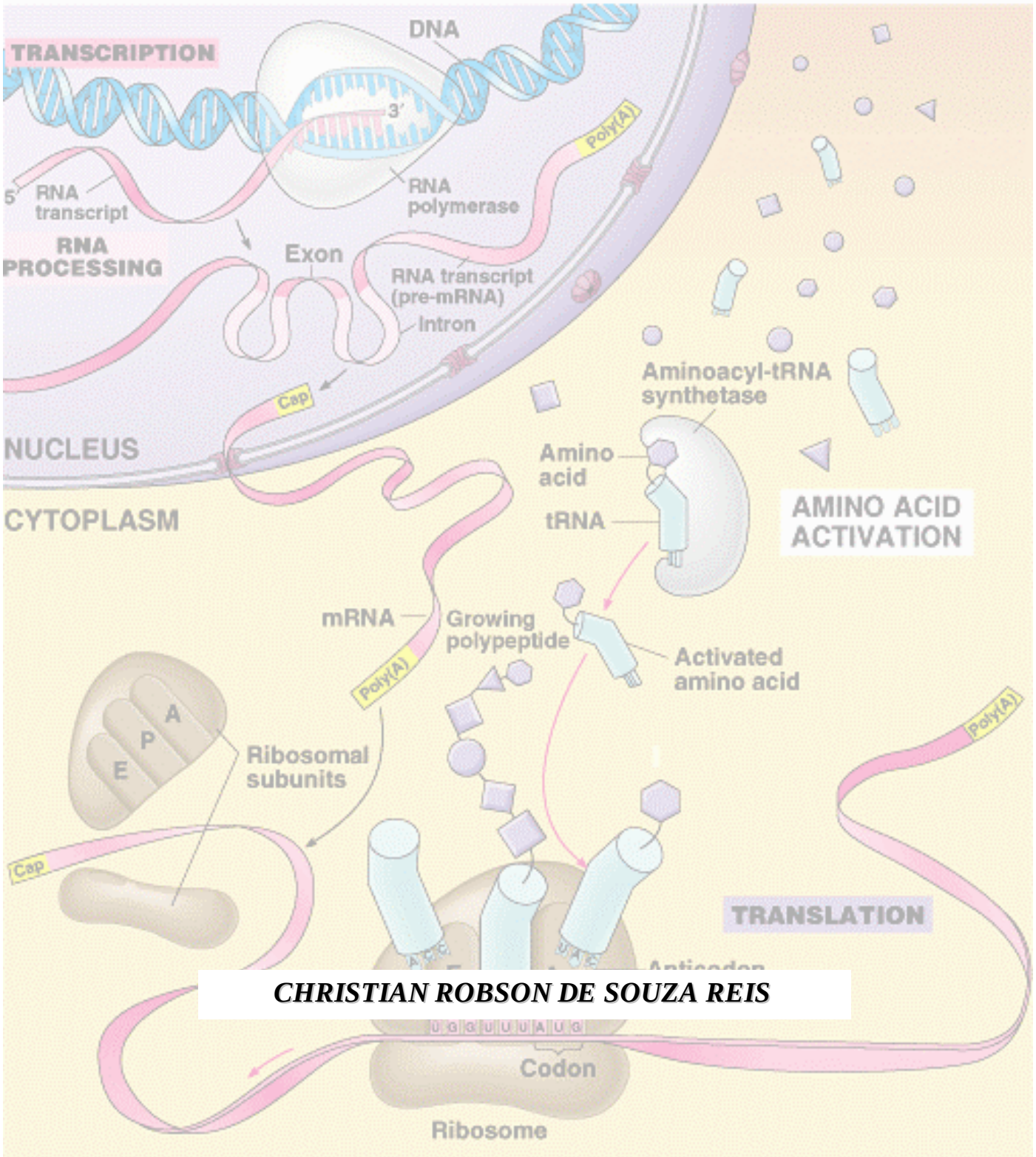
UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
CURSO DE PÓS-GRADUAÇÃO EM GENÉTICA



CHRISTIAN ROBSON DE SOUZA REIS

RECIFE, FEVEREIRO DE 2004

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
CURSO DE PÓS-GRADUAÇÃO EM GENÉTICA

CARACTERIZAÇÃO DE POSSÍVEIS HOMÓLOGOS AOS FATORES DE INICIAÇÃO DA TRADUÇÃO eIF4G E eIF4A DE *Leishmania major*

Dissertação apresentada ao Curso de Pós-graduação em Genética da Universidade Federal de Pernambuco - UFPE para obtenção do grau de Mestre em Genética, área de concentração Biologia Molecular

Orientador: Dr. Osvaldo Pompílio de Melo Neto

CHRISTIAN ROBSON DE SOUZA REIS

RECIFE, FEVEREIRO DE 2004

Reis, Christian Robson de Souza

Caracterização de possíveis homólogos a fatores de iniciação da tradução eIF4G e eIF4A *Leishmania major* / Christian Robson de Souza Re – Recife : O Autor, 2004.

102 folhas : il., fig.

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCB. Genética, 2004.

Inclui bibliografia.

1. Genética molecular – Síntese protéica. 2. Proteínas (eIF4G – eIF4A) – Biossíntese. 3. Parasitologia – Tripanosomatídeos (*Leishmania major*) – Iniciação da tradução. I. Título.

577.217	**CDU (2.ed.)**	**UFPE**
572.645	**CDD (21.ed.)**	**BC2003-259**

BANCA EXAMIDADORA

Titulares:

Dra. Alzira Maria Paiva de Almeida – CPqAM-FIOCRUZ/MICROBIOLOGIA

Dra. Andréa Queiroz Maranhão – UNB/BIOLOGIA CELULAR

Dr. Marcel Ivan Ramirez Araya – CPqAM-FIOCRUZ/IMUNOLOGIA

Suplentes:

Dra. Norma Lucena Cavalcanti Licinio da Silva – CPqAM-FIOCRUZ/IMUNOLOGIA

Dra. Constância Ayres – CPqAM-FIOCRUZ/ENTOMOLOGIA

DEFESA PÚBLICA DE DISSERTAÇÃO

Local: Auditório do Centro de Pesquisas Aggeu Magalhães – FIOCRUZ/PE

Data: 26/02/2004

Horário: 14:30

DEDICATÓRIA

A minha família (Painho, mainhã, Lelé, Rose, Rafa, Brother, vós Julieta e Laura, meu avô Sebastião, tias, primos e demais familiares) por terem compartilhado comigo os momentos de dificuldade e os momentos de alegria, por terem apostado em meu potencial. Agradeço, todos os dias por ter tido uma educação balizada em ótimas virtudes e bons preceitos. O que aprendi certamente passarei a meus filhos, certo que serão pessoas de garra, persistentes no que abraçarem. Obrigado Pai, Mãe e todos os "Reis", uma família que luta, vence, e também perde porém nunca desiste.

AGRADECIMENTOS

Ao senhor criador de toda a perfeição manifestada de forma criativa na natureza e muito superficialmente observada através da ciência.

A mim mesmo por ter enfrentado de maneira corajosa todos os desafios e obstáculos durante estes dois anos, transpondo-os com sabedoria, perseverança e garra.

A meu orientador Osvaldo Pompílio de Melo Neto por ter influenciado positivamente (desde 1999) minha carreira acadêmica.

A Rafael Dhalia, um bom exemplo a ser seguido. Um bom colega e um ótimo pesquisador.

As minhas duas co-orientandas Tami e Dani. Duas excelentes alunas com um futuro brilhante.

Aos sêniors do grupo de Osvaldo: Éden, Alessandra, Cheila, Patrícia, Rodolfo e Polly. No que abraçarem em suas vidas desejo sincero sucesso.

Ao grupo dos eIFs: Comparem-se a fatores eIFs, cada um tem um papel, mas todos trabalham em conjunto. Obrigado para aqueles que me ajudaram direta ou indiretamente!

A todas as pessoas do departamento de Microbiologia, no presente e no passado – Bolivar, Bruno (calção), Mirna, Élcio (frescura...), Emanuelle, Mariana Lacerda, Felipe, Raquel, Cláudia, Henrique, Rodrigo, Mariana Marques, Marina, Cariri, Eleonora, Gerlane, Soraya, *et al*.

Ao apoio técnico: Isaac (martinau) Pedrada, Nelson, Rita, Silvana, Yara, Conceição.

As Doutoras Nilma, Alzira, Cris, Marise, Raquel Coimbra.

A todos meus amigos (Antônio, Fábio, Rogério, Sérgio), meus familiares, colegas de graduação (Edmar, Fábio Neri, Janielle, Joelma, Sidney) e de mestrado, obrigado!

A frase "Não precisa correr tanto; o que tiver de ser seu às mãos lhe há de ir." Machado de Assis. (Dom Casmurro) Cap. LXVI (Intimidade)

SUMÁRIO

LISTA DE FIGURAS DO ANEXO

LISTA DE ABREVIATURAS

ATP – Adenosina Trifosfato

cDNA – DNA complementar

DEPC – Dietil pirocarbonato

DTT – Ditiotreitol

ECL – Deteção quimiofluorescente

EDTA – ácido etilenodiaminotetracético

eIF – Fator de iniciação da tradução eucariótico

GST – Glutationa Sefarose Transferase

GTP – Guanosina Trifosfato

HEAT – ***H****untingtin,* ***E****longation factor 3,* ***A*** *subunit of protein phophatase 2A, and* ***T****arget of rapamycin*

IPTG – Isopropil-tio-â-D-galactopiranosídeo

kDa – kilodálton

LB –Luria Bertani

mRNA – ácido ribonucléico mensageiro

nt- nucleotídeo

NTP – Nucleotídeo trifosfato

ORF – Seqüência aberta de leitura

PABP – proteína de ligação a cauda poli-A

PBS – solução salina tamponada com fosfato

PCR – Reação em cadeia da polimerase

p/v – peso volume

RNA pol – RNA polimerase

RRMs – Motivos de reconhecimento de RNA

rRNA – ácido ribonucléico ribossômico

SDS-PAGE – Gel de poliacrilamida em condições desnaturantes

SL – seqüência *spliced leader* ou mini-éxon

snRNA – pequeno RNA nuclear

VSG – Glicoproteína de superficie de *Trypanosoma brucei*

UTR – Região não traduzível

TE – Tampão Tris-EDTA

tRNA – ácido ribonucléico transportador

tRNAi – tRNA iniciador

RESUMO

Em eucariotos, a iniciação da tradução é um processo essencial de regulação pós-transcricional da expressão gênica. Neste processo, atuam proteínas designadas eIFs (fatores de iniciação da tradução). Destas destaca-se o complexo eIF4F – eIF4E, eIF4A e eIF4G – que permite associar o mRNA ao ribossomo. Identificamos no genoma de *Leishmania major* seqüências que apresentam homologia aos componentes do eIF4F. Este trabalho contempla a caracterização de um homólogo do eIF4A (LmeIF4A2) e três homólogos eIF4G (LmeIF4G1-3). O gene LmeIF4A2 foi clonado, expresso e utilizado na produção de anticorpos. Ensaios de *Western-blot* sugeriram não haver expressão do LmeIF4A2 na fase promastigota de *Leishmania major*. Em seguida realizamos construções com os LmeIF4G1-3, fusionando-os a GST, permitindo a realização de ensaios de *pull down* visando investigar sua associação com as proteínas LmeIF4A1-2. O LmeIF4A2 não interage com nenhum dos LmeIF4G1-3 e o LmeIF4A1 interage especificamente com o LmeIF4G3. Em outra etapa produzimos anticorpos contra as proteínas LmeIF4G1-3 para avaliar sua expressão na forma promastigota do parasita. A expressão do LmeIF4G3 foi confirmada, todavia não detectamos a expressão dos LmeIF4G1-2. A interação eIF4G/eIF4E foi investigada em ensaios onde homólogos LmeIF4G1-3 foram incubados a homólogos LmeIF4E1-3, para tentar reconstituir parcialmente o complexo eIF4F *in vitro*, e testados quanto a sua afinidade pelo cap sintético. Estes resultados se mostraram inconclusivos. A utilização de novas abordagens e a caracterização dos demais fatores será importante na elucidação da tradução nestes protozoários.

ABSTRACT

In eukaryotes, translation initiation is a critical process for the post-transcriptional control of gene expression. In this process, act proteins called eIFs (from "eukaryotic initiation factors"). Within the eIFs the eIF4F complex – eIF4E, eIF4A e eIF4G – allows the recognition of the mRNAs by the small ribosomal subunit. We have identified, in the *Leishmania major* genome, sequences that show homology to the components of eukariotic eIF4F. Here we describe the preliminary characterization of one eIF4A (LmeIF4A2) and three eIF4Gs homologues (LmeIF4G1-3). The LmeIF4A2 gene was cloned, expressed and used to produce antibodies. Western-blot experiments indicate that there is no expression of LmeIF4A2 in *L. major* promastigotes, in contrast with LmeIF4A1 which is expressed as a very abundant protein. We then subcloned previously clone fragments from the LmeIF4G1-3 genes into the pGEX4T3 vector so as produce the corresponding proteins fusioned to GST. These proteins were used in pull down assays to investigate their interactions with both *Leishmania* eIF4A homologues. The LmeIF4A2 doesn't seem to bind to any of the eIF4G homologues whilst LmeIF4A1 binds specifically to LmeIF4G3. We also produced antibodies to the LmeIF4G1-3 proteins to analyze their expression in the parasite promastigotes. We confirmed the LmeIF4G3 expression, however we were not able to detect the expression of LmeIF4G1-2. Finally, the interaction eIF4G/eIF4E was investigated in assays where the LmeIF4G1-3 homologues were mixed individually with each of the three *Leishmania* eIF4Es available so as to partially reconstitute candidate *Leishmania* eIF4Fs *in vitro*. The complexes were then tested for their affinity for synthetic cap. These results so far were inconclusive. New approaches and the characterisation of others translation factors will be important in understanding translation and its regulation in this protozoan family.

I. INTRODUÇÃO

Os protozoários tripanosomatídeos são eucariotos causadores de enfermidades comuns a muitas comunidades humanas, inclusive a brasileira. Estas doenças tem impacto significativo no sistema de saúde dos países acometidos por colocarem em risco milhares de pessoas e por levarem muitas vezes a problemas permanentes de saúde, constituindo assim também um problema econômico. As Tripanosomíases e as Leishmanioses infectam e matam milhares de pessoas por ano e milhões estão em área de risco (www.who.org, 2002).

Este grupo de microorganismos possuem características próprias em comparação aos demais eucariotos, como a transcrição policistrônica, *trans-splicing*, editoramento de RNAs mitocôndriais entre outros. A transição entre dois estágios evolutivos em hospedeiros específicos requer uma intensa mudança no repertório de expressão gênica. Os tripanosomatídeos parecem perder a canônica regulação da expressão dos genes, o controle transcricional, típico de eucariotos. Portanto, o controle da expressão genética deve acontecer predominantemente por mecanismos pós-transcricionais, e a síntese de proteínas parece ser o principal alvo de regulação nestes parasitas (Clayton, 2002).

A tradução é o processo pelo qual a informação contida na molécula de RNA é utilizada na formação de proteínas. Este evento acontece nos ribossomos e pode ser dividido didaticamente em Iniciação, Elongação e Terminação. A Iniciação é, sem dúvida, o principal alvo de regulação, uma vez que constitui uma etapa crítica no qual ocorre a ligação do mRNA ao ribossomo (Gingras *et al.*, 1999 ; Prévôt *et al*., 2003). Neste processo, encontramos como principal mediador um conjunto de três fatores de iniciação da tradução – eIF4E, eIF4G e eIF4A – que unidos formam um complexo chamado *cap binding complex* ou eIF4F. Este complexo protéico, bem como o seu papel vêm na iniciação da tradução, vêm sendo estudado em eucariotos, tais como mamíferos e leveduras; contudo, na literatura existe poucas informações relativas a esse assunto em tripanosomatídeos. Num esforço de esclarecer o processo de tradução nestes protozoários, nosso grupo procurou estudar possíveis homólogos das subunidades do eIF4F de *Leishmania major*. Para tal, genes que codificam homólogos aos fatores eIF4E, eIF4A, eIF4G foram identificados no genoma de *L. major* e as respectivas proteínas estão sendo caracterizadas bioquimicamente.

Dentro deste contexto, este trabalho teve por objetivo inicial caracterizar homólogos ao fator de iniciação da tradução eIF4G em *L.major* (neste trabalho homólogos serão designados por LmeIFs). Inicialmente, regiões conservadas dos genes LmeIF4G1-3 expressas em *Escherichia coli* (Dhalia, resultados não publicados) foram purificadas e utilizadas para imunizar

coelhos visando a obtenção de soro policlonal específico para cada isoforma. Com a identificação de um segundo homólogo ao eIF4A em *L.major* (LmeIF4A2), partiu-se também para a clonagem de seu gene, expressão da proteína recombinante em bactéria e imunização em coelhos para produção de soro policlonal. De posse dos anticorpos avaliou-se o perfil de expressão por *Western-blot* de algumas dessas proteínas. Também foram obtidas diferentes construções com as várias proteínas que permitiram sua utilização em ensaios de interação proteína/proteína, visando investigar a capacidade de ligação entre as mesmas. E por fim, realizamos um ensaio de afinidade entre os homólogos do eIF4E (LmeIF4E1-3) e o cap de mamíferos na presença de homólogos do eIF4G. Resultados deste trabalho foram utilizados na confecção do artigo intitulado *Translation initiation in Leishmania major: characterisation of multiple eIF4F subunit homologues*, o qual será submetido a revista *Molecular and Biochemical Parasitology*.

II. OBJETIVOS

⇒ **Objetivo Geral**:

- Caracterização de possíveis homólogos aos fatores de iniciação da tradução eIF4G e eIF4A de *Leishmania major*

⇒ **Objetivos Específicos**:

i)- Clonar o gene LmeIF4A2;

ii)- Expressar, purificar e imunizar coelhos utilizando as proteínas LmeIF4A2 e LmeIF4G1-3;

iii)- Analisar a expressão das referidas proteínas na forma promastigota de *Leishmania major* por *Western-blot*.

iv)- Subclonar a região central dos homólogos eIF4G no pGEX4T3 para análise por *pull down*

v)- Analisar ligação das proteínas LmeIF4E1-3 a cap sintético na presença dos homólogos eIF4Gs

III. REVISÃO BIBLIOGRÁFICA

1. Os tripanosomatídeos, doenças e impacto

Os tripanosomatídeos são protozoários flagelados da ordem Kinetoplastida, cujas principais espécies, representantes dos gêneros *Trypanosoma* e *Leishmania,* apresentam importância médica e veterinária. Estes organismos são considerados eucariotos primitivos, por constituírem um ramo que divergiu cedo da linhagem evolutiva que originou os demais eucariotos. Caracterizam-se por apresentarem dois estágios evolutivos bem definidos em hospedeiros diferentes, um hospedeiro invertebrado e um hospedeiro vertebrado. São causadores de enfermidades com distribuição endêmica em várias regiões do mundo, tais como a doença do sono, doença de Chagas e as Leishmanioses.

As Leishmanioses compreendem um grupo de doenças causadas por mais de 20 diferentes espécies e subespécies do gênero *Leishmania*. Estas patologias apresentam um amplo espectro de manifestações clínicas, destacando-se os comprometimentos cutâneo e ou visceral. São encontradas de forma endêmica em 88 países e no último levantamento sobre Leishmanioses, a Organização Mundial de Saúde registrou aproximadamente 2 milhões e 400 mil novos casos e 59000 mortes (WHO, 2002). No Brasil são notificados aproximadamente 2000 casos anuais de Leishmaniose visceral, sendo a maior parte dos registros de ocorrência na região Nordeste (FUNASA, 2002).

Estas parasitoses são de difícil controle. As drogas utilizadas no tratamento possuem efeitos colaterais e exigem um esquema terapêutico demorado. Um fenômeno freqüentemente associado à terapêutica inadequada é o mecanismo de resistência do parasita a droga (Leandro & Campino, 2003). Desta forma, o entendimento de processos biológicos fundamentais nestes protozoários – como a tradução – pode contribuir na elaboração de melhores alternativas de tratamento, e identificação de novos alvos para a ação de quimioterápicos.

2. Transcrição policistrônica

O arranjo de genes em tripanosomatídeos se assemelha em alguns pontos aos conhecidos operons bacterianos. Os genes deste grupo de eucariotos são transcritos em unidades policistrônicas, onde um único promotor é utilizado na transcrição de múltiplos genes. Assim como em procariotos, os genes dos tripanosomatídeos não possuem íntrons, com uma única exceção até o momento (Mair *et al*., 2000). Todavia, diferentemente do operon encontrado em

bactérias e nematódas, as proteínas codificadas por um mesmo policístron não parecem ser relacionadas e o número de sequências codificadoras por RNA pode ser bem maior (Blumenthal *et al*., 2002). Também, ao contrário das bactérias esses RNAs sofrem um processo de maturação através de um mecanismo chamado *trans-splicing* (Sutton *et al*., 1986; revisto por Campbell, 2003). Através deste processamento segmentos do mRNA policistrônico precursor contendo uma única sequência codificadora são clivados gerando mRNAs maduros monocistrônicos. Neste processo, os diferentes RNAs mensageiros recebem em suas extremidades 5' uma seqüência de 39 nt denominada de *Spliced-Leader* – SL ou *mini-exon* (Parsons *et al*., 1986; revisto por Campbell, 2003), que é característica de todos os mRNAs de tripanosomatídeos.

2.1 Promotores e RNA polimerases

A transcrição em eucariotos, incluindo os kinetoplastídeos, ocorre pela ação de três RNA polimerases que são farmacologicamente identificadas com base em sua sensibilidade (RNA pol-I) ou resistência (RNA pol-II e RNA pol-III) a droga ∝-amanitina (Kooter & Borst, 1984; Earnshaw *et al*., 1987). Além disso, homólogos de subunidades da RNA polimerase também estão presentes em tripanosomatídeos, assim como homólogos a fatores de transcrição (Clayton, 2002).

Nestes eucariotos, a RNA pol-I transcreve os genes que codificam para o RNA ribossomal (rRNA) e, além disso, em *Trypanosoma brucei* transcreve também os genes de antígenos de superfície estágio específicos (VSG e PARP) (Lee & Van der Ploeg, 1997). A RNA pol-II transcreve genes que codificam proteínas, pequenos RNAs nucleares (snRNAs) e o principal RNA envolvido no processo de maturação do mRNA: o RNA-SL (Stiles *et al*., 1999; Gilinger & Bellofatto, 2001). A RNA pol-III transcreve o RNA transportador (tRNA), o RNA ribossômico 5S (rRNA 5S) e também snRNAs (Lee & Van der Ploeg, 1997).

A procura por promotores em tripanosomatídeos tem sido alvo de inúmeras pesquisas, porém os resultados obtidos não são plenamente satisfatórios. Assim, promotores para RNA pol-I e para RNA pol-III foram encontrados com êxito (Campbell, 2003). Já promotores para RNA pol II ainda não foram descritos. Entretanto, recentes evidências sugerem que uma região presente no cromossomo 1 de *L.major* poderia ser um promotor para RNA pol-II (Martínez-Cavillo *et al*., 2003). Todos os fatos sugerem que estes parasitas perdem a capacidade de regular a expressão dos genes na transcrição, fenômeno típico de outros eucariotos. Propõem-se que a RNA pol-II transcreva os genes em baixos níveis e após processamento e formação dos mRNAs maduros, se dê o controle da expressão gênica (Clayton, 2002).

3. Trans-splicing

O *trans-splicing* é um mecanismo de processamento de mRNA que difere do *cis-splicing* convencional – típico de células eucarióticas – pelo fato de dois exons provenientes de diferentes transcritos serem fusionados para a formação do mRNA maduro (Harris *et al*., 1995) (Bruzik *et al*., 1998). Através deste evento, uma seqüência de 39 nucleotídeos (*SL*) é retirada da região 5' de um RNA precursor (RNA SL de 85 a 140 nt) e adicionada à extremidade 5′ de todos os mRNAs de tripanosomatídeos (Schnare & Gray, 1999).

O *trans-splicing* é um mecanismo de processamento de mRNA típico da família tripanosomatidae na qual foi primeiramente descrito (Nielsen, 2001). Este processo também foi encontrado em tremátodas (Bektesh *et al*., 1988); nemátodas (Liou & Blumenthal, 1990; Maroney *et al*., 1995); cestódeos (Brehm *et al*., 2000); *Euglena* (Frantz *et al*., 2000); cnidários (Stover & Steele, 2001); e no homem (Flouriot *et al*., 2002). Em pelo menos dois filos, Tremátoda e Nematoda, os processos de *cis-splicing* e *trans-splicing* conhecidos coexistem nas mesmas células. Até pouco tempo julgava-se que em tripanosomatídeos apenas ocorria *trans-splicing*. Contudo descobriu-se que genes que codificam para a poli-A polimerase em *T. brucei e T.cruzi*, sofrem *cis-splicing*, constituindo o primeiro caso descrito deste tipo de processamento, nestes protozoários (Mair *et al*., 2000).

Em tripanosomatídeos, o processamento em *trans* parece ser acoplado à clivagem da região 3' e aos eventos de poliadenilação (LeBowitz *et al*., 1993). A poliadenilação tem múltiplas funções notadamente na estabilidade do mRNA. Um sinal típico para a poliadenilação em eucariotos, a seqüência AAUAAA, está ausente em tripanosomatídeos (Ullu *et al*., 1993). Trabalhos têm sugerido que *trans-splicing* e poliadenilação são intimamente relacionados e acontecem acopladamente. Seqüências presentes no RNA policistrônico conhecidas como tratos polipirimidínicos (ricas em uridinas) servem como sinais para o *trans-splicing* e parecem funcionar, na ausência de sinais típicos, na sinalização para poliadenilação (Clayton, 2002).

3.1 *Spliced Leader e cap*

Como dito anteriormente, os genes SL RNA são transcritos pela RNA pol-II. Existe de 100 a 200 cópias destes genes por genoma, cuja transcrição representa 6% do total da síntese de RNA da célula (Campbell, 2003).

Após a transcrição, o SL RNA de ~140 nt, assume uma estrutura secundária complexa contendo três estruturas. Estas são definidas pelo pareamento interno de seqüências complementares intercaladas por regiões não pareadas (alças) associadas a ribonucleoproteínas. As estruturas recebem a classificação de grampos, ou *stem-loops*, I, II e III (Cross *et al.*, 1991). O grampo I contém a SL (extremidade 5' do RNA SL) e os grampos I e II são importantes na formação do cap4 (Sturm & Campbell, 1999).

O cap é um nucleotídeo metilado adicionado à extremidade 5' dos mRNAs de eucariotos com o objetivo de estabilizar, transportar e sinalizar para o início da tradução (Lewis & Izaurralde, 1997). Sua adição ocorre através da condensação da guanosina trifosfato (GTP) com a terminação 5' trifosfato do mRNA transcrito, catalisada por uma enzima nuclear denominada guanililtransferase. De acordo com o grupo de eucariotos envolvidos o cap apresenta modificações, por metilação, em seus nucleotídeos adjacentes. Assim, o cap típico de mamíferos é o cap monometilado (cap 7-metil-guanosina); em nematóides o cap presente na seqüência SL é trimetilado (2,2,7-tri-metil-guanosina). Em tripanosomatídeos, o cap é uma estrutura complexa denominada de cap4, que além do radical 7-metil guanosina inclui outros 4 nucleotídeos metilados (m^7guanosina (5′) ppp (5′) -N^6, N^6, 2′-*O*-tri-metil-adenosina-p-2′-*O*-metil-adenosina-p-2′-*O*-metil-citosina-p-3,2′-*O*-di-metil-uridina) (Bangs *et al.*, 1992; Ullu & Tschudi, 1995). O cap4 é adicionado aos pré-mRNAs juntamente com a SL através do *trans-splicing* (Campbell, 2003).

4. O processo de síntese protéica em eucariotos

4.1 Visão geral

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. A regulação da tradução é clara em eventos celulares tais como crescimento, proliferação e desenvolvimento (Gingras *et al*., 1999). O processo de tradução pode ser dividido didaticamente em três etapas: iniciação, elongação e terminação. Destes, a iniciação é o principal alvo de regulação; o ponto crítico, uma vez que é nessa fase que o ribossomo irá ser recrutado ao RNA mensageiro e posicionado de forma correta no códon de iniciação da tradução, AUG (Prévôt *et al*., 2003).

4.2 Iniciação da Tradução

A tradução do mRNA em proteínas se inicia após a montagem do ribossomo 80S através da interação entre tRNAi, mRNA e as subunidades 40S e 60S. O complexo processo de iniciação que leva a formação do ribossomo 80S consiste em vários estágios relacionados que são mediados pelos fatores de iniciação eucarióticos (eIFs) (Pestova, *et al*., 2001).

O início da síntese de proteínas tem como componentes essenciais ácidos ribonucléicos (mRNA, tRNA) a subunidade menor ribossomal 40S e fatores de iniciação da tradução (eIFs). Em mamíferos são descritos seis fatores: eIF1, eIF2, eIF3, eIF4, eIF5 e eIF6. Este complexo processo bioquímico tem como cenário o ribossomo. Na iniciação da tradução ocorre formação do ribossomo 80S junto ao mRNA, pela união das subunidades 40S e 60S, no correto sítio de inicio da tradução por ação do eIFs.

As principais etapas da iniciação da tradução e os fatores associados são:

1- Associação do complexo eIF4F (formado pelas subunidades eIF4E, eIF4A e eIF4G) ao cap presente na extremidade 5' do mRNA eucariótico.

