os complexos eIF4F (eIF4E/eIF4G) e eIFiso4F (eIFiso4E/eIFiso4G). eIF4G e eIFiso4G diferem em tamanho (165 e 86 kDa, respectivamente) e apresentam apenas 30% de identidade (Browning *et al.*, 1992; Gallie e Browning, 2001), enquanto as isoformas presentes em leveduras e mamíferos são mais conservadas (Goyer *et al.*, 1993; Gradi *et al.*, 1998). Apesar destas diferenças significativas, ambas possuem domínios para interação com o eIF4A, eIF3 e PABP (Browning, 1996). Entretanto, estes dois complexos apresentam diferenças funcionais, enquanto o eIF4F pode traduzir mRNAs contendo ou não cap, com estruturas secundárias ou codificando mais de uma proteína, o eIFiso4F atua na tradução de mRNAs com cap e sem estruturas secundárias na região 5' (Gallie e Browning, 2001). Recentemente foi mostrado que estes dois complexos estão envolvidos na tradução de mRNAs particulares, ou classes de mRNAs, sugerindo que a tradução diferencial pode requerer complexos compostos de isoformas específicas de fatores de iniciação da tradução (Mayberry *et al.*, 2009). Assim, mRNAs individuais ou classes de mRNAs podem responder à abundância relativa de um fator de iniciação da tradução específico afetando a quantidade final de proteínas traduzidas. É provável que complexos de iniciação, formados por proteínas específicas, permitam a tradução eficiente de mRNAs nas diferentes variedades de processos celulares.

A família eIF4G não se restringe apenas ao eIF4G típico, mediador da iniciação da tradução, mas por um conjunto de proteínas que apresentam homologia ao eIF4G por conter o domínio MIF4G e ou MA3 (Figura 10). Entre esse homólogos podemos citar: a proteína de ligação a PABP (Paip1) que possui 39% de similaridade com o MIF4G do eIF4G, interage com o eIF4A e estimula a tradução (Craig *et al.*, 1998) possivelmente através de uma interação com o eIF3 (Martineau *et al.*, 2008); a proteína humana p97 (Levy-Strumpf *et al.*, 1997; Imataka *et al.*, 1997; Yamanaka *et al.*, 1997), descrita como repressor traducional e que está relacionada a tradução de mRNAs contendo IRES (Henis-Korenblit *et al.*, 2002; Nevins *et al.*, 2003); a proteína nucleolar NOM1 (Simmons *et al.*, 2005), de papel funcional

desconhecido, mas que contém um domínio MIF4G e um domínio MA3, que interage com eIF4AI e com a proteína Fostase I direcionado-a para o nucléolo (Gunawardena *et al.*, 2008); a PDCD4 (proteína de morte celular programada 4 - *programmed cell death protein 4*), uma proteína supressora de tumor que inibe a tradução (Yang *et al.*, 2003; Yang *et al.*, 2004), através de seus dois domínios MA3 que competem com eIF4G e RNA pela ligação ao eIF4A (Suzuki *et al.*, 2008). Finalmente, uma outra proteína, cuja estrutura e domínios funcionais são semelhantes ao eIF4G, é a subunidade maior do complexo de ligação ao cap nuclear (CBP80). Juntamente com a subunidade CBP20, a CBP80 está envolvida no processamento e maturação dos mRNAs nascentes, assim como em um provável teste traducional destes mRNAs, no núcleo, antes de serem enviados para a tradução citoplasmática (Marintchev e Wagner, 2005).

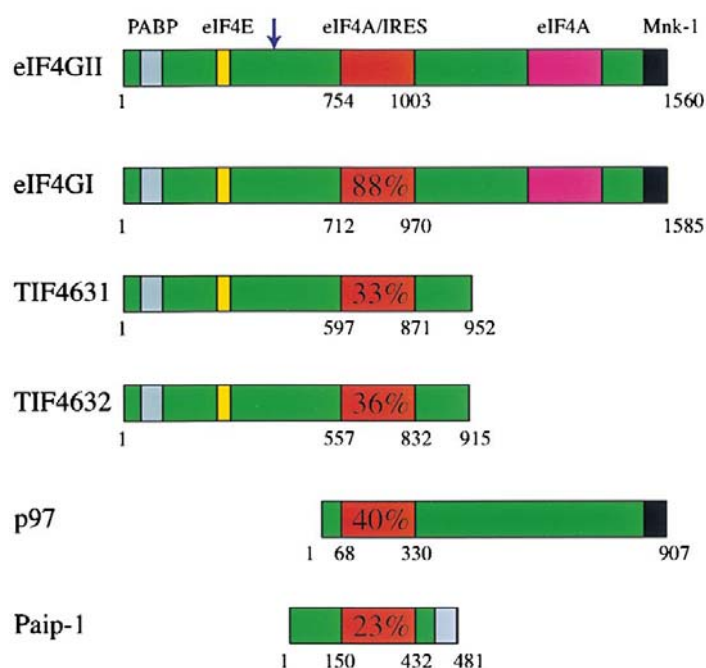


Figura 11. Motivos em comum entre eIF4GI, eIF4GII e proteínas da família eIF4G. Na figura está representado, percentualmente, a identidade entre o eIF4GII humano e os demais eIF4Gs: eIF4GI; as duas cópias em levedura (TIF4631 e TIF4632); os homólogos humanos p97; Paip-1 (Marcotrigiano *et al.*, 2001).

2.3.2.3.3 - Modificações pós-traducionais do eIF4G

As proteínas eIF4G podem ter sua função modulada por reações de fosforilação. As fosforilações em resíduos de aminoácidos específicos, presentes na região C-terminal do eIF4GI, podem ocorrer em resposta aos componentes do soro e mitógenos em células humanas. eIF4GII e p97 são também fosfoproteínas, entretanto as regiões C-terminal de ambas as proteínas não são fosforiladas (Raught *et al.*, 2000). A proteína eIF4GII é fosforilada durante a mitose (Pyronnet *et al.*, 2001), por uma quinase dependente de Calmodulina I na Ser 1156 (Qin *et al.*, 2003). A fosforilação do eIF4GII reduz sua afinidade com eIF4E, e somado a desfosforilação da 4EBP-1, promove o rompimento do complexo eIF4F e a supressão da tradução cap dependente (Pyronnet *et al.*, 2001). O eIF4G pode ainda ser fosforilado, entre outras proteínas, pela proteína quinase ativada por p21 (Pak2). A Pak2 é ativada por estresse inibindo a divisão celular, podendo ainda ser ativada no processo de apoptose promovendo a morte celular. Foi demonstrado que a Pak2 inibe diretamente a tradução em reticulócito de coelho e cultura de células 293T, fosforilando eIF4G e competindo com eIF4E pelo mesmo sítio de ligação ao eIF4G. A fosforilação do eIF4G, pela Pak2, está envolvida na regulação negativa da iniciação da tradução em resposta a condições de estresse celular (Ling *et al.*, 2005).

A atividade do eIF4G também pode ser regulada por proteólise, em resposta a processos apoptóticos ou infecção viral (Prevot *et al.*, 2003). A clivagem do eIF4GI e eIF4GII por proteases virais, como a protease 2A de picornavírus, que promove a separação da região envolvida na ligação ao eIF4E da região central responsável pela ligação ao ribossomo, abolindo assim a ligação do eIF4F ao cap (Gingras *et al.*, 1999). Como consequência ocorre uma severa e seletiva inibição da síntese de proteínas, e o fragmento C-terminal do eIF4G funciona em um complexo de pré-iniciação modificado que permite a tradução de mRNAs virais que não possuem cap e mRNAs celulares contendo IRES (Prevot *et al.*, 2003).

2.3.2.3.4 - Bases moleculares da interação eIF4G-eIF4A

Os dois primeiros domínios HEATs do eIF4GI humano, HEAT-1/MIF4G e HEAT-2/MA3, mediam sua ligação ao eIF4A (Lamphear *et al.*, 1995; Imataka e Sonenberg, 1997). HEAT-1 estimula a atividade helicase do eIF4A, enquanto HEAT-2 tem um papel modulatório (Morino *et al.*, 2000; Marintchev e Wagner, 2004). O HEAT-1 interage principalmente o domínio DCT do eIF4A, e embora seja capaz de interagir apenas com este domínio, sua interação com o eIF4A completo é mais forte (Oberer *et al.*, 2005).

Vários estudos vem sendo realizados para avaliar como ocorre a ligação entre o eIF4G e eIF4A, assim com a moléculas acessórias como o eIF4H e os mecanismos de ligação a ATP e RNA. O domínio MIF4G/HEAT-1 do eIF4G liga-se a ambos os domínios eIF4A-DNT e ao eIF4A-DCT (Oberer *et al.*, 2005), enquanto e o domínio HEAT2/MA3 liga-se ao eIF4A-DCT. Adicionalmente, foi demonstrado que o eIF4A-DCT liga-se também a região entre os domínios HEAT-1 e HEAT-2 (Marintchev *et al.*, 2009). Por outro lado, foi verificado que não há sobreposição entre os sítios de ligação ao eIF4H e os domínios HEAT-1/HEAT-2, no eIF4A. O eIF4H liga-se ao eIF4A-DCT, entretanto a atividade de helicase do eIF4A é apenas estimulada pelo eIF4H apenas quando é utilizada a proteína eIF4A completa, indicando que de alguma forma o eIF4H associa-se também ao eIF4A-DNT (Feng *et al.*, 2005). Esses dados são consistentes com um recente trabalho que mostra que o eIF4A-DCT liga-se ao eIF4H na presença de RNA e AMP-PNP. Este modelo de interação parece ser aplicado ao eIF4B, homólogo do eIF4H, visto que eles se ligam ao mesmo sítio no eIF4A e podem estimular a atividade helicase do eIF4A (Rozovsky *et al.*, 2008).

As interações entre eIF4A com HEAT-1 e HEAT-2 são antikooperativas e mediadas pela conformação do eIF4A. O HEAT-1 interage com a conformação fechada, ativa, ligada ao ATP, enquanto o HEAT-2 interage com o estado aberto, não-ativo, livre do nucleotídeo. Assim, na ausência de ATP, o eIF4A liga-se ao HEAT-2 com uma afinidade maior que ao HEAT-1, favorecendo o estado aberto do eIF4A. ATP e RNA ligam-se kooperativamente ao

HEAT-1 e antiooperativamente ao HEAT-2. O eIF4A ligado ao ATP tem sua afinidade de ligação aumentada por eIF4H, na presença do HEAT-1 do eIF4G, corroborando a habilidade do eIF4H de estimular a atividade helicase do eIF4A (Rogers, Jr. *et al.*, 2002).

A Figura 12 ilustra o modelo proposto de ligação do eIF4A aos seus parceiros. Segundo esse modelo, a hidrólise de ATP e liberação de ADP reduz a afinidade do eIF4A pelo RNA, eIF4H e o domínio HEAT-1 do eIF4G, enquanto aumenta a afinidade de ligação ao domínio HEAT-2, conduzindo o eIF4A para seu estado aberto. Após a ligação de uma nova molécula de ATP, o eIF4A é novamente modificado para seu estado fechado. As mudanças de orientação dos domínios do eIF4A, durante o ciclo de ligação/hidrólise de ATP, provavelmente causam movimentos nos domínios do eIF4G e eIF4H que interagem com a helicase.

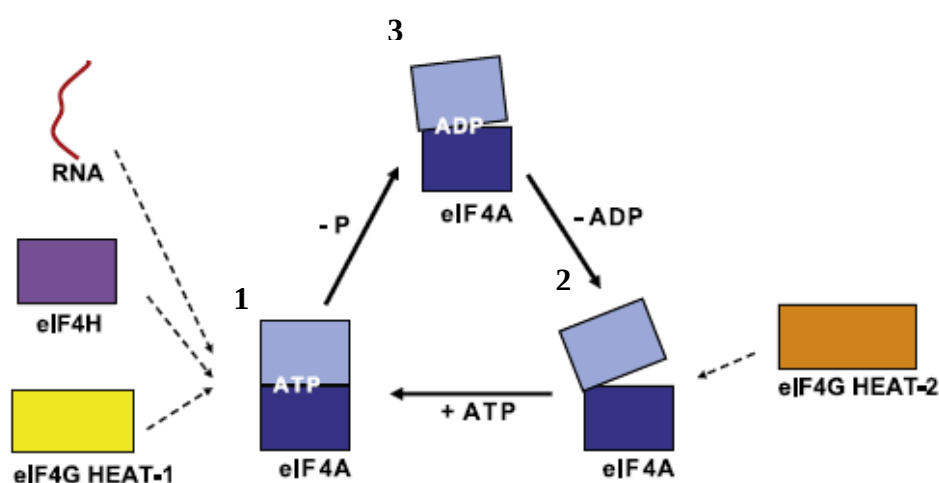


Figura 12. Interações do eIF4G, eIF4H e RNA com o eIF4A. RNA, eIF4H e HEAT1 (esquerda) ligam-se a conformação fechada do eIF4A na presença de ATP (1), enquanto HEAT-2 (direita) liga-se preferencialmente à conformação aberta do eIF4A (2). Intermediando estas duas formas o eIF4A está ligado a ADP, na transição entre os estados fechado e aberto (3) (Marintchev *et al.*, 2009).

A estequiometria da interação entre o eIF4G e o eIF4A no modelo humano tem gerado algumas controvérsias, sendo descrito na literatura como 1:2 assim como 1:1 (Li *et al.*, 2001; Korneeva *et al.*, 2001). O modelo acima proposto sugere que na realidade não são duas

moléculas de eIF4A que interagem com o eIF4G, mas que os domínios HEAT do eIF4G apresentam diferentes comportamentos de ligação (figura 11) corroborando a proporção de 1:1. A região central do eIF4G contendo o HEAT-1 (682-1138aa) liga-se ao eIF4A com um afinidade de 17 nM em humanos (Korneeva *et al.*, 2001), sendo em levedura esse valor estimado em 30 nM (Dominguez *et al.*, 1999). Estes valores provavelmente são diferentes no contexto da proteína eIF4G completa e na presença dos parceiros eIF4H/eIF4B, assim como dos substratos ATP e RNA (Marintchev *et al.*, 2009).

2.3.2.3.5 - Modo de ligação do eIF4E ao eIF4G

O domínio no eIF4G de ligação ao eIF4E foi mapeado a partir da observação que o eIF4E permanecia associado à região N-terminal do eIF4G, após clivagem com proteases virais (Lamphear *et al.*, 1995). Estes dados foram ainda corroborados através de ensaios de interação entre eIF4GI mutantes e o eIF4E, por imunoprecipitação e *far western blot* (Mader *et al.*, 1995). Neste trabalho foi caracterizada uma região de 49 aminoácidos, presente em mamíferos e leveduras. Dentro desta região o motivo YXXXXLΦ (onde X pode ser qualquer aminoácido e Φ é um resíduo hidrofóbico – L, M ou F) é altamente conservado nas proteínas eIF4G de vários organismos (Figura 12), e mutações pontuais em Y ou L abolem a interação entre o eIF4G e o eIF4E (Mader *et al.*, 1995). Estudos estruturais mostram que este motivo assume a estrutura de uma hélice curta que se associa à superfície dorsal, convexa do eIF4E (Marcotrigiano *et al.*, 1999). O motivo YX₄Φ também está presente em uma família de proteínas que competem com o eIF4GI pela interação com o eIF4E, as proteínas 4E-BPs, que são reguladas pelo seu estado de fosforilação. As 4E-BPs atuam como repressores da tradução, por afetarem a formação do complexo 48S (Richter e Sonenberg, 2005).

Em mamíferos, um peptídeo de 16 aminoácidos do eIF4GII, contendo o motivo YX₄Φ, liga-se a face dorsal do eIF4E com uma alta afinidade de 26±6 nM e a conformação do eIF4E não se altera quando ligado a este peptídeo (Marcotrigiano *et al.*, 1999). Em

contraste, em levedura o peptídeo do eIF4G correspondente liga-se ao eIF4E com baixa afinidade (2 μ M), embora um fragmento maior do eIF4G (393-490AA) ligue-se ao eIF4E com afinidade relativamente maior (3 η M) (Gross *et al.*, 2003). A ligação do eIF4G ao eIF4E causa uma mudança de conformação em ambas as proteínas, conduzindo a formação de um complexo estável. Neste complexo, o eIF4G envolve o N-terminal do eIF4E, formando um bracelete molecular (Gross *et al.*, 2003) e o complexo eIF4E-eIF4G de levedura aumenta a afinidade da interação do eIF4E pelo cap (Gross *et al.*, 2003). A utilização de pequenas moléculas que competem com o motivo YXXXXL Φ , pela ligação ao eIF4E, pode ter diversas aplicações como no desenvolvimento de drogas que impeçam a interação eIF4G/eIF4E. Muitos tipos de câncer apresentam super-expressão do eIF4E e eIF4G e a inibição da interação entre os dois fatores pode constituir um ótimo alvo para o tratamento deste grupo de doenças (Moerke *et al.*, 2007).

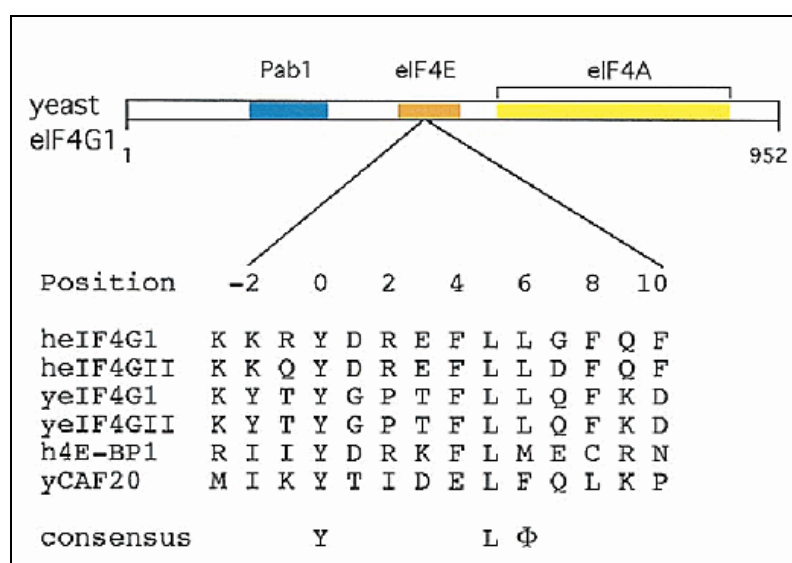


Figura 13. Esquema do eIF4G1 de levedura ressaltando o domínio de ligação ao eIF4E, também presente nas 4E-BPs. O domínio de ligação do eIF4G1 ao eIF4E é caracterizado pela sequência YXXXXL Φ , onde X pode ser qualquer aminoácido e Φ um resíduo hidrofóbico: normalmente L, M ou F (Sachs e Varani, 2000).

Diversos trabalhos tem mostrado que o eIF4G aumenta a afinidade do eIF4E pelo cap. Inicialmente, foi descrito que a indução da ligação química entre o eIF4E de mamíferos e o cap é mais eficiente quando o eIF4E faz parte do complexo eIF4F (Lee *et al.*, 1985) ou quando ele está complexado ao eIF4GI (Haghighat e Sonenberg, 1997). Um aumento similar da afinidade do eIF4E pelo cap, também foi verificado em leveduras (von der Haar *et al.*, 2000). Porém dois diferentes mecanismos foram propostos para explicar as duas situações. Em mamíferos é postulado que o segmento central do eIF4G, que liga-se ao RNA (RRM), estabiliza a ligação do eIF4E ao cap. Este modelo foi baseado na infecção de células com poliovírus, onde o eIF4GI é clivado liberando suas regiões N-terminal e central, e o eIF4E permanece associado ao fragmento N-terminal e nessa situação a afinidade de ligação do eIF4E ao cap é baixa (Pelletier e Sonenberg, 1985; Lee *et al.*, 1985). Entretanto, em levedura, uma forte interação entre cap e eIF4E é encontrada usando um fragmento de eIF4G contendo o sítio de ligação ao eIF4E na ausência de RRM (Ptushkina *et al.*, 1998; von der Haar *et al.*, 2000). Provavelmente estas diferenças, no modo de ligação eIF4G/eIF4E, refletem características específicas dos modelos avaliados. Recentemente, em mamíferos, foi mostrado que o eIF4G aumenta fortemente a interação do eIF4E pelo cap por mobilizar o eIF4E ao mRNA via RRM do eIF4G (Yanagiya *et al.*, 2009).

2.2.3.6 - O sítio de ligação para eIF3 no eIF4G

O fator eIF3 é o maior fator de iniciação da tradução em mamíferos, sendo composto por 13 subunidades nomeadas de eIF3a-m. O principal papel do eIF3 é fazer a conexão entre a subunidade menor ribossomal e o mRNA, através da ligação específica entre sua subunidade eIF3e e o eIF4G (LeFebvre *et al.*, 2006). O sítio de ligação para o eIF3 foi mapeado no eIF4G humano, entre os domínios HEAT-1 e HEAT-2, compreendendo entre os aminoácidos 1015 e 1118 (Korneeva *et al.*, 2000). O domínio HEAT-1 do eIF4G também liga-se ao eIF3, uma vez

que mutações pontuais e deleções neste domínio afetam a ligação ao eIF3 (Imataka e Sonenberg, 1997; Morino *et al.*, 2000). Como recentemente foi demonstrado que o eIF4A liga-se na região entre os domínios HEAT-1 e HEAT-2, e próximo ao local onde eIF3 se liga, há múltiplas oportunidades para acoplamento das interações eIF4G/eIF4A e eIF4G/eIF3. Apesar de nenhuma ligação direta ter sido ainda descrita entre eIF3 e eIF4A, a proximidade de seus sítios de interação no eIF4G sugere que eles poderiam contactar-se entre si, ao menos fracamente, no contexto de um complexo ternário eIF4A/eIF4G/eIF3 (Marintchev *et al.*, 2009).

2.3.3 - A proteína de ligação à cauda poli-A (PABP), parceira do eIF4F

A PABP liga-se especificamente à cauda poli-A presente na região 3' dos mRNAs eucarióticos, e participa de vários processos celulares em eucariotos relacionados ao mRNA tais como biogênese, processamento, transporte, degradação e tradução (Chekanova e Belostotsky, 2003; Chang *et al.*, 2004; Dunn *et al.*, 2005; Brune *et al.*, 2005; Woods *et al.*, 2005; Hosoda *et al.*, 2006). Inicialmente, dados bioquímicos indicaram que a adição do cap e da cauda poli-A, aos mRNAs, resultavam em um efeito sinérgico de estimulação da tradução (Gallie, 1991). Este efeito foi observado em *Saccharomyces cerevisiae* como consequência da interação entre PABP e o eIF4G, promovendo a ligação entre os terminais 5' e 3' do mRNA (Tarun, Jr. e Sachs, 1996). O domínio de interação foi mapeado numa região de 114 aminoácidos da porção N-terminal do eIF4G de levedura, que também media a ligação da PABP ao poli-A *in vitro* (Tarun, Jr. *et al.*, 1997). Assim foi proposto um modelo onde a interação PABP-eIF4G promoveria a circularização do mRNA, durante a tradução (Tarun, Jr. e Sachs, 1996). Este modelo foi formalmente demonstrado através da visualização, por microscopia de força atômica, da formação de complexos eIF4E/eIF4G/PABP (utilizando proteínas recombinantes) em mRNAs capeados e poliadenilados (Wells *et al.*, 1998). A

interação direta entre PABP e eIF4G foi também demonstrada em plantas (Le *et al.*, 1997) e em mamíferos (Imataka *et al.*, 1998). O principal sítio de interação com a PABP, em mamíferos, é localizado dentro de uma região mínima de 29 aminoácidos na parte N-terminal do eIF4GI que é conservada no eIF4GII, mas não exibe nenhuma homologia de sequência com as duas isoformas de eIF4G de levedura, exceto por um rico conteúdo de aminoácidos básicos (Tarun, Jr. *et al.*, 1997; Imataka *et al.*, 1998).

A estrutura da PABP é altamente conservada em diferentes organismos (Kuhn e Wahle, 2004). A PABP possui uma região N-terminal que apresenta quatro RRMs (1-4), uma região rica em prolina/glutamina e uma região C-terminal contendo um domínio chamado PABC. Os dois primeiros RRMs são suficientes para a ligação específica ao poli-A (Kuhn e Pieler, 1996) e tem sido relacionados à interação entre PABP e eIF4G (Kessler e Sachs, 1998; Imataka *et al.*, 1998; Cheng e Gallie, 2007). A interação PABP/eIF4G parece reforçar a iniciação da tradução e a estabilidade do mRNA, permitindo uma verificação da integridade do mRNA que precederia a sua tradução (Svitkin e Sonenberg, 2006; Amrani *et al.*, 2008). Já os RRMs 3 e 4 não tem especificidade por poli-A e podem mediar o reconhecimento de sequências não poli-A (Sladic *et al.*, 2004; Khanam *et al.*, 2007).

A região C-terminal da PABP não se liga ao RNA, e contém um domínio conservado denominado PABC. Este domínio parece ser responsável pela ligação da PABP ao fator de terminação da tradução eRF3 e, às proteínas parceiras Paip1 e Paip2 (Kozlov *et al.*, 2001; Albrecht e Lengauer, 2004). A Paip1 é uma proteína que liga-se à PABP e estimula a tradução provavelmente por ligar-se também a eIF3, estabilizando a ligação PABP-eIF4G, tornando mais eficiente a circularização dos mRNAs (Craig *et al.*, 1998; Martineau *et al.*, 2008). Já o impedimento da circularização dos mRNAs, pela ligação de proteínas inibitórias como a Paip2 à PABP, parece constituir um mecanismo de regulação negativa da tradução (Khaleghpour *et al.*, 2001). Estas proteínas regulatórias possuem dois sítios distintos, dos quais um deles se liga ao domínio PABC (Kozlov *et al.*, 2001). A interação do eRF3 com a

PABC resulta na circularização da extremidade 3' UTR dos mRNAs, facilitando a reciclagem de ribossomos da extremidade 3' para a 5' de iniciação (Uchida *et al.*, 2002; Sonenberg e Dever, 2003; Hosoda *et al.*, 2003). De fato já foi demonstrado que a interação entre eRF3 e PABP aumenta a eficiência da tradução em mamíferos (Uchida *et al.*, 2002). Conectando os RRM's ao domínio PABC, existe uma região rica em prolina e glutamina na PABP, que é importante para a multimerização de moléculas de PABP na cauda poli-A (Melo *et al.*, 2003).

2.4 - Complexo eIF4F e PABP nos tripanossomatídeos

Apesar do grande conhecimento acumulado nos últimos anos com relação ao processo de tradução e, notadamente, à iniciação da tradução em eucariotos, as informações obtidas até o momento em tripanossomatídeos são escassas (Clayton e Shapira, 2007). Os dados obtidos, a partir de seqüências de projetos genomas de tripanossomatídeos (El-Sayed *et al.*, 2005a), tem sido o ponto de partida para análises mais refinadas em relação à busca da função de proteínas homólogas, a fatores de iniciação da tradução, já estudadas em outros eucariotos. Uma vez que a regulação da expressão gênica nestes parasitas é predominantemente pós-transcricional, e que a tradução pode ter um papel crítico neste processo, alguns trabalhos vêm se detendo na caracterização funcional de homólogos das subunidades do eIF4F e da PABP no controle da expressão gênica em tripanossomatídeos.

2.4.1 - PABP nos tripanossomatídeos

Homólogos da PABP em *T. cruzi*, *T. brucei* e *L. major* já foram descritos há algum tempo (Batista *et al.*, 1994; Pitula *et al.*, 1998; Bates *et al.*, 2000). O primeiro homólogo identificado foi a PABP1 de *T. cruzi*, codificada por dois genes adjacentes, e constitutivamente expressa durante o ciclo de vida do parasita (Batista *et al.*, 1994). O

ortólogo em *T. brucei*, também é codificado por dois genes, apresenta 86.4% de identidade à proteína de *T. cruzi* (Hotchkiss *et al.*, 1999). A proteína PABP1 de *L. major*, designada aqui *LmPABP1*, é codificada por apenas um gene e também constitutivamente expressa durante o ciclo de vida do parasita. Ela é representada por duas isoformas, uma das quais fosforilada, com predominância de localização citoplasmática. Ensaio de complementação demonstraram que *LmPABP1* não é capaz de substituir a letalidade da deleção da *pab1*, PABP de levedura, em *S. cerevisiae*. Surpreendentemente *LmPABP1* é apenas 35% idêntica aos homólogos de PABP descritas nos tripanossomas, um nível de identidade comparado ao observado entre a proteína de *Leishmania* e proteínas humanas (36% de identidade) (Bates *et al.*, 2000). Análises de bioinformática confirmaram que tripanosomatídeos apresentam dois homólogos de PABP (PABP1-2), e que *Leishmania major* apresenta um terceiro homólogo (*LmPABP3*) adicional (Costa Lima, Dissertação de Mestrado).

No tripanossomatídeo *Crithidia fasciculata*, o ortólogo da PABP2 faz parte de um complexo que reconhece seqüências presentes nas regiões não-traduzíveis de mRNAs que acumulam-se, periodicamente, durante o ciclo celular (Mittra e Ray, 2004). O ortólogo da PABP2 de *T. cruzi* interage diretamente com a *TcUPB-1*, uma proteína que liga-se na 3'-UTR de mRNAs ricos em ARES que são seletivamente desestabilizados durante o estágio tripomastigota do parasito (D'Orso e Frasch, 2002). As novas PABPs encontradas em *L. major* (*LmPABP2-3*) apresentam ainda diferenças funcionais em relação a *LmPABP1*, o homólogo funcional da PABP de mamíferos e leveduras (Costa lima, manuscrito em preparação).

2.4.2 - eIF4AI e eIF4AIII nos tripanossomatídeos

Um homólogo ao eIF4A, denominado de LeIF, foi descrito inicialmente em *L. brasiliensis* e *L. major*. O LeIF é constitutivamente expresso nos estágios parasitários do inseto e de mamífero e embora tenha sido considerado imunologicamente relevante, seu papel na tradução não foi investigado (Skeiky *et al.*, 1998). Além do LeIF (também denominado de

LmEIF4AI) foi identificado um segundo homólogo ao eIF4A, o *LmEIF4AIII*. (Dhalia *et al.*, 2005). Os dois homólogos do eIF4A, *LmEIF4AI* e *LmEIF4AIII*, apresentam uma identidade de 59 e 52% respectivamente em relação ao eIF4AI humano e possuem ortólogos em *T. brucei*. Estas proteínas variam quanto sua abundância na forma promastigota: *LmEIF4AI* é muito abundante com níveis acima de 10^5 moléculas/célula enquanto *LmEIF4AIII* é ausente ou está presente em níveis abaixo de 10^4 moléculas/célula. Em outro trabalho, a proteína *LmEIF4AI* mostrou ter função ATPase dependente de RNA *in vitro*, e apesar da alta homologia em eucariotos, esta proteína não pode complementar a função do eIF4A de levedura *in vivo* e de mamífero *in vitro*. Entretanto, o *LmEIF4AI* inibe o crescimento celular em levedura, por provavelmente seqüestrar proteínas envolvidas com a tradução, sendo demonstrado sua interação direta com o eIF4G de levedura (Barhoumi *et al.*, 2006). Os dois ortólogos do *LmEIF4AI* e *LmEIF4AIII* foram estudados em *T. brucei*, *TbEIF4AI* e *TbEIF4AIII*, e apenas o primeiro parece estar envolvido com a formação do complexo eIF4F da iniciação da tradução. Enquanto *TbEIF4AI* apresenta localização citoplasmática, o *TbEIF4AIII* apresenta-se concentrado no núcleo. Diante destes dados, especula-se que o *TbEIF4AIII* faça parte de algum mecanismo de degradação de mRNA, sendo funcionalmente equivalente ao eIF4AIII humano (Dhalia *et al.*, 2006). Recentemente foram isolados dois componentes funcionais (Mago e Y14) do Complexo de Junção entre Éxons (EJC), envolvido no mecanismo de degradação de mRNA prematuro em humanos, nos tripanossomatídeos. Surpreendentemente, o homólogo ao eIF4AIII não foi detectado junto ao complexo, necessitando estudos mais aprofundados (Bercovich *et al.*, 2009b).

2.4.3 - Isoformas do eIF4E em *Leishmania major*

Com relação aos homólogos do eIF4E, quatro proteínas têm sido investigadas em tripanossomatídeos (Yoffe *et al.*, 2004; Dhalia *et al.*, 2005; Yoffe *et al.*, 2006). Nenhum dos

quatro homólogos (*LmEIF4E1-4*), entretanto, é capaz de complementar a função do eIF4E de levedura (Yoffe *et al.*, 2006), sendo que o *LmEIF4E1* também não pôde complementar a função do eIF4E de mamíferos (Yoffe *et al.*, 2004). *LmEIF4E1* e *LmEIF4E4* ligam ao cap monometilado clássico (Yoffe *et al.*, 2004; Dhalia *et al.*, 2005) e ao cap-4 (Yoffe *et al.*, 2004), enquanto que o *LmEIF4E2* liga principalmente ao cap-4 e o *LmEIF4E3* fracamente ao cap (Yoffe *et al.*, 2004). Recentemente a partir de extratos de *L.major*, utilizando-se a resina m7 GTP Sefarose, foi purificado um possível complexo eIF4F contendo as proteínas *LmEIF4E1* e *E4* associadas ao *LmEIF4G3* e ao *LmEIF4A1* (Yoffe *et al.*, 2009).

2.4.4 - Candidatos a eIF4G em *Leishmania major*

A subunidade eIF4G do eIF4F é a proteína menos conservada ao nível de sequência, sendo a região central conservado em diferentes sequências de eIF4G de diferentes organismos (Gingras *et al.*, 1999). Usando a proteína humana eIF4G1 como parâmetro de busca com o auxílio da ferramenta BLAST em banco de dados do genoma de *L. major* foram encontrados cinco proteínas contendo o domínio central do eIF4G (*LmEIF4G1-5*). As cinco proteínas variam significativamente em tamanho (variando de 1016 a 1425, 635, 765 e 782 aminoácidos para o *LmEIF4G1-5*), carga geral (positiva no *LmEIF4G1* e negativo nas outras quatro) e na relativa localização do domínio conservado MIF4G (central no *LmEIF4G1*, 2, 5 e N-terminal no *LmEIF4G3-4*). Ambos os genes *LmEIF4G1* e *LmEIF4G2* são localizados no cromossomo 15 enquanto os genes *LmEIF4G3-5* estão localizados nos cromossomos 16, 36 e 10 respectivamente (Figura 14; Tabela I). Nenhum domínio óbvio para ligação a eIF3, eIF4E ou PABP foi identificado baseado apenas na homologia. Exceto pelo domínio HEAT-1/MIF4G, as sequências parecem não relacionadas, com exceção dos *LmEIF4G3* e 4. Estes duas proteínas mostram uma região N-terminal similar e uma região conservada de aproximadamente 120 aminoácidos localizada a cerca de 180 aminoácidos a jusante do

domínio HEAT, indicando que eles podem ser funcionalmente relacionados entre si e também aos domínios humanos do eIF4G, HEAT-2/MA3 e ou HEAT-3/W2 (Figuras 10; Figura 14) (Dhalia *et al.*, 2005). Os níveis intracelulares do *LmEIF4G3* foram determinados e seu domínio HEAT foi capaz de se ligar especificamente à proteína *LmEIF4AI* (Dhalia *et al.*, 2005). Recentemente foi mostrado que a incubação de extrato celulares de *L. major* a resina m⁷GTP Sefarose permitiu a purificação das proteínas *LmEIF4E1* e *LmEIF4E4*, assim como a co-eluição das proteínas *LmEIF4G3* e *LmEIF4A1*. Estas interações foram confirmadas a partir da imunoprecipitação da proteína *LmEIF4G3* em fusão com FLAG expressa em células transgênicas de *L. major*. As proteínas *LmEIF4G3*, *LmEIF4E1* e *LmEIF4E4* foram encontradas, ainda, em frações polissomais indicando para a formação de pelo menos um complexo eIF4F nos tripanossomatídeos (Yoffe *et al.*, 2009).

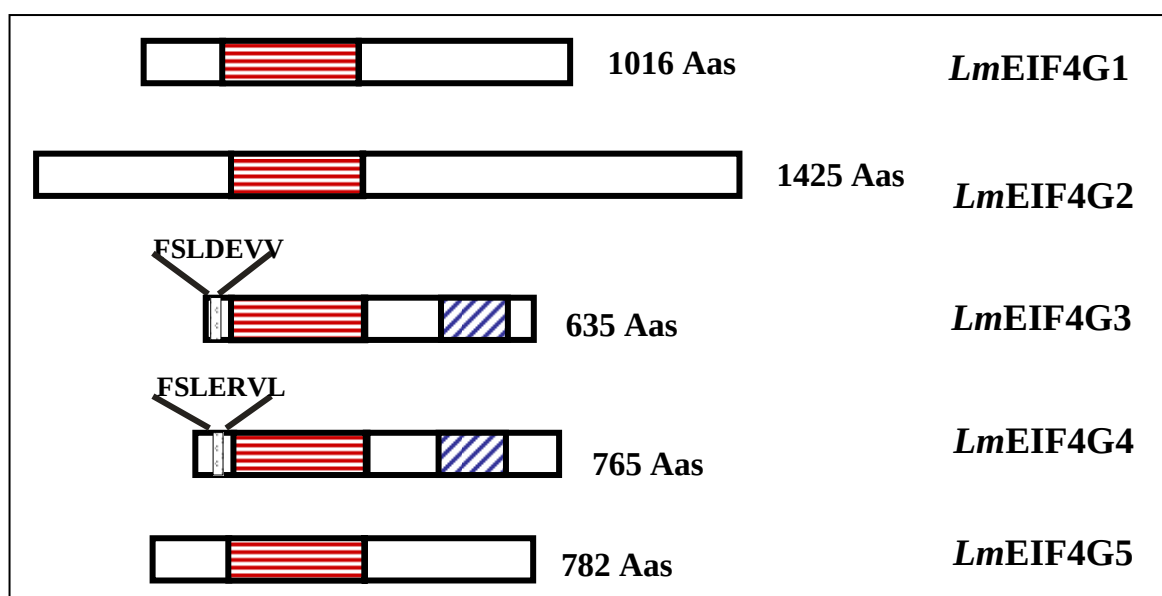


Figura 14. Esquema dos homólogos ao eIF4G de *L. major* e seus principais domínios. A região mais conservada das proteínas, semelhante a outros organismos, é o domínio MIF4G (listras vermelhas). Os homólogos *LmEIF4G3* e *LmEIF4G4* aparentemente são funcionalmente relacionados, uma vez que além do domínio HEAT-1/MIF4G, ambos possuem uma região N-terminal curta e uma região conservada de 120 aminoácidos na região C-terminal (listras azuis) que poderia ser um segundo domínio presente em eucariotos, o HEAT-2/MA3 ou o HEAT-3/W2 ou ambos.

Tabela II. Sumário das sequências dos homólogos eIF4G de *L. major* descrevendo as identidades (similaridades) entre as seqüências de *L.major*/ humano e *L.major*/*T.brucei*.

Seqüência	Número de identificação no GeneDB	Tamanho (Nº de aminoácidos)	Massa molecular predita (kDa)	Localização cromossômica	Identidade (similaridade) ao homólogo humano (Q04637)*	Identidade (similaridade) aos ortólogos de <i>T. brucei</i>
<i>LmEIF4G1</i>	LmjF15.0060	1016	114	15	25% (43)	31% (47)
<i>LmEIF4G2</i>	LmjF15.1320	1425	145.9	15	21% (37)	33% (49)
<i>LmEIF4G3</i>	LmjF16.1600	635	71.2	16	26% (39)	38% (55)
<i>LmEIF4G4</i>	LmjF36.6060	765	84.6	36	22% (38)	26% (45)
<i>LmEIF4G5</i>	LmjF10.1080	782	88.8	10	21% (37)	43% (60)

* Similaridade relativa apenas ao domínio HEAT central. Modificado a partir de Dhalia *et al.*, 2005.

3. OBJETIVOS

3.1 Objetivo Geral:

- Identificar interações entre homólogos do eIF4G e os parceiros eIF4E, eIF4A e PABP, em *Leishmania major*, mapeando domínios e motivos de interação com outros fatores e adicionalmente contribuir para o estudo da função dos respectivos homólogos em *T. brucei*.

3.2 Objetivos Específicos:

- Analisar a interação dos homólogos de eIF4G com homólogos de outras subunidades do complexo eIF4F e PABP
- Identificar nos homólogos eIF4G os domínios envolvidos com a interação a outros fatores.
- Realizar mutagênese sítio-dirigida nos resíduos de aminoácidos identificados nos homólogos do eIF4G de *Leishmania major*, visando definir os motivos envolvidos nas suas interações protéicas.
- Analisar a formação de complexos protéicos, *in vitro*, entre os homólogos do eIF4G e outros parceiros em *Leishmania major*.
- Verificar as interações eIF4G/eIF4E dos ortólogos de *Trypanosoma brucei* com as proteínas de *Leishmania major* e vice-versa.
- Determinar os níveis intracelulares das proteínas EIF4AI e EIF4AIII nas formas procíclica e sanguínea de *T.brucei*.

4. MATERIAIS E MÉTODOS

4.1. PCR e clonagem/subclonagem dos homólogos ao eIF4F de *L. major* e *T. brucei*.

Sequências codificantes para os domínios (MIF4G, N-Terminal+MIF4G, C-terminal+MIF4G) e ou proteínas completas *LmEIF4G1-5* [*LmEIF4G1*₁₋₄₇₅ (os números indicam os aminoácidos correspondentes da proteína selvagem), *LmEIF4G1*₁₂₈₋₁₀₁₆, *LmEIF4G2*₁₋₇₀₅, *LmEIF4G3*₁₋₃₁₀, *LmEIF4G3*₂₆₋₆₃₅, *LmEIF4G3* completa, *LmEIF4G4*₁₋₃₆₂, *LmEIF4G4*₃₇₋₃₆₂, *LmEIF4G4*₃₇₋₇₆₅, *LmEIF4G4* completa, *LmEIF4G5*₁₋₄₅₅, *LmEIF4G5*₁₃₈₋₇₈₂] foram amplificadas por PCR, flanqueadas pelos sítios de restrição *Afl* III/ *Not* I (*LmEIF4G1*), *Nco* I/*Xho* I (*LmEIF4G2*₁₋₇₀₅), *Nco* I/*Not* I (*LmEIF4G3*/*LmEIF4G5*), *Bam*H I/*Not* I (*LmEIF4G4*). Os fragmentos gênicos do *LmEIF4G1* foram clonados nos sítios *Nco* I/*Not* I do vetor procariótico de expressão pET21d (Novagen®). A construção *LmEIF4G1*₁₂₈₋₁₀₁₆ (C-terminal+MIF4G) foi digerida com as enzimas de restrição *Sal* I/*Not* I e a subclonagem do fragmento obtido de 2123pb nos mesmos sítios na construção *LmEIF4G1*₁₋₄₇₅ (N-terminal+MIF4G) permitiu a obtenção do gene *LmEIF4G1* completo, no plasmídeo pET21d. Em seguida, a partir da construção pET21d-*LmEIF4G1* foi reamplificado o gene *LmEIF4G1* completo permitindo sua clonagem nos sítios *Bam*H I/*Not* I do vetor procariótico de expressão pGEX4T3 (GE Healthcare®). O fragmento *LmEIF4G2*₁₋₇₀₅ foi clonado nos sítios *Nco* I/*Xho* I do vetor pET21d e reamplificado para clonagem nos sítios *Bam*H I/*Xho* I no vetor pGEX4T3. Os fragmentos do *LmEIF4G3* foram clonados nos sítios *Nco* I/*Not* I dos vetores pET21d e pGEX4T3_{NcoI} (pGEX modificado contendo o sítio de *Nco* I no sítio múltiplo de clonagem – Tamara D’Carli Costa Lima, 2007, Dissertação de Mestrado em Genética - UFPE). Os fragmentos gênicos do *LmEIF4G4* foram clonados nos sítios *Bam*H I/*Not* I dos plasmídeos pGEX4T3 e pET21a (Danielle Maria Nascimento Moura, 2007, Dissertação de

Mestrado em Genética - UFPE). Após sequenciamento, foi necessária a subclonagem dos fragmentos gênicos *LmEIF4G4*₃₇₋₇₆₅ e do gene *LmEIF4G4* completo no vetor pET21a, obtidos a partir da digestão com as enzimas *BamH I*/*Not I* das respectivas seqüências clonadas no vetor pGEX4T3. Os fragmentos gênicos do *LmEIF4G5* foram clonados nos sítios *Nco I*/*Not I* do vetor pGEX4T3_{*NcoI*}. Através da digestão da construção *LmEIF4G5*₁₋₄₅₅ com as enzimas de restrição *Sal I*/*Not I* e subclonagem do fragmento obtido de 1878pb nos mesmos sítios na construção *LmEIF4G5*₁₃₈₋₇₈₂, foi obtida a seqüência completa do gene *LmEIF4G5*₁₋₇₈₂ (*LmEIF4G5* completo) no vetor pGEX4T3 (Danielle Maria Nascimento Moura, 2007, dados não publicados). Em seguida, a partir da digestão da construção pGEX4T3-*LmEIF4G5* com as enzimas *Nco I*/*Not I*, foi realizada a subclonagem do gene completo nos mesmos sítios no plasmídeo pET21d. A clonagem dos genes *LmEIF4E1*, *LmEIF4E2*, *LmEIF4E3*, *LmEIF4AI* e *LmEIF4AIII* (previamente denominados de *LmEIF4A1* e *LmEIF4A2*) nos plasmídeos pET21d, pRSETA (*LmEIF4AIII*) e pGEX4T3 foi descrita previamente (Dhalia *et al.*, 2005). O gene *LmEIF4E4* foi clonado de forma similar: *Nco I*/*Not I* no vetor pET21d e *BamH I*/*Not I* no vetor pGEX4T3 (Éden Ribeiro Freire, 2005, Dissertação de Mestrado em Genética - UFPE). A partir da reamplificação do gene *LmEIF4E3*, usando primers 5' internos, foram gerados fragmentos menores de 888pb e 570pb que foram clonados nos sítios *Nco I*/*Not I* dos plasmídeos pET21d e pGEX4T3. Os genes homólogos da PABP de *L.major* clonados nos sítios *BamH I*/*Hind III* (*LmPABP1*) e *BamH I*/*Not I* (*LmPABP2* e *LmPABP3*), foram clonados nos mesmos sítios do plasmídeo pET21a e subclonado no pGEX4T3 (Tamara D'Carli Costa Lima, 2007, Dissertação de Mestrado em Genética - UFPE). O gene *TbEIF4G3* foi clonado no vetor pGEM-T Easy (Promega®), flanqueado pelos sítios *Hind III* e *Not I*. A partir da clivagem com as enzimas de restrição *Hind III* (digestão parcial) e *Not I*, o *TbEIF4G3* foi subclonado nos mesmos sítios no vetor pET21a. A construção pET21a-*TbEIF4G3* foi digerida com *BamH I* (digestão parcial) e *Not I*, e o referido gene foi também subclonado nos mesmos sítios do vetor pGEX4T3. Já o gene *TbEIF4G4* foi clonado no vetor pGEM-T Easy,

flanqueado pelos sítios *Bam*H I e *Not* I. Em seguida, o gene *TbEIF4G4* foi clivado do vetor pGEM-T Easy, através de digestão com *Bam*H I (digestão parcial) e *Not* I, e subclonado nos mesmos sítios nos plasmídeos pET21a e pGEX4T3 (Danielle Maria Nascimento Moura, 2008, dados não publicados). Os genes *TbEIFAI-III* e *TbEIF4E1-4* foram amplificados por PCR e clonados nos sítios *Hind* III e *Bam*H I. Em seguida as construções dos genes *TbEIF4AI-III* foram digeridas com as enzimas *Eco* RI (parcial) e *Not* I, e os referidos genes foram subclonados no vetor pGEX4T3 (Rafael Dhalia, 2005, Tese de Doutorado em Biologia Molecular – UNB). As construções *TbEIF4E1-4* foram digeridas do vetor pGEM-T Easy, com *Eco*R I e *Bam*H I (digestão parcial com *Bam*H I para o genes *TbEIF4E1* e *TbEIF4E3*), e subclonados nos mesmos sítios do plasmídeo pGEM3zf+ (Promega®) (Éden Ribeiro Freire, 2005, Dissertação de Mestrado em Genética - UFPE). O gene *TbMago* foi clonado nos sítios *Bam*H I e *Hind* III no vetor pGEM3zf+ e *Bam*H I e *Not* I no vetor pGEX4T3.

Para a realização da clonagem gênica, os plasmídeos e os diferentes produtos de PCR foram digeridos com enzimas de restrição, separados por eletroforese em gel de agarose 0.8%, purificados com o *QIAEX II Gel Extraction Kit* (Qiagen®) e ligados utilizando a enzima T4 DNA ligase (New England BioLabs®). Células de *Escherichia coli* DH5α (Invitrogen®) foram transformadas com o plasmídeo recombinante. As colônias obtidas foram crescidas em meio LB contendo 100 µg/mL ampicilina e utilizadas para extração de DNA plasmidial segundo o método de lise alcalina (Sambrook e Russell, 2001).

As construções plasmídias, pET21a/d e pGEX4T3, foram linearizadas com enzimas de restrição de corte único. A quantificação dos DNAs foi realizada através de diluições dos plasmídeos digeridos comparadas com concentrações conhecidas do marcador *1Kb DNA ladder plus* (Invitrogen®). As amostras foram visualizadas após corrida em gel de agarose 0.8% corado com brometo de etídio e quantificadas no programa da câmara digital Kodak 1D Image Analysis Software®. Todos os plasmídeos obtidos neste trabalho foram quantificados e seqüenciados como descritos a seguir.

Tabela III. Oligonucleotídeos usados para amplificação e clonagem dos genes ou fragmentos gênicos nos vetores plasmidiais utilizados neste trabalho.