2- Formação do complexo ternário entre o tRNAi, o fator eIF2 e a molécula de GTP (eIF2 / GTP / tRNAi). A associação destes e outros eIFs, como o eIF3, a subunidade ribossomal 40S forma o complexo de pré-iniciação 43S.

3- Recrutamento do complexo de pré-iniciação 43S ao cap dos mRNA por ação das subunidade eIF4G e eIF4E do eIF4F e do fator eIF3. Em alguns mRNAs celulares e virais, o ribossomo se liga diretamente a segmentos internos no mRNA (IRES – *Internal Ribosome Entry Segment*), também através do eIF4G.

4- Movimento do complexo ribossomal ao longo da região 5' não traduzida do mRNA (5' UTR), a partir do cap até o códon de iniciação da tradução. Nessa etapa atuam a subunidade eIF4A do complexo eIF4F e o fator eIF4B. No complexo resultante, 48S, o códon de iniciação tem suas bases pareadas com as bases do anticódon do tRNAi.

5- Liberação dos fatores de iniciação associados ao complexo 48S (desencadeados pela hidrólise de GTP, promovida pela ação do eIF5 e eIF5B liberando eIF2-GDP) e recrutamento da subunidade 60S para formar o ribossomo 80S, deixando o tRNAi no sítio P ribossomal.

6- Regeneração do eIF2-GTP em reação catalisada por eIF2B. O eIF2-GTP se associará novamente ao tRNAi, recompondo o complexo ternário (eIF2 / GTP / tRNAi).

O principal papel do eIF4F é promover a ligação do ribossomo ao mRNA via interação cap-eIF4E-eIF4G-eIF3. Além disso, o fator eIF4F interage com a região 3' do mRNA via proteína de ligação a cauda poli A (PABP) e este evento promove a circularização do mRNA (Wells *et al*., 1998).

4.3 Complexo eIF4F e proteínas relacionadas

4.3.1 eIF4E ou proteína de ligação ao cap

A proteína eIF4E (24 kDa) é o fator de iniciação responsável pelo reconhecimento do cap presente nos mRNAs eucarióticos (Prévôt *et al*., 2003). A forte ligação entre eIF4E-cap ocorre pela interação entre dois aminoácidos triptofanos altamente conservados e o radical 7-metil guanosina (Marcotrigiano *et al*., 1997; Matsuo *et al*., 1997). Dentro do complexo eIF4F, eIF4E interage com eIF4G e media a tradução cap dependente por promover a formação do complexo 48S no terminal 5' do mRNA. O fator eIF4E pode ser isolado sozinho ou fazendo parte do complexo eIF4F. Curiosamente, o complexo eIF4F exibe uma afinidade dez vezes maior pelo cap em relação ao eIF4E sozinho (revisto por Prévôt *et al*., 2003). Experimentos *in vitro* utilizando proteínas recombinantes mostram que o eIF4GI aumenta a afinidade do eIF4E pelo cap, e isso poderia estar relacionado ao fato de que eIF4GI poderia se ligar diretamente ao mRNA mensageiro via motivos de ligação a RNA (RRMs) (Goyer et al., 1993).

4.3.2 eIF4A

A proteína eIF4A tem sido caracterizada em diferentes organismos incluindo plantas, mamíferos e leveduras. Provavelmente o fator de iniciação da tradução eIF4A é mais abundante em uma célula eucariótica (von der Haar & McCarthy, 2002). Ele pode ser isolado sozinho ou fazendo parte do complexo eIF4F. Os genes para RNA helicase eIF4A são conservados entre espécies distintas (similaridade acima de 60% entre leveduras e mamíferos). Todavia, o eIF4A de mamíferos não é capaz de substituir o homólogo em levedura, ao contrário de outros fatores de iniciação da tradução bem menos conservados, como o eIF4E (Sonenberg *et al*., 2000).

O eIF4A é considerado um protótipo da família de proteínas DEAD-box, um grupo de RNA helicases caracterizado por apresentar oito motivos, dos quais o mais conservado é o domínio V conhecido como DEAD (L-D-E-A-D-X-X-L: Leu-Asp-Glu-Ala-Asp-X-X-Leu). As

RNAs helicases são encontradas de organismos mais simples (procariotos) a organismos mais estruturados (eucariotos superiores). Elas estão envolvidas em vários processos celulares que incluem RNA tais como: transcrição, *splicing*, *turnover*, editoramento, transporte e tradução (Lüking *et al*., 1998).

Como mecanismo de ação é postulado que o eIF4A liga-se a ATP e tem a capacidade de remover estruturas secundárias que poderiam dificultar o deslocamento do ribossomo ao longo do mRNA até o códon de iniciação da tradução (Gingras *et al*., 1999; Hershey & Merrick, 2000). O modelo mais aceito mostra que, inicialmente, o eIF4A (no contexto do eIF4F) se associa fracamente com o ATP. Essa interação modifica a conformação do eIF4A tornando-o capaz de interagir com o mRNA, aumentando sua afinidade por ATP e hidrólise deste. A hidrólise de ATP resulta em um aumento da afinidade do eIF4A pelo mRNA, e após a ligação do eIF4B tem inicio o processo de translocação da subunidade ribossomal menor ao longo da 5'-UTR do mRNA. Acreditava-se que o eIF4A tivesse afinidade por RNA fita simples e o eIF4B por fita simples e dupla. Com base nestes fatos postulava-se que o papel do eIF4B seria associar o eIF4A a RNAs dupla fita (Lüking *et al*., 1998). Porém, evidências recentes mostram de forma surpreendente que o eIF4A é capaz de remover curtos segmentos de dupla fita de RNA na ausência do eIF4B e sem a necessidade de extremidades de RNA fita simples (Rogers *et al*., 1999, 2000). Assim, o papel do eIF4B na iniciação da tradução de mamíferos e leveduras, bem como o mecanismo de ação das helicases, permanece uma grande incógnita.

4.3.3 eIF4G

O fator eIF4G é uma proteína de alto peso molecular cujo papel crucial é coordenar a formação do complexo de iniciação da tradução. Nessa tarefa ela interage, via domínios presentes em sua estrutura, com outros eIFs: eIF3, eIF4A, eIF4E, PABP entre outros.

Genes para o fator eIF4G tem sido isolados e clonados em diferentes organismos incluindo leveduras (Goyer *et al*., 1993), *Drosophila* (Zapata *et al*., 1994), plantas (Browning 1996) e mamíferos (Yan *et al*., 1992; Gradi *et al*., 1998; Imataka *et al*., 1998). Em plantas, leveduras e mamíferos existem duas isoformas de eIF4G. No homem elas são denominados de eIF4GI e eIF4GII. Estas isoformas exibem discreta homologia entre espécies diferentes, mas mostram semelhantes atividades bioquímicas. Todavia, algumas diferenças foram observadas em leveduras (Gallie & Browning, 2001; Tarun *et al* ., 1997).

Inicialmente, o eIF4GI de mamíferos foi denominado de p220, devido a co-migração de quatro proteínas diferentes em SDS-PAGE, onde a maior delas correspondia aproximadamente ao marcador molecular de 220 kDa (Tahara *et al*., 1981). Porém, a clonagem de seu gene revelou a presença de uma seqüência de leitura (ORF) que codificaria para uma proteína de 154 kDa (Yan *et al*., 1992). Este gene está presente em cópia única (Yan & Rhoads, 1995). Com a clonagem do gene para um segundo homólogo, o eIF4GII, e comparando-se sua seqüência com a do eIF4GI suspeitou-se que esta seqüência continha um íntron. A homologia entre as duas seqüências parava abruptamente em um sítio aceptor de *splicing* no gene eIF4GI (Gradi *et al*, 1998). Com base nestas evidências, utilizando-se a técnica de 5' RACE, um novo cDNA contendo uma extensão N-terminal de 156 aminoácidos foi obtido. Esta nova seqüência não só era altamente homológa a sua congênere, como também podia se associar a PABP, tendo funções na tradução dependente de poli-A (Imataka *et al*., 1998). Mais recentemente, uma seqüência de 340 nucleotídeos foi adicionada a região 5' (Bryd, 2002). A tradução *in vitro* de RNAs construídos a partir deste novo cDNA deu origem ao conjunto de isoformas resultantes de alternativos códons de iniciação da tradução AUGs. Este resultado foi confirmado por espectrometria de massa, que permitiu a caracterização de várias isoformas de eIF4GI que se diferenciam no tamanho de sua região N-terminal. Assim suspeita-se que dependendo do requerimento celular, o gene eIF4GI seja transcrito em diferentes tamanhos que serviriam a funções especializadas, evitando gasto de energia desnecessário na elaboração de uma proteína de grande tamanho (Bryd, 2002).

O eIF4G interage com vários componentes, entre eles: 1) fatores de iniciação da tradução - eIF4A, eIF4E, eIF3 (Gingras *et al*., 1999); 2) CBP80 ou proteína de ligação ao cap nuclear (Fortes *et al*., 2000); 3) proteínas virais NS1 e NSP3 dos vírus influenza e rotavírus (Aragon *et al*., 2000; Piron *et al*., 1998); 4) proteína de choque térmico hsp27 (Cuesta *et al*., 2000); 5) proteínas envolvidas no metabolismo de RNA como PABP (Imataka *et al*., 1998; Le *et al*., 1997; Tarun & Sachs, 1996), Dcp1– enzima que remove o cap de RNAs (Vilela *et al*., 2000) e Mnk – proteína quinase de eIF4E (Pyronnet *et al*., 1999). Assim, esta proteína está envolvida em vários processos celulares.

O eIF4G humano pode ser dividido em três distintas regiões de tamanhos similares com base em sua clivagem por proteases virais. Na região amino-terminal há domínios para a interação da PABP e o fator eIF4E, associando a região 5' e 3' do mRNA (Imataka *et al*., 1998). Uma região mínima para tradução dependente de cap tem sido mapeada entre os resíduos de aminoácidos 550-1090 (Morino *et al*., 2000), este fragmento inclui o sítio de ligação ao fator eIF4E e o domínio central. O domínio central inclui conservadas repetições HEAT, comuns em

proteínas envolvidas na formação de complexos multiméricos e constituídas por arranjos de cinco pares de repetições de alfa-hélices antiparalelas (Andrade & Bork, 1995). Este domínio constitui a região mais conservada evolutivamente de homólogos do eIF4G em diferentes organismos. Na região central, se ligam o eIF3 e o eIF4A (Imataka & Sonenberg 1997) e possui um motivo para ligação de RNA (Pestova *et al*., 1996). A região carboxi-terminal, ausente em não metazoários, contém um segundo sítio de ligação ao eIF4A (Imataka & Sonenberg 1997) e um sítio de ligação a uma proteína quinase que atua sobre o eIF4E, a Mnk1 (Pyronnet *et al*., 1999).

Em mamíferos existem proteínas agrupadas na família eIF4G por possuírem sequências homólogas: DAP-5, Paip-1 e DUG. A DAP-5 possui homologia com a região carboxi-terminal do eIF4GI (28% de identidade), assim como com o domínio central do eIF4G (39% de identidade). Dessa forma, a DAP-5 pode interagir com os fatores eIF4A, eIF3, mas perde a capacidade de interação com o eIF4E e PABP (Levy-Strumpf *et al.*, 1997; Gingras *et al.*, 1999). Experimentos de super-expressão da DAP-5 em cultura sugerem que a mesma atue como um repressor da tradução, todavia o papel desta proteína ainda não é plenamente conhecido (Imataka *et al*., 1997; Yamanaka *et al*., 1997). Outras duas proteínas possuem uma região com homologia ao domínio central do eIF4GI e ambas podem se ligar ao fator eIF4A. Outras duas proteínas possuem homologia ao domínios central e carboxi-terminal respectivamente do eIF4GI, Paip-1 (*Poly A binding protein Interacting Protein-1*) e DUG (Death-upregulated gene). Experimentos evidenciam que a Paip-1 funciona como um coativador traducional. Entretanto, este papel não é claro, uma vez que esta proteína possui domínios de interação ao eIF4A, PABP mas não possui domínio de ligação ao eIF4E (Craig *et al.*, 1998). Já DUG é descrita como uma proteína capaz de regular a apoptose em células submetidas a diferentes condições de estresse (Goke *et al*., 2002).

4.3.3.1 Sítio de ligação ao eIF4E

Este domínio foi mapeado a partir da observação que o eIF4E permanecia associado a região N-terminal do eIF4G após clivagem com proteases virais (Lamphear *et al*., 1995) e confirmado através da interação de eIF4GI mutantes e eIF4E em ensaios de imunoprecipitação e *far western blot* (Mader *et al*., 1995). Este trabalho conduziu a caracterização de uma região de 49 aminoácidos que está presente em mamíferos quanto em leveduras. Dentro desta região, um pequeno motivo YXXXXLΦ (onde Φ usualmente é uma leucina, mas pode também ser uma metionina ou uma fenilalanina), é altamente conservado nas proteínas eIF4G de levedura, *Drosophila*, plantas e mamíferos. Este motivo também está presente em uma família de proteínas

que competem com o eIF4GI pela interação com o eIF4E, as proteínas ligantes de eIF4E (4E-BPs). As 4E-BPs atuam como repressores da tradução (Mader *et al*., 1995). Estudos estruturais deste domínio mostram que o eIF4GI se associa à superfície dorsal convexa do eIF4E. A estrutura deste domínio mostra que a Tirosina 73 no eIF4E interage com o peptídeo presente no eIF4GI e que mutações nesta tirosina abolem a interação eIF4G/eIF4E no modelo humano (Marcotrigiano *et al*., 1999; Pyronnet *et al*., 1999) e em leveduras (Ptsushkina *et al*., 1998).

4.3.3.2 Sítio de ligação ao eIF4A

No fator eIF4GI existe dois sítios de interação para o fator eIF4A, o primeiro domínio, presente na região central, tem sido mapeado entre os aminoácidos 672-970 (Morino *et al*., 2000), 672-876 (Korneeva *et al*., 2000, 2001) e 722-949 (Lomakin *et al*., 2000). Já o segundo domínio de interação ao eIF4A, região carboxi-terminal, foi mapeado no eIF4GI entre os resíduos 1201-1411 (Morino *et al*., 2000).

Uma mutação pontual no eIF4GI que elimina a ligação do eIF4A ao domínio central impede a tradução, enquanto o eIF4GI com uma mutação semelhante no sítio no domínio carboxi-terminal diminui a eficiência de tradução em seis vezes frente ao eIF4GI selvagem (Morino *et al*., 2000). Assim, o domínio carboxi-terminal (ausente em leveduras) não é absolutamente requerido, mas tem um papel modulatório. Já o domínio central é critico para a formação do complexo traducional (Morino *et al*., 2000; Prevot et al., 2003).

Dois modelos tentam explicar como o fator eIF4A interage com o fator eIF4GI, devido ao fato de existir dois domínios de interação para o eIF4A. Num primeiro modelo é proposto que a interação ocorra na proporção de duas moléculas de eIF4A para uma molécula de eIF4GI (Korneeva *et al*., 2001). Já um segundo modelo sugere uma proporcionalidade de uma molécula de eIF4A para uma molécula de eIF4G (Li *et al*., 2001). Em leveduras, o complexo eIF4F purificado não contem o eIF4A, enquanto o eIF4F de mamíferos purificado possui esta proteína. Isso pode acontecer porque o procedimento de purificação deve abolir a fraca ligação em leveduras (apenas um domínio de interação eIF4A), porém não é capaz de desfazer a ligação entre eIF4A / eIF4G em mamíferos (dois domínios para eIF4A).

4.3.3.3 Sítio de ligação ao eIF3 , eIF1, eIF5 e PABP

O domínio de interação ao eIF3 está presente entre os aminoácidos 975-1065 do eIF4GI na região central. Esta região não se sobrepõem ao domínio de interação para o fator eIF4A.

Recentemente, foi visto em leveduras que o eIF4G interage com o fator eIF1 e eIF5 *in vivo e in vitro* através do domínio HEAT (He *et al*., 2003). A interação com o eIF4G parece ser mutuamente exclusiva para os duetos eIF1/eIF5 e eIF4A/eIF1, o que poderia ser explicado pela sobreposição de domínios para as respectivas proteínas (He *et al*., 2003). O eIF1 é conhecido por ser um monitor da interação códon-mRNA/anticódon-Met-tRNA. Simultâneas interações do eIF1 com o eIF4G e outros fatores são importantes na específica associação Met-tRNA, parte do complexo 48S, ao correto códon de iniciação de tradução AUG (He *et al*., 2003).

No fator eIF4G o domínio responsável pela interação com a proteína de ligação a cauda poli-A (PABP) foi mapeado na região N-terminal, precisamente entre os aminoácidos 132-160.

4.3.4 eIF3

O fator de iniciação da tradução eIF3 é o maior dos fatores de iniciação encontrado em mamíferos, consistindo de 11 subunidades cujo peso molecular varia entre 35 e 170 kDa.

São apontadas três funções fisiológicas importantes do eIF3 no processo de tradução:

1) Estabilização do complexo ternário (eIF2/GTP/tRNAi) por associação física entre o eIF3 e a subunidade menor
2) Ligação do mRNA ao ribossomo
3) Evitar formação do ribossomo 80S

O complexo eIF4F, via eIF4G, e o fator eIF3 promovem a associação entre a subunidade 40S do ribossomo e o mRNA. Interessantemente, experimentos demonstraram que o eIF4A e o eIF3 incrementam mutuamente suas afinidades pelo eIF4GI (Korneeva *et al*., 2000).

4.3.5 PABP

A PABP é a proteína que se associa a cauda poli-A presente na extremidade 3' dos mRNAs. Todas as PABPs caracterizadas até o momento podem ser divididas em uma porção N-terminal constituída de 4 domínios de ligação ao RNA altamente conservados (RRMs), uma porção central pouco conservada e uma extremidade C-terminal contendo um domínio de interação proteína-proteína também conservado (Bates *et al.*, 2000). A PABP é envolvida em processos como estabilidade de mRNA, processamento da extremidade 3' e no processo de tradução. A interação entre as regiões 5' e 3' do mRNA foi inicialmente descrita em leveduras,

envolvendo a associação PABP-eIF4G (Tarun & Sachs, 1996), que propunha um modelo de tradução onde o RNA mensageiro seria traduzido circularizado. Ensaios de reconstituição de sistemas de tradução *in vitro*, utilizando eIF4E/eIF4G/PABP recombinantes, demonstraram a circularização do RNA mensageiro em células eucarióticas através de microscopia de varredura (Wells *et al*., 1998).

4.4 Tradução em tripanosomatídeos

Poucos trabalhos vêm se detendo na caracterização de fatores protéicos necessários a tradução em tripanosomatídeos. O conhecimento obtido é inferido indiretamente a partir de dados provenientes da comparação de proteínas destes organismos com seus homólogos encontrados em outros eucariotos.

Assim, entre os fatores caracterizados até o momento incluem subunidades dos fatores de elongação tais como eEF1 e eEF2 (Nakamura, 1996). Homólogos da PABP de *T. cruzi*, *T. brucei* e *L. major* também já foram caracterizados (Batista *et al*, 1994; Hotchkiss *et al*, 1999 e Bates *et al*, 2000). O fator eIF4A, componente do complexo eIF4F também foi identificado em *L. major* (Skeiky *et al.*, 1998) e em *L. braziliensis*, como um antígeno que induz a produção de IL-12 e resposta tipo Th1 em humanos (Skeiky *et al.*, 1995). Entretanto, muito pouco é conhecido com relação ao papel destes fatores na tradução.

4.5 Complexo eIF4F em *Leismania major*

Tentando compreender os processos básicos da expressão gênica em tripanosomatídeos bem como estabelecer semelhanças e diferenças no processo de iniciação da tradução entre estes e eucariotos mais complexos, o grupo de biossíntese protéica em tripanosomatídeos do Departamento de Microbiologia do Centro de Pesquisas Aggeu Magalhães resolveu estudar o complexo eIF4F nestes protozoários. Inicialmente, tentou-se isolar, sem muito êxito, a partir de coluna de afinidade contendo o cap o fator eIF4E de tripanosomatídeos. Com o advento dos projetos de seqüenciamento do genoma de tripanosomatídeos, resolvemos adotar o modelo da *L.major* em função da disponibilidade de seqüências de DNA geradas pelo processo de seqüenciamento do genoma deste parasito conduzido pelo Sanger Centre (www.sanger.ac.uk). Com base em análises de bioinformática utilizando seqüências descritas em vertebrados e leveduras do complexo eIF4F, foram encontradas seqüências homólogas ao eIF4E, eIF4A e ao

eIF4G em *L.major* (as seqüências encontradas em *L.major* foram denominadas a partir do acréscimo do prefixo Lm, por exemplo LmeIF4G).

Em um primeiro momento, deu-se início a caracterização funcional de dois homólogos ao fator eIF4E (LmeIF4E1, LmeIF4E2) (Dhalia, 2001). Posteriormente incluiu-se neste estudo um terceiro homólogo de eIF4E identificado (LmeIF4E3) e o homólogo ao fator eIF4A já descrito na literatura (LmeIF4A1 – Skeiky, 1995). Dando prosseguimento ao processo de elucidação do complexo eIF4F em *L.major*, este trabalho se deteve na caracterização de três homólogos do eIF4G (LmeIF4G1, LmeIF4G2 e LmeIF4G3) e uma segunda seqüência homóloga do eIF4A (LmeIF4A2). Com a caracterização funcional de todos os homólogos do complexo eIF4F em *L.major*, espera-se compreender os eventos que conduzem a iniciação da tradução em tripanosomatídeos, algo que deve ser muito importante para esta classe de organismos devido à ausência de importantes pontos de controles da expressão gênica como a regulação transcricional.

IV. REFERÊNCIAS BIBLIOGRÁFICAS

Andrade MA and Bork P (1995) HEAT repeats in the Huntington's disease protein. Nat Genet. 11(2):115-6.

Aragon T, de la Luna S, Novoa I, Carrasco L, Ortin J and Nieto A (2000) Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus. Mol Cell Biol. 20(17):6259-68.

Bangs JD, Crain PF, Hashizume T, McCloskey JA and Boothroyd JC (1992) Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. J. Biol. Chem. 267:9805-15.

Bates EJ, Knuepfer E and Smith DF (2000) Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites. Nucleic. Acids. Res. 28:1211-1220.

Batista JA, Teixeira SM, Donelson JE, Kirchhoff LV and de Sa CM (1994) Characterization of a Trypanosoma cruzi poly(A)-binding protein and its genes. Mol. Biochem. Parasitol. 67:301-12.

Bektesh S, Van Doren K and Hirsh D (1988) Presence of the *Caenorhabditis elegans* spliced leader on different mRNAs and in different genera of nematodes. Genes Dev. *2*: 1277-1283.

Blumenthal T, Evans D, Link CD, Guffanti A, Lawson D, Thierry-Mieg J, Thierry-Mieg D, Chiu WL, Duke K, Kiraly M and Kim SK (2002) A global analysis of Caenorhabditis elegans operons. Nature 20;417(6891):851-4.

Brehm K, Jensen K, Frosch M. (2000) mRNA trans-splicing in the human parasitic cestode Echinococcus multilocularis.J Biol Chem. 8;275(49):38311-8

Browning KS (1996) The plant translational apparatus. Plant Mol Biol. 32(1-2):107-44.

Bruzik JP, Van Doren K, Hirsh D and Steitz JA (1988) Trans splicing involves a novel form of small nuclear ribonucleoprotein particles. Nature. 335(6190):559-62.

Byrd MP, Zamora M and Lloyd RE (2002) Generation of multiple isoforms of eukaryotic translation initiation factor 4GI by use of alternate translation initiation codons. Mol Cell Biol. 22(13):4499-511.

Campbell DA, Thomas S, Sturm NR (2003) Transcription in kinetoplastid protozoa: why be normal? Microbes Infect. 5(13):1231-40.

Clayton CE (2002) Life without transcriptional control? From fly to man and back again. EMBO J. 21:1881-1888.

Craig AW, Haghighat A, Yu AT and Sonenberg N (1998) Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. Nature. 2;392(6675):520-3.

Cross M, Gunzl A, Palfi Z, Bindereif A (1991) Analysis of small nuclear ribonucleoproteins (RNPs) in *Trypanosoma brucei*: structural organization and protein components of the spliced leader RNP. Mol. Cell. Biol. 11(11):5516-26.

Cuesta R, Laroia G and Schneider RJ (2000) Chaperone hsp27 inhibits translation during heat shock by binding eIF4G and facilitating dissociation of cap-initiation complexes. Genes Dev. 15;14(12):1460-70.

Dhalia R. Caracterização dos fatores protéicos envolvidos no reconhecimento do RNA mensageiro nos tripanosomatídeos. Dissertação de Mestrado, Pós-Graduação em Biologia Celular e Molecular, IOC-FIOCRUZ, 2001.

Earnshaw DL, Beebee TJC and Gutteridge WE (1987) Demonstration of RNA polymerase multiplicity in Trypanosoma brucei. Characterization and purification of alpha-amanitin-resistant and -sensitive enzymes. Biochem J. 1;241(3):649-55.

Flouriot G, Brand H, Seraphin B, Gannon F (2002) Natural trans-spliced mRNAs are generated from the human estrogen receptor-alpha (hER alpha) gene. J Biol Chem. 19;277(29):26244-51.

Fortes P, Inada T, Preiss T, Hentze MW, Mattaj IW and Sachs AB (2000) The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol Cell. 6(1):191-6.

Frantz C, Ebel C, Paulus F, Imbault P. Characterization of trans-splicing in Euglenoids. Curr Genet 37(6):349-55. 2000

FUNASA - Fundação Nacional De Saúde (2002) Guia de Vigilância Epidemiológica. Vol II. pp 503-539

Gilinger G & Bellofatto V (2001) Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Res. 1;29(7):1556-64.

Gingras AC, Raught B and Sonenberg N (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu. Rev. Biochem. 68:913-63.

Goke A, Goke R, Knolle A, Trusheim H, Schmidt H, Wilmen A, Carmody R, Goke B and Chen YH (2002). DUG is a novel homologue of translation initiation factor 4G that binds eIF4A. Biochem Biophys Res Commun. 13;297(1):78-82.

Goyer C, Altmann M, Lee HS, Blanc A, Deshmukh M, Woolford JL Jr, Trachsel H, 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(8):4860-74.

Gradi A, Imataka H, Svitkin YV, Rom E, Raught B, Morino S and Sonenberg N (1998) A novel functional human eukaryotic translation initiation factor 4G. Mol Cell Biol 18(1):334-42.

Harris KA Jr., Crothers DM, and Ullu E (1995) In vivo structural analysis of spliced leader RNAs in *Trypanosoma brucei* and *Leptomonas collosoma*: a flexible structure that is independent of cap4 methylations. RNA. *1*, 351-362.

He H, von der Haar T, Hinnebusch AG, McCarthy JE and Asano K (2003) The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. Mol Cell Biol. 23(15):5431-45.

Hotchkiss TL, Nerantzakis GE, Dills SC, Shang L and Read LK (1999) Trypanosoma brucei poly(A) binding protein I cDNA cloning, expression, and binding to 5 untranslated region sequence elements. Mol. Biochem. Parasitol. 98, 117-29.

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(12):6940-7.

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,:7480-9.

Joshi B, Yan R and Rhoads RE (1994) In vitro synthesis of human protein synthesis initiation factor 4 gamma and its localization on 43 and 48 S initiation complexes. J. Biol. Chem. 269:2048-55.

Kooter JM and Borst P (1984) Alpha-amanitin-insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. Nucleic Acids Research. 21;12(24):9457-72.

Korneeva NL, Lamphear BJ, Hennigan FL and Rhoads RE (2000) Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. J. Biol. Chem. 275: 41369-41376.

Korneeva NL, Lamphear BJ, Hennigan FL, Merrick WC and Rhoads RE (2001) Characterization of the two eIF4A-binding sites on human eIF4G-1. J Biol Chem. 26;276(4):2872-9.

Lamphear BJ, Kirchweger R, Skern T and Rhoads RE (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. 15;270(37):21975-83.

Leandro C & Campino L (2003) Leishmaniasis: efflux pumps and chemoresistance. Int J Antimicrob Agents.22(3):352-7.

LeBowitz JH, Smith HQ, Rusche L, and Beverley SM (1993) Coupling of poly(A) site selection and trans-splicing in *Leishmania*. Genes Dev. *7*, 996-1007..

Lee MG and Van der Ploeg LH (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annu. Rev. Microbiol. 51:463-89.

Levy-Strumpf N, Deiss LP, Berissi H and Kimchi A. 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. 1997 Mar;17(3):1615-25.

Li Q, Imataka H, Morino S, Rogers GW Jr, Richter-Cook NJ, Merrick WC and Sonenberg N (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. Mol. Cell Biol. 19: 7336-46.

Li W, Belsham GJ and Proud CG (2001). Eukaryotic initiation factors 4A (eIF4A) and 4G (eIF4G) mutually interact in a 1:1 ratio in vivo. J Biol Chem. 3;276(31):29111-5.

Linder P (2003) Yeast RNA helicases of the DEAD-box family involved in translation initiation. Biology of the Cell 95:157-167.

Lomakin IB, Hellen CUT and Pestova T (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. Mol Cell Biol.20(16):6019-29.

Lu J, Aoki H and Ganoza MC (1999). Molecular characterization of a prokaryotic translation factor homologous to the eukaryotic initiation factor eIF4A. Int. J. Biochem. Cell Biol. 31, 215-29.

Lüking A, Stahl U and Schmidt U (1998) The protein family of RNA helicases. Critical Ver. Biochem Mol Biol. 33(4), 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(9):4990-7.

Mair G, Shi H, Li H, Djikeng A, Aviles HO, Bishop JR, Falcone FH, Gavrilescu C, Montgomery JL, Santori MI, Stern LS, Wang Z, Ullu E and Tschudi C (2000) A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. RNA. 6(2): 163-9.

Marcotrigiano J, Gingras AC, Sonenberg N and Burley SK (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 89, 951-61.

Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CU and Burley SK (2001) A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell. 7(1):193-203.

Maroney PA, Denker JA, Darzynkiewicz E, Laneve R and Nilsen TW (1995) Most mRNAs in the nematode Ascaris lumbricoides are trans-spliced: a role for spliced leader addition in translational efficiency. RNA. 1:714-23.

Martinez-Calvillo S, Yan S, Nguyen D, Fox M, Stuart K and Myler PJ (2003) Transcription of Leishmania major Friedlin chromosome 1 initiates in both directions within a single region. Mol Cell. 11(5):1291-9.

Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N and Wagner G (1997) Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nat. Struct. Biol. 4, 717-24.

Morino S, Hazama H, Ozaki M, Teraoka Y, Shibata S, Doi M, Ueda H, Ishida T and Uesugi S (1996) Analysis of the mRNA cap-binding ability of human eukaryotic initiation factor-4E by use of recombinant wild-type and mutant forms. Eur. J. Biochem. 239, 597-601.

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.

Muhich ML and Boothroyd JC (1988) Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol. Cell Biol. 8, 3837-46.

Nakamura Y, Hashimoto T, Kamaishi T, Adachi J, Nakamura F, Okamoto K and Hasegawa M (1996) Phylogenetic position of kinetoplastid protozoa inferred from the protein phylogenies of elongation factors 1alpha and 2. J. Biochem. (Tokyo) 119, 70-9.

Nilsen TW (1995). trans-splicing: an update. Mol. Biochem. Parasitol. 73, 1-6.

Pascolo E, Blonski C. Shire D and Toulme J.J (1993) Antisense effect of oligodeoxynucleotides complementary to the mini-exon sequence of the protozoan parasite Leishmania amazonensis. Biochimie 75, 43-7.

Pestova TV, Hellen CU and Shatsky IN (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol. Cell Biol. 16, 6859-69.

Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI and Hellen CU (2001) Molecular mechanisms of translation initiation in eukaryotes PNAS. Proc Natl Acad Sci U S A. 19;98(13):7029-36.

Piron M, Vende P, Cohen J and Poncet D (1998) Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F. EMBO J. 1;17(19):5811-21.

Ponting CP (2000) Novel eIF4G domain homologues linking mRNA translation with nonsense-mediated mRNA decay. TIBS 25 423-426

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

Ptushkina M, von der Haar T, Vasilescu S, Frank R, Birkenhager R and McCarthy JE (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, 4798-808.

Pyronnet S, Imataka H, Gingras AC, Fukunaga R, Hunter T and Sonenberg N (1999) Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. EMBO J. 4;18(1):270-9.

Ramazeilles C, Mishra RK, Moreau S, Pascolo E and Toulme JJ (1994).Antisense phosphorothioate oligonucleotides: selective killing of the intracellular parasite Leishmania amazonensis. Proc. Natl. Acad. Sci. U. S. A. 91, 7859-63.

Rau, M, O, T , Morley SJ, Pain VM (1996) A reevaluation of the cap binding protein eIF4E, as a rate limiting factor for initiation of translation in reticulocyte lysate JBC 271, 8983-8990

Raught B, Gingras, AC (1999) eIF4E activity is regulated at multiple levels. International Journal of Biochemistry & Cell Biology 31, 43-57.

Rogers GW Jr, Richter NJ and Merrick WC (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. J Biol Chem. 30;274(18):12236-44.

Rogers GW Jr, Richter NJ, Iima WF, Merrick WC (2001) Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. J Biol Chem. 17;276(33):30914-22

Saxena A, Worthey EA, Yan S, Leland A, Stuart KD and Myler PJ (2003) Evaluation of differential gene expression in Leishmania major Friedlin procyclics and metacyclics using DNA microarray analysis. Mol. Biochem. Parasitol. 129, 103-114.

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

Schnare MN, Gray MW (1999) A candidate U1 small nuclear RNA for trypanosomatid protozoa. J. Biol. Chem. 274(34):23691-4.

Skeiky YA, Guderian JA, Benson DR, Bacelar O, Carvalho EM, Kubin M, Badaro R, Trinchieri G and Reed SG (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, 1527-37.

Skeiky YA, Kennedy M, Kaufman D, Borges MM, Guderian JA, Scholler JK, Ovendale PJ, Picha KS, Morrissey PJ, Grabstein KH, Campos-Neto A and Reed SG (1998). LeIF: a recombinant Leishmania protein that induces an IL-12-mediated Th1 cytokine profile. J. Immunol. 161, 6171-9.

Sonenberg N, Hershey JWB and Mathews MB (2000) Translational Control of Gene Expression. The Pathway and Mechanism of Initiation of Protein Synthesis. Cold Spring Harbor Laboratory Press. pp 33-88

Sonenberg N and Gingras AC (1998) The mRNA 5'cap binding protein eIF4E and control of cell growth. Current Opinio in Cell Biology 10: 268-275

Sonenberg N and Dever TE (2003) Eukaryotic translation initiation factors and regulators. Current Opinion in Structural Biology 13, 56-63

Stiles JK, Hicock PI, Shah PH and Meade JC (1999). Genomic organization, transcription, splicing and gene regulation in *Leishmania*. Ann. Trop. Med. Parasitol. 93(8):781-807.

Story, RM, Li, H and Abelson, JN (2001) Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. PNAS 98 (4), 1465-1470

Sturm NR and Campbell DA (1999) The role of intron structures in trans-splicing and cap 4 formation for the Leishmania spliced leader RNA. J Biol Chem. 2;274(27):19361-7.

Sutton RE & Boothroyd JC (1986) Evidence for trans splicing in trypanosomes. Cell. 21;47(4):527-35.

Tarun SZ Jr and Sachs AB (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J. 16;15(24):7168-77.

Tarun SZ Jr and Sachs AB (1997) Binding of eukaryotic translation initiation factor 4E (eIF4E) to eIF4G represses translation of uncapped mRNA. Mol Cell Biol. 17(12):6876-86.

Vilela C, Velasco C, Ptushkina M and McCarthy JE (2000) The eukaryotic mRNA decapping protein Dcp1 interacts physically and functionally with the eIF4F translation initiation complex. EMBO J. 15;19(16):4372-82.

von der Haar T and McCarthy JE (2002) Intracellular translation initiation factor levels in Saccharomyces cerevisiae and their role in cap-complex function. Mol Microbiol. 46(2):531-44.

Ullu E, Matthews KR, and Tschudi C (1993) Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. Mol. Cell. Biol. *13*, 720-725.

Zapata JM, Martinez MA and Sierra JM (1994) Purification and characterization of eukaryotic polypeptide chain initiation factor 4F from Drosophila melanogaster embryos. J Biol Chem. Jul 8;269(27):18047-52.

Wells SE, Hillner PE, Vale RD and Sachs AB (1998) Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell. 2(1):135-40.

WHO - World Health Organization 2002. Leishmaniasis
http://www.who.int/tdr/publications/tdrnews/default.htm

Yamanaka S, Poksay KS, Arnold KS and Innerarity TL (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. 1;11(3):321-33.

Yan R, Rychlik W, Etchison D and Rhoads RE (1992) Amino acid sequence of the human protein synthesis initiation factor eIF-4 gamma. J Biol Chem 267(32):23226-31.

Os resultados obtidos durante o trabalho experimental dessa dissertação foram utilizados na confecção de um manuscrito entitulado **"Translation Initiation In *Leishmania Major*: Characterisation Of Multiple eIF4F Subunit Homologues."** Este manuscrito, de múltiplos autores, deverá ser submetido ao Molecular and Biochemical Parasitology e se encontra a seguir, já na sua forma semi-definitiva em inglês. Tendo em vista a necessidade de se avaliar individualmente o trabalho experimental realizado pelo autor desta dissertação, decidimos discriminar os experimentos realizados exclusivamente pelo mesmo (exceto quando mencionado), durante o período de vigência do Mestrado. Este trabalho está descrito detalhadamente no anexo dividido sob a forma de Metodologia, Resultados e Discussão.

V . ARTIGO CIENTÍFICO

April 2004

TRANSLATION INITIATION IN *Leishmania major*: CHARACTERISATION OF MULTIPLE eIF4F SUBUNIT HOMOLOGUES.

Rafael Dhalia[#$], Christian R. S. Reis[+$], Eden R. Freire[$], Polyanna O. Rocha[$], Rodolfo Katz[$], João R. C. Muniz[@], Nancy Standart* and Osvaldo P. de Melo Neto[$]

[#] Departamento de Biologia Celular, Universidade de Brasilia, Brasilia DF Brazil 70910-900

[+]Departamento de Genética, Universidade Federal de Pernambuco, Avenida Professor Moraes Rego, s/n, Cidade Universitária Recife, PE Brazil 50732-970

* Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge UK CB2 1GA

[@] Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, São Carlos SP Brazil 13560-970

[$]Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Avenida Professor Moraes Rego s/n, Cidade Universitária, Recife PE Brazil 50670-420

Keywords: eIF4F, translation initiation, *Leishmania major*, protein-protein interaction

Corresponding author : Osvaldo P. de Melo Neto

e-mail: opmn@cpqam.fiocruz.br

Abstract

In yeast, plants and animals, initiation of protein synthesis starts with the binding of the multimeric translation initiation complex eIF4F - eIF4E, eIF4A and eIF4G – to the monomethylated cap present on the 5' end of mRNAs. eIF4E interacts directly with the cap nucleotide, while eIF4A is a highly conserved RNA helicase and eIF4G acts as a scaffold for the complex with binding sites for both eIF4E and eIF4A. To initiate protein synthesis, eIF4F binds to the mRNA and recruits the small ribosomal subunit to its 5' end. Its eIF4A subunit is then thought to melt secondary structures in the mRNA 5'UTR allowing the small ribosomal subunit to scan the mRNA and find the initiation codon. Little is known in detail of protein synthesis in the protozoan parasites belonging to the family *Trypanosomatidae*. However, the presence of the highly modified cap structure, cap4, and the spliced leader sequence on the 5' ends of all mRNAs suggests possible differences in mRNA recruitment by ribosomes. To investigate translation initiation in *Trypanosomatidae*, we searched databases for eIF4F homologues in *Leishmania major*. Here we describe the identification of several potential homologues: 4 eIF4Es (*LmeIF*4E1-4), 2 eIF4As (*LmeIF*4A1-2) and 5 eIF4Gs (*LmeIF*4G1-5). We also report the initial characterisation of several of these homologues (*LmeIF*4E1-3, *LmeIF*4A1-2, *LmeIF*4G3) by different approaches. First we sought to quantitate the expression of some of these proteins in *L. major* promastigotes through western-blots and found out that *LmeIF*4A1 is very abundant (> 10^5 molecules/cell), *LmeIF*4E2-3 and *LmeIF*4G3 are moderately abundant (10^3-10^4 molecules/cell) and *LmeIF*4E1 / *LmeIF*4A2 are rare or not detected (< 10^3 molecules/cell). Next, we investigated the cap-binding activity of the eIF4E homologues and found that only *LmeIF*4E2 was capable of binding to the 7-methyl-GTP-Sepharose resin, with an affinity similar to *Xenopus* eIF4E. Molecular modelling of the *LmeIF*4E2 structure also confirmed that it has all the features of a true cap binding protein. Pull-down assays were then used to investigate the interaction between the different eIF4A and eIF4G (*LmeIF*4G1-3 only) homologues and observed that only *LmeIF*4G3 bound specifically both to the *LmeIF*4A1 parasite homologue as well as to the human eIF4A factor. Therefore for each factor one of the *L. major* forms seems to fulfil, in part at least, the expected characteristics of a translational initiation factor. These results are consistent with roles for these factors in parasite mRNA translation.

1. Introduction

Flagellate protozoans belonging to the order Kinetoplastida, family Trypanosomatidae, are responsible for a series of important diseases, such as the Leishmaniasis, Sleeping Sickness and Chagas' Disease, which affect millions of people world-wide and put a great number of lives at risk (www.who.int/tdr). These are unique eukaryotes which are characterised by unusual processes in the formation of mature mRNAs, such as polycistronic transcription and *trans*-splicing (Muhich, 1988; Ullu, 1991; for recent reviews see Campbell *et al*., 2003; Liang *et al*., 2003). These parasites are also distinguished by the lack of known mechanisms for the transcriptional control of the expression of their protein coding genes (reviewed in Clayton, 2002). After transcription, processing of the polycistronic precursor mRNAs occurs through the *trans*-splicing mechanism which allows for the capped 39 nt long spliced leader (SL) or mini-exon sequence to be added to the 5' end of each monocistronic mature mRNA. The first 4 nt of the SL sequence are subjected to various modifications and constitute, with the 7 methyl-GTP nucleotide (cap) at its 5'end, the structure called cap4 (Bangs, 1992; Nilsen, 1995). Roles for the capped SL sequence in mRNA metabolism have been proposed, based largely on cap functions in metazoans in mRNA transport, stability and/or translation, but no firm conclusions have been reached yet. Experiments using antisense oligonucleotides complementary to the SL sequence in heterologous cell-free translation systems are suggestive of a role for this sequence in translation (Pascolo, 1993; Ramazeilles, 1994). Indirect evidence from nematodes, using message-dependent homologous protein synthesis systems, also reinforce a role in translation (Maroney, 1995). Indeed recent data suggests that either the SL sequence or the cap4 is responsible for polysome association of mRNA in *L. tarentolae* (Zheiner, 2003). Since the 5'end of most eukaryotic mRNAs is required for translation initiation (see below), it is expected that their influence would be exerted at this stage of the translation process.

In animals, yeast and plants, the highly regulated initiation stage of protein synthesis starts with the binding of the multimeric translation initiation complex eIF4F (formed by the translation initiation factors eIF4E, eIF4A and eIF4G) to the monomethylated cap present on the 5'end of the mRNAs (Gingras, 1999). With the help of other initiation factors, such as eIF3, eIF4F allows the recognition of the mRNAs by the 40S ribosomal subunit and the start of the translation process. Its interaction with the poly-A binding protein (PABP), via eIF4G, may also promote the circularization of the mRNA and enhance translation reinitiation [reviewed in (Hershey, 2000; Pestova, 2001; Sonenberg, 2003)]. The eIF4E, the cap binding protein, is a 24 kDa polypeptide responsible for specific cap recognition. It binds to both the cap and eIF4G and

its activity can be regulated by phosphorylation or through the eIF4E interacting proteins [reviewed in (Sonenberg, 1998; Raught, 1999; Scheper, 2002)]. The eIF4A is an ATP dependent RNA helicase which has been classified as a member of the DEAD box family of RNA helicases [reviewed in (Lüking, 1998; Tanner, 2001; Linder, 2003)]. During translation initiation eIF4A binds eIF4G and, with the help of another translation initiation factor eIF4B, is responsible for melting secondary structures along the mRNA 5'UTR. This unwinding of the mRNA 5'end allows the recruitment of the small ribosomal subunit which then scans the mRNA 5'UTR and finds the translation initiation codon. The large subunit eIF4G (approx. 200 kDa – reviewed in Prevot, 2003) provides the scaffold for the eIF4F complex. Through its N-terminus eIF4G binds both eIF4E [Mader, 1995] and the poly(A) binding protein (PABP) (Tarun & Sachs, 1996; Imataka *et al*., 1998), providing the link between the 5'and 3'ends of the mRNA. The eIF4G's central core region includes the conserved HEAT repeat domain (Marcotrigiano, 2001), which is responsible for the interaction with eIF4A and the binding to RNA, and also mediates the interaction with eIF3 (Korneeva, 2000; Lomakin, 2000), which recruits the 40S ribosomal subunit. Mammalian eIF4G contain a C-terminal domain with a second eIF4A binding site as well as the binding site for eIF4E kinase Mnk. Yeast eIF4G lacks this C-terminal domain (Imataka, 1997; Morino, 2000; Pyronnet, 1999).

Knowledge about the process of protein synthesis in trypanosomatid protozoans is inferred by indirect evidence such as sequence similarities between individual translation factors with homologues from higher eukaryotes. Translation factors characterised so far from these organisms include subunits of elongation factors such as eEF1 and eEF2 (Nakamura, 1996) and the PABP from *T. cruzi*, *T. brucei* and *L. major* (Batista, 1994; Hotchkiss, 1999; Bates, 2000]. The eIF4A component of the eIF4F complex has also been identified both in *L. major* (Skeiky, 1998) and in *L. braziliensis*, as an antigen that elicits IL-12 production and Th1-type response in humans (Skeiky, 1995). However, little is known regarding the role of these factors in translation.

As part of an effort to understand translation initiation in the trypanosomatids we set out to characterise the various eIF4F subunits from *L. major*. We started by using the sequences of the known vertebrate and yeast factors in BLAST searches with the *L. major* genome sequences available online. Relevant *L. major* sequences identified in these searches were amplified, cloned and the respective proteins subjected to functional analysis. Multiple potential homologues for the three eIF4F components were identified and arbitrarily named *LmeIF*4E1-4, *LmeIF*4A1-2 and *LmeIF*4G1-5. These sequences show different degrees of similarity to the vertebrate factors

but similar isoforms have also been found in *T. brucei*, which indicates that they all play conserved relevant roles in the parasites. The multiple homologues seem to vary in different aspects such as RNA binding affinity (for the eIF4Es), levels of expression and interaction with other components of the eIF4F factor. These results suggest a high degree of complexity in translation in these parasites, which may reflect an adaptation to their complex life cycle.

2. Materials and methods

2.1. Parasites

Promastigotes of *L. major*, *L. braziliensis* and *L. chagasi* as well as epimastigotes of *T. cruzi*, were cultivated in modified LIT medium (0.2% sucrose w/v; 0.3% liver broth w/v; 0.1% tryptose w/v; 0.002% haemin w/v) containing 10% fetal bovine serum and 1% ampicillin/streptomycin. Lysates were obtained from log phase haemocytometer quantified parasite pellets.

2.2 Sequence analysis and modelling.

At the start of this work, the complete sequences of *LmeIF*4E1-2, a partial sequence of *LmeIF*4E3 as well as the published *LmeIF*4A1 sequence (Skeiky, 1998) were available at the NCBI database (www.ncbi.nlm.nih.gov). The full length *LmeIF*4E3 ORF was a kind gift from Al Ivens. For the other homologues initially BLAST searches were carried out with the non-annotated draft *L. major* genome database (produced by the *L. major* Sequencing Group and available at the Sanger Center Website - www.sanger.org), using as query the sequences from the human and yeast (*Saccharomyces cerevisae*) eIF4F homologues. Selected sequences and preliminary contigs from this database were processed when necessary in order to find the appropriate full length ORF encompassing the complete protein. Later, the sequence refinements were confirmed with the availability of the Gene DB website of the Sanger Institute Pathogen Sequencing Unit (www.genedb.org). Similar procedures were performed to find *T. brucei* 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). Occasionally manual refinement of the alignments was performed. For the eIF4A sequence analysis, the alignment produced with CLUSTAL W was followed by phylogenetic and molecular evolutionary analyses conducted with the program MEGA version 2.1 (Kumar *et al*., 2001) using the Neighbour-Joining method (Saitou & Nei, 1987). Gaps were treated as missing data. Bootstrap analysis (10,000 replicates) was done by interior branch test.

For the molecular modelling of *LmeIF*4E2, first a structural alignment was performed with the GenTHREADER program (Jones, 1999) comparing its secondary structure with that of available proteins with resolved structure and the best match was the mouse eIF4E bound to 7-methyl-GDP (Marcotrigiano *et al*., 1997). The atomic coordinates from the mouse structure were submitted with the alignment results to the program MODELLER (Sali & Blundell, 1993) in order to produce the models. The best models were then validated using the programs PROCHECK (Laskowski *et al*., 1998), Verify 3D (Luthy *et al*., 1993) and WHAT IF (Vriend, 1990). Similar procedures were done to obtain the model of the *LmeIF*4G3 HEAT domain using the published structure from the human eIF4GII HEAT domain (Marcotrigiano *et al*., 2001).

2.3. PCR and cloning methods for the various L. major factors.

eIF4Es:

The *LmeIF4E1* coding sequence was amplified using primers flanked by sites for the restriction enzymes *Kpn* I/*Xba* I, for the full length sequence (5' primer - CCT GGT ACC CCG ACG CTA TGG ATC CG; 3'primer – CCC TCT AGA TCA CAC ACC TTC ATT CCT C; the restriction sites are underlined), or alternatively flanked by sites for *Bam*H I/*Xba* I but missing the 5'end coding for the first 7 amino acids of the protein (*LmeIF4E1*b; 5'primer - CCT GGA TCC CCG GCG AGT GCT GTG AC; 3'primer – same as above). For the full length ORFs from *LmeIF4E2* (5' primer – CCT GGA TCC TCG CAT TAC AGC ATG TCA G; 3'primer – CCC TCT AGA CTA AGA CGC CTC GCC GTG C) and *LmeIF*4E3 (5' primer - GTG GGA TCC GGG AAT AGA ACG ATG AAC G; 3'primer - AGG TCT AGA CTA GAA CGT GTG ATC GGG), they were also amplified using primers flanked by sites for *Bam*H I and *Xba* I. As PCR template, DNA from the cosmids L2185 or L5515 (a kind gift from Al. Ivens from the Sanger Centre) was used to amplify *LmeIF4E1* and *2* respectively. For *LmeIF4E3*, total *L. major* DNA was used as template. The various amplified fragments were then cloned into the corresponding restriction sites of the pGEM3zf+ vector (Promega). The resulting plasmids containing the full length sequences were then used as template in PCR reactions designed to add *Nco* I/*Not* I sites to the constructs *LmeIF4E1* (5'primer - TGC CAT GGA TCC GAA TAC ATG TG; 3'primer - TGG CGG CCG CCA CAC CTT CAT TCC TCA C) and *LmeIF4E3* (5'primer - TGC CAT GGA CCC GTC TGC CGC TG; 3'primer - TGG CGG CCG CGA ACG TGT GAT CGG GCG; here a N to D mutation was introduced in the second codon to accommodate the *Nco* I site) or *Afl* III/*Not* I sites to *LmeIF4E2* (5'primer - TGA CAT GTC AGC CCC GTC TTC AG; 3'primer - TGG CGG CCG CAG ACG CCT CGC CGT GCT T). The PCR fragments were then cloned into

the *Nco* I/*Not* I sites of the expression vector pET21D (Novagen). The resulting plasmids were used for the production of the [^{35}S] labelled proteins and for the expression of recombinant His-tagged proteins (tag at the C-terminus) in *E. coli*. In order to express the respective recombinant proteins as fusions with Glutathione S-transferase (GST) on their N-terminus, the fragments *LmeIF4E1*b and *LmeIF4E2* were recovered from the pGEM derived plasmids by digestion with *Bam*H I and *Sal* I and subcloned into the vector pGEX4T3 (Amersham Biosciences). For *LmeIF4E3*, it was reamplified using the T7 promoter primer (anneals 5'end to the fragment in the plasmid vector) and the 3'end primer flanked by *Not* I shown above and cloned into the *Bam*H I/*Not* I sites of the pGEX4T3 vector.

eIF4As:

To clone the *LmeIF4A1* sequence, it was amplified from total *L. major* DNA using oligonucleotides flanked by restriction sites for the enzymes *Bam*H I/*Xba* I (5'primer - GTG GGA TCC TGT CGT CCT CCA TCA TGG; 3'primer - CGG TCT AGA CTT ACT CGC CAA GGT AGG C). The PCR product was initially cloned into the same sites of the pGEM3zf+ vector and the resulting plasmid used as template in a second PCR reaction designed to add the *Nco* I and *Xho* I sites to its 5'and 3'ends respectively (5'primer - TGC CAT GGC GCA GAA TGA TAA G; 3'primer - TGC TCG AGC TCG CCA AGG TAG GCA G). The new amplified fragment was then cloned into the *Nco* I / *Xho* I sites of the vector pET21D. Alternatively another PCR reaction was performed using as template the pGEM-*LmeIF4A1* plasmid to generate *Bam*H I and *Xho* I sites flanking the *LmeIF4A1* gene (5'primer - GTG GGA TCC ATG GCG CAG AAT GAT AAG; 3'primer – same as above). The resulting fragment was cloned into the same sites of pGEX4T3. The pGEX-*LmeIF4A1* was then digested with *Bam*H I/*Xho* I and the *LmeIF4A1* fragment cloned into the same sites of the pRSETA plasmid (Invitrogen). These plasmids allowed the expression of recombinant *LmeIF*4A1 fused to a tag of 6 Histidines on its C or N-terminus or GST on the N-terminus. The second putative *L. major* eIF4A gene (*LmeIF4A2*) was amplified from total *L. major* DNA using primers flanked by sites for *Bam*H I/*Xho* I (5'primer - CTC GGA TCC ATG GAG ACC GAG CAA GTA G; 3'primer - TG CTC GAG AAG CGA AAG GTG GAG AG) and cloned into the same sites of the vectors pRSETA and pGEX4T3 to produce N-terminal His tagged and GST fusions proteins respectively.

eIF4Gs:

Fragments coding for the HEAT repeats domain from the *LmeIF*4G1-3 homologues were amplified and cloned from total *L. major* DNA as follows. For *LmeIF4G1*, the sequence coding

for amino acids 128-475 (construct *LmeIF*4G1$_{128-475}$) of the putative protein was amplified using primers flanked by sites for *Afl* III/*Not* I and cloned into the *Nco* I/*Not* I sites of the pET21D plasmid (5' Primer – GGA CAT GTT TAT GTC GGT CCG GAA GG; 3' Primer – TGG CGG CCG CGC TCA TGA TGG AGG ACT GCA G). An equivalent region from *LmeIF4G2*, coding for amino acids 387-705 (*LmeIF*4G2$_{387-705;}$ ending in DGLRAML and missing the last 22 amino acids from the HEAT domain in figure 7), was also amplified using primers flanked by *Nco* I and *Xho* I restriction sites and cloned into the same sites of the pET21D vector (5' Primer – TGC CAT GGA CGG CGG CGG GTT CAG CTC; 3' Primer – TGC TCG AGC AGC ATG GCA CGC AGG CCA TC). Likewise, for *LmeIF4G3* the sequence coding for amino acids 26-310 of the predicted protein (*LmeIF*4G3$_{26-310}$) was amplified and cloned into the *Nco* I/*Not* I sites of the pET21D plasmid (5' Primer – TGC CAT GGA TGA AGT GGT GCG ACG CC; 3' Primer – TGG CGG CCG CGG CAG TGG CGT ACT TTT TAT CC). For both *LmeIF*4G2$_{387-705}$ and *LmeIF*4G3$_{26-310}$ an AUG start codon was introduced in the constructs with the *Nco* I site. Later the full length *LmeIF4G3* (*LmeIF*4G3$_{1-636}$) as well as its first half (*LmeIF*4G3$_{1-343}$) were also amplified from total *L. major* DNA flanked by sites for *Nco* I/*Not* I (5' Primer – TGC CAT GGA GTT CAC CGT GGA GCA G; 3'Primers - TCG CGG CCG CAT TAC TTG GGG AAG and T GGC GGC CGC TGC TGC CGC AGT ACC TGT C for *LmeIF*4G3$_{1-636}$ and *LmeIF*4G3$_{1-343}$ respectively) and cloned into the same sites of pET21D. The resulting plasmids were also used as templates in new PCR reactions with a new 5'primer (GTG GGA TCC ATG GAG TTC ACC GTG GAG) and the same 3'primers so as to have both fragments flanked by *Bam*H I/*Not* I. These fragments were cloned into pGEX4T3 to have both proteins as GST fusions.

2.4. Cap binding assay.

[^{35}S] Met-labelled *L. major* eIF4E proteins were synthesised in the TnT T7 Coupled Reticulocyte Lysate System (Promega) using the pET derived plasmids (pET-*LmeIF4E1-3*) as templates. As a positive control, a *Xenopus laevis* eIFE cDNA, cloned into the pSP64TEN plasmid (a kind gift from Simon Morley), was recovered by digestion with *Hin*D III/*Bgl* II and subcloned into the *Hin*D III/*Bam*H I sites of the pGEM2 vector (Promega) under control of the T7 promoter. Labelled *Xenopus* eIF4E protein was obtained from this plasmid as described above. For the cap binding assay, 35 µl of 7-Methyl-GTP Sepharose 4B beads (Amersham Biosciences), previously equilibrated with buffer A (50 mM Hepes; 1 mM EDTA; 0.1 mM GTP; 14 mM 2-mercaptoethanol; 100 mM KCl), was mixed with 25 µl of the translation products (in 200 µl final volume – completed with buffer A) and incubated for 30 minutes in ice. The beads were then washed

3 times with 500 µl buffer A and, for the elution, 3 times with 50 µl buffer A containing 50 µM cap analogue (m^7G(5')ppp(5')G - NEB) instead of the GTP. Any labelled protein still remaining on the beads was recovered with two final washes with 2 M KCl and SDS-PAGE sample buffer respectively. Aliquots from the various washes as well as the initial translation products and the post-elution beads were mixed with SDS-PAGE sample buffer and analysed by SDS-PAGE and autoradiography.