Genes de <i>L.major</i> / <i>T.brucei</i>	Plasmídeo	Sítio restrição	Orientação	Oligonucleotídeo
<i>LmEIF4G1</i> ₁₋₄₇₅	pET21d	<i>Afl</i> III	Forward	GAACATGTTTCATGGAAACACAGATTG
<i>LmEIF4G1</i> ₁₋₄₇₅	pET21d	<i>Not</i> I	Reverse	TGGCGGCCGCGCTCATGATGGAGGACTGCAG
<i>LmEIF4G1</i> ₁₂₈₋₁₀₁₆	pET21d	<i>Afl</i> III	Forward	GGACATGTTTATGTTCGGTCCGGAAGG
<i>LmEIF4G1</i> ₁₂₈₋₁₀₁₆	pET21d	<i>Not</i> I	Reverse	TGGCGGCCGCCGATAAGTATGTGAGGACGG
<i>LmEIF4G2</i> ₁₋₇₀₅	pET21d	<i>Nco</i> I	Forward	GTAAACCATGGAGCACTCAGGTC
<i>LmEIF4G2</i> ₁₋₇₀₅	pGEX4T3	<i>Bam</i> HI	Forward	GTGGGATCCATGGAGCACTCAGGTCGC
<i>LmEIF4G2</i> ₁₋₇₀₅	pET21d/ pGEX4T3	<i>Xho</i> I	Reverse	TGCTCGAGCAGCATGGCACGCAGGCCATC
<i>LmEIF4G3</i> ₂₆₋₆₃₅	pET21d	<i>Nco</i> I	Forward	TGCCATGGATGAAGTGGTGCGACGCC
<i>LmEIF4G3</i> ₁₋₆₃₅	pET21d	<i>Nco</i> I	Forward	TGCCATGGAGTTCACCGTGGAGCAG
<i>LmEIF4G3</i> ₂₆₋₆₃₅ / <i>LmEIF4G3</i> ₁₋₆₃₅	pET21d/ pGEX4T3	<i>Not</i> I	Reverse	TCGCGGCCGCCTTGGGGAAGCGCTCCTTG
<i>LmEIF4G4</i> ₁₋₃₆₂ / <i>LmEIF4G4</i> ₁₋₇₆₅	pET21a/ pGEX4T3	<i>Bam</i> HI	Forward	TTCGGATCCCCAATTATGCTCTTCAACCTG
<i>LmEIF4G4</i> ₃₇₋₃₆₂ / <i>LmEIF4G4</i> ₃₇₋₇₆₅	pET21a/ pGEX4T3	<i>Bam</i> HI	Forward	CTTGGATCCGAGCCGAAGTCTCGCTAGAG
<i>LmEIF4G4</i> ₁₋₃₆₂ / <i>LmEIF4G4</i> ₃₇₋₃₆₂	pET21a/ pGEX4T3	<i>Not</i> I	Reverse	TTGGCGGCCGCGTCGTTGTTGGTCGTGTTATT G
<i>LmEIF4G4</i> ₃₇₋₇₆₅ / <i>LmEIF4G4</i> ₁₋₇₆₅	pET21a/ pGEX4T3	<i>Not</i> I	Reverse	TTCGCGGCCGCCGTGTCGGGGTACATCTG CCTCCATGGACAAGAAGTACCCGGGCCGCAT C
<i>LmEIF4G5</i> ₁₋₄₅₅	pGEX4T3	<i>Nco</i> I	Forward	TTCGCGGCCGCCTCGTAGTCAGGAATCCGTG
<i>LmEIF4G5</i> ₁₋₄₅₅	pGEX4T3	<i>Not</i> I	Reverse	CGACCATGGACGACATGGAGAACAAC
<i>LmEIF4G5</i> ₁₃₈₋₇₈₂	pGEX4T3	<i>Nco</i> I	Forward	TTCGCGGCCGCCTTCCGCTCCTTGATGGCCTC
<i>LmEIF4G5</i> ₁₃₈₋₇₈₂	pGEX4T3	<i>Not</i> I	Reverse	TTCGCGGCCGCCTTCCGCTCCTTGATGGCCTC
<i>LmEIF4E3</i> ₅₄₋₃₄₉	pET21d/ pGEX4T3	<i>Nco</i> I	Forward	TGCCATGGCCGCGACATCGACGC
<i>LmEIF4E3</i> ₁₆₀₋₃₄₉	pET21d/ pGEX4T3	<i>Nco</i> I	Forward	TGCCATGGCGTCTAATAGCTCCC
<i>LmEIF4E3</i> ₅₄₋₃₄₉ / <i>LmEIF4E3</i> ₁₆₀₋₃₄₉	pET21d/ pGEX4T3	<i>Not</i> I	Reverse	TGGCGGCCGCGAACGTGTGATCGGGCG
<i>TbEIF4G3</i>	pGEM T- Easy	<i>Hind</i> III	Forward	AAGCTTCCGCCACCATGCACGTCTACACCATC G
<i>TbEIF4G3</i>	pGEM T- Easy	<i>Not</i> I	Reverse	GGATCCAGTCAAGGTCCCCTTCCC TCTGGATCCCCCGCCACCATGCTGTTCAAACC T
<i>TbEIF4G4</i>	pGEM T- Easy	<i>Bam</i> HI	Forward	CGTGG
<i>TbEIF4G4</i>	pGEM T- Easy	<i>Not</i> I	Reverse	GGATCCCGCATATATAACTGAGAGTTC
<i>TbMago</i>	pGEM3zf+/ pGEX4T3	<i>Bam</i> HI	Forward	TCTGGATCCCCCGCCACCATGCAGTCTTGTC TTCAC
<i>TbMago</i>	pGEM3zf+	<i>Hind</i> III	Reverse	TCTAAGCTTCTACCACGCCTTGATTTTAAAAT G
<i>TbMago</i>	pGEX4T3	<i>Not</i> I	Reverse	TCGCGGCCGCCACGCCTTGATTTTAAAATG

4.2 - Sequenciamento das construções plasmidiais

Para a reação de sequenciamento dos plasmídeos pET21a/d e pGEX4T3 contendo os diferentes insertos de *L. major* ou *T. brucei*, foi utilizado o kit *big Dye* reagentes da Applied Biosystems (*bigDye Terminator Cycle Sequencing Ready Reaction v3.1®*). A reação se deu com a adição de 1µL [3.2 pmoles/ µL] de oligonucleotídeos iniciadores de DNA (M13F e M13R), 0.5µL de *bigDye*, 1µL de tampão [200mM Tris-HCl e 5mM MgCl₂ pH 9.0], 300 a 500ng de DNA plasmidial e quantidade suficiente de água para completar 10µL. As misturas foram incubadas a 95° C por 5 minutos, e imediatamente incubadas no gelo. Após 35 ciclos [desnaturação (96° C por 45 segundos), anelamento (50° C por 30 segundos) e extensão (60° C por 4 minutos)]. Os DNAs foram precipitados com isopropanol com concentração final de 50% (v/v), por 30 minutos a temperatura ambiente, lavados com etanol a 60% (v/v) e solubilizados em 15µL de formamida Hi-Di (Applied Biosystems®).

No sequenciador automático, modelo ABI Prism 3100 (Applied Biosystems), foi utilizado o polímero POP6 em capilares de 50 cm, e a análise do sequenciamento realizada utilizando os programas 3100 *Data Collection* versão 1.1 e *Sequencing Analysis* versão 3.7 também da Applied Biosystems. As seqüências obtidas foram analisadas através dos programas SeqManTM II, EditSeqTM, MegalignTM, MapDrawTM, pacote de programas Lasergene, versão 4.01, DNASTAR® Inc., Madison, WI).

4.3. Expressão e purificação de proteínas recombinantes

Para expressar as proteínas recombinantes fusionadas à uma sequência de polihistidinas (genes clonados nos vetores de expressão pET21a, pET21d) ou Glutathione-S-Transferase (GST) (genes clonados no vetor de expressão pGEX4T3), as construções plasmidiais foram inseridas em cepas BL21*star* (Invitrogen®) ou BLR (Novagen®) de *Escherichia coli*. As bactérias foram transformadas por choque térmico, sendo incubadas com os DNAs a 4°C/30 minutos antes de sofrerem choque térmico a 37°C/5 minutos. As bactérias transformadas foram semeadas em placas de Petri contendo meio ágar LB sólido [na presença de ampicilina (100 µg/ml), para BL21*star*, ou ampicilina (100µg/ml)/tetraciclina (12 µg/ml)/cloranfenicol (20 µg/ml), para BLR], e em seguida foram incubadas a 37°C por 18 horas. As colônias obtidas foram inoculadas em 15 ml de meio LB líquido contendo antibióticos, como descrito anteriormente, e foram crescidas a 37°C, sob agitação de 150 rpm por 18 horas. Para a indução da expressão das proteínas recombinantes, alíquotas de 1-10 ml das culturas obtidas foram inoculadas separadamente em 500 ml de meio LB com ampicilina (100µg/ml) crescidas a 37°C sob agitação de 150 rpm, e o crescimento bacteriano foi monitorado pela aferição em espectrofotômetro (comprimento de onda de 600 nm) até atingir a densidade óptica (DO) entre 0.5 ou 0.8. Após a adição de IPTG na concentração final de 0.1mM, a cultura permaneceu sob agitação por 4 horas a 30°C (DO igual a 0.5) ou alternativamente por 1 hora a 25°C (DO igual a 0.8). As culturas foram centrifugadas a 9600 x g durante 10 minutos, e o sedimento bacteriano obtido foi ressuspensionado em 20 ml de tampão PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.47 mM KH₂PO₄, pH 7.3) acrescido de inibidor de proteases EDTA-*free*, de acordo com as instruções do fabricante (Roche®). Os extratos bacterianos foram lisados por ultra-sonicação (6 pulsos de 30 segundos), acrescidos de 1% final de Triton-X 1%, e centrifugados a 9600 x g por 10 minutos, a 4°C. O sobredanante pós-lise resultante foi incubado com 200µl de Resina Ni-NTA Agarose

(Quiagen®) ou Glutathiona Sefarose (GE Healthcare®) (lavadas 3 vezes com 10 ml de PBS (por 1h a 4°C sob agitação) sendo em seguida as transferidas para tubos de 1.5 ml. Para as proteínas fusionadas à cauda de poli-Histidinas, a resina foi lavada 3 vezes com 500µl de PBS, 3 vezes com 500µl de tampão de A (50mM NaH₂PO₄, 300mM e NaCl e 10 mM imidazol), e em seguida as proteínas recombinantes foram recuperadas da resina a partir da obtenção de eluatos resultantes de duas lavagens com 500µl de tampão de lavagem, contendo 0.5M de imidazol. Em todas as etapas da purificação foram usadas soluções com concentração de 10mM de imidazol, para evitar ligação inespecífica de proteínas à resina Ni-NTA Agarose. As proteínas ligadas à GST foram eluídas com 500µl de 20mM glutathiona reduzida (em 50 mM Tris-HCl pH 8.0) por 30 minutos e 500µl de 50mM glutathiona reduzida por mais 30 minutos. Aliquotas de todas as fases da purificação foram verificadas em gel 15% SDS-PAGE e, em seguida, os eluatos obtidos foram submetidos a diálise. As amostras foram transferidas para sacos de diálise de 6mm de diâmetro (Sigma®) e incubadas em 700 ml de tampão de diálise (44 mM Na₂HPO₄; 5,2 mM NaH₂PO₄; 50 mM NaCl; 5%Glicerol; 0,1% Triton X-100), sob leve agitação por 6 horas a 4°C, sendo o tampão trocado 3 vezes a cada intervalo de 2 horas. As proteínas purificadas foram analisadas em gel 15% SDS-PAGE corados com azul de Coomassie R-250 (Sigma), sendo a quantificação realizada através de diluições seriadas das proteínas recombinantes comparadas com diluições seriadas de concentrações conhecidas de BSA no programa da câmara digital Kodak 1D Image Analysis Software®

4.4. Mutagênese sítio-dirigida

Os genes *LmEIF4G3* e *LmEIF4G4*, clonados respectivamente nos vetores pET21a e pET21d foram submetidos a reações de mutagênese sítio-dirigida utilizando o Kit *QuikChange II Site-Directed Mutagenesis* (Stratagene®).

4.4.1 Desenho dos Oligonucleotídeos

Para desenhar os oligonucleotídeos algumas recomendações foram seguidas, de acordo com as recomendações do fabricante:

1. Ambos os oligonucleotídeos, contendo as mutações de interesse, se anelam nas mesmas regiões em ambas as fitas de DNA.
2. Tamanho entre 25 e 45 bases, com temperatura de anelamento $\geq 78^\circ\text{C}$.
3. A temperatura de anelamento foi obtida através da fórmula:
$$T_m = 81.5 + 0.41 (\%GC) - 675/N$$
4. As mutações foram posicionadas na região central do oligonucleotídeo distando de 10 a 15 bases de sua extremidade para ambas as direções, sendo estas bases preservadas na sequência original.
5. Os oligonucleotídeos foram otimizados para terem um conteúdo de GC acima de 40%, terminando em uma ou mais bases G ou C.

Tabela IV. Características dos oligonucleotídeos obtidos, seguindo as recomendações do fabricante.

Oligonucleotídeos	Nº bases	Nº GC	% GC	Bases não pareadas	% bases não pareadas	T _m
Mut1	37	27	73	5	13,5	79,7°C
Mut2	37	27	73	5	13,5	79,7°C
Mut3	37	27	73	5	13,5	79,7°C
Mut4	37	27	73	5	13,5	79,7°C
Mut5	42	29	69	6	14,3	79,5°C
Mut6	42	29	69	6	14,3	79,5°C
Mut7	34	26	76,5	5	14,7	78,3°C
Mut8	34	26	76,5	5	14,7	78,3°C
Mut9	36	26	72,2	5	13,9	78,5°C
Mut10	36	26	72,2	5	13,9	78,5°C
Mut11	37	28	75,7	6	16,2	78,1°C
Mut12	37	28	75,7	6	16,2	78,1°C
Mut13	41	21	51,2	2	4,9	81,2°C
Mut14	41	21	51,2	2	4,9	81,2°C
Mut15	43	27	62,8	4	9,3	82,2 °C
Mut16	43	27	62,8	4	9,3	82,2 °C

Tabela V. Descrição dos oligonucleotídeos utilizados para mutagênese sítio-dirigida dos genes *LmEIF4G3* e *LmEIF4G4*.

Nome	Orientação	Oligonucleotídeo	Mutação
Mut1	Forward	CGCCGTACCCGGGCG GCTGCTG CGGATGAAGTGGTGCG	<i>LmEIF4G3</i> - FSL/AAA
Mut2	Reverse	CGCACCACCTTCATCC GCAGCAG CGCCCGGGTACGGCG	<i>LmEIF4G3</i> - FSL/AAA
Mut3	Forward	GCTTGACGCAGACG GCGGCGG CGCGCGGTGAGAATGC	<i>LmEIF4G3</i> - KLV/AAA
Mut4	Reverse	GCATTCTCACCGCG CGCCGCC CGCTCTGCGTCAAGC	<i>LmEIF4G3</i> - KLV/AAA
Mut5	Forward	CGACTGCTGCACGGCACAG CGGCCG CGCTCACAGAGGAGAAC	<i>LmEIF4G3</i> - LNK/AAA
Mut6	Reverse	GTTCTCCTCTGTGAG CGCGCCG CTGTGCCGTGCAGCAGTCG	<i>LmEIF4G3</i> - LNK/AAA
Mut7	Forward	CCTGAGCCGAAC GCCGCGG CAGAGCGCGTGCTGC	<i>LmEIF4G4</i> - FSL/AAA
Mut8	Reverse	GCAGCACGCGCTCT GCCGCGG CGTTCCGGCTCAGG	<i>LmEIF4G4</i> - FSL/AAA
Mut9	Forward	GGCGGAGATTCCT GCGGCCG CCACGTCCGAGAACGG	<i>LmEIF4G4</i> - KLV/AAA
Mut10	Reverse	CCGTTCTCGGACGTGG CGGCCG CAGGAATCTCCGCC	<i>LmEIF4G4</i> - KLV/AAA
Mut11	Forward	CCGCGTTCAGGGCTCG GCGGCCG CGCTCACCGACAAG	<i>LmEIF4G4</i> - LNK/AAA
Mut12	Reverse	CTTGTCGGTGAG GCGGCCG CCGAGCCCTGAACGCGG	<i>LmEIF4G4</i> - LNK/AAA
Mut13	Forward	GGAGAAGGTGAAGAACCAG KMT ACGCTTGCCGACATTCTCG	<i>LmEIF4G4</i> -M20Y,S,D ou A- Primer degenerado
Mut14	Reverse	CGAGAATGTCGGCAAGCGTA KM CTGGTTCTTCACCTTCTCC	<i>LmEIF4G4</i> -M20Y,S,D, ou/A- Primer degenerado
Mut15	Forward	CCAGATGACGCTTGCCGAC GCTGCCG CTTTTCGCGACACATGG	<i>LmEIF4G4</i> - I25A/L26A
Mut16	Reverse	CCATGTGTGCGGAAAAGCG G CAGCGTCGGCAAGCGTCATCTGG	<i>LmEIF4G4</i> - I25A/L26A

4.4.2 Reação de mutagênese

Inicialmente os oligonucleotídeos liofilizados, sintetizados em um total de 25nM, foram dissolvidos em água deionizada para uma concentração estoque de 1µg/µl. Após a confirmação da concentração e integridade em gel de agarose 1% corado com brometo de etídio, estes oligonucleotídeos foram diluídos para a concentração de uso de 25ng/µl. Maxipreparações das construções plasmídias pET-*LmEIF4G3-4* foram linearizadas com *Nco* I

(*LmEIF4G3*) e *BamH I* (*LmEIF4G4*) e quantificadas em gel, corado com brometo de etídio, como descrito no tópico 4.1. As quantificações foram ainda confirmadas por espectrofotometria. A mutagênese sítio-dirigida foi realizada numa reação contendo: 5µL de tampão de reação 10 vezes concentrado, 5µL de plasmídeo (6ng/µl - 30ng), 5µL do oligonucleotídeo 1 (25ng/µL - 125 ng), 5µL do oligonucleotídeo 2 (25 ng/µl -125 ng), 1µL de dNTPs (Stratagene®), 28µL de água deionizada e 1µL de DNA polimerase *PfuUltra* HF (2.5 U/µL). As condições utilizadas na reação de PCR foram: 95°C por 30 segundos, 55°C por 1 minuto e 68°C por 7 minutos (16 ciclos). Como controle da reação foi utilizado o plasmídeo pWhitescriptTM, que contém um códon de parada (TAA) dentro do gene da β-galactosidase na posição onde deveria haver uma Glutamina (CAA - pBluescript), e oligonucleotídeos que permitem reverter esta mutação. Após a reação, 10µL de cada amplificação foram resolvidos em gel de agarose 0.8%, contendo brometo de etídio, e visualizados em transiluminador UV. Em seguida, as reações foram digeridas com 1µL da enzima *Dpn I* (10 U/µL) a 37 °C/1h, que digere os plasmídeos não-mutagenizados. Bactérias *XL1-Blue* eletrocompetentes, previamente preparadas em nosso laboratório, foram transformadas com 1-5µL das reações de mutagênese (tratadas com *Dpn I*) através de um choque elétrico de 2kV. Após a descarga elétrica, as bactérias foram inoculadas em 1 ml de meio LB puro por 1 hora (para recuperação) e semeadas em placas contendo LB sólido com ampicilina (100 µg/ml), a 37° C por 18 horas. Para avaliar a eficiência de transformação, 1µl do plasmídeo pUC19 (0,1ng) foi transformado nas mesmas condições. Alguns clones transformantes, oriundos da mutagênese, foram inoculados em 2ml de meio LB contendo ampicilina, crescidos a 37°C por 18 horas e submetidos à minipreparações de DNA (Sambrook e Russell, 2001). Os plasmídeos resultantes, obtidos foram quantificados e seqüenciados, para a confirmação das mutações inseridas.

4.5 - Transcrição e expressão de mRNAs sintéticos

Para transcrição *in vitro*, com a enzima T7 RNA polimerase, primeiro as diferentes construções plasmidiais foram linearizadas com as seguintes enzimas de restrição: [*LmEIF4G1*, *LmEIF4G3-5*, *LmEIF4E1-4*, *LmEIF4AI*, *LmEIF4G3_{FSL/AAA}*, *LmEIF4G3_{KLIV/AAA}*, *LmEIF4G4_{FSL/AAA}*, *LmEIF4G4_{M20Y}*, *LmEIF4G4_{M20S}*, *LmEIF4G4_{M20D}*, *LmEIF4G4_{M20A}* e *LmEIF4G4_{IL}* (clonados nos vetores pET21a ou pET21d, com *Not* I)]; [*LmEIF4G2*, *LmEIF4AI* e *LmEIF4AIII* (clonados nos vetores pET21d ou pRSETA, com *Xho* I)]; [*LmEIF4G3_{LNK/AAA}*, *LmEIF4G4_{KLIV/AAA}*, *LmEIF4G4_{LNK/AAA}* (clonado no pET21a ou pET21d com *Pvu* I)]; [*TbEIF4E1* e *TbEIF4E3* (clonados no vetor pGEM3zf+ com *Xba* I)]; [*TbEIF4E2* e *TbEIF4E4* (clonados no vetor pGEM3zf+ com *BamH* I)]. O gene codificando o eIF4G humano foi obtido a partir de uma construção no pBluescript KS (Melo et al., 2003), linearizada com *Sac* I. As reações de transcrição foram realizadas em um volume final de 50µL contendo: 5µL tampão de transcrição 10 vezes concentrado (400mM Tris-HCl pH 8.0, 150mM MgCl₂), 2µL de solução de rNTPs [100mM-ATP, CTP, UTP e 10mM-GTP (GE Healthcare®)], 5µL de 10mM m⁷-GTP (New England Biolabs®) a 5 mM na reação, 0.5µL de 1M DTT, 1.5 µL de inibidor de RNase (GE Healthcare®), 2µL T7 RNA polimerase 5U/µL (Usb®), 3µg de DNA plasmidial e 3 µg de DNA plasmidial cortado com enzima de restrição 3' imediatamente após o gene, e água deionizada, tratada com DEPC, para 50µL final. A reação foi incubada por 2 horas a 37°C, após os 30 minutos iniciais foi adicionado 0.5 µl de 100mM GTP, e após mais 30 minutos foi adicionado 0.5µl de T7 RNA polimerase. Ao término das reações de transcrição, 1µL de cada transcrito sintetizado foi analisado em gel de 1% agarose corado com brometo de etídio. Os mRNAs sintéticos obtidos foram extraídos em fenol/cloroformio 1:1, precipitados com 100% etanol / 0,3M NaCl e ressuspensos em 20µL água deionizada tratada com DEPC. Aliquotas de 0.1µL, 0.2µL e 1 µL dos RNAs purificados foram traduzidas

em lisado de reticulócito de coelho (RRL), de acordo com as recomendações do fabricante (Promega®): 3.6µL de lisado, 0.2µL de solução de aminoácidos livre de metionina, 0.2µL de metionina marcada com ^{35}S -10µCi/µL e 1µL de RNA. As reações de tradução foram realizadas na temperatura de 30°C, durante 90 minutos. Após a tradução, alíquotas de 5µL de cada amostra foi acrescida de tampão de amostra para, gel de SDS-PAGE, e submetida à migração por eletroforese. Após a migração, o gel foi corado com azul de Coomassie R-250 (Sigma®) e submetido a secagem em papel de filtro a 70°C, durante uma hora, sob vácuo. Para a visualização das proteínas recombinantes marcadas ^{35}S , o gel foi exposto a um filme de autoradiografia (BioMax MR, Kodak®), que foi em seguida revelado.

4.6 - Ensaio de interação proteína-proteína (*Pull-Down*)

Para os ensaios de interação proteína-proteína (*pull-down*), as resinas Ni-NTA Agarose ou Glutathione Sefarose 4B foram previamente equilibradas com o tampão de ligação B (100mM KCl, 1 mM MgCl₂, 50mM Hepes pH 7.2, 0.2% NP-40, 5% glicerol). O tampão B foi acrescido de 10mM imidazol nos ensaios com as proteínas de fusão com cauda de poli-Histidinas, para evitar ligações inespecíficas com a resina Ni-NTA Agarose, enquanto nos experimentos com as proteínas de fusão GST a resina Glutathione Sefarose 4B foi previamente bloqueada com 10µg/µL de BSA, com a mesma finalidade. Para cada ensaio, aproximadamente 10µL de cada resina bloqueada foi incubada com 2µg de proteína recombinante em um volume final de 200µL (em tampão B), durante 1 hora/sob agitação/ a 4°C. As resinas foram lavadas duas vezes com tampão B e incubadas com 5µL de proteínas marcadas radioativamente (no volume final de 200µL em tampão B), durante 2 horas/sob agitação/a 4°C. Após três lavagens com 500 µL de tampão B, as proteínas ligadas às resinas foram eluídas com 40µL de tampão de amostra para SDS-PAGE e 10µL de cada amostra

foram submetidas a eletroforese em gel 15%. Os géis foram corados com azul de Coomassie R-250, para a visualização das proteínas recombinantes e submetidos à autoradiografia (como descrito anteriormente) para visualização das proteínas radioativas marcadas com ^{35}S (Einarson, 2001; Dhalia *et al.*, 2005). A figura 4 ilustra um ensaio de interação proteína-proteína, *pull-down*, entre uma proteína recombinante fusionada à GST e uma proteína marcada radioativamente com ^{35}S .

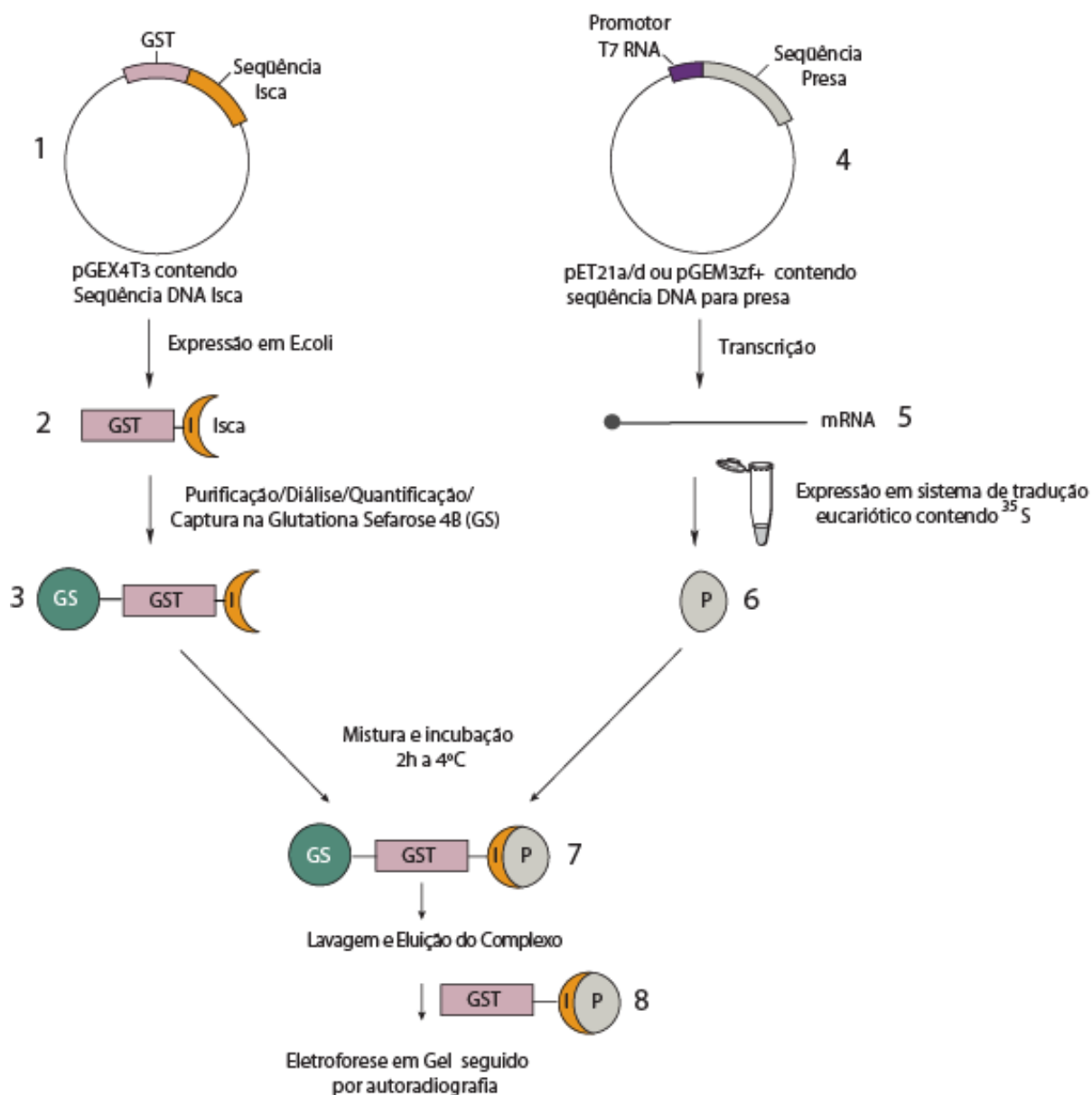


Figura 15. Esquema da técnica de *Pull-Down* usando a proteína isca (I) expressa em bactéria e a proteína presa (P) expressa em sistema de tradução de mamíferos (marcada com ^{35}S). 1- A sequência codificando para a isca (I) foi clonada no vetor pGEX4T3; 2- Expressão da proteína I fusionada a GST, em *E. coli*. 3- A proteína GST-I purificada,

dialisada e quantificada foi imobilizada na resina Glutathione Sefarose 4B (GS); 4- A sequência codificante para a proteína presa (P) foi clonada nos vetores pET21a, pET21d ou pGEM3zf+, que permitem a sua transcrição *in vitro*; 5- Síntese da proteína ^{35}S -P, em sistema de tradução de mamíferos; 6- Uma alíquota da proteína GST-I, imobilizada na resina GS, é incubada com a proteína ^{35}S -P por duas horas a 4°C; 7- Após lavagens, para remover ligações inespecíficas, as proteínas presa e isca são eluídas da GS. As proteínas são em seguida fracionadas em SDS-PAGE 15%, corado com azul de Coomassie R-250 para visualização das proteínas recombinantes e as proteínas radioativas foram detectadas por autoradiografia.

4.7. Ensaios de Western-blot

Anticorpos contra as proteínas TbEIF4AI-III (Dhalia R, tese de doutorado em biologia molecular), , foram utilizados em ensaios de *Western-blot*. Proteínas recombinantes e extratos de parasitas foram fracionadas em SDS-PAGE 12,5% e transferidas para membrana de transferência Immobilon-P PVDF (Millipore®), para a realização de ensaios de detecção utilizando os referidos anticorpos. A membrana foi bloqueada em solução de TBS (20mM Tris, 500mM NaCl, pH 7,5), 5% leite desnatado e 0,05% Tween-20. em temperatura ambiente/sob agitação/ por uma hora. Em seguida, foi incubada com os respectivos soros policlonais imunopurificados (Dhalia *et al.*, 2005), em diluições variadas, em solução de TBS, 1% leite desnatado e 0,05% Tween-20 nas mesmas condições já relatadas. Após lavagem com TBS, a membrana foi incubada com o segundo anticorpo (anti-IgG de coelho conjugado à peroxidase), numa diluição de 1:10,000, nas mesmas condições anteriores. Após nova lavagem com TBS, a membrana foi incubada (1 minuto/temperatura ambiente) com uma solução contendo 1,2mM luminol, 0,4mM iodofenol (diluído em DMSO) e 0,03% peróxido de hidrogênio para a reação de quimioluminescência (ECL - *Enhanced Chemiluminescent*). A membrana foi seca com papel de filtro, exposta contra filme específico para ECL (GE Healthcare®), em intervalos de 1, 5 e 10 minutos, que foi revelado através de sua incubação em solução 1:2 dektol (Kodak®), seguida de fixação em solução de ácido acético.

5. RESULTADOS

Os resultados deste trabalho podem ser agrupados em quatro tópicos distintos devido as diferentes abordagens utilizadas para responder a questionamentos específicos, sendo cada um deles relacionado à produção de um artigo científico diferente. O primeiro, foco central desta tese, reúne um conjunto de experimentos de interação do tipo *pull-down* que teve como objetivo, identificar e mapear domínios nos homólogos do eIF4G responsáveis pela interação com homólogos de outras subunidades do complexo eIF4F (eIF4E/eIF4A). Aminoácidos importantes para estas interações foram identificados, a partir de proteínas mutagenizadas em sítios específicos. Estes resultados, juntamente com dados *in vivo* das proteínas *TbEIF4G3-5*, serão em breve publicados no artigo científico “**Two eIF4G Homologues Conserved in Trypanosomatid species display functional differences compatible with roles in two diverged eIF4F complexes**”, no qual serei primeiro autor, junto com a doutoranda Danielle Nascimento Moura, uma vez que ambos os autores contribuíram igualmente para o trabalho. O segundo tópico, aborda os resultados de interação entre os homólogos do eIF4G/domínios e três homólogos da PABP em *Leishmania major*. Estes resultados, juntamente com dados funcionais das PABPs de *L. major*, serão publicados no artigo científico “**Functional Characterization of Three Different *Leishmania* PABP Homologues with Distinct RNA Binding Properties**”, em fase final de redação, onde serei co-autor. No terceiro tópico, foram realizadas análises comparativas das interações entre os homólogos EIF4G3-4 aos EIF4E3-4, de *L. major* e *T. brucei*, e os resultados obtidos farão parte do manuscrito “**Characterization of Multiple Homologues of the cap binding protein (eIF4E) from *Trypanosoma brucei***”, em fase final de redação, onde também serei co-autor. No quarto, e último tópico, foi realizado um estudo da função dos eIF4As em *T. brucei*. Neste trabalho foram determinados os níveis intracelulares das proteínas *TbEIF4AI* e *TbEIF4AIII* nas suas formas procíclicas e sanguíneas, além da avaliação do efeito causado pela depleção da proteína *TbEIF4AIII* (em

extratos da forma procíclica) após interferência de RNA, que permitiram minha participação, como co-autor, no trabalho **“The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct”**, publicado na revista “Nucleic Acids Research” no ano de 2006. Adicionalmente, as proteínas eIF4A de *T. brucei* foram testadas quanto sua capacidade de interação com a proteína *TbMago*, possivelmente relacionada a algum mecanismo de degradação de mRNA em *T.brucei*, abrindo a possibilidade de início de uma nova linha de pesquisa em nosso laboratório.

5.1. Estudo da interação entre homólogos do eIF4G com as demais subunidades do complexo eIF4F e PABP, em *L. major*

Em animais, plantas e leveduras para que o processo de tradução seja iniciado é importante a montagem do complexo eIF4F, através de interações específicas entre as subunidades eIF4E, eIF4A e eIF4G (Marintchev *et al.*, 2009). A ocorrência de múltiplos homólogos de cada uma destas subunidades nos tripanosomatídeos indica a existência de mais de um complexo eIF4F nestes parasitas, ou então que apenas um dos homólogos de cada subunidade participe na composição de um único eIF4F. Neste último caso, fica em aberto qual seria a função dos demais homólogos de cada subunidade e como estes interagiriam entre si. Do ponto de vista experimental, o foco principal deste tópico foi a determinação das interações existentes entre os homólogos de eIF4G encontrados em *L. major* e seus possíveis parceiros funcionais dentro do complexo eIF4F. Além disso, quando possível, buscou-se mapear os sítios/motivos de ligação responsáveis pelas interações identificadas, de forma a verificar elementos conservados e ou divergentes em relação ao que se sabe sobre essas interações em outros eucariotos.

Como resultado de trabalhos anteriores (Reis C., Dissertação de Mestrado 2004; Moura D., Monografia de Conclusão de Curso 2004; Dhalia et al., 2005; Dhalia R., Tese de Doutorado 2005), parte dos genes codificando para os 5 homólogos de eIF4G de *L. major* já

estavam total ou parcialmente clonados, em diversos plasmídeos, quando do início deste doutorado. Neste trabalho, foram selecionados os vetores de expressão pET21a/d e/ou pGEX4T3 para a expressão dos homólogos em questão, em *Escherichia coli*, como proteínas recombinantes fusionadas a uma sequência de poli-histidinas (HIS) ou a proteína Glutathione S-transferase (GST), respectivamente. Os fragmentos clonados no vetor pET21a/d também permitiram, por transcrição e tradução *in vitro*, a produção das respectivas proteínas marcadas com ^{35}S com ou sem a sequência de poli-histidinas. Ao longo deste trabalho novas construções foram realizadas, de forma a se obter o perfil completo dos homólogos de eIF4G em *L. major* em ambos os plasmídeos e, quando possível, foram obtidos os genes completos e domínios relevantes destas proteínas. A figura 16 ilustra, de uma forma esquemática, o conjunto de construções obtidas para os diferentes ensaios de interação proteína-proteína, do tipo *pull-down*, realizados. Todas estas construções foram confirmadas por sequenciamento, e as sequências obtidas foram comparadas com àquelas geradas pelos projetos genomas dos tripanossomatídeos. A figura 17 mostra um alinhamento do domínio MIF4G, das cinco proteínas candidatas a eIF4G de *L. major*. Foi a presença deste domínio que possibilitou encontrar os homólogos ao eIF4G no banco de dados de sequências genômicas do parasita, visto ser a região mais conservada. Nos ensaios de afinidade investigou-se então a interação entre os homólogos de eIF4G, e/ou regiões contendo o domínio MIF4G com os homólogos ao eIF4A, eIF4E e PABP de *L. major*.

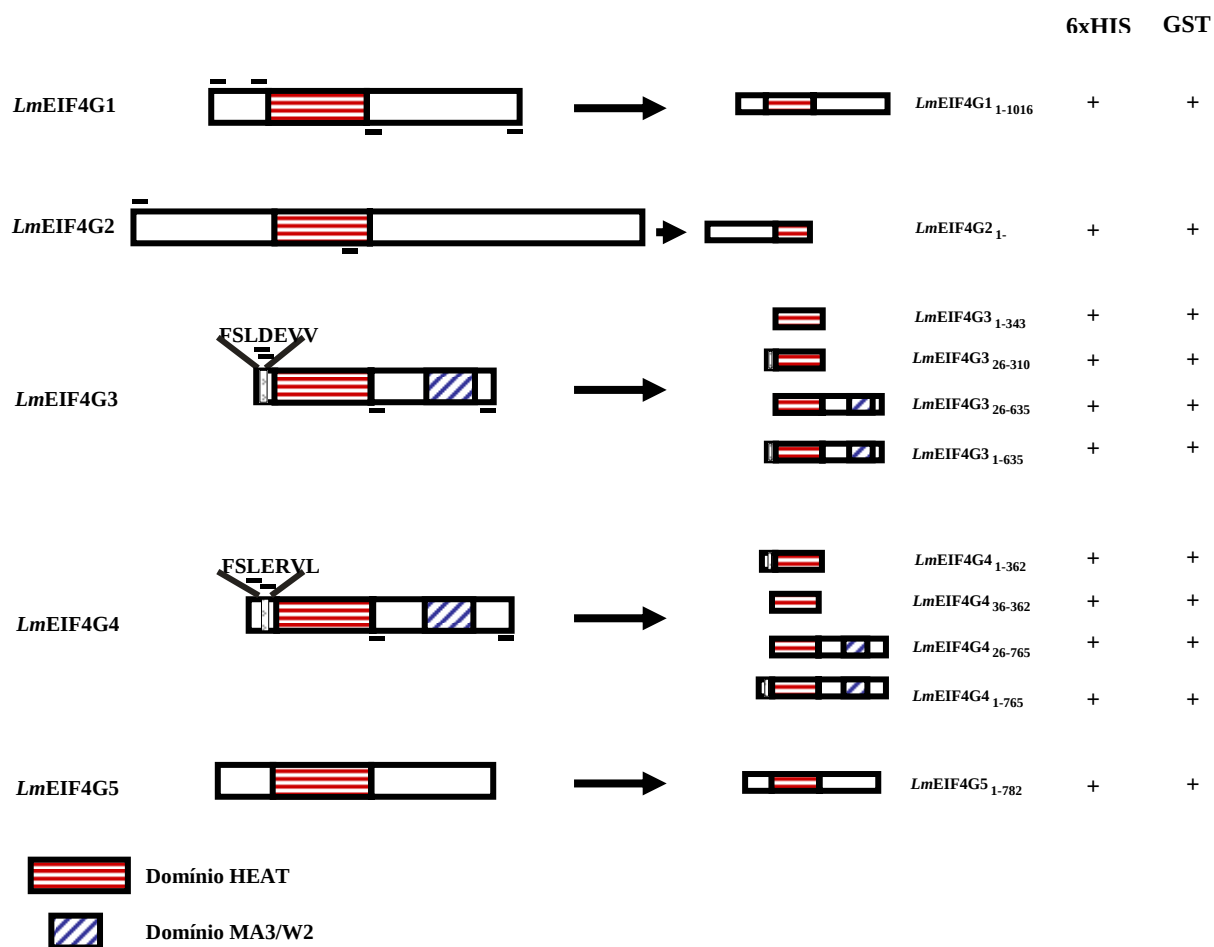


Figura 16. Esquema das proteínas *LmEIF4Gs*, e dos vários fragmentos utilizados nesse trabalho, destacando possíveis domínios de ligação a outras proteínas. Esta figura demonstra todos os fragmentos obtidos para cada uma das proteínas, bem como as proteínas de fusão obtidas de acordo com as clonagens realizadas. Todos os genes completos foram clonados, com exceção do gene codificante a proteína *LmEIF4G2*. Os traços pretos representam os primers utilizados para a obtenção dos diferentes fragmentos gênicos, a partir da reação em cadeia da polimerase, utilizando DNA total de *L. major*. Para obtenção dos genes completos *LmEIF4G1* e 5 foram realizadas subclonagens utilizando diferentes construções contendo diferentes regiões dos genes em questão.

[illegible]

5.1.1. Estudo da interação entre homólogos do eIF4G/eIF4A em *L. major*

Com o objetivo de investigar a presença de um ou mais complexos eIF4F em *L. major*, e definir quais homólogos das subunidades participariam de sua composição, proteínas recombinantes e marcadas com ^{35}S foram produzidas e utilizadas em ensaios de interação proteína-proteína do tipo *pull-down*. Através desta técnica já foi validado o eIF4A humano recombinante se liga eficientemente ao eIF4G humano radioativo, e que a região central (MIF4G) da proteína *LmEIF4G3* é capaz de interagir com o *LmEIF4AI*, enquanto o mesmo não ocorre para a região similar das proteínas *LmEIF4G1* e 2 (Dhalia *et al.*, 2005). Para avaliar a interação entre homólogos do eIF4G e eIF4A de *L. major*, as proteínas *LmEIF4G* foram sintetizadas com ^{35}S em, sistema de tradução eucariótico e as proteínas recombinantes *LmEIF4A* foram produzidas em *Escherichia coli*, em fusão à GST.

As proteínas *LmEIF4G1-5*, marcadas com ^{35}S , foram testadas quanto a sua capacidade de ligar-se as proteínas GST-*LmEIF4AI* e/ou GST-*LmEIF4AIII*. Nos ensaios de *pull-down*, as proteínas fusionadas a GST são pré-incubadas com a resina Glutathiona Sefarose 4B e em seguida incubadas com as proteínas marcadas com ^{35}S . A resina é então coletada, lavada para remoção de interações inespecíficas e as proteínas retidas são analisadas por eletroforese desnaturante, em gel de poliacrilamida. Apenas no caso de interações específicas, entre as proteínas com GST e àquelas marcadas com ^{35}S é que as proteínas radioativas vão ser detectadas junto à resina. Como pode ser visto na Figura 18, apenas a proteína *LmEIF4G3* é capaz de se ligar eficientemente ao *LmEIF4AI*. As proteínas *LmEIF4G1*, 2 e 4 são capazes de interagir com o *LmEIF4AI*, porém de forma muito mais fraca quando comparados ao *LmEIF4G3*. A proteína *LmEIF4AIII* é capaz de interagir fracamente com o homólogo *LmEIF4G12₁₋₇₀₅*. Vale ressaltar a total ausência de ligação entre o *LmEIF4G5* e os dois homólogos de eIF4A.

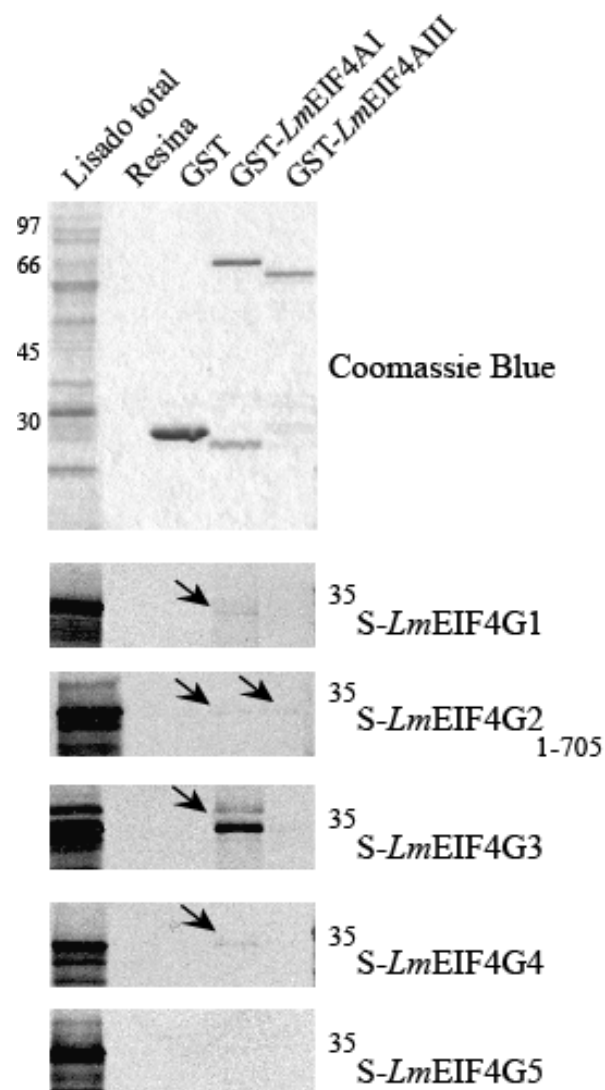


Figura 18. Análise de interação entre os homólogos ao eIF4A e os homólogos ao eIF4G de *L. major*. As proteínas recombinantes GST, GST-*LmEIF4AI* e GST-*LmEIF4AIII* foram imobilizadas em resina Glutathione Sefarose 4B e incubadas com as proteínas *LmEIF4G1-5* marcadas com ³⁵S. O painel superior representa as proteínas recombinantes fracionados em gel 15% SDS-PAGE e corados com azul de Coomassie. Os painéis inferiores representam as autoradiografias mostrando as proteínas marcadas com ³⁵S. Ligações específicas, entre as proteínas marcadas com ³⁵S e as proteínas fusionadas com GST, estão indicadas por setas nas autoradiografias. Os números à esquerda dos painéis são referentes a marcadores de peso molecular, em kDa.

5.1.2. Estudo da Interação entre os homólogos eIF4G/eIF4E em *L. major*

Para a validação dos ensaios de interação eIF4G/eIF4E, inicialmente foi imobilizado na resina Glutathiona-Sefarose a proteína GST-eIF4E de *Xenopus laevis* que foi incubada com a proteína eIF4G humana marcada com ^{35}S . Como esperado, o eIF4G humano ligou-se eficientemente ao eIF4E de *Xenopus laevis* (Figura 19). O próximo passo foi incubar, nas mesmas condições, cada um dos homólogos do eIF4E (GST-*LmEIF4E1-4*) com cada um dos homólogos do eIF4G (^{35}S -*LmEIF4G1-5*). O *LmEIF4G4* ligou-se eficientemente ao *LmEIF4E3* e fracamente ao *LmEIF4E4* (Figura 20). A interação ocorreu independentemente do fato da proteína GST-*LmEIF4E3* ter apresentado um perfil de degradação específico, indicando que um dos fragmentos protéicos (oriundo da degradação) deve conter o motivo íntegro para interação com a proteína ^{35}S -*LmEIF4E3*. A proteína *LmEIF4G3* parece ligar-se a todos os homólogos do eIF4E, preferencialmente ao *LmEIF4E3*. Em relação aos demais homólogos do eIF4G (*LmEIF4G1*, *LmEIF4G2*₁₋₇₀₅ e *LmEIF4G5*) não foi detectada nenhuma interação específica quando incubados com os homólogos do eIF4E no ensaio realizado (Figura 20).

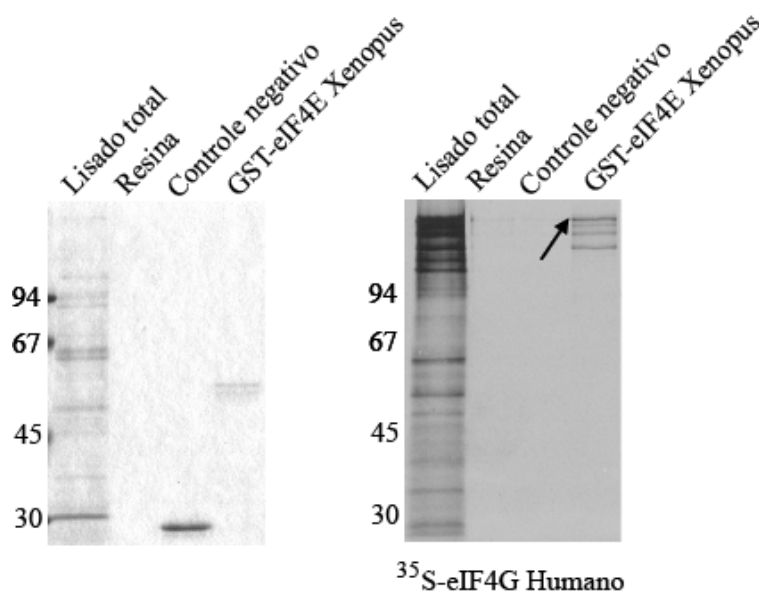


Figura 19. Análise da interação entre o eIF4E de *Xenopus laevis* e o eIF4G humano. As proteínas recombinantes GST (controle negativo) e GST-eIF4E *Xenopus* foram imobilizadas em resina Glutathiona Sefarose 4B e incubadas com o eIF4G humano marcado com ^{35}S . O painel da esquerda representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE

corado com azul de coomassie. O painel da direita representa a autoradiografia mostrando a ligação específica entre o eIF4E de *Xenopus* e o eIF4G humano (seta).