2.5. Expression and purification of recombinant proteins.

For the expression of either His or GST-tagged recombinant proteins, the various plasmids were transformed into *E. coli* BLR or BL21 cells. The transformed bacteria were grown in LB medium and induced with IPTG. Induced cells were sedimented, ressuspended in PBS and lysed by ultra-sonication or French press. Protein purification was performed as described (de Melo Neto *et al*., 1995) with either Ni-NTA Agarose (Qiagen) or glutathione-Sepharose (Amersham Biosciences). Protein products were analysed in 15% SDS-PAGE stained with Coomassie-Blue R-250. For the quantification of the recombinant proteins serial dilutions were compared in Coomassie stained gels with serial dilutions of defined concentrations of BSA.

2.6. Antibody production and Western-blots.

Rabbit antisera were raised against *LmeIF*4A1-2, *LmeIF*4E1-2 and *LmeIF*4G3$_{26-310}$ by immunising adult New Zealand White rabbits with the His-tagged recombinant forms. For *LmeIF*4E3 (which did not express well as His fusion) the GST-fusion was first cleaved with thrombin and the fragment corresponding to the protein excised from SDS-PAGE gels and used for the immunization. The various antibodies were then affinity purified. Briefly 100 µg of the various recombinant protein were run on a SDS-PAGE, transferred to a Immobilon-P PVDF membrane (Millipore) and the membranes stained briefly with 0.2% (w/v) Ponceau S (in 1% [w/v] trichloroacetic acid) to identify the protein bands. The membrane fragments containing the proteins were then cut into pieces, washed 3x with 1 ml phosphate-buffered saline (PBS), blocked with 5% (w/v) skimmed milk in PBS-T (PBS with 0.05% Tween-20) for 1 h at room temperature prior to incubation with 500 µl of the specific antisera plus 500 µl PBS at 4^{o}C overnight. The depleted serum was removed and the membranes washed 3x again with PBS. Bound antibodies were eluted in 200 µl 0.1 M Glycine pH 2.5 at room temperature for 5 minutes, neutralised by addition of 20 µl of Tris-HCl 1M pH 8.0 and diluted with 220 µl of PBS. The purified antibodies were kept at 4^{o}C with 0.03% sodium azide and used in a dilution of 1/500. Western-blots were performed with the Immobilon-P membrane, using as second

antibody peroxidase conjugated Goat anti-rabbit IgG serum (Jackson ImmunoResearch Laboratories) diluted 1/15000. The reactions were detected by enhanced chemiluminescence (ECL).

2.7. Pull-down assays.

The Ni-NTA Agarose or glutathione-Sepharose beads used for the pull-down assays were initially equilibrated with binding buffer B (100 mM KCl, 1 mM $MgCl_2$, 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. For the GST-tagged protein experiments, the glutatione-Sepharose beads were initially saturated with 10 mg/ml of BSA, followed by 2 washes with Buffer B prior to their use in the assays. Approximately 10 µl of the equilibrated beads was then incubated with 2 µg of the His or GST tagged constructs 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 [^{35}S-Met]-labelled proteins, in Buffer B in a final volume of 200 µl, for 2 hours at 4°C. After 3 more washes, the bound proteins were eluted by addition of SDS-PAGE sample buffer and the samples run on 15% SDS-PAGE. The gels were then stained with Coomassie Blue (to visualize the recombinant proteins), and subjected to autoradiography. Labelled proteins were obtained through the linearizations of the various plasmids described previously (with *Not* I for pET-*LmeIF*4G1$_{128-475}$ and the *LmeIF*4G3 variants and with *Xho* I for pET-*LmeIF*4G2$_{387-705}$ and pRSET-*LmeIF*4A1-2), followed by transcription with T7 RNA polymerase in the presence of the cap analogue and translation with the rabbit reticulocyte lysate supplemented with [^{35}S] methionine. To obtain the labelled human eIF4A, the human eIF4A cDNA from plasmid pET(His$_6$-eIF4A) (Pestova *et al*., 1996), a kind gift from I. Ali and R. Jackson, was first recovered by digestion with *Nde* I/ *Not* I and subcloned into the same sites of the plasmid pET21A. The resulting plasmid was linearized with *Not* I and the eIF4A cDNA transcribed and translated as above. Labelled human eIF4G was obtained as described (Melo *et al*., 2003).

3. Results

3.1. Sequence analysis of four eIF4E homologues from L. major.

To characterise eIF4F homologues from *L. major*, we started by focusing on the four available *L. major* eIF4E sequences which we named as *LmeIF*4E1, *LmeIF*4E2, *LmeIF*4E3 and *LmeIF*4E4 according to the chronological order of their identification by our group and following

the proposed nomenclature for the *Leishmania* and *Trypanosoma* proteins (Clayton *et al*., 1998). The potential eIF4E sequences code for proteins with predicted molecular weights of 31.5, 24, 37.7 and 46.1 kDa respectively and with homologies varying from 41 to 45% similarity to the human eIF4E homologue (see Table I). *LmeIF*4E1 is encoded by two identical genes contained within a small duplicated region on chromosome 19 which also includes one unrelated ORF (a non-identified trans-membrane protein) and are separated by approximately 4.5 kb. *LmeIF*4E2-4 are encoded by single genes present on chromosomes 27, 28 and 30 respectively.

In order to compare conserved features in the four *Leishmania* eIF4E homologues we aligned these sequences with characterised human, yeast and plant eIF4E proteins (Figure 1). Two distinct *Caenorhabiditis elegans* homologues with contrasting cap binding affinities – IFE-1 which binds both monomethylated and trimethylated cap and IFE-3 which only binds a monomethylated cap (Jankowska-Anyszka, 1998)– were also included in the alignment. For simplicity all the numbers mentioned in Figure 1 will refer to the amino acids' position in the human eIF4E sequence which is representative of the mammalian protein (only 3 substitutions in the N-terminus distinguish the human and mouse sequences). eIF4E is characterised by eight tryptophan residues located at conserved positions along the protein (Marcotrigiano, 1997; Gingras, 1999). *LmeIF*4E1 contains all eight tryptophan residues, *LmeIF*4E2 contains six, *LmeIF*4E3 only four and *LmeIF*4E4 five. Three-dimensional structures of mouse eIF4E and its yeast homologue, both bound to 7-methyl-GDP, were solved by X-ray crystallography (Marcotrigiano, 1997), and by solution NMR spectroscopy (Matsuo, 1997), respectively. The complex is shaped like a cupped hand, with the cap analogue located in a narrow cap-binding slot on the concave side of the protein. Recognition of the 7-methylguanine moiety is mediated by base sandwich-stacking between W56 and W102, formation of three Watson-Crick-like hydrogen bonds with a side-chain carboxylate of a conserved E103 and a backbone NH of W102, and a van der Waals contact of the N(7)-methyl group with W166. The four residues making contacts with 7-methyl guanine (W56, W102, E103 and W166) are conserved among all eIF4E sequences, with the exception of *C. elegans* IFE-4 and human 4E-HP (eIF4E Homologous Protein) which contain the equivalent of Y56 [(Marcotrigiano, 1997; Rom, 1998; Keiper, 2000) and Figure 1]. All four are present in both *LmeIF*4E1 and *LmeIF*4E2, while in *LmeIF*4E3-4 the W56 is replaced by a methionine and a glutamate respectively. Of the three residues that interact with the two phosphate groups, R157 is absolutely conserved in eIF4E sequences, and the remaining two residues (R112 and K162) make either direct or water mediated contacts with phosphates and are either arginine or lysine (Marcotrigiano, 1997). The two exceptions again are *C. elegans* IFE-4 and human 4E-HP (eIF4E Homologous Protein) which contain the equivalent

of I162 (Rom 1998; Keiper 2000). In the *Leishmania* proteins, *LmeIF*4E1 and 2 contain K112, R157 and K/R162 respectively, while both *LmeIF*4E3-4 only contain R157, and lack basic residues at 112 and 162. On this basis, *LmeIF*4E1 and 2 possess all the features expected of a 7-methyl-GDP-binding protein, while *LmeIF*4E3-4 have some, but not all, the predicted features.

The eIF4G binds the convex dorsal side of eIF4E primarily due to the interaction between W73 of eIF4E and three amino acid side chains of the eIF4E binding peptide, YXXXL ô (where X can be any amino acid and ô is usually a hydrophobic amino acid such as L, M or F – Mader *et al.*, 1995), also present in eIF4E-binding proteins (Marcotrigiano, 1999; Niedzwiecka, 2002). Mutation of W73 to A or R, respectively, in the mouse and yeast eIF4Es abolishes binding to eIF4G, while the yeast F73 eIF4E mutant maintains binding to eIF4G (Ptushkina, 1998; Pyronnet, 1999). *LmeIF*4E1, 3 and 4 contain W73, while *LmeIF*4E2 contains F73 (Figure 1); thus on this basis all four proteins would be predicted to interact with eIF4G.

The most striking differences between the *Leishmania* proteins and known eIF4Es include the *LmeIF*4E3-4 N-terminal extensions of about 80 and 190 amino acids respectively and the significant insertions of variable number of amino acids at specific positions in the first three *Leishmania* proteins. Noteworthy of mention is a short insertion in both *LmeIF*4E1 and *LmeIF*4E3 in the loop which according to the vertebrate structure model would be placed between strands S1/S2 of the predicted β-sheet structure (Marcotrigiano, 1997). This loop, which in yeast is structured as a small α-helix (Matsuo, 1997), includes the W56 residue required for cap binding. A second insertion of 12 amino acids in *LmeIF*4E1 is near another tryptophan residue required for cap binding (W166) as well as the positively charged residues required for binding to the nucleotide phosphate groups. *LmeIF*4E1 has yet another long insertion on its C-terminus while *LmeIF*4E2 differs from all the other eIF4E sequences in having a long 21 amino acids insertion between the S4 strand and helix 2 of the predicted structure. Comparison with the two nematode eIF4E sequences included in Figure 1 does not indicate any motifs which might be involved in differential cap recognition and indeed it seems that the multiple eIF4E homologues evolved independently in both systems. Nevertheless, a comparison between the *L. major* and *T. brucei* eIF4E sequences (Table I; see discussion) confirm that all the specific features mentioned above for the various *L. major* proteins, with minor variations, are also conserved in the *T. brucei* homologues. Such conservation indicates that the multiple eIF4E proteins are conserved within the family *Trypanosomatidae* and play significant roles during the parasites' life cycle.

3.2 Expression analysis and quantitation of LmeIF*4E1-3 in* L. major *promastigotes.*

The first three *L. major* eIF4E sequences (*LmeIF*4E4 was only identified later in this study) were amplified, cloned and expressed in *E. coli* His-tagged or as GST-fusions. Isoform-specific antibodies were generated in rabbit against the three proteins. These antibodies were affinity purified (See Methods for details) and tested in Western-blots against the recombinant GST fusions to confirm their specificity. The antibodies were specific for each isoform and no cross-reaction was observed to the other recombinant *L. major* eIF4Es (data not shown), which was expected considering the low homology between the three isoforms (less than 35% identity between any two isoforms). These antibodies were then used to analyse the expression of the three eIF4E homologues in extracts of exponentially growing *L. major* promastigotes. These assays were designed in order to allow the quantitation of the proteins so that an estimate of their cellular levels could be determined (Figure 2). All three proteins were detected in the *L. major* promastigotes. Their levels however varied significantly, with *LmeIF*4E3 being the most abundant (>5 x 10^4 molecules/cell) and *LmeIF*4E1 the least one (<10^3 molecules/cell). Levels of *LmeIF*4E2 are intermediate between the other two proteins (~7 x 10^3 molecules/cell - Table II). When compared to the levels obtained for yeast eIF4E [3.2-3.6 x 10^5 molecules/cell – (von der Haar, 2002)] even *LmeIF*4E3 is expressed at levels lower than those obtained for the yeast protein. However, considering that in yeast a significant fraction of the eIF4E pool is not complexed to eIF4G (present at only 1.5-2 x 10^4 molecules/cell) these levels could account for the production of enough eIF4F for the survival of the organism. The same might even be true for *LmeIF*4E2 but it is unlikely that the levels of *LmeIF*4E1 would be sufficient to support translation in these parasites, at least in the promastigote stage.

3.3 Cap binding assay for the putative LmeIF*4Es.*

*LmeIF*4E1-3 were then labelled with [^{35}S] and tested for their ability to recognise the mammalian monomethylated cap. As positive control the *Xenopus* eIF4E homologue was also produced labelled with [^{35}S] and used in the same assay. All four proteins were tried in binding assays with the resin 7-methyl-GTP Sepharose 4B, followed by washes with GTP and the cap analogue. The *Xenopus* eIF4E bound to the resin as expected and was only eluted with the cap analogue. In contrast, of the three *L. major* proteins, only *LmeIF*4E2 was capable of binding to the resin (Figure 3). Binding was specific since GTP did not release the protein and in parallel experiments performed with CL6B Sepharose no binding was observed (not shown). *LmeIF*E1 and 3

lacked any affinity with the resin and both labelled proteins came down with the flow-through. These results confirm that at least one of the *Leishmania* putative eIF4E homologues (*LmeIF*4E2) has all the hallmarks, including the molecular weight, sequence and cap-binding ability of a typical eIF4E.

3.4 Modelling of LmeIF*4E2*.

In order to better define the conservation of the structure/function of the first three *L. major* eIF4E homologues we attempted to model their structure based on the mammalian eIF4E structure (Marcotrigiano *et al*., 1997). Unfortunately, the insertion next to W56 in both *LmeIF*4E1 and *LmeIF*4E3 prevented the adequate modelling of the neighbouring β-strands (S1 and S2 in Figure 1) which are essential to define the cap-binding slot. In contrast, the 21 amino acids long insertion in *LmeIF*4E2, which is located in a loop between a β-strand and an α-helix (S4 and H2), could be removed from the model without causing major interferences with the β-sheet or the general structure allowing a model of this protein to be produced. The overall stereochemical quality of the final model was assessed by the program PROCHECK (Laskowski *et al*., 1993), which confirmed that 88.8% of the residues are in the most favoured region and 10.5% in additional allowed regions. The programs Verify 3D and WHAT IF contributed to attest the model as a good one (data not shown) (Figure 4).

The final *LmeIF*4E2 model obtained confirms the predictions derived from the alignment in that it is possible to observe that the interactions mediated by the amino acids W56, D90, W102, E103, R157, K162, R112 and W166 in mammalian eIF4E, which are required for specific binding to the 7-methyl-GDP, are taken over by the amino acids W37, Q71, W83, E84, K93, R167, R172 and W176 in *LmeIF*4E2 (Figure 4B). Noteworthy of mention is the substitution of an aspartate (D90) in mammalian eIF4E for a glutamine (Q71) in the *Leishmania* protein. This position is not conserved in any of the *L. major* eIF4E homologues (see Figure 1), however the model indicates that at least for *LmeIF*4E2 the substitution does not modify the interactions required for the anchoring of GDP. The potential to bind eIF4G was also investigated through the analysis of the convex side of the molecule. The amino acid F54, substituting for the crucial W73 in mammalian eIF4E, is placed in the expected position for the eIF4E/eIF4G interaction (Figure 4C). Also shown in the figure are the amino acids H20 and V50 equivalent to H37 and V69 in the mammalian protein, as well as amino acids E134 and M138 substituting for L128 and L135 (Fig. 1). In mammalian eIF4E these amino acids have been shown to be involved in the eIF4E/eIF4G interaction (Marcotrigiano *et al*., 1999). Their positioning in the *LmeIF*4E2 model confirms that, despite the minor variations in sequence, not only *LmeIF*4E2 but possibly the other *L. major* eIF4E homologues as well have the potential to bind eIF4G. Finally, an

analysis of the electrostatic potential of both the cap binding slot as well as the eIF4G binding side in the *LmeIF*4E2 model, using the program GRASP (Nicholls *et al*., 1991), does not show any significant differences in overall charges when compared with the mouse homologue (not shown). Once again these results confirm that *LmeIF*4E2 has all the features required for a functional eIF4E homologue.

3.5 LmeIF*4As sequence analysis.*

The RNA helicase subunit of the eIF4F complex, eIF4A, is a highly conserved eukaryote protein. The similarity between the described *L. major* eIF4A (LeiF) (Skeiky, 1998) here called *LmeIF*4A1, and the human eIF4A1 homologue is 74% (Table I). *LmeIF*4A1 is a 403 amino acids long protein, with a predicted molecular weight of 45.3 kDa, encoded by two identical genes placed *in tandem* roughly 1.8 kb apart within chromosome 1. Blast searches with the human eIF4A as query using the raw sequence data from the *L. major* genome yielded both *LmeIF*4A1 as well as a possible second eIF4A homologue from *L. major* (*LmeIF*4A2), with 71% similarity to the human protein (Table I). The *LmeIF4A2* gene is located on chromosome 28 and codes for a 389 amino acids long protein with a predicted molecular weight of 43.9 kDa. Homologues to both *Leishmania* candidate eIF4A proteins were also identified in *T. brucei*, indicating that they are conserved within the *Trypanosomatids* (similarities of 91 and 89% between the *L. major* and *T. brucei* eIF4A1 and 2 respectively).

Figure 5 shows a sequence alignment comparing both *L. major* eIF4A sequences with those of human, yeast and plant eIF4A homologues. We have also included in the alignment human eIF4AIII [a negative regulator of translation – (Li, 1999)], a related Eubacteria protein (Ec) which is apparently also involved in translation initiation (Lu, 1999), an Archae DEAD-box protein (Mj) whose structure has been solved (Story, 2001), as well as two unrelated eukaryotic RNA helicases, *Lm*DHH1 and *Spisula* p47 (see below). eIF4A and other related DEAD-box RNA helicases are characterised by nine motifs (motifs I, Ia, Ib, II, III, IV, V and VI plus the recently identified Q motif), conserved among the various members of this protein family, with roles in ATP binding, and hydrolysis, and RNA-binding (Tanner, 2001; Tanner, 2003; Linder, 2003). The crystal structure of yeast eIFA (Caruthers, 2000) indicates that it assumes a "dumbbell" shape with two globular domains connected by a flexible linker. The two domains correspond to the ATPase (N-terminus) and RNA binding (C-terminus) activities of the protein and the structure also implicates several conserved arginine residues as important for eIF4A/helicase function. It is thought that these domains interact so that the binding and hydrolysis of ATP influences the RNA binding activity, and comparison with the structure of

the related protein from the Archae *Methanococcus jannaschii* supports this model (Story, 2001; Tanner, 2001).

Overall, the alignment in Figure 5 indicates that both *L. major* proteins share all the conserved motifs from RNA helicases, as well as several minor sequences typical of eIF4A. To support our identification of *LmeIF*4A2 as a possible eIF4A homologue, we searched the *L. major* sequence databases for any further eIF4A/RNA helicase homologues. The sequence with the nearest match identified (*Lm*Dhh1 in Figure 5) contain all the motifs which characterise the DEAD-box RNA helicases yet it shares a similarity of only 50% with the human eIF4A1 sequence (Table I). Remarkably this protein has a much higher degree of similarity (82%) to the clam *Spisula solidissima* p47 RNA helicase (Minshall, 2001), a member of the dhh1/p54 family of RNA helicases, with roles in decapping and translational repression (Coller, 2001; Fischer, 2002; Minshall, 2001; Nakamura, 2001), suggesting that it is a dhh1 homologue. The sequences aligned in Figure 5, plus a few others which we judged relevant, were used in a neighbour-joining analysis to compare their phylogenetic relations. The resulting tree (Figure 6A) confirms that both *L. major* eIF4A homologues are more closely related to each other and to other eIF4A proteins involved in translation than to non-related RNA helicases including *Lm*DHH1 and bacterial proteins. It also highlights the close proximity between the *Leishmania* and *T. brucei* homologues.

3.6. LmeIF*4As quantitation and expression analysis in L. major.*

Both the *LmeIF*4A1 and the *LmeIF*4A2 coding sequences were amplified from *L. major* genomic DNA, and expressed and purified in *E. coli* in His- or GST-tagged forms. Figure 6B shows a sample of the representative recombinant proteins. His-tagged *LmeIF*4A1 and 2 were then used to immunize rabbits to obtain specific antisera. The resulting antibodies were affinity purified and used to confirm their specificity as well as to determine their cellular levels in total *L. major* extracts as described for the *LmeIF*4Es (Figure 6C-D). Both antibodies were found to be specific for their respective proteins and no cross-reaction can be observed (Figure 6C). *LmeIF*4A1 is readily detected and its quantitation indicates that it is a very abundant protein with about 8 x 10^5 molecules per cell (Figure 6D, Table 2). This result is similar to what is observed in yeast eIF4A (von der Haar, 2002), and is also consistent with results in HeLa and reticulocyte lysate where eIF4A has been found to be a very abundant translation initiation factor present at levels approximately 10 fold higher than eIF4E (Duncan *et al*., 1987; Pause *et al*., 1994). In contrast we could not detect the presence of *LmeIF*4A2 even in 2 x 10^7 promastigote cells (Figure 6D) indicating that it is either a very rare protein, which could not possibly have a

dominant role in translation, or that it is expressed only at specific stages of development. The sensitivity of the *LmeIF*4A2 antibody prevents it from detecting the protein if it is present at levels lower than 2 x 10^3 molecules/cell as observed for *LmeIF*4E1. However in such circumstances its abundance would be at least 400 fold lower than that of *LmeIF*4A1 which makes it very unlikely for *LmeIF*4A2 to play a significant role in general translation at the *L. major* promastigote stage examined. We conclude then that *LmeIF*4A1 is the functional homologue of eIF4A in *L. major* promastigotes and that *LmeIF*4A2 is either expressed in other stages of the parasite life cycle or is necessary as a rare factor which cannot possibly have a prominent role in general translation.

3.7. Identification of candidate LmeIF*4G homologues.*

In contrast to eIF4A, the eIF4G subunit of eIF4F is a much less conserved protein at the sequence level. Even the two functionally similar eIF4G homologues in mammals (eIF4GI and II) are only 46% identical and similar levels of identity can be observed between the two yeast homologues (reviewed in Gingras, 1999). However, the central eIF4A/ RNA binding domain is conserved in the eIF4G sequences from divergent organisms, and its recently solved structure from human eIF4GII was shown to consist mostly of alpha-helices organised into HEAT repeats (Marcotrigiano, 2001). This eIF4G HEAT domain not only mediates the interaction with eIF4A and RNA (Marcotrigiano, 2001), but also in yeast it was implicated, together with flanking sequences, in the binding to eIF5 and eIF1 which may bridge the interaction between eIF3 and eIF4G (Asano *et al* 2001; He *et al* 2003). A second motif conserved in eIF4G sequences from distantly related organisms is the eIF4E binding peptide (Mader, 1995), whilst the PABP binding region is not to conserved in sequence from yeast to mammals (Tarun *et al*., 1997; Imataka *et al*., 1998).

Using the human eIF4GI as a BLAST query, we found five *L. major* ORFs containing the conserved central eIF4G domain (*LmeIF*4G1-5). The five proteins vary significantly in size (predicted molecular weights of 114, 152, 71.2, 84.6 and 88.8 kDa for *LmeIF*4G1, 2, 3, 4 and 5 respectively, corresponding to 1016, 1425, 635, 765 and 782 amino acids), overall charge (positive in *Lme*IEFG1 and negative in the other four) and in the relative location of the conserved eIF4G domain (central in *LmeIF*4G1,2,5 and N-terminal in *LmeIF*4G3-4). The genes coding for *LmeIF*4G1 and *LmeIF*4G2 are both located on chromosome 15 while *LmeIF*4G3-5 are encoded by genes located on chromosomes 16, 36 and 10 respectively (see Table I). No additional similarities with human, yeast or plant homologues can be seen anywhere else within their sequences and no putative eIF3, eIF4E or PABP binding domains can be clearly identified

based on homology analysis alone (not shown). Apart from the HEAT domain no conserved features can be seen between the various sequences with the exception of *LmeIF*4G3 and 4. These two proteins share a short similar N-terminus and a conserved region (approximately 120 Aas long) located about 180 amino acids after the HEAT domain, indicating that they may be functionally related.

Figure 7 shows an alignment comparing the sequences of the HEAT domain from the five putative *L. major* eIF4G proteins with the equivalent sequence from human eIF4GI, *S. cerevisae* Tif4631p, and wheat eIFiso4G as well as the translation regulator PAIP1. Overall the domain from *LmeIF*4G1 is the most similar to human eIF4GI whilst the ones from *LmeIF*4G2 and 5 are the least similar (Table I), but the differences in similarity are small. All five *L. major* domains contain most of the various conserved amino acids which have been shown in mammalian eIF4G homologues to be required for binding to eIF4A (Imataka & Sonenberg, 1997; Morino *et al*., 2000; Marcotrigiano *et al*., 2001). Minor exceptions are the K731 (human eIF4GI numbering), in *LmeIF*4G1, and the R781/K787 which are not conserved in any of the *L. major* homologues or even in yeast and plant. Strikingly, other residues which have been shown not to be required for eIF4A binding (such as F737, K765, E769, P770, F812 and R855) are also well conserved and are present in most *Leishmania* proteins. It is noteworthy mentioning that the doublet E769/P770 (strictly conserved in *LmeIF*4G1, 3 and 4 and partially conserved in the other 2 proteins) has been implicated in 43S recruitment and formation of the 48S complex (Marcotrigiano, 2001), suggesting that all five proteins could function in translation. At this stage however it is not possible to identify which of these proteins are true eIF4G orthologues with major roles in translation initiation.

3.8 Analysis of the interaction between L. major *eIF4A/eIF4G through pull-down assays.*

In animals, yeast and plants the specific binding between the eIF4A and eIF4G subunits of eIF4F needs to occur so that both proteins can function in translation. In order to functionally demonstrate a possible role in translation for the various *L. major* candidate eIF4A/eIF4G proteins, we decided to investigate if these factors are able to interact in a eIF4F like manner through pull-down assays. First, to validate the assay we initially immobilised in Ni-NTA agarose recombinant *LmeIF*4A1-2 and human eIF4A (His_6-eIF4A - Pestova *et al*., 1996), all containing a his-tag at their N-terminus (Figure 8A – left panel), and incubated the beads with ^{35}S-labelled human eIF4G (Joshi *et al*., 1994). Human eIF4A efficiently bound to human eIF4G, however neither *LmeIF*A1 or 2 was capable of binding to the human protein (Figure 8A – right panel). Fragments coding for the HEAT domain from three of the *L. major* candidate eIF4G proteins (*LmeIF*4G1-3, the ones initially identified), were selected to assay their ability to bind

eIF4A. These fragments (*LmeIF*4G1$_{128-475}$, *LmeIF*4G2$_{387-705}$, *LmeIF*4G3$_{26-310}$) were produced labelled with [^{35}S] methionine and tested for eIF4A binding. All three fragments include the region equivalent to aas 722-949 from human eIF4G1 which has been shown to be sufficient to promote 48S complex formation on the EMCV IRES (Lomakin *et al*., 2000). Only labelled *LmeIF*4G3$_{26-310}$ was able to bind efficiently to both *LmeIF*4A1 and human eIF4A as shown in Figure 8A (middle panel). No specific binding by the *LmeIF*4G1-2 constructs to the various recombinant proteins was observed (data not shown). Also, despite the fact that *LmeIF*4A1 and 2 share a similarity of 71%, no significant binding by any of the three *LmeIF*4G proteins to *LmeIF*4A2 was observed (Figure 8A and data not shown).

To confirm the specific interaction between *LmeIF*4G3$_{26-310}$ with *LmeIF*4A1, we expressed the first half of *LmeIF*4G3 (*LmeIF*4G3$_{1-310}$) as well as the full length *LmeIF*4G3 (*LmeIF*4G3$_{1-636}$) in *E. coli* fusioned to GST and tested them in a reverse pull-down assay. In this second pull-down strategy both GST fusions were immobilised in the glutathione-Sepharose resin and incubated with ^{35}S-labelled *LmeIF*4A1 and human eIF4A (Figure 8B). As negative controls we used GST on its own and a murine GST-cdc2 fusion (a kind gift from M. Carrington) which has a molecular weight very similar to the GST-*LmeIF*4G3$_{1-310}$ protein. This reverse assay confirmed the interaction between the GST-*LmeIF*4G3$_{1-310}$ construct and both labelled *LmeIF*4A1 and human eIF4A. However no binding by the full length GST-*LmeIF*4G3$_{1-636}$, or the negative controls, to the labelled proteins could be seen (Figure 8B – middle and right panels). This lack of binding observed for the full length *LmeIF*4G3$_{1-636}$ could be a consequence of inappropriate folding of the GST fusion. *LmeIF*4G3$_{1-636}$ was also used labelled with ^{35}S in pull-down assays with the his-tagged *LmeIF*4A1, as in figure 8A, however a degradation product from the full-length protein, similar in size to the *LmeIF*4G3$_{26-310}$, competes for the binding to *LmeIF*4A1 and prevents binding by the intact factor (not shown). In summary the pull-down experiments indicate that of the three *L. major* putative eIF4Gs tested only *LmeIF*4G3 interacts specifically with *LmeIF*4A1 as well as with human eIF4A.

So far our results are consistent with roles for both *LmeIF*4A1 and *LmeIF*4G3 in the process of parasite RNA translation. In order to compare the ratio of the 4A/4G homologues in *L. major* with those described for other eukaryotes we then decided to quantify the endogenous *LmeIF*4G3 levels as done previously for *LmeIF*4E1-3 and *LmeIF*4A1-2. Figure 8C shows a representative figure of this quantitation which is included in Table II. This quantitation indicate that, at 3.6 x 10^3 molecules per cell, *LmeIF*4G3 is present at about half the level observed for *LmeIF*4E2 in *L. major* promastigotes and is about 20 to 200 fold less abundant than *LmeIF*4E3 and *LmeIF*4A1 respectively. The large differences in concentration between the eIF4A/eIF4G or

eIF4E/eIF4G homologues are not very much different from what was observed in yeast where the eIF4G levels are about 45 and 20 fold less abundant than those of the eIF4A and eIF4E homologues respectively (von der Haar, 2002). However, considering that the overall number of eIF4G molecules in *L. major* is still one fifth of the levels in yeast, at the moment it is unclear whether the intracellular levels of *LmeIF*4G3 would be sufficient to support translation on its own in *L. major*. It does seem to have a prominent role in translation or translation regulation but we can't rule out contributions from the other eIF4G homologues identified.

3.9 Molecular modelling of the LmeIF*4G3 HEAT domain*

So far the full length eIF4G has been refractory to structural studies. Only recently have individual domains within this protein, such as the central HEAT domain from human eIF4GII (Marcotrigiano *et al*., 2001) and the eIF4E interacting region from yeast eIF4G1 (Gross *et al*., 2003), had their structure solved. The solution structure of the latter, bound to eIF4E, was solved by multidimensional NMR spectroscopy and shown to consist of a segment of about 90 amino acids in the protein N-terminus which includes the short conserved YXXXXLô motif. Upon binding eIF4E, this segment, positioned about 100 amino acids before the start of the HEAT domain, folds into an unusual ring-shaped structure consisting of five helices which wrap around the eIF4E N-terminus like a molecular bracelet (Gross *et al*., 2003). Given the evidence obtained above implicating *LmeIF*4G3 in translation in *L. major*, we decided to attempt the molecular modelling of *LmeIF*4G3 and study it at the structural level. However, its very short N-terminus, only 50 amino acids long, plus the lack of the consensus eIF4E binding motif, indicates that it is unlikely for it to bind to any of the *L. major* eIF4E homologues as has been shown for the yeast protein. Nevertheless, the *LmeIF*4G3 HEAT domain could be modelled and its structural features associated with eIF4A binding compared with the structure of the human eIF4GII HEAT domain.

Figure 9A shows the ribbon drawing of the predicted three dimensional structure of the *LmeIF*4G3 HEAT domain. Absent from the drawings are the two segments corresponding to the 2b-3a and 3a-3b loops (amino acids 128 to 140 and 161 to 178) since they were not structured in the original human protein. As for the *LmeIF*4E2 model, the overall stereochemical quality of the final model was assessed by the program PROCHECK which confirmed that 96.9% of the residues are in the most favoured region. Again the programs Verify 3D and WHAT IF also contributed to attest the model as a good one (data not shown).

The *LmeIF*4G3 HEAT domain folds into the same crescent-shaped molecule described for the human protein, with the 5 pairs of antiparallel α helices forming the same double layer

with the convex and concave surfaces formed by the a and b helices, respectively (Figs. 9A). A major difference between the *Leishmania* and mammalian protein arises from the analysis of the surface electrostatic potential of the model performed with the program GRASP (Figs. 9B-C). This difference is concentrated in the region which in the human protein has been implicated in the interaction with the EMCV IRES. In human eIF4GII this region has an overall positive charge but in the *LmeIF*4G3 Heat domain the equivalent region is negatively charged (Fig. 9B). In contrast the region involved in the eIF4A interaction presents a surface electrostatic potential similar to the human model and all the conserved amino acids implicated in the interaction with eIF4A (R61, H64, K69, L70, R270 and F273 equivalent to R756, R759, K764, L765, R968 and F971 in the mammalian protein) are positioned adequately for the interaction to occur (Fig. 9C and Marcotrigiano *et al*., 2001). Again these results give support for a role for *LmeIF*4G3 in translation, although the lack of a possible eIF4E interacting region and the differences in the putative RNA binding side of the protein's HEAT domain might indicate a regulatory role in protein synthesis or a different mechanism for its function in translation initiation.

4. Discussion

The three eIF4F initiation factors have multiple isoforms in *L. major* protozoan. This complexity is reflected in *Trypanosoma brucei* (which have similar isoforms) and indicates that they all have conserved roles in the parasites. Our results indicate that for each factor, one of the *L. major* forms seems to fulfil, in part at least, the expected characteristics of a translational initiation factor. Thus: *LmeIF*4E2, of 24kDa, can bind cap-Sepharose; *LmeIF*4A1, of 45 kDa, is an abundant cytoplasmic protein with intracellular levels similar to yeast eIF4A; *LmeIF*4G3 binds *LmeIF*4A1 and human eIF4A. However we believe that under more stringent assays many of the various isoforms may be seen to have specific roles in translation. For instance, it is possible that any of the other eIF4E homologues might have stronger affinities for the parasite cap4 than the *LmeIF*4E2. Alternatively, they may require association to the eIF4G homologues in order to bind efficiently to the cap, since in other eukaryotes the complex eIF4E/4G binds with higher affinity to cap than eIF4E alone (Haghighat & Sonemberg 1997).

The existence of multiple isoforms for the eIF4F subunits in other eukaryotes, specially pluricellular organisms, may be associated with different patterns of tissue expression and during development. Insights into the role of the multiple isoforms come from the nematode system which in many aspects regarding mRNA metabolism can be comparable to the trypanosomatids. In nematodes a 22 nt SL sequence, containing a trimethylated cap, is trans-spliced to about 70% of all mRNAs. It has been shown in *Ascaris lumbricoides* that the SL sequence and its associated

cap functionally collaborate to enhance translation, very likely at the level of initiation (Maroney *et al*., 1995). It is not known how these features influence protein synthesis, but it seems likely that they do so via trans-acting factors. The description of multiple homologues for eIF4E in *Caenorhabditis elegans* is also reminiscent of what we see in the parasite system (Jankowska-Anyszka *et al*., 1998). The nematode homologues differ in cap binding affinity, requirement for viability (Keiper *et al*., 2000) and possible roles in development (Amiri *et al*., 2001). In the case of the tripanosomatid protozoans, unicellular organisms, the multiple eIF4F isoforms could be associated with their different life stages or be required for the translation of different classes of mRNAs. Indirect evidence for the second hypothesis comes from wheat germ, where two different eIF4F isoforms have been described, with distinct eIF4E and eIF4G subunits, which differ in their ability to translate mRNAs containing structured regions in their 5'UTR as well as uncapped mRNAs and dicistronic messages (Gallie & Browning 2001).

Our own results tend to indicate a stage specific expression for at least some of the eIF4E orthologues since in *T. brucei* we have observed that the *Tb*EIF4E3 protein is expressed at high levels in the procyclic stage and is absent from bloodstream forms (Dhalia *et al*., unpublished results). Also it has been described that the mRNA coding for the protein we called *LmeIF*4E2 contains in its 3'UTR a sequence similar (68% identity) to the regulatory element found in the *Leishmania* amastin mRNA (Wu *et al*., 2000; Boucher *et al*., 2002). This element is found in a number of *Leishmania* mRNAs, several of which are differentially expressed in *L. donovani* amastigotes, and can confer amastigote-specific expression to a reporter mRNA possibly by regulating translation instead of stability (Boucher *et al*., 2002). It seems plausible to suggest that *LmeIF*4E2 could have its expression enhanced at the amastigote stage specially considering that, assuming a similar behaviour to the *T. brucei* orthologue, *LmeIF*4E3 (the most abundant of the three eIF4E proteins quantitated in *L. major* promastigotes), could be absent from the *Leishmania* mammalian form as well. Another interesting observation is suggested by the presence of 2 genes coding for *LmeIF*4E1 which is one of the least abundant proteins in promastigotes.

Regarding the eIF4A subunit the differences in levels between the two *Leishmania* eIF4A orthologues leaves no doubt as to the relevant role for *LmeIF*4A1 in translation in promastigotes. Since equivalent levels of expression have been reported for *LmeIF*4A1in both promastigotes and amastigotes stages of (Skeiky *et al*., 1995) it is unlikely that it would be replaced to a significant extent by *LmeIF*4A2 in the mammalian stage. Recent additional evidence from *T. brucei* indicates that *LmeIF*4A2 has a predominantly nuclear localisation (Dhalia *et al*., manuscript in preparation) ruling out any significant role in translation and suggesting possible

functions in mRNA metabolism. Strikingly it has been reported that human eIF4AIII, previously identified as a negative regulator of translation (Li, 1999), is a component of the exon junction complex in the nucleus, with roles in mRNA export, cytoplasmic localization and nonsense mediated decay (Chan *et al*., 2004; *Palacios et al*., 2004). It is tempting to speculate that *LmeIF*4A2 might have a related function in mRNA metabolism in trypanosomatids.

The multiple candidate eIF4G homologues identified in *L. major* add yet a new level of complexity to the study of translation initiation in this parasite. None of the proteins identified can be clearly assigned to the role of a translation factor. Although *LmeIF*4G3 (and possibly the related protein *LmeIF*4G4 since they share a few features in common) seems to be clearly involved in translation, its very short N-terminus may not accommodate binding sites to both eIF4E and PABP as in other eukaryotes. It is possible that both *LmeIF*4G3-4 could behave as a translational regulator/inhibitor as has been proposed for the mammalian protein p97/DAP-5/NAT1 (reviewed in Gingras *et al*., 1999), which is homologous to the C-terminal two thirds of mammalian eIF4G but lacks the N-terminal one third including the eIF4E- and PABP-binding sites. However none of the other eIF4G homologues from *L. major* share any homology to *LmeIF*4E3-4 outside the HEAT domain and despite having much longer N-terminus they do not have any motifs resembling the eIF4E binding consensus either.

In order to continue the comparison with the nematode system, we did a brief search looking for eIF4G homologues in the *C. elegans* genome sequences. Strikingly we found only one clear homologue, with long N-terminal and C-terminal regions, but containing what is seems to be a possible modified eIF4E binding sequence **F**GR**D**F**MV**. Similar sequences **F**SL**D**E**VV** and **F**SL**E**R**VL** are present in the short N-terminus of both *LmeIF*4G3 and *LmeIF*4G4 respectively. Their similarity to the nematode sequence could mean that all three sequences would be able to bind eIF4E-like proteins and then both *LmeIF*4G3 and *LmeIF*4G4 could be true eIF4G orthologues.

Acknowledgements

This work was supported by CNPq, CAPES, FIOCRUZ, FACEPE and The Wellcome Trust. Technical assistance to grow the *Leishmania* cells was provided by Lucas Rocha. We thank Al. Ivens for the gift of the cosmid DNAs and help with some preliminary sequence data. Sequencing of *L. major* was accomplished as part of the *Leishmania* genome network with support by The Wellcome Trust. We also acknowledge E. Cupolillo for some of the *Leishmania* strains. Sequence data for *L. major* was obtained in part from the Sanger Institute Website at

http://www.sanger.ac.uk/Projects/L_major/. M. Carrington for access to the *T. brucei* eIF4E sequences.

References

Amiri A, Keiper BD, Kawasaki I, Fan Y, Kohara Y, Rhoads RE and Strome S (2001) An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in C. elegans. Development 128, 3899-3912.

Asano K, Phan L, Valasek L, Schoenfeld LW, Shalev A, Clayton J, Nielsen K, Donahue TF, Hinnebusch AG (2001) A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNA(i)Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. Cold Spring Harb Symp Quant Biol. 66:403-15.

Bangs JD, Crain PF, Hashizume T, McCloskey JA and Boothroyd JC (1992) Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. J. Biol. Chem. 267, 9805-15.

Bates EJ, Knuepfer E and Smith DF (2000) Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites. Nucleic. Acids. Res. 28, 1211-1220.

Batista JA, Teixeira SM, Donelson JE, Kirchhoff LV and de Sa CM (1994) Characterization of a Trypanosoma cruzi poly(A)-binding protein and its genes. Mol. Biochem. Parasitol. 67, 301-12.

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

Campbell DA, Thomas S, Sturm NR (2003) Transcription in kinetoplastid protozoa: why be normal? Microbes and Infection 5: 1231-1240.

Caruthers JM, Johnson ER and McKay DB (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. Proc. Natl. Acad. Sci. U. S. A. 97, 13080-13085.

Chan CC, Dostie J, Diem MD, Feng W, Mann M, Rappsilber J, Dreyfuss G (2004) eIF4A3 is a novel component of the exon junction complex. RNA. 10(2):200-9.

Clayton C, Adams M, Almeida R, Baltz T, Barrett M, Bastien P, Belli S, Beverley S, Biteau N, Blackwell J, Blaineau C, Boshart M, Bringaud F, Cross G, Cruz A, Degrave W, Donelson J, El-Sayed N, Fu G, Ersfeld K, Gibson W, Gull K, Ivens A, Kelly J, Vanhamme L, et al. Genetic

nomeclature for *Trypanosoma* and *Leishmania*. 1998 Molecular and Biochemical Parasitology 97: 221-224.

Clayton C (2002) Life without transcriptional control? From fly to man and back again. EMBO J. 21, 1881-1888.

Coller JM, Tucker M, Sheth U, Valencia-Sanchez MA and Parker R (2001) The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. RNA. 2001 7(12):1717-27.

de Melo Neto OP, Standart N and Martins De Sa C (1995) Autoregulation of poly(A)-binding protein synthesis in vitro. Nucleic. Acids. Res. 23, 2198-205.

Duncan R, Milburn SC and Hershey JWB (1987) Regulated phosphorylation and low abundance of HeLa cell initiation factor eIF-4F suggest a role in translational control. J. Biological Chemistry 262 (1): 380-388.

Fischer N and Weis K (2002) The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. EMBO J. 2002 3;21(11):2788-97.

Futcher B, Latter GI, Monardo P, McLaughlin CS and Garrels JI (1999) A sampling of the yeast proteome. Mol. Cell Biol. 19, 7357-68.

Gallie DR and Browning KS (2001) eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. Journal of Biological Chemistry 276 (40), 36951-36960.

Gingras AC, Raught B and Sonenberg N (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu. Rev. Biochem. 68, 913-63.

Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G (2003) Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. Cell. 12;115(6):739-50.

Haghighat A and Sonenberg N (1997) eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. J Biol Chem. 29;272(35):21677-80.

He H, von der Haar T, Hinnebusch AG, McCarthy JE and Asano K (2003) The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. Mol Cell Biol. 23(15):5431-45.

Hershey and Merrick . Book chapter in translational control of gene expression 2001.

Hotchkiss TL, Nerantzakis GE, Dills SC, Shang L and Read LK (1999) Trypanosoma brucei poly(A) binding protein I cDNA cloning, expression, and binding to 5 untranslated region sequence elements. Mol. Biochem. Parasitol. 98, 117-29.

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(12):6940-7.

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,:7480-9.

Jankowska-Anyszka M, Lamphear BJ, Aamodt EJ, Harrington T, Darzynkiewicz E, Stolarski R and Rhoads RE (1998) Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in Caenorhabditis elegans can distinguish between mono-and trimethylated mRNA cap structures. J. Biol. Chem. 273, 10538-42.

Jones DT (1999) GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J. Molecular Biology 276:265-285.

Joshi B, Yan R and Rhoads RE (1994) In vitro synthesis of human protein synthesis initiation factor 4 gamma and its localization on 43 and 48 S initiation complexes. J. Biol. Chem. 269, 2048-55.

Keiper BD, Lamphear BJ, Deshpande AM, Jankowska-Anyszka M, Aamodt EJ, Blumenthal T and Rhoads, RE (2000) Functional characterization of five eIF4E isoforms in Caenorhabditis elegans. J. Biol. Chem. 275, 10590-10596.

Korneeva NL, Lamphear BJ, Hennigan FL and Rhoads RE (2000) Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. J. Biol. Chem. 275, 41369-41376.

Kumar S, Tamura K, Jakobsen IB, Nei M. Mega 2: molecular evolutionary analysis software. 2001. Bioinformatics 17(12):1244-1245.

Laskowski RA, MacArthur MW, Moss DS and Thornton JM. Procheck: A program to check the stereochemical quality of protein structures. 1993 J. Appl. Cryst. 26: 283-291.

Li Q, Imataka H, Morino S, Rogers GW Jr, Richter-Cook NJ, Merrick WC and Sonenberg N (1999) Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. Mol. Cell Biol. 19, 7336-46.

Liang XH, Haritan A, Uliel S and Michaeli, S. Trans and cis splicing in trypanosomatids: mechanism, factors and regulation. 2003 Eukaryot Cell. 2(5): 830-840.

Linder P (2003) Yeast RNA helicases of the DEAD-box family involved in translation initiation. Biology of the Cell 95, 157-167.

Lomakin IB, Hellen CUT and Pestova T (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. Mol Cell Biol.20(16):6019-29.

Lu J, Aoki H and Ganoza MC (1999) Molecular characterization of a prokaryotic translation factor homologous to the eukaryotic initiation factor eIF4A. Int. J. Biochem. Cell Biol. 31, 215-29.

Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. 1993 Nature 356: 83-85

Luking A, Stahl, U and Schmidt U (1998) The protein family of RNA helicases. Critical Ver. Biochem Mol Biol. 33(4), 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(9):4990-7.

Marcotrigiano J, Gingras AC, Sonenberg N and Burley SK (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell 89, 951-61.

Marcotrigiano J, Gingras AC, Sonenberg N and Burley SK (1999) Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell. 1999 3(6):707-16.

Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CU and Burley SK (2001) A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell. 7(1):193-203.

Maroney PA, Denker JA, Darzynkiewicz E, Laneve R and Nilsen TW (1995) Most mRNAs in the nematode Ascaris lumbricoides are trans-spliced: a role for spliced leader addition in translational efficiency. RNA. 1:714-23.

Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N and Wagner G (1997) Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nat. Struct. Biol. 4, 717-24.

Melo EO, Dhalia R, Martins de Sa C, Standart, N and de Melo Neto OP (2003) Identification of a C-terminal PABP-PABP interaction domain: Role in cooperative binding to poly(A) and efficient cap distal translational repression. J Biol Chem. 21;278(47):46357-68.

Minshall N, Thom G and Standart N (2001) A conserved role of a DEAD box helicase in mRNA masking. RNA. 7, 1728-1742.

Morino S, Hazama H, Ozaki M, Teraoka Y, Shibata S, Doi M, Ueda H, Ishida T and Uesugi S (1996) Analysis of the mRNA cap-binding ability of human eukaryotic initiation factor-4E by use of recombinant wild-type and mutant forms. Eur. J. Biochem. 239, 597-601.

Morino S, Imataka H, Svitkin YV, Pestova TV 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.

Muhich ML and Boothroyd JC (1988) Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol. Cell Biol. 8, 3837-46.

Nakamura Y, Hashimoto T, Kamaishi T, Adachi J, Nakamura F, Okamoto K and Hasegawa M (1996) Phylogenetic position of kinetoplastid protozoa inferred from the protein phylogenies of elongation factors 1alpha and 2. J. Biochem. (Tokyo) 119, 70-9.

Nicholls A, Sharp KA and Honig B (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins. 1991;11(4):281-96.

Niedzwiecka A, Marcotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszynska A, Dadlez M, Gingras AC, Mak P, Darzynkiewicz E, Sonenberg N, Burley SK, Stolarski R (2002)

Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. J Mol Biol. 2002 7;319(3):615-35.

Nilsen TW (1995) trans-splicing: an update. Mol. Biochem. Parasitol. 73, 1-6.

Pascolo E, Blonski C, Shire D and Toulme JJ (1993) Antisense effect of oligodeoxynucleotides complementary to the mini-exon sequence of the protozoan parasite *Leishmania amazonensis*. Biochimie 75, 43-7.

Pause A, Méthot N, Svitkin Y, Merrick WC 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 (5): 1205-1215.

Pestova TV, Hellen CU and Shatsky IN (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol. Cell Biol. 16, 6859-69.

Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI and Hellen CU (2001) Molecular mechanisms of translation initiation in eukaryotes PNAS. Proc Natl Acad Sci U S A. 19;98(13):7029-36.

Ponting CP (2000) Novel eIF4G domain homologues linking mRNA translation with nonsense-mediated mRNA decay. TIBS 25 423-426

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

Ptushkina M, von der Haar T, Vasilescu S, Frank R, Birkenhager R and McCarthy JE (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, 4798-808.

Pyronnet S, Imataka H, Gingras AC, Fukunaga R, Hunter T and Sonenberg N (1999) Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. EMBO J. 4;18(1):270-9.

Ramazeilles C, Mishra RK, Moreau S, Pascolo E and Toulme JJ (1994). Antisense phosphorothioate oligonucleotides: selective killing of the intracellular parasite Leishmania amazonensis. Proc. Natl. Acad. Sci. U. S. A. 91, 7859-63.

Rau MOT, Morley SJ and Pain VM (1996) A reevaluation of the cap binding protein eIF4E, as a rate limiting factor for initiation of translation in reticulocyte lysate JBC 271, 8983-8990

Raught B and Gingras AC (1999) eIF4E activity is regulated at multiple levels. International Journal of Biochemistry & Cell Biology 31, 43-57.

Rom E, Kim HC, Gingras AC, Marcotrigiano J, Favre D, Olsen H, Burley SK and Sonenberg N (1998) Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem. 22;273(21):13104-9.

Sali A and Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234: 779-815.

Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Saxena A, Worthey EA, Yan S, Leland A, Stuart KD and Myler PJ (2003) Evaluation of differential gene expression in Leishmania major Friedlin procyclics and metacyclics using DNA microarray analysis. Mol. Biochem. Parasitol. 129, 103-114.

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

Skeiky YA, Guderian JA, Benson DR, Bacelar O, Carvalho EM, Kubin M, Badaro R, Trinchieri G and Reed SG (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, 1527-37.

Skeiky YA, Kennedy M, Kaufman D, Borges MM, Guderian JA, Scholler JK, Ovendale PJ, Picha KS, Morrissey PJ, Grabstein KH, Campos-Neto A and Reed SG (1998). LeIF: a recombinant Leishmania protein that induces an IL-12-mediated Th1 cytokine profile. J. Immunol. 161, 6171-9.

Sonnenberg N and Gingras AC (1998) The mRNA 5'cap binding protein eIF4E and control of cell growth. Current Opinio in Cell Biology 10: 268-275

Sonnenberg N and Dever TE (2003) Eukaryotic translation initiation factors and regulators Current Opinion in Structural Biology 13, 56-63

Story RM, Li H and Abelson JN (2001) Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. PNAS 98 (4), 1465-1470

Strudwick S and Borden KL (2002) The emerging roles of translation factor eIF4E in the nucleus. Differentiation.;70(1):10-22.

Tanner NK and Linder P. (2001) DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol Cell. 8(2):251-62.

Tanner NK, 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. 2003 11(1):127-38.

Tarun SZ Jr and Sachs AB (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J. 16;15(24):7168-77.

Tarun SZ Jr, Wells S, Deardorff JA and Sachs A (1997) Translation initiation factor eIF4G mediates *in vitro* poly(A) tail-dependent translation. PNAS 94, 9046-9051

Ullu E and Tschudi C (1991) Trans splicing in trypanosomes requires methylation of the 5' end of the spliced leader RNA. Proc Natl Acad Sci U S A. 15;88(22):10074-8.

von der Haar T and McCarthy JE (2002) Intracellular translation initiation factor levels in Saccharomyces cerevisiae and their role in cap-complex function. Mol Microbiol. 46(2):531-44.

Vriend G (1990) WHAT IF: A molecular modeling and drug design program. J. Mol. Graph. 8: 52-56

Zeiner GM, Sturm NR and Campbell DA (2003) Exportin 1 mediates nuclear export of the kinetoplastid spliced leader RNA. Eukaryot Cell. 2(2):222-30.

Wells SE, Hillner PE, Vale RD amd Sachs AB (1998) Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell. 1998 2(1):135-40.

Wu Y, El Fakhry Y, Sereno D, Tamar S and Papadopoulou B (2000) A new developmentally regulated gene family in Leishmania amastigotes encoding a homolog of amastin surface proteins. Mol Biochem Parasitol. 110(2):345-57.

Table I: Summary of the sequence analysis of the different *L. major* eIF4F homologues, depicting identities (similarities) between the *Leishmania* and human sequences and also between the *L. major* and *T. brucei* homologues. Results obtained using the NCBI BLAST with the BLOSUM 62 Matrix.

L major sequence	GeneDB accession	Predicted molecular weight (kDa)	Chromosome localization	Identity (similarity) to Human homologue #	Identity (similarity) to nearest *T. brucei* homologue
***LmeIF*4E1**[@]	**LmjF19.1500/ LmjF19.1480**	**31.5**	**19 (2 genes)**	**27% (41)**	**45% (57)**
***LmeIF*4E2**[@]	**LmjF27.1620**	**24**	**27**	**22% (42)**	**48% (61)**
***LmeIF*4E3**	LmjF28.2500	**38**	**28**	**27% (43)**	**49% (61)**
***LmeIF*4E4**	LmjF30.0450	**33.8**	**30**	**28% (45)**	**37% (50)**
***LmeIF*4A1**[@]	**LmjF1.0780/ LmjF1.0770**	**45.3**	**1 (2 genes)**	**56% (74)**	**85% (91)**
***LmeIF*4A2**	**LmjF28.1530**	**43.9**	**28**	**50% (71)**	**79% (89)**
***Lm*DHH1**	LmjF35.0370	**46.4**	**35**	**30% (50)**	**86% (92)**
***LmeIF*4G1**	LmjF15.0060	**114**	**15**	**25% (43)***	**31% (47)**
***LmeIF*4G2**	**LmjF15.1320**	**145.9**	**15**	**21% (37)***	**33% (49)**
***LmeIF*4G3**	LmjF16.1600	**71.2**	**16**	**26% (39)***	**38% (55)**
***LmeIF*4G4**	LmjF36.6060	**84.6**	**36**	**22% (38)***	**26% (45)**
***LmeIF*4G5**	LmjF10.1080	**88.8**	**10**	**21% (37)***	**43% (60)**

* These sequences show similarity to human eIF4G only at the level of the central HEAT domain.

@ NCBI acessions avalilable: *LmeIF*4E1 – CAB94109 and CAB94111; *LmeIF*4E2 – CAB77676; *LmeIF*4A1 – NP_047099 # eIF4A1, eIF4E1 or eIF4G1

Table II: Summary of the quantitation of the various *L. major* eIF4F homologues analysed in this study.

L major sequence	fentomoles/ 10^6 cell	No. of molecules / cell	No. of molecules / cell in yeast *
***LmeIF*4E1**	**1.4**	**8.4×10^2**	**$3.2\text{-}3.6 \times 10^5$**
***LmeIF*4E2**	**12**	**7.2×10^3**	
***LmeIF*4E3**	**145**	**8.7×10^4**	
***LmeIF*4A1**	**230**	**8.4×10^5**	**$7\text{-}9 \times 10^5$**
***LmeIF*4A2**	**ND**	**ND**	
***LmeIF*4G3**	**6**	**3.6×10^3**	**$1.5\text{-}2 \times 10^4$**

*** Numbers based on [von der Haar, 2002 #61] for yeast eIF4E, eIF4A and eIF4G respectively. ND – Not determined.**

Figure Legends

Figure 1: Sequence comparison of the putative *L. major* eIF4E homologues with the human, yeast, plant and nematode sequences.

Clustal W alignment of different eIF4E homologues from selected organisms. Amino acids identical in the majority of the sequences are highlighted in dark gray, while amino acids similar to the consensus are shown in pale gray. When necessary spaces were inserted within the various sequences (dashes) to allow better alignment. The structural elements of the mammalian protein are shown numbered S1-S8 and H1-H4 (from Marcotrigiano *et al*., 1997) * indicate the conserved tryptophan residues. Double arrows highlight amino acids found to be relevant for the interaction with cap. Single arrows indicate conserved non-tryptophan residues shown to be involved in eIF4G binding (Marcotrigiano *et al*., 1999). # indicates the amino acid D90, replaced by a Q in *LmeIF*4E2 and described in Fig. 4. Relevant NCBI accession numbers: human (*Hs*) – P06730; yeast (*Saccharomyces cerevisae*, *Sc*) – P07260; plant (*Triticum aestivum*, *Ta*) P29557; nematode (*C. elegans*, *Ce*) IF4E1 and IF4E3 – NP_499751 and NP_503124 respectively.

Figure 2: Expression analysis and quantitation of *LmeIF*4E1-3 in exponentially grown *L. major* promastigotes.

(**A**) Serial dilutions of recombinant GST-tagged *LmeIF*4E1-3 compared in Coomassie Blue stained gels with known concentrations of BSA. The dots indicate the position of the full length proteins (**B**) Quantitation of the recombinant proteins (done at least twice). The bands from the BSA curve shown in A were quantified by densitometric scanning and plotted as a function of the BSA concentration. To calculate the concentration of recombinant *LmeIF*4E1-3, the densitometry of representative bands from these proteins (highlighted in the figure) were determined and spotted on the BSA curve. (**C**) Expression analysis of *LmeIF*4E1-3 in extracts of *L. major* promastigotes. The different recombinant GST-fusions, diluted to adequate concentrations (in fentomoles), and whole parasite extracts were fractionated in 15% SDS-PAGE and transferred to Immobilon-P membranes. These membranes were then incubated with the isoform specific sera raised against the three proteins, followed by incubation with goat anti-rabbit IgG conjugated with peroxidase. The Western-Blots were visualised by ECL. (**D**) Procedure for the quantitation of the endogenous levels of the various factors using *LmeIF*4E1 as an example. The Western results shown in C for GST-*LmeIF*4E1 was quantified by densitometry and plotted as a function of the number of fentomoles used in the assay. The densitometry obtained for the endogenous factor at $2.5x10^6$ cells was then spotted in the curve and used to calculate its concentration in fentomoles (3.1 fentomoles in the example). The same procedure was used for the different factors including *LmeIF*4A1 and *LmeIF*4G3. All the results presented are representative of at least four different experiments using a minimum of three independently grown cell cultures. The data obtained from the various experiments were processed as described above and used to calculate the values shown in Table II.