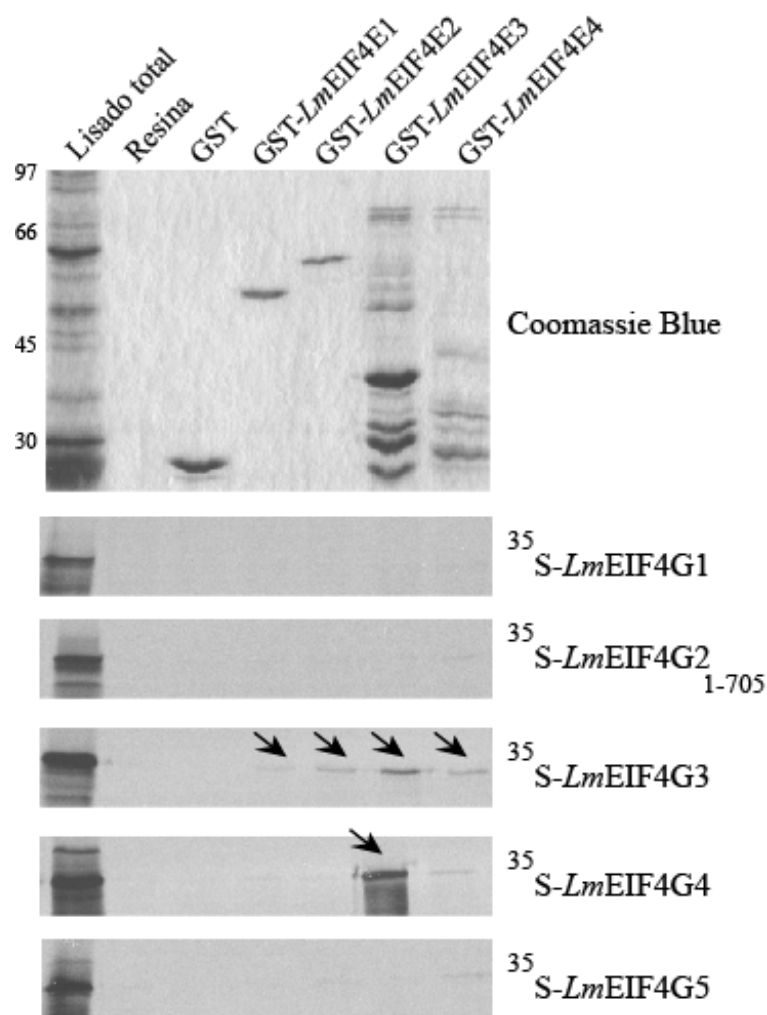


Figura 20. Análise da interação entre os homólogos ao eIF4E e os homólogos ao eIF4G de *L. major*. As proteínas recombinantes GST-*LmEIF4E1-4* foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com as proteínas *LmEIF4G1-5* marcadas com ^{35}S . O painel superior representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE e corados com azul de Coomassie. Os painéis abaixo representam as autoradiografias mostrando as ligações específicas entre os homólogos ao eIF4E e os homólogos ao eIF4G marcados com ^{35}S (setas).

Para confirmar a interação obtida entre as proteínas *LmEIF4G4* e *LmEIF4E3*, a proteína *LmEIF4G4* foi sintetizada contendo uma cauda de seis Histidinas (HIS) fusionadas à sua região C-terminal, marcada com ^{35}S , e imobilizada na resina Níquel-NTA Agarose. A

resina contendo *LmEIF4G4* marcada com ^{35}S foi em seguida incubada com a proteína *LmEIF4E3* também marcada radioativamente. Este ensaio corroborou a interação entre as proteínas *LmEIF4E3/LmEIF4G4* (Figura 21).

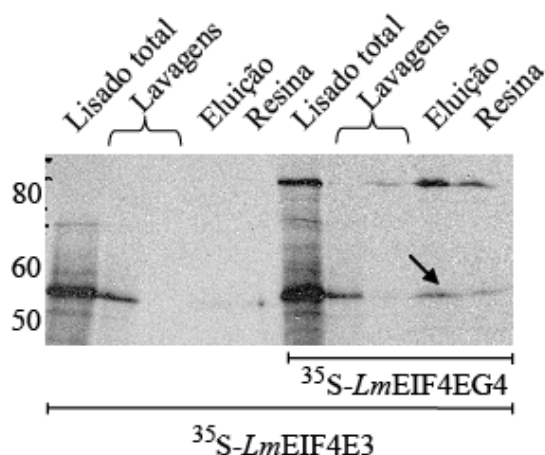


Figura 21. Análise de interação entre o homólogo *LmEIF4G4* com o homólogo *LmEIF4E3* de *L. major*. A proteína ^{35}S -LmEIF4G4-HIS foi imobilizada na resina Ni-NTA agarose, incubada com ^{35}S -LmEIF4E3 e eluída com 0,5M do competidor Imidazol. As amostras foram fracionadas em gel 15% SDS-PAGE e submetidas à autoradiografia. O painel representa a autoradiografia mostrando que ao desligar a proteína *LmEIF4G4*-HIS da resina, a proteína *LmEIF4E3* também se desliga (setas).

5.1.3. Estudo dos domínios de Interação entre os homólogos eIF4G/eIF4A e eIF4G/eIF4E e em *L. major*.

Uma vez identificadas as interações específicas entre os homólogos das subunidades do eIF4F de *L. major*, tornou-se importante estabelecer que região/domínios das proteínas eIF4G poderiam estar mediando estas interações, com seus parceiros, tentando correlacionar com semelhanças e/ou diferenças das interações já descritas em outros eucariotos. Foram escolhidas as proteínas *LmEIF4G3-4*, cujos resultados de interação às proteínas *LmEIF4AI* e *LmEIF4E3* foram considerados mais consistentes, para iniciar os ensaios de localização de domínios. Fragmentos gênicos codificando para o domínio MIF4G (*LmEIF4G3*₂₆₋₃₁₀ e *LmEIF4G4*₃₇₋₃₆₂) e as proteínas *LmEIF4G3-4* completas foram sintetizados com ^{35}S , e as

respectivas proteínas radioativas foram analisadas quanto a sua capacidade de ligar-se as proteínas recombinantes GST-*LmEIF4AI* e GST-*LmEIF4AIII* (Figura 22A). Como já descrito anteriormente, o *LmEIF4G3* liga-se ao GST-*LmEIF4AI* através de seu domínio MIF4G (Dhalia et al., 2005), e aqui é mostrado que a ligação também ocorre com a proteína completa. Já no caso do *LmEIF4G4*, embora a proteína inteira interaja com o *LmEIF4AI* (Figura 20B), o domínio MIF4G isoladamente não parece ser capaz de realizar esta interação.

Para identificar o domínio na proteína *LmEIF4G4* envolvido na ligação às proteínas *LmEIF4AI* e *LmEIF4E3*, construções plasmidais contendo as regiões codificantes para as regiões N-terminal+MIF4G (*LmEIF4G4*₁₋₃₆₂), MIF4G+C-terminal (*LmEIF4G4*₃₇₋₇₆₅) e o gene completo do *LmEIF4G4* foram utilizados para a obtenção das respectivas proteínas marcadas com ³⁵S. As proteínas obtidas foram incubadas (em ensaios de *pull down*) com as proteínas GST, GST-*LmEIF4AI* e GST-*LmEIF4E3*. A proteína GST-*LmEIF4E3* é capaz de interagir com o fragmento protéico N-terminal+MIF4G e com a proteína completa, entretanto essa interação não acontece com o domínio MIF4G+C-terminal (Figura 22B). Esse resultado indica que o motivo de ligação da proteína *LmEIF4G4* à *LmEIF4E3* se encontra na sua curta região N-terminal, ou que esta região seja necessária para que a proteína assuma o dobramento apropriado para que ocorra a ligação. Outro dado relevante, obtido neste ensaio, é que a proteína GST-*LmEIF4AI* interage com as proteínas C-terminal+MIF4G e proteína completa *LmEIF4G4*, todavia o mesmo não ocorre com o N-terminal+MIF4G (Figura 22B), corroborando o resultado da figura 22A que mostra a necessidade além do MIF4G, do segmento C-terminal para que ocorra a interação com a proteína *LmEIF4A1*. Esta região, como mostrado anteriormente, pode abrigar o domínio MA3 que também ligase ao eIF4A em mamíferos.

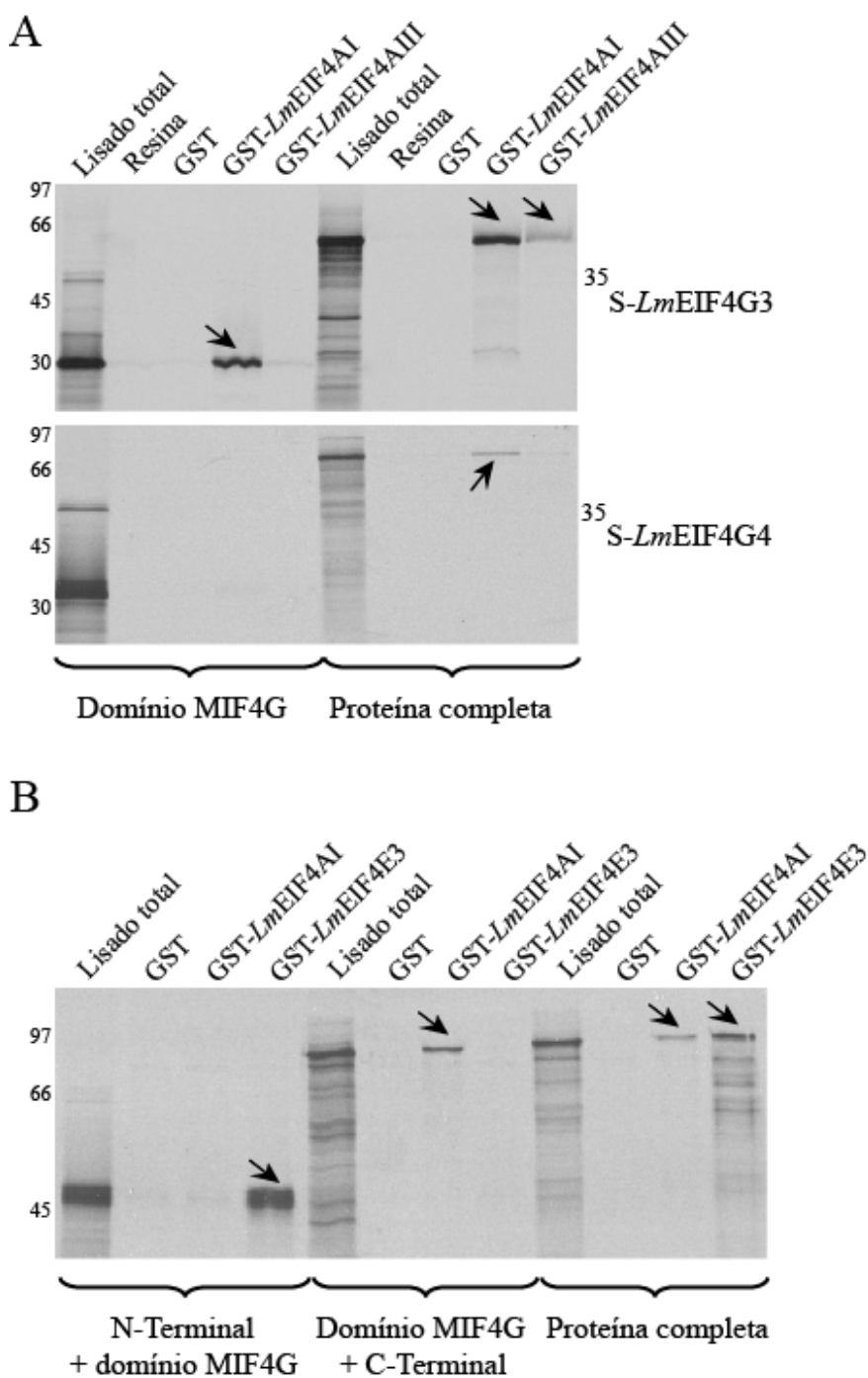


Figura 22. Análise da interação entre os homólogos do eIF4A e eIF4E com diferentes domínios dos homólogos EIF4G3-4 de *L. major*. As proteínas recombinantes fusionadas à GST (*LmEIF4AI*, *LmEIF4AIII* e *LmEIF4E3*) foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com as diferentes proteínas EIF4G3-4 marcadas com ^{35}S . (A) O domínio MIF4G do *LmEIF4G3* e as proteínas completas *LmEIF4G3-4* ligam-se à *LmEIF4AI*. A proteína *LmEIF4G3* completa também liga-se, mais fracamente, a *LmEIF4AIII* (setas). (B) A região N-terminal+MIF4G e a proteína completa *LmEIF4G4* interagem com a proteína *LmEIF4E3*, enquanto o segmento protéico C-terminal+MIF4G e a proteína inteira se ligam a proteína *LmEIF4AI*(setas).

5.1.4. Alinhamento múltiplo das seqüências *LmEIF4G3-4* e identificação de motivos conservados

Das cinco proteínas homólogas ao eIF4G, encontradas em *L. major*, as proteínas *LmEIF4G3* e *LmEIF4G4* foram as únicas capazes de ligar-se de forma satisfatória aos homólogos eIF4A e eIF4E. Além disso, elas são as únicas que apresentam homologia e arquitetura semelhante, fora do seu domínio MIF4G, com os demais homólogos de eIF4G descritos em outros eucariotos (Dhalia *et al.*, 2005). O alinhamento da figura 23 mostra vários segmentos conservados na região N-terminal, o domínio central MIF4G e uma região C-terminal com alguns trechos conservados. Dentro da região N-terminal destas proteínas, algumas seqüências se assemelham ao motivo de ligação [YXXXXLΦ (onde X pode ser qualquer aminoácido e Φ é um resíduo hidrofóbico – L, M ou F)] ao eIF4E presente no eIF4G de outros eucariotos (Mader *et al.*, 1995). Dois segmentos conservados nos homólogos EIF4G3-4 de *L.major*, assemelham-se ao motivo YX₄LΦ. Estes são os motivos **F/Y/MXXXXI/LR/L** (aminoácidos 3-9 no *LmEIF4G3* – FTVEQIR) e **FSLXXV/IV/L** (aminoácidos 23-29 no *LmEIF4G3* – FSLDEVV), com a segunda seqüência mais conservada quando comparada ao peptídeo de ligação descrito. Um terceiro possível motivo de ligação ao eIF4E é encontrado apenas na região N-terminal dos homólogos do EIF4G4, **FN/KL/PRGI/VV/T** (aminoácidos 3-9 no *LmEIF4G4* – FNLRGIV), porém assumindo que ambos homólogos EIF4G3-4 deveriam interagir com parceiros do eIF4E, usando seqüências similares, consideramos esta seqüência não comum mas potencialmente candidata a ser um motivo de ligação ao eIF4E. Outro aspecto importante, relativo à curta região N-terminal das proteínas EIF4G3-4, é que ela pode não abrigar a região de ligação à PABP, diferentemente do que tem sido mostrado para os homólogos do eIF4G de levedura e humano (Tarun *et al.*, 1997; Imataka *et al.*, 1998).

O alinhamento também destaca várias seqüências conservadas entre os homólogos EIF4G3-4, posicionados dentro da região C-terminal destas proteínas (Figura 23). Análises de predição das estruturas secundárias sugerem que estas regiões são principalmente composta de α -hélices, em contraste como os outros homólogos ao eIF4G nos tripanossomatídeos. No eIF4G de mamíferos, o C-terminal das proteínas contém dois domínios HEATs compostos de α -hélices também envolvidas em interações proteína-proteína, MA3 e W2 (Bellssolell *et al.*, 2006). Nenhuma homologia significativa, porém, foi observada entre o C-terminal do eIF4G humano e os homólogos EIF4G3-4, quando comparadas as seqüências de proteínas em análises utilizando a ferramenta BLAST. No entanto a presença de dois resíduos conservados de triptofano (W), na região C-terminal das proteínas de tripanossomatídeos, nos orientou a estudar algumas similaridades funcionais com o domínio W2 do eIF4G de mamíferos, considerando que ele também é caracterizado pela presença de dois resíduos de W (Bellssolell *et al.*, 2006). De fato estes resíduos de W, somados a aminoácidos vizinhos, assemelham-se as duas caixas aromáticas acídicas conservadas previamente observadas no C-terminal no eIF4G de mamíferos e proteínas relacionadas (Bellssolell *et al.*, 2006; Liberman *et al.*, 2008). Adicionalmente, quando tentamos encontrar resíduos conservados presentes nos domínios MA3 e W2 depositados no banco de dados do pfam, dentro das regiões equivalentes das seqüências EIF4G3-4, várias coincidências com elementos conservados dentro destas proteínas foram observadas. Assim, no contexto geral parece que os domínios MA3 e W2, divergentes em seqüência devido à grande distância evolutiva entre tripanossomatídeos e eucariotos superiores, estão presentes nos homólogos EIF4G3 e 4 (Figura 23).

LmEIF4G3	1				M	QFTVEQLRSV	RNNYLEP-PY	PGFSLDEVVR	RRRLTQT---	---KLVRGEN	AWVAKG-TAQ	TTEEWVQRL	HGTLNKLTEE	*****
TbEIF4G3	1				MH	VYITDQILEL	RSLYPEP-PY	PGFSLLEEACR	RKKQTQT---	---KLVRGPN	AWAARG-SAK	TTAEWVERLV	YGLNKLSTAA	
TcEIF4G3	1				MH	VYSIQQILEV	RSMYKDA-PY	PGFSLLEEACR	RKKMTQT---	---KLVRGAN	AWVARG-SAK	TTEEWVERLV	QGLNKLSTAA	
LmEIF4G4	1	MLFNLRGIVP	QKEEKVKM--	QMTLADILAF		RDWTAT-PE	PNFSLERVLL	TARLEKQKKA	EIPKLVTSN	GFRVRDKKDM	DASELELRV	QGLNKLSTOK	QGLNKLSTOK	
TbEIF4G4	1	MLFKPRGVTS	NDPRYAGGSR	LMSVSDLLAY		RDWTWRGPT	ESFSLAQILR	DARAADCKPV	VTEKLVMSN	GFKVKSRDSI	DSLGRGVRAV	QGLNKLSTEW	QGLNKLSTEW	
TcEIF4G4	1	MLFKPRGVTS	NDPKYAGNSR	IMTIPDLLAY		RDWTWKR-LS	DDFSLGRILF	AARAASKSPA	APEKLVMTSN	GFKVKDRDAI	APSERGVRAV	QSTLNKLSTES	QSTLNKLSTES	
MIF4G - HEAT DOMAIN														
LmEIF4G3	74	NKDIMIDKL	L	KELFATEDI		MNMVVMIRFK	KALDEPENSK	LYAGVCHSLA	LYEAKVLDRG	HK-----	-----	GE-RPQSQR	DATISTAQHE	
TbEIF4G3	75	NFNEIVSQLQ	INTTFSSDEM	LKKTVSIIFN		KALGEPENSN	VYAGLCYKLA	EYEVSLNVVQ	KQ-----	-----	-----	KEGKRLSKLR	NAVVGIAQTE	
TcEIF4G3	75	NFDDMVAKLQ	IEVLFSTKET	LNIAVRIIFK		KALDEPECCK	SYAGVCYKLA	EFEVGLTAAG	QK-----	-----	-----	AEGKKYSKLR	NAVVSIAQEE	
LmEIF4G4	98	NFDVAVVEEA	SPDLVLNPLV	LKGAVDIIFN		KALAEVPVFSG	IYAQLCORIK	VYEQDLVAEA	AENPTSELGQ	LMAAVQAGDE	DQKNVNGVRV	NALVKRCQEF	NALVKRCQEF	
TbEIF4G4	101	NFDVAVVQT	TPDIIINNEV	VKDVVRLIYE		KALMEVPVFAV	LYARMCFISV	RFYEYRRTKF	LP-----	-----	-----	GTCQVPSAVR	VAIVEKCOQM	
TcEIF4G4	100	NFDLIAKAIL	VPEIILNANV	VADVRLIYD		KALMEVPVFAV	LYARLCYMLV	RYEYDYRSTQ	SG-----	-----	-----	EVDAAKSEVR	VAIVEKCOQM	
MIF4G - HEAT DOMAIN														
LmEIF4G3	155	FRTLSKELSK	MEKSEEELE	YERSNVMRRK		RSNMRFTIGEL	YLCTALTHST	MFTVMDLTMR	CS---DKTGF	PSSSENTILLA	ALLNTIGDRL	DKN-----	DKN-----	
TbEIF4G3	157	FQNR-RNVPS	SEGLSEEEVE	QQRSSFMRRK		VANMRFTIGEL	FLHKLVSHT	MMDIINIIMQ	P---EKGGY	PASEDLEFLT	VLFTIVGKSL	DSIA-----	DSIA-----	
TcEIF4G3	157	FLSR-RKMPS	MEGLTEEEEM	LRRITFMRRK		RANMRFTIGEL	FHMKVLSCNT	MMNIITQTIMQ	EA---ERGGY	PTSEDTLEFLT	ELFTIGESL	DAVP-----	DAVP-----	
LmEIF4G4	198	HDSFVKQP--	PPRNAEAE	QLR---KRN		MANKFVIGEL	YMRSLSHRV	ILAVCSTALV	LGFIQSPKLV	TTDADLEMMV	SLMNVIGKRF	DEENAGMALNH	DEENAGMALNH	
TbEIF4G4	183	FTNAVEESK-	QTMSEEAQAE	RLR---KRN		VNNIKFAGEL	FKTLITQKI	IDRLTNERIY	E-----MV	PNDMELEVIV	NLEVVGKLY	EENK-----	EENK-----	
TcEIF4G4	182	FDGATKAVR-	ETASEEEAE	KVR---KRN		VNNIKFAGEL	FLKSLITQKI	IDRLTGEKLF	N-----IT	PNDMDLEVIVV	NLEVVGKMY	EEKH-----	EEKH-----	
MIF4G - HEAT DOMAIN														
LmEIF4G3	244	-----HKK	TLDPYFTSL	NFNKSDCPY		PPRIRFKIMD	LDDLRL-KRGW	GLPKPKPV	---	-----	-----	-----	-----KA	
TbEIF4G3	245	-----ELRP	KLDAYFKYLE	GLK-DQK-IY		PPRIRFKIMD	LTELRRDRNW	ESR-----	-----	-----	-----	-----ET	-----ET	
TcEIF4G3	246	-----ELRV	RLNGYFELLE	KLK-DQKEVY		PPRIRFKIMD	LTELRSKFNW	ERR-----	-----	-----	-----	-----AT	-----AT	
LmEIF4G4	291	VPKDAAAAAA	PHDTIWRALN	SSM--NDPRF		SRIRFLIQN	LLEWR-NDGW	KPKESPAQGS	SGGSADDSNN	TTNDRSNHR	SSNACNHSN	TGGFSGPNEQ	TGGFSGPNEQ	
TbEIF4G4	263	-----PD	AQEQLWKMLL	SMQ--ENRRY		SMIRFLIQN	LIDRR-NGGW	KPREPEQVVI	EDIQSQQPQ	QSQSQQQQQQ	QSQSQQQQQQ	QSQSQQQQQQ	QSQSQQQQQQ	
TcEIF4G4	262	-----PA	AQPALWKLLG	ELQ--EERRF		SNIRFLIQN	LTIERR-NGGW	KPREPEQVIV	EDPQ-----	QSQSQQQQQQ	QSQSQQQQQQ	QSQSQQQQQQ	QSQSQQQQQQ	
MIF4G - HEAT DOMAIN														
LmEIF4G3	296	ALPPSHGKDK	KYATAPPPKK	GGRDQGSAAA		A-----	-----	--SKSWRDA	TTTGTAAAP	APSADKGLNG	RGTSAAAPAA	NTAALSSSSA	NTAALSSSSA	
TbEIF4G3	293	VMPKTSAAATQ	K-----EPPRA	SDRAVRSSTT		P-----	-----	--SSNASSS	GSKKKGAGAP	CDALG----	-QRAAASTVG	KDAREGLKSR	KDAREGLKSR	
TcEIF4G3	295	AAPKTSSQTQ	R-----EPTRV	TDKVATRS-		-----	-----	--SNASSS	GSKKGRQDSV	SDPAS----	-QAAAAAA	KTAMGEAGGR	KTAMGEAGGR	
LmEIF4G4	388	MHRNTAPNNF	SGSGSWQQQM	QQQHSGRGP		IIGSSNAGGN	RGGTRPPSMA	DRGDEKYGRG	GPGGLRQGP	VSSFNDLSQF	PASNDHMMP	PPFPAAPPPP	PPFPAAPPPP	
TbEIF4G4	335	QQSQQQPQQQ	QQGQGHYSQ	QQRHQQSFP		HNQPQSHGGY	HQHPPPPPPP	PRNSEYQHTP	GKRGGSYHE	VSSSLGMGGR	CSQSTNDLTR	GGSYNNMPPP	GGSYNNMPPP	
TcEIF4G4	324	QQQQQQHQHQ	RYHFHHHHHQ	QQHQHQHHQ		HHQQQ-----	-----	QRNSDQLQTP	IGKRAKSYQD	VSNPYGVGAR	RSTSHLDLPG	YGAYAMPPPP	YGAYAMPPPP	
POSSIBLE MA3 DOMAIN														
LmEIF4G3	374	-----STAA	ATAPAGASFR	DEASQVPLV		KFELRVASMF	QEWVADRTND	VILHVDQFN	TCDRFFESES	ELCVAVAOEV	IHSACTTTRK	IHSACTTTRK	IHSACTTTRK	
TbEIF4G3	362	-----AWRD	VVKSEIVACD	DHEMTVTESV		AFBERVRSLF	QEWLSGYRTG	CLPNWQEEFR	DCGARIPDM	DLPTAVAAQV	VREACMTTRK	VREACMTTRK	VREACMTTRK	
TcEIF4G3	360	-----SWSN	VVKSPQSPRD	ATVARAVETV		NFBAVRSLF	QEWVLECSND	FIPELMNEFR	HCRHFDSFD	ALCKAAAEV	VREACMTTRK	VREACMTTRK	VREACMTTRK	
LmEIF4G4	488	FPGEGIMPPS	MAPQAASAPG	PQMSDEDRKL		MLNSTPPKCV	DADLSSRIIS	CIRDAHTDGD	WAAAGKAIVE	A---VQDAA	HMCRCMAVVF	IACKVTETSS	IACKVTETSS	
TbEIF4G4	434	-----AHGFS	SRLAPEERKR	LNLFARPPAQ		LEEFKKNILF	IVRDAEEGM	ADPEGMKVQL	N-ALVPQDSN	APISEVSVVV	TLIRALMDAN	TLIRALMDAN	TLIRALMDAN	
TcEIF4G4	409	PP-----P	PPTLYQAGVE	EQLSVDERKL		IVLARPPETL	TEEVNRNIFR	IARDAVEDGL	PNRDMITNDL	SRVLGNES	SQSLMTSVV	ILLRALMDPK	ILLRALMDPK	
POSSIBLE MA3 DOMAIN														
LmEIF4G3	459	DAOREAFSFL	VVGLYMDTE	VFDGFASALA		SATEDGLLED	VPKFGERFMS	MLRLTSTEQ-	-ETRADVYFD	AANVLRITYN	RLESPDSSAV	DTLMSFWRV	DTLMSFWRV	
TbEIF4G3	447	DAQCEASRLM	VIGLFLEDDQ	VFNFGAAALS		SAIEEGILED	VPKFSERFAN	MLRFTSGD-	-DVKTDDVYD	AVRVLCTAYG	MLRDPDEMVL	STLMEFVGKI	STLMEFVGKI	
TcEIF4G3	445	EAQREASSFL	VLDGFALALV	VLDGFALALV		SAIEEGILED	VPKFSERFAN	MLRFTSGE-	-NTVADVYD	TARVLCCTAYG	LMREPDEMVL	DTMMEFWEKI	DTMMEFWEKI	
LmEIF4G4	585	EADRELFMNS	LNSGMFDIKE	LTRGYSWCLT		SAYAYNVKED	FPKVVSRFVA	CVTATKSMDF	LSMVSVMAR	TANYLDALYV	PLEGAQMEWE	EDFLEWTSV	EDFLEWTSV	
TbEIF4G4	520	ESERKLLLSV	LQKGSFENHI	LTRGFSWALA		KIISDRDCDD	CPRIYARFAE	AVWSIPSNF	RCVTKDIVSR	TAIHLGALSV	EYE-SEGWE	EDFIWVENI	EDFIWVENI	
TcEIF4G4	502	ESERTLCTFA	LESGNWERSI	LGRGFANCLT		KIISAEERDAD	CPRIYSRFDV	VVTRVTELFN	LCVTKDIVSR	TARYLDVLQV	VVD-ETEEWE	EDFIWVENI	EDFIWVENI	
POSSIBLE W2 DOMAIN														
LmEIF4G3	557	PLP-STDEEN	MIRMDLDVVY	SLCNPEGVDQ		GLEKLLSRIT	HSLQMLQMD	AEVLDEFCLC	DVEDGLCAKV	IADYKERFPK				635
TbEIF4G3	544	PPP-EKKED-	-AIFSLPVVQ	SLVTMTKL--		GRAQITGHII	SSLHANGLVD	DATLHEMLGT	PGGE-VDAEV	KDAFRKATVG	KGTLT			622
TcEIF4G3	542	PRP-ATDEN-	-AVLPLQVVQ	SLVEMSTR--		GQAPLTGRII	ASLHSICLVH	DETREWELAT	PHEE-TSAEV	VEAFKANK				614
LmEIF4G4	685	LKGWREAHGP	D-KSTGSDIL	NCIASIKQRP		FMKDIVSDFM	MELSQQGYLD	DAETKTWIRE	NRTNEKYKVF	VDGMEQMPD	TA			765
TbEIF4G4	619	VTASETRGQ	ESKPSVADMM	ESIAPSCTVP		FMCNVLPDFV	AVMVQARFFT	EEELGAWRNK	NKDNSKVFSL	LEELSVIYA				697
TcEIF4G4	601	LVA---RQA	PNKVTAESM	ESMSFTRLGS		FLRGILPDFV	ASMVQARFFT	EDELQAWRNK	NESNPKFRAI	TEELAQLYP				675

Figura 23. Alinhamento múltiplo, por Clustal W, comparando as seqüências dos homólogos *LmEIF4G3* e *LmEIF4G4* com seus ortólogos em *T. brucei* e *T. cruzi*. Notar a presença dos motivos M/F/YXXXXI/LL e FSLXXXX na porção N-terminal (setas simples), que poderiam constitui um putativo motivo de ligação à subunidade eIF4E. A região de conservação indicada pelos asteriscos é referente ao domínio conservado MIF4G, comum aos cinco homólogos identificados, e de duas regiões candidatas aos domínios MA3 e W2 identificadas apenas nos EIF4G3-4 nos Tryps.

5.1.5. Estudo da interação entre mutantes *LmEIF4G3-4* com homólogos das subunidades do eIF4F de *L. major*

Para localizar os motivos responsáveis pela interação entre as proteínas *LmEIF4G3* e *LmEIF4G4* e homólogos das demais subunidades do eIF4F em *L. major*, foram realizados ensaios de mutagênese sítio-dirigida nos respectivos genes codificantes. A partir do alinhamento das seqüências dos aminoácidos destas proteínas em *L. major*, *T. brucei* e *T. cruzi*, e de dados de outros modelos de eucariotos melhor caracterizados como o homem e levedura, foram escolhidos os seguintes motivos conservados de aminoácidos nas proteínas para serem modificados: FSL/AAA (*LmEIF4G3-4*), KLV/AAA (*LmEIF4G3-4*), LNK/AAA (*LmEIF4G3-4*), M20Y/S/D/A (*LmEIF4G4*) e IL/AA (*LmEIF4G4*). O motivo FSL (FSLXXVL) das proteínas *LmEIF4G3-4* foi escolhido por se assemelhar ao motivo de interação YX₄LΦ, descrito em vários organismos (Mader *et al.*, 1995). A seqüência KLV foi escolhida por ser conservada em todas as seqüências de tripanossomatídeos, na curta região N-terminal, e o motivo LNK foi escolhido como alvo porque a mudança do aminoácido equivalente da L (L729A), no modelo humano, reduz a interação com o eIF4A (Imataka *et al.*, 1997). As modificações M20 por Y, S, D, ou A e IL/AA estão dentro de outro provável motivo de ligação a eIF4E, e foram escolhidos após experimentos preliminares. As proteínas GST-*LmEIF4AI*, GST-*LmEIF4E3* e GST-*LmPABP1* foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com as proteínas *LmEIF4G3* e *LmEIF4G3*_{LNK/AAA}, marcadas radioativamente com ³⁵S. Como esperado a proteína *LmEIF4G3* foi capaz de se ligar ao *LmEIF4AI*, entretanto a proteína *LmEIF4G3*_{LNK/AAA} teve sua capacidade de se ligar ao *LmEIF4AI* praticamente abolida (Figura 24). As proteínas selvagem e mutante ligam-se da mesma forma ao *LmEIF4E3* e *LmPABP1*, indicando que o

modo de interação a estas duas proteínas não foi alterado pela mutação LNK/AAA (Figura 24). Em um segundo momento, a interação entre os demais mutantes do *LmEIF4G4* e o *LmEIF4E3* foi melhor investigada. A maioria das variantes da proteína *LmEIF4G4* ligam-se ao *LmEIF4E3*, de uma forma equivalente à proteína selvagem (Figura 25). Apenas a proteína mutante IL/AA têm sua capacidade de associação ao GST-*LmEIF4E3* drasticamente reduzida, sem prejuízo de sua interação com o GST-*LmEIF4AI*. Estes resultados indicam que os aminoácidos mutados (LNK/AAA e IL/AA) participam nas interações entre o *LmEIF4G4* e os parceiros *LmEIF4AI* (Figura 24) e *LmEIF4E3* (Figura 23), respectivamente.

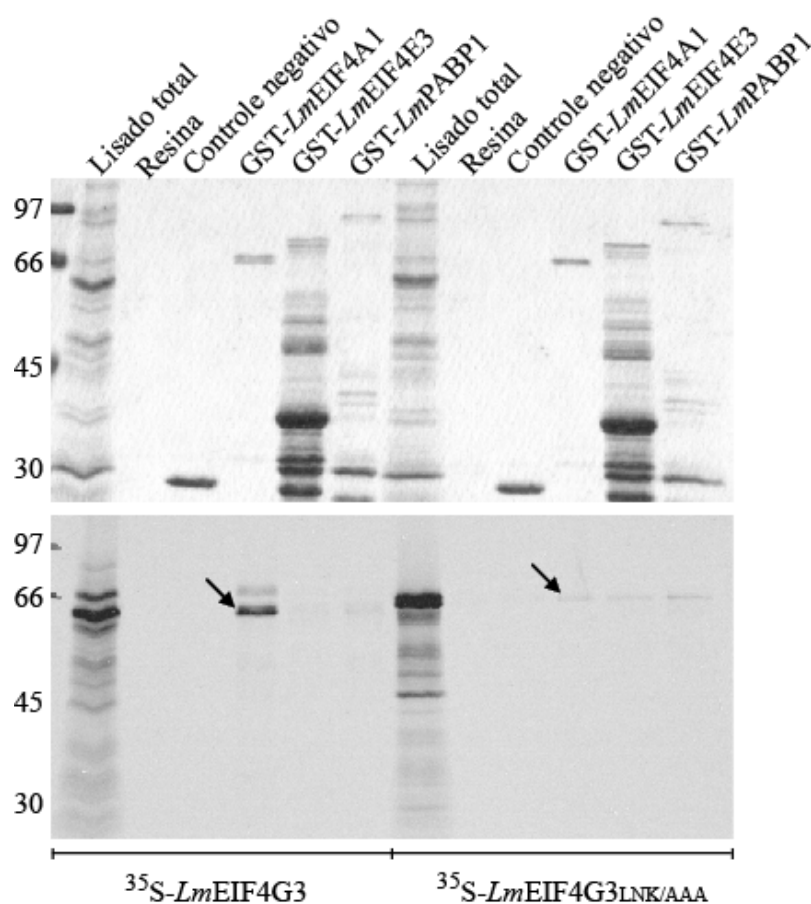


Figura 24. Análise da interação entre a proteína mutante *LmEIF4G3*_{LNK/AAA} com a proteína *LmEIF4AI* selvagem. As proteínas recombinantes fusionadas à GST (*LmEIF4AI*, *LmEIF4E3* e *LmPABP1*) foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com as proteínas *LmEIF4G3* e *LmEIF4G3*_{LNK/AAA}, marcadas radioativamente. O painel de cima representa as proteínas recombinantes fracionados em gel 15% SDS-PAGE corado com azul de Coomassie. O painel inferior representa a autoradiografia, deste mesmo gel, evidenciando as interações entre as proteínas *LmEIF4G3* selvagem e mutante marcadas com ³⁵S e as proteínas GST-*LmEIF4EAI*, GST-*LmEIF4E3* e GST-*LmPABP1*. As setas indicam a interação *LmEIF4G3*/*LmEIF4AI* que é afetada pela referida mutação no *LmEIF4G3*.

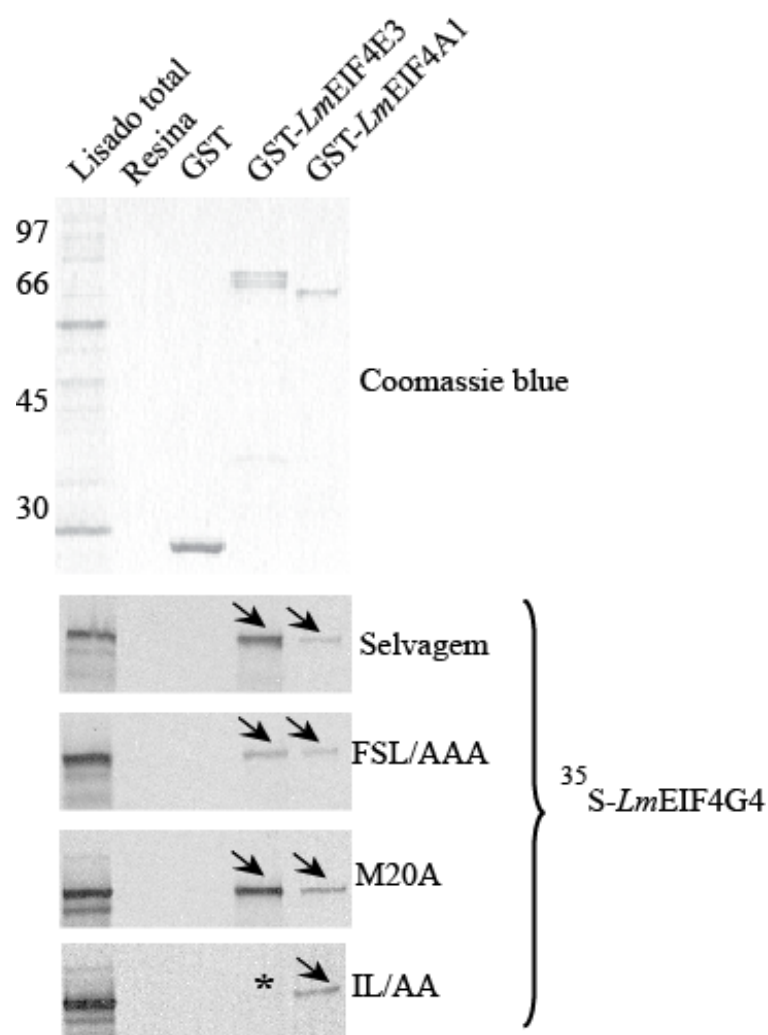


Figura 25. Análise da interação entre mutantes da proteína *LmEIF4G4* com as proteínas *LmEIF4AI* e *LmEIF4E3*. As proteínas recombinantes fusionadas a GST (*LmEIF4AI*, *LmEIF4E3*) imobilizadas na resina Glutathione Sefarose 4B foram incubadas com as diferentes proteínas *LmEIF4G4* (Selvagem, FSL/AAA, M20A, IL/AA), marcadas radioativamente. O painel superior representa as proteínas recombinantes fracionadas em 15% gel SDS-PAGE e corados com azul de Coomassie. O painel inferior representa a autoradiografia, deste mesmo gel, evidenciando as interações observadas entre as diferentes proteínas *LmEIF4G4* com as proteínas GST-*LmEIF4AI* e GST-*LmEIF4E3* (setas), exceto para a mutação IL/AA que abole a interação *LmEIF4G4*/*LmEIF4E3* (asterisco).

5.1.6. Estudo reverso da interação entre os mutantes *LmEIF4G3-4* com homólogos das subunidades do eIF4F de *L. major*

Para confirmar a ligação específica entre homólogos do eIF4G e as demais subunidades do eIF4F de *L. major*, foi estabelecido um ensaio reverso a partir da expressão em *E. coli* de todas as variantes das proteínas *LmEIF4G3* e *LmEIF4G4* (proteína completa, regiões parciais e proteínas mutadas) em fusão com GST. Estas proteínas recombinantes foram incubadas com as demais subunidades do complexo eIF4F de *L. major*, marcadas com ^{35}S , em ensaios de *pull-down*. No primeiro ensaio, imobilizamos as variantes da proteína *LmEIF4G3* na resina Glutathione Sefarose 4B e em seguida incubamos com *LmEIF4AI* marcado radioativamente. Duas variantes da proteína *LmEIF4G3* [(N-terminal+MIF4G (*LmEIF4G3*₁₋₃₁₀) e MIF4G+C-terminal (*LmEIF4G3*₂₆₋₆₃₅)] e proteína *LmEIF4G3* completa interagem com o *LmEIF4AI*. Não ocorreu interação das proteínas mutantes do *LmEIF4G3* (FSL/AAA e LNK/AAA) com *LmEIF4AI* (Figura 26). Comparando os resultados reversos, com os resultados diretos, é possível que a proteína GST-*LmEIF4G3* recombinante não se ligue tão bem ao ^{35}S -*LmEIF4AI* como a proteína GST-*LmEIF4AI* se liga ao ^{35}S -*LmEIF4G3*, entretanto a mutação dos aminoácidos LNK/AAA no *LmEIF4G3*, em ambos os ensaios, abolem a interação ao *LmEIF4AI* (Figura 24 e 26). Já a mudança dos aminoácidos FSL/AAA, pode induzir alteração na conformação de proteína GST-*LmEIF4G3*_{FSL/AAA}, justificando sua não interação à proteína *LmEIF4AI* (Figura 26).

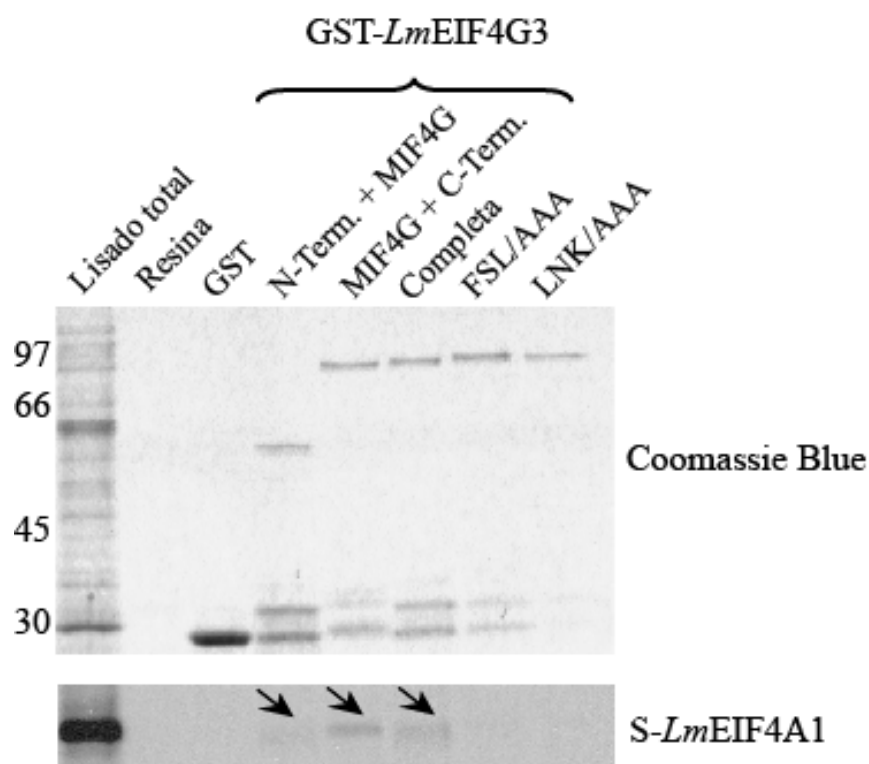


Figura 26. Análise da interação entre a proteína *LmEIF4AI* e as diferentes proteínas *LmEIF4G3*. As diferentes proteínas recombinantes *LmEIF4G3* fusionadas a GST (N-terminal+MIF4G, MIF4G+C-terminal, completa, FSL/AAA e LNK/AAA), foram imobilizadas na resina Glutathione Sepharose 4B e incubadas com a proteína ^{35}S -*LmEIF4AI*. O painel superior representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE corado com azul de Coomassie. O painel inferior representa a autoradiografia mostrando as ligações específicas entre as proteínas N-terminal+MIF4G (mais fracamente), MIF4G+C-terminal e *LmEIF4G3* completa com a proteína *LmEIF4AI* (setas).

Em seguida, testamos a capacidade das proteínas GST-*LmEIF4G3-4* interagirem com os diferentes homólogos *LmEIF4E1-4*. Confirmamos a interação específica entre a proteína GST-*LmEIF4G4* e a proteína ^{35}S -*LmEIF4E3*, como já esperado, e verificamos ainda a interação específica da proteína GST-*LmEIF4G3* com ^{35}S -*LmEIF4E4* (Figura 27). Como descrito anteriormente, a proteína ^{35}S -*LmEIF4G3* foi capaz de interagir com todos os homólogos *LmEIF4E1-4* numa afinidade baixa (Figura 20), todavia o resultado obtido no

ensaio reverso sinaliza para uma interação específica apenas entre as proteínas *LmEIF4G3* e *LmEIF4E4* (Figura 27). Para estudar mais detalhadamente a interação *LmEIF4G3/LmEIF4E4*, as variantes da proteína *LmEIF4G3* (em fusão com GST) foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com a proteína ^{35}S -*LmEIF4E4*. Neste ensaio foi confirmada a interação entre a proteína *LmEIF4E4* e proteínas N-terminal+MIF4G, completa e mutante LNK/AAA, embora não tenha ocorrido interação com MIF4G+C-terminal e a proteína mutante FSL/AAA (Figura 28). Estes resultados demonstram que o motivo de ligação no *LmEIF4G3* à proteína *LmEIF4E4* está localizado em sua curta região N-terminal (apenas 26 aminoácidos) e envolve um ou mais aminoácidos da trinca FSL, pois quando FSL é modificada para AAA a interação *LmEIF4G3/LmEIF4E4* é abolida (Figura 28).

Visando a identificação de domínios de interação e aminoácidos envolvidos na ligação da proteína *LmEIF4G4* aos parceiros *LmEIF4E3* e *LmEIF4AI*, diferentes regiões/ou mutações da proteína *LmEIF4G4* foram expressas fusionadas à GST (N-terminal+MIF4G, MIF4G+C-terminal, completa, e os mutantes FSL/AAA, KLV/AAA, LNK/AAA e IL/AA) e testadas quanto a sua capacidade de interagir com ^{35}S -*LmEIF4E3* (Figura 29) e ^{35}S -*LmEIF4A1* (dados não mostrados). Foi confirmado que a proteína *LmEIF4E3* interage especificamente com a maioria das variedades da proteína *LmEIF4G4*, exceto com as proteínas MIF4G+C-terminal e mutante IL/AA (Figura 29). Este resultado confirma que o motivo de interação com *LmEIF4E3* está presente na curta região N-terminal da proteína *LmEIF4G4*, e que a modificação da IL/AA abole a interação *LmEIF4G4/LmEIF4E3* (Figura 29). Com relação ao estudo da interação entre as variantes de *LmEIF4G4* com a proteína *LmEIF4AI*, nenhuma interação foi detectada, provavelmente devido a baixa afinidade do

LmEIF4G4 pelo *LmEIF4A1*, (quando comparada a afinidade do *LmEIF4G3* pelo *LmEIF4A1*).

Estes resultados obtidos sinalizam para a presença de pelo menos dois possíveis complexos eIF4F em *L. major*. Um dos complexos seria formado pelas proteínas *LmEIF4E4*, *LmEIF4G3* e *LmEIF4A1*, e o outro seria formado pelas proteínas *LmEIF4E3*, *LmEIF4G4* e *LmEIF4A1*. Estes complexos apresentam propriedades diferentes, quanto a ligação de suas subunidades: a ligação ao *LmEIF4A1* é dependente da região C-terminal do *LmEIF4G4*, diferentemente do *LmEIF4G3* que necessita apenas do MIF4G. Além disso, apresentam diferentes motivos de ligação aos seus parceiros eIF4E. O *LmEIF4E4* se associa ao *LmEIF4G3* através do trio de aminoácidos FSL, dentro do contexto FSLXXV/IV/L. Enquanto a proteína *LmEIF4E3* liga-se à região envolvendo os aminoácidos hidrofóbicos IL, do motivo (F/Y/MXXXXI/LR/L), presente na região N-terminal do *LmEIF4G4*. Este modo diferencial de interação sinaliza para mecanismos únicos, encontrados nestes protozoários, que precisam ser melhor investigados. Atualmente, os motivos de ligação equivalentes nos ortólogos EIF4G3-4 de *T.brucei* foram modificados por mutações sítio-dirigida, e o efeito destas mutações na interação entre os homólogos ao eIF4F de *T. brucei*, estão sendo investigados (Danielle do Nascimento Moura, 2009, dados não publicados).

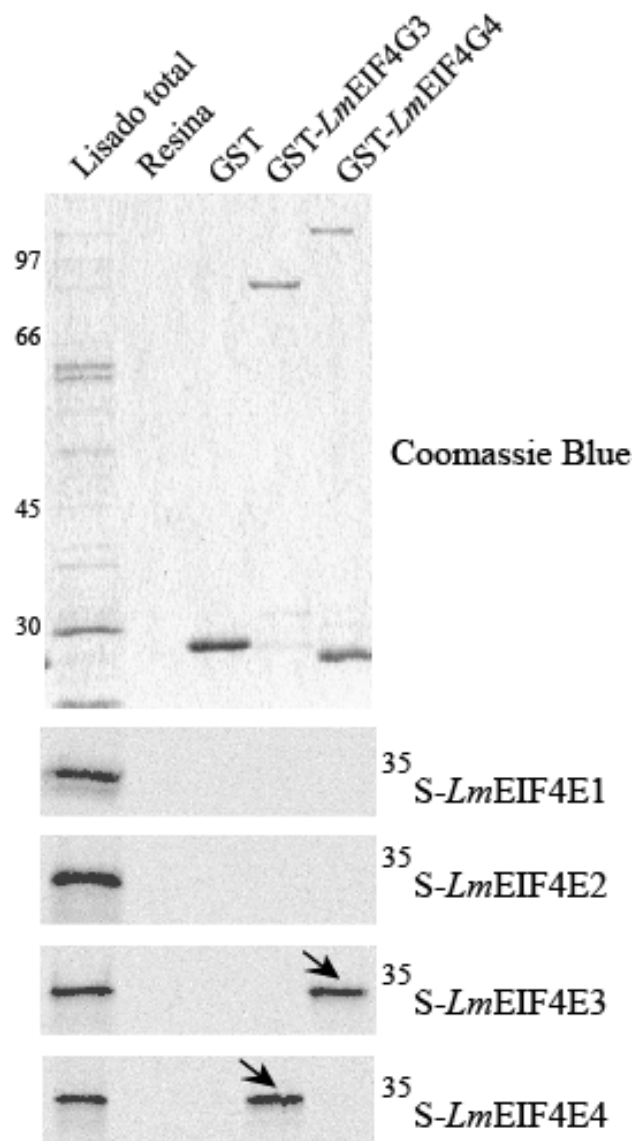


Figura 27. Análise da interação entre as proteínas *LmEIF4G3-4* e as proteínas *LmEIF4E1-4* (ensaio reverso). As proteínas recombinantes GST (controle negativo), GST-*LmEIF4G3* e GST-*LmEIF4G4* foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com as proteínas *LmEIF4E1-4*, marcadas radioativamente com ³⁵S. As autoradiografias demonstram as interações específicas entre os pares de proteínas *LmEIF4G3/LmEIF4E4* e *LmEIF4G4/LmEIF4E3* (setas).

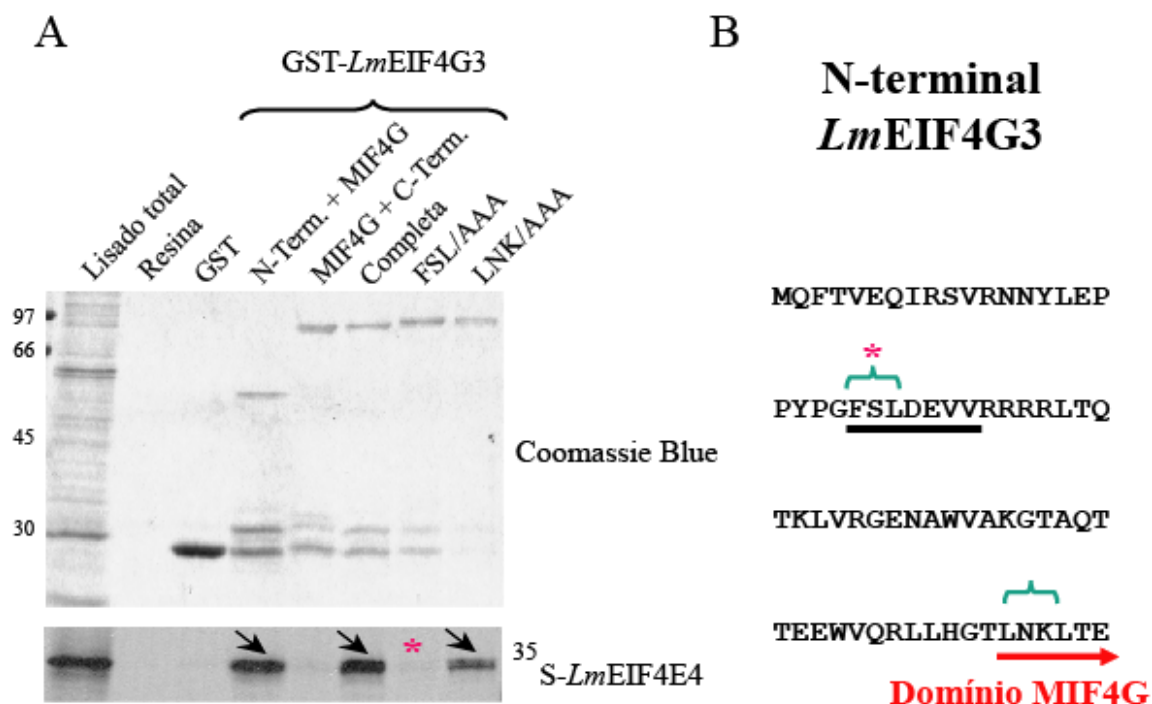


Figura 28. Análise da interação entre a proteína *LmEIF4E4* e as diferentes proteínas *LmEIF4G3*. As proteínas recombinantes *LmEIF4G3* fusionadas a GST (N-terminal+MIF4G, MIF4G+C-terminal, completa e mutantes FSL/AAA e LNK/AAA), foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com a proteína ³⁵S-*LmEIF4E4*. (A) O painel superior representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE corado com azul de Coomassie. O painel inferior representa a autoradiografia demonstrando ligações específicas das proteínas N-terminal+MIF4G, completa, e mutante LNK/AAA com a proteína *LmEIF4E4* (setas). Entretanto não ocorre ligação entre a proteína *LmEIF4E4* e as proteínas C-terminal +MIF4G e mutante FSL/AAA (asterisco). (B) Representação esquemática da região N-terminal da proteína *LmEIF4G3*, indicando que o motivo de ligação ao *LmEIF4E4* está presente na região N-terminal e parece envolver pelo menos um dos aminoácidos da trinca FSL (asterisco).

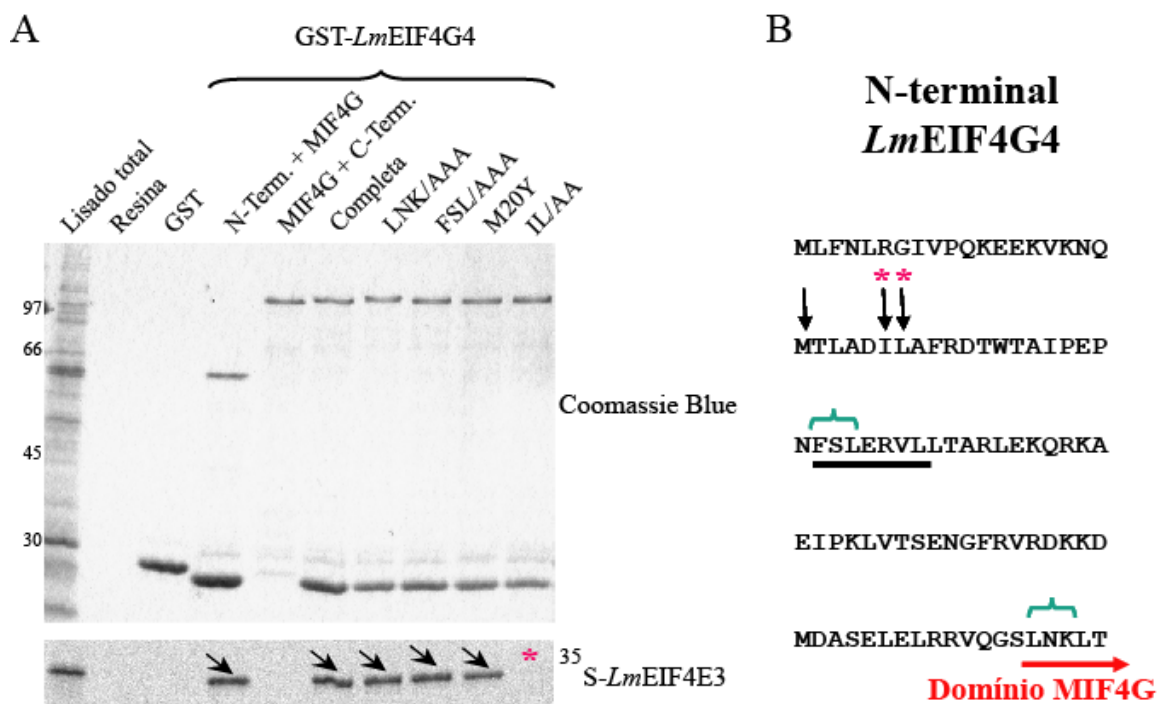


Figura 29. Análise da interação entre a proteína *LmEIF4E3* e as diferentes proteínas *LmEIF4G4*. As proteínas recombinantes *LmEIF4G4* fusionadas a GST (N-terminal+MIF4G, C-terminal+MIF4G, completa, e mutantes LNK/AAA, FSL/AAA, M20Y e IL/AA) foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com a proteína ³⁵S-*LmEIF4E3*. (A) O painel superior representa as proteínas recombinantes fracionados em gel 15% SDS-PAGE corado com azul de Coomassie. O painel inferior representa a autoradiografia demonstrando ligações específicas de variantes da proteína *LmEIF4G4* com a proteína *LmEIF4E3* (setas). Neste ensaio, também é evidenciada a ausência de ligação entre proteínas C-terminal +MIF4G e mutante IL/AA (asterisco). (B) Representação esquemática da região N-terminal da proteína *LmEIF4G4*, indicando que o motivo de ligação ao *LmEIF4E3* está presente na região N-terminal e parece envolver pelo menos um dos aminoácidos IL (asterisco).

5.1.7 Estudo da interação entre os homólogos de eIF4G e de PABP de *L. major*

A PABP (proteína de ligação ao poli-A) é uma das principais parceiras do eIF4G eucariótico e a interação entre ambas pelo menos *in vitro* media a circularização do mRNA, durante o processo de iniciação da tradução, além de participar na própria regulação da tradução. Dados experimentais, obtidos por colaboradores em nosso laboratório, indicam a presença de 3 homólogos à PABP *L. major*. Destes apenas a *LmPABP1* parece ser a PABP envolvida na tradução em *L. major*, pois além de ser fosforilada, é única que liga-se especificamente a seqüências de RNA ricas em A. Esses resultados, mais detalhados, estão sendo compilados no manuscrito “**Functional Characterization of three different *Leishmania* PABP homologues with distinct RNA binding properties**”, em fase final de redação.

5.1.7.1 Estudo da interação entre os homólogos do eIF4G e da PABP de *L. major*

Tendo em vista a disponibilidade dos genes homólogos à PABP de *Leishmania*, já clonados no vetor de expressão pGEX4T3 em nosso laboratório, decidimos então averiguar a possibilidade de interação entre os homólogos do eIF4G e da PABP de *L. major*. Como controle, a proteína recombinante GST-PABP humana foi imobilizada na resina Glutathione Sefarose 4B e incubada com a proteína eIF4G humana, marcada radioativamente como descrito previamente (Melo *et al.*, 2003). Como esperado, o eIF4G humano ligou-se eficientemente à PABP humana (Figura 30). Inicialmente, nas mesmas condições, incubamos dois dos homólogos da PABP de *L. major* (GST-*LmPABP1-2*) com as proteínas *LmEIF4G1-5* marcadas com ³⁵S. As proteínas *LmEIF4G1* e *LmEIF4G4* ligam-se mais

fracamente às proteínas *LmPABP1-2* quando comparados com a proteína *LmEIF4G3* (Figura 31). Não foi detectada interação entre as proteínas *LmPABP1-2* e as proteínas *LmEIF4G2* (resíduos 1-705) e *LmEIF4G5* (dados não mostrados). Para identificarmos qual região da proteína *LmEIF4G3* estaria envolvida na ligação com os dois homólogos da PABP, foram sintetizadas algumas variantes da proteína *LmEIF4G3* marcadas com ^{35}S [N-terminal+MIF4G (*LmEIF4G3*₁₋₃₁₀), MIF4G (*LmEIF4G3*₂₆₋₃₁₀) e a proteína completa *LmEIF4G3*] e estas foram incubadas com as proteínas GST-*LmPABP1-2*. Todas as diferentes regiões analisadas são capazes de mediar a interação com os homólogos da PABP (Figura 31), indicando que provavelmente a região de ligação esteja presente no domínio MIF4G, diferente de outros eucariotos onde o domínio de ligação à PABP tem localização N-terminal. Em outro ensaio, as três PABPs de *L. major*, GST-*LmPABP1-3*, foram em seguida incubadas com a proteína ^{35}S - *LmEIF4G3* na presença das caudas poli-A ou poli-C sintéticas. Como controle, o ensaio também foi repetido na ausência de caudas sintéticas. Na presença da caudas poli-C não há diferença significativa de interação em relação ao controle sem ambas as caudas. Entretanto, na presença das caudas poli-A, ao contrário do que se esperava, houve uma diminuição da afinidade das três *LmPABPs* pela proteína *LmEIF4G3* (Figura 32).

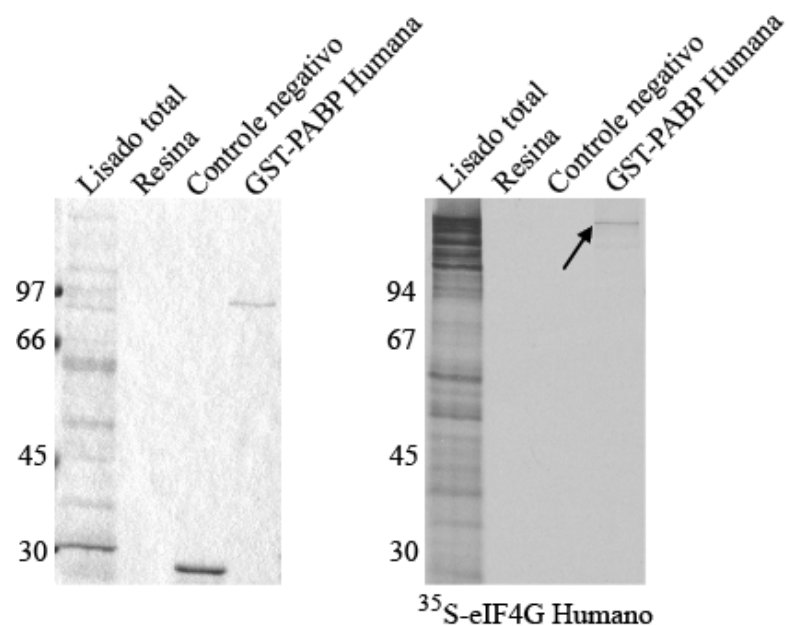


Figura 30. Análise da interação entre a PABP e o eIF4G humanos. As proteínas recombinantes GST e GST-PABP humana foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com o eIF4G humano, marcado com ^{35}S . O painel da esquerda representa as proteínas recombinantes fracionados em gel 15% SDS-PAGE e corado com azul de Coomassie. O painel da direita representa a autoradiografia mostrando a ligação específica entre a PABP e eIF4G humanos (seta).

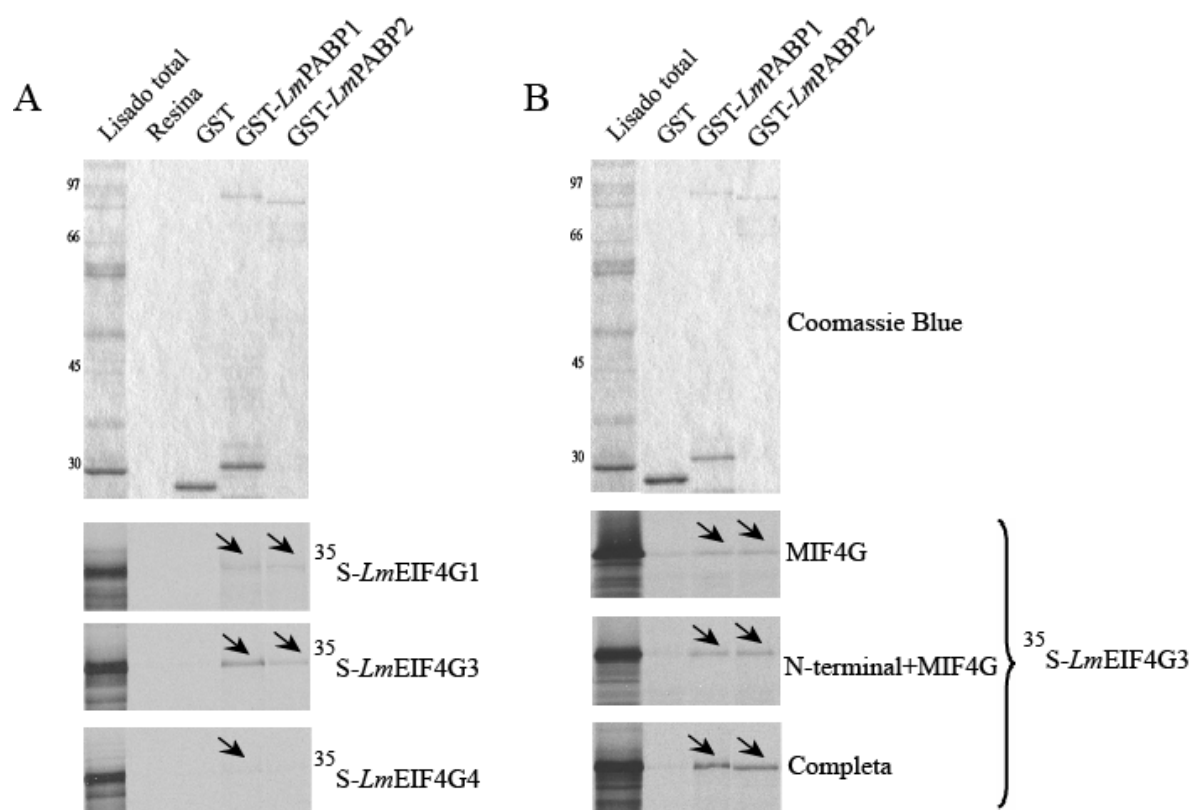


Figura 31. Análise da interação entre os homólogos PABP1-2 com homólogos do eIF4G de *L. major*. As proteínas recombinantes GST-*LmPABP1-2* foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com variantes da proteína *LmEIF4G*, marcadas com ³⁵S. Os painéis superiores representam as proteínas recombinantes fracionados em gel 15% SDS-PAGE corado com azul de Coomassie. A-Autoradiografia demonstrando as interações das proteínas *LmEIF4G1*, *LmEIF4G3* e *LmEIF4G4*, marcadas com ³⁵S, com as proteínas *LmPABP1-2* (setas). B-Autoradiografia demonstrando a interação das variantes do *LmEIF4G3*, marcadas com ³⁵S, com as proteínas *LmPABP1-2*.

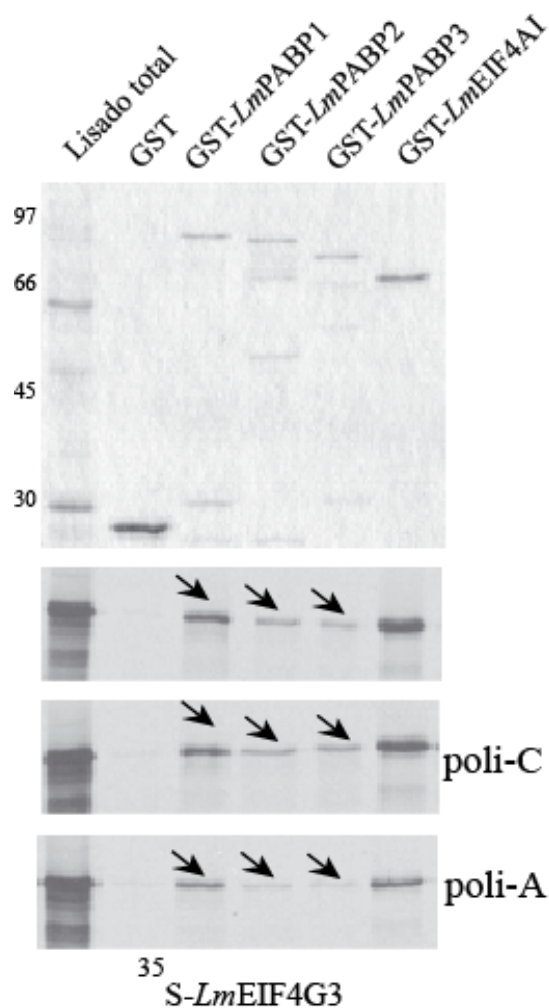


Figura 32. Análise da interação entre os homólogos da PABP e o homólogo *LmEIF4G3* de *L. major* na presença de caudas de poli-C e poli-A sintéticos. As proteínas recombinantes GST-*LmPABP1-3* foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com a proteína *LmEIF4G3*, marcada com ^{35}S . O painel superior representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE e corado com azul de Coomassie. Os painéis inferiores representam as autoradiografias demonstrando as interações observadas entre o ^{35}S -*LmEIF4G3* e os homólogos da PABP de *L. major* (setas). Ao contrário do esperado, nos ensaios onde a cauda poli-A foi adicionada a afinidade do *LmEIF4G3* pelas PABPs foi reduzida. Como controle positivo para validação do ensaio, utilizamos a interação entre o GST- *LmEIF4AI* com *LmEIF4G3*.

5.1.7.2 Estudo reverso da interação entre homólogos do eIF4G e da PABP de *L. major*

Para o ensaio reverso da interação, as proteínas GST-*LmEIF4G3* e GST-*LmEIF4G4* foram incubadas com as três proteínas *LmPABP* radioativas. Foram identificadas interações apenas das proteínas *LmEIF4G3* e *LmEIF4G4* à proteína *LmPABP1* (Figura 33). Para identificar a região responsável pela interação *LmEIF4G3/LmPABP1* e comparar com resultados obtidos no ensaio direto (Figura 31), variantes da proteína *LmEIF4G3* (N-terminal+MIF4G, C-terminal+MIF4G, proteína completa, mutantes FLS/AAA e LNK/AAA) fusionadas a GST foram testadas quanto a sua capacidade de interagir com a proteína *LmPABP1* (Figura 34). Todas as variedades da proteína *LmEIF4G3* apresentaram a capacidade de interagir com *LmPABP1*, o que pode significar numa análise preliminar que o domínio MIF4G da proteína *LmEIF4G3* é responsável pela ligação à proteína *LmPABP1* (Figura 34), corroborando os resultados obtidos previamente (Figura 31). Uma contribuição da região C-terminal não pode ser descartada, já que as diferentes variantes contendo essa região interagem mais fortemente com a proteína *LmPABP1* (Figura 34).

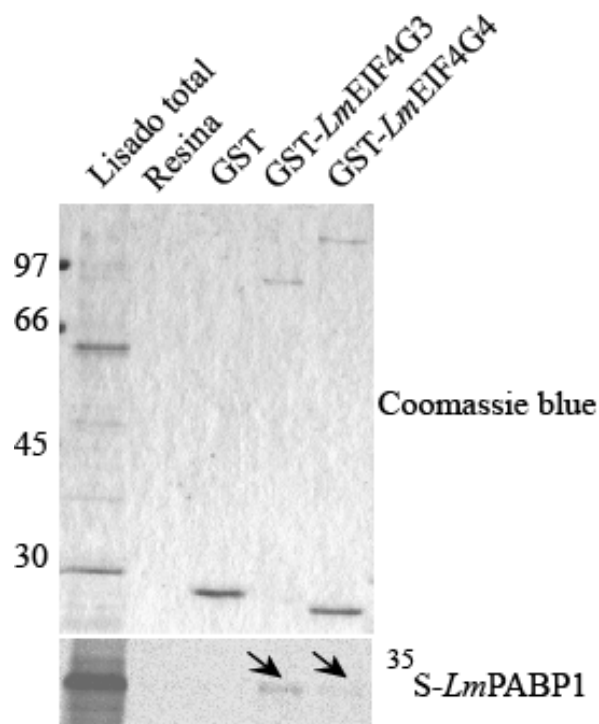


Figura 33. Análise da interação entre as proteínas *LmEIF4G3-4* com *LmPABP1*. As proteínas recombinantes *LmEIF4G3* e *LmEIF4G4* fusionadas à GST foram imobilizadas na resina Glutathiona Sefarose 4B e incubadas com a proteína ^{35}S -*LmPABP1*. O painel superior representa as proteínas recombinantes fracionadas em gel 15% SDS-PAGE e corado com azul de Coomassie. O painel inferior representa a autoradiografia mostrando ligação específica das proteínas *LmEIF4G3-4* com a proteína *LmPABP1* (setas).

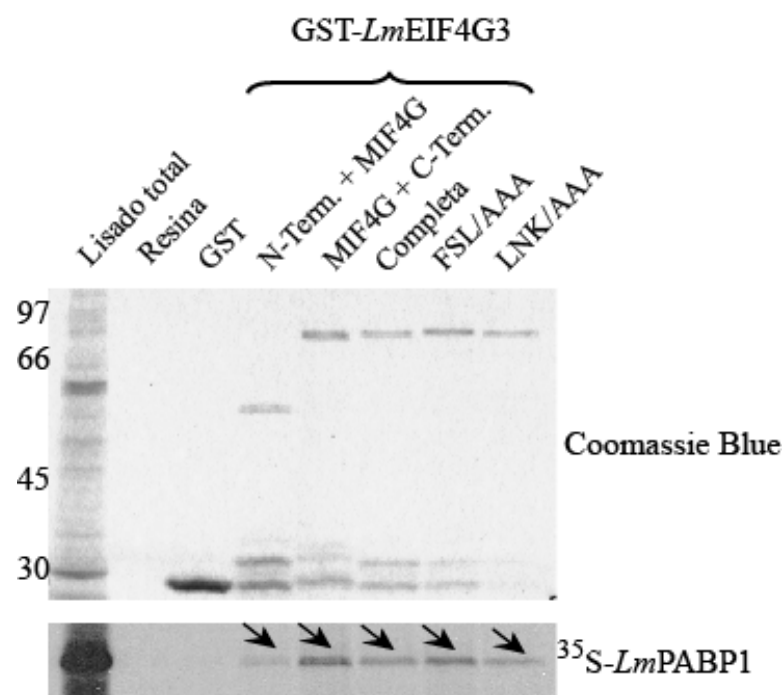


Figura 34. Análise da interação entre a proteína *LmPABP1* e as diferentes proteínas *LmEIF4G3*. As diferentes proteínas recombinantes *LmEIF4G3* fusionadas à GST (N-terminal+MIF4G, C-terminal+MIF4G, proteína inteira, FSL/AAA e LNK/AAA), foram imobilizadas na resina Glutathione Sepharose 4B e incubadas com a proteína ^{35}S -*LmPABP1*. O painel superior representa as proteínas recombinantes fracionadas em SDS-PAGE 15% e corado com azul de Coomassie. O painel inferior representa a autoradiografia mostrando ligação específica das diferentes proteínas *LmEIF4G3* com a proteína *LmPABP1* (setas).

5.1.8 Formação *in vitro* do complexo eIF4F de *L. major*

Como todas as interações entre os homólogos do eIF4F identificadas em ensaios de *pull-down*, foram utilizados pares de proteínas, decidimos reconstituir os dois prováveis complexos eIF4F de *L. major*, *in vitro*, utilizando todos os parceiros que ligam-se aos homólogos *LmEIF4G3* ou *LmEIF4G4*. Com esse experimento buscamos também avaliar se na presença de parceiros como o *LmEIF4AI*, ou a *LmPABP1*, a interação com os diferentes

homólogos ao eIF4E se alterava. Para esse objetivo, As proteínas GST-*LmEIF4G3* ou GST-*LmEIF4G4* foram imobilizadas na resina Glutathione Sefarose 4B e incubadas, ao mesmo tempo, com os homólogos ^{35}S -*LmEIF4AI*/ ^{35}S -*LmPABP1*. Aos sistemas contendo GST-*LmEIF4G3* + ^{35}S -*LmEIF4AI* + ^{35}S -*LmPABP1* ou GST-*LmEIF4G4* + ^{35}S -*LmEIF4AI* + ^{35}S -*LmPABP1* foram adicionados, individualmente, cada um dos homólogos do eIF4E marcados com ^{35}S (^{35}S -*LmEIF4E1-4*). A proteína *LmEIF4G3* liga-se especificamente com as proteínas *LmEIF4AI* e *LmPABP1*, em todas as condições testadas, e também com o *LmEIF4E4* (Figura 35). Já a proteína *LmEIF4G4* liga-se especificamente com as proteínas *LmPABP1* e *LmEIF4E3* (Figura 35). Mesmo na presença das proteínas *LmEIF4AI* e *LmPABP1*, não foram observadas novas interações das proteínas *LmEIF4G3-4* com outros homólogos do eIF4E, diferente daquelas já mostradas anteriormente. Em todas as combinações com cada um dos homólogos do eIF4E ou na sua ausência, a ligação da proteína *LmEIF4G4* com a proteína *LmPABP1* acontece de forma semelhante, exceto quando temos a ligação concomitante com o *LmEIF4E3* onde a ligação parece ser mais eficiente (Figura 35). Não foi possível detectar de forma clara a interação do GST-*LmEIF4G4* com a proteína *LmEIF4AI*, na maioria das combinações testadas, entretanto essa interação foi evidenciada no ensaio direto (GST-*LmEIF4AI*/ ^{35}S -*LmEIF4G4* – Figuras 18 e 22) e, como observado anteriormente, pode ser um reflexo da menor afinidade do *LmEIF4G4* por esta molécula. Estes resultados sugerem a presença de dois complexos eIF4F nos tripanossomatídeos, um deles que tem como proteína mantenedora o *LmEIF4G3* e no outro *LmEIF4G4*. Estes complexos vêm sendo validados em ensaios de imunoprecipitação em *L. major* e *T. brucei* (Tamara D'Carli Costa Lima, 2008, dados não publicados). Uma vez identificados é importante estabelecer o papel destes complexos na

iniciação da tradução dos tripanossomatídeos, bem como na a regulação da expressão gênica destes protozoários.

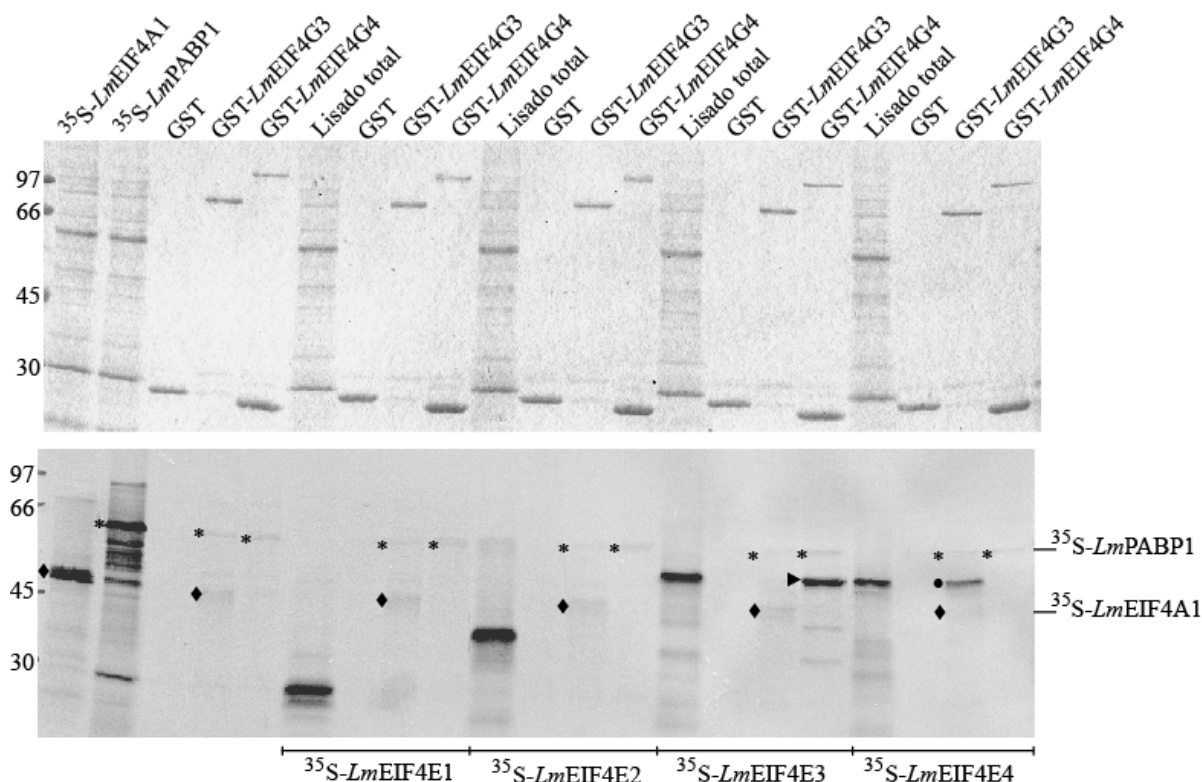


Figura 35. Reconstituição, “in vitro”, do complexo eIF4F de *Leishmania major*. As proteínas GST-*LmEIF4G3* e GST-*LmEIF4G4* foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com as proteínas ^{35}S -*LmEIF4A1* e ^{35}S -*LmPABP1* e réplicas dessa combinação foram misturados isoladamente com cada um dos homólogos do eIF4E de *L.major* (*LmEIF4E1-4*). O painel superior representa as proteínas recombinantes fracionadas em SDS-PAGE 15% e corados com azul de Coomassie. O painel inferior representa a autoradiografia mostrando ligação específica entre a proteína *LmEIF4G3* com as proteínas *LmEIF4A1* (♦) e *LmPABP1*(*) (em todas as combinações) e ao mesmo tempo com a proteína *LmEIF4E4* (•); e da proteína *LmEIF4G4* com *LmPABP1* (*) (em todas as combinações) e também com a proteína *LmEIF4E3*(►).

5.2 Estudo comparativo da interação dos homólogos EIF4G3-4 com os homólogos EIF4E3-4 de *L. major* e *Trypanosoma brucei*

Além dos ensaios reversos de interação entre os homólogos ao eIF4F de *L. major*, decidimos realizar um estudo comparativo para avaliar se os homólogos do EIF4G3-4 são capazes de interagir com os homólogos ao EIF4E3-4 de *L. major* e *T. brucei*. As proteínas de *T. brucei* foram incluídas dos ensaios de *pull-down* não só para analisar se elas são funcionalmente relacionadas às proteínas de *L. major*, mas também pelo fato de que ensaios adicionais de interação ao cap e interferência de RNA indicam diferentes comportamentos. As proteínas *TbEIF4E1*, *TbEIF4E2* e *TbEIF4E4* são capazes de ligar a resina cap Sefarose, entretanto apenas a proteína *TbEIF4E3* é requerido para a viabilidade celular das formas procíclicas e sanguíneas. Estes resultados adicionais serão melhor explanados na discussão e estão sendo compilados em um manuscrito, em fase final de redação, intitulado “**Characterization of multiple homologues of the cap binding protein (eIF4E) from *Trypanosoma brucei***”.

5.2.1 Ensaios para investigar a conservação das interações entre homólogos EIF4G3-4 e homólogos EIF4E3-4 em *L. major* e *Trypanosoma brucei*

As proteínas do eIF4F de *L. major* possuem ortólogos em *T. brucei* e *T. cruzi*, bem como em outras espécies da família Trypanosomatidae, sinalizando que o processo da iniciação da tradução destes protozoários seja comum e tenha aparecido evolutivamente a partir de um ancestral comum. Para investigar a conservação das interações observadas

entre as proteínas de *L. major* em seus ortólogos de *T. brucei*, diferentes proteínas EIF4G3 fusionadas à GST (*TbEIF4G3*, *LmEIF4G3* e mutante *LmEIF4G3* FSL/AAA) foram incubadas com as proteínas *TbEIF4E4* ou *LmEIF4E4* marcadas com ^{35}S . A proteína *TbEIF4E4* não se liga às proteínas EIF4G3 de *T. brucei* e *L. major*, embora a proteína *LmEIF4E4* tenha sido capaz de ligar-se especificamente ao seu parceiro *LmEIF4G3* e ao seu ortólogo *TbEIF4G3*. Corroborando a figura 28, o *LmEIF4E4* não foi capaz de ligar-se ao mutante *LmEIF4G3* FSL/AAA (Figura 36A). Apesar da interação entre as proteínas *TbEIF4G3* com a proteína *TbEIF4E4* não ter sido equivalente a interação entre as respectivas proteínas ortólogas em *L. major*, a forte interação *TbEIF4G3/LmEIF4E4* reforça a hipótese de que essa interação seja preservada entre as espécies, assim como o motivo de ligação entre ambas. Para avaliar a interação entre homólogos e ortólogos EIF4G4/EIF4E3, as proteínas EIF4G4 fusionadas à GST (*TbEIF4G4*, *LmEIF4G4* e mutante IL/AA) foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com *TbEIF4E3* e *LmEIF4E3* marcados com ^{35}S . A proteína *TbEIF4E3* ligam-se especificamente às proteínas *TbEIF4G4* e *LmEIF4G4* (setas), entretanto não foi capaz de ligar-se a proteína mutante IL/AA do *LmEIF4G4* (Figura 36B), sinalizando que o modo e o motivo de ligação entre eIF4G e eIF4E seja conservado nestas duas espécies e provavelmente nos tripanosomatídeos como um todo. Corroborando estas interações, os complexos obtidos neste trabalho estão sendo validados em ensaios de imunoprecipitação em *L. major* e *T. brucei* (Tamara D'Carli Costa Lima, 2009, dados não publicados).

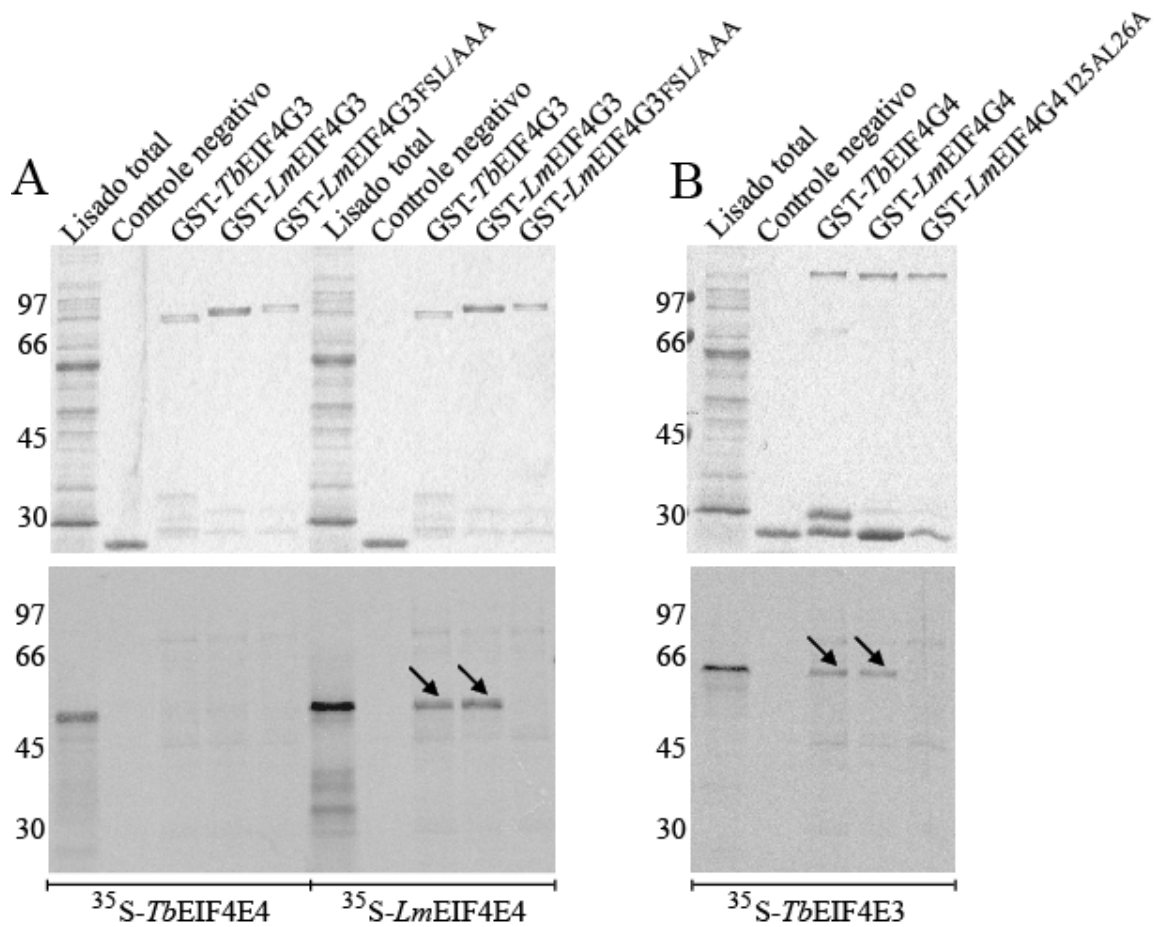


Figura 36. Análise da interação entre as proteínas EIF4G3-4 e as proteínas EIF4E3-4 de tripanossomatídeos. As proteínas GST-*TbEIF4G3*, GST-*LmEIF4G3*, GST-*LmEIF4G3_{FSL/AAA}*, GST-*TbEIF4G4*, GST-*LmEIF4G4* e GST-*LmEIF4G4_{IL/AA}* foram imobilizadas na resina Glutathione Sepharose 4B e incubadas com ³⁵S-*TbEIF4E4*/³⁵S-*LmEIF4E4* (*TbEIF4G3* e *LmEIF4G3*) e com ³⁵S-*TbEIF4E3*/³⁵S-*LmEIF4E3* (*TbEIF4G4* e *LmEIF4G4*). Os painéis superiores representam as proteínas recombinantes fracionados em gel 15% SDS-PAGE e corados com azul de Coomassie. O painel superior (A) representa a autoradiografia mostrando as ligações específicas da proteína *LmEIF4E4* com as proteínas EIF4G3 de *T.brucei* e *L.major* (setas). O painel inferior (B) representa a autoradiografia demonstrando as ligações específicas entre a proteína *TbEIF4E3* com as proteínas *LmEIF4G4* e *TbEIF4G4* (setas).

6. DISCUSSÃO

6.1-Identificação de dois complexos eIF4F em *L. major*.

No modelo humano, a proteína eIF4G interage com o eIF4A através dos domínios HEAT-1/MIF4G, HEAT-2/MA3 e de outras regiões presentes no eIF4G (Marintchev *et al.*, 2009). Dos cinco homólogos do eIF4G em *L. major*, apenas o *LmEIF4G5* não se ligou a proteína *LmEIF4AI* pela metodologia testada, entretanto apenas as proteínas EIF4G3-4 foram capazes de ligar-se aos homólogos *TbEIF4E3-4*, mostrando que a princípio formam-se pelo menos dois complexos eIF4F nestes parasitas. Apesar das proteínas *LmEIF4G1* e *LmEIF4G2* ligarem-se ao *LmEIF4AI*, ambas não são capazes de interagir com nenhum dos homólogos de eIF4E descritos (Dhalia *et al.*, 2005; Yoffe *et al.*, 2006). Entretanto, a participação deles na formação de outros complexos eIF4F não pode ser descartada, uma vez que mais recentemente foram encontrados dois homólogos adicionais do eIF4E [(EIF4E5-6) (Éden Ribeiro Freire, resultados não publicados)], cuja capacidade de interação com os eIF4Gs de *L.major* ainda não foi avaliada.

A proteína *LmEIF4G3* interage forte, e especificamente, com as proteínas *LmEIF4AI*, *LmEIF4E4* e *LmPABP1*, todas interações esperadas para a proteína eIF4G mantenedora do complexo eIF4F. De fato, a depleção do ortólogo *TbEIF4G3* na forma procíclica de *T. brucei* é essencial e inibe a síntese protéica apenas seis horas após a indução de RNAi (Danielle Moura, Resultados não publicados). O parceiro *TbEIF4AI* em *T. brucei* também é essencial e a remoção de pequenas quantidades, na forma procíclica, reduz fortemente a síntese de proteínas (Dhalia *et al.*, 2006). Desta forma, a interação mais forte entre essas proteínas, *LmEIF4G3/LmEIF4AI*, pode refletir uma relação de

dependência entre ambas. Entretanto, a remoção da proteína *TbEIF4E4* não é essencial para a viabilidade celular e não inibe a síntese de proteínas, em células procíclicas (Éden Ribeiro Freire, manuscrito em preparação). Em um trabalho independente em *L. major*, foi demonstrado que o *LmEIF4G3* além de interagir com *LmEIF4E4* também liga-se com menor afinidade ao *LmEIF4E1*, um resultado não observado nos nossos ensaios de *pull-down* usando os homólogos de eIF4G fusionados à GST. A interação entre o *LmEIF4E1* e a região N-terminal do *LmEIF4G3* foi demonstrada a partir da técnica de ressonância magnética nuclear (NMR- *Nuclear Magnetic Resonance*) (Yoffe *et al.*, 2009). Entretanto, o NMR permite identificar interações muito fracas [$K_d > 3\text{mM}$] (Vaynberg, J., Qin, J., 2006) que podem não ser detectadas em ensaios de *pull-down*, como os descritos nesta tese, justificando assim a não-observação de interação entre *LmEIF4E1* e *LmEIF4G3* nos nossos experimentos.

Em *T. brucei*, foi demonstrado que a proteína *TbEIF4E1* não é essencial à viabilidade celular. Entretanto a depleção simultânea de *TbEIF4E1* e *TbEIF4E4* levam as células a pararem de crescer, embora não morram, indicando que estas proteínas possuem funções complementares e/ou estejam relacionadas ao processo de divisão celular (Éden Ribeiro Freire, manuscrito em preparação). O fenótipo da depleção dupla *TbEIF4E1/TbEIF4E4* ainda é mais fraco que o observado para o *TbEIF4G3*, de forma é que possível que outros parceiros do EIF4G3, como por exemplo os recém identificados homólogos ao EIF4E (EIF4E5-6), tenham relevância na tradução e precisem ser melhor estudados. Nesse cenário, coerente com nossos primeiros resultados de *pull-down* onde o *LmEIF4G3* marcado parece ligar-se de forma equivalente a todos os homólogos ao EIF4E testados. A ligação diferencial entre *LmEIF4G3* e os respectivos homólogos do eIF4E de *L. major* poderia ser um alvo para a regulação da tradução nos tripanossomatídeos.

Embora a proteína *LmEIF4G4* apresente uma ligação mais reduzida ao *LmEIF4AI* e a *LmPABP1*, sua interação com o *LmEIF4E3* foi considerada forte e específica. Os ortólogos do *LmEIF4G4* e *LmEIF4E3* em *T. brucei* são essenciais à viabilidade celular, entretanto enquanto o *TbEIF4E3* parece estar envolvido com a síntese de proteínas, a depleção do *TbEIF4G4* não parece afetar o processo de tradução (Danielle Nascimento Moura e Éden Ribeiro Freire, resultados não publicados). As proteínas EIF4G3 e EIF4G4 ligam-se respectivamente às proteínas EIF4E4 e EIF4E3, assim paradoxalmente as proteínas que estariam relacionadas à tradução (EIF4G3 e EIF4E3) se ligariam às proteínas que aparentemente não seriam essenciais (EIF4E4 e EIF4G4), de acordo com os dados de RNAi em *T. brucei*. A presença de dois complexos homólogos ao eIF4F, nos tripanossomatídeos, poderia ser correlacionada com a descrição de dois complexos eIF4F distintos, eIF4F (eIF4G/eIF4E) e eIFiso4F (eIFiso4G/eIFiso4E), descritos em plantas. Da mesma forma como ocorre em plantas, onde dois complexos eIF4F podem ser ativados em diferentes tecidos, traduzindo diferencialmente mRNAs específicos ou classes de mRNAs (Browning et al, 1992; Mayberry et al., 2009), é possível que a ocorrência de pelo menos dois complexos eIF4F, coordenados por subunidades eIF4G distintas, esteja relacionada com a regulação da expressão gênica nos tripanossomatídeos.

6.2-Mapeamento dos motivos de ligação eIF4G/eIF4E

Em vários organismos foi demonstrado que o eIF4E reconhece o eIF4G através do motivo YXXXXLΦ (Rhoads, 2008) localizado na sua região N-terminal. Recentemente também foi descrito, por meio de ensaios de duplo híbrido, que o *LmEIF4G3* liga-se ao *LmEIF4E4* através de sua região N-terminal e que o motivo YPGFSLD seria responsável

por essa interação, visto que a mudança do Y, F, ou L por A, individualmente abolem a interação com o *LmEIF4E4*. Este motivo sobrepõe-se, porém não é idêntico ao motivo investigado nesta tese (Figura 37). Entretanto, não foi realizado o ensaio duplo híbrido reverso para corroborar se as mutações do motivo YPGFSLD, em outro contexto de dobramento protéico, continuariam abolindo a interação do *LmEIF4G3* ao *LmEIF4E4* (Yoffe et al., 2009). Por outro lado, para evitar este problema, nos nossos experimentos de interação entre os homólogos *LmEIF4G3/LmEIF4E4* e *LmEIF4G4/LmEIF4E3*, foram realizados ensaios reversos.

Nos nossos ensaios, as mutações implementadas no motivo FSL abolem a interação do *LmEIF4G3* com o *LmEIF4E4* por alterar dois resíduos de aminoácidos, chaves, de acordo com Yoffe *et al*, 2009. Entretanto, o Y do motivo YPG**F**SLD do EIF4G3 não é conservado nos homólogos do EIF4G4, indicando que essa região poderia não estar envolvida diretamente na interação com os homólogos do eIF4E. De acordo com as nossas análises de sequência e experimentos de *pull-down* o motivo de ligação do EIF4G3 seria o **FSLXXV/IV/L**, e nesse caso o efeito observado pela mutação do Y seria indireto devido a mudanças conformacionais.

Por outro lado, apesar da região FSL ser conservado nos homólogos EIF4G3 e EIF4G4, dos Tritryps, mutações pontuais nesta sequência abolem a interação entre *LmEIF4G3/LmEIF4E4* embora não afete a interação *LmEIF4G4/LmEIF4E3*. Estes dados indicam que motivos idênticos podem ser reconhecidos de maneiras diferentes, como é o caso dos motivos de ligação dos dois eIF4Gs de plantas que são diferencialmente reconhecidos pelos fatores eIF4E e eIFiso4E por mecanismos desconhecidos (Gallie, 2002).