Figure 3: *LmeIF*4E2, but not *LmeIF*4E1 or 3, binds specifically to the eukaryotic cap analogue.

The DNAs coding for the three *L. major* eIF4E homologues (*LmeIF*4E1-3), as well as for the *X. laevis* eIF4E (*Xenopus* in the figure) used as positive control, were transcribed and translated *in vitro* in the presence of [^{35}S] methionine. The labelled proteins were tested for their ability to bind the resin 7-methyl-GTP Sepharose. Non-specific binding was removed by three washes with GTP and specific elution was performed with three more washes of the soluble cap analogue. Aliquots of the various washes were ran on SDS-PAGE and compared with samples from the original translation

reaction (Whole extract) as well as the non-bound fraction (Flow-through). The residual protein bound to the beads were recovered by washes with 2M KCl and SDS-PAGE sample buffer (Beads) and compared also. Arrows indicate proteins eluted by cap analogue. On the left are indicated sizes in kDa of protein molecular weight markers.

Figure 4: Model of the predicted *LmeIF*4E2 structure bound to 7-methyl GDP.

Ribbon diagrams of the overall predicted *LmeIF*4E2 structure and detail of the cap binding slot created using the program PyMol (http://www.pymol.org). (**A**) Model of the structure of *LmeIF*4E2 bound to 7-methyl GDP, based on the mouse eIF4E structure (Marcotrigiano *et al*., 1997). The aas W37 and W83 (equivalent to W56 and W102 in mammalian eIF4E) are shown binding the cap. The arrow indicates the loop specific to *LmeIF*4E2 which could not be modelled adequately. (**B**) Figure showing the details of the cap binding slot in the predicted structure, highlighting the interactions between amino acids W37, Q71, W83, E84, R167, R172 and W176 and the cap nucleotide. The interactions mediated through water molecules are not shown so that the K93 interaction (which correspond to R112 in mouse) is omitted. (**C**) Side view of the *LmeIF*4E2 model showing details of the predicted eIF4G binding surface. Amino acids H20, K21, V50, F54, E134 and M138 equivalent to H37, P38, V69, W73, L128 and L135in the mammalian protein are highlighted. The cap binding slot with the W37 and W83 amino acids is also shown for comparison.

Figure 5: Sequence comparison of the *L. major* eIF4A homologues with various selected sequences.

Clustal W alignment performed as described in Fig. 1 comparing the sequences of the *L. major* eIF4A homologues with selected sequences from several different organisms. * indicate the nine motifs typical of DEAD box RNA helicases (Tanner & Linder 2001; Linder 2003). Relevant NCBI accession numbers: human (Hs) P04765; yeast (*Saccharomyces cerevisae*, Sc) – NP_012397; plant (*Triticum aestivum*, Ta) – P41378; human eIF4AIII (HseIF4A3) – P38919; Eubacteria (*Escherichia coli*, Ec) – AAA23674; Archae (*Methanococcus jannaschii*, Mj) Dead box protein – NP_247653; Clam (*Spisula solidissima*, Ss) p47 – AAK85400.

Figure 6: *LmeIF*4A1, but not *LmeIF*4A2, is expressed as a very abundant protein in *L. major* promastigotes.

(**A**) Evolutionary relationship of the *L. major* eIF4A sequences. Neighbour-Joining tree based on the alignment of various eIF4A sequences including *LmeIF*4A1-2. Boostrap values are shown next to the respective branches (10,000 replicates). When compared with figure 5 we have included for this tree the sequences from: human eIF4A2 – acession AAH12547; *S. cerevisae* ded1p - NP_014847; *T. brucei* eIF4A1 (GeneDB) - Tb09.160.3270; *T. brucei* eIF4A2 (GeneDB) - TRYP_xi-153a09.q1c_17. (B-D) Expression analysis of the two putative *L. major* eIF4As. Recombinant *LmeIF*4A1-2 and whole parasite extract were fractionated in 15% SDS-PAGE and transferred to nitrocellulose membrane for Western-Blots. The membrane was then incubated with polyclonal sera against the two proteins followed by incubation with goat anti-rabbit IgG conjugated with peroxidase. The Western-Blot was visualised by ECL. (**B**) Recombinant proteins used for the antibody production (HIS) and quantitation assays (GST). The recombinant proteins were quantitated as described for the eIF4E homologues and about 0.5 (HIS-*LmeIF*4A2 and both GST proteins) or 1 μg (HIS-*LmeIF*4A1) of each were loaded on the gels and compared with known concentrations of BSA. (**C**) Analysis of the specificity of the antibody obtained against the recombinant proteins. Both GST and His-tagged fusions of *LmeIF*4A1-2 were tested with either one of the two antisera. For the *LmeIF*4A1 assay 12.5 μg of each recombinant protein were loaded on the gel whilst for *LmeIF*4A2 6.25 μg of the HIS and 25 μg of the GST fusions were used. Notice that the antibody produced against HIS-*LmeIF*4A2 cross-reacts with HIS-*LmeIF*4A1 since both recombinant proteins share conserved epitopes introduced by the plasmid vector. Those epitopes are absent from the GST constructs or the protein used to produce the anti-*LmeIF*4A1 antibody. (D) Analysis of the expression of both *LmeIF*4A1 and *LmeIF*4A2 in total *L. major* extracts. The quantitation analysis of *LmeIF*4A1 was processed as described before for the *LmeIF*4E homologues.

Figure 7: Sequence comparison of the HEAT domain from the putative *L. major* eIF4G homologues with selected sequences from various organisms.

Clustal W alignment performed as described in Fig. 1 comparing the sequences of the HEAT domain from the putative *L. major* eIF4G homologues (*LmeIF*4G1-5) with the equivalent domain from human, plant and yeast eIF4G homologues as well as the human PAIP1 protein. The predicted five antiparallel alfa helical pairs (1-5, A and B) are indicated (Marcotrigiano *et*

al., 2001). Selected mutations in mammalian eIF4G which have been shown to reduce the binding to eIF4A are shown as follows: A – R723D, R726D and K731D (Marcotrigiano *et al*., 2001); B – R781D and K787D (Marcotrigiano *et al*., 2001); 1 (M-1) – L729A, L732A and F737A (Imataka & Sonenberg 1997); and 4 (M-4) – R935A and F938A (Imataka & Sonenberg 1997). The R801D/K802D mutant (K in the figure ?) does not prevent eIF4A binding although it does abolish IRES binding (Marcotrigiano *et al*., 2001). Double arrows highlight amino acids which abolish the binding to eIF4A when individually mutated to alanine (Morino *et al*., 2000). * indicate conserved amino acids which do not interfere with eIF4A binding when mutated to alanine (Morino *et al*., 2000; Marcotrigiano *et al*., 2001). Relevant NCBI accession numbers: human eIF4GI (*Hs*) – Q04637; yeast Tif4631p (*Saccharomyces cerevisae*, *Sc*) – NP_011678; plant eIFiso4G (*Triticum aestivum*, *Ta*) Q03387; human PAIP1 – NP_006442.

Figure 8: Analysis of the interaction between human and *L. major* eIF4A homologues with the *LmeIF*4G3 protein.

(**A**) Pull-down assay using recombinant his-tagged *LmeIF*4A1-2 and human eIF4A immobilized in Ni-NTA beads and incubated with ^{35}S-labelled human eIF4G or ^{35}S-labelled *LmeIF*4G3$_{26-310}$ (the eIF4A binding region only from *LmeIF*4G3). Bound proteins were eluted with SDS-PAGE sample buffer, fractionated in 15% SDS-PAGE (for *LmeIF*4G3$_{26-310}$) or 10% SDS-PAGE (for human eIF4G) and stained by comassie-blue R-250 (to visualise the recombinant proteins). 15% SDS-PAGE showing total translation extract (Whole extract) and recombinant His-tagged proteins (left panel). Autoradiography showing specific binding (arrows) between *LmeIF*4A1-HIS/ labelled *LmeIF*4G3$_{26-310}$ and Human eIF4A-HIS/ labelled *LmeIF*4G3$_{26-310}$ (middle panel). Positive control autoradiography of a 10% SDS-PAGE showing the specific binding between human eIF4A and labelled human eIF4G (right panel). As negative control murine cdc2 cloned into the *Bam*H I/*Eco*R I sites of pGEX2T (Amersham Biosciences) and expressed as GST-fusion was used. (**B**) Reverse pull-down assay using GST-tagged *LmeIF*4G3$_{1-343}$/ *LmeIF*4G3$_{1-636}$ and ^{35}S-Met-labelled *LmeIF*4A1/ Human eIF4A. 15% SDS-PAGE showing ^{35}S-Met-labelled whole extract translation and recombinant GST-tagged proteins (left panel). Autoradiography showing specific binding (arrows) between *LmeIF*4G3$_{1-343}$-GST/ labelled *LmeIF*4A1 (middle panel) and *LmeIF*4G3$_{1-343}$-GST/ labelled Human eIF4A (right panel). (**C**) Quantitation of *LmeIF*4G3 in *L. major* promastigotes. Recombinant GST-*LmeIF*4G3$_{1-636}$ and total *L. major* extracts were assayed on Western-Blots with isoform specific polyclonal sera against

*LmeIF*4G3. The quantitation of the endogenous protein levels was performed as described in Fig. 2 for the eIF4E homologues.

Figure 9: Model of the structure of the *LmeIF*4G3 HEAT domain.

(**A**) Ribbon diagram of the structure of the predicted *LmeIF*4G3 HEAT domain with the concave surface on the right and the convex surface on the left. The α helices are labelled as described for human eIF4GII (Marcotrigiano *et al*., 2001) and as shown in Fig. 7. (**B-C**) Grasp representation of the surface of the domain colored coded for electrostatic potential (red < -8kBT;blue > 8kBT). A and B represent identical views of the molecule whilst in C the opposite side of the molecule (rotated 180^{o} about the long axis relative to A and B), predicted to be involved in the interaction with eIF4A, is shown. In B the dotted circle delimits the region in *LmeIF*4G3 which differs significantly from the human eIF4GII HEAT domain. In the latter this region is charged predominantly positive (Marcotrigiano *et al*., 2001) but in the *Leishmania* protein it is negative. Highlighted in C are the amino acids R61, H64, K69, L70, R270, F273, equivalent to amino acids implicated in the binding to eIF4A in the mammalian protein. The disordered 2b-3a and 3a-3b loops which could not be modelled are not shown.

Figure 1 – Dhalia et al.

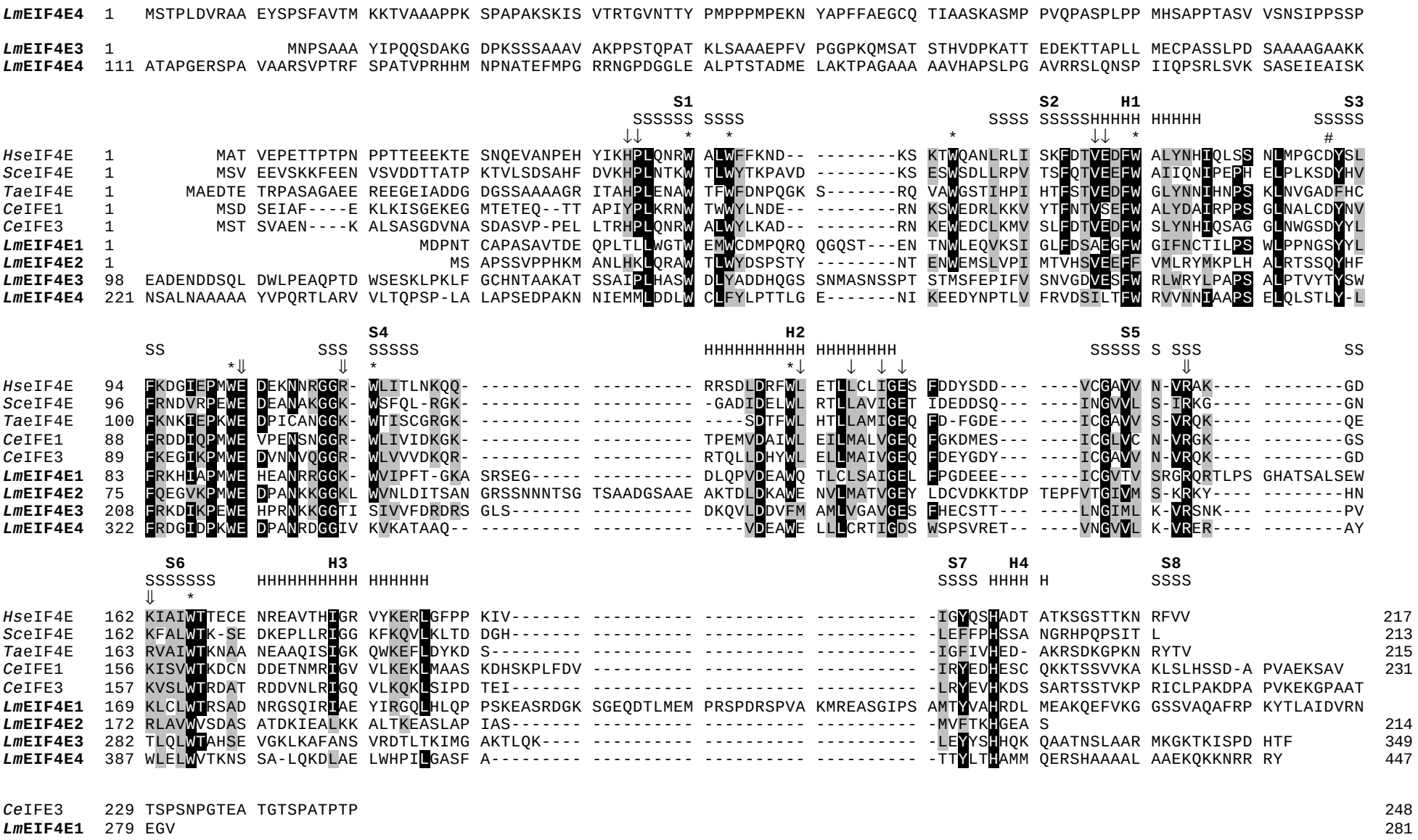


Figure 2 – Dhalia et al.

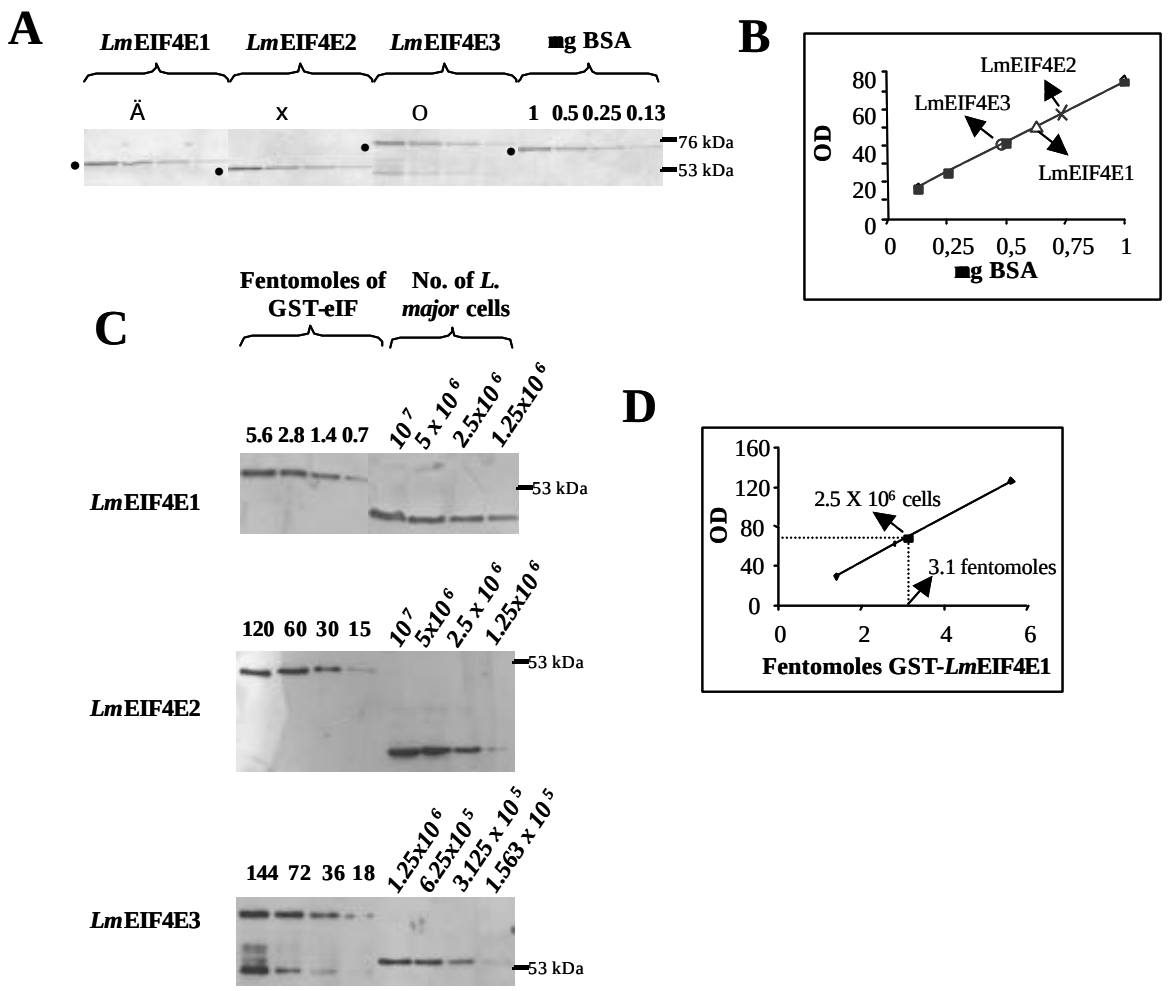


Figure 3 – Dhalia et al.

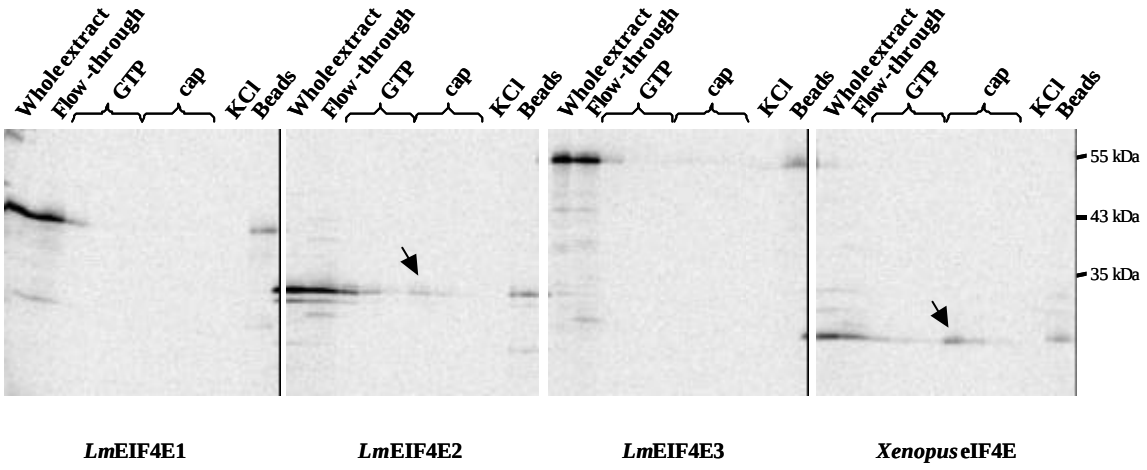


Figure 4 – Dhalia et al.

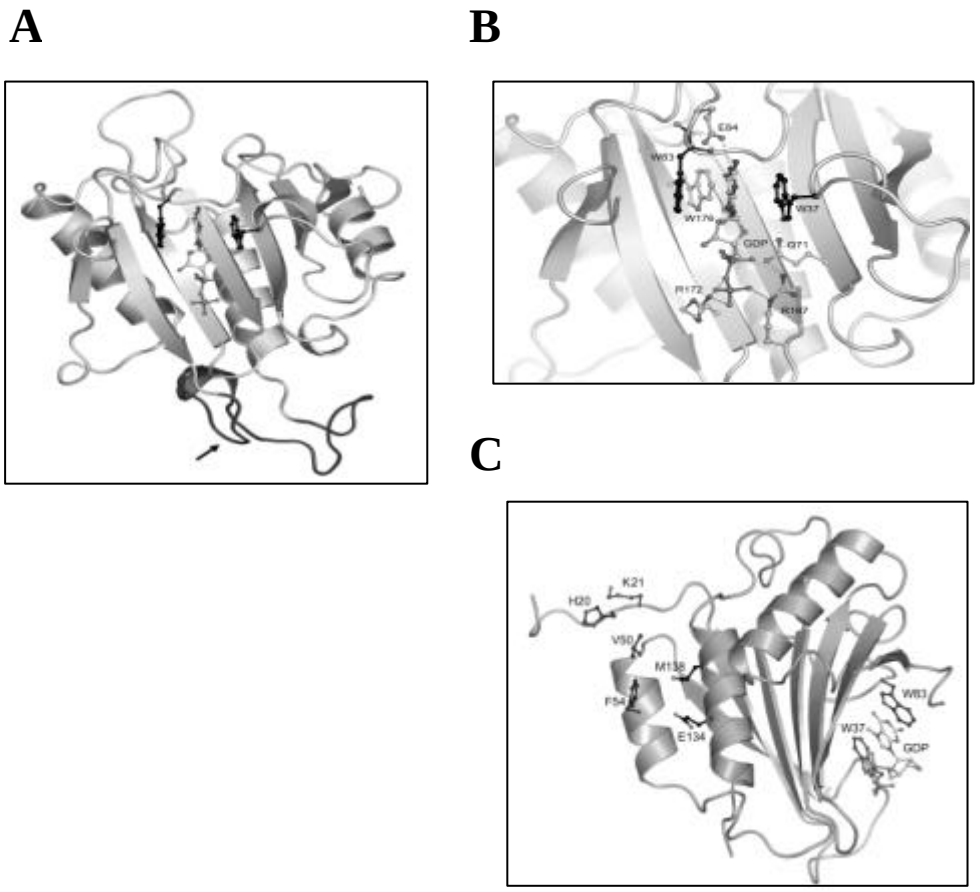


Figure 5 – Dhalia et al.

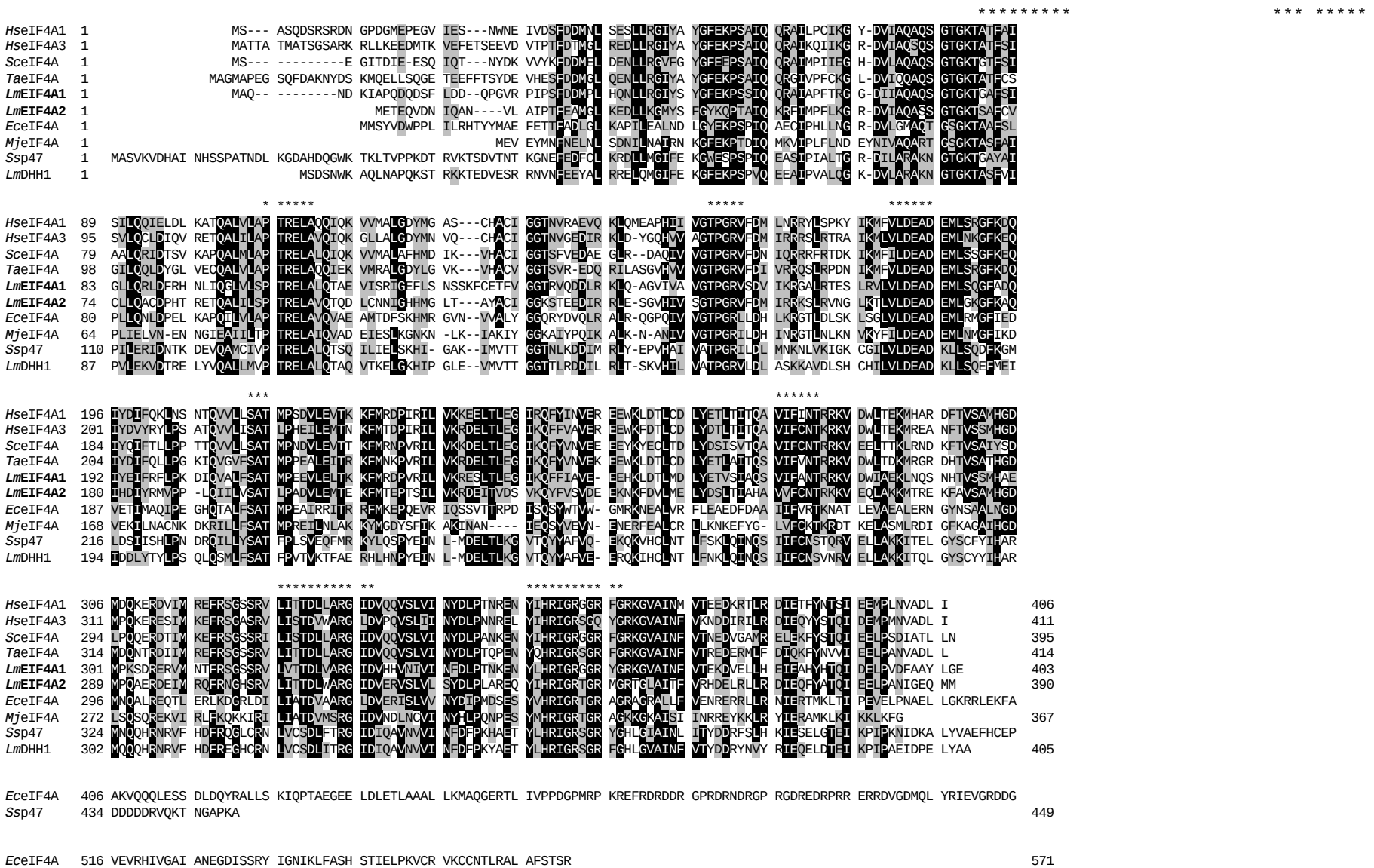


Figure 6 – Dhalia et al.

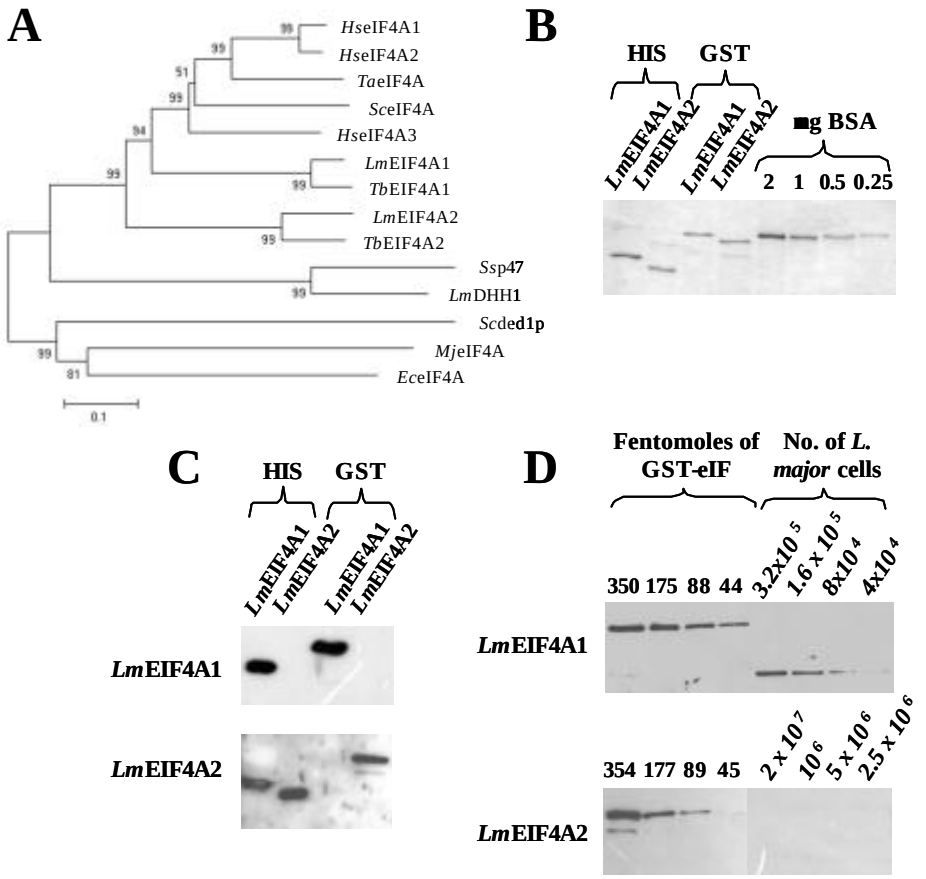


Figure 7 – Dhalia et al.