A análise detalhada da sequência dos ortólogos do EIF4G4 através dos alinhamentos das figura 23 e figura 37, indicam um novo motivo M/F/YXXXXI/LL, onde

os resíduos IL quando modificados abolem a interação *LmEIF4G4/LmEIF4E3*. Entretanto, o aminoácido M no *LmEIF4G4* não parece ser importante para a ligação, uma vez que sua mutação não afeta a interação com a proteína *LmEIF4E3*. Assim, o resíduo M/F/Y não parece funcionar como substituto de Y do motivo típico, mostrando que este motivo em tripanossomatídeos pode tolerar variações maiores, que seriam espécie-específicas ou que esta região não represente o verdadeiro motivo de ligação aos homólogos do eIF4E. Ao mesmo tempo no *LmEIF4G3*, o L final é substituído por um R, reforçando a hipótese de que esse motivo não estaria envolvido na interação eIF4G/eIF4E. A ligação dos homólogos do EIF4E3-4 de *L. major* aos homólogos do EIF4G em *T. brucei*, demonstra a conservação dos motivos de ligação entre ambas as espécies. A substituição de L por R no *LmEIF4G3* não afeta a interação deste com a proteína *TbEIF4E4*, demonstrando que essa mutação pode de fato não influenciar na ligação EIF4G3/EIF4E4.

Apesar do motivo YXXXXLΦ ser altamente conservado em vários organismos analisados, para alguns outros organismos, parece que variações podem ser toleradas desde que mantenham-se resíduos de aminoácidos de mesmo grupo funcional. Uma variação, neste consenso de ligação, é encontrada também no nematódeo *Caenorhabditis. elegans* que possui o motivo FGRDFM. Muitos parceiros de ligação ao eIF4E, identificados em vários organismos, usam em sua maioria motivos similares (Rhoads, 2009). A possível presença de dois diferentes motivos de ligação para interação eIF4G/eIF4E nos tripanossomatídeos, sinaliza para funções no contexto de dois complexos eIF4F que poderiam desempenhar funções diferentes, como tradução diferencial de classes de mRNAs, como descrito em plantas (Mayberry *et al.*, 2009).

Até o momento, embora as evidências experimentais sugiram, não se pode concluir que existam dois motivos funcionais nos homólogos do eIF4G de *L. major*. Entretanto, é

(Tarun *et al.*, 1997; Le *et al.*, 1997; Imataka *et al.*, 1998). Em humano, o motivo de interação à PABP1 envolve 45 resíduos de aminoácidos e o sítio de interação ao eIF4E se estende além do motivo principal (YXXXXLΦ) por mais aproximadamente 120 resíduos de aminoácidos (Marintchev *et al.*, 2009). Apesar da região N-terminal dos EIF4G3-4 não apresentar as características para abrigar um sítio de ligação à PABP, a reconstituição dos complexos eIF4F/PABP indica que não há sobreposição dos motivos de ligação aos homólogos EIF4E e PABP. O estudo mais detalhado dos diferentes domínios do *LmEIF4G3* indica que o MIF4G pode mediar a interação à *LmPABP1-2*. Corroborando essa hipótese, a região N-terminal do *LmEIF4G3* de *L. major* não é capaz de associar-se à *LmPABP1* através de ensaios de duplo híbrido (Yoffe *et al.*, 2009). Pelo menos dois conjuntos de experimentos, independentes, sinalizam para um novo conceito na interação eIF4G/PABP, sendo aqui feito o primeiro registro indicando o MIF4G como domínio envolvido nessa interação. Por outro lado, a interação do *LmEIF4G1* com a PABP1 poderia constituir, em tripanossomatídeos, um exemplo clássico de interação EIF4G/PABP, uma vez que a sua região N-terminal e conteúdo de aminoácidos básicos são maiores. Entretanto, não foi demonstrada interação entre *LmEIF4G1* e *LmEIF4E1-4*, fazendo-se necessários novos experimentos de interação com os recém identificados *LmEIF4E5-6* na busca de um terceiro complexo eIF4F/PABP nestes protozoários.

6.4 - Considerações finais

Considerando que os ensaios de ligação entre as proteínas homólogas ao eIF4F foram realizados *in vitro*, é possível que algumas interações não tenham ocorrido devido a diferenças fisiológicas e/ou problemas conformacionais. Através de sistemas de

purificação, capazes de simular condições fisiológicas, será possível inferir se estas ligações ocorrem da mesma forma “*in vivo*”, assim como identificar possíveis outros ligantes. Ensaio de imunoprecipitação já permitiram confirmar muitas das interações observadas a partir dos experimentos de *pull-down* (como as interações EIF4G3/EIF4E4, EIF4G3/PABP e EIF4G4/EIF4E3 – Tamara De’Carli, resultados não publicados). Experimentos de super-expressão das proteínas *TbEIF4G3-4*, selvagens e mutantes, permitirá avaliar os efeitos dos diferentes mutantes sobre a síntese de proteínas, em *T. brucei*, e a capacidade ou não de interagir *in vivo* com outras subunidades do eIF4F, que será verificada por imunoprecipitação (Danielle Nascimento Moura, Tese de Doutorado em andamento). Ensaio funcionais complementares, para definir melhor a participação de cada um dos fatores homólogos ao eIF4F nos tripanosomatídeos na sua síntese protéica, também estão previstos: (1) TAP-Tag das proteínas EIF4AIII, EIF4G3, EIF4G4 de *T. brucei*; (2) identificação de novos parceiros, através do fracionamento de complexos protéicos obtidos pelo TAP-Tag, em eletroforese bidimensional e análise por espectrometria de massas dos *spots* obtidos.

Os resultados obtidos aqui indicam a formação de pelo menos dois complexos eIF4F em tripanosomatídeos (Figura 38), indicando que estes parasitas possuem uma maquinaria complexa de iniciação da síntese protéica, com características próprias, mas que em muitos aspectos assemelha-se aos modelos presentes em outros organismos. O Estudo do papel funcional destes fatores permitirá relacioná-los a mecanismos de regulação da expressão gênica, assim como à função dos mesmos em relação à sobrevivência, destes parasitas, em diferentes hospedeiros. A investigação dos motivos de interação eIF4G/eIF4A e eIF4G/eIF4E podem dar suporte a terapias mais específicas contra doenças como as

Leishmanioses que apresentam tratamentos dolorosos, inespecíficos e associados à geração de resistência pelo uso inadequado de drogas (Santos *et al.*, 2008). A inibição da interação eIF4G/eIF4E tem sido investigada em humanos como possibilidade do tratamento câncer, onde uma pequena molécula, que compete com o eIF4E pela ligação ao eIF4G, exibe atividade contra várias linhagens de células cancerígenas (Moerke *et al.*, 2007). Além disso, o bloqueio da interação do eIF4G/eIF4A tem sido obtida utilizando uma molécula inibidora que liga-se especificamente ao eIF4A (Lindqvist *et al.*, 2008).

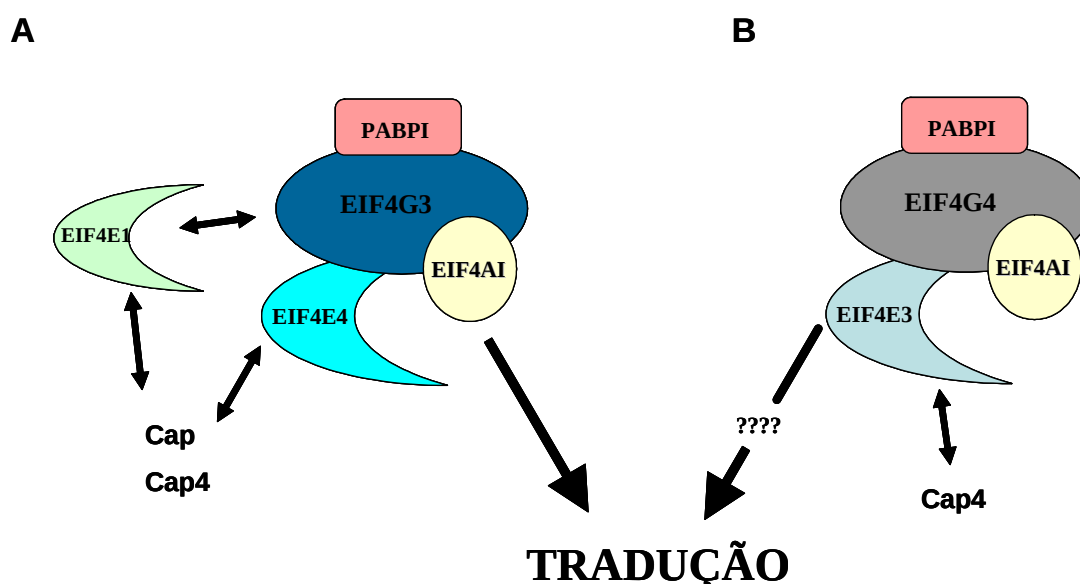


Figura 38. Os dois possíveis complexos homólogos ao eIF4F nos tripanossomatídeos, com base nos ensaios de *pull-down* em *L. major*. (A) *LmEIF4G3* interage especificamente com *LmEIF4A1*, *LmEIF4E4* (e talvez o *LmEIF4E1*) e *LmPABP1*. (B) *LmEIF4G4* interage especificamente com o *LmEIF4AI*, *LmEIF4E3* e *LmPABP1*.

Em síntese, os conhecimentos gerados neste trabalho com relação aos homólogos ao eIF4F nos tripanossomatídeos, pode fornecer subsídios para o desenvolvimento de novas terapias e vacinas, capazes de interromper a proliferação destes parasitas nos seus hospedeiros definitivos.

7. RESULTADOS COMPLEMENTARES OBTIDOS DURANTE O DOUTORADO

7.1 Estudo complementar acerca do papel funcional de homólogos de eIF4A em *T. brucei*

A proteína helicase eIF4A, fator de tradução responsável pela remoção de estruturas secundárias durante a iniciação da tradução, apresenta dois homólogos nos tripanossomatídeos, tendo sua função sido inicialmente estudada em *L. major* (Dhalia *et al.*, 2005). Contribuindo para elucidação da função destas proteínas, os ortólogos do eIF4A de *T. brucei* (*TbEIF4AI* e *III*) foram estudados quanto à localização subcelular, determinação de seus níveis intracelulares e efeito da sua depleção (por interferência de RNA) na viabilidade/morfologia celular. Assim foi mostrado que o *TbEIF4AI* é citoplasmático, expresso em altos níveis e uma depleção mínima de 10% reduz drasticamente a síntese de proteínas, sendo essencial à viabilidade celular. Em contraste, o *TbEIF4AIII* é nuclear, moderadamente expresso e sua depleção só inibe a proliferação celular apenas quatro ciclos de divisão celular após sua interferência de RNA. De acordo com os dados obtidos, concluímos que as proteínas *TbEIF4AI* e *TbEIF4AIII* são respectivamente os homólogos funcionais do eIF4AI e eIF4AIII de mamíferos. A minha contribuição para este trabalho foi determinar os níveis intracelulares de ambas as proteínas e avaliar a depleção do *TbEIF4AIII* após a interferência de RNA em extratos prociclios de *T. brucei*. Adicionalmente, ainda realizei ensaios de interação entre as proteínas *TbEIF4AI* e *TbEIF4AIII* com a proteína de *T. brucei* homóloga a proteína humana Mago (Magoh), *TbMago*, associada ao mecanismo de degradação de mRNA em humanos. Estes resultados

fazem parte do artigo “**The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct**” publicado na revista *Nucleic Acids Research*, em 2006 (em anexo).

7.1.1 Análise da expressão das proteínas TbEIF4AI-III nas fases procíclicas e sanguíneas de *T. brucei*

Os anticorpos contra as proteínas TbEIF4AI-III, produzidos anteriormente (Rafael Dhalia, 2005 - Tese de Doutorado em Biologia Molecular, UNB) foram purificados por imunoabsorção e utilizados para a determinação dos níveis intracelulares das respectivas proteínas endógenas nas fases procíclica e sanguínea de *T. brucei*. As proteínas recombinantes, TbEIF4AI e TbEIF4AIII, fusionadas a uma cauda de seis Histidinas, foram expressas e quantificadas, através de curvas de diluição das mesmas comparadas com concentrações conhecidas de albumina sérica bovina – BSA, em gel SDS-PAGE corado com azul de Coomassie. As proteínas recombinantes foram reveladas em ensaios de *Western-blot* comparando o sinal de detecção das mesmas com diluições seriadas de extratos protéicos totais de *T. brucei* nas fases procíclica e sanguínea (Figura 39). Os níveis de TbEIF4AI e III endógenos foram então estimados através de análises densitométricas, dos filmes de *Western-blot*, revelados por quimioluminescência (Dhalia *et al.*, 2005). Os valores obtidos em fentomoles das proteínas TbEIF4AI e TbEIF4AIII para 10^5 e 10^6 células respectivamente, foram convertidos em número de moléculas por célula. A proteína TbEIF4AI foi quantificada em 2.4×10^5 e 6×10^4 moléculas por célula, respectivamente na forma procíclica e sanguínea *T. brucei*. A proteína TbEIF4AIII foi estabelecida em 2×10^4 e 4.5×10^3 moléculas por célula respectivamente na forma procíclica e sanguínea de *T. brucei*. TbEIF4AI foi considerada abundante e seus níveis intracelulares são compatíveis com os

níveis de seu ortólogo em *L. major* (Dhalia *et al.*, 2005), assim como com o eIF4A de levedura (von der Haar e McCarthy, 2002). Os resultados da quantificação indicam que a concentração intracelular de *TbEIF4AI* é, pelo menos, 10 vezes maior que a de *TbEIF4AIII*. Adicionalmente, a concentração das proteínas *TbEIF4AI-III* na forma procíclica, foi considerada 4 vezes maior que na forma sanguínea.

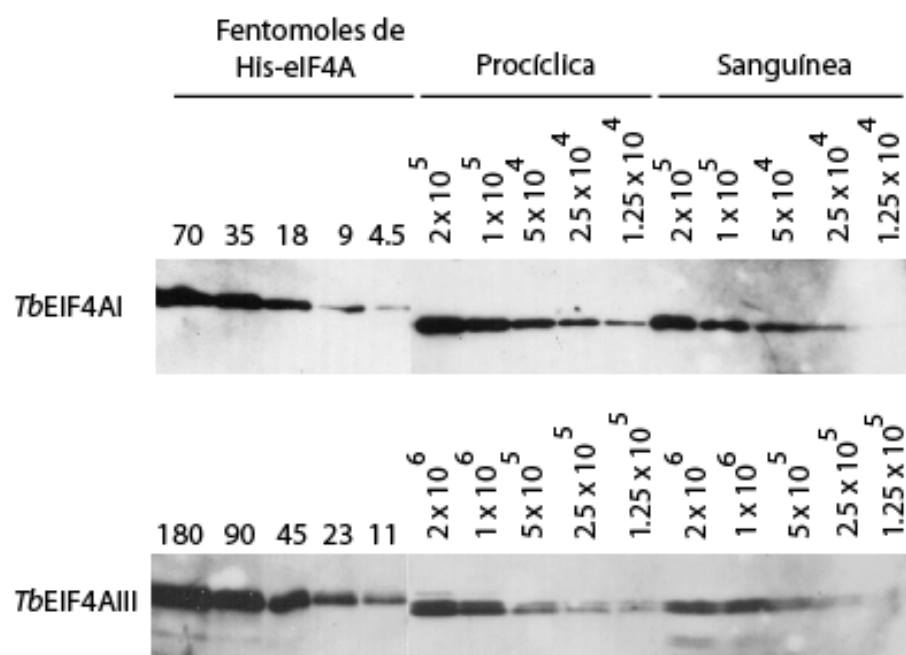


Figura 39: Quantificação dos níveis intracelulares das proteínas *TbEIF4AI-III* nas formas procíclicas e sanguíneas de *T. brucei*. Diluições seriadas das proteínas *TbEIF4AI-III*, em concentrações definidas (em fentomoles), foram fracionadas em gel 15% SDS-PAGE, com uma curva de diluição de células de ambas as formas procíclica e sanguínea ($1.25 \times 10^4 - 2 \times 10^5$ para o *TbEIF4AI* e $1.25 \times 10^5 - 2 \times 10^6$ para o *TbEIF4AIII*). Os valores obtidos em fentomoles das proteínas *TbEIF4AI* e *TbEIF4AIII* para 10^5 e 10^6 células respectivamente, foram convertidos em número de moléculas por célula.

7.1.2 Efeito da interferência de RNA nos níveis intracelulares da proteína TbEIF4AIII.

Para investigar a viabilidade dos parasitos *in vivo*, na ausência dos homólogos de *TbEIF4AIII*, foi utilizada a técnica de interferência de RNA (RNAi). O gene *TbEIF4AIII* foi clonado no cassete de expressão p2T7-177 (LaCount DJ et al., 2000), construído para indução de RNAi. Os genes clonados no plasmídeo p2T7-177, sob controle de dois promotores T7 em orientação reversa (regulados pela tetraciclina), são flanqueados por seqüências de *T. brucei* para integração genômica estável, nos seus mini-cromossomos. Os parasitos transfectados com a construção p2T7-177/*TbEIF4AIII* foi selecionado pela marca de resistência a fleomicina, para posterior análise do efeito da interferência de RNA através de curvas de crescimento, na presença e ausência de tetraciclina (+TET/- TET). Aliquotas de extratos protéicos, obtidos durante a realização do RNAi do *TbEIF4AIII* na concentração de 10^5 células/ μ L foram retiradas nos tempos 24, 48, 72 horas. O perfil de expressão da proteína *TbEIF4AIII* nas amostras do RNAi foi avaliado por *Western-blot*. Após 24 horas de indução do RNAi, a proteína *TbEIF4AIII* já não foi mais detectada no extrato total de *T. brucei* (Figura 40). Como esperado, os níveis da proteína *TbEIF4AIII* permanecem inalterados na interferência de RNA do *TbEIF4AI*.

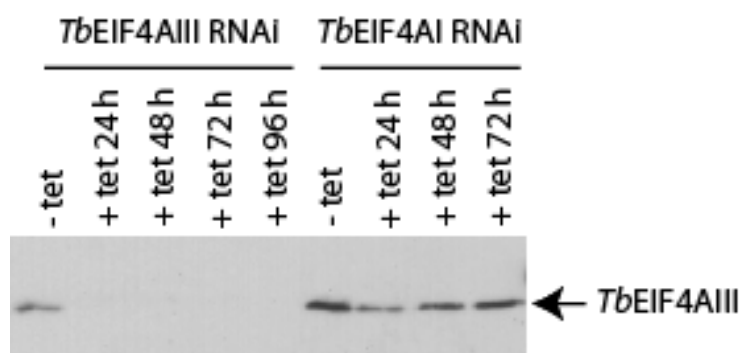


Figura 40. Interferência do RNA *TbEIF4AIII*. Parasitas da forma procíclica foram transfectados com a construção p2T7-177/*TbEIF4AIII*, monitorados quanto ao crescimento celular e analisados em relação à expressão do *TbEIF4AIII*. Análises dos experimentos de RNAi do *TbEIF4AI* e *TbEIF4AIII* por *Western-blot*, utilizando os anticorpos contra *TbEIF4AIII*, mostram a depleção específica da proteína *TbEIF4AIII* a partir das primeiras 24h de interferência.

7.1.3 - Interação entre homólogos do EIF4A com o homólogo à proteína Magoh em *T. brucei*

Para avaliar a possível formação de um complexo de degradação de RNA nos tripanossomatídeos, similar ao complexo de Junção entre Éxons (EJC) descrito em humanos, as proteínas recombinantes *TbEIF4AI* e *III* fusionadas à proteína GST foram imobilizadas na resina Glutathione Sefarose 4B e incubadas com *TbMago*, marcada com ^{35}S . Ambos homólogos do eIF4A ligam-se, fracamente, a proteína *TbMago* (Figura 41). Como esperado, a proteína *TbEIF4AIII* liga-se a proteína *TbMago*, como descrito para o fator eIF4AIII humano. A interação entre *TbMago* com *TbEIF4AI* pode acontecer em função da alta similaridade deste com *TbEIF4AIII*, entretanto pode não ser ocorrer *in vivo* devido as diferentes localizações das proteínas eIF4A em *T.brucei*. Estas interações serão

confirmadas por ensaios reversos de *pull-down*, imunoprecipitação, e por purificação por cromatografia de afinidade dupla (TAP-Tag). Para os ensaios de TAP-tag, onde é possível investigar a formação de complexos protéicos *in vivo*, já temos uma linhagem de *T. brucei* modificada para expressar a proteína *TbEIF4AIII* fusionada à um tag de purificação. Com estes estudos complementares, esperamos contribuir não só para o estudo da iniciação da tradução, como também para a investigação de processos celulares, relacionados a estes dois homólogos funcionais de eIF4A, em *T. brucei*.

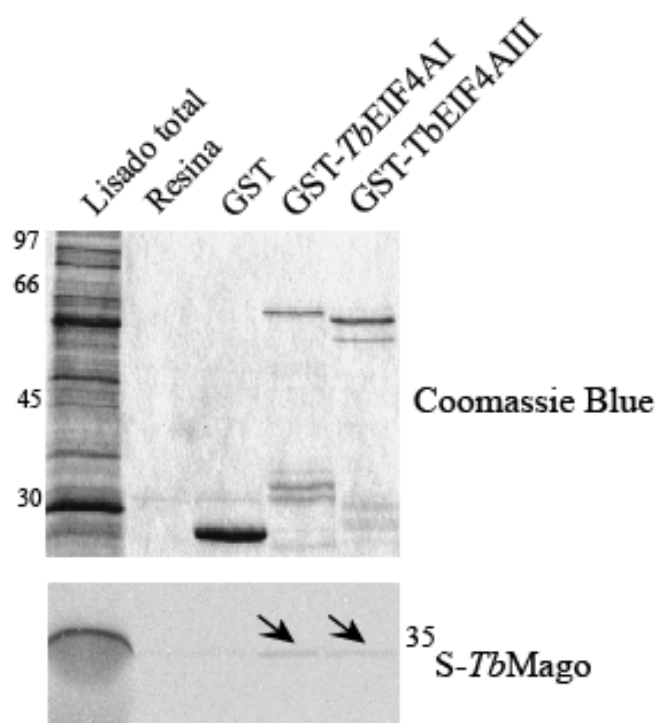


Figura 41. Análise de interação entre os homólogos ao eIF4A e o homólogo Mago de *T. brucei*. As proteínas recombinantes GST-*TbEIF4AI-III* foram imobilizadas em resina Glutathiona Sefarose 4B e incubadas com a proteína ³⁵S-*TbMago*.. O painel superior representa o *pull-down* as proteínas GST-*TbEIF4As* recombinantes fracionadas em gel 15% SDS-PAGE e corado com azul de Coomassie. O painel inferior representa as ligações entre o *TbMago* e as proteínas *TbEIF4AI* e *III* (setas).

7.2 – Discussão dos resultados complementares obtidos durante o Doutorado

A quantificação das proteínas *TbEIF4AI* e *TbEIF4AIII* em *T. brucei* demonstrou que o *TbEIF4AI* é bem mais abundante em ambas as formas do ciclo de vida do parasita, em relação à proteína *TbEIF4AIII*, corroborando os resultados obtidos para os respectivos ortólogos de *L. major* (Dhalia *et al.*, 2005). Apesar das proteínas *TbEIF4AI* e *TbEIF4AIII* apresentarem 69% de identidade entre si, o fator *TbEIF4AI* foi detectado no citoplasma e o *TbEIF4AIII* no núcleo, apesar do mesmo não possuir nenhum sinal claro de importação para o núcleo. Em humanos, os homólogos *eIF4AI* e *eIF4AIII* apresentam 66% de identidade entre si e localizações também distintas, o *eIF4AI* sendo citoplasmático e o *eIF4AIII* nuclear. O *eIF4AIII* humano foi caracterizado como um fator envolvido em um complexo formado nas junções entre os éxons (EJC), gerado após o processamento dos mRNAs por *cis-Splicing*. O complexo EJC, formado pelas proteínas *eIF4AIII*, Magoh e Y14, está associado ao transporte de mRNA e a degradação de mRNAs terminados precocemente (*Non-sense Mediated Decay* – NMD), sem qualquer envolvimento no processo de iniciação da síntese protéica (Palacios *et al.*, 2004; Shibuya *et al.*, 2004). Em conjunto com outros resultados descritos no trabalho publicado (Dhalia *et al.*, 2006, em anexo) as evidências obtidas para os homólogos de *eIF4A* de *T. brucei* apontam para o *TbEIF4AI* como um fator de tradução enquanto o *TbEIF4AIII* poderia ser um homólogo funcional ao *eIF4AIII* humano, que pode ser componente do EJC.

Em tripanossomatídeos praticamente não existem íntrons, assim o processamento do mRNA em *trans* assegura a formação de um mRNA maduro, contendo um mini-éxon (seqüência SL) fusionado ao éxon codificante para proteína. A região cicatriz entre mini-

éxon/éxon poderia fornecer condições para a formação de um complexo EJC nos tripanossomatídeos (Dhalia *et al.*, 2006). Esta hipótese foi ainda reforçada pela identificação de um homólogo de *T. brucei* da Magoh, denominada aqui denominado *TbMago*. Ensaio de ligação proteína-proteína demonstraram a associação entre as duas proteínas *TbEIF4AI* e *III* à proteína *TbMago*, indicando para a provável formação de um complexo EJC em tripanossomatídeos. Recentemente foi publicado a descrição de um provável EJC em tripanossomatídeos, onde as subunidades *TbMago* e a uma proteína divergente Y14 de *T. brucei* se associam *in vitro* (também avaliado para os homólogos de *T. cruzi*) e *in vivo*. Entretanto estas proteínas não parecem ligar-se a proteína *TbEIF4AIII*. Especula-se que mutações em dois dos três resíduos na proteína *TbY14*, que em humanos abolem a interação a *eIF4AIII*, estejam envolvidas nessa ausência de interação ao *TbEIF4AIII*, todavia a implicação biológica destas interações precisa ser melhor caracterizada (Bercovich *et al.*, 2009b) .

8. CONCLUSÕES

- As proteínas EIF4G3 e EIF4G4 possuem dois possíveis motivos de ligação às proteínas EIF4E4 e EIF4E3 em *L. major*;
- As interações entre as proteínas EIF4G3-4 e EIF4E3-4 de *L. major* e de *T. brucei*, sugerem a conservação evolutiva da iniciação da tradução nos tripanossomatídeos;
- O motivo de interação das proteínas EIF4G3 e EIF4G4 com a PABP1, em *L. major*, parece se localizar no domínio MIF4G e não na parte N-terminal, como descrito em vários outros eucariotos;
- O motivo de ligação das proteínas EIF4G3 e EIF4G4 ao EIF4AI é semelhante aos descritos em outros eucariotos, pois quando mutagenizados abolem a interação eIF4G/eIF4A;
- Em *T. brucei*, a proteína EIF4AI parece está envolvida com a iniciação da tradução, enquanto EIF4AIII possivelmente esta associada a algum mecanismo de degradação de mRNAs;
- Existem dois prováveis complexos EIF4F/PABP nos tripanossomatídeos: EIF4E4/EIF4G3/EIF4AI/PABP1 e EIF4E3/EIF4G4/EIF4AI/PABP1;

9. REFERÊNCIAS BIBLIOGRÁFICAS

- Acker, M.G., Shin, B.S., Nanda, J.S., Saini, A.K., Dever, T.E., and Lorsch, J.R. (2009). Kinetic analysis of late steps of eukaryotic translation initiation. *J. Mol. Biol.* 385, 491-506.
- Albrecht, M. and Lengauer, T. (2004). Survey on the PABC recognition motif PAM2. *Biochem. Biophys. Res. Commun.* 316, 129-138.
- Amrani, N., Ghosh, S., Mangus, D.A., and Jacobson, A. (2008). Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* 453, 1276-1280.
- Andersen, C.B., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L., Pedersen, J.S., Seraphin, B., Le, H.H., and Andersen, G.R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* 313, 1968-1972.
- Anderson, P. and Kedersha, N. (2006). RNA granules. *J. Cell Biol.* 172, 803-808.
- Anderson, P. and Kedersha, N. (2008). Stress granules: the Tao of RNA triage. *Trends Biochem. Sci.* 33, 141-150.
- Andrade, M.A. and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. *Nat. Genet.* 11, 115-116.
- Balogopal, V. and Parker, R. (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* 21, 403-408.
- Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B., and Le, H.H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat. Struct. Mol. Biol.* 12, 861-869.
- Bangs, J.D., Crain, P.F., Hashizume, T., McCloskey, J.A., and Boothroyd, J.C. (1992). Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J. Biol. Chem.* 267, 9805-9815.
- Barhoumi, M., Tanner, N.K., Banroques, J., Linder, P., and Guizani, I. (2006). *Leishmania infantum* LeIF protein is an ATP-dependent RNA helicase and an eIF4A-like factor that inhibits translation in yeast. *FEBS J.* 273, 5086-5100.
- Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* 33, 7138-7150.
- Barrett, M.P., Burchmore, R.J., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J., and Krishna, S. (2003). The trypanosomiases. *Lancet* 362, 1469-1480.

- Bates,E.J., Knuepfer,E., and Smith,D.F. (2000). Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites. *Nucleic Acids Res.* 28, 1211-1220.
- Batista,J.A., Teixeira,S.M., Donelson,J.E., Kirchhoff,L.V., and de Sa,C.M. (1994a). Characterization of a Trypanosoma cruzi poly(A)-binding protein and its genes. *Mol. Biochem. Parasitol.* 67, 301-312.
- Bellsolell,L., Cho-Park,P.F., Poulin,F., Sonenberg,N., and Burley,S.K. (2006). Two structurally atypical HEAT domains in the C-terminal portion of human eIF4G support binding to eIF4A and Mnk1. *Structure.* 14, 913-923.
- Benz,C., Nilsson,D., Andersson,B., Clayton,C., and Guilbride,D.L. (2005). Messenger RNA processing sites in Trypanosoma brucei. *Mol. Biochem. Parasitol.* 143, 125-134.
- Bercovich,N., Levin,M.J., and Vazquez,M.P. (2009a). The FIP-1 like polyadenylation factor in trypanosomes and the structural basis for its interaction with CPSF30. *Biochem. Biophys. Res. Commun.* 380, 850-855.
- Bercovich,N., Levin,M.J., Clayton,C., and Vazquez,M.P. (2009b). Identification of core components of the exon junction complex in trypanosomes. *Mol. Biochem. Parasitol.* 166, 190-193.
- Bradley,C.A., Padovan,J.C., Thompson,T.L., Benoit,C.A., Chait,B.T., and Rhoads,R.E. (2002). Mass spectrometric analysis of the N terminus of translational initiation factor eIF4G-1 reveals novel isoforms. *J. Biol. Chem.* 277, 12559-12571.
- Bringaud,F., Muller,M., Cerqueira,G.C., Smith,M., Rochette,A., El-Sayed,N.M., Papadopoulou,B., and Ghedin,E. (2007). Members of a large retroposon family are determinants of post-transcriptional gene expression in Leishmania. *PLoS. Pathog.* 3, 1291-1307.
- Browning,K.S. (1996). The plant translational apparatus. *Plant Mol. Biol.* 32, 107-144.
- Browning,K.S., Webster,C., Roberts,J.K., and Ravel,J.M. (1992). Identification of an isozyme form of protein synthesis initiation factor 4F in plants. *J. Biol. Chem.* 267, 10096-10100.
- Brune,C., Munchel,S.E., Fischer,N., Podtelejnikov,A.V., and Weis,K. (2005). Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA.* 11, 517-531.
- Buttner,K., Wenig,K., and Hopfner,K.P. (2006). The exosome: a macromolecular cage for controlled RNA degradation. *Mol. Microbiol.* 61, 1372-1379.
- Byrd,M.P., Zamora,M., and Lloyd,R.E. (2002). Generation of multiple isoforms of eukaryotic translation initiation factor 4GI by use of alternate translation initiation codons. *Mol. Cell Biol.* 22, 4499-4511.

- Byrd,M.P., Zamora,M., and Lloyd,R.E. (2005). Translation of eukaryotic translation initiation factor 4GI (eIF4GI) proceeds from multiple mRNAs containing a novel cap-dependent internal ribosome entry site (IRES) that is active during poliovirus infection. *J. Biol. Chem.* 280, 18610-18622.
- Campbell,D.A., Sturm,N.R., and Yu,M.C. (2000). Transcription of the kinetoplastid spliced leader RNA gene. *Parasitol. Today* 16, 78-82.
- Campbell,D.A., Thomas,S., and Sturm,N.R. (2003). Transcription in kinetoplastid protozoa: why be normal? *Microbes. Infect.* 5, 1231-1240.
- Caron,S., Charon,M., Cramer,E., Sonenberg,N., and Dusanter-Fourt,I. (2004). Selective modification of eukaryotic initiation factor 4F (eIF4F) at the onset of cell differentiation: recruitment of eIF4GII and long-lasting phosphorylation of eIF4E. *Mol. Cell Biol.* 24, 4920-4928.
- Caruthers,J.M., Johnson,E.R., and McKay,D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl. Acad. Sci. U. S. A* 97, 13080-13085.
- Cassola,A., De Gaudenzi,J.G., and Frasch,A.C. (2007). Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Mol. Microbiol.* 65, 655-670.
- Chakraborty,K. (2001). Translational regulation by ABC systems. *Res. Microbiol.* 152, 391-399.
- Chang,T.C., Yamashita,A., Chen,C.Y., Yamashita,Y., Zhu,W., Durdan,S., Kahvejian,A., Sonenberg,N., and Shyu,A.B. (2004). UNR, a new partner of poly(A)-binding protein, plays a key role in translationally coupled mRNA turnover mediated by the c-fos major coding-region determinant. *Genes Dev.* 18, 2010-2023.
- Chappuis,F., Sundar,S., Hailu,A., Ghalib,H., Rijal,S., Peeling,R.W., Alvar,J., and Boelaert,M. (2007). Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? *Nat. Rev. Microbiol.* 5, 873-882.
- Chekanova,J.A. and Belostotsky,D.A. (2003). Evidence that poly(A) binding protein has an evolutionarily conserved function in facilitating mRNA biogenesis and export. *RNA.* 9, 1476-1490.
- Cheng,S. and Gallie,D.R. (2007). eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains. *J. Biol. Chem.* 282, 25247-25258.
- Clayton,C. and Shapira,M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.* 156, 93-101.
- Clayton,C.E. (2002). Life without transcriptional control? From fly to man and back again. *EMBO J.* 21, 1881-1888.

- Coldwell,M.J. and Morley,S.J. (2006). Specific isoforms of translation initiation factor 4GI show differences in translational activity. *Mol. Cell Biol.* 26, 8448-8460.
- Costa Lima, T.D. (2007). Caracterização funcional de homólogos à Proteína de Ligação a Cauda Poli-A (PABP) de *Leishmania major*. Dissertação de Mestrado, Universidade Federal de Pernambuco, Recife- PE.
- Craig,A.W., Haghighat,A., Yu,A.T., and Sonenberg,N. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* 392, 520-523.
- D'Orso,I., De Gaudenzi,J.G., and Frasch,A.C. (2003). RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol.* 19, 151-155.
- D'Orso,I. and Frasch,A.C. (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.
- Das,A., Banday,M., and Bellofatto,V. (2008). RNA polymerase transcription machinery in trypanosomes. *Eukaryot. Cell* 7, 429-434.
- De,G.J., Frasch,A.C., and Clayton,C. (2005). RNA-binding domain proteins in Kinetoplastids: a comparative analysis. *Eukaryot. Cell* 4, 2106-2114.
- Dever,T.E. (1999). Translation initiation: adept at adapting. *Trends Biochem. Sci.* 24, 398-403.
- Dhalia,R., Marinsek,N., Reis,C.R., Katz,R., Muniz,J.R., Standart,N., Carrington,M., and de Melo Neto,O.P. (2006). The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. *Nucleic Acids Res.* 34, 2495-2507.
- Dhalia,R., Reis,C.R., Freire,E.R., Rocha,P.O., Katz,R., Muniz,J.R., Standart,N., and de Melo Neto,O.P. (2005). Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. *Mol. Biochem. Parasitol.* 140, 23-41.
- Dhalia,R. (2005). Caracterização da iniciação da síntese protéica em tripanosomatídeos: análise funcional de polipeptídeos homólogos às subunidades do complexo eIF4F de iniciação da tradução. Tese de Doutorado, Universidade de Brasília, Brasília-DF.
- Dominguez,D., Altmann,M., Benz,J., Baumann,U., and Trachsel,H. (1999). Interaction of translation initiation factor eIF4G with eIF4A in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 274, 26720-26726.
- Dunn,E.F., Hammell,C.M., Hodge,C.A., and Cole,C.N. (2005). Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev.* 19, 90-103.

Einarson, M.B. (2001). Detection of Protein-Protein Interactions Using the GST Fusion Protein Pulldown Technique. *In* Molecular Cloning: A Laboratory Manual, 3rd Edition, Cold Spring Harbor Laboratory Press, pp.18.55-18.59.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.N., Wortman, J.R., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., and Hall, N. (2005a). Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404-409.

El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van Aken S, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B. (2005b). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309, 409-415.

Estevez, A.M., Kempf, T., and Clayton, C. (2001). The exosome of *Trypanosoma brucei*. *EMBO J.* 20, 3831-3839.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol. Cell Biol.* 27, 3970-3981.

Feng, P., Everly, D.N., Jr., and Read, G.S. (2005). mRNA decay during herpes simplex virus (HSV) infections: protein-protein interactions involving the HSV virion host shutoff protein and translation factors eIF4H and eIF4A. *J. Virol.* 79, 9651-9664.

Freire, E.R. (2005). Caracterização funcional de homólogos ao fator de iniciação da tradução eIF4E de tripanossomatídeos. Dissertação de Mestrado, Universidade Federal de Pernambuco, Recife- PE.

Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.* 5, 2108-2116.

Gallie,D.R. and Browning,K.S. (2001). eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. *J. Biol. Chem.* 276, 36951-36960.

Ganoza,M.C., Kiel,M.C., and Aoki,H. (2002). Evolutionary conservation of reactions in translation. *Microbiol. Mol. Biol. Rev.* 66, 460-85, table.

Ginger,M.L. (2005). Trypanosomatid biology and euglenozoan evolution: new insights and shifting paradigms revealed through genome sequencing. *Protist.* 156, 377-392.

Gingras,A.C., Raught,B., and Sonenberg,N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* 68, 913-963.

Goodfellow,I.G. and Roberts,L.O. (2008). Eukaryotic initiation factor 4E. *Int. J. Biochem. Cell Biol.* 40, 2675-2680.

Goyer,C., Altmann,M., Lee,H.S., Blanc,A., Deshmukh,M., Woolford,J.L., Jr., Trachsel,H., and Sonenberg,N. (1993). TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function. *Mol. Cell Biol.* 13, 4860-4874.

Gradi,A., Imataka,H., Svitkin,Y.V., Rom,E., Raught,B., Morino,S., and Sonenberg,N. (1998). A novel functional human eukaryotic translation initiation factor 4G. *Mol. Cell Biol.* 18, 334-342.

Gross,J.D., Moerke,N.J., von der,H.T., Lugovskoy,A.A., Sachs,A.B., McCarthy,J.E., e Wagner,G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739-750.

Gunawardena,S.R., Ruis,B.L., Meyer,J.A., Kapoor,M., and Conklin,K.F. (2008). NOM1 targets protein phosphatase I to the nucleolus. *J. Biol. Chem.* 283, 398-404.

Haghighat,A. and Sonenberg,N. (1997). eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J. Biol. Chem.* 272, 21677-21680.

Haile,S., Dupe,A., and Papadopoulou,B. (2008). Deadenylation-independent stage-specific mRNA degradation in *Leishmania*. *Nucleic Acids Res.* 36, 1634-1644.

Haile,S., Estevez,A.M., and Clayton,C. (2003). A role for the exosome in the in vivo degradation of unstable mRNAs. *RNA.* 9, 1491-1501.

Haile,S. and Papadopoulou,B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.* 10, 569-577.

Hendriks,E.F., bdul-Razak,A., and Matthews,K.R. (2003). tbCPSF30 depletion by RNA interference disrupts polycistronic RNA processing in *Trypanosoma brucei*. *J. Biol. Chem.* 278, 26870-26878.

Henis-Korenblit,S., Shani,G., Sines,T., Marash,L., Shohat,G., and Kimchi,A. (2002). The caspase-cleaved DAP5 protein supports internal ribosome entry site-mediated translation of death proteins. *Proc. Natl. Acad. Sci. U. S. A* 99, 5400-5405.

Hershey JWB and Merrick WC (2000): The Pathway and Mechanism of Initiation of Protein Synthesis. *Translational Control of Gene Expression* (Cold Spring Harbor Laboratory Press); pp. 33-88.

Holetz,F.B., Correa,A., Avila,A.R., Nakamura,C.V., Krieger,M.A., and Goldenberg,S. (2007). Evidence of P-body-like structures in *Trypanosoma cruzi*. *Biochem. Biophys. Res. Commun.* 356, 1062-1067.

Holzmann,K., Gerner,C., Poltl,A., Schafer,R., Obrist,P., Ensinger,C., Grimm,R., and Sauermann,G. (2000). A human common nuclear matrix protein homologous to eukaryotic translation initiation factor 4A. *Biochem. Biophys. Res. Commun.* 267, 339-344.

Hong,Y. and Sarge,K.D. (1999). Regulation of protein phosphatase 2A activity by heat shock transcription factor 2. *J. Biol. Chem.* 274, 12967-12970.

Hosoda,N., Kobayashi,T., Uchida,N., Funakoshi,Y., Kikuchi,Y., Hoshino,S., and Katada,T. (2003). Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. *J. Biol. Chem.* 278, 38287-38291.

Hosoda,N., Lejeune,F., and Maquat,L.E. (2006). Evidence that poly(A) binding protein C1 binds nuclear pre-mRNA poly(A) tails. *Mol. Cell Biol.* 26, 3085-3097.

Hotchkiss,T.L., Nerantzakis,G.E., Dills,S.C., Shang,L., and Read,L.K. (1999). *Trypanosoma brucei* poly(A) binding protein I cDNA cloning, expression, and binding to 5 untranslated region sequence elements. *Mol. Biochem. Parasitol.* 98, 117-129.

Imataka,H., Gradi,A., and Sonenberg,N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* 17, 7480-7489.

Imataka,H., Olsen,H.S., and Sonenberg,N. (1997). A new translational regulator with homology to eukaryotic translation initiation factor 4G. *EMBO J.* 16, 817-825.

Imataka,H. and Sonenberg,N. (1997). Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol. Cell Biol.* 17, 6940-6947.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchettin G, Borzym K, Bothe G, Bruschi CV, Collins M,

Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Müller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinoiwitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schäfer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442.

Joshi,B., Lee,K., Maeder,D.L., and Jagus,R. (2005). Phylogenetic analysis of eIF4E-family members. *BMC. Evol. Biol.* 5, 48.

Vaynberg, J., and Qin Jun. (2006). Weak protein–protein interactions as probed by NMR spectroscopy. *Trends Biotechnol.* 24, 22-27.

Kennedy,P.G. (2008). The continuing problem of human African trypanosomiasis (sleeping sickness). *Ann. Neurol.* 64, 116-126.

Kessler,S.H. and Sachs,A.B. (1998). RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol. Cell Biol.* 18, 51-57.

Khaleghpour,K., Svitkin,Y.V., Craig,A.W., DeMaria,C.T., Deo,R.C., Burley,S.K., and Sonenberg,N. (2001). Translational repression by a novel partner of human poly(A) binding protein, Paip2. *Mol. Cell* 7, 205-216.

Khanam,T., Rozhdestvensky,T.S., Bundman,M., Galiveti,C.R., Handel,S., Sukonina,V., Jordan,U., Brosius,J., and Skryabin,B.V. (2007). Two primate-specific small non-protein-coding RNAs in transgenic mice: neuronal expression, subcellular localization and binding partners. *Nucleic Acids Res.* 35, 529-539.

Kieft,J.S. (2008). Viral IRES RNA structures and ribosome interactions. *Trends Biochem. Sci.* 33, 274-283.

Kissinger,J.C. (2006). A tale of three genomes: the kinetoplastids have arrived. *Trends Parasitol.* 22, 240-243.

Korneeva,N.L., Lamphear,B.J., Hennigan,F.L., Merrick,W.C., and Rhoads,R.E. (2001). Characterization of the two eIF4A-binding sites on human eIF4G-1. *J. Biol. Chem.* 276, 2872-2879.

Korneeva,N.L., Lamphear,B.J., Hennigan,F.L., and Rhoads,R.E. (2000). Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J. Biol. Chem.* 275, 41369-41376.

Kozlov,G., Trempe,J.F., Khaleghpour,K., Kahvejian,A., Ekiel,I., and Gehring,K. (2001). Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proc. Natl. Acad. Sci. U. S. A* 98, 4409-4413.

Kramer,S., Queiroz,R., Ellis,L., Webb,H., Hoheisel,J.D., 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.

Kuhn,U. and Pieler,T. (1996). Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J. Mol. Biol.* 256, 20-30.

Kuhn,U. and Wahle,E. (2004). Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* 1678, 67-84.

Lamphear,B.J., Kirchweger,R., Skern,T., and Rhoads,R.E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J. Biol. Chem.* 270, 21975-21983.

Le,H., Tanguay,R.L., Balasta,M.L., Wei,C.C., Browning,K.S., Metz,A.M., Goss,D.J., and Gallie,D.R. (1997). Translation initiation factors eIF-iso4G and eIF-4B interact with the poly(A)-binding protein and increase its RNA binding activity. *J. Biol. Chem.* 272, 16247-16255.

Le,H.H., Izaurralde,E., Maquat,L.E., and Moore,M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* 19, 6860-6869.

Le,H.H., Nott,A., and Moore,M.J. (2003). How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* 28, 215-220.

Le,H.H. and Seraphin,B. (2008). EJCs at the heart of translational control. *Cell* 133, 213-216.

LeBowitz,J.H., Smith,H.Q., Rusche,L., and Beverley,S.M. (1993). Coupling of poly(A) site selection and trans-splicing in Leishmania. *Genes Dev.* 7, 996-1007.

Lee,K.A., Edery,I., and Sonenberg,N. (1985). Isolation and structural characterization of cap-binding proteins from poliovirus-infected HeLa cells. *J. Virol.* 54, 515-524.

LeFebvre,A.K., Korneeva,N.L., Trutschl,M., Cvek,U., Duzan,R.D., Bradley,C.A., Hershey,J.W., and Rhoads,R.E. (2006). Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J. Biol. Chem.* 281, 22917-22932.

Lejeune,F. and Maquat,L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.* 17, 309-315.

- Levy-Strumpf,N., Deiss,L.P., Berissi,H., and Kimchi,A. (1997). DAP-5, a novel homolog of eukaryotic translation initiation factor 4G isolated as a putative modulator of gamma interferon-induced programmed cell death. *Mol. Cell Biol.* 17, 1615-1625.
- Li,C.H., Irmer,H., Gudjonsdottir-Planck,D., Freese,S., Salm,H., Haile,S., Estevez,A.M., and Clayton,C. (2006). Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation. *RNA.* 12, 2171-2186.
- Li,Q., Imataka,H., Morino,S., Rogers,G.W., Jr., Richter-Cook,N.J., Merrick,W.C., and Sonenberg,N. (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell Biol.* 19, 7336-7346.
- Li,W., Belsham,G.J., and Proud,C.G. (2001). Eukaryotic initiation factors 4A (eIF4A) and 4G (eIF4G) mutually interact in a 1:1 ratio in vivo. *J. Biol. Chem.* 276, 29111-29115.
- Liang,X.H., Haritan,A., Uliel,S., and Michaeli,S. (2003). Trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot. Cell* 2, 830-840.
- Linder,P. (2003). Yeast RNA helicases of the DEAD-box family involved in translation initiation. *Biol. Cell* 95, 157-167.
- Ling,J., Morley,S.J., and Traugh,J.A. (2005). Inhibition of cap-dependent translation via phosphorylation of eIF4G by protein kinase Pak2. *EMBO J.* 24, 4094-4105.
- Lindqvist L, Oberer M, Reibarkh M, Cencic R, Bordeleau ME, Vogt E, Marintchev A, Tanaka J, Fagotto F, Altmann M, Wagner G, Pelletier J. 2008. Selective Pharmacological Targeting of a DEAD Box RNA Helicase. *PLoS One.* 13;3(2):e1583.
- Lorsch,J.R. and Herschlag,D. (1998). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry* 37, 2180-2193.
- Lukes J, Mauricio IL, Schönián G, Dujardin JC, Soteriadou K, Dedet JP, Kuhls K, Tintaya KW, Jirků M, Chocholová E, Haralambous C, Pratlong F, Oborník M, Horák A, Ayala FJ, Miles MA. (2007). Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. *Proc. Natl. Acad. Sci. U. S. A* 104, 9375-9380.
- Luking,A., Stahl,U., and Schmidt,U. (1998). The protein family of RNA helicases. *Crit Rev. Biochem. Mol. Biol.* 33, 259-296.
- Mader,S., Lee,H., Pause,A., and Sonenberg,N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol. Cell Biol.* 15, 4990-4997.
- Mair,G. *et al.* (2000). A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA.* 6, 163-169.

- Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol. Cell* 3, 707-716.
- Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89, 951-961.
- Marcotrigiano, J., Lomakin, I.B., Sonenberg, N., Pestova, T.V., Hellen, C.U., and Burley, S.K. (2001). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell* 7, 193-203.
- Marintchev, A., Edmonds, K.A., Marintcheva, B., Hendrickson, E., Oberer, M., Suzuki, C., Herdy, B., Sonenberg, N., and Wagner, G. (2009). Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell* 136, 447-460.
- Marintchev, A. and Wagner, G. (2004). Translation initiation: structures, mechanisms and evolution. *Q. Rev. Biophys.* 37, 197-284.
- Marintchev, A. and Wagner, G. (2005). eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G. *Biochemistry* 44, 12265-12272.
- Martineau, Y., Derry, M.C., Wang, X., Yanagiya, A., Berlanga, J.J., Shyu, A.B., Imataka, H., Gehring, K., and Sonenberg, N. (2008). Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation. *Mol. Cell Biol.* 28, 6658-6667.
- Martinez-Calvillo, S., Nguyen, D., Stuart, K., and Myler, P.J. (2004). Transcription initiation and termination on Leishmania major chromosome 3. *Eukaryot. Cell* 3, 506-517.
- Matsuo, H., Li, H., McGuire, A.M., Fletcher, C.M., Gingras, A.C., Sonenberg, N., and Wagner, G. (1997). Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat. Struct. Biol.* 4, 717-724.
- Mayberry, L.K., Allen, M.L., Dennis, M.D., and Browning, K.S. (2009). Evidence for variation in the optimal translation initiation complex: Plant eIF4B, eIF4F and eIF(iso)4F Differentially Promote Translation of mRNAs. *Plant Physiol.*
- Melo, E.O., Dhalia, R., Martins de, S.C., Standart, N., and de Melo Neto, O.P. (2003). Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: role in cooperative binding to poly (A) and efficient cap distal translational repression. *J. Biol. Chem.* 278, 46357-46368.
- Milone, J., Wilusz, J., and Bellofatto, V. (2002). Identification of mRNA decapping activities and an ARE-regulated 3' to 5' exonuclease activity in trypanosome extracts. *Nucleic Acids Res.* 30, 4040-4050.

Mitra,B. and Ray,D.S. (2004). Presence of a poly(A) binding protein and two proteins with cell cycle-dependent phosphorylation in *Crithidia fasciculata* mRNA cycling sequence binding protein II. *Eukaryot. Cell* 3, 1185-1197.

Moerke,N.J. *et al.* (2007). Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* 128, 257-267.

Moore,M.J. and Proudfoot,N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* 136, 688-700.

Morino,S., Imataka,H., Svitkin,Y.V., Pestova,T.V., and Sonenberg,N. (2000). Eukaryotic translation initiation factor 4E (eIF4E) binding site and the middle one-third of eIF4GI constitute the core domain for cap-dependent translation, and the C-terminal one-third functions as a modulatory region. *Mol. Cell Biol.* 20, 468-477.

Moura, D.M.N. (2007). Utilização de abordagens moleculares *in vivo* no estudo de homólogos do fator de iniciação da tradução 4G (eIF4G) de tripanosomatídeos. Dissertação de Mestrado, Universidade Federal de Pernambuco, Recife-PE.

Nevins,T.A., Harder,Z.M., Korneluk,R.G., and Holcik,M. (2003). Distinct regulation of internal ribosome entry site-mediated translation following cellular stress is mediated by apoptotic fragments of eIF4G translation initiation factor family members eIF4GI and p97/DAP5/NAT1. *J. Biol. Chem.* 278, 3572-3579.

Nielsen,P.J. and Trachsel,H. (1988). The mouse protein synthesis initiation factor 4A gene family includes two related functional genes which are differentially expressed. *EMBO J.* 7, 2097-2105.

Oberer,M., Marintchev,A., and Wagner,G. (2005). Structural basis for the enhancement of eIF4A helicase activity by eIF4G. *Genes Dev.* 19, 2212-2223.

Palacios, I.M., Gatfield, D., St Johnston., D., and Izaurralde E (2004): An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature*; 427(6976); pp. 753-757.

Palenchar,J.B. and Bellofatto,V. (2006). Gene transcription in trypanosomes. *Mol. Biochem. Parasitol.* 146, 135-141.

Parker,R. and Song,H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* 11, 121-127.

Parra-Palau,J.L., Scheper,G.C., Wilson,M.L., and Proud,C.G. (2003). Features in the N and C termini of the MAPK-interacting kinase Mnk1 mediate its nucleocytoplasmic shuttling. *J. Biol. Chem.* 278, 44197-44204.

Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream MA, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead

S, Rabinowitsch E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LR, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M. (2007). Comparative genomic analysis of three Leishmania species that cause diverse human disease. *Nat. Genet.* 39, 839-847.

Pelletier,J. and Sonenberg,N. (1985). Photochemical cross-linking of cap binding proteins to eucaryotic mRNAs: effect of mRNA 5' secondary structure. *Mol. Cell Biol.* 5, 3222-3230.

Pestova, T., Lorsch, J.R., Hellen, C.U.T. (2007): The Mechanism of Translation Initiation in Eukaryotes. *Translational Control in Biology and Medicine* (Cold Spring Harbor Laboratory Press); pp. 87-128.

Pitula,J., Ruyechan,W.T., and Williams,N. (1998). Trypanosoma brucei: identification and purification of a poly(A)-binding protein. *Exp. Parasitol.* 88, 157-160.

Prevot,D., Darlix,J.L., and Ohlmann,T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biol. Cell* 95, 141-156.

Ptushkina,M., von der,H.T., Karim,M.M., Hughes,J.M., and McCarthy,J.E. (1999). Repressor binding to a dorsal regulatory site traps human eIF4E in a high cap-affinity state. *EMBO J.* 18, 4068-4075.

Ptushkina,M., von der,H.T., Vasilescu,S., Frank,R., Birkenhager,R., and McCarthy,J.E. (1998). Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *EMBO J.* 17, 4798-4808.

Pyronnet,S., Dostie,J., and Sonenberg,N. (2001). Suppression of cap-dependent translation in mitosis. *Genes Dev.* 15, 2083-2093.

Pyronnet,S., Imataka,H., Gingras,A.C., Fukunaga,R., Hunter,T., and Sonenberg,N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J.* 18, 270-279.

Qin,H., Raught,B., Sonenberg,N., Goldstein,E.G., and Edelman,A.M. (2003). Phosphorylation screening identifies translational initiation factor 4GII as an intracellular target of Ca(2+)/calmodulin-dependent protein kinase I. *J. Biol. Chem.* 278, 48570-48579.

Raught,B., Gingras,A.C., Gygi,S.P., Imataka,H., Morino,S., Gradi,A., Aebersold,R., and Sonenberg,N. (2000). Serum-stimulated, rapamycin-sensitive phosphorylation sites in the eukaryotic translation initiation factor 4GI. *EMBO J.* 19, 434-444.

Reithinger,R. (2008). Leishmaniases' burden of disease: ways forward for getting from speculation to reality. *PLoS. Negl. Trop. Dis.* 2, e285.

Reithinger,R., Dujardin,J.C., Louzir,H., Pirmez,C., Alexander,B., and Brooker,S. (2007). Cutaneous leishmaniasis. *Lancet Infect. Dis.* 7, 581-596.

Rhoads,R.E. (2009). eIF4E: New Family Members, New Binding Partners, New Roles. *J. Biol. Chem.* 284, 16711-16715.

Richter,J.D. and Sonenberg,N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477-480.

Rogers,G.W., Jr., Komar,A.A., and Merrick,W.C. (2002). eIF4A: the godfather of the DEAD box helicases. *Prog. Nucleic Acid Res. Mol. Biol.* 72, 307-331.

Rozovsky,N., Butterworth,A.C., and Moore,M.J. (2008). Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. *RNA*. 14, 2136-2148.

Ruan,J.P., Arhin,G.K., Ullu,E., and Tschudi,C. (2004). Functional characterization of a *Trypanosoma brucei* TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol. Cell Biol.* 24, 9610-9618.

Sambrook, J. and Russell, D.W. (2001). *Molecular Cloning, a laboratory manual*. Cold Spring Harbor Laboratory Press.

Sachs,A.B. and Varani,G. (2000). Eukaryotic translation initiation: there are (at least) two sides to every story. *Nat. Struct. Biol.* 7, 356-361.

Santos DO, Coutinho CE, Madeira MF, Bottino CG, Vieira RT, Nascimento SB, Bernardino A, Bourguignon SC, Corte-Real S, Pinho RT, Rodrigues CR, Castro HC. (2008). Leishmaniasis treatment--a challenge that remains: a review. *Parasitol. Res.* 103, 1-10.

Scheper,G.C. and Proud,C.G. (2002). Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *Eur. J. Biochem.* 269, 5350-5359.

Shibuya, T., Tange, T.O., Sonenberg, N., Moore, M.J. (2004): eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nat. Struct. Mol. Biol.*;11(4); pp. 346-351.

Shibuya,T., Tange,T.O., Stroupe,M.E., and Moore,M.J. (2006). Mutational analysis of human eIF4AIII identifies regions necessary for exon junction complex formation and nonsense-mediated mRNA decay. *RNA*. 12, 360-374.

Simmons,H.M., Ruis,B.L., Kapoor,M., Hudacek,A.W., and Conklin,K.F. (2005). Identification of NOM1, a nucleolar, eIF4A binding protein encoded within the chromosome 7q36 breakpoint region targeted in cases of pediatric acute myeloid leukemia. *Gene* 347, 137-145.

Simpson,A.G., Stevens,J.R., and Lukes,J. (2006). The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol.* 22, 168-174.

Skeiky YA, Kennedy M, Kaufman D, Borges MM, Guderian JA, Scholler JK, Ovendale PJ, Picha KS, Morrissey PJ, Grabstein KH, Campos-Neto A, Reed SG. (1998). LeIF: a

recombinant *Leishmania* protein that induces an IL-12-mediated Th1 cytokine profile. *J. Immunol.* 161, 6171-6179.

Sladic,R.T., Lagnado,C.A., Bagley,C.J., and Goodall,G.J. (2004). Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. *Eur. J. Biochem.* 271, 450-457.

Smith,D.F., Peacock,C.S., and Cruz,A.K. (2007). Comparative genomics: from genotype to disease phenotype in the leishmaniasis. *Int. J. Parasitol.* 37, 1173-1186.

Sonenberg,N. (2008). eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem. Cell Biol.* 86, 178-183.

Sonenberg,N. and Dever,T.E. (2003). Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.* 13, 56-63.

Stiles,J.K., Hicock,P.I., Kong,L., Xue,L., and Meade,J.C. (1999). *Leishmania donovani* proton translocating P-type adenosine triphosphatases LDH1A and LDH1B: trans-splicing and polyadenylation of transcripts in amastigotes and promastigotes. *Mol. Biochem. Parasitol.* 103, 105-109.

Stoecklin, G., Mayo, T., and Anderson, P. (2006). ARE-mRNA degradation requires the 5'-3' decay pathway. *EMBO Rep*;7:72-7.

Stuart,K., Brun,R., Croft,S., Fairlamb,A., Gurtler,R.E., McKerrow,J., Reed,S., and Tarleton,R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. *J. Clin. Invest* 118, 1301-1310.

Suzuki,C., Garces,R.G., Edmonds,K.A., Hiller,S., Hyberts,S.G., Marintchev,A., and Wagner,G. (2008). PDCD4 inhibits translation initiation by binding to eIF4A using both its MA3 domains. *Proc. Natl. Acad. Sci. U. S. A* 105, 3274-3279.

Svitkin,Y.V. and Sonenberg,N. (2006). Translational control by the poly(A) binding protein: a check for mRNA integrity. *Mol. Biol. (Mosk)* 40, 684-693.

Tange,T.O., Nott,A., and Moore,M.J. (2004). The ever-increasing complexities of the exon junction complex. *Curr. Opin. Cell Biol.* 16, 279-284.

Tarleton,R.L., Reithinger,R., Urbina,J.A., Kitron,U., and Gurtler,R.E. (2007). The challenges of Chagas Disease-- grim outlook or glimmer of hope. *PLoS. Med.* 4, e332.

Tarun,S.Z., Jr. and Sachs,A.B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* 15, 7168-7177.

Tarun,S.Z., Jr., Wells,S.E., Deardorff,J.A., and Sachs,A.B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proc. Natl. Acad. Sci. U. S. A* 94, 9046-9051.

- Uchida,N., Hoshino,S., Imataka,H., Sonenberg,N., and Katada,T. (2002). A novel role of the mammalian GSPT/eRF3 associating with poly(A)-binding protein in Cap/Poly(A)-dependent translation. *J. Biol. Chem.* 277, 50286-50292.
- von der,H.T., Ball,P.D., and McCarthy,J.E. (2000). Stabilization of eukaryotic initiation factor 4E binding to the mRNA 5'-Cap by domains of eIF4G. *J. Biol. Chem.* 275, 30551-30555.
- von der, H.T., and McCarthy, J.E. (2002): Intracellular translation initiation factor levels in *Saccharomyces cerevisiae* and their role in cap-complex function. *Mol. Microbiol.* ;46; pp. 531-544.
- Weinstein,D.C., Honore,E., and Hemmati-Brivanlou,A. (1997). Epidermal induction and inhibition of neural fate by translation initiation factor 4AIII. *Development* 124, 4235-4242.
- Wells,S.E., Hillner,P.E., Vale,R.D., and Sachs,A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* 2, 135-140.
- Woods,A.J., Kantidakis,T., Sabe,H., Critchley,D.R., and Norman,J.C. (2005). Interaction of paxillin with poly(A)-binding protein 1 and its role in focal adhesion turnover and cell migration. *Mol. Cell Biol.* 25, 3763-3773.
- Yamanaka,S., Poksay,K.S., Arnold,K.S., and Innerarity,T.L. (1997). A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev.* 11, 321-333.
- Yan,R., Rychlik,W., Etchison,D., and Rhoads,R.E. (1992). Amino acid sequence of the human protein synthesis initiation factor eIF-4 gamma. *J. Biol. Chem.* 267, 23226-23231.
- Yanagiya,A., Svitkin,Y.V., Shibata,S., Mikami,S., Imataka,H., and Sonenberg,N. (2009). Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol. Cell Biol.* 29, 1661-1669.
- Yang,H.S., Cho,M.H., Zakowicz,H., Hegamyer,G., Sonenberg,N., and Colburn,N.H. (2004). A novel function of the MA-3 domains in transformation and translation suppressor Pcd4 is essential for its binding to eukaryotic translation initiation factor 4A. *Mol. Cell Biol.* 24, 3894-3906.
- Yang,H.S., Jansen,A.P., Komar,A.A., Zheng,X., Merrick,W.C., Costes,S., Lockett,S.J., Sonenberg,N., and Colburn,N.H. (2003). The transformation suppressor Pcd4 is a novel eukaryotic translation initiation factor 4A binding protein that inhibits translation. *Mol. Cell Biol.* 23, 26-37.
- Yoder-Hill,J., Pause,A., Sonenberg,N., and Merrick,W.C. (1993). The p46 subunit of eukaryotic initiation factor (eIF)-4F exchanges with eIF-4A. *J. Biol. Chem.* 268, 5566-5573.

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, 3243-3253.

Yoffe,Y., Zuberek,J., Lerer,A., Lewdorowicz,M., Stepinski,J., Altmann,M., Darzynkiewicz,E., and Shapira,M. (2006). Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania. *Eukaryot. Cell* 5, 1969-1979.

Yoffe,Y., Zuberek,J., Lewdorowicz,M., Zeira,Z., Keasar,C., Orr-Dahan,I., Jankowska-Anyszka,M., Stepinski,J., Darzynkiewicz,E., and Shapira,M. (2004). Cap-binding activity of an eIF4E homolog from Leishmania. *RNA*. 10, 1764-1775.

10. ANEXOS (ARTIGOS CIENTÍFICOS)

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 Carrington¹ 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 *TbEIF4AI* and *TbEIF4AIII*). 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. *TbEIF4AI* is a very abundant predominantly cytoplasmic protein (over 1×10^5 molecules/cell) and depletion to $\sim 10\%$ of normal levels through RNA interference dramatically reduces protein synthesis one cell cycle following double-stranded RNA induction and stops cell proliferation. In contrast, *TbEIF4AIII* is a nuclear, moderately expressed protein ($\sim 1\text{--}2 \times 10^4$ molecules/cell), and its depletion stops cellular proliferation after approximately four cell cycles. Ectopic expression of a dominant negative mutant of *TbEIF4AI*, but not of *TbEIF4AIII*, induced a slow growth phenotype in transfected cells. Overall, our results suggest that only *TbEIF4AI* is involved in protein synthesis while the properties and sequence of *TbEIF4AIII* 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

eIF4A 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, eIF4A 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 eIF4A (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 implicated 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, eIF4A binds to the central region of eIF4G, via the eIF4G HEAT domain (15) and, in mammals at least, also to the eIF4G C-terminus (16,17). eIF4A 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 eIF4A have been described. Both eIF4AI and II (90% identity between the two proteins) are able to reconstitute the eIF4F subunit and presumably have similar roles in translation (20,21). In contrast, eIF4AIII, only 66% identical to mammalian eIF4AI, is functionally distinct. While eIF4AIII 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). eIF4AIII 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* eIF4A 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 eIF4F subunits, all of which are conserved in *T. brucei* (33). We characterized two putative *L. major* eIF4A homologues, *LmEIF4A1* (LeiF) and *LmEIF4A2*, with 59 and 52% identities to human eIF4AI, respectively. When assayed with isoform specific antibodies these two factors differ significantly in abundance in *L. major* promastigotes. *LmEIF4A1* is very abundant with over 10^5 molecules/cell whilst *LmEIF4A2* is either absent or present at levels below 10^4 molecules/cell. Furthermore, only *LmEIF4A1* was found to bind specifically to the HEAT domain of one of the *Leishmania* eIF4G 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 eIF4A homologues. Initially, the mRNA and protein levels of the two *T. brucei* eIF4A 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 *LmEIF4A1* (named as *TbEIF4AI*) is the functional homologue of the eIF4A present in eIF4F. As for the orthologue of *LmEIF4A2*, it seems to be the functional homologue of the nuclear eIF4AIII present in higher eukaryotes and has been named here as *TbEIF4AIII*.

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 *TbEIF4AI* coding sequence was amplified from *T. brucei* Lister 427 genomic DNA (5'primer, AAG CTT 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 *TbEIF4AIII* sequence (5'primer, AAG CTT CCG CCA CCA TGA CAG CAA CCG CAA GG; and 3'primer, GGA TCC AGA ACC 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 *TbEIF4AI*-EYFP and *TbEIF4AIII*-EYFP constructs, the two eIF4A 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 pLEW82. On expression, the resultant fusion protein had the sequence: eIF4A C-terminal residue-GSGSGGG-EYFP. For the RNAi experiments the same two eIF4A DNA fragments were also subcloned into the HindIII/BamHI sites of the transfection vector p2T7-177 (35). Dominant negative mutants were made by altering the sequence of the DEAD box of the helicase (motif II in Figure 1) to DQAD (11,36). Tetracycline-inducible expression of wild type and dominant negative forms of eIF4A was performed using p2280, a derivative of pLEW100 made by introducing a BamHI/BglII DNA segment encoding three tandem myc epitope tags to its BamHI site. The two HindIII/BamHI *T. brucei* eIF4A fragments were cloned into the same sites of p2280 resulting in the expression of fusion proteins with the myc epitope tags on their C-terminus giving the sequence

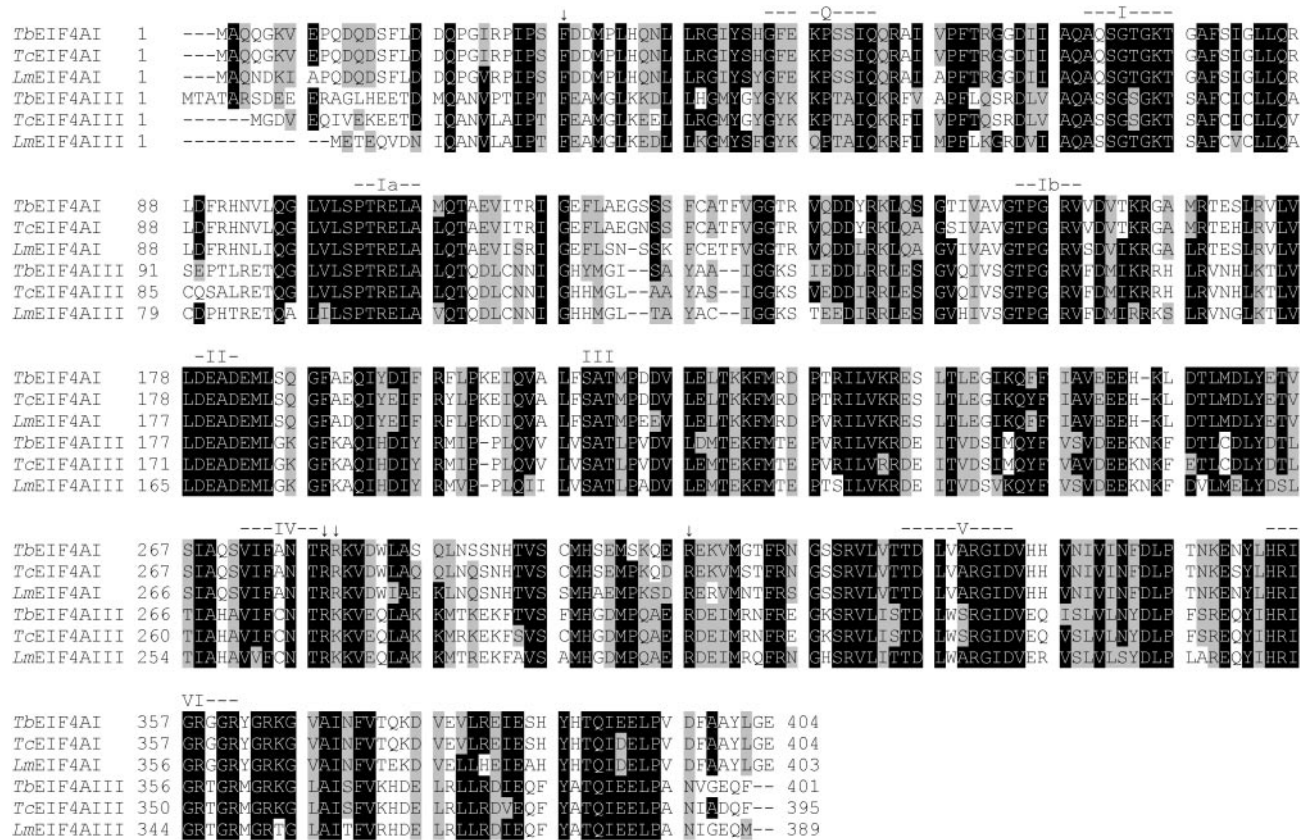


Figure 1. Sequence alignment comparing the *T. brucei*, *T. cruzi* and *L. major* eIF4A homologues. Sequences were aligned with the Clustal W program, from the Centre for Molecular and Biomolecular Informatics (<http://www.cmbi.kun.nl/bioinf/tools/clustalw.shtml>). Amino acids identical in >60% of the sequences are highlighted in dark gray, while amino acids defined as similar, based on the BLOSUM 62 Matrix, on >60% of the sequences, are shown in pale gray. When necessary, gaps were inserted within the various sequences (dashes) to allow better alignment. The nine motifs typical of DEAD-box RNA helicases (10,11) are highlighted. The single arrows indicate other individual amino acids which seems to be relevant for eIF4A function or RNA binding (12,42). Relevant GenBank accession numbers: *LmEIF4AI*, AAC24684/AAC24685; *LmEIF4AIII*, CAJ05468; *TbeIF4AI*, EAN76544; *TbeIF4AIII*, EAN79829; *TcEIF4AI*, EAN98527; *TcEIF4AIII*, EAN88971.

eIF4A-GSGSGPREQKLISEEDLPREQKLISEEDLPREQKLISEEDLP.

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 µg/ml) and hygromycin (25 µg/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 µg/ml). For the RNAi experiments 1 µg/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 *TbeIF4As* 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 *TbeIF4AI* and *TbeIF4AIII* 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 polyclonal serum (Covalab). Prior to their use, both antibodies were first affinity purified as described elsewhere (40) with their respective recombinant proteins. Cross-reacting antibodies were eliminated by previous incubation of the anti-*TbeIF4AI* antisera with *TbeIF4AIII* recombinant protein and vice versa. To estimate the levels of the eIF4A proteins, first the recombinant proteins were quantified by serial dilutions in SDS-PAGE by comparison

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 *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 $-20^\circ\text{C}/15$ min. Antibody detection of *TbEIF4AI* and *III* followed standard procedures. DNA was stained using Hoechst 33258. For the analysis of the cells expressing *TbEIF4AI-EYFP* and *TbEIF4AIII-EYFP*, aliquots of 5×10^6 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, [^{35}S]methionine (10 $\mu\text{Ci}/\text{ml}$) 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 $\mu\text{g}/\text{ml}$ 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 $\mu\text{Ci}/\text{ml}$ [^{35}S]methionine and incubated for 1 h at 28°C prior to harvesting and analysis by SDS-PAGE and autoradiography.

RESULTS

Identification of the *T.brucei* eIF4A homologues

The *T.brucei* homologues of eIF4A were identified in searches of the genome sequence using human eIF4AI as well as the two *Leishmania* eIF4A sequences. At the amino acid level, the two *T.brucei* proteins, *TbEIF4AI* and *TbEIF4AIII*, are very similar to their *L.major* orthologues with identities of 88% for the eIF4AI and 82% for the eIF4AIII pair (to avoid confusion and in view of the data presented below the *L.major* eIF4A homologues, previously called *LmEIF4A1* and 2, have also been renamed to *LmEIF4AI* and *LmEIF4AIII* and this nomenclature will be used when needed). The assignment as eIF4A 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 *T. brucei* and *L. major* databases has been assigned to another group of RNA helicases, Dhh1 (33), with an identity of only $\sim 40\%$ to human eIF4AI. In order to analyse the conservation of the putative eIF4A homologues within an additional member of the family *Trypanosomatidae*, we performed similar searches using the *T.cruzi* genome database. Again, orthologues to both

proteins could be found in *T.cruzi* with the third nearest match to human eIF4AI being Dhh1.

Figure 1 shows a sequence alignment comparing the two eIF4A sequences from *T.brucei* with those from *T.cruzi* and *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 eIF4AI and three 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 (V328W) and VI (G359T). So far however the functional significance of these substitutions is unknown.

Expression of *TbEIF4A* mRNAs in *T.brucei* bloodstream and procyclic forms

To begin the functional characterization of *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 *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 *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 *TbEIF4AI* mRNA produces a much stronger signal than *TbEIF4AIII*. Since both probes used were of similar specific activity, and the exposures times for the films were similar as well, it seems that the *TbEIF4AIII* mRNA is far less abundant than *TbEIF4AI*. Remarkably, although the ORFs for both proteins are similar, 1215 versus 1206 bp for *TbEIF4AI* and *TbEIF4AIII*, respectively, their mRNAs differ significantly, with the *TbEIF4AI* message, at ~ 3 kb, being nearly twice the length of *TbEIF4AIII* (~ 1.6 kb), probably reflecting a considerable difference in the length of the 3'-UTR.

TbEIF4AI is about 10-fold more abundant than *TbEIF4AIII* in both procyclic and bloodstream forms

Recombinant His-*TbEIF4AI* and *III* were expressed in *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

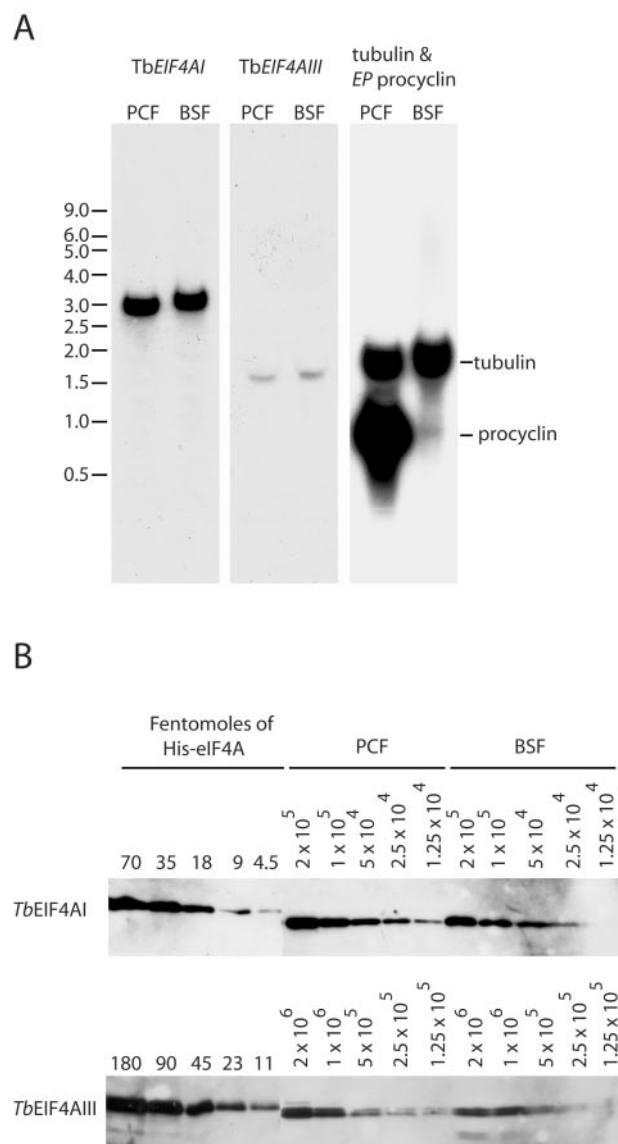


Figure 2. Expression analysis of *TbEIF4AI* and *III*. (A) Total RNA from both procyclic (PCF) and bloodstream (BSF) *T. brucei* forms was separated on denaturing gels and used in northern blot assays to detect the expression of *TbEIF4AI* and *III*. One of the blots was overprobed with tubulin (ubiquitously expressed) and EP procyclin (expressed in procyclics only) as controls. The migration of RNA size markers is indicated on the left in kilobases. (B) Quantification of *TbEIF4AI* and *TbEIF4AIII* in the procyclic and bloodstream forms of *T. brucei*. Recombinant His-tagged *TbEIF4AI* and *III* were quantified, diluted to defined concentrations (in fmol) and ran on SDS-PAGE gels with whole parasite extract obtained from known number of cells from both procyclic and bloodstream forms (1.25×10^4 – 2×10^5 for *TbEIF4AI* and 1.25×10^5 – 2×10^6 for *TbEIF4AIII*). 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 femtomoles/ 10^5 or 10^6 cells were then converted in number of molecules/cell.

antibodies were then used in western blots to analyse the expression of both proteins as well as to estimate their intracellular levels. *TbEIF4AI* is very abundant (Figure 2B) and although the quantification is only approximate, its levels were estimated at ~ 2 – 5×10^5 and 0.8 – 1.5×10^5 molecules/cell in procyclic and bloodstream forms, respectively, this difference

being a reflection of the relative volumes of the two cell types. These levels are compatible with what has been observed with the *L. major* orthologue (33) as well as yeast eIF4A (44). In contrast, *TbEIF4AIII* levels were estimated at $\sim 2 \times 10^4$ and 1×10^4 molecules/cell in procyclic and bloodstream forms, respectively (Figure 2B). These data indicate that *TbEIF4AI* is present at levels at least 10-fold higher than *TbEIF4AIII*, a difference which is reminiscent of the situation with the *L. major* orthologues (33). Since there are estimated to be $\sim 50\,000$ mRNAs per procyclic cell (Supplementary Data), *TbEIF4AI* is in excess relative to mRNA, in contrast to *TbEIF4AIII*. Overall we conclude that both proteins are expressed constitutively and that only the *TbEIF4AI* levels are compatible with a role in translation.

Subcellular localization of *TbEIF4AI* and *III*

To determine the subcellular localization of the *T. brucei* eIF4A homologues we used two different experimental approaches. First the *TbEIF4AI* and *TbEIF4AIII* ORFs were cloned into the vector p2215 and the construct integrated into the non-transcribed spacer of a ribosomal RNA gene locus in the procyclic cell line Lister 427 29-13. This resulted in a tetracycline-inducible transgene encoding the eIF4A 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 *TbEIF4AI* 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. *TbEIF4AI*-EYFP is found predominantly in the cytoplasm, whilst *TbEIF4AIII*-EYFP is only found in the nucleus. These results were confirmed for the endogenous proteins through indirect immunofluorescence using isoform specific antibodies: again *TbEIF4AI* was mainly found in the cytoplasm whilst *TbEIF4AIII* was only detected in the nucleus (Figure 3).

RNAi of *TbEIF4AI* and *TbEIF4AIII*

The function of the two eIF4A 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 procyclic 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 *TbEIF4AI* 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 *TbEIF4AI* 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 *TbEIF4AI* 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.

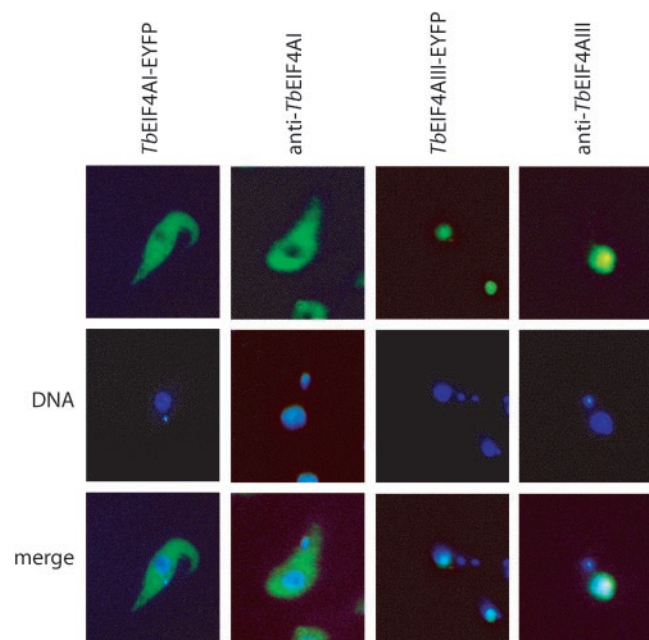


Figure 3. Subcellular localization of *TbEIF4AI* and *III* in *T.brucei* procyclic forms. Subcellular localization of the *TbEIF4AI* and *III* /EYFP fusion proteins in transfected *T.brucei* cells was examined with a fluorescence microscope. The localization of native *TbEIF4AI* and *III* was also confirmed in wild-type procyclic cells (WT 427) by indirect immunofluorescence using the *TbEIF4AI* 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.

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, residuals 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 *TbEIF4AI* since no reduction in levels of *TbEIF4AIII* was observed in cells submitted to the *TbEIF4AI* RNAi procedure (Figure 5B). Likewise the *TbEIF4AIII* RNAi does not lead to any reduction in the levels of *TbEIF4AI* (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 *TbEIF4AI*, 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 *TbEIF4AI* and *III* in transfected procyclic cells

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

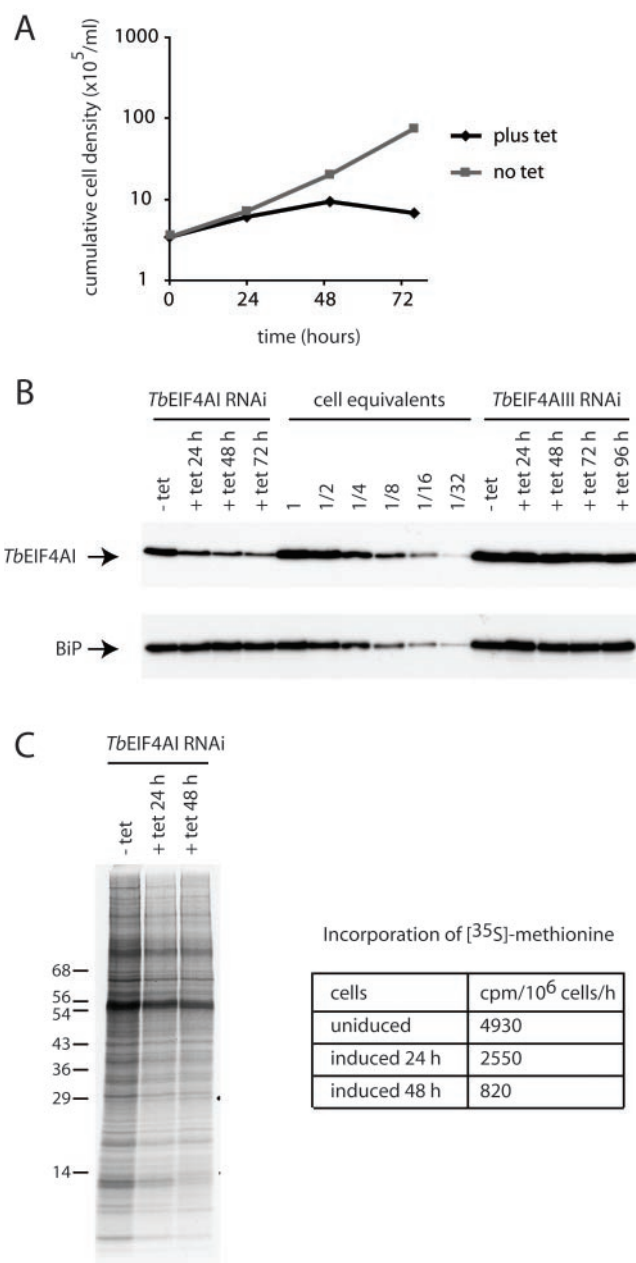


Figure 4. RNAi of *TbEIF4AI*. Procyclic *T.brucei* cells were transfected with the p2T7-177 derived plasmid containing the *TbEIF4AI* gene. Transfected cells were selected after growth in the presence of phleomycin and RNA interference induced after tetracycline addition. At regular intervals, cellular growth was monitored by counting the number of viable cells, expression of *TbEIF4AI* 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 10⁶ cells and was used in the various RNAi lanes). *TbEIF4AI* was detected with the affinity purified antisera 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 *TbEIF4AIII* 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 *TbEIF4AI* 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.

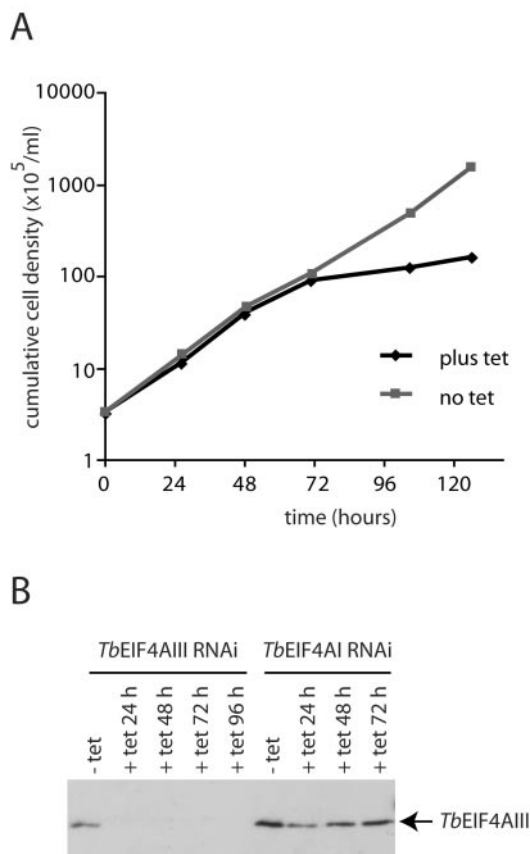


Figure 5. RNAi of *TbEIF4AIII*. Procyclic *T.brucei* cells were transfected with the p2T7-177/ *TbEIF4AIII* construct as described for Figure 4, monitored for cellular growth and assayed for expression of *TbEIF4AIII*. (A) Cell density of transfected cultures at different time points with and without tetracycline addition. (B) Western blot analysis of the time course for both the *TbEIF4AIII* and *TbEIF4AI* RNAi experiments using the *TbEIF4AIII* antibodies. Samples from the same experiment assayed in Figure 4B were assayed for *TbEIF4AIII* expression.

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 potent inhibition 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 *TbEIF4AIII* transgenes was several fold higher than the endogenous protein. All the myc-tagged transgenes localized correctly (data not shown).

Cell growth and transgene expression was monitored over a time course (Figure 6B). Expression from the *TbEIF4AI* transgene reduced over the time course and was barely detectable by 104 h. We have observed this diminution of expression over time with other, but not all, transgenes expressed from vectors derived from pLEW100 and are unsure of the cause. The only 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 *TbEIF4AIII* mutant transgene had no effect on growth of the culture. Interestingly, the expression of the *TbEIF4AIII* 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 *TbEIF4AIII*, 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 *TbEIF4AIII* is not active in translation.

Mapping of isoform specific amino acids

The results described above for *TbEIF4AIII* are reminiscent of what is known of mammalian eIF4AIII (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*, Fallp, a nucleolar protein shown to be required for 40S ribosomal subunit formation (45). Fallp, however, does not seem to be involved in EJC

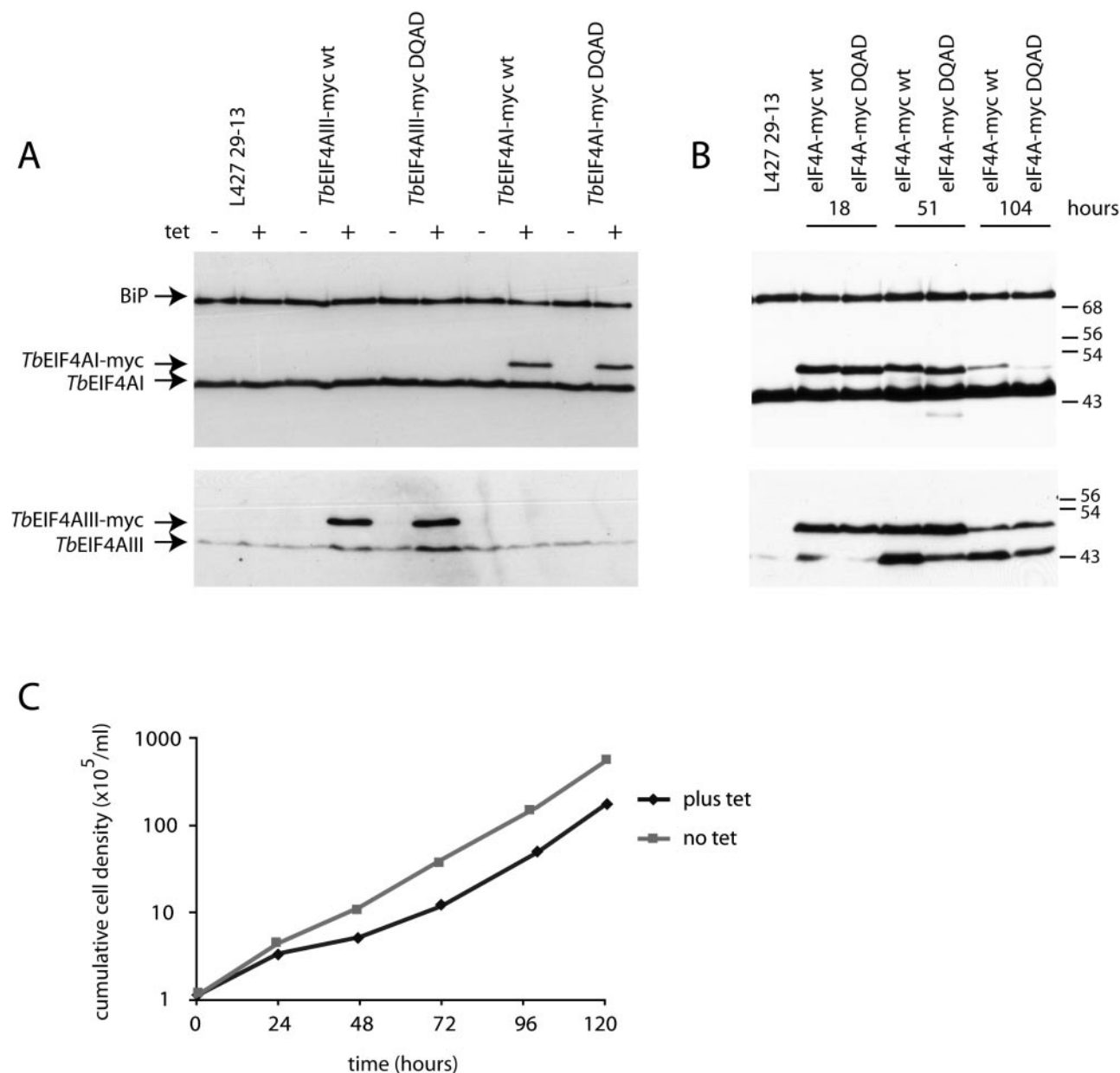


Figure 6. Expression of myc-tagged dominant negative mutants of *TbEIF4AI* and *III* in procyclic cells. (A) Western blot analysis of the expression of the various *TbEIF4AI* and *III*/myc fusions in transfected cells in the absence or after exposure to tetracycline for 18 h. In each case the expression was detected using antibodies specific to each of the eIF4A homologues. The *TbEIF4AI* western blot was simultaneously probed with anti-BiP as a loading control. (B) Time course expression of the different versions of *TbEIF4A-myc* after tetracycline addition to the culture. The *TbEIF4AI* western blot was simultaneously probed with anti-BiP as a loading control. (C) Effect of the expression of the dominant negative form of *TbEIF4AI-myc* on the growth of the transfected cells in culture.

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 (*TbEIF4A* I 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 eIF4AI/III function in general (see below).

Molecular modelling of *TbEIF4AI* and *TbEIF4AIII*

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

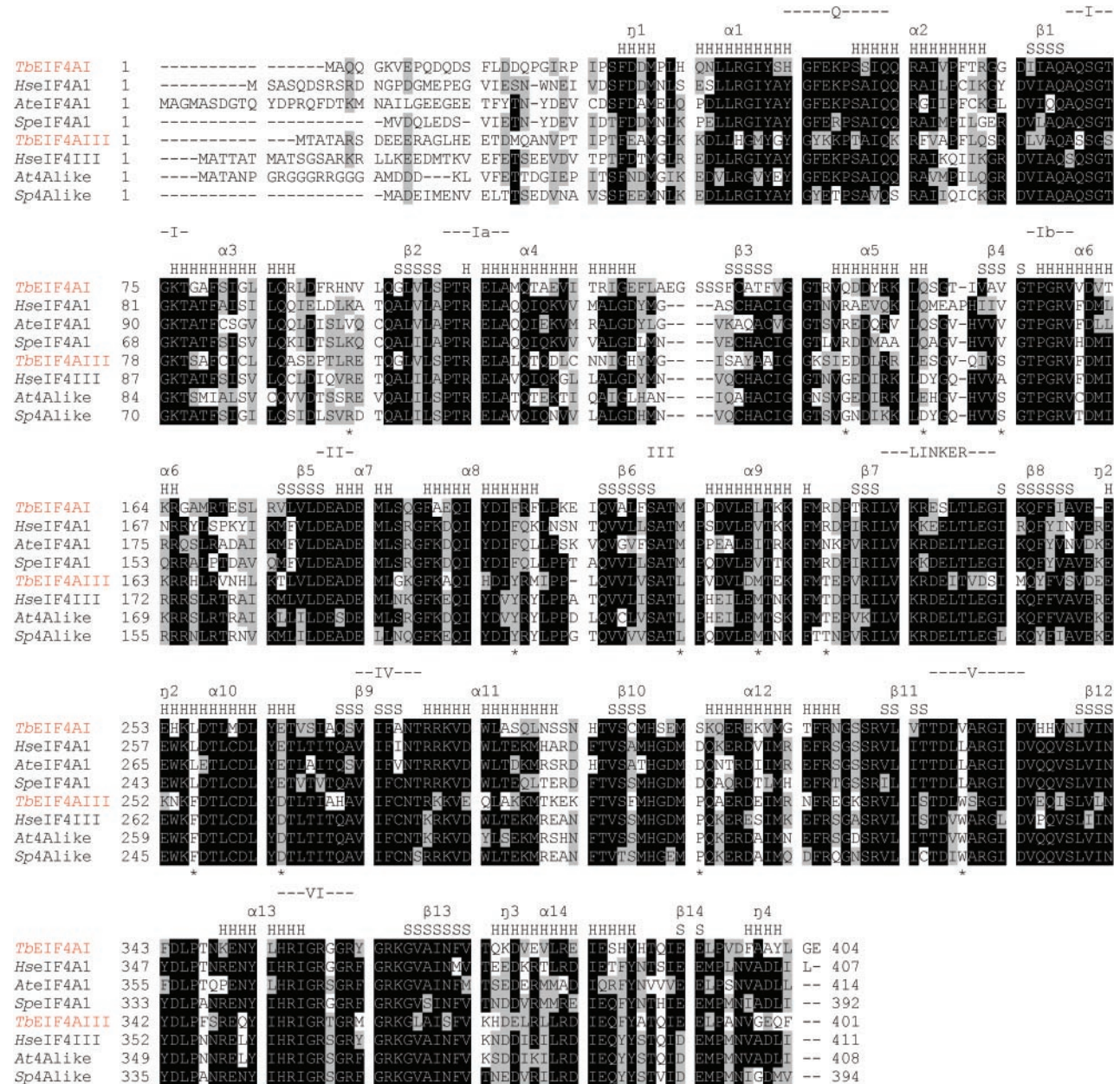


Figure 7. Sequence alignment comparing *TbeIF4AI* and *III* with the putative eIF4AI and eIF4AIII from selected organisms. (A) Sequences were aligned as described in Figure 1 and the various DEAD-box motifs are shown as indicated previously. The predicted secondary structural elements derived from the modelling shown in Figure 8 and from Ref. (46) are indicated numbered $\alpha 1$ – $\alpha 13/\eta 1$ – $\eta 4$ (alpha-helices—H) and $\beta 1$ – $\beta 14$ (beta-strands—S). Asterisk indicates amino acids which distinguish between the eIF4AI and eIF4AIII homologues. Further relevant GenBank accession numbers: human (*Hs*) eIF4AI, AAX43035; human eIF4AIII (*Hs*eIF4A3), P38919; *S.pombe* (*Sp*) eIF4AI, CAA56772; *S.pombe* eIF4A-like protein (*Sp4*Alike), CAA92238; *A.thaliana* (*At*) eIF4AI, NP_177417; *A.thaliana* eIF4A-like protein (*At4*Alike), NP_188610.

trypanosomatids but also in eukaryotes in general, we modeled the structures of both *TbeIF4AI* and *III* based on the solved structure of either the yeast *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 *TbeIF4AI* and *TbeIF4AIII*. 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.

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

Position: TbEIFAI	Substitution: 4AI > 4AIII	Secondary structure	Domain	Overall position in predicted tertiary structure
93	K/N/V > R	Loop	N-terminal	Near $\alpha 5$ /exposed
139	Q/R > E/G	$\alpha 5$	N-terminal	Exposed
146	Q/A > E/D	$\alpha 5$	N-terminal	Exposed
153	V > S/A ^a	$\beta 4$	N-terminal	Buried
197	F > Y	$\alpha 8$	N-terminal	Partially exposed
213	M > L	Loop	N-terminal	Interface/next to Motif III/partially buried
220	L/V/I > M	$\alpha 9$	N-terminal	Buried
226	R/N > T	Loop	C-terminal	Exposed
256	L > F ^a	$\alpha 10$	C-terminal	Partially buried
264	E > D ^a	$\alpha 10$	C-terminal	Exposed
303	S/D > P ^a	Loop	C-terminal	Exposed
328	V/L > W ^a	Loop	C-terminal	Interface/Motif V/Buried
388	H/N > A/S ^a	Loop	C-terminal	Exposed

^aThese substitutions but not the others are present in the *S.cerevisiae* nucleolar protein Fal1p [may be related to the eIF4AIII proteins (45)].

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 helicase's 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 EJC 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 eIF4AIII to bind spliced mRNA and to rescue NMD in eIF4AIII depleted cells (30). In

summary, we have identified several individual amino acids conserved in either eIF4AI 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; LmEIF4AI binds strongly to at least two eIF4G homologues whereas LmEIF4AIII 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 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

homologues but so far without success. However, both Y14 and Aly/REF are small RNA-binding proteins with single RRM, 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 eIF4AIII-like proteins may be active within this complex.

In a very recent study eight eIF4AIII specific motifs (named A to H) were identified in an alignment comparing various eIF4AIII homologues with the human eIF4AI and II proteins. Selected amino acids in some of these motifs, as well as in the canonical eIF4A motifs I, Ia and VI, were then mutated in recombinant or *in vivo* overexpressed eIF4AIII to investigate their requirement for eIF4AIII function (30). In the alignment provided here, which includes both *T.brucei* eIF4A homologues, as well as eIF4AI sequences from divergent organisms, no continuous set of amino acids were found to be typical of either eIF4AI or eIF4AIII proteins. However, unique amino acid substitutions were identified which distinguish eIF4AIII-like proteins from eIF4AI homologues in all sequences investigated so far. Several of these substitutions not only coincide with some of the proposed eIF4AIII specific motifs (motifs C, E and H), but also are included in two of those motifs found to be involved in specific eIF4AIII 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 eIF4AI/eIF4AIII substitutions identified here do not coincide with the remaining proposed motifs. These might be involved in mediating other aspects of eIF4A function and should be considered as targets for further investigation.

Very few protein coding genes in trypanosomatids contains 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)

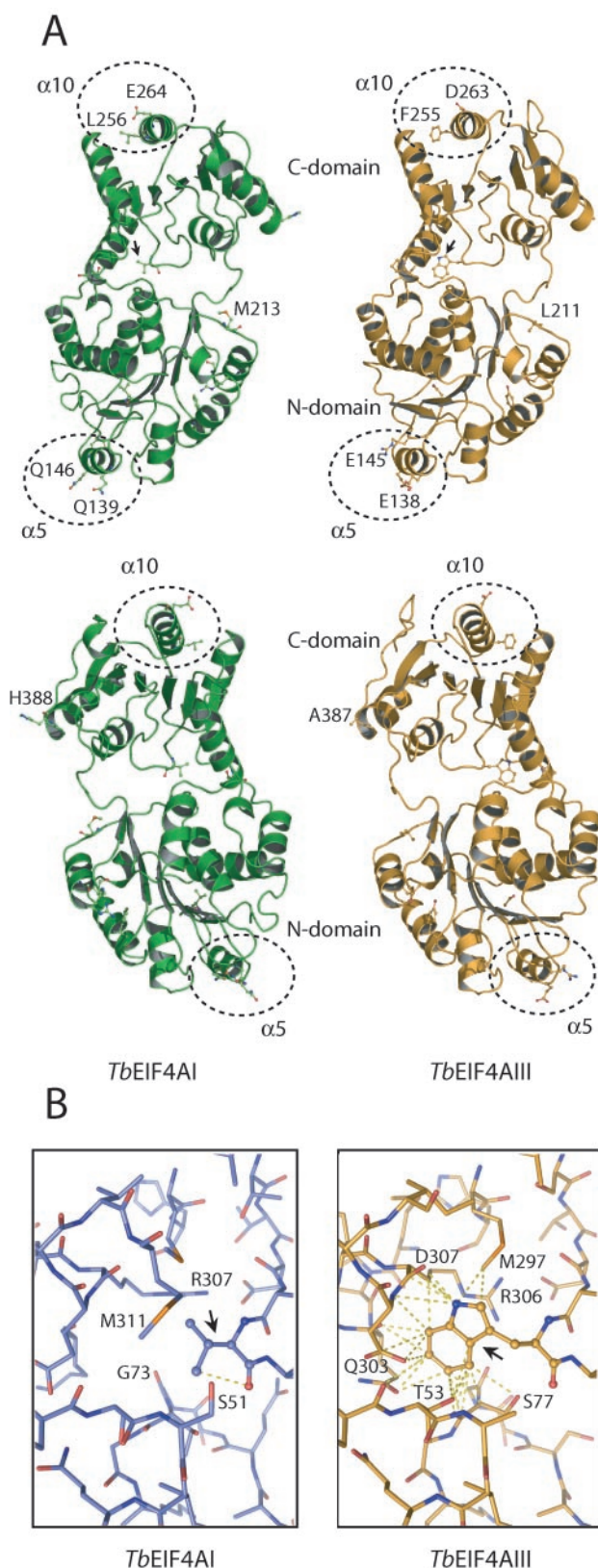


Figure 8. Molecular modelling of *TbEIF4AI* and *III* highlighting the position of amino acids unique to the eIF4AI or eIF4AIII homologues. Diagrams were created with the program PyMol (<http://www.pymol.org>). (A) Ribbon diagrams of the overall structure of both *TbEIF4AI* and *III* viewed as in (46) (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, $\alpha 5$ and $\alpha 10$. The H/N388A/S and M213L substitutions are also indicated (their numbering differ however from the eIF4AI/eIF4AIII sequences—for instance, H388 in *TbEIF4AI* is equivalent to A387 in *TbEIF4AIII* and so on). (B) Balls and sticks representation showing the neighbourhood of the L328W substitution in both *TbEIF4AI* 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 *TbEIF4AIII* 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 were supported by grants/studentships from CNPq and CAPES. Funding for the laboratory in Recife came in part from a grant from FACEPE/CNPq/CT INFRA.

Conflict of interest statement. None declared.

REFERENCES

- Campbell,D.A., Thomas,S. and Sturm,N.R. (2003) Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect.*, **5**, 1231–1240.
- Liang,X.H., Haritan,A., Uliel,S. and Michaeli,S. (2003) *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot. Cell*, **2**, 830–840.
- Perry,K.L., Watkins,K.P. and Agabian,N. (1987) Trypanosome mRNAs have unusual 'cap 4' structures acquired by addition of a spliced leader. *Proc. Natl Acad. Sci. USA*, **84**, 8190–8194.
- Hershey,J.W.B. and Merrick,W.C. (2000) Pathway and mechanism of initiation of protein synthesis. In Sonenberg,N., Hershey,J.W.B. and Mathews,M.B. (eds), *Translational Control Of Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 33–88.
- Pestova,T.V., Kolupaeva,V.G., Lomakin,I.B., Pilipenko,E.V., Shatsky,I.N., Agol,V.I. and Hellen,C.U. (2001) Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl Acad. Sci. USA*, **98**, 7029–7036.
- Preiss,T. and Hentze,M.W. (2003) Starting the protein synthesis machine: eukaryotic translation initiation. *Bioessays*, **25**, 1201–1211.
- Sonenberg,N. and Dever,T.E. (2003) Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.*, **13**, 56–63.
- Gingras,A.C., Raught,B. and Sonenberg,N. (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.*, **68**, 913–963.
- Gorbalenya,A.E. and Koonin,E.V. (1993) Helicases-amino-acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.*, **3**, 419–429.
- Linder,P. (2003) Yeast RNA helicases of the DEAD-box family involved in translation initiation. *Biol. Cell*, **95**, 157–167.
- Rocak,S. and Linder,P. (2004) DEAD-box proteins: the driving forces behind RNA metabolism. *Nature Rev. Mol. Cell Biol.*, **5**, 232–241.
- Caruthers,J.M., Johnson,E.R. and McKay,D.B. (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl Acad. Sci. USA*, **97**, 13080–13085.
- Story,R.M., Li,H. and Abelson,J.N. (2001) Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proc. Natl Acad. Sci. USA*, **98**, 1465–1470.
- Caruthers,J.M. and McKay,D.B. (2002) Helicase structure and mechanism. *Curr. Opin. Struct. Biol.*, **12**, 123–133.
- Marcotrigiano,J., Lomakin,I.B., Sonenberg,N., Pestova,T.V., Hellen,C.U. and Burley,S.K. (2001) A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell*, **7**, 193–203.
- Imataka,H. and Sonenberg,N. (1997) Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol. Cell Biol.*, **17**, 6940–6947.
- Yang,H.S., Cho,M.H., Zakowicz,H., Hegamyer,G., Sonenberg,N. and Colburn,N.H. (2004) A novel function of the MA-3 domains in transformation and translation suppressor Pdc4 is essential for its binding to eukaryotic translation initiation factor 4A. *Mol. Cell Biol.*, **24**, 3894–3906.
- Svitkin,Y.V., Pause,A., Haghighat,A., Pyronnet,S., Witherell,G., Belsham,G.J. and Sonenberg,N. (2001) The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA*, **7**, 382–394.
- Pestova,T.V. and Kolupaeva,V.G. (2002) The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev.*, **16**, 2906–2922.
- Nielsen,P.J. and Trachsel,H. (1988) The mouse protein synthesis initiation factor 4A gene family includes two related functional genes which are differentially expressed. *EMBO J.*, **7**, 2097–2105.
- Yoder-Hill,J., Pause,A., Sonenberg,N. and Merrick,W.C. (1993) The p46 subunit of eukaryotic initiation factor (eIF)-4F exchanges with eIF-4A. *J. Biol. Chem.*, **268**, 5566–5573.
- Li,Q., Imataka,H., Morino,S., Rogers,G.W., Jr, Richter-Cook,N.J., Merrick,W.C. and Sonenberg,N. (1999) Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell Biol.*, **19**, 7336–7346.
- Holzmann,K., Gerner,C., Poltl,A., Schafer,R., Obrist,P., Ensinger,C., Grimm,R. and Saueremann,G. (2000) A human common nuclear matrix protein homologous to eukaryotic translation initiation factor 4A. *Biochem. Biophys. Res. Commun.*, **267**, 339–344.
- Ferraiuolo,M.A., Lee,C.S., Ler,L.W., Hsu,J.L., Costa-Mattioli,M., Luo,M.J., Reed,R. and Sonenberg,N. (2004) A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proc. Natl Acad. Sci. USA*, **101**, 4118–4123.
- Chan,C.C., Dostie,J., Diem,M.D., Feng,W., Mann,M., Rappsilber,J. and Dreyfuss,G. (2004) eIF4A3 is a novel component of the exon junction complex. *RNA*, **10**, 200–209.
- Palacios,I.M., Gatfield,D., St.J.D. and Izaurralde,E. (2004) An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature*, **427**, 753–757.
- Shibuya,T., Tange,T.O., Sonenberg,N. and Moore,M.J. (2004) eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nature Struct. Mol. Biol.*, **11**, 346–351.
- Ballut,L., Marchadier,B., Baguet,A., Tomasello,C., Seraphin,B. and Le,H.H. (2005) The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nature Struct. Mol. Biol.*, **12**, 861–869.
- Tange,T.O., Shibuya,T., Jurica,M.S. and Moore,M.J. (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA*, **11**, 1869–1883.
- Shibuya,T., Tange,T.O., Stroupe,M.E. and Moore,M.J. (2006) Mutational analysis of human eIF4AIII identifies regions necessary for exon junction complex formation and nonsense-mediated mRNA decay. *RNA*, **12**, 360–374.
- Skeiky,Y.A., Guderian,J.A., Benson,D.R., Bacelar,O., Carvalho,E.M., Kubin,M., Badaro,R., Trinchieri,G. and Reed,S.G. (1995) A recombinant *Leishmania* antigen that stimulates human peripheral blood mononuclear cells to express a Th1-type cytokine profile and to produce interleukin 12. *J. Exp. Med.*, **181**, 1527–1537.

32. Skeiky, Y.A., Kennedy, M., Kaufman, D., Borges, M.M., Guderian, J.A., Scholler, J.K., Ovendale, P.J., Picha, K.S., Morrissey, P.J., Grabstein, K.H. *et al.* (1998) LeIF: a recombinant *Leishmania* protein that induces an IL-12-mediated Th1 cytokine profile. *J. Immunol.*, **161**, 6171–6179.
33. Dhaliya, R., Reis, C.R., Freire, E.R., Rocha, P.O., Katz, R., Muniz, J.R., Standart, N. and de Melo Neto, O.P. (2005) Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. *Mol. Biochem. Parasitol.*, **140**, 23–41.
34. Wirtz, E., Leal, S., Ochatt, C. and Cross, G.A. (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **99**, 89–101.
35. Wickstead, B., Ersfeld, K. and Gull, K. (2002) Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **125**, 211–216.
36. Pause, A., Methot, N., Svitkin, Y., Merrick, W.C. and Sonenberg, N. (1994) Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. *EMBO J.*, **13**, 1205–1215.
37. Hirumi, H. and Hirumi, K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J. Parasitol.*, **75**, 985–989.
38. Webb, H., Burns, R., Ellis, L., Kimblin, N. and Carrington, M. (2005) Developmentally regulated instability of the GPI-PLC mRNA is dependent on a short-lived protein factor. *Nucleic Acids Res.*, **33**, 1503–1512.
39. Carrington, M., Roditi, I. and Williams, R.O. (1987) The structure and transcription of an element interspersed between tandem arrays of mini-exon donor RNA genes in *Trypanosoma brucei*. *Nucleic Acids Res.*, **15**, 10179–10198.
40. Minshall, N., Thom, G. and Standart, N. (2001) A conserved role of a DEAD box helicase in mRNA masking. *RNA*, **7**, 1728–1742.
41. Cordin, O., Tanner, N.K., Doere, M., Linder, P. and Banroques, J. (2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J.*, **23**, 2478–2487.
42. Tanner, N.K., Cordin, O., Banroques, J., Doere, M. and Linder, P. (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell*, **11**, 127–138.
43. Roditi, I., Carrington, M. and Turner, M. (1987) Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. *Nature*, **325**, 272–274.
44. von der Haar, T. and McCarthy, J.E. (2002) Intracellular translation initiation factor levels in *Saccharomyces cerevisiae* and their role in cap-complex function. *Mol. Microbiol.*, **46**, 531–544.
45. Kressler, D., de la, C.J., Rojo, M. and Linder, P. (1997) Fallp is an essential DEAD-box protein involved in 40S-ribosomal-subunit biogenesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **17**, 7283–7294.
46. Cheng, Z., Collier, J., Parker, R. and Song, H. (2005) Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA*, **11**, 1258–1270.
47. Oberer, M., Marintchev, A. and Wagner, G. (2005) Structural basis for the enhancement of eIF4A helicase activity by eIF4G. *Genes Dev.*, **19**, 2212–2223.
48. De Gaudenzi, J., Frasch, A.C. and Clayton, C. (2005) RNA-binding domain proteins in kinetoplastids: a comparative analysis. *Eukaryot. Cell*, **4**, 2106–2114.
49. Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H.O., Bishop, J.R., Falcone, F.H., Gavrilescu, C., Montgomery, J.L., Santori, M.I. *et al.* (2000) A new twist in trypanosome RNA metabolism: *cis*-splicing of pre-mRNA. *RNA*, **6**, 163–169.
50. Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R. *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science*, **309**, 436–442.
51. Conti, E. and Izaurralde, E. (2005) Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.*, **17**, 316–325.
52. Lejeune, F. and Maquat, L.E. (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr. Opin. Cell Biol.*, **17**, 309–315.
53. Lejeune, F., Ishigaki, Y., Li, X. and Maquat, L.E. (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.*, **21**, 3536–3545.
54. Ishigaki, Y., Li, X., Serin, G. and Maquat, L.E. (2001) Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, **106**, 607–617.
55. Li, H. and Tschudi, C. (2005) Novel and essential subunits in the 300-kilodalton nuclear cap binding complex of *Trypanosoma brucei*. *Mol. Cell Biol.*, **25**, 2216–2226.

**TWO eIF4G HOMOLOGUES CONSERVED IN TRYPANOSOMATID SPECIES
DISPLAY FUNCTIONAL DIFFERENCES COMPATIBLE WITH ROLES IN TWO
DIVERGED eIF4F COMPLEXES**

Reis, C.R.S. *, Moura, D. M. N.*, Lima, R. P., Figueiredo, R. C. B. Q., and de Melo Neto, O. P.

*These authors contributed equally to this work.

Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Av. Moraes Rego s/n, Campus UFPE, Recife PE 50670-420.

Keywords: protein synthesis, eIF4G, eIF4F, Trypanosomatid, *Leishmania major*, *Trypanosoma brucei*

Corresponding author: Osvaldo P. de Melo Neto

e-mail: opmn@cpqam.fiocruz.br

ABSTRACT

The trypanosomatids are pathogenic protozoan which include species belonging to the genera *Trypanosoma* and *Leishmania* and are characterized by unique molecular mechanisms such as regulation of their gene expression, achieved mainly at the post-transcriptional level. Translation initiation is not well known but may include potentially novel processes not yet observed in other organisms. In higher eukaryotes, the scaffolding protein eIF4G, part of the eIF4F complex, performs many functions in translation initiation related to ribosome recruitment to the mRNA. Here we describe work aimed at the functional characterization of the trypanosomatid eIF4G homologues, five of which have been described (*LmEIF4G1-5* in *L. major*). Through pull-down assays, two of these (*LmEIF4G3-4*) were found to bind to *Leishmania* homologues of the remaining eIF4F subunits, eIF4A and eIF4E, and were selected for further studies. The two proteins differ in the binding to eIF4A and bind to different eIF4E homologues. The eIF4E binding sequences were located to their short N-terminal regions, but site directed mutagenesis of selected residues uncovered binding motifs which differ for each protein and which are also distinct from the consensus described from other eukaryotes. To study these proteins *in vivo*, we opted to use *T. brucei* and selected their orthologues (*TbEIF4G3-4*), plus *TbEIF4G5*, for further analysis. The three proteins were found to be cytoplasmic and all were essential for cellular viability after knock-down through RNAi. Depletion of *TbEIF4G3* led to a quick reduction in translation with subsequent cellular death whilst depletion of *TbEIF4G4-5* produced a more delayed growth retardation effect prior to cell death. Cell lacking *TbEIF4G4* also exhibited changes in morphology but no substantial inhibition in protein synthesis. In all our results are consistent with the existence of at least two distinct eIF4F complexes, with the one containing *TbEIF4G3* having a major role in translation.

INTRODUCTION

The trypanosomatids are pathogenic protozoan responsible for relevant human illnesses such the Leishmaniasis, Sleeping Sickness and Chagas' Disease (*Leishmania* and *Trypanosoma* species). These organisms are characterized by unique molecular mechanisms involved in gene expression and its control (Ivens et al., 2005). Lack of defined promoters for protein encoding genes and polycistronic transcription strongly implicates post-transcriptional mechanisms to be required for regulating gene expression (for reviews see (Clayton, 2002; Clayton and Shapira, 2007; Haile and Papadopoulou, 2007). A likely stage for such regulation and a possible target for chemotherapy designed to inhibit the growth of these pathogens is the initiation stage of protein synthesis. So far however not much is known about specific details of the translation initiation process in these organisms or how it differs from what is known in higher eukaryotes.

In plants, yeast and animals, translation initiation is a very complex process mediated by a number of translation initiation factors (eIFs) which help the ribosome bind to the mRNA and find the AUG start codon (Hershey and Merrick, 2000; Pestova et al., 2001; Preiss and Hentze, 2003). Paramount within these factors is eIF4F, a heterotrimeric complex formed by the subunits eIF4E, eIF4A and eIF4G (Gingras et al., 1999). In translation eIF4F binds to the monomethylated cap present on the 5' end of the mRNAs and mediates the recruitment of the 43S complex containing the small ribosomal subunit and other translation factors. eIF4E is the cap binding protein, which specifically recognizes the cap, whilst eIF4A is a RNA helicase which melts secondary structures at the 5' UTR of the messages, facilitating ribosome binding and scanning prior to finding the start codon. The eIF4G protein is a very large polypeptide (approximately 200 kDa), which provides a scaffold for the other eIF4F subunits, and has binding sites for other proteins such as the eIF3 initiation factor (which binds the 40S ribosomal subunit) (reviewed in Keiper et al., 1999; Prevot et al., 2003). The eIF4F activity is enhanced by the poly (A) binding protein (PABP), which acts through a direct interaction with eIF4G (Tarun Jr., et al., 1996; Imataka et al., 1998; Wells et al., 1998; reviewed in Gallie 1998; Cheng and Gallie 2007). In metazoans eIF4G is the subject of multiple regulation events required to control protein synthesis and is able to bind RNA on its own (See also Contreras et al., 2008).

The prototype eIF4G protein (from mammals) can be schematically divided into three regions: an N-terminal domain, containing binding sites for eIF4E and PABP; a central HEAT domain (also known as MIF4G), which binds eIF4A, eIF3 and can also bind to RNA (Marcotrigiano et al., 2001; Prevot et al., 2003); and the C-terminus which, in mammalian eIF4G, contains a second HEAT domain (named MA3), also involved in binding to eIF4A, and a W2 domain which mediates the binding to the eIF4E kinase, Mnk (Pyronnet et al., 1997; Hinton et al., 2005; Bellsollet et al., 2006). This arrangement seems to be conserved in different groups of eukaryotes (metazoans, plants and fungi) with the exception of the eIF4G from *Saccharomyces cerevisiae* which lacks a proper C-terminal domain. In mammals, yeast and plants two different eIF4G isoforms are generally observed with complementary roles in translation (Gallie and Browning 2001; Coldwell and Morley 2006). However several eIF4G related proteins have also been detected in different organisms which have some, but not all features of a true eIF4G, and which may participate as regulators of the protein synthesis apparatus or, else, performs other functions related to mRNA metabolism (Imataka et al., 1997; reviewed in Hernandez et al., 2005; Marintchev et al., 2005). The activity of eIF4G can be regulated through different means such as phosphorylation and expression of different isoforms and its association with binding partners, such as eIF4E and PABP, can also be controlled (Fraser et al., 1999).

In trypanosomatids, five putative eIF4G homologues were originally identified within the *L. major* genome database (*LmEIF4G1-5*), all sharing a central conserved eIF4A binding domain (HEAT-MIF4G) and with orthologues clearly defined in other *Leishmania* and *Trypanosoma* species (Dhalia et al., 2005). Apart from *LmEIF4G3* and 4, which seem to be related and share conserved features both in their N and C-terminus, the remaining proteins lack any significant similarities outside this domain and no putative binding sites for the other eIF4G partners can be identified. Likewise, homology to described eIF4G proteins from other organisms is also restricted to the eIF4A binding domain. Preliminary experimental evidence for a role in translation for one of these homologues was put forward with the demonstration that *LmEIF4G3* could bind to the *Leishmania* eIF4AI homologue (Dhalia et al., 2005). More recently this same protein has been shown to be able to bind to two different *L. major* eIF4E

homologues, an interaction which depends on its very short N-terminus, reinforcing its role as a possible major player in translation in trypanosomatids (Yoffe et al., 2009).

Several homologues to the eIF4F subunits and PABP from *Leishmania major* have been previously cloned and expressed as radioactive labelled proteins or recombinant His-tagged or GST-fusions. Here we have investigated possible interactions between several of these homologues through pull-down assays. So far, we have confirmed direct interactions between two eIF4G homologues (*LmEIF4G3-4*) and homologues of the remaining subunits of the eIF4F complex, as well as PABP. Site directed mutagenesis of selected residues in *LmEIF4G3-4* reduced drastically its interaction with eIF4E/eIF4A homologues as identified in others eukaryotes. These experiments confirm the existence of at least two eIF4F-like complexes in *L. major* and indicate both novel and conserved interactions and interaction-motifs when compared with other eukaryotes. In order to characterize these proteins *in vivo*, we opted for using *T. brucei* and selected *TbEIF4G3-5* for further analysis, including cellular localization and RNA interference. The subcellular localization of all three proteins confirmed their presence in the cytoplasm but their knock-down produced different phenotypes. Although all three are essential for viability, only the knock-down of *TbEIF4G3*, but not *TbEIF4G4-5*, impacts on protein synthesis.

MATERIALS AND METHODS

Cloning procedures

Cloning of the *Leishmania major* genes

The cloning of the *LmEIF4E1-3*, *LmEIF4AI* (*LmEIF4A1*) and full length-*LmEIF4G3* genes into the pET21D (Novagen) and pGEX4T3 (GE Healthcare) vectors has been described previously (Dhalia et al., 2005). Likewise, the cloning of *LmEIF4AIII* (previously called *LmEIF4A2* – see Dhalia et al., 2006) into the pRSETA (Invitrogen) and pGEX4T3 has also been reported. The cloning of the three genes encoding the *L. major* PABP homologues (*LmPABP1-3*) into both pET21D and pGEX4T3 vectors are reported in a separate manuscript (da Costa Lima et al., in prep). For the *LmEIF4E4* homologue, the cloning of its gene into the pET21D is also described elsewhere (Freire et al., manuscript in prep.). Here, for cloning into pGEX4T3,

the *LmEIF4E4* gene was reamplified using a new 5' primer (TCT GGA TCC ATG TCT ACC CCT CTC GAT GTG) and the original 3' primer, flanked by *Bam*H I/*Not* I and cloned into the *Bam*H I/*Not* I sites of the vector.

DNA fragments encoding full-length sequences and/or truncated proteins containing the HEAT-MIF4G domain on its own or plus the N or C-terminuses from *LmEIF4G1* (*LmEIF4G1*₁₋₄₇₅ and *LmEIF4G1*₁₂₈₋₁₀₁₆ – the number indicate the amino acids remaining from the wild-type protein), *LmEIF4G2* (*LmEIF4G2*₁₋₇₀₅), *LmEIF4G4* (*LmEIF4G4*₁₋₃₆₂, *LmEIF4G4*₃₇₋₃₆₂, *LmEIF4G4*₃₇₋₇₆₅, *LmEIF4G4*₁₋₇₆₅) and *LmEIF4G5* (*LmEIF4G5*₁₋₄₅₅, *LmEIF4G5*₁₃₈₋₇₈₂) were amplified by PCR from total *L. major* (MHOM/IL/81/Friedlin) genomic DNA flanked by sites for the restriction enzymes *Afl* III/ *Not* I (*LmEIF4G1*), *Nco* I/*Xho* I (*LmEIF4G2*₁₋₇₀₅), *Bam* HI/*Not* I (*LmEIF4G4*) or *Nco* I/*Not* I (*LmEIF4G5*). The *LmEIF4G1* fragments were first cloned into the *Nco* I/*Not* I sites of pET21d. To obtain the *LmEIF4G1*₁₋₁₀₁₆ full length sequence in pET21d, the *LmEIF4G1*₁₂₈₋₁₀₁₆ construct was digested with *Sal* I/*Not* I and the resulting fragment, 2123 bp, subcloned into the same sites of the pET21d-*LmEIF4G1*₁₋₄₇₅ plasmid. From the resulting construct the full-length *LmEIF4G1* sequence was reamplified flanked by sites for *Bam* HI/*Not* I and the resulting fragment cloned into the same sites of pGEX4T3. The *LmEIF4G2*₁₋₇₀₅ fragment was first cloned into the *Nco* I/*Xho* I sites of pET21d and the resulting plasmid used as template to reamplify the same fragment flanked by sites for e *Bam* HI/*Xho* I and subsequently cloned into the same sites of pGEX4T3. Fragments encoding the various *LmEIF4G4* variations were directly cloned into the *Bam* HI/*Not* I sites of both pET21a (Novagen – similar to pET21d) and pGEX4T3 plasmids. Finally, for the *LmEIF4G5* fragments, these were first cloned into the *Nco* I/*Not* I sites of pET21d and pGEX4T3_{NcoI} (derived from pGEX4T3 by the addition of a *Nco* I site immediately after the *Bam* HI site originally in the vector). The pGEX4T3-*LmEIF4G5*₁₋₄₅₅ construct was then digested with *Sal* I/*Not* I and the 1878 bp insert subcloned into the same sites of plasmid pGEX4T3-*LmEIF4G5*₁₃₈₋₇₈₂. The resulting plasmid contained the full length *LmEIF4G5* sequence (*LmEIF4G5*₁₋₇₈₂) in pGEX4T3 which was subsequently recovered by digestion with *Nco* I/*Not* I and also subcloned into the same sites of pET21d. All amplified fragments were sequenced and the resulting sequences confirmed by comparison with those derived from the *L. major* genome sequencing project. Table I lists all the primers used for the amplifications listed above.

Cloning of the *Trypanosoma brucei* genes

The amplification and cloning of the *TbEIF4G3-4* coding sequences, flanked by sites for *Hind* III and *Bam*H I (*TbEIF4G3*) or *Bam*H I only (*TbEIF4G4*), into the pGEM-T Easy vector (Promega) have been described elsewhere (Freire et al., manuscript in prep.) as well as the preparation of the isoform specific sera for both proteins. The *TbEIF4G5* coding sequences was amplified from *T. brucei* Lister 427 total genomic DNA flanked by sites for the enzymes *Hind* III and *Bam*H I (5' primer – AAG CTT CCG CCA CCA GGA GCA CCG CGC CGC AC e 3' primer – GGA TCC TTT AGA TTT CTC AAA CGC AG). The amplified fragment was then cloned into the vector pGEM-T Easy (Promega), sequenced and the resulting sequence confirmed by comparison with the one derived from the *T. brucei* genome sequencing project. To generate the EYFP fusions and for the RNAi experiments, the *Hind* III and *Bam*H I fragments from *TbEIF4G3* and 5 were recovered and subcloned respectively into the same sites of the transfection vectors p2216 (Kelly et al., 2007) or p2T7-177 (Wickstead et al., 2002). Likewise, for *TbEIF4G4*, it was recovered from the pGEM-T Easy derived plasmid through digestion with *Bam*H I and *Xho* I (which cuts internally at the end of the sequence) and subcloned into p2216 and p2T7-177. To produce the serum anti-*TbEIF4G5*, a protein fragment consisting of the protein's central HEAT-MIF4G domain plus a C-terminal His-tag was produced (*TbEIF4G5*₁₋₄₀₁). This was accomplished by amplifying the *TbEIF4G5*₁₋₄₀₁ gene fragment flanked by sites for *Bam*H I and *Xho* I and cloning into the same sites of pET21a (5' primer – TCT GGA TCC CCC GCC ACC ATG GAG CAC CGC GCC GCA C e 3' primer – CCC CTC GAG ATC ACC CGT CCC TGT CAC). The resulting plasmid was sequenced to confirm the identity of the inserts prior to transformation into *Escherichia coli* to produce the corresponding recombinant his-tagged protein. The serum was produced using the his-tagged recombinant *TbEIF4G5*₁₋₄₀₁ which was used to immunize rabbits resulting in the isoform specific polyclonal serum.

Site-directed mutagenesis

The mutagenesis of the *LmEIF4G3* and *LmEIF4G4* genes cloned into the pET21d and pET21a vectors, respectively, were carried out using the QuickChange II Site-Directed

Mutagenesis kit from Stratagene® following the manufacturer's recommendations. Table II lists all the primers used for the mutagenesis reactions. Following mutagenesis, and when necessary the various *LmEIF4G3-4* mutants were subcloned into the pGEX4T3 plasmid to produce the respective proteins as GST fusions.

Pull-down assays

For the pull-down assays glutathione-Sepharose (GE Biosciences) beads used were initially equilibrated with binding buffer B (100 mM KCl, 1 mM MgCl₂, 50 mM Hepes pH 7.2, 0.2% NP-40, 5% glycerol). Buffer B was supplemented with 5 mM of imidazole during all the steps of the His-tagged pull-downs. Labeled proteins were obtained through the linearization of the various plasmids described previously (with *Not* I for those cloned into the pET vectors), followed by transcription with T7 RNA polymerase in the presence of the cap analogue and translation in the rabbit reticulocyte lysate supplemented with ³⁵S-methionine.

For the GST-tagged protein experiments, the glutathione-Sepharose beads were initially saturated with 10 mg/ml of BSA, followed by two washes with Buffer B prior to their use in the assays. Approximately 10 µl of the beads was then incubated with 2 µg of the GST tagged proteins in a final volume of 200 µl (completed with Buffer B), for 1 hour rotating at 4°C. The beads were washed twice with Buffer B and incubated with 10 µl of translation lysate containing the ³⁵S-Met-labelled proteins, in Buffer B in a final volume of 200 µl, for 2 hours at 4°C. After three more washes, the bound proteins were eluted by addition of SDS sample buffer and the samples ran on 15% SDS-PAGE. The gels were stained with Coomassie Blue (to visualize the recombinant proteins), and subjected to autoradiography.

Parasite growth, transfection, RNA interference and metabolic labeling

Procyclic *T. brucei* cells (29-13) were transfected with the p2T7-177 (RNAi) and p2216 (EYFP fusions) derived plasmids containing the genes of interest. Transfected cells were selected after growth in the presence of phleomycin (2.5 µg/mL). RNA interference or expression of fluorescent fusion proteins was induced after tetracycline addition (1 µg/mL). At regular intervals, cellular growth was monitored by counting the number of viable cells. To evaluate the protein expression levels, western blots were performed using specific antisera.

The fusion fluorescent proteins were detected through confocal microscopy using a Leica SPII-AOBS microscope after 24h of induction with tetracycline.

To measure the rate of protein synthesis, [³⁵S]-methionine (10 µCi/ml) was added to mid-log cultures which were incubated for one hour prior to the determination of trichloroacetic acid precipitable incorporation into protein. For metabolic labeling, cultures were washed twice with methionine-free medium and then resuspended at 1×10^7 cells/ml in the methionine-free medium containing 50 µCi/ml [³⁵S]-methionine and incubated for one hour at 28°C prior to harvesting and analysis by SDS-PAGE and autoradiography.

Fluorescence microscopy

For the indirect immunofluorescence assay, wild type procyclic cells grown to mid-log phase were harvested and washed with SDM-79 minus serum. The cells were fixed at a density of 5×10^6 /ml with 3% paraformaldehyde, washed once in PBS and adhered to poly-L-lysine coated slides. Permeabilization was carried out with 0.1% Triton X-100 followed by blocking with 1% BSA. Antibody detection of *Tb*EIF4G3-5 followed standard procedures using goat anti-rabbit IgG Alexa Fluor 488 and DNA was stained using TOTO-3 (both from Invitrogen – Molecular Probes, USA). The cells were visualized through confocal microscopy.

RESULTS

Identification of interactions reminiscent of eIF4F function performed by the trypanosomatid eIF4G homologues

Previously, when the central HEAT-MIF4G domain from the first three *L. major* eIF4G homologues (*Lm*EIF4G1-3) were assayed for their ability to bind the parasite eIF4AI homologue (*Lm*EIF4A1 or *Lm*EIF4AI), through pull-down assays, only *Lm*EIF4G3 was found to be capable to bind *Lm*EIF4AI. Here, as a first approach to verify which of the five homologues perform any relevant roles in translation, new pull-downs were performed, first investigating their interaction with the two *L. major* eIF4A homologues. For this assays GST fusions of the eIF4A proteins were incubated with ³⁵S-labelled full length *Lm*EIF4G1, 3, 4 and 5 and a *Lm*EIF4G2 construct consisting of the proteins' N-terminus plus its HEAT-MIF4G

domain (*LmEIF4G2*₁₋₇₀₅). The two trypanosomatid eIF4A homologues (*LmEIF4AI* and *III* for the *L. major* proteins) have been shown respectively to be orthologues of mammalian eIF4AI, which is part of the eIF4F complex and participates in translation, and eIF4AIII, a nuclear protein which is part of the exon-junction complex with roles in mRNA splicing and decay (Dhalia et al., 2006). Samples containing glutathione sepharose beads only or bound to GST were used as negative controls on the assays. As shown in Figure 1A, *LmEIF4G1*, 3 and 4 were found to be able to bind preferentially to the GST-*LmEIF4AI* protein and, to a minor extent, GST-*LmEIF4AIII*. Binding by the labelled *LmEIF4G3* was substantially higher than the other two proteins.

Next, we used a similar assay to investigate the binding of the five eIF4G proteins to the four eIF4E homologues previously described from *L. major* (Dhalia et al., 2005, Yoffe et al., 2006). In a recent report, different approaches have been used to investigate the interaction between the *L. major* EIF4G3 protein and the four eIF4E homologues and specific interactions were observed with the EIF4E1 and 4 (Yoffe et al., 2009). Here, when the four GST-tagged eIF4Es were incubated with the same ³⁵S-labelled eIF4G proteins tested in Figure 1A, significant interactions were observed only with *LmEIF4G3* and 4 (Figure 1B). *LmEIF4G3* was found to interact with the four eIF4E homologues, albeit weakly, whilst *LmEIF4G4* bound strongly to *LmEIF4E3*. Overall our results are consistent with both *LmEIF4G3* and 4 being able to perform interactions reminiscent of the eIF4F complex and implying relevant roles in translation.

Sequence analysis of the trypanosomatid eIF4G3-4 homologues

As well as their binding properties to eIF4A and eIF4E homologues, *LmEIF4G3-4* also share similarities at the sequence level absent from other trypanosomatid homologues. These include a very short N-terminus and a C-terminal domain which contain features conserved between the two proteins. In order to investigate their similarities in more detail we opted to perform an alignment comparing the full length sequences of EIF4G3 and 4 orthologues from three different trypanosomatids *L. major*, *T. brucei* and *T. cruzi* (Figure 2). The alignment highlights several conserved segments within the proteins' N- and C-terminal domains as well as within the central HEAT-MIF4G domain. Noteworthy are sequences within the proteins' N-

terminus which resembles the consensus eIF4E binding peptide present in eIF4G homologues from other eukaryotes (Mader et al., 1995). This sequence is defined as YXXXXLΦ (where X can be any amino acid and Φ is usually a hydrophobic amino acid such as L, M or F) and two of the conserved segments shown in the alignment resemble this segment (boxed in the figure). These are the **F/Y/MXXXXI/LR/L** (amino acids 3-9 in *LmEIF4G3* – FTVEQIR) and **FSLXXV/IV/L** (amino acids 23-29 in *LmEIF4G3* – FSLDEVV), with the second sequence being more conserved and in agreement with the described consensus peptide. A third possible consensus is found only on the very N-terminus of the EIF4G4 homologues, **FN/KL/PRGI/VV/T** (amino acids 3-9 in *LmEIF4G4* – FNLRGIV), however assuming that both EIF4G3-4 homologues would interact with their eIF4E partners using similar sequences, we consider it unlikely to be a possible eIF4E binding motif. Another point worth mentioning is the possibility that, due to the very short N-terminus of the trypanosomatid EIF4G3-4 homologues, it may not harbour a PABP binding region, as has been shown for eIF4G homologues from yeast and human (Tarun Jr., et al., 1997; Imataka et al., 1998).

The alignment also highlights several sequences conserved between the EIF4G3-4 homologues positioned within the proteins' C-terminus. Secondary structure prediction analyses suggest that these are mainly α -helical in nature. In mammalian eIF4G, the proteins' C-terminus contain two α -helical HEAT domains also involved in protein-protein interactions, MA3 and W2 (Bellsollell et al., 2006). No significant homologies, however, were observed between the C-terminus of human eIF4G and those of the trypanosomatid EIF4G3-4 when they were compared through BLAST analysis. Nevertheless, the presence of two conserved tryptophan residues in the C-terminus of the latter proteins led us to investigate further any similarities with the W2 domain of mammalian eIF4G, considering that it is also characterized by the presence of two W residues. Indeed, these residues, plus neighbouring amino acids, resemble the two conserved aromatic and acidic boxes previously observed in the C-terminus of mammalian eIF4G and related proteins (Bellsollell et al., 2006; Liberman N et al., 2008). Further, when we attempted to find conserved residues present in the MA3 and W2 domains deposited at the pfam database within the equivalent regions of EIF4G3-4 sequences, several coincided with conserved elements within these proteins (highlighted in Figure 2). Overall it seems likely then

that both MA3 and W2 domains, divergent in sequence due to the greater evolutionary distance between trypanosomatids and higher eukaryotes, are present in the EIF4G3 and 4 homologues.

Defining conserved interactions between eIF4G and eIF4A-eIF4E homologues

Considering the similarities identified between the trypanosomatid EIF4G3-4 sequences and the prototype mammalian eIF4G, we decided to investigate in more detail the eIF4G/eIF4A and eIF4G/eIF4E interactions using the *L. major* proteins in order to define how much these are conserved in such divergent organisms and to what extent they differ between the EIF4G3 and 4 homologues. By continuing to use the pull-down assays with the GST-tagged eIF4A homologues and the ³⁵S-labelled eIF4G proteins, we first compared the binding to eIF4A of either the full length proteins or truncations containing their HEAT-MIF4G domain alone (Figure 3A). Both the HEAT-MIF4G domain of *LmEIF4G3* as well as the full length protein bound very strongly to *LmEIF4AI* whilst only the full length proteins seemed to bind efficiently to *LmEIF4AIII*, although not as strongly to the *LmEIF4AI* homologue. In contrast, for *LmEIF4G4*, no binding by the HEAT-MIF4G domain of the protein could be detected for either of the eIF4A homologues and, as shown above, binding by the full-length protein seemed much reduced as compared to the *LmEIF4G3* protein. We then investigated the *LmEIF4G4/LmEIF4E3* interaction, with the purpose of confirming the localisation of the eIF4E binding site within the very short N-terminus of the eIF4G homologue (at this stage we opted not to test the *LmEIF4G3*/EIF4E interactions due to their lower binding in this assay). As shown in Figure 3B, the N-terminus of *LmEIF4G4* indeed harbours its binding site for *LmEIF4E3* since its removal specifically abolishes binding to the latter protein without interfering with the binding to *LmEIF4AI*. That the removal of the proteins' C-terminus does not prevent binding to the eIF4E homologue, although interfering with the eIF4A interaction, is an indication that this construct is folded correctly and that the proteins' requirement for the C-terminus to bind *LmEIF4AI* is real.

Next, we focused on specific motifs within the proteins' N-terminus which might be mediating the *LmEIF4G4/LmEIF4E3* interaction. For this we selected the two sets of amino acid triplets strictly conserved within the N-terminus of the EIF4G3-4 proteins (highlighted in Figure 2): FSL, part of a conserved segment which resembles the described eIF4E binding

sequence (amino acids 40-46, FSLERVL in *LmEIF4G4*), as mentioned above; and KLV, also conserved within the various sequences. Through site-directed mutagenesis, these were alternatively replaced from *LmEIF4G4* with three consecutive Alanine residues. We also replaced the LNK triplet, at the start of the HEAT-MIF4G domain and implicated in the binding to eIF4A in other eIF4G homologues, for three Alanine residues in yet another independent mutagenesis event. Surprisingly, all three of the resulting proteins bound efficiently to *LmEIF4E3* (data not shown). We then opted to mutate a second putative motif, partially conserved in the EIF4G3-4 sequence, which partially resembles the eIF4E binding peptide (amino acids 20-26, MTLADIL, in *LmEIF4G4*). To do this, the first amino acid of the motif (M20) was alternatively mutated to an alanine, a tyrosine or an aspartate in the wild type protein prior to its use in the pull-down assays. No change in binding of the labelled protein to *LmEIF4E3* was observed (Figure 3C and data not shown). Independently, we also replaced the IL doublet within this motif (I25, L26) with two alanines and tested once again the recombinant protein in the assay and this time binding to *LmEIF4E3* was abolished without affecting the binding to *LmEIF4A1*. So far these results are consistent with a unique eIF4E binding site in *LmEIF4G4* when compared to other known eIF4G homologues.

Confirming the interaction between the *Leishmania* eIF4G and eIF4E homologues

The previous results regarding the possible binding motif for *LmEIF4E3* in *LmEIF4G4* are striking, considering what is known regarding this interaction in other organisms. To confirm them, and also to study the *LmEIF4G3* interaction with the eIF4E homologues, we opted to establish a second pull-down assay, this time using GST-tagged *LmEIF4G3-4* and the ³⁵S-labelled *LmEIF4E1-4*. As shown in Figure 4A, this “reverse” assay confirms the interaction between *LmEIF4G4/LmEIF4E3* as well as the lack of interaction between *LmEIF4G4* and any of the other three eIF4E homologues. More important, this assay also highlights a very specific interaction between *LmEIF4G3* and *LmEIF4E4* with no binding of *LmEIF4G3* to the remaining eIF4E homologues. Next, we used the same assay to investigate the binding of the several *LmEIF4G4* deletions and mutants tested previously in Figure 3 to labelled *LmEIF4E3*. The results shown in Figure 4B confirm that the eIF4E binding site in *LmEIF4G4* is restricted to its N-terminus and that the partially conserved IL doublet is important for the interaction whilst

neither the FSL motif nor M20 seems to be required. We then performed similar deletions for *LmEIF4G3* as well as obtained mutants where alternatively the FSL or LNK amino acid triplets were replaced for three Alanines. The resulting GST-tagged proteins were incubated with ³⁵S-labelled *LmEIF4E4* and the results obtained (Figure 4C) confirmed also for *LmEIF4G3* the localization of its eIF4E binding motif to its very short N-terminus. However, in contrast to what was seen for *LmEIF4G4*, mutation of the FSL motif abolished binding to *LmEIF4E4*. The pull-down results then imply significant differences in the way the two protein pairs, *LmEIF4G4/LmEIF4E3* and *LmEIF4G3/LmEIF4E4*, interact. They also indicate that the binding between eIF4G and eIF4E homologues might be more variable than what has been shown so far for other eukaryotes.

Subcellular localisation of *TbEIF4G3-5*

As shown in the alignment in Figure 1, most major features which distinguish the EIF4G3 and 4 factors are conserved among their orthologues in the various trypanosomatid species. Taking into account this conservation, which indicates similar functions within the trypanosomatid family, and to continue the characterization of these two proteins functionally, we decided to change our model organism in order to fully exploit the tools available for genetic manipulation of *T. brucei*. For these experiments we decided to include for comparison a third unrelated eIF4G homologue, EIF4G5. Considering the lack of affinity of its *Leishmania* orthologue for homologues to the other eIF4F subunits, as well as its lack of homology outside of the HEAT-MIF4G domain to EIF4G3 and 4, we expected it to behave different from the two proteins which are the major focus of this work. First we started by investigating their subcellular localization through the cloning of the genes encoding their trypanosome counterparts (*TbEIF4G3*, 4 and 5) and expression as enhanced yellow fluorescent fusion proteins (EYFP) in transgenic cells. Expression of the fusion proteins was induced with tetracycline and their subcellular localization visualized through confocal microscopy (Figure 5A). As can be observed, all three proteins localized predominantly to the cytoplasm with very little localization in the nucleus. Rabbit polyclonal antisera were then raised against the three *T. brucei* proteins, after expression as his-tagged fusions in *Escherichia coli*, and used to validate the localization experiments. First, through western-blotting with extracts from transfected cells

after tetracycline addition, we confirmed the expression of the EYFP fusion proteins in levels roughly similar to the endogenous protein. We then performed immunocytochemistry experiments with wild-type *T. brucei* cell using affinity purified antibodies and confirmed the strict cytoplasmic localization of two of the three proteins (*LmEIF4G4-5* – Figure 5B). Within the cytoplasm these proteins did not show any specific restriction in localization although some perinuclear concentration may be observed. In all, the cytoplasmic localization of the three proteins rules out any role in mRNA transcription or processing in the nucleus and is consistent with major or minor roles in mRNA metabolism and translation in the cytoplasm.

RNA interference of *TbEIF4G3-5* in *T. brucei* procyclic cells

Considering the similarities observed between the EIF4G3-4 homologues and assuming a possible overlap in function, we went on to perform RNA interference analysis for the *T. brucei* proteins in order to evaluate cellular viability after their knock-down. The genes encoding *TbEIF4G3-5* were cloned into the p2T7-177 vector and the resulting plasmid transfected into *T. brucei* cells and double stranded RNA synthesis induced after tetracycline addition. The results obtained for the subsequent growth curves are summarized in Figure 6A. All three proteins were essential for viability but the phenotype observed was unique to each one of them. For *TbEIF4G3*, cells expressing the dsRNAs stopped growing very soon after induction and began to die before 24hs of RNAi. Cellular death preceded total depletion of *TbEIF4G3*, evaluated through western-blotting of total cellular extracts from the various time-points (Figure 6B), with substantial amounts of the protein still remaining at 24hs. In contrast, cells expressing *TbEIF4G4*-dsRNAs just showed growth retardation after 72h of induction, when the protein levels were totally depleted, with cell death occurring after the 5th day. As for the *TbEIF4G5* RNAi, it displayed an even more delayed phenotype with the reduction in cellular growth taking twice as long as observed for *TbEIF4G4*, despite the rapid and efficient knock-down of the protein. A relevant observation regarding the RNAi of the three EIF4G homologues has to do with the morphology of the cells prior to their death. For both *TbEIF4G3* and 5, and despite the very large differences in the RNAi response for these two proteins, no major changes in cell morphology was observed prior to their death. In contrast, for the

TbEIF4G4 RNAi, the cells assumed a rounded morphology which persisted for some time before they died (Figure 7A).

To evaluate the effect of the RNAi of the three proteins on protein synthesis, we performed metabolic labeling with ^{35}S -methionine of the cells at selected time points after RNAi. Samples were investigated for changes in total protein synthesis as well as in overall synthesis profile. No significant differences were observed in protein synthesis for the RNAi of *TbEIF4G4* or 5 even up to 72 hours after tetracycline addition. However, the knock down of *TbEIF4G3* inhibited protein synthesis as soon as 6 hours after RNAi induction with more than 50% inhibition after 24 hours (Figure 7B and data not shown), although no selective changes in profile were observed. So far these results are consistent with a major role for *TbEIF4G3* in protein synthesis in trypanosomatids. *TbEIF4G4-5*, despite being essential for viability, may not be directly required for protein synthesis or may perform minor roles in this process.