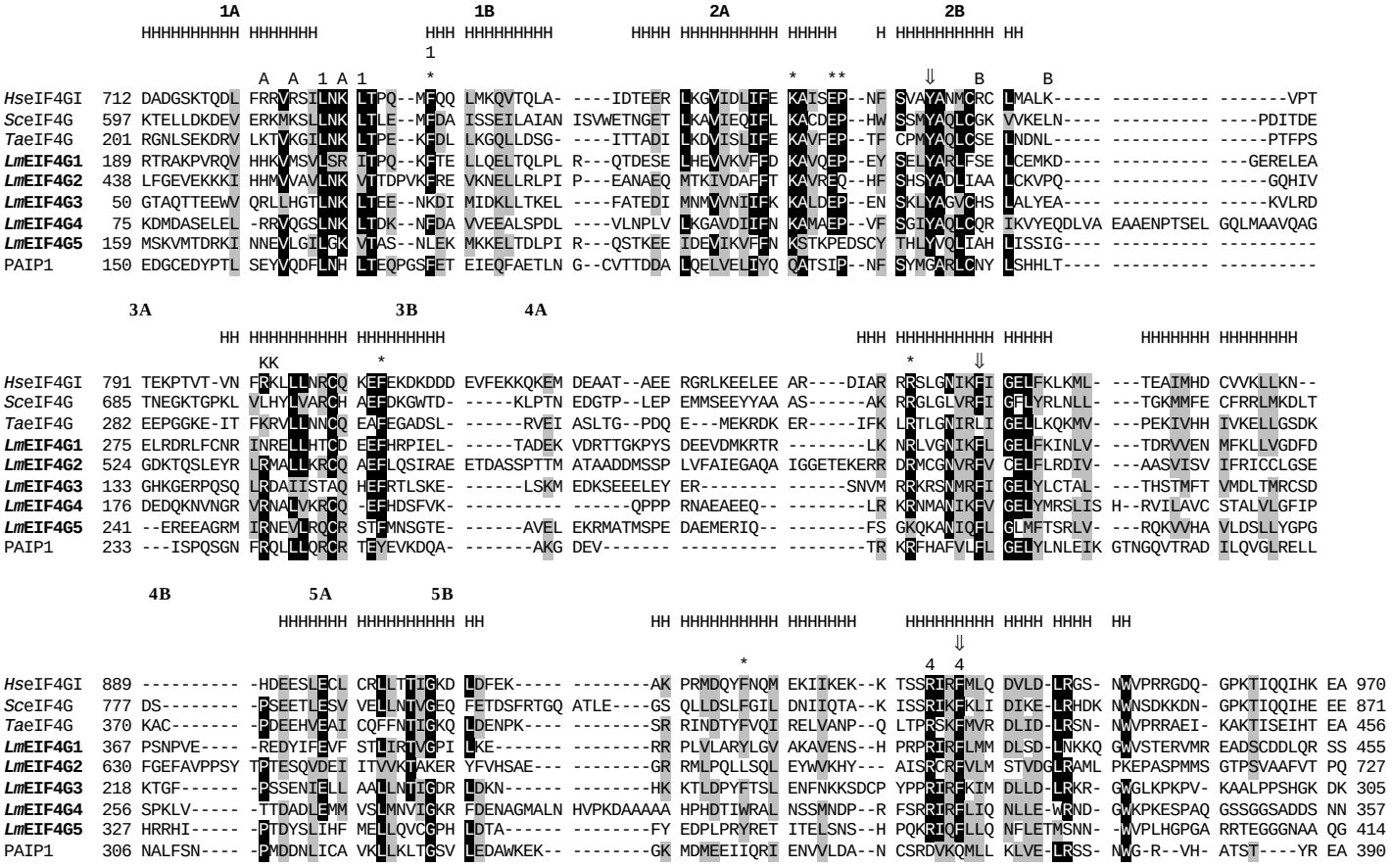


Figure 8 – Dhalia et al.

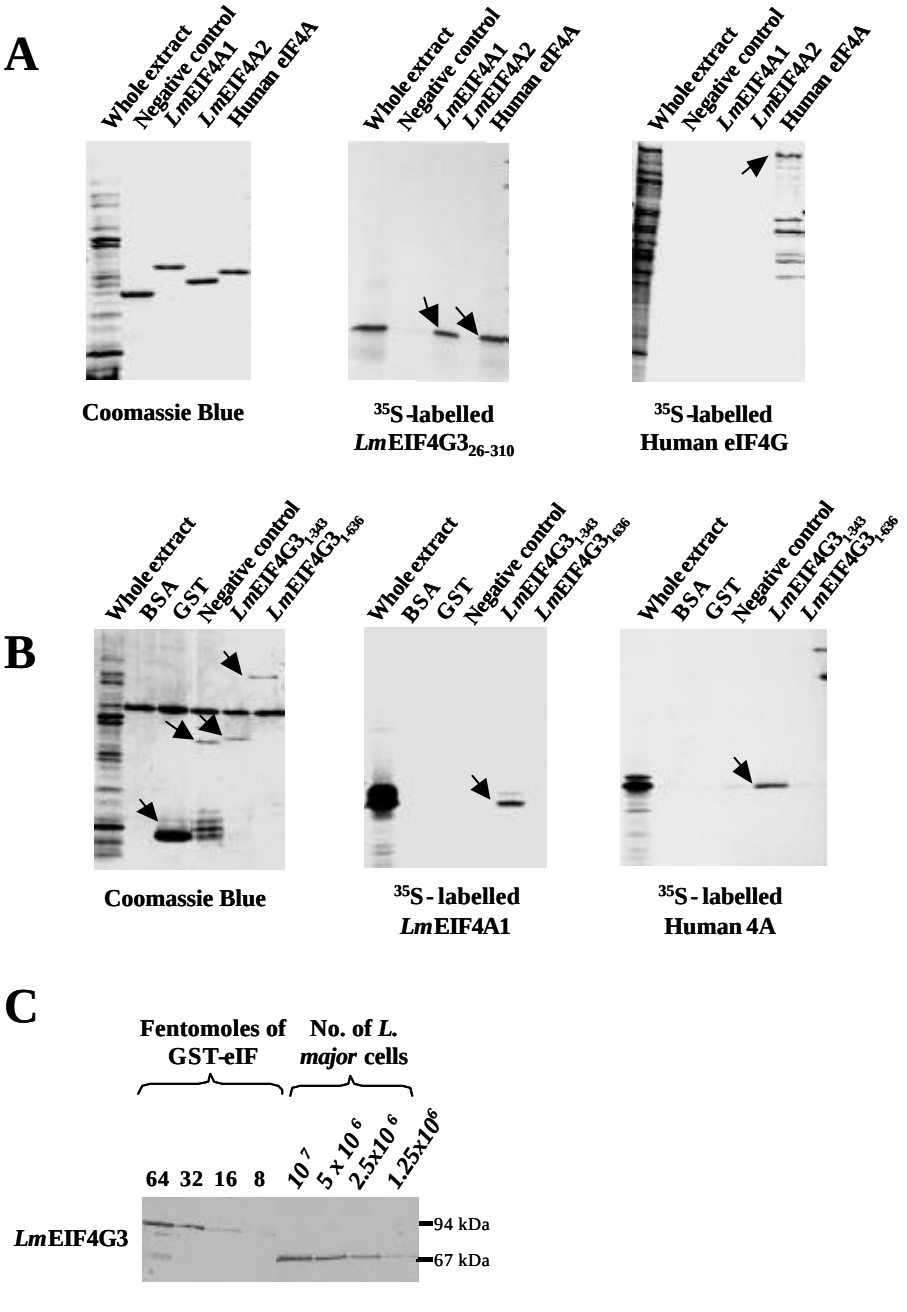
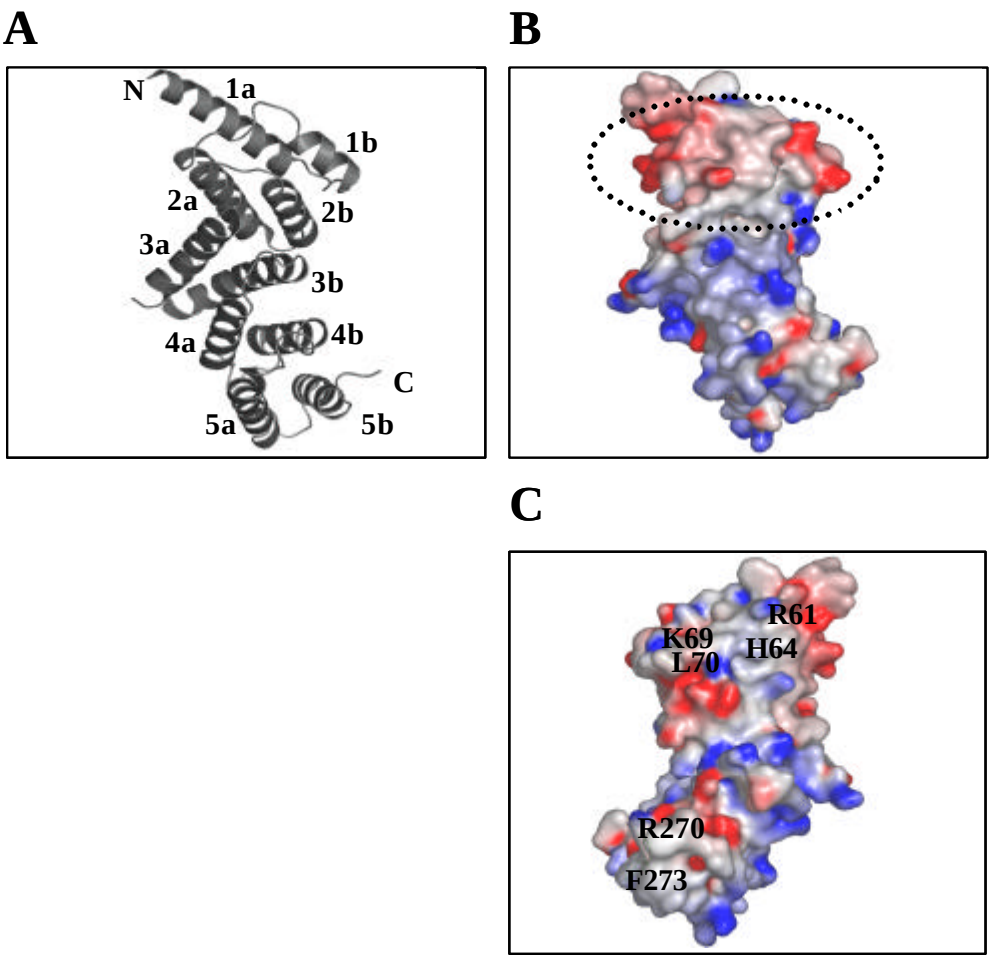


Figure 9 – Dhalia et al.



VI. ANEXO

1. METODOLOGIA

1.1 Cultivo de parasitas

Culturas celulares de *Leishmania major* foram obtidas através de cultivo em meio LIT - *Liver Infusion Triptose*) - modificado (0.2% sacarose p/v; 0.3% infuso de fígado p/v; 0.1% triptose p/v; 0.002% hemina p/v) contendo 10% de soro fetal bovino e 1% de ampicilina/estreptomicina. Na fase exponencial de crescimento, as culturas foram contadas em câmara de Newbauer e submetidas a duas lavagens com PBS pH 7.4 para posterior lise. As células de *L.major* foram lisadas para obtenção de DNA cromossomal segundo Sambrook & Russel (2001) ou com tampão Laemini 2X para análises de *Western-blot*.

1.2 PCR / Clonagem / Subclonagens

1.2.1 Clonagem do LmeIF4A2

O gene LmeIF4A2 foi amplificado a partir de DNA total de *L. major* usando um conjunto de *primers* que permitiriam que a seqüência amplificada fosse flanqueada pelos sítios *Bam* HI/*Xho* I (5'primer - CTC GGA TCC ATG GAG ACC GAG CAA GTA G; 3'primer - TG CTC GAG AAG CGA AAG GTG GAG AG). Este gene foi clonado nos sítios *Bam* HI/*Xho* I do vetor plasmidial pGEX4T3 (Amersham) e subclonado nos mesmos sítios no plasmídeo pRSETA (Invitrogen). A indução em *E.coli* gerou proteínas fusionadas em sua extremidade amino-terminal a uma cauda de histidinas ou à proteína GST.

1.2.2 Subclonagens LmeIF4G1-3

Os domínios centrais contendo as repetições HEAT dos 3 candidatos a eIF4G em *L. major* foram clonadas por nosso grupo em trabalhos prévios ($LmeIF4G1_{128-475}$, $LmeIF4G2_{387-705}$ e $LmeIF4G3_{26-310}$) (Dhalia, resultados não publicados).

Neste trabalho realizou-se subclonagens com estes cDNAs, utilizados em alguns dos experimentos. A seqüência $LmeIF4G1_{1-475}$ foi amplificada por PCR a partir de DNA total de *L.major* utilizando primers que permitiriam que a seqüência fosse flanqueada pelos sítios *Afl* III

/*Not* I (5' *Primer* – G AAC ATG TTC ATG GAA ACA CAG ATT TG; 3'*Primer* - T GGC GGC CGC CGA TAA GTA TGT GAG GAC GG) e clonado nos sítios *Nco* I/*Not* I no vetor pET21D (Novagen). O cDNA LmeIF4G1$_{1-1016}$ foi reconstituído a partir da inserção de fragmento oriundo da digestão do LmeIF4G1$_{128-1016}$ com as enzimas *Sal* I/ *Not* I e clonagem deste nos mesmos sítios na construção LmeIF4G1$_{1-475}$.

Outra etapa importante foi a subclonagem dos fragmentos HEAT no vetor de expressão pGEX (Amersham). Para o LmeIF4G1$_{128-475}$, utilizou-se um par de oligonucleotídeos (5' Primer – TGG AAT TCT ATG TCG GTC CGG AAG GAT G; 3'*Primer* – o mesmo utilizado para a obtenção do LmeIF4G1$_{1-475}$ no pET21D) que permitiriam obter, a partir da construção LmeIF4G1$_{128-1016}$/pET21D, o LmeIF4G1$_{128-475}$ flanqueados pelos sítios *Bam* HI/*Not* I. Este fragmento foi subclonado nos mesmos sítios no vetor pGEX. De forma semelhante aconteceu para o LmeIF4G2$_{387-705}$. O par de *primers* utilizados (5' *Primer* – GTG GGA TCC GAC GGC GGC GGG TTC AGC; 3'*Primer* – TG CTC GAG CAG CAT GGC ACG CAG GCC ATC) permitiu a subclonagem do LmeIF4G2$_{387-705}$ nos sítios *Bam* HI/*Not* I do pGEX4T3. Para o LmeIF4G3, a subclonagem foi do cDNA completo como descrito abaixo.

A seqüência completa do LmeIF4G3 (LmeIF4G3$_{1-636}$) foi amplificada por PCR a partir de DNA total de *L.major* utilizando *primers* que permitiram que a seqüência fosse flanqueada pelos sítios *Nco* I/*Not* I (5' *Primer* – TGC CAT GGA GTT CAC CGT GGA GCA G; 3' *Primer* - TCG CGG CCG CAT TAC TTG GGG AAG) e clonado nos mesmos sítios no vetor pET21D; um segundo conjunto de *primers* foi utilizado para inserir os sítios *Bam* HI/*Not* I no LmeIF4G3$_{1-636}$ usando o fragmento de PCR LmeIF4G3$_{1-636}$ obtido para clonagem no pET21D (5' *Primer* - GTG GGA TCC ATG GAG TTC ACC GTG GAG; 3' *Primer* – mesmo utilizado para obtenção da construção LmeIF4G3$_{1-636}$ no pET21D) permitindo a clonagem desta seqüência nos mesmos sítios do vetor pGEX4T3.

1.3 Transcrição e tradução *in vitro*

As construções LmeIF4G1 / LmeIF4G3 foram linearizadas com a enzima de restrição *Not* I, e o LmeIF4G2 foi digerido por *Xho* I para serem transcritas *in vitro* com a enzima T7 RNA polimerase (Amersham). A reação de transcrição se processou num volume final de 50 µL com: 5µL tampão de transcrição (tris-HCl 400mM, pH 7, $MgCl_2$ 150mM), 2µL de solução de NTPs a 100 mM, 0,5 µL DTT 1M, 1,5 µL inibidor de RNAse, 2,0 µL T7 RNA polimerase 5U/µL (Pharmacia), 3 µg de DNA plasmidial linearizado e H_2O (destilada, deionizada, estéril e

tratada com DEPC). Essa reação foi incubada a 37 ºC por 60 minutos. Após esse tempo foi adicionado mais 0,5 µL de T7 RNA polimerase e incubado por mais 60 minutos a 37 ºC. Ao término da incubação adicionou-se ao sistema 50 µL de tampão TE 10X (Tris 10 mM, EDTA 10 mM) para parar a reação. Da reação final de transcrição, com o volume final de 100 µL, foi analisado 1 µL dos transcritos sintetizados em gel de agarose 1%.

Para as transcrições com cap seguiu-se o mesmo protocolo descrito acima com adição de 5 µL de m7-GTP (cap) 5mM (Amersham) na reação. A solução (contendo NTPs) foi alterada para conter GTP na concentração de 10 mM, enquanto os outros NTPs continuaram a 100 mM, sem alteração. Após a primeira incubação de 30 minutos a 37 ºC, as moléculas de m7 GTP já estavam incorporadas. Em seguida 0,5 µL de GTP 100 mM foi adicionado para compensar a concentração menor desse nucleotídeo posta no início.

Os RNAs sintéticos obtidos foram extraídos em fenol/clorofórmio 1:1, precipitados com etanol 100% / NaCl 0,3 M e ressupendidos em 20 µL de H_2O. Uma alíquota de 1 µL RNA purificado (~1µg/µL) foi traduzido no lisado de reticulócito de coelho, de acordo com o fabricante (sistema de tradução - Promega) contendo metionina marcada com ^{35}S (2.5µCi/reação). As reações foram submetidas a temperatura 30 ^{o}C por 90 minutos. Em seguida, foi adicionado tampão de amostra Laemili 2X e os produtos das traduções foram separados por eletroforese em gel de poliacrilamida 15% em condições desnaturantes e exposto ao filme β-Max (Amersham) por autoradiografia.

1.4 Expressão e purificação das proteínas recombinantes

Para expressar as proteínas fusionadas a uma seqüência de poli-histidinas ou a proteína GST, transformou-se as construções plasmidiais em células *Escherichia coli* competentes das cepas BLR ou BL21. As bactérias transformadas foram crescidas a 37ºC em meio LB líquido na presença de ampicilina (100µg/mL) para BL21 ou LB ampicilina (100µg/mL)/ tetraciclina (12µg/mL) / cloranfenicol (20µg/mL) para a cepa BLR. O crescimento foi acompanhado pela dosagem da cultura, em diferentes intervalos de tempo, no espectrofotômetro. Ao atingirem a densidade óptica (DO) de 0,5 foram induzidas a 30ºC com IPTG na concentração final de 0,1mM por 4 horas. As células foram coletadas por centrifugação (a 10000rpm por 10 minutos a 4ºC) e ressuspendidas em PBS e lisadas por ultra-sonicação. Ao lisado foi adicionado o

detergente Triton-X 1% e centrifugado nas mesmas condições, o sobrenadante resultante foi incubado durante 1h com sua resina específica.

Para purificação das proteínas fusionadas à cauda de histidinas utilizou-se a resina Ni-NTA Agarose (Qiagen) e para as proteínas fusionadas a GST usou-se a glutationa-Sefarose 4B (Amersham). Para ambos os casos, 200µL de resina foram equilibrados em PBS, após três lavagens de 10 mL. As resinas foram incubadas com o lisado bacteriano resultante da ultrasonicação durante 1h a temperatura de 4ºC sob agitação. As proteínas recombinantes ligadas à glutationa-sefarose-4B, foram recuperadas da resina por lavagens desta com 500 µL de glutationa reduzida em concentrações de 20 mM e 50 mM por 1 h cada a 4ºC sob agitação. A 10µl das proteínas purificadas eluídas e da resina foram adicionados 10µL de tampão Laemmli 2X, fracionado em SDS-PAGE 15% e visualizado após coloração com Comassie blue R-250. Os polipeptídeos fusionados a uma cauda N-terminal ou C-terminal de seis histidinas ligados a resina Ni-NTA Agarose foram eluídos da resina por lavagens com 500 µL de tampão de lavagem (50 mM tampão fosfato de sódio pH 6,0, 300 mM NaCl e 10% glicerol) contendo imidazol na concentração de 20 mM e 500 mM. Cada lavagem foi realizada por 10 minutos a 4ºC sob agitação. Um volume de 10 µL de cada lavagem foi processado e visualizado, nas mesmas condições descritas para a purificação das fusões contendo GST.

1.5 Produção de anticorpos e *Western-blot*

As proteínas recombinantes com a cauda de His, LmeIF4G1-3 e LmeIF4A2 foram utilizadas para imunizar coelhos brancos adultos *New Zealand* de acordo com Coligan *et al.*, (1995). Assim, as proteínas recombinantes LmeIF4G1-3-His e LmeIF4A2-His foram isoladas diretamente de gel preparativo SDS-PAGE 10% e usadas em 4 imunizações diferentes. Em cada utilizando aproximadamente 100 µg e com um intervalo de 15 dias (entre a 1ª e a 3ª) e 30 dias na última (4ª). Em cada uma a região do gel de acrilamida contendo a proteína de interesse foi excisada, macerada num sistema de válvulas acoplados a seringas de 5mL. Adicionado de 200µL de adjuvante completo de Freund (1ª) ou incompleto (2ª a 4ª) (Amersham) e 600µL de PBS, os componentes foram misturados e aplicados pela via subcutânea em coelhos (3 regiões diferentes). Após a 3º exposição dos coelhos às proteínas, alíquotas de sangue foram coletadas da veia marginal da orelha dos animais. Confirmada a produção de anticorpos por *Western-blot* (descrito adiante), 1 semana após a 4ª exposição os animais foram sacrificados sob anestesia, tendo sido realizada a punção cardíaca para a retirada do maior volume de sangue possível. Em etapas subsequentes, os anticorpos foram imunopurificados. Neste procedimento

aproximadamente 100 µg das proteínas recombinantes foram separadas em gel SDS-PAGE 15%, transferidas para membrana de nitrocelulose Immobilon-P (Millipore) previamente tratada com metanol 100% durante 15 segundos e hidratada em água destilada. A membrana foi corada com Rouge Ponceau S 0.2% (p/v) (em 1% de ácido tricloroacético), para visualização das proteínas na membrana, descorado com água destilada e as bandas contendo as proteínas foram excisadas da membrana, picotadas e colocadas em tubos de 1,5 mL. Estes fragmentos de membrana contendo as proteínas imobilizadas foram então incubados com os respectivos soros (durante a noite a 4^{o}C) e os anticorpos purificados eluídos em pH ácido.

Nas reações de *Western-blot*, a membrana utilizada foi a mesma utilizada na purificação de anticorpos. As proteínas de interesse bem, como extratos viáveis de *L.major* (fase promastigota), foram fracionadas em gel SDS-PAGE 15% e transferidos para membranas. Estas, por sua vez, foram bloqueadas com leite desnatado 5% por 1h e, em seguida, incubadas por mais 1h em solução de TBS 1X, leite 1% e Tween-20 0,05% e um dos anticorpos específicos (imunoadsorvido) na diluição de 1:500. O segundo anticorpo, anti Ig-G conjugado a peroxidase, foi utilizado na diluição de 1:3000 e a revelação do *Western-blot* foi realizado por ECL.

1.6 Ensaio de afinidade ao cap associando eIF4Es a eIF4Gs.

As construções LmeIF4G1-3 e LmeIF4E1-3, clonadas no vetor pET21D, foram linearizadas com a enzima de restrição *Not* I, para o LmeIF4G2 utilizou-se *Xho* I. Em seguida estes DNAs, extraídos em fenol, foram transcritos e traduzidos *in vitro*, conforme já descrito. Os RNAs foram combinados dois a dois, isto é, cada homólogo de eIF4G com cada homólogo de eIF4E de *L.major*. Como controle positivo de ligação foram utilizadas as proteínas radioativas eIF4E de *Xenopus laevis*, obtidas a partir da construção do pGEM1 (cedido pela Dra Nancy Standart do Departamento de Bioquímica, Universidade de Cambridge – UK) e eIF4GI humano, obtido no pBluescript KS (Stratagene); cuja construção foi realizada pela subclonagem, nos sítios *Xba* I/ *Hind* III, do eIF4GI humano, oriundo da construção pSK-HFC1 (Joshi *et al*., 1994). Cada conjunto de proteínas marcadas com S^{35} foi incubado com a resina 7-metil-GTP-sefarose por 1 hora a 4 °C. A resina foi lavada com PBS, GTP 100 µM e cap solúvel 50 µM. Alíquotas coletadas em todas as fases do ensaio foram analisadas em SDS-PAGE 15% e o gel após secagem a vácuo, foi exposto ao filme β-MAX, sendo o resultado visualizado por auto-radiografia.

2. RESULTADOS

2.1 Análise da expressão do homólogo LmeIF4A2

Análises de bioinformática indicaram a presença de um segundo homólogo ao eIF4A, o LmeIF4A2 (de Melo Neto, resultados não publicados). O alinhamento dos genes LmeIF4A1-2 pode ser visualizado na figura 1. Com o objetivo de esclarecermos se o LmeIF4A2 seria um homólogo funcional do eIF4A, tendo um papel importante na tradução, seu gene foi clonado.

Experimentos anteriores desenvolvidos pelo grupo de Biossíntese Protéica em Tripanosomatídeos no departamento de microbiologia do CPqAM mostraram que LmeIF4A1 é uma proteína presente em grande quantidade (Rocha *et al*., resultados não publicados). De posse do fator LmeIF4A2, o questionamento inicial era se o perfil de expressão deste seria semelhante ou não aos obtidos para o LmeIF4A1. Para solucionar este problema utilizou-se anticorpos produzidos para a proteína LmeIF4A2 em ensaios de *Western-blot* (figura 2a). A partir de concentrações determinadas de proteína recombinante (50ng, 25ng, 12,5ng e 6,25ng) e do extrato total de *L.major* contendo $2x10^7$ e 10^6 células, deduziu-se que o fator LmeIF4A2 parece não ser expresso na fase promastigota de *L.major* (figura 2b).

Outra indagação era saber se algum dos LmeIF4Gs poderia interagir com o LmeIF4A2 e, assim, fornecer indícios de seu possível papel na tradução. Para responder esta questão utilizou-se a técnica de *pull down*. Inicialmente tivemos que subclonar os fragmentos LmeIF4G1-3 no pGEX, necessários na confecção do ensaio de *pull down* reverso. O produto da expressão/purificação destes genes foi observada em SDS-PAGE 15% (dados não mostrados). Ensaios preliminares mostraram que a proteína $LmeIF4G3_{26-310}$ interage com o LmeIF4A1. Assim, procurou-se incluir também a obtenção do gene $LmeIF4G3_{1-636}$, correspondente ao gene completo. Este gene foi clonado no pET21D e subclonado no pGEX, cuja expressão pode ser visualizada na figura 5. Todas as construções acima foram utilizadas pelo doutorando Rafael Dhalia nos ensaio de *pull down* encontrado na figura 3. O LmeIF4A2 não interage com nenhum dos LmeIF4G1-3 *in vitro*. O LmeIF4A1 interage com o fragmento $LmeIF4G3_{26-310}$ conforme já mencionado, mas surpreendentemente não interage com a proteína recombinante inteira $LmeIF4G3_{1-636}$.

2.2 Análise da expressão do homólogo LmeIF4G3

O domínio heat de 3 dos homólogos do eIF4G de *L.major* (LmeIF4G1$_{128-475}$, LmeIF4G2$_{387-705}$ e LmeIF4G3$_{26-310}$), clonados previamente, foram expressos e as proteínas purificadas utilizadas na imunização de coelhos. Os anticorpos obtidos foram testados quanto a sua capacidade de reconhecer especificamente suas respectivas proteínas (figura 4). A próxima etapa seria avaliar o perfil de expressão de cada uma delas em *L.major*. Foram obtidos dados, até o momento, para o LmeIF4G3, entretanto ainda não fomos capazes de observar a expressão dos homólogos LmeIF4G1-2. Anticorpos anti-LmeIF4G3 foram imunoadsorvidos com a proteína recombinante LmeIF4G3$_{1-636}$GST e utilizados em reações de *Western-blot*. A proteína recombinante utilizada no ensaio também foi o LmeIF4G3$_{1-636}$GST (figura 5a). Experimentos preliminares de *Western-blot* mostraram que o LmeIF4G3 é expresso em *L.major* e que o tamanho da proteína corresponde a proteína recombinante LmeIF4G3$_{1-636}$. Para quantificá-la foram utilizados concentrações da proteína LmeIF4G3$_{1-636}$GST recombinante e de extrato de *L.major*. Comparando as curvas de proteína recombinante e extrato total concluímos que o LmeIF4G3 é expresso na forma promastigota de *L.major* na concentração de 1 nanograma de proteína em 10^6 células (figura 5b)

2.3 Análise da interação ao cap de homológos eIF4Es / eIF4Gs

Como mostrado nos ensaios de *pull down* o LmeIF4A1 interage com o LmeIF4G3 *in vitro*. Outra dúvida levantada era se homólogos LmeIF4Es poderiam interagir com homólogos LmeIF4Gs. Para responder esta pergunta realizou-se um ensaio de afinidade dos diferentes LmeIF4Es pelo cap na presença dos diferentes LmeIF4Gs. Optou-se por este experimento por dois motivos: 1) é descrito na literatura que o eIF4E interage com a porção amino-terminal do eIF4G (Mader *et al*., 1995); 2) o eIF4E de mamíferos interage com maior afinidade pelo cap na presença de eIF4G (Haghighat *et al*., 1996).