DISCUSSION

The results presented here builds upon previous published work by us and others (Yoffe et al., 2004; Dhalia et al., 2005; Yoffe et al., 2006; Yoffe et al., 2009 and manuscripts in prep.) which are beginning to unravel the complex events involved in regulating translation initiation in trypanosomatids and which may have implications into how they regulate their gene expression. So far out of the five eIF4G homologues, two (EIF4G3-4) seem to be clearly involved in events associated with protein synthesis although the EIF4G3 protein may have a more dominant role. As to the remaining homologues, it is striking that despite having many of the conserved residues implicated in eIF4A binding (Dhalia et al., 2005), no evidence for an interaction was observed for at least one of the eIF4G homologues, *LmEIF4G5*, and the same may apply to *LmEIF4G2*. Even the binding by *LmEIF4G1* is very weak and may or may not reflect a true functional role and will need complementary data perhaps showing a real interaction *in vivo*. Nevertheless, the lack of binding by *LmEIF4G1*, 2 and 5 to any of the eIF4E homologues clearly rule out any function in translation within an eIF4F like complex.

The results presented clearly show the very specific interaction between the *LmEIF4G4/LmEIF4E3* pair as well as highlighting the unique differences in the eIF4E binding

motif within the N-terminus of the eIF4G homologue, distinct from anything so far seen in related proteins from other organisms. The *LmEIF4G3* binding results confirm data recently published showing the *LmEIF4G3/LmEIF4E4* interaction, through two hybrid assays mainly, but also through pull downs (Yoffe et al., 2009), however our results are in disagreement in that we haven't seen any specific interaction between *LmEIF4G3/LmEIF4E1*. The minor binding seen in the first assay, with the GST-tagged eIF4E homologues and ³⁵S labelled *LmEIF4G3* is not selective for any of the eIF4E homologues or even enhanced for *LmEIF4E4* or *LmEIF4E1*. In the future, we hope to investigate further the molecular basis for the specificity in binding between the two protein pairs, *LmEIF4G3/LmEIF4G4* and *LmEIF4G4/LmEIF4E3*, and identify the basis for selectivity in binding.

As to the RNAi results, they clearly implicate *TbEIF4G3*, but no *TbEIF4G4* or 5 as being required for overall protein synthesis in *T. brucei* at least. Indeed, it seems that endogenous levels of *TbEIF4G3* are strictly regulated and any depletion can stop translation and prevent cellular growth. The observation that knock-down of *TbEIF4G4* induces a change in morphology of the cells prior to their death may indicate a selective role in the translation of specific mRNAs, whose products, however, are not visible through unidimensional SDS-PAGE after metabolically labeling. As for *TbEIF4G5*, the delayed phenotype showed after its RNAi may be due to a secondary or modulatory role of this protein in *T. brucei* metabolism although its cytoplasmic localization is a relevant point to be considered. In other organisms different eIF4G isoforms may have specific roles in translation of different mRNAs at the single cell stage or within different cells of a pluricellular organism (Castelo et al., 2006; Lee & McCormick 2006; Baker and Fuller 2007) and this indeed maybe what is being observed for the *TbEIF4G4/TbEIF4E3* pair.

In summary our results constitute a major breakthrough in the solving of the function of at least two of the trypanosomatid eIF4G homologues, at the same time identifying novel features not yet observed in other eukaryotes. However much has yet to be done in order to fully understand what their roles are and also what are the functions of the other trypanosomatid eIF4E and eIF4G homologues and how they may impact on the regulation of gene expression in these organisms.

ACKNOWLEDGEMENTS

This work was partially funded with grants provided from the Brazilian funding agencies FACEPE and CNPq. CRSR, DMNM and RPL and received studentships from CAPES and FACEPE. Technical assistance to grow the *Leishmania* and *Trypanosoma* cells was provided by B. S. Lima. We would like to thank C. Docena for help with the confocal microscopy and sequencing.

REFERENCES

- Baker,C.C. and Fuller,M.T. (2007). Translational control of meiotic cell cycle progression and spermatid differentiation in male germ cells by a novel eIF4G homolog. *Development* *134*, 2863-2869.
- Bellsolell,L., Cho-Park,P.F., Poulin,F., Sonenberg,N., and Burley,S.K. (2006). Two structurally atypical HEAT domains in the C-terminal portion of human eIF4G support binding to eIF4A and Mnk1. *Structure*. *14*, 913-923.
- Castello,A., Alvarez,E., and Carrasco,L. (2006). Differential cleavage of eIF4GI and eIF4GII in mammalian cells. Effects on translation. *J. Biol. Chem.* *281*, 33206-33216.
- Cheng,S. and Gallie,D.R. (2007). eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains. *J. Biol. Chem.* *282*, 25247-25258.
- Clayton,C. and Shapira,M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol. Biochem. Parasitol.*
- Clayton,C.E. (2002). Life without transcriptional control? From fly to man and back again. *EMBO J.* *21*, 1881-1888.

- Coldwell,M.J. and Morley,S.J. (2006). Specific isoforms of translation initiation factor 4GI show differences in translational activity. *Mol. Cell Biol.* 26, 8448-8460.
- Contreras,V., Richardson,M.A., Hao,E., and Keiper,B.D. (2008). Depletion of the cap-associated isoform of translation factor eIF4G induces germline apoptosis in *C. elegans*. *Cell Death. Differ.* 15, 1232-1242.
- Dhalia,R., Reis,C.R., Freire,E.R., Rocha,P.O., Katz,R., Muniz,J.R., Standart,N., and de Melo Neto,O.P. (2005). Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. *Mol. Biochem. Parasitol.* 140, 23-41.
- Dhalia,R., Marinsek,N., Reis,C.R., Katz,R., Muniz,J.R., Standart,N., Carrington,M., and de Melo Neto,O.P. (2006). The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. *Nucleic Acids Res.* 34, 2495-2507.
- Fraser,C.S., Pain,V.M., and Morley,S.J. (1999). The association of initiation factor 4F with poly(A)-binding protein is enhanced in serum-stimulated *Xenopus* kidney cells. *J. Biol. Chem.* 274, 196-204.
- Gallie,D.R. (1998). A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene* 216, 1-11.
- Gallie,D.R. and Browning,K.S. (2001). eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. *J. Biol. Chem.* 276, 36951-36960.
- Gingras,A.C., Raught,B., and Sonenberg,N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* 68, 913-963.
- Haile,S. and Papadopoulou,B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.* 10, 569-577.
- Hernandez,G. and Vazquez-Pianzola,P. (2005). Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. *Mech. Dev.* 122, 865-876.

Hershey, J.W.B. and Merrick, W.C. (2000). Pathway and mechanism of initiation of protein synthesis. In *Translational control of gene expression.*, N.Sonenberg, J.W.B.Hershey, and M.B.Mathews, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 33-88.

Hinton, T.M., Coldwell, M.J., Carpenter, G.A., Morley, S.J., and Pain, V.M. (2007). Functional analysis of individual binding activities of the scaffold protein eIF4G. *J. Biol. Chem.* 282, 1695-1708.

Imataka, H., Olsen, H.S., and Sonenberg, N. (1997). A new translational regulator with homology to eukaryotic translation initiation factor 4G. *EMBO J.* 16, 817-825.

Imataka, H., Gradi, A., and Sonenberg, N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* 17, 7480-7489.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchetti, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De, G.J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., and Myler, P.J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442.

- Keiper,B.D., Gan,W., and Rhoads,R.E. (1999). Protein synthesis initiation factor 4G. *Int. J. Biochem. Cell Biol.* *31*, 37-41.
- Kelly,S., Reed,J., Kramer,S., Ellis,L., Webb,H., Sunter,J., Salje,J., Marinsek,N., Gull,K., Wickstead,B., and Carrington,M. (2007). Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol. Biochem. Parasitol.* *154*, 103-109.
- Lee,S.H. and McCormick,F. (2006). p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. *EMBO J.* *25*, 4008-4019.
- Lieberman,N., Dym,O., Unger,T., Albeck,S., Peleg,Y., Jacobovitch,Y., Branzburg,A., Eisenstein,M., Marash,L., and Kimchi,A. (2008). The crystal structure of the C-terminal DAP5/p97 domain sheds light on the molecular basis for its processing by caspase cleavage. *J. Mol. Biol.* *383*, 539-548.
- Mader,S., Lee,H., Pause,A., and Sonenberg,N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol. Cell Biol.* *15*, 4990-4997.
- Marcotrigiano,J., Lomakin,I.B., Sonenberg,N., Pestova,T.V., Hellen,C.U., and Burley,S.K. (2001). A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell* *7*, 193-203.
- Marintchev,A. and Wagner,G. (2005). eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G. *Biochemistry* *44*, 12265-12272.
- Pestova,T.V., Kolupaeva,V.G., Lomakin,I.B., Pilipenko,E.V., Shatsky,I.N., Agol,V.I., and Hellen,C.U. (2001). Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. U. S A* *98*, 7029-7036.

- Preiss,T. and Hentze,M.W. (2003). Starting the protein synthesis machine: eukaryotic translation initiation. *Bioessays* 25, 1201-1211.
- Prevot,D., Darlix,J.L., and Ohlmann,T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. *Biol. Cell* 95, 141-156.
- Prevot,D., Decimo,D., Herbreteau,C.H., Roux,F., Garin,J., Darlix,J.L., and Ohlmann,T. (2003). Characterization of a novel RNA-binding region of eIF4GI critical for ribosomal scanning. *EMBO J.* 22, 1909-1921.
- Pyronnet,S., Imataka,H., Gingras,A.C., Fukunaga,R., Hunter,T., and Sonenberg,N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J.* 18, 270-279.
- Tarun,S.Z., Jr. and Sachs,A.B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* 15, 7168-7177.
- Wells,S.E., Hillner,P.E., Vale,R.D., and Sachs,A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* 2, 135-140.
- Wickstead,B., Ersfeld,K., and Gull,K. (2002). Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 125, 211-216.
- Yoffe,Y., Zuberek,J., Lewdorowicz,M., Zeira,Z., Keasar,C., Orr-Dahan,I., Jankowska-Anyszka,M., Stepinski,J., Darzynkiewicz,E., and Shapira,M. (2004). Cap-binding activity of an eIF4E homolog from *Leishmania*. *RNA.* 10, 1764-1775.
- Yoffe,Y., Zuberek,J., Lerer,A., Lewdorowicz,M., Stepinski,J., Altmann,M., Darzynkiewicz,E., and Shapira,M. (2006). eIF4E isoforms of *Leishmania* - binding specificities and potential roles. *Eukaryot. Cell.*

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, 3243-3253.

FIGURE LEGENDS

Figure 1: Analysis of the interaction between the *L. major* LmEIF4G1-5 proteins and homologues to the remaining subunits of the eIF4F complex. Pull-down assays were carried out using GST-tagged versions of *L. major* eIF4A and eIF4E homologues incubated with the ³⁵S-labelled *LmEIF4G1-5*. As controls, in each assay, GST on its own or only the glutathione sepharose beads were also added to test their ability to bind non-specifically to any of the labelled protein. **(A) Upper panel:** Coomassie Blue stained gel showing total translation extract (**Input**) as well as recombinant GST or GST-*LmEIF4AI* or III. **Lower panels:** autoradiography showing the specific binding between GST-*LmEIF4AI* and III to labeled *LmEIF4G1-4* (arrows). **(B) Upper panel:** Coomassie Blue stained gel showing total translation extract (**Input**) as well as recombinant GST or GST-*LmEIF4E1-4*. **Lower panels:** autoradiography showing the binding between GST-*LmEIF4E1-4* to labeled *LmEIF4G3* and between GST-*LmEIF4E3-4* to labeled *LmEIF4G4* (arrows).

Figure 2: Sequence comparison of the EIF4G3-4 homologues from different trypanosomatids showing possible regions of homology outside the HEAT domain. Clustal W alignment of the different homologues from *L. major*, *T. brucei* and *T. cruzi*. Amino acids identical in more than 60% of the sequences are highlighted in dark gray, while amino acids defined as similar, based on the BLOSUM 62 Matrix, on more than 60% of the sequences, are shown in pale gray. When necessary, spaces were inserted within the various sequences (dashes) to allow better alignment. The two possible eIF4E binding segments highlighted in the text are boxed. * denotes the central MIF4G HEAT domain, the region which defines the various eIF4G homologues, as well as the putative α -helical rich regions which may constitute variants MA3 and W2 domains. ↓ indicates the conserved amino acid segments which have been mutated. ↓↓

indicates conserved amino acid residues conserved within the proteins' N and C-terminal domains. The ▲ symbol highlights amino acid conserved in most of the trypanosomatid EIF4G3 and/or EIF4G4 sequences and which are similar to amino acids found in the same position in the human eIF4G1 MA3 and W2 domains and in related proteins.

Figure 3: Binding assays to define the domains and/or binding motifs of *LmEIF4AI* and *LmEIF4E3*. Pull-down assays were carried out using GST-tagged *LmEIF4AI* and III and *LmEIF4E3* and the ³⁵S-labelled full-length or truncated *LmEIF4G3-4*. As in Figure 1, assays always included GST and, in occasions, the glutathione sepharose beads. **(A) Upper panel:** autoradiography of the assay comparing the binding between full-length and truncated labeled *LmEIF4G3* with GST-*LmEIF4AI* and III. **Lower panel:** same as above but comparing the binding between full-length and truncated labeled *LmEIF4G4* with GST-*LmEIF4AI* and III. **(B)** Autoradiography comparing the binding between full-length and truncated versions of labeled *LmEIF4G4* with GST-*LmEIF4AI* and GST-*LmEIF4E3*. **(C)** Autoradiography showing the effect of mutations of selected residues on the N-terminus of labeled *LmEIF4G4* on its binding to GST-*LmEIF4AI* and GST-*LmEIF4E3*. **Upper panel:** Single mutation of the M20 residue (replaced by an Alanine). **(Lower panel).** Double mutations of the IL doublet to two Alanines.

Figure 4: Fine mapping of the *LmEIF4E3-4* binding sites within the N-terminuses of *LmEIF4G3-4*. Reverse pull-down assays carried out using GST-tagged *LmEIF4G3-4* incubated with selected ³⁵S-labelled *LmEIF4E* sequences. The GST-fusions included truncated versions of *LmEIF4G3-4* or full-length mutants with specific residues mutated after site directed mutagenesis. Again, GST on its own and the glutathione sepharose beads were also added as controls. **(A) Upper panel:** Coomassie Blue stained gel showing total translation extract **(Input)** as well as recombinant GST or GST-*LmEIF4G3-4*. **Lower panels:** autoradiography showing the specific binding between GST-*LmEIF4G3* and labeled *LmEIF4E4* and between GST-*LmEIF4G4* and *LmEIF4E3*. **(B)** Fine mapping of the interaction between GST-*LmEIF4G4* and *LmEIF4E3* **Upper panel:** Coomassie Blue stained gel showing the full-length mutants or truncated versions of *LmEIF4G4* used in the assay. **Lower panels:** autoradiography showing lack of binding between the truncated GST-*LmEIF4EG4* lacking its N-terminus, as well as the

full-length mutant with I25L26 replaced by Alanines, to labeled *LmEIF4E3*. (C) Fine mapping of the interaction between GST-*LmEIF4G3* and *LmEIF4E4* **Upper panel:** Coomassie Blue stained gel showing the full-length mutants or truncated versions of *LmEIF4G3* used in the assay. **Lower panels:** autoradiography showing lack of binding between the truncated GST-*LmEIF4EG3* lacking its N-terminus, as well as the full-length mutant with the FSL triplet replaced by Alanines, to labeled *LmEIF4E4*.

Figure 5: Subcellular localization of the *TbEIF4G3-5* homologues. (A) *T. brucei* cells were transfected with the p2216/ *TbEIF4G3-5* constructs and the expression was induced after tetracycline addition to the medium. The fluorescent proteins were detected through confocal microscopy. The three proteins showed cytoplasmic distribution with the regions correspondent to the nucleus emitting no fluorescence. Below are western-blot results confirming the expression of the EYFP fusions. (B) Immunofluorescence of *TbEIF4G4-5* using the isoform specific antibodies followed by incubation with the fluorescein-conjugated secondary antibody. Where indicated, the cells were counterstained to locate the nuclear and kinetoplast DNA.

Figure 6: RNA interference of *TbEIF4G3-5*. (A) Cell density of transfected cultures with and without tetracycline addition. (B) Western blot analysis of the time course. *TbEIF4G3-5* were detected with the affinity purified antisera and anti-*TbEIF4AI* was used as a loading control. The same blot was probed with both antibodies.

Figure 7: Phenotype of cells after RNA interference of *TbEIF4G3-5*. (A) Microscopy with altered morphology of *TbEIF4G4* RNAi. (C) Phenotype of RNAi of *TbEIF4G3-5* through effect on ³⁵S labeling.

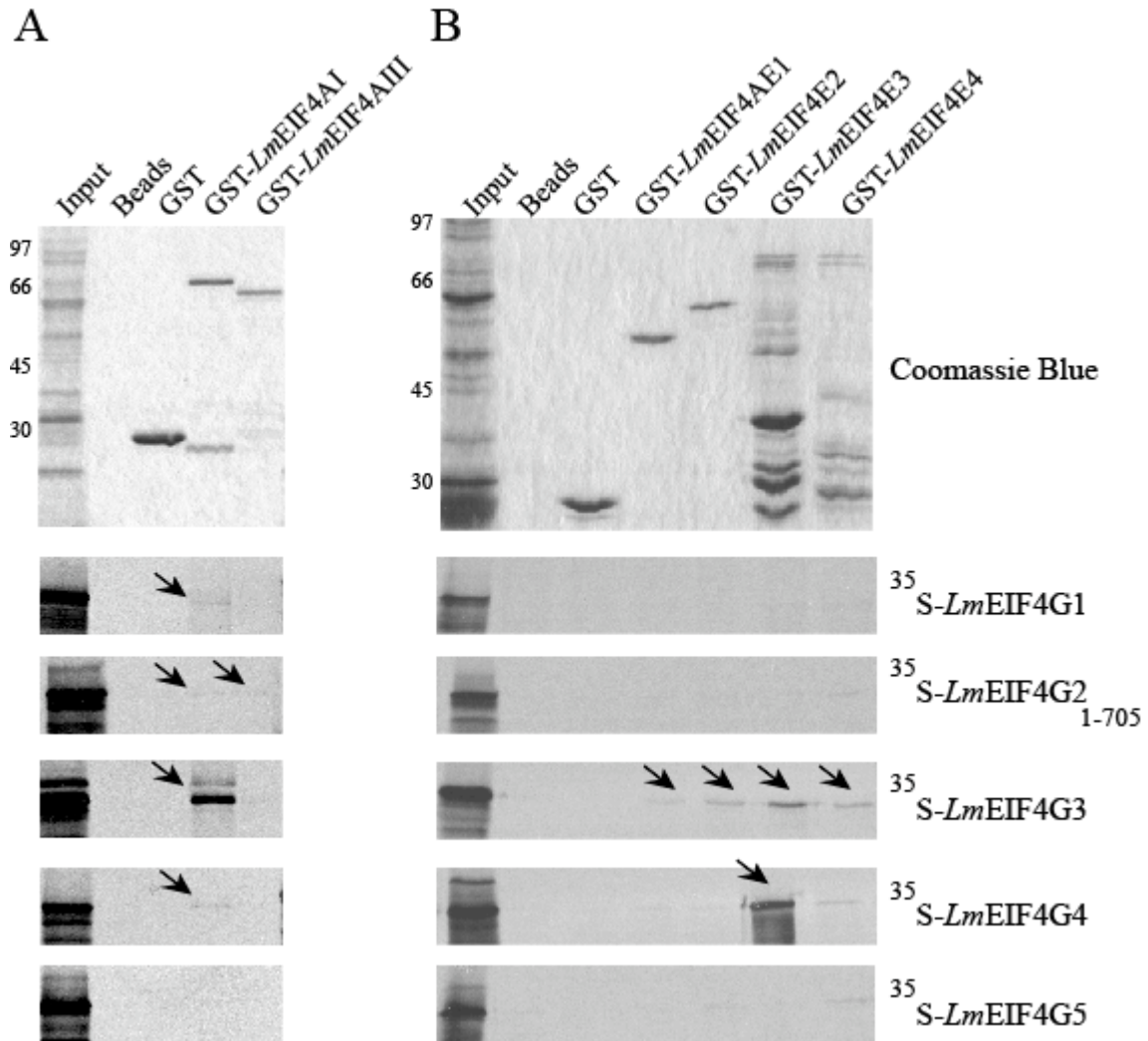
TABLE I – Primers used for the amplification and cloning of the various *L. major* eIF4G genes and gene fragments.

Protein	Vector	Oligonucleotide
<i>LmEIF4G1</i> ₁₋₄₇₅	pET21d	5' primer – GAACATGTTTCATGGAAACACAGATTTG
<i>LmEIF4G1</i> ₁₋₄₇₅	pET21d	3' primer – TGGCGGCCGCGCTCATGATGGAGGACTGCAG
<i>LmEIF4G1</i> ₁₂₈₋₁₀₁₆	pET21d	5' primer – GGACATGTTTATGTCCGGTCCGGAAGG
<i>LmEIF4G1</i> ₁₂₈₋₁₀₁₆	pET21d	3' primer – TGGCGGCCGCCGATAAGTATGTGAGGACGG
<i>LmEIF4G2</i> ₁₋₇₀₅	pET21d	5' primer – GTAAACCATGGAGCACTCAGGTC
<i>LmEIF4G2</i> ₁₋₇₀₅	pGEX	5' primer – GTGGGATCCATGGAGCACTCAGGTCGC
	pET21d/	
<i>LmEIF4G2</i> ₁₋₇₀₅	pGEX4T3	3' primer – TGCTCGAGCAGCATGGCACGCAGGCCATC
<i>LmEIF4G3</i> ₂₆₋₆₃₅	pET21d	5' primer – TGCCATGGATGAAGTGGTGCGACGCC
<i>LmEIF4G3</i> ₁₋₆₃₅	pET21d	5' primer – TGCCATGGAGTTCACCGTGGAGCAG
<i>LmEIF4G3</i> ₂₆₋₆₃₅ /	pET21d/	
<i>LmEIF4G3</i> ₁₋₆₃₅	pGEX4T3	3' primer – TCGCGGCCGCCTTGGGGAAGCGCTCCTTG
<i>LmEIF4G4</i> ₁₋₃₆₂ /	pET21a/	
<i>LmEIF4G4</i> ₁₋₇₆₅	pGEX4T3	5' primer – TTCGGATCCCCAATTATGCTCTTCAACCTG
<i>LmEIF4G4</i> ₃₇₋₃₆₂ /	pET21a/	
<i>LmEIF4G4</i> ₃₇₋₇₆₅	pGEX4T3	5' primer – CTTGGATCCGAGCCGAAGTTCTCGCTAGAG
<i>LmEIF4G4</i> ₁₋₃₆₂ /	pET21a/	
<i>LmEIF4G4</i> ₃₇₋₃₆₂	pGEX4T3	3' primer – TTGGCGGCCGCGTCGTTGTTGGTCGTGTTATTG
<i>LmEIF4G4</i> ₃₇₋₇₆₅ /	pET21a/	
<i>LmEIF4G4</i> ₁₋₇₆₅	pGEX4T3	3' primer – TTCGCGGCCGCCGTGTCGGGGTACATCTG
<i>LmEIF4G5</i> ₁₋₄₅₅	pGEX4T3	5' primer – CCTCCATGGACAAGAACTACCCGGGCCGCATC
<i>LmEIF4G5</i> ₁₋₄₅₅	pGEX4T3	3' primer – TTCGCGGCCGCCTCGTAGTCAGGAATCCGTG
<i>LmEIF4G5</i> ₁₃₈₋₇₈₂	pGEX4T3	5' primer – CGACCATGGACGACATGGAGAACAAC
<i>LmEIF4G5</i> ₁₃₈₋₇₈₂	pGEX4T3	3' primer – TTCGCGGCCGCCTCCGCTCCTTGATGGCCTC

TABLE II – Primers used for the site directed mutagenesis of the various *L. major* eIF4G genes and gene fragments. Mutated nucleotides are highlighted in bold.

Protein/ Mutation	Oligonucleotide
<i>LmEIF4G3</i> – FSL/AAA	5' primer – CGCCGTACCCGGG CGCTGCTG CGGATGAAGTGGTGCG
<i>LmEIF4G3</i> – FSL/AAA	3' primer – CGCACCATTTCATCC GCAGCAGC CCCCGGGTACGGCG
<i>LmEIF4G3</i> – KLV/AAA	5' primer – GCTTGACGCAGAC GGCGGGCGGCG CGGTGAGAATGC
<i>LmEIF4G3</i> – KLV/AAA	3' primer – GCATTCTCACCGCG CGCCGCCG CCGTCTGCGTCAAGC
	5' primer –
<i>LmEIF4G3</i> – LNK/AAA	CGACTGCTGCACGGCACAG CGGCCGCG CTCACAGAGGAGAAC
	3' primer –
<i>LmEIF4G3</i> – LNK/AAA	GTTCTCCTCTGTGAG CGCGGCCG CTGTGCCGTGCAGCAGTCG
<i>LmEIF4G4</i> – FSL/AAA	5' primer – CCTGAGCCGAAC GCCGCGGC AGAGCGCGTGCTGC
<i>LmEIF4G4</i> – FSL/AAA	3' primer – GCAGCACGCGCTCT GCCGCGGCG TTTCGGCTCAGG
<i>LmEIF4G4</i> – KLV/AAA	5' primer – GGCGGAGATTCCT GCGGCCGCC ACGTCCGAGAACGG
<i>LmEIF4G4</i> – KLV/AAA	3' primer – CCGTTCTCGGACGT GCGGCCGC AGGAATCTCCGCC
<i>LmEIF4G4</i> – LNK/AAA	5' primer – CCGCGTTCAGGGCT CGGCGGCCGCG CTCACCGACAAG
<i>LmEIF4G4</i> – LNK/AAA	3' primer – CTTGTGCGGTGAG CGCGGCCGCC GAGCCCTGAACGCGG
	5' primer –
<i>LmEIF4G4</i> – M20Y/S/D/A	GGAGAAGGTGAAGAACCAG KMT ACGCTTGCCGACATTCTCG
	3' primer –
<i>LmEIF4G4</i> – M20Y/S/D/A	CGAGAATGTCGGCAAGCGTA KMCT GTTCTTCACCTTCTCC
	5' primer –
<i>LmEIF4G4</i> – I25A/L26A	CCAGATGACGCTTGCCGAC GCTGCCG CTTTTCGCGACACATGG
	3' primer –
<i>LmEIF4G4</i> – I25A/L26A	CATGTGTCGCGAAAAGCG GCAGCG TCGGCAAGCGTCATCTGG

Reis et al., 2009 - Figure 1



Reis et al., 2009 - Figure 2

LmEIF4G3	1	M	QFTVEQIRSV	RNNYLEP-PY	PGFSIDEVVR	RRRLTQT---	---KLVVRGEN	AWVAKG-TAQ	TTBEWVQRLL	HCLLNKLTTEE	↓
TbEIF4G3	1	MH	VYTIQDQIEL	RSLYPEP-PY	PGFSIEEACR	RKKQTQT---	---KLVVRGFN	AWAARG-SAK	TTAEWVERIV	YCSLNKLSAA	
TcEIF4G3	1	MH	VYSIQDQIEE	RSMYKDA-PY	PGFSIEEACR	RKKMTQT---	---KLVVRGAN	AWARG-SAK	TTBEWVERIV	YCSLNKLSA	
LmEIF4G4	1	MLFNLRGIVP	QKEEKVKN--	QMTLADILAF	ROTWTAI-PE	PNFSLEVRLL	TARLEKQRKA	EIPKILVTSEN	GFRVRDKKDM	DASELEIRRV	QCSLNKLTDK
TbEIF4G4	1	MLFKPRGVTS	NDPRYAGGSR	IMSVSDILAY	ROTWGRREP	ESFSLAQIIR	DARAADCKPV	VTEKLVMSFN	GFKVKSRDSI	DSIERGVRAV	QASLNKLTETW
TcEIF4G4	1	MLFKPRGVTS	NDPKYAGNSR	IMTIPDILAY	ROTWKER-LS	DDPSLGRILF	AARAAKSKPA	APEKLVMTSN	GFKVKDRDAI	APSERGVRAV	QSTLNKLTES

MIF4G - HEAT DOMAIN

LmEIF4G3	74	NKDIMIDKIL	TKELFATEDI	MMNVNIIIRK	KALDEPENSK	LYAGVCHSLA	LYBAKVLDRG	HK-----	-----	GE-RPOSQR	DALISTACHE
TbEIF4G3	75	NFNEIVSQLQ	TNTIFSSDEM	LKKTVISIIR	KALCEPENSN	YVAGLCYKLA	EYEVSLNVVQ	KQ-----	-----	KEGKRISKLR	NAVVGIAQIE
TcEIF4G3	75	NFNDMAKLO	TEVIFSTKET	LNIARIRIRK	KALDEPECSC	SYAGVCYKLA	EKEVGITAAK	QK-----	-----	AEGKRSKLR	NAVVSIAQEE
LmEIF4G4	98	NFDADVVEAL	SPDLVLNPLV	LKGAVDIIRN	KAMAEPVFSG	IYAQLCQRIK	VVEQDLVAEA	AENPTSELGQ	LMAAVQAGDE	DQKNVNGVRV	NALVKRCQEF
TbEIF4G4	101	NFDVVTQTVL	TPDIIINNEV	VKDVVRLIYE	KALMEPVFAG	LYARMCFISV	RREYERTXKF	LP-----	-----	GTCQVPSAVR	VATVEKCOQM
TcEIF4G4	100	NFDLIAKAL	VETIILNANV	VADVRLIYVD	KALVEPVFAG	LYARLCVMIV	RKEYDYRSTQ	SG-----	-----	EVDAQSEVR	VATVEKCOQM

MIF4G - HEAT DOMAIN

LmEIF4G3	155	FRTLSEKLSK	MDDKSEEELE	YERSNVMMRK	RSNMRIFIGEL	YLCTALHTST	MFTVMDLTMR	CS---DKTGF	PSSNIEELLA	ALNTTICDRL	DKN-----
TbEIF4G3	157	FQNR-RNVPS	SEGLSEEEVE	QORSSFMRKK	VANMRIFIGEL	FLHKVLHSLT	MMDIINIIMQ	P---EKGGY	PASDELEFLT	VLFTIVGSKL	DSIA-----
TcEIF4G3	157	FLSR-RKMPS	MEGLTEEEEM	LRRITFMRRK	RANMKFIFIGEL	FMEKVLSCNT	MMNIQTIMQ	EA---ERGGY	PTSDIEFLT	ELFLTIGESL	DAVP-----
LmEIF4G4	198	HDSFVKQP--	-PPRNAPEAE	QLR---KRN	MANIKFVGEL	YMSRLSHRV	ILAVCSTAIV	LGFIPIPKLV	TTDADLEMMV	SLNMVIGKRF	DENAGMALNH
TbEIF4G4	183	ETNAVEESK-	-QTMSEEQAE	RLR---KRN	VNNKIFAGEL	FLKTLITQKI	IDRLINERIY	E-----MV	ENDMELEVII	NLEVVCKLY	EENK-----
TcEIF4G4	182	FDGATKAVER	-ETASEEEAE	KVR---KRN	VNNKIFAGEL	FLKSLITQKI	IDRLILGELF	N-----IT	ENDMDLEVIV	NLEVVCKMY	EENK-----

MIF4G - HEAT DOMAIN

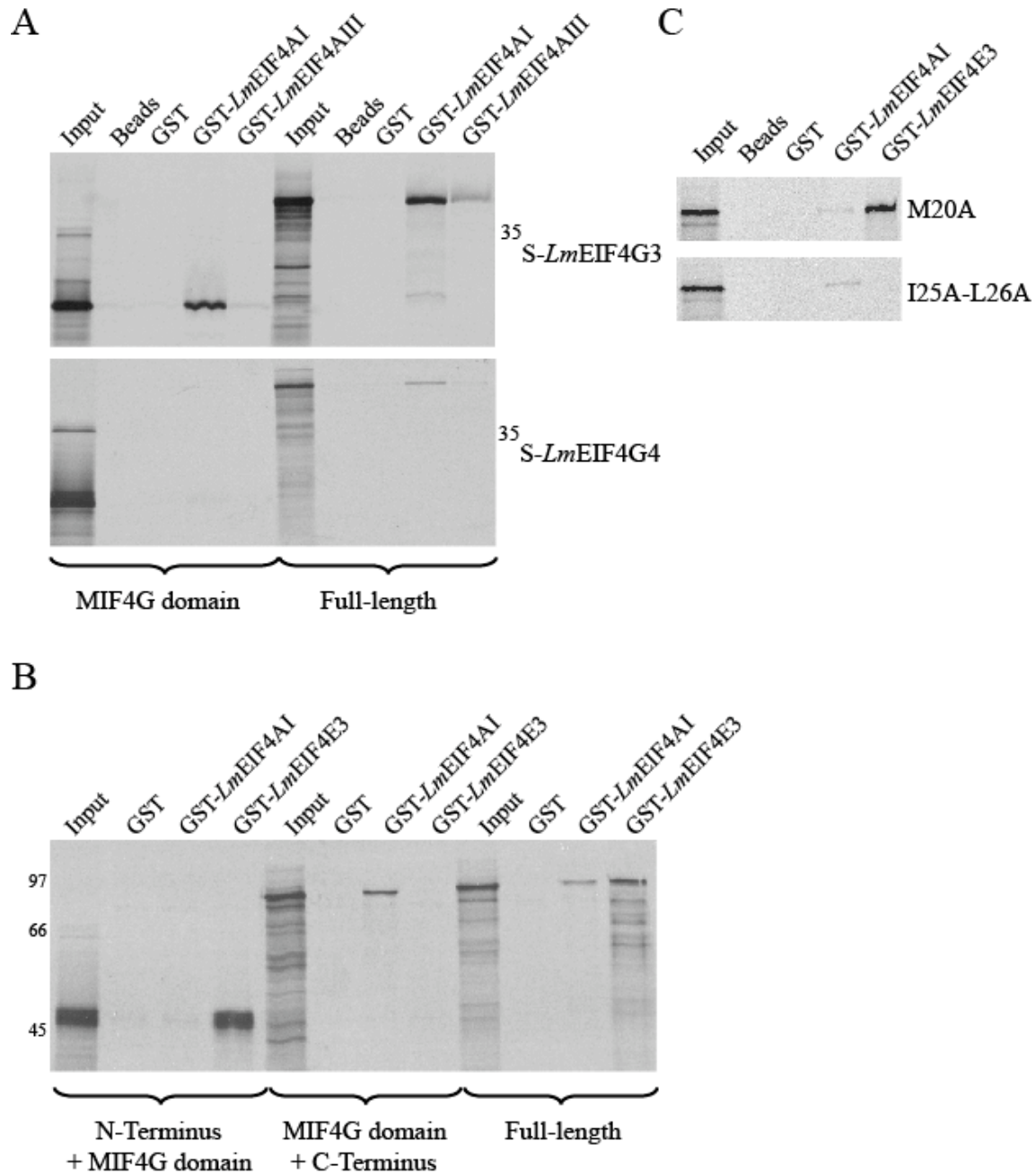
LmEIF4G3	244	-----HKK	TLDPYTSE	NFNKSDCHYV	PPRIREFKIMD	LIEDLR-KRGW	GLKPKPV---	-----	-----	-----	-----KA
TbEIF4G3	245	-----ELRP	KLDAYFKVLE	GLK-DQK-IY	PPRIREFKIMD	LIEDLRDRNW	ESR-----	-----	-----	-----	-----ET
TcEIF4G3	246	-----ELRV	RLNGYFELLE	KLK-DQKEVY	PPRIREFKIMD	LIEDLRSKFNW	ERR-----	-----	-----	-----	-----AT
LmEIF4G4	291	VPKDAAAAAH	PHDTIWRAPN	SSM---NDPRF	SNRIREFLLQN	LIEDLR-NDGW	KPKESPAQGS	SGGSADDSNN	TTNDRSNHR	SSNNACNHSN	TGGFSGPNEQ
TbEIF4G4	263	-----PD	AQQLWKMS	SMQ---ENRRY	SMRIREFLLQN	LIEDRR-NGGW	KPREPEQVVI	EDIQSQQQPQ	QQSQQQQQQQ	-----	-----QQ
TcEIF4G4	262	-----PA	AQPALWKLIG	ELQ---EERRF	SNRIREFLLQN	LIEDRR-NGGW	KPREPEQVIV	EDPQ-----	-----	-----	-----QQ

MIF4G - HEAT DOMAIN

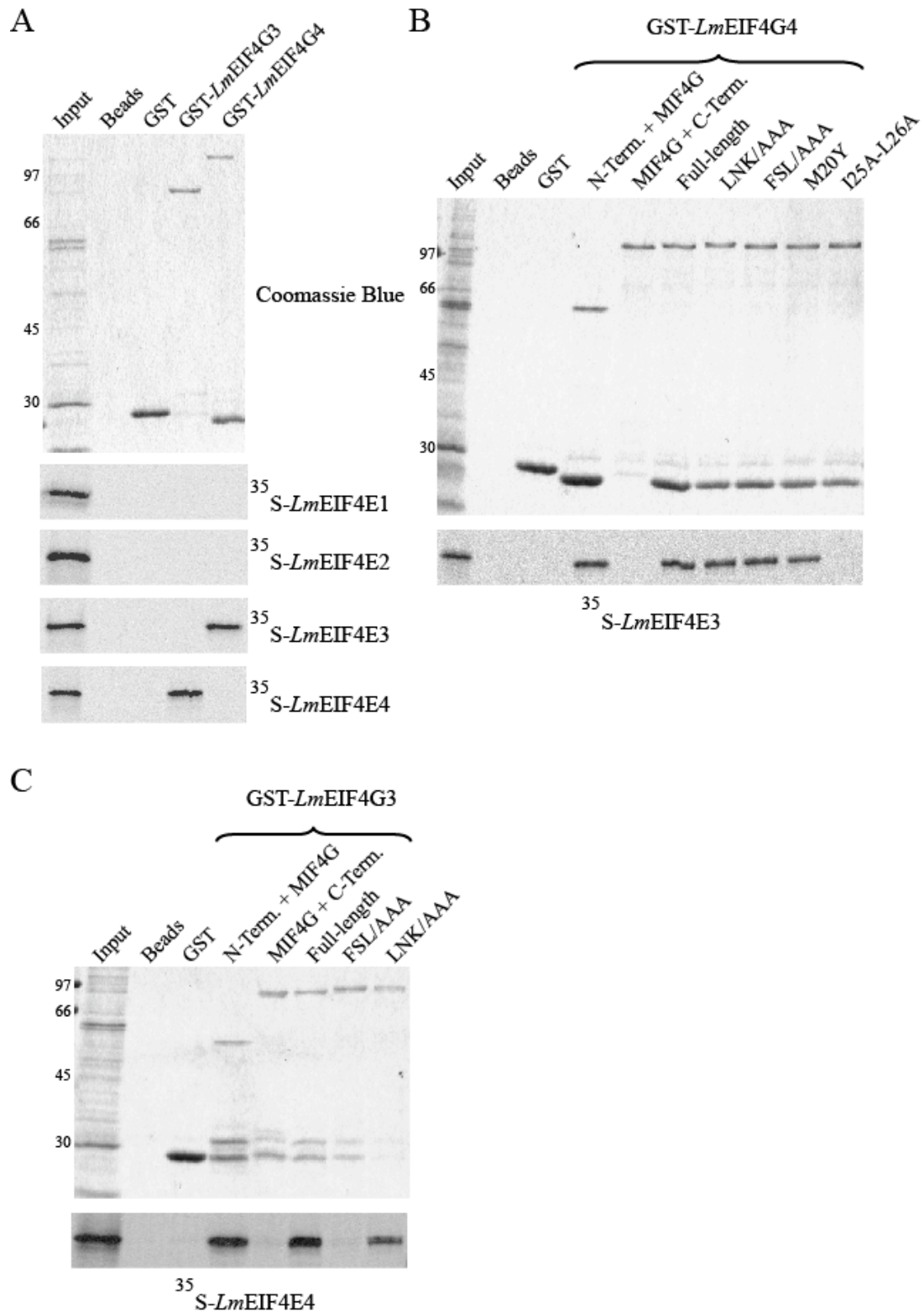
LmEIF4G3	296	ALPPSHGKDK	KYATAPPPKK	GGRDQGSAAA	A-----	-----	--SKSWRDAA	TKTKCTAAAP	AFSADKGLNG	RGTSAAAPAA	NTHALSSSSA
TbEIF4G3	293	VMPKTSAAQT	K-----EPPRA	SDRAVRSSTT	P-----	-----	--NSSNASSS	CSKKCKGAGP	CDSALG----	-QRAAASTVG	KDAREGLKSR
TcEIF4G3	295	AAPKTSSTQT	R-----EPTVR	TDKVVATRS-	-----	-----	--SMNASSS	CSKKCRQDSV	SDPAS----	-QAAAAAAA	KTAMEAGGR
LmEIF4G4	388	MHRNTAPNPF	SGSGSWQQQM	QQQHSGGRG	IIGSSNAGGN	RGGTRPPSMA	DRGDEKYGRG	CGPGGLRQGP	VSSFNDLSQF	PASNDHMMPP	PPFFAAPPPP
TbEIF4G4	335	QQSQQQPQQ	QQQGQKHYSQ	QQRHQQSFP	HNQPSHGGY	HQHPPPPPPP	PRNSEYQHTP	CMKRCGSYHE	VSSSLGMGGR	CSQSTNDLTR	GGSYNNMPPP
TcEIF4G4	324	QQQQQQHQHQ	RYFHHHHHHQ	QQQHQQHHQQ	HHQQQ----	-----	QRNSDQLOTP	IGKRAKSYQD	VSNPYGVGAR	RSTSHLDLPG	YGAYAMPPPP

POSSIBLE MA3 DOMAIN

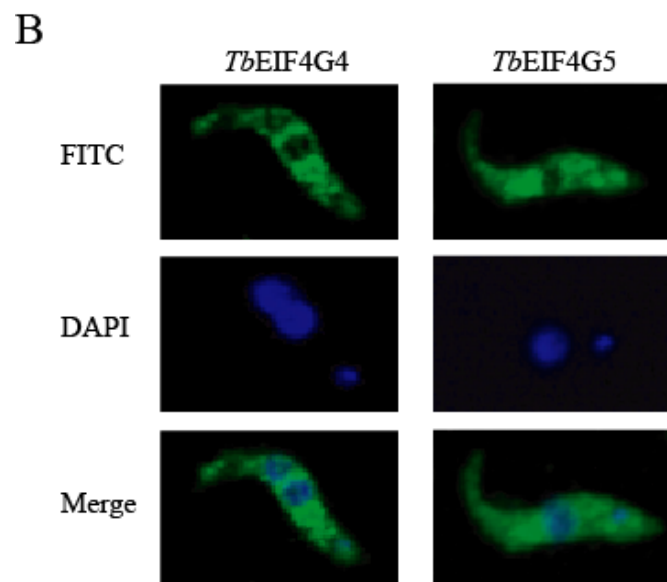
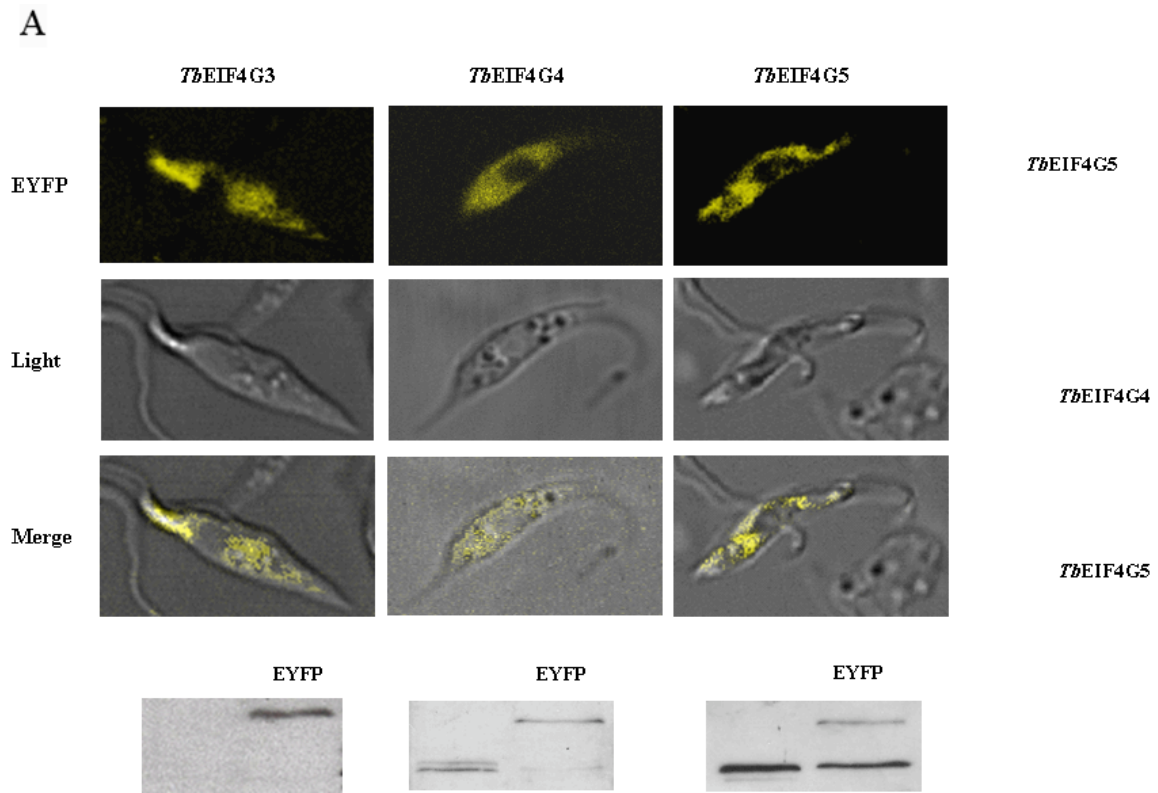
LmEIF4G3	374	-----STAA	ATAPAGASF	DEASPOVPLV	KFELRLVASME	QEWVADRTND	VILHVDQFN	TCDRFFESES	ELCVAVQEV	IHSACTTTRK	
TbEIF4G3	362	-----AWRD	VVKSEIVACD	DHEMTVTESV	AFEBEVRSLF	QEWLSGYRTG	CLPNQEEFR	DCGARFEPDM	DLPTAVAAQV	VREACMTTRK	
TcEIF4G3	360	-----SWSN	VVKSPQSPRD	ATVARAVETV	NFBARVSRLE	QEWVARECSND	FIPETMNEFR	HCQRHFDSED	ALCKAAAEV	VREACMTTRK	
LmEIF4G4	488	FPGEGIMPPS	MAPQAASAPG	PQMSDEDRLK	MASTPPKCV	DADLSSRIIS	CIRDAHTDGD	WAAAGKAIVE	A---VFQDAA	HMCRCMAEVV	IACKVTETSS
TbEIF4G4	434	-----AHGFS	SRLAPEERKR	LNFAIRPPAQ</							



Reis et al., 2009 - Figure 4

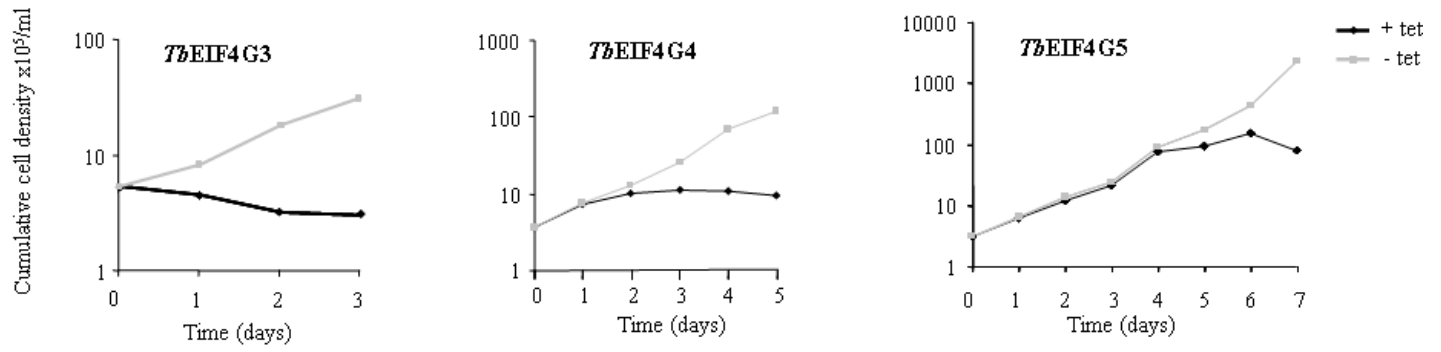


Reis et al., 2009 - Figure 5

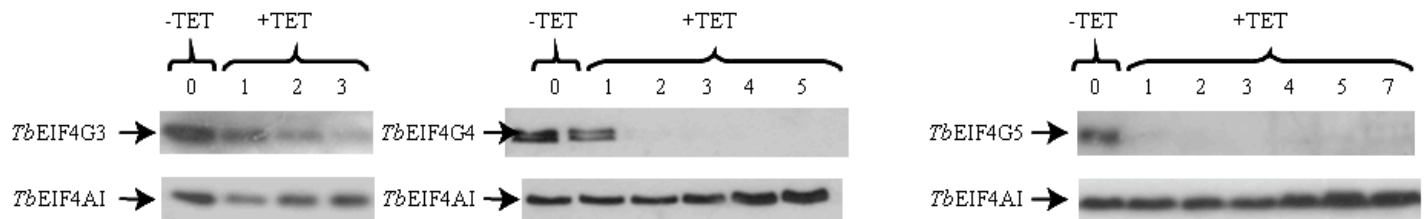


Reis et al., 2009 - Figure 6

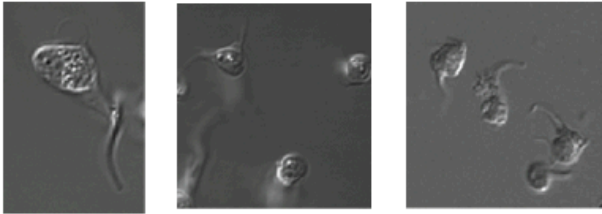
A



B



A



B

TbEIF4G3

0 6 24



TbEIF4G4

0 48 72



TbEIF4G5

0 72