Para o ensaio de afinidade, inicialmente foi reconstituída a porção N-terminal do LmeIF4G1. Em trabalhos prévios, a região clonada do LmeIF4G1 correspondia ao domínio central e a região carboxi-terminal da proteína. Para este ensaio seria importante obter a parte codificante para a porção N-terminal no LmeIF4G1, uma vez que esta região poderia abrigar um potencial domínio para interação de algum dos homólogos LmeIF4Es. O fragmento foi inicialmente clonado no vetor pET21D. Em seguida utilizando-se sítios para enzimas de restrição presentes em comum no LmeIF4G1$_{1-475}$ e LmeIF4G1$_{128-1016}$ reconstituímos o LmeIF4G1$_{1-1016}$.

De posse dos cDNAs LmeIF4G1-3 / LmeIF4E1-3 aptos ao ensaio de afinidade, foi realizada a transcrição e tradução *in vitro*. Nessa etapa também fizemos a expressão *in vitro* do eIF4G humano e eIF4E de *Xenopus*, proteínas que seriam utilizadas como controle do ensaio. Confirmados a produção das proteínas (figura 6), iniciamos o ensaio de afinidade. RNAs sintéticos de cada um dos LmeIF4G1-3 foram combinados a cada um dos LmeIF4E1-3. Em seguida foram incubados com a resina como já descrito em materiais e métodos. A autoradiografia resultante mostrou que não ocorria aumento de afinidade do LmeIF4E1 (figura 7a), LmeIF4E2 (figura 7b) e LmeIF4E3 (figura 7c) pelo cap na presença de quaisquer LmeIF4G. O ensaio foi validado pelo controle positivo, o cap solúvel foi capaz de remover especificamente o eIF4E de *Xenopus* que estava associado ao eIF4G humano (figura 7d). Como esperado o LmeIF4E2 reconheceu o cap de mamífero (figura 7d) (Dhalia *et al*., 2001). O LmeIF4G3 ligou inespecificamente ao LmeIF4E3.

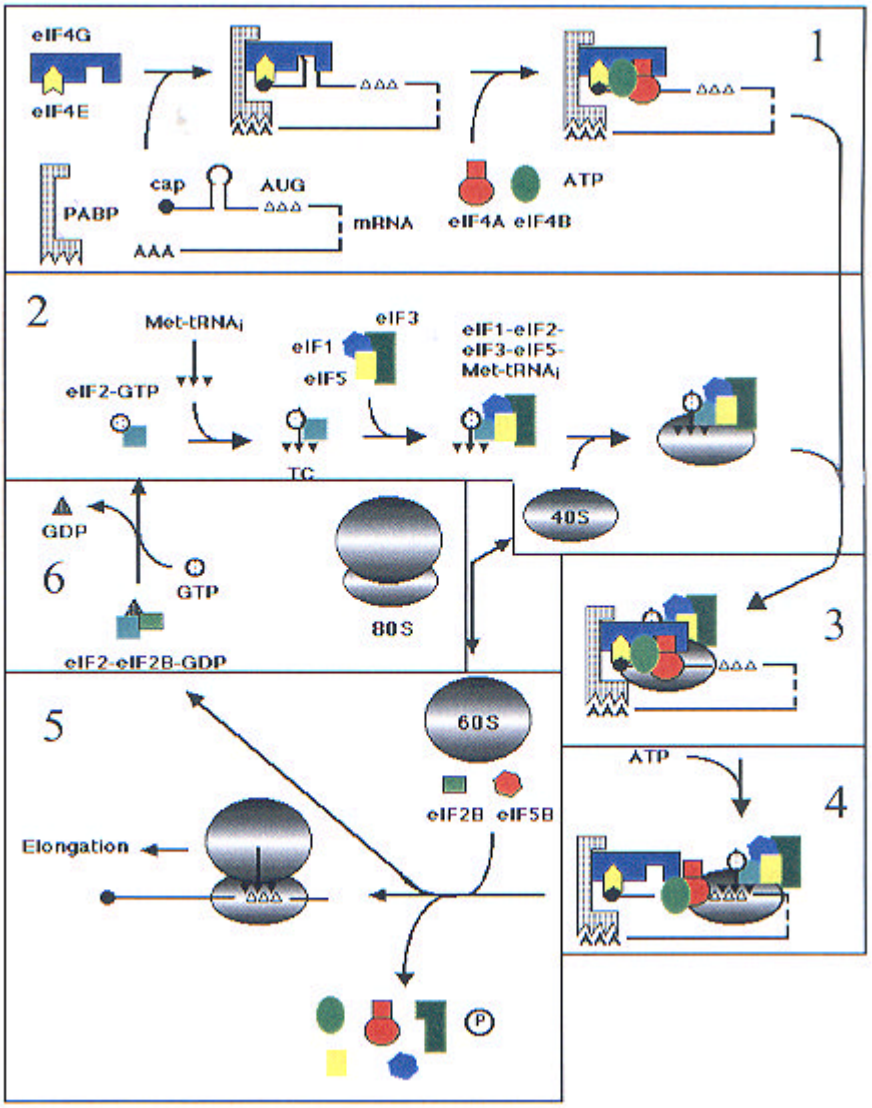


Figura 1. Modelo da iniciação da tradução cap dependente. O fator eIF4G se liga ao fator eIF4E que se associa ao cap do mRNA e a PABP promovendo a circularização do mRNA. O recrutamento dos fatores eIF4A e eIF4B prepara o mRNA para ligação do ribossomo por remoção de estruturas secundárias num processo dependente de ATP(1). O ribossomo 80S é dissociado nas subunidades ribossomais 40S e 60S. A subunidade 40S liga ao complexo protéico contendo eIF1, eIF3, eIF5 e ao complexo ternário eIF2-GTP-Met-tRNAi (TC) (2). O complexo resultante de pré-iniciação então se associa com o mRNA através da interação do eIF3 com o eIF4G no cap do mRNA(3). O complexo de iniciação da tradução se movimenta no sentido 5'-3' no mRNA (scanning) e reconhece o códon através do pareamento correto AUG-Met-tRNAi (4). Os fatores eIF5 e eIF5B desencadeiam a hidrólise de GTP, e a liberação de eIF2-GDP que promove a liberação dos demais fatores e ligação da subunidades 60S e 40S para formar o ribossomo 80S competente para elongação polipeptídica (5). Regeneração do eIF2-GTP em reação catalisada por eIF2b (6).

Figura 2. Alinhamento no programa Clustal W das seqüências dos homólogos eIF4A de *Leishmania major* (LmeIF4A1-2). Amioácidos idênticos estão marcados em preto. Amioácidos similares são mostrados em cinza. Quando necessário foram inseridos espaços representados por traços para permitirem um melhor alinhamento. * indicam os nove motives típicos da família de RNA helicases *DEAD box* (**Tanner & Linder 2001; Linder 2003**).

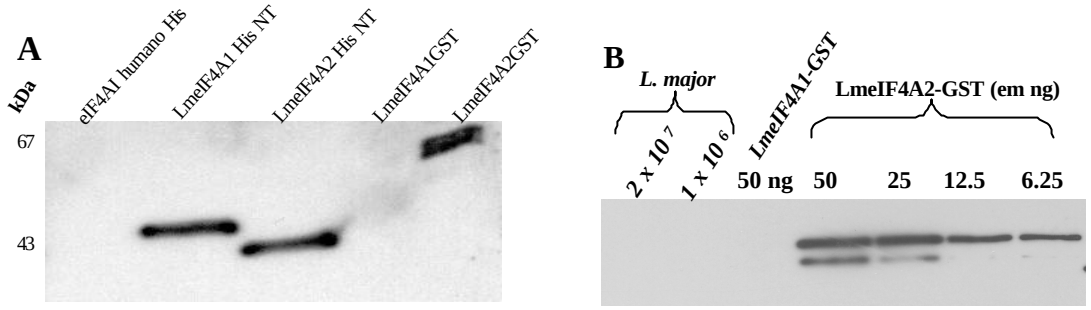


Figura 3. Análise da especificidade do Anti-LmeIF4A2 (A) e de expressão do LmeIF4A2 (B). (**A**) As proteínas eIF4A I humano (fusionada a His), LmeIF4A1 e LmeIF4A2 (fusionadas His ou GST) foram fracionadas em gel SDS-PAGE 15% e transferidos para membrana de nitrocelulose. Em seguida, utilizando-se o anticorpo anti-LmeIF4A2 testou-se a capacidade de reconhecer especificamente LmeIF4A2. As proteínas recombinantes LmeIF4A1-HisNT e possuem um epítopo em comum. O anti-LmeIF4A2 foi produzido utilizando-se a proteína LmeIF4A2-HisNT. Em outro experimento (**B**), uma curva da proteína LmeIF4A2-GST purificada (50ng; 25ng; 12.5ng; 6,25ng) e extratos de *L.major* (contendo $2x10^7$; 10^6 células) foram analisados pela mesma técnica. Para o *Western-blot* utilizou-se o soro anti-LmeIF4A2 imunoadsorvido na concentração 1:500, e em seguida anti-IgG conjugado a peroxidase, 1:3000. O ensaio de *Western-blot* e a detecção por ECL foram realizados como descritos na metodologia

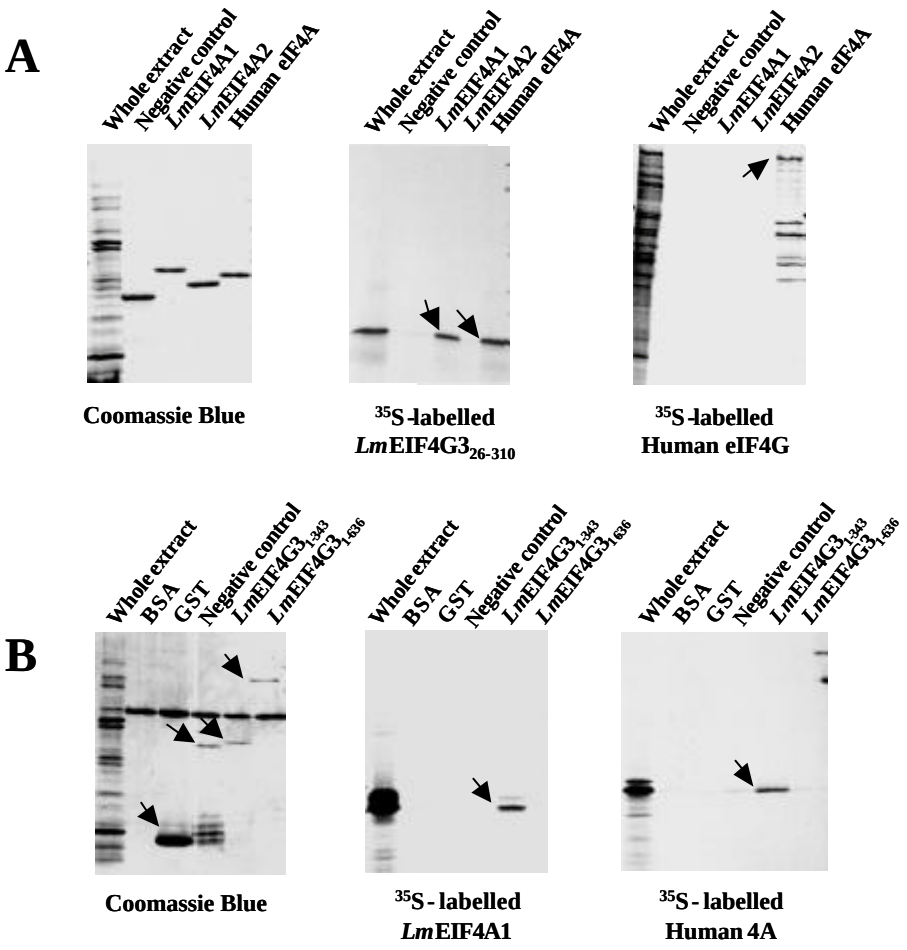


Figura 4. Ensaio de interação *in vitro* LmeIF4A1 / LmeIF4G3. (**A**) Ensaio de Pull-down usando LmeIF4A1-2 e eIF4A humano recombinantes fusionadas a cauda de Histidinas imobilizadas na resina Ni-NTA agarose e incubadas com eIF4G humano ou LmeIF4G3$_{26-310}$ marcados com ^{35}S. Proteínas ligadas foram eluídas com tampão de amostra Laeminni 2X, fracionados em gel SDS-PAGE 15 % (para *Lme*IF4G3$_{26-310}$) ou 10% (para eIF4G humano) e corado com comassie-blue R-250 (para visualizar as proteínas recombinantes). SDS-PAGE 15% mostrando extrato total da tradução (Whole extract) e proteínas recombinantes fusionadas a histidinas (figura a esquerda). Autoradiografia mostrando ligação específica (setas) entre *Lm*EIF4A1-HIS/ *Lm*EIF4G3$_{26-310}$ radioativo e eIF4A-HIS humano / *Lm*EIF4G3$_{26-310}$ radioativo (figura central). Controle positivo da autoradiografia de um Gel SDS-PAGE 10% mostrando ligação específica entre o eIF4A humano e o eIF4G humano marcado com ^{35}S (figura a direita). Como controle negativo foi utilizado a proteína cdc2 de murino fusionada a GST. (**B**) Ensaio reverso de Pull-down usando *Lm*EIF4G3$_{1-343}$/ *Lm*EIF4G3$_{1-636}$ fusionadas a GST e *Lm*EIF4A1/ Human eIF4A marcados com ^{35}S. SDS-PAGE 15% mostrando extrato total da tradução (Whole extract) e proteínas recombinantes fusionadas a GST (figura a esquerda). Autoradiografia mostrando ligação específica (setas) entre *Lm*EIF4G3$_{1-343}$-GST e *Lm*EIF4A1marcado com ^{35}S (figura central) e *Lm*EIF4G3$_{1-343}$-GST e eIF4A humano marcado com ^{35}S (figura a direita).

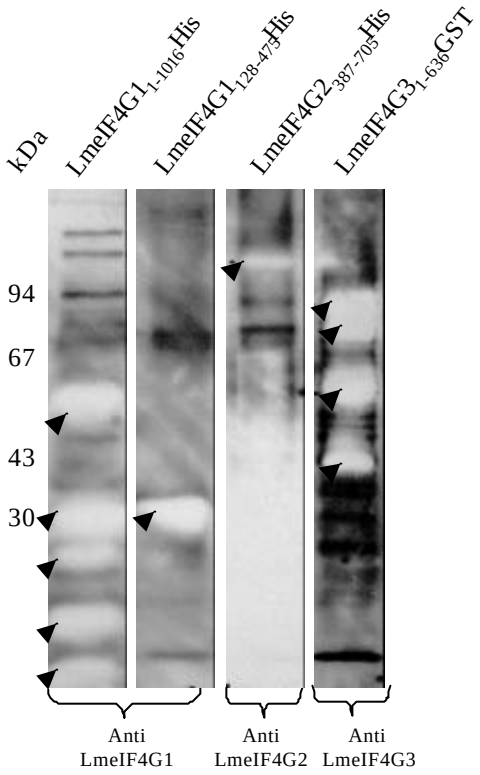


Figura 5. Produção de anticorpos anti-LmeIF4G1-3. As proteínas recombinantes LmeIF4G1$_{128-475}$ e LmeIF4G2$_{1-705}$ (com cauda de His) e LmeIFG3$_{1-636}$ (fusionada a GST) foram utilizadas em ensaios de *Western-blot* para confirmar a produção de anticorpos. Para os soros anti-LmeIF4G1 e anti-LmeIF4G3 utilizou-se a diluição de 1:7500, para o soro anti-LmeIF4G2 1:1000. A reação ECL foi intensa devido a grande quantidade de proteína usada (~1µg), assim elas ficaram brancas. As setas ilustram o reconhecimento das proteínas LmeIF4G1-3 através de seus anticorpos específicos (Anti LmeIF4G1-3). A presença de múltiplas bandas no LmeIF4G1$_{1-1016}$His e LmeIF4G3$_{1-636}$GST se deve ao fato destas proteínas possivelmente terem sofrido degradação.

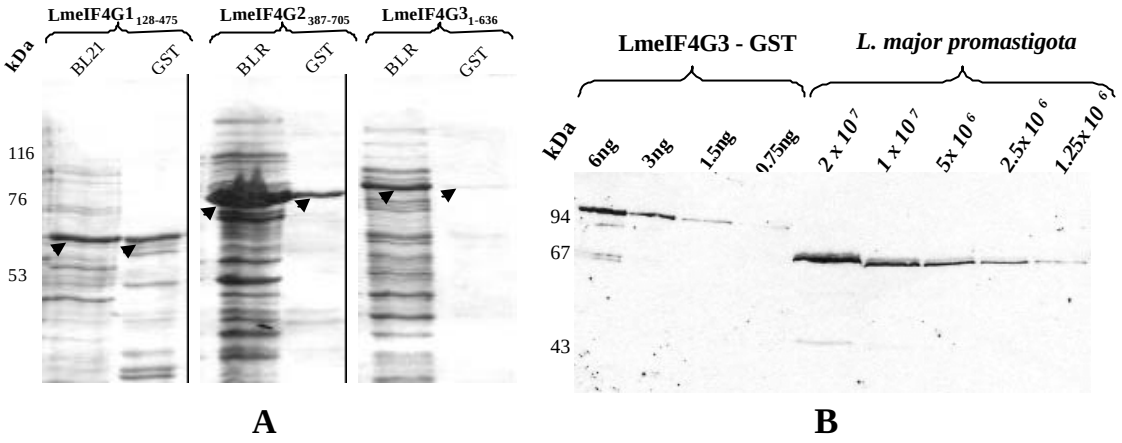


Figura 6. Expressão das construções LmeIF4G1-3 em *E.coli* e análise da expressão da proteína LmeIF4G3 em *Leishmania major.* Os genes LmeIF4G1-3 foram clonados e as respectivas proteínas foram expressas em *E.coli* fusionadas a GST ou uma cauda histidinas na região carboxi-terminal. As proteínas contendo histidinas foram utilizadas para imunizar coelhos para a produção de anticorpos específicos. As proteínas recombinantes ou extrato total do parasita foram analisados por *Western-blot* como descrito na figura 2. Nós não fomos capazes de observar a expressão dos homólogos LmeIF4G1-2, mas o terceiro homólogo é mostrado nesta figura. (**A**) SDS-PAGE (15%) corado com Coomassie blue mostrando as proteínas LmeIF4G1-3 fusionadas a GST em extrato total de *E.coli* (BL21 ou BLR) ou purificadas em resina GST, marcador de peso molecular indica o tamanho das proteínas (**B**) *Western-blots* comparando diferentes concentrações da proteína recombinante LmeIF4G3-GST com extrato total de proteínas obtidos de promastigotas de *L. major* (diferentes quantidades de célula).

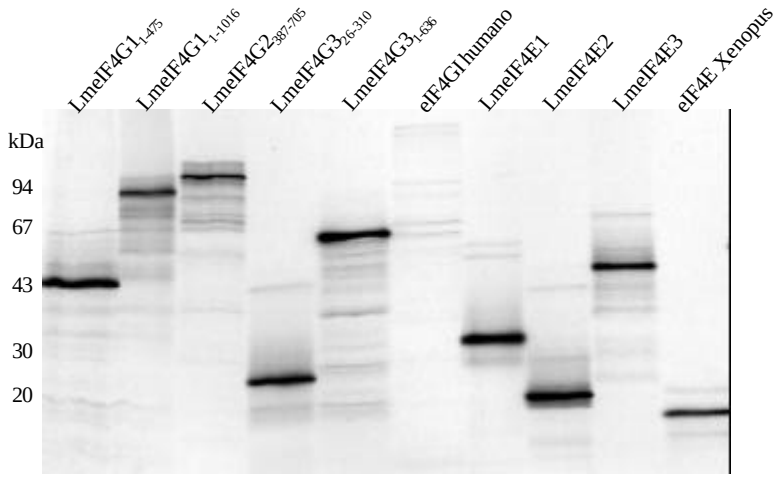


Figura 7. Expressão *in vitro* de diferentes construções do LmeIF4G e do LmeIF4E. cDNAs para os homólogos do LmeIF4G1-3 e LmeIF4E1-3, no pET21D, foram linearizados com enzimas de restrição posicionados após o fim da matriz de leitura e transcritos *in vitro* na presença de cap. Após constatação da integridade destes RNAs, o mesmo foi purificado e submetido a tradução em RRL suplementado com metionina radioativa. As respectivas traduções foram separadas em gel SDS-PAGE 15%, coradas com *Comassie blue*, o gel foi seco e exposto a filme β-Max (Amersham). Traduzimos também o eIF4G humano e o eIF4E de *Xenopus*.

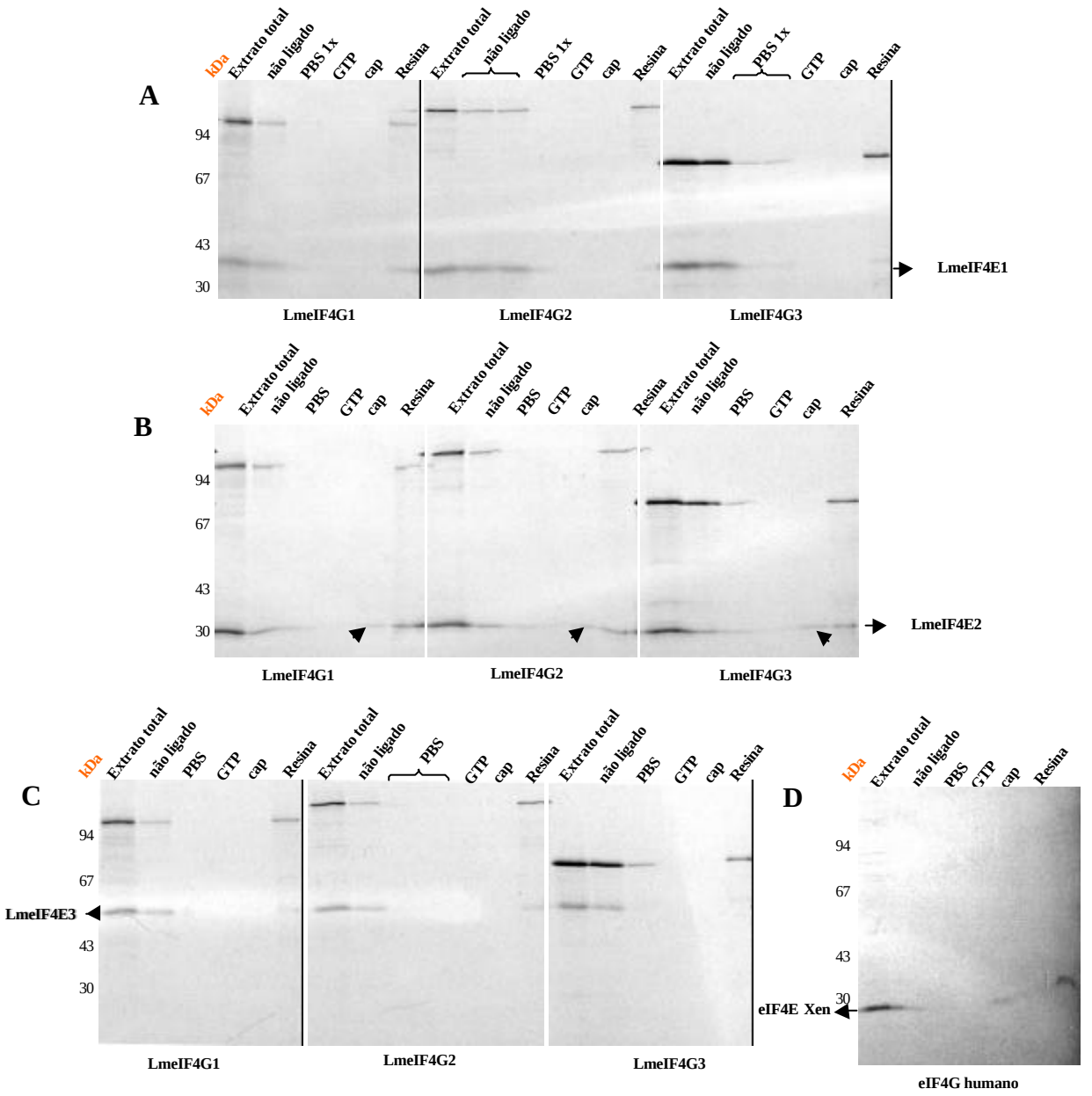


Figura 8. Análise da interação LmeIF4E/LmeIF4G ao cap. Para este experimento traduzimos *in vitro* RNAs sintéticos dos LmeIF4E1-3 combinados a RNAs dos LmeIF4G1-3 em RRL suplementado com ^{35}S. Como controle positivo fizemos a co-tradução dos RNAs eIF4G humano e eIF4E *X. laevis* . Após a tradução, cada conjunto foi incubado com a resina cap sefarose (AMERSHAM) por 1 hora no gelo, sendo levemente agitada a cada 10 minutos. Ao termino da incubação e após pulso de centrifugação, foi obtido o sobrenadante da ligação, aqui referido como não ligado. Em seguida a resina foi lavada com PBS 1x, GTP e cap solúvel. Em cada etapa, foram coletados amostras que permitiram obter as figuras acima. As amostras foram diluídas em LE2x para ficarem em quantidades semelhantes e autoradiografia foi obtida com descrito. Em **A** um ensaio de afinidade LmeIF4E1 pelo cap eucariótico na presença das diferentes proteínas LmeIF4G1-3. Em **B** e **C** de maneira análoga, enfocando respectivamente a proteína LmeIF4E2 e LmeIF4E3. Em **D** pode ser visualizado o ensaio de afinidade do eIF4E de *X. laevis* na presença do eIF4GI humano

3. DISCUSSÃO

Dados da literatura mostram que apenas um único fator de iniciação é descrito em *L. major*, o fator LmeIF4A1 (Skeiky *et al*., 1998). O segundo homólogo do eIF4A encontrado neste parasito (LmeIF4A2) aparentemente não é expresso na fase promastigota de *L.major*. Ensaios de *Western-blot* utilizando o Anti-LmeIF4A2 não foram capazes de detectar a referida proteína no extrato total mesmo em uma alta quantidade de células ($2x10^7$!). O fator LmeIF4A2 pode ser expresso em pequenas somas ou ser expresso na outra fase evolutiva do parasita (amastigota). Esta questão poderá ser respondida analisando o perfil de expressão desta proteína na outra fase do ciclo de vida do protozoário considerado. O fator LmeIF4A2 também não é capaz de interagir *in vitro* com o $LmeIF4G3_{26-310}$, Com o homólogo LmeIF4A1 foi observado o contrário: 1º) Este fator é expresso em alta quantidade, constituindo por certo uma das proteínas de maior expressão na forma promastigota de *L.major* (ver manuscrito em anexo). Este resultado é condizente com o que é descrito para eIF4A de outros organismos (von der Haar & McCarthy, 2002). 2º) O LmeIF4A1 também é capaz de interagir com um dos homólogos do LmeIF4G, $LmeIF4G3_{26-310}$. Esse resultado sugere que o fator LmeIF4A1 seja importante na tradução destes protozoários. Assim deduzimos que o LmeIF4A1 é certamente uma proteína envolvida na tradução e que o LmeIF4A2 não parece estar relacionado a este fenômeno.

A proteína LmeIF4G3 é expresso na forma promastigota de *L.major*. Em ensaios de *Western-blot*, utilizando-se soro anti-LmeIFG3, identificou-se que o respectivo gene se expressa, no parasita, em uma proteína idêntica em tamanho ao $LmeIF4G3_{1-636}$. Comparando diferentes concentrações da recombinante LmeIF4G3-GST com diferentes quantidades de células promastigotas de *L. major* estimamos a concentração total LmeIF4G3 em 1 nanograma/ 10^6 células. Ainda não fomos capazes de analisar a expressão dos homólogos LmeIF4G1-2, mais experimentos nestes sentido estão em andamento.

A afinidade dos homólogos LmeIF4Es por cap não é incrementada na presença dos homólogos LmeIF4Gs. Em outros sistemas eIF4E tem sua afinidade por cap aumentada quando o eIF4G ou eIF4F estão presentes (Haghighat & Sonenberg, 1997). Previamente, sabíamos que o LmeIF4E2 interage com o cap de mamíferos (ver manuscrito). Então fizemos um ensaio de afinidade pelo cap eucariótico combinando cada um LmeIF4Es com cada um dos LmeIF4Gs. Nenhum dos fatores LmeIF4Gs foi capaz de aumentar a afinidade de qualquer um dos LmeIF4Es na presença de cap. Ensaios de *pull down* podem ser outra estratégia utilizada para evidenciar se alguns deles são capazes de associarem-se entre si.

Uma possível justificativa para não lograrmos êxito no ensaio poderia ser porque não é encontrado nas proteínas LmeIF4Gs o consenso típico de ligação para o eIF4E descrito em vários organismos YXXXXLΦ (Mader *et al.*, 1995) e a região amino-terminal parece ser muito curta para comportar dois domínios de interação (eIF4E e PABP). Assim, parece que os LmeIF4Gs possuem um consenso diferente do habitual e ou arquitetura diferente de alguns de seus domínios, com regiões de interação para fatores como os LmeIF4Es situados em outra posição na proteína.

No processo de caracterização dos fatores LmeIF4G1-3 ainda serão necessários: 1) a conclusão do perfil de expressão dos homólogos LmeIF4G1 e LmeIF4G2 através de *Western-blot*, inclusive confirmando o tamanho das proteínas expressas no parasita; 2) realizar experimentos tais como *pull down*, imunoprecipitação, transfecção avaliando possível(is) interação(es) destes fatores a homólogos da PABP e eIF4E de *L major in vitro e in vivo.* A análise funcional destes fatores será algo que fornecerá dados relevantes para a compreensão do mecanismo de tradução em tripanosomatídeos, certamente o ponto chave de regulação da expressão gênica nestes protozoários.